

  
**CASEIB**  
**2022**

XL Congreso  
Anual de la  
Sociedad  
Española de  
Ingeniería  
Biomédica

**40** AÑOS UNIENDO FUERZAS  
PARA IMPULSAR LA  
INGENIERÍA BIOMÉDICA

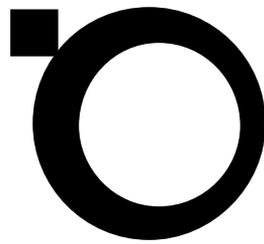
23-25 NOVIEMBRE 2022 | VALLADOLID

# LIBRO DE ACTAS



GRUPO DE  
INGENIERÍA  
BIOMÉDICA  
UNIVERSIDAD DE VALLADOLID





**CASEIB  
2022**

**XL Congreso Anual  
de la Sociedad  
Española de  
Ingeniería  
Biomédica**

**23-25 NOVIEMBRE 2022 | VALLADOLID**

*XL Congreso Anual de la Sociedad Española de Ingeniería Biomédica CASEIB 2022*

Los trabajos incluidos en esta publicación han sido evaluados por el Comité Científico que en ella se relaciona y según el procedimiento que se recoge en [www.caseib.es/2022/](http://www.caseib.es/2022/)

**Presidente del Comité Organizador:**

Roberto Hornero Sánchez

**Secretarios del Comité Organizador:**

Jesús Poza Crespo

Carlos Gómez Peña

**Editores:**

Roberto Hornero Sánchez

Jesús Poza Crespo

Carlos Gómez Peña

María García Gadañón

**Maquetación:**

Aarón Maturana Candelas

Carlos Gómez Peña

**Diseño de portada:**

Sandra Martínez Cagigal

Víctor Martínez Cagigal

ISBN: 978-84-09-45972-8



*Libro de Actas del XL Congreso Anual de la Sociedad Española de Ingeniería Biomédica*

Se distribuye bajo una licencia de Creative Commons Reconocimiento-NoComercial-SinObraDerivada 4.0 Internacional.

## ORGANIZADORES:



## COLABORADORES:



**PATROCINADORES:**



**AVALADO POR:**



## BIENVENIDA

Estimados congresistas,

En nombre del Comité Organizador, es un honor daros la bienvenida al XL Congreso Anual de la Sociedad Española de Ingeniería Biomédica (**CASEIB 2022**), que se celebra en Valladolid del 23 al 25 de noviembre de 2022. Esta edición ha sido organizada por el Grupo de Ingeniería Biomédica de la Universidad de Valladolid con la colaboración de la Sociedad Española de Ingeniería Biomédica (SEIB). Desde el Comité Organizador queremos destacar un importante matiz que hace a esta edición ciertamente especial, pues, tras dos años de celebración telemática, nuestro congreso recupera su carácter presencial. La organización de este evento ha sido apadrinada por la *International Federation for Medical and Biological Engineering* (IFMBE) y la *European Alliance for Medical and Biological Engineering and Sciences* (EAMBES).

**CASEIB 2022** es la cuadragésima edición de este congreso, lo que demuestra el magnífico estado de salud del que goza esta reunión anual y su importante papel en la difusión de la investigación científica relacionada con la Ingeniería Biomédica. Uno de los principales objetivos de **CASEIB 2022** es ser una herramienta de difusión científica de calidad, así como un punto de encuentro de expertos y profesionales de las diferentes ramas científicas y técnicas que interactúan en el área de la Ingeniería Biomédica. Puesto que la transferencia de conocimiento es imprescindible para crecer y evolucionar como profesionales en particular y como sociedad en general, hemos organizado dos mesas redondas para fortalecer la relación entre los principales actores que integran nuestra disciplina: el sector sanitario, la industria y la investigación. Con esta misma intención de garantizar el crecimiento y el futuro de la Ingeniería Biomédica, en esta edición se continuarán celebrando los concursos destinados a alumnos de Grado, Máster y Doctorado, para fomentar la participación de nuestros investigadores más jóvenes. Es nuestra intención que **CASEIB 2022**, de nuevo en formato presencial, continúe siendo nuestro lugar de encuentro y el congreso de referencia de la Ingeniería Biomédica en España.

En esta nueva edición de CASEIB hemos programado una serie de actividades científicas que esperamos sean de vuestro agrado. Con el objetivo de que el evento sea un lugar de enriquecimiento profesional y personal, **CASEIB 2022** presenta los siguientes actos de interés:

- Conferencia Inaugural del Congreso a cargo de Prof. Dr. Thomas Penzel, *Scientific Chair of the Interdisciplinary Center of Sleep Medicine* del Charité – Universitätsmedizin de Berlín (Alemania), con el título “*Sleep Medicine and Biomedical Engineering*”.
- Conferencia de Clausura del Congreso a cargo del Prof. Dr. Lluís Blanch, Director de Investigación e Innovación de la Corporació Sanitària Parc Taulí (Sabadell), con el título “*Traslación de la Innovación en Ingeniería Biomédica a los Hospitales*”.
- Mesa Redonda titulada “*Investigación y Transferencia en la Ingeniería Biomédica*”, para debatir sobre la importancia de la transferencia de conocimiento entre diferentes áreas del sector.
- Mesa Redonda de Sociedades Médicas, titulada “*El Ingeniero Biomédico y el Médico, Aliados en la Transformación hacia un Nuevo Modelo de Asistencia Sanitaria*”, que abordará desde un enfoque clínico el pronóstico de futuro de la práctica sanitaria y el rol de los ingenieros biomédicos en ella.
- Asamblea General de la Sociedad Española de Ingeniería Biomédica, para debatir las futuras actividades y objetivos de la Sociedad.
- Premio José María Ferrero Corral, en la que se presentan los seis mejores trabajos realizados por estudiantes de Máster y Doctorado.

- Competición FENIN-SEIB de estudiantes de Grado en Ingeniería Biomédica, en la que se exponen los mejores trabajos de estudiantes de este Grado en diferentes Universidades españolas.
- Premio EIT Health-SEIB, donde se presentarán los trabajos más emprendedores e innovadores de recientes graduados en Ingeniería Biomédica, en las modalidades “Patient-focused”, “Business-oriented” y “Audience Choice”.

**CASEIB 2022** también acoge una amplia muestra de trabajos científicos en diversas líneas de investigación de la Ingeniería Biomédica. Se han aceptado 130 artículos de conferencia y 15 resúmenes extendidos, revisados por el Comité Científico y repartidos entre las diferentes áreas temáticas del congreso. De ellos, 92 se presentan como comunicaciones orales y 53 en formato de póster.

- Comunicaciones orales. Se han planificado 12 sesiones orales organizadas en 4 bloques de 3 sesiones paralelas cada uno. En concreto, tres sesiones se dedican a Señales Biomédicas, dos a Neurotecnologías, una a Imágenes Biomédicas, una a Sistemas de Ayuda a la Decisión Clínica, una a Bioinstrumentación, una a Biomecánica, una a Informática Biomédica, una a Simulación y Planificación Quirúrgica y una a Telemedicina y e-Salud.
- Comunicaciones tipo póster. Se han organizado tres sesiones de pósteres en las que se presentan un total de 53 comunicaciones en diversas áreas de trabajo en Ingeniería Biomédica.

Para finalizar, nos gustaría mostrar nuestro agradecimiento a las instituciones que han colaborado en la realización del congreso. También estamos especialmente agradecidos a todos los participantes y asistentes al evento; sin ellos no sería posible la celebración de CASEIB un año más. Deseamos asimismo reconocer a los miembros del Comité Científico su labor en la revisión de los trabajos científicos y agradecer a los Conferenciantes invitados y Ponentes de las Mesas Redondas su generosidad a la hora de compartir con nosotros su experiencia. Por supuesto, queremos mostrar nuestra gratitud un año más a la SEIB por su confianza al encargarnos la organización de esta edición de CASEIB. También nuestro agradecimiento a la colaboración de la Federación de Empresas de Tecnologías Sanitarias (FENIN) y el nodo español del *European Institute of Innovation & Technology in Health* (EIT Health Spain). Finalmente, no podemos dejar de destacar el arduo trabajo de gestión del Comité Organizador, formado por miembros del Grupo de Ingeniería Biomédica de la Universidad de Valladolid. Gracias por vuestra diligencia y profesionalidad, que han hecho que **CASEIB 2022** sea una realidad.

Deseamos que el congreso esté a la altura de vuestras expectativas y lo disfrutéis enormemente. Se ha destinado mucho tiempo y un gran esfuerzo para que así sea. Para cualquier duda o sugerencia que pueda surgir, nos tenéis a vuestra entera disposición.

Un abrazo de todo el Comité Organizador.

**Dr. Roberto Hornero Sánchez**

*Presidente del Comité Organizador del CASEIB 2022*

**Dr. Jesús Poza Crespo**

*Secretario del Comité Organizador del CASEIB 2022*

**Dr. Carlos Gómez Peña**

*Secretario del Comité Organizador del CASEIB 2022*

## COMITÉ ORGANIZADOR

### **Presidente:**

Roberto Hornero Sánchez

### **Secretarios:**

Carlos Gómez Peña

Jesús Poza Crespo

### **Vocales:**

Daniel Álvarez González

María García Gadañón

Javier Gómez Pilar

Gonzalo C. Gutiérrez Tobal

Víctor Martínez Cagigal

Fernando Vaquerizo Villar

### **Colaboradores:**

Verónica Barroso García

Clara García Vicente

Enrique Gurdíel González

Víctor Gutiérrez de Pablo

María Herrero Tudela

Jorge Jiménez García

Diego Marcos Martínez

Adrián Martín Montero

Aarón Maturana Candelas

Selene Moreno Calderón

Sergio Pérez Velasco

Marcos Revilla Vallejo

Víctor Rodríguez González

Roberto Romero Oraá

Eduardo Santamaría Vázquez

## COMITÉ CIENTÍFICO

### Presidente:

Enrique J. Gómez Aguilera  
*Universidad Politécnica de Madrid*

### Vocales:

Begoña Acha Piñero  
*Universidad de Sevilla*

Raúl Alcaraz Martínez  
*Universidad de Castilla-La Mancha*

Erik Alonso González  
*Universidad del País Vasco*

Daniel Álvarez González  
*Universidad de Valladolid*

Juan Carlos Antony  
*GE Healthcare*

Elisabete Aramendi Ecenarro  
*Universidad del País Vasco*

José Luis Arce Diego  
*Universidad de Cantabria*

M<sup>a</sup> Gloria Bueno García  
*Universidad de Castilla-La Mancha*

Ana de los Reyes Guzmán  
*Servicio de Salud de Castilla-La Mancha*

Félix del Campo Matías  
*Universidad de Valladolid*

Manuel Desco Menéndez  
*Universidad Carlos III de Madrid*

Félix Fanjul Vélez  
*Universidad de Cantabria*

Javier García Casado  
*Universidad Politécnica de Valencia*

María García Gadañón  
*Universidad de Valladolid*

Beatriz F. Giraldo Giraldo  
*Universidad Politécnica de Cataluña*

### Secretarios:

Carlos Gómez Peña  
*Universidad de Valladolid*

Jesús Poza Crespo  
*Universidad de Valladolid*

Javier Gómez Pilar  
*Universidad de Valladolid*

Gonzalo C. Gutiérrez Tobal  
*Universidad de Valladolid*

María Elena Hernando Pérez  
*Universidad Politécnica de Madrid*

Roberto Hornero Sánchez  
*Universidad de Valladolid*

Raimon Jané Campos  
*Universidad Politécnica de Cataluña*

Pablo Laguna Lasasosa  
*Universidad de Zaragoza*

María Isabel López Gálvez  
*Hospital Clínico Universitario de Valladolid*

Víctor Martínez Cagigal  
*Universidad de Valladolid*

Gema Prats Boluda  
*Universidad Politécnica de Valencia*

Javier Reina Tosina  
*Universidad de Sevilla*

José Joaquín Rieta Ibáñez  
*Universidad Politécnica de Valencia*

Isabel Román Martínez  
*Universidad de Sevilla*

Francisco Javier Saiz Rodríguez  
*Universidad Politécnica de Valencia*

Patricia Sánchez González  
*Universidad Politécnica de Madrid*

Daniel Sánchez Morillo  
*Universidad de Cádiz*

Andrés Santos Lleó  
*Universidad Politécnica de Madrid*

Carmen Serrano Gotarredona  
*Universidad de Sevilla*

Ruth Solana Gracia  
*Hospital Universitario Infanta Leonor*

Abel Torres Cebrián  
*Universidad Politécnica de Cataluña*

Fernando Vaquerizo Villar  
*Universidad de Valladolid*

## PROGRAMA – MIÉRCOLES 23 DE NOVIEMBRE

MIÉRCOLES · 23 de NOVIEMBRE de 2022			
9:30 – 11:00	<b>Imágenes Biomédicas</b> Salón de Grados	<b>Señales Biomédicas I</b> Aula Magna	<b>Telemedicina y e-salud</b> Anfiteatro López-Prieto
11:00 – 12:00	<b>Sesión Pósteres I – Café</b> Vestíbulo Planta 3		
12:00 – 13:00	<b>Conferencia Plenaria Inaugural: Prof. Thomas Penzel</b> Aula Magna		
13:00 – 13:30	<b>Acto Inaugural</b> Aula Magna		
13:30 – 15:00	<b>Comida</b> Cafetería		
15:00 – 16:30	<b>Bioinstrumentación</b> Salón de Grados	<b>Informática Biomédica</b> Aula Magna	<b>Neurotecnologías I</b> Anfiteatro López-Prieto
16:30 – 19:00	<b>Premios FENIN-SEIB para alumnos Grado en IB</b> Anfiteatro López-Prieto		

### Acto Inaugural

Acto inaugural de CASEIB 2022 con la presencia de autoridades.

### Competición de Estudiantes de Grado en Ingeniería Biomédica

Presentaciones de los catorce finalistas de la Competición FENIN-SEIB de Estudiantes de los Grados en Ingeniería Biomédica.

### Conferencia Plenaria Inaugural

“*Sleep Medicine and Biomedical Engineering*”, Prof. Dr. Thomas Penzel.

## PROGRAMA – JUEVES 24 DE NOVIEMBRE

JUEVES · 24 de NOVIEMBRE de 2022			
8:30 – 10:00	<b>Biomecánica</b> Salón de Grados	<b>Señales Biomédicas II</b> Aula Magna	<b>Neurotecnologías II</b> Anfiteatro López-Prieto
10:00 – 11:00	<b>Sesión Pósteres II – Café</b> Vestíbulo Planta 3		
11:00 -12:00	<b>Mesa Redonda “Investigación y Transferencia en la Ingeniería Biomédica”</b> Aula Magna		
12:00 – 13:30	<b>Premio José María Ferrero Corral</b> Anfiteatro López-Prieto		
13:30 – 15:00	<b>Comida</b> Cafetería		
15:30 – 17:00	<b>Premio EIT Health-SEIB</b> Anfiteatro López-Prieto		
17:00 – 18:30	<b>Asamblea SEIB</b> Anfiteatro López-Prieto		

### Mesa Redonda

“*Investigación y Transferencia en la Ingeniería Biomédica*”, moderada por el Prof. Dr. Francisco Javier Saiz.

### Premio José María Ferrero Corral

Presentaciones de los seis finalistas del concurso de comunicaciones de estudiantes de Máster y Doctorado.

### Premio EIT Health-SEIB

Presentaciones de los trabajos más emprendedores e innovadores de recientes graduados en Ingeniería Biomédica, en las modalidades “Patient-focused”, “Business-oriented” y “Audience Choice”

### Asamblea de la SEIB

Asamblea Anual de la Sociedad Española de Ingeniería Biomédica.

## PROGRAMA – VIERNES 25 DE NOVIEMBRE

VIERNES · 25 de NOVIEMBRE de 2022		
8:30 – 10:00	<b>Sistemas de Ayuda a la Decisión</b> Salón de Grados	<b>Señales Biomédicas III</b> Aula Magna
		<b>Sim. y Planificación Quirúrgica</b> Anfiteatro López-Prieto
10:00 – 11:00	<b>Sesión Pósteres III – Café</b> Vestíbulo Planta 3	<b>Reunión de coordinadores de Grado en IB</b> Sala de Juntas
11:00 – 12:00	<b>Mesa Redonda “Sociedades Médicas”</b> Aula Magna	
12:00 – 13:00	<b>Conferencia de Clausura: Prof. Lluís Blanch</b> Aula Magna	
13:00 – 14:00	<b>Acto de Clausura/Entrega de Premios</b> Aula Magna	
14:00 – 15:30	<b>Cóctel de Clausura</b> Vestíbulo Planta 3	

### Mesa Redonda de Sociedades Médicas

“*El Ingeniero Biomédico y el Médico, Aliados en la Transformación hacia un Nuevo Modelo de Asistencia Sanitaria*”, moderada por la Dra. Ruth Solana Gracia y el Prof. Dr. Enrique Gómez Aguilera.

### Conferencia de Clausura

“*Traslación de la Innovación en Ingeniería Biomédica a los Hospitales*”, Prof. Dr. Lluís Blanch.

### Acto de clausura / Entrega de Premios

Acto de clausura de CASEIB 2022 y entrega del Premio José María Ferrero Corral y de las distinciones de la Competición FENIN-SEIB.

## CONFERENCIA PLENARIA INAUGURAL

### Sleep Medicine and Biomedical Engineering

Miércoles 23 de noviembre, 12:00 h – Aula Magna

Moderador: Dr. Gonzalo César Gutiérrez Tobal. *Universidad de Valladolid.*



#### **Prof. Dr. Thomas Penzel**

*Scientific Chair of the Interdisciplinary Center of Sleep Medicine*

Thomas Penzel se graduó en física (1986), biología humana (1991) y fisiología (1995) en la Universidad de Marburg, Alemania. En 2006 se trasladó a Berlín donde es director de investigación del Centro Interdisciplinario de Medicina del Sueño de la *Charité - Universitätsmedizin Berlin* (Alemania). En 2001, recibió el premio Bial de Medicina Clínica en Portugal, en 2008 el Premio Bill Gruen a la Innovación en la Investigación del Sueño de la Sociedad Estadounidense de Investigación del Sueño y en 2014 el distinguido premio de investigación de la Sociedad China

de Investigación del Sueño. Fue presidente de la conferencia IEEE EMBC 2019, conferencista distinguido y miembro del IEEE. Es presidente de la Sociedad Alemana del Sueño y miembro Adcom de IEEE EMBS. Es editor jefe de la revista *Sleep and Breathing* y miembro del consejo editorial de otras revistas. Ha publicado 400 artículos en revistas (Pubmed), 80 capítulos de libro y ha editado varios libros. Sus intereses de investigación son la medicina del sueño, las señales biomédicas y los dispositivos portátiles para el registro del sueño.

## CONFERENCIA PLENARIA DE CLAUSURA

### Traslación de la Innovación en Ingeniería Biomédica a los Hospitales

**Viernes 25 de noviembre, 12:00 h – Aula Magna**

Moderador: Prof. Dr. Enrique Javier Gómez Aguilera. *Universidad Politécnica de Madrid.*



#### **Prof. Dr. Lluís Blanch**

*Director de Investigación e Innovación de la Corporació Sanitària Parc Taulí (Sabadell)*

El Dr. Lluís Blanch es especialista en Medicina Intensiva y Premio Extraordinario de Doctorado por la Universitat Autònoma de Barcelona (1994). Actualmente es Director de Investigación e Innovación de la Corporació Sanitària Parc Taulí (Sabadell), Director del Instituto de Investigación e Innovación Parc Taulí (I3PT), Jefe del Grupo 33 del CIBER de Enfermedades Respiratorias (CIBERES) del Instituto de Salud Carlos III (ISCIII) y Coordinador de la Plataforma

ISCIII de Dinamización e Innovación de las Capacidades industriales del SNS (ITEMAS). El Dr. Lluís Blanch ha realizado 2 estancias de un año en la Universidad McGill de Montreal (1987-88) y en la Universidad de Minnesota en St. Paul, MN (1997-98). Ha sido Presidente de la Sociedad Española de Medicina Intensiva y Unidades Coronarias (SEMICYUC). Sus ámbitos de interés son la investigación traslacional en insuficiencia respiratoria aguda y la innovación tecnológica relacionada con conectividad, interoperabilidad y análisis avanzado de los datos generados por el paciente grave. Tiene más de 240 publicaciones indexadas, un índice h de 46, es revisor de numerosas revistas del primer cuartil de su especialidad y ha obtenido numerosas ayudas para proyectos de investigación y de innovación. Consta como inventor en 3 patentes y es socio fundador de la empresa Better Care S.L. *spin off* de la Corporació Sanitària Parc Taulí para el desarrollo de una plataforma de monitorización inteligente de señales biomédicas en pacientes críticos.

## MESAS REDONDAS

### Mesa Redonda: *‘Investigación y Transferencia en la Ingeniería Biomédica’*

**Jueves 24 de noviembre, 11:00 h – Aula Magna**

#### **Moderador**

- Dr. Francisco Javier Saiz, *Catedrático de la Universidad Politécnica de Valencia (UPV), Coordinador del Subárea “Ingeniería Biomédica” del Plan Nacional de Investigación*

#### **Participantes**

- Dr. Cristobal Belda-Iniesta, *Director General del Instituto de Salud Carlos III*
- D. Juan Luis Romera Santiago, *CDTI, Departamento de Promoción Institucional y Cooperación Territorial, Dirección de Programas de la UE y Cooperación Territorial*
- D<sup>a</sup>. Ángeles Barrios, *EIT Health Spain, Public Affairs & Stakeholder Relations Lead*

## **Mesa Redonda de Sociedades Médicas: ‘El Ingeniero Biomédico y el Médico, Aliados en la Transformación hacia un Nuevo Modelo de Asistencia Sanitaria’**

**Viernes 25 de noviembre, 11:00 h – Aula Magna**

### **Moderadores**

- Dra. Ruth Solana Gracia. *Vocal de la Junta Directiva de la SEIB. Especialista en Cardiología Infantil en el Hospital Infanta Leonor de Madrid*
- Prof. Dr. Enrique Gómez Aguilera. *Presidente de la Sociedad Española de Ingeniería Biomédica. Catedrático de la Universidad Politécnica de Madrid*

### **Participantes**

- Dr. José Alberto San Román. *Jefe de Servicio de Cardiología y Director del Instituto de Ciencias del Corazón (ICICOR) del Hospital Clínico de Valladolid, en representación de la Sociedad Española de Cardiología (SEC)*
- Dr. Javier García Fernández. *Presidente de la Sociedad de Anestesiología, Reanimación y Terapéutica del Dolor (SEDAR)*
- Dr. Carlos Casillas. *Radiólogo del Hospital Vithas Rey Don Jaime de Castellón, anterior Vicepresidente de la filial valenciana de la SERAM y Presidente Saliente de la SEDIA (sección de abdomen de la SERAM)*
- Dra. Patricia Alonso Fernández. *Médico Adjunto de Servicio de Admisión del Hospital Clínico San Carlos, Vocal de la Junta Directiva de la Sociedad Española de Directivos de la Salud (SEDISA)*

## COMPETICIONES

### Premio José María Ferrero Corral

**Jueves 24 de noviembre, 12:00 h – Anfiteatro López-Prieto**

CASEIB 2022 organiza con la colaboración de la SEIB el Premio José María Ferrero Corral para estudiantes de Máster y Doctorado en el área de la Ingeniería Biomédica. Los seis finalistas presentarán sus trabajos durante 10 minutos, a los que le seguirán 5 minutos de debate con el tribunal.

#### Tribunal

- Prof. Dr. Enrique Gómez Aguilera. *Universidad Politécnica de Madrid*
- Prof. Dr. Luis Javier Reina Tosina. *Universidad de Sevilla*
- Prof. Dra. Gema Prats Boluda. *Universidad Politécnica de Valencia*

#### Trabajos seleccionados

- Ander Loidi & Jose M. Ferrero. “*Análisis computacional de las características ideales de fármacos en isquemia miocárdica aguda*”
- Marcos Revilla-Vallejo, *et al.* “*Acoplamiento cruzado amplitud-amplitud: un nuevo método para caracterizar la actividad neuronal y las alteraciones provocadas por la enfermedad de Alzheimer*”
- Álvaro Huerta Herraiz, *et al.* “*Efecto de la congelación de capas en redes neuronales convolucionales entrenadas para la evaluación de la calidad del ECG*”
- Juan Miguel Mira-Tomás, *et al.* “*Assessment of common information in surface electromyography recordings with adhesive electrodes and an intravaginal probe*”
- Víctor Gutiérrez-de Pablo, *et al.* “*Alteraciones patológicas de la estructura neurofisiológica en los diferentes subtipos de migraña en mujeres*”
- Carlos Gallardo, *et al.* “*Glucose prediction based on deep learning models and objective physical activity data*”

## Competición FENIN-SEIB de Estudiantes de Grado en Ingeniería Biomédica

**Miércoles 23 de noviembre, 16:30 h – Anfiteatro López-Prieto**

CASEIB 2022 acoge un año más la Competición de Estudiantes de Grado en Ingeniería Biomédica coordinada por la FENIN y la SEIB. En esta edición del CASEIB, hay catorce finalistas seleccionados de entre todos los trabajos enviados con resultados relevantes en el campo de la Ingeniería Biomédica de diferentes Universidades, ya sea como parte de su Trabajo Fin de Grado o de trabajos de asignaturas de la titulación. Durante la sesión, los estudiantes expondrán oralmente su trabajo durante 7 minutos y debatirán con el tribunal durante 3 minutos.

### Tribunal

- Dra. Elisabete Aramendi Ecenarro. *Universidad del País Vasco*
- Dra. Beatriz F. Giraldo Giraldo. *Universidad Politécnica de Cataluña*
- D. Juan Carlos Antony García. *GE Healthcare*

### Trabajos seleccionados

- Antoni Ignasi Cànaves Llabrés, *et al.* “*Tumor-on-a-chip: new strategies of in-vitro cancer cell culture*”, Universitat de Girona
- Eliana Millaray Solar Risueño & D. Marcos Jorquera. “*Sistema de asistencia y monitorización para personas con Deterioro Cognitivo Leve*”, Universidad de Alicante
- Blanca Romero Milá, *et al.* “*EEG-Based Functional Connectivity during Progression from Infantile Spasms to Lennox Gastaut Syndrome*”, Universidad de Barcelona
- Adri Gómez, *et al.* “*Creación y optimización automática de arquitecturas de aprendizaje profundo aplicadas a imagen médica*”, Universidad Carlos III de Madrid
- Judith Zubia Aranburu, *et al.* “*Quantification of the adhesion force of E. coli on Ti via single-cell force spectroscopy*”, Universidad de Mondragón
- Auba Fuster Palà, *et al.* “*In vitro neuronal cultures and network analysis to build a proof-of-concept biological computing AI device*”, Universitat Pompeu Fabra
- Blanca Gutiez-Bravo, *et al.* “*Design and implementation of a virtual reality training environment for health emergency technicians*”, Universidad Politécnica de Madrid
- María Rollano Corroto. “*Design, Development and Validation of an Orthosis for a Brachial Plexus Injury*”, Universidad CEU San Pablo
- Clara Sánchez del Valle, *et al.* “*Determinación de requisitos biomecánicos y funcionales en población sana con el software RehabHand*”, Universidad Europea de Madrid
- Francesca Aguilar Paredes, *et al.* “*Análisis multicanal de sonidos respiratorios mediante entropía muestral fija para la evaluación de la COPD*”, Universidad Politécnica de Cataluña

- Agnés Rigo-Vidal, *et al.* “*Análisis de un nuevo método de posicionamiento de electrodos en el cuero cabelludo para las pruebas electrofisiológicas*”, Universitat Rovira I Virgili
- Sara Valiente Jaén, *et al.* “*Localización de Drivers en Fibrilación Auricular con Redes Convolucionales LSTM*”, Universidad Rey Juan Carlos
- María Martínez Donázar, *et al.* “*Development of strategies for the generation and characterisation of a polymeric scaffold for in vitro bladder reconstruction*”, TECNUN Universidad de Navarra
- Beatriz Lafuente Alcázar. “*Análisis de un Algoritmo de Deep Learning para Detectar Picos R en Electrocardiogramas*”, Universitat Politècnica de València

## ACTOS

### Acto inaugural

#### Miércoles 23 de noviembre, 13:00 h – Aula Magna

La XL edición del Congreso Anual de la Sociedad Española de Ingeniería Biomédica (CASEIB 2022) contará en su inauguración con la presencia de las siguientes autoridades, en representación de la Universidad de Valladolid, la Facultad de Medicina, la Sociedad Española de Ingeniería Biomédica y el Comité Organizador de CASEIB 2022:

- Prof. Dr. Antonio Largo Cabrerizo, *Excmo. Sr. Rector Magnífico de la Universidad de Valladolid*
- Prof. Dr. José María Fidel Fernández Gómez, *Decano de la Facultad de Medicina de la Universidad de Valladolid*
- Prof. Dr. Enrique J. Gómez Aguilera, *Presidente de la Sociedad Española de Ingeniería Biomédica*
- Prof. Dr. Roberto Hornero Sánchez, *Presidente del Comité Organizador del XL Congreso Anual de la Sociedad Española de Ingeniería Biomédica (CASEIB 2022)*

### Asamblea de la SEIB

#### Jueves 24 de noviembre, 17:00 h – Anfiteatro López-Prieto

Durante el CASEIB 2022 se celebrará la asamblea anual de la Sociedad Española de Ingeniería Biomédica.

### Acto de clausura

#### Viernes 25 de noviembre, 13:00 h – Aula Magna

El XL Congreso Anual de la Sociedad Española de Ingeniería Biomédica será clausurado por los representantes de la Universidad de Valladolid, de la SEIB y del Comité Organizador de CASEIB 2022. Durante la ceremonia de clausura se hará entrega del Premio José María Ferrero Corral y de las distinciones de la Competición de Estudiantes de Grado en Ingeniería Biomédica:

- Prof. Dr. Enrique Baeyens Lázaro. *Vicerrector de Investigación de la Universidad de Valladolid*
- Prof. Dr. Enrique J. Gómez Aguilera. *Presidente de la Sociedad Española de Ingeniería Biomédica*
- Dr. Jesús Poza Crespo. *Secretario del Comité Organizador de CASEIB 2022*

## ÍNDICE DE SESIONES CIENTÍFICAS

Miércoles, 23 de noviembre de 2022, a las 9:30

---

### Sesión oral: Imágenes Biomédicas

**Moderadores:** M<sup>a</sup> Gloria Bueno García y María García Gadañón

Breast cancer classification based on hyperspectral imaging .....	2
Metodologías de cuantificación de imágenes [ <sup>18</sup> F]FDG PET/CT para la detección de infiltración medular en pacientes con linfoma folicular .....	6
Corrección automática del contraste en imagen radiográfica mediante aprendizaje profundo ..	10
Mecanismo de atención y <i>deep learning</i> en la evaluación automática de la calidad de las retinografías .....	14
Uso de la Termografía para el Análisis y Seguimiento de los Hemangiomas Infantiles .....	18
Análisis textural de imágenes [ <sup>18</sup> F]FDG PET/CT para la predicción de enfermedad mínima residual en pacientes con sintomatología de mieloma múltiple .....	22

### Sesión oral: Señales Biomédicas I

**Moderadores:** Raimon Jané y Gonzalo César Gutiérrez Tobal

Análisis de los cambios en la conectividad funcional tras un entrenamiento cognitivo mediante Neurofeedback .....	27
Metodología Robusta Basada en los Fundamentos del Machine Learning Para la Clasificación de Señales Biomédicas. Aplicación a 3 Desafíos de la Ingeniería Biomédica Moderna .....	31
Analysis of surface electromyographic parameters for the assessment of muscle fatigue during moderate exercises .....	35
Estimación de la severidad de apnea del sueño pediátrica a partir de la señal de flujo aéreo usando redes neuronales convolucionales.....	39
Medición automática del "jitter" en registros electromiográficos patológicos .....	43
Impacto de la expresión del gen <i>MAPT</i> en las relaciones inter-banda de la actividad eléctrica cerebral .....	47

## Sesión oral: Telemedicina y E-Salud

**Moderadores:** Luis Javier Reina Tosina y Daniel Álvarez González

Implementación y evaluación de una plataforma tecnológica para estrategias nutricionales personalizadas en personas mayores .....	52
Percepción de la ciberseguridad en usuarios sobre soluciones de salud digital con dispositivos HIoT .....	56
Herramienta de valoración de aplicaciones móviles de eSalud: programadores, pacientes y profesionales de salud.....	60
Entrenamiento de la memoria declarativa episódica mediante serious games en personas mayores con deterioro cognitivo leve .....	64
Plataforma mHealth ALIBIRD®: Descripción y resultados preliminares de su evaluación en la asistencia clínica de pacientes oncológicos .....	68
Sistema de ayuda a la prescripción farmacológica .....	72

Miércoles, 23 de noviembre de 2022, a las 11:00

---

## Sesión de pósteres I

**Moderador:** Aarón Maturana Candelas

Implementación de Sistema de Cribado para la Audición y la Visión de Neonatos .....	77
Fabricación de fibras de alta resistencia obtenidas a partir de glándulas sericígenas de <i>Samia cynthia ricini</i> . Caracterización y comparación con <i>Bombyx mori</i> y potenciales aplicaciones en Ingeniería Tisular .....	81
Simulación paralela multidispositivo de secuencias de imagen por resonancia magnética .....	85
Análisis hemodinámico en dos localizaciones de la Aorta .....	89
Influencia del tipo de muerte del donante en la supervivencia del trasplante renal: un estudio centrado en la población andaluza .....	93
Realidad Virtual Inmersiva para rehabilitación del miembro superior en patología neurológica	97
A review on the use of sEMG for the analysis of motor cues in Parkinson's disease: The AI4HealthyAging project .....	101
Diseño de un dispositivo implantado con transmisión inalámbrica de energía para tratamiento del Glioblastoma mediante TTF .....	105
Estudio de las alteraciones en la arquitectura temporal de la actividad neuronal provocadas por la enfermedad de Alzheimer .....	109
Technological platforms using bioengineering techniques for studies of the alternative vasculature in neuroblastoma .....	113
Diseño de modelos automáticos basados en inteligencia artificial para predicción del Síndrome Metabólico .....	117
Transformación OMOP-CDM del repositorio para usos secundarios de la cohorte de la Red Española de Investigación en SIDA (CoRIS) .....	121
Validación preliminar de ITACA: Un entorno novedoso para estudios de Neurofeedback .....	125

Miércoles, 23 de noviembre de 2022, a las 15:00

---

### Sesión oral: Bioinstrumentación

**Moderadores:** Francisco Javier García Casado y José Luis Arce Diego

Sistema integral de pupilometría binocular para roedores .....	130
Nuevo Método para el Análisis Independiente de la Orientación en Matrices Multielectrodo Equiespaciadas .....	134
Estudio Comparativo con Señales Epicárdicas de las Limitaciones del Omnipolo con Multielectrodos de Alta Densidad.....	138
Development of techniques to minimize the interference of muscle activity on the electroencephalographic signal during swallowing.....	142
Altered pelvic floor muscle activity associated with the use of intravaginal probes in severe vulvodynia.....	146
Cuantificación de la Heterogeneidad del Sustrato Electrofisiológico Cardíaco en Registros Obtenidos mediante Multielectrodos de Alta Densidad .....	150

### Sesión oral: Informática Biomédica

**Moderadores:** Carmen Serrano Gotarredona y Erik Alonso González

Segmentación de la capnografía durante la resucitación cardiopulmonar mediante el uso de redes neuronales.....	155
Simulación de la actividad electrofisiológica auricular mediante autómatas celulares.....	159
Subject-based association networks for clustering analysis and phenotyping of pediatric sleep apnea .....	163
A Cross-Platform Editor and Simulator of Magnetic Resonance Imaging Sequences: Design and Implementation .....	167
ARforIOERT: Una aplicación de Realidad Aumentada para el entrenamiento en la administración de Radioterapia Intraoperatoria .....	171
Aprendizaje profundo para la segmentación de ventilaciones en impedancia durante la resucitación cardiopulmonar.....	175

## Sesión oral: Neurotecnologías I

**Moderadores:** José María Azorín Poveda y Eduardo Rocón De Lima

An Overview of M3Rob, a Robotic Platform for Neuromotor and Cognitive Rehabilitation Using Augmented Reality .....	180
Precisión de movimientos y gestos de la mano en terapias virtuales inmersivas.....	184
Análisis de Imaginación Motora durante pedaleo a partir de señales EEG .....	188
Propuesta de un sistema domótico de fácil configuración controlado por la actividad cerebral y aplicado a personas afectadas con ELA.....	192
Un nuevo método de parada temprana no paramétrico para sistemas Brain–Computer Interface basados en c-VEP.....	196

Miércoles, 23 de noviembre de 2022, a las 16:30

---

**Sesión oral: Competición FENIN-SEIB de Estudiantes de Grado en Ingeniería Biomédica**

**Tribunal:** Elisabete Aramendi Ecenarro, Juan Carlos Antony y Beatriz F. Giraldo

Tumor-on-a-chip: new strategies of in-vitro cancer cell culture .....	201
Sistema de asistencia y monitorización para personas con Deterioro Cognitivo Leve.....	205
EEG-Based Functional Connectivity during Progression from Infantile Spasms to Lennox Gastaut Syndrome.....	209
Creación y optimización automática de arquitecturas de aprendizaje profundo aplicadas a imagen médica .....	213
Quantification of the adhesion force of <i>E. coli</i> on Ti via single-cell force spectroscopy .....	217
In vitro neuronal cultures and network analysis to build a proof-of-concept biological computing AI device.....	221
Design and implementation of a virtual reality training environment for health emergency technicians.....	225
Design, Development and Validation of an Orthosis for a Brachial Plexus Injury.....	229
Determinación de requisitos biomecánicos y funcionales en población sana con el software RehabHand .....	233
Análisis multicanal de sonidos respiratorios mediante entropía muestral fija para la evaluación de la COPD .....	237
Análisis de un nuevo método de posicionamiento de electrodos en el cuero cabelludo para las pruebas electrofisiológicas .....	241
Localización de Drivers en Fibrilación Auricular con Redes Convolucionales LSTM .....	245
Development of strategies for the generation and characterisation of a polymeric scaffold for in vitro bladder reconstruction.....	249
Análisis de un Algoritmo de Deep Learning para Detectar Picos R en Electrocardiogramas ...	253

Jueves, 24 de noviembre de 2022, a las 8:30

---

### Sesión oral: Biomecánica

**Moderadores:** Abraham Otero Quintana y Pablo F. Viñas Nosedá

Design and development of a robotic hand for finger spelling .....	258
A virtual reality application for the assessment and training of neck sensorimotor control .....	262
Efecto del entrenamiento de la anastomosis microquirúrgica experimental en la calidad del procedimiento y la ergonomía del cirujano .....	266
Automatización del método RULA para análisis ergonómico con sensores inerciales vestibles	270
Biomechanical modelling of the behaviour of the human breast under the effect of gravity for tumor detection .....	274
Estudio preliminar de la variación del rango de movimiento y la actividad muscular en pacientes de avanzada edad usando el andador SWalker .....	278

### Sesión oral: Señales Biomédicas II

**Moderadores:** Javier Gómez Pilar y Víctor Martínez Cagigal

Improving the quality of combined TMS-EEG neural recordings: artefact removal and time analysis .....	283
Análisis espectral de la actividad electroencefalográfica basal a nivel de fuente para la diferenciación de subtipos de migraña .....	287
Protocolo de registro de EEG para identificación de patrones cerebrales asociados a la anosmia por COVID-19.....	291
Sincronía, metaestabilidad y dimensión fractal del EEG durante diferentes estados conscientes	295
Cognitive stimulation leads to a long-term brain activation with a spectral shift to higher frequencies in Rett syndrome patients .....	299
Aplicación de <i>Deep Learning</i> para el procesado automático de componentes ICA de registros de magnetoencefalografía .....	303

## Sesión oral: Neurotecnologías II

**Moderadores:** Eduardo Rocón De Lima y Ana de los Reyes Guzmán

Evaluación del nivel de atención durante el entrenamiento de una BMI con un exoesqueleto de miembro inferior .....	308
Neuroprótesis para evaluar el efecto de estimulación vibratoria mecánica en el freezing de pacientes con Parkinson .....	312
Caracterización espacio-temporal de la clasificación de imaginación motora con herramientas de <i>explainable artificial intelligence</i> (XAI) .....	316
Diseño de una BCI mediante redes convolucionales profundas para la detección de la intención de parar ante un obstáculo inesperado .....	320
Evaluación de un videojuego multijugador basado en Brain Computer Interfaces utilizando c-VEPs .....	324
Rehabilitación de la marcha asistida por un sistema híbrido personalizable en dos casos de lesión medular incompleta .....	328

Jueves, 24 de noviembre de 2022, a las 10:00

---

## Sesión de pósteres II

**Moderador:** Roberto Romero Oraá

Estudio de las especificaciones energéticas para un dispositivo cerebral implantado recargable inalámbricamente .....	333
Deep Learning methodologies for brain image reconstruction in Positron Emission Tomography .....	337
The Identification of Patients with Type 1 Diabetes Mellitus Subgroups Applying Cluster Analysis Techniques .....	341
Evaluación del potencial de las imágenes OCT maculares para el diagnóstico de la enfermedad de Parkinson .....	345
POWERUP: Design and preliminary evaluation of a 3D-printed upper-limb exoskeleton for children with cerebral palsy .....	349
5 years of the Master of Technological Innovation in Health (EIT Health): A critical analysis of its strengths and weaknesses.....	353
Efecto de la distancia sonda – muestra en la aplicación de la espectroscopia de reflectancia difusa .....	357
Intra-variability in IB4- and IB4+ DRG neurons: a population of models study .....	361
La Complejidad Preoperatoria de la Distribución Espectral de las Ondas Fibrilatorias de Superficie Predice el Resultado de la Ablación por Catéter de Fibrilación Auricular Persistente	365
Análisis de Biomarcadores Electrocardiográficos en Pacientes con Cardiopatía Arritmogénica de Ventrículo Izquierdo .....	369
Desarrollo de un modelo de colon para formación en cirugía laparoscópica a partir de técnicas de procesamiento semiautomático de imagen médica e impresión 3D .....	373
Simulador clínico para el entrenamiento de la técnica de ecografía torácica y tratamiento de neumotórax y derrame pleural en neonatos .....	377
Development of a software for radiopharmaceutical management in nuclear medicine.....	381

Jueves, 24 de noviembre de 2022, a las 12:00

---

**Sesión oral: Premio José María Ferrero Corral**

**Tribunal:** Enrique Javier Gómez Aguilera, Luis Javier Reina Tosina y Gema Prats Boluda

Análisis computacional de las características ideales de fármacos en isquemia miocárdica aguda.....	386
Acoplamiento cruzado amplitud-amplitud: un nuevo método para caracterizar la actividad neuronal y las alteraciones provocadas por la enfermedad de Alzheimer .....	390
Efecto de la Congelación de Capas en Redes Neuronales Convolucionales Entrenadas para la Evaluación de la Calidad del ECG.....	394
Assessment of common information in surface electromyography recordings with adhesive electrodes and an intravaginal probe .....	398
Alteraciones patológicas de la estructura neurofisiológica en los diferentes subtipos de migraña en mujeres .....	402
Glucose prediction based on deep learning models and objective physical activity data .....	406

Viernes, 25 de noviembre de 2022, a las 8:30

---

### Sesión oral: Sistemas de Ayuda a la Decisión

**Moderadores:** Begoña Acha Piñero y Fernando Vaquerizo Villar

Modelo de Deep Learning basado en redes convolucionales 1D para el diagnóstico de la apnea del sueño mediante señales respiratorias .....	411
Clasificación explicada de las lesiones pigmentadas mediante la detección de los patrones dermatoscópicos de carcinoma .....	415
Diseño de un clasificador interpretable de ruido clínico en ECG de monitorización prolongada mediante autoencoders .....	419
Análisis de factores de riesgo y predicción de enfermedades crónicas usando métodos de aprendizaje automático .....	423
Combinación de redes neuronales convolucionales y recurrentes para la detección de la apnea obstructiva del sueño en niños empleando las señales de flujo aéreo y oximetría .....	427
Clasificación de ruido clínico en ECG de monitorización prolongada mediante máquinas de vectores soporte .....	431

### Sesión oral: Señales Biomédicas III

**Moderadores:** José Joaquín Rieta y Raúl Alcaraz Martínez

Simulación de señales ECG incluyendo dinámica del intervalo PQ con el ritmo cardíaco y ruido muscular variante en el tiempo .....	436
Simulación de los efectos de fármacos antiarrítmicos en la electrofisiología auricular .....	440
Detección automática de complejos QRS en pacientes con actividad eléctrica sin pulso durante la parada cardiorrespiratoria .....	444
Caracterización de los Cambios en las Medidas de la HRV durante Episodios de Apnea del Sueño .....	448
Métodos de filtrado adaptativo para la restauración de la forma de onda del ECG durante la reanimación cardiopulmonar .....	452
Análisis de la variabilidad del ritmo cardíaco durante la realización de apneas controladas en sujetos sanos .....	456

## Sesión oral: Simulación y Planificación Quirúrgica

**Moderadores:** Patricia Sánchez González y Félix Fanjul Vélez

Estudio <i>in Silico</i> de la Inducibilidad de Arritmias Reentrantes en Modelos Personalizados de Ventrículos Infartados .....	461
Detección del grado de limpieza de la mucosa basado en inteligencia artificial como sistema de apoyo para la evaluación de calidad en colonoscopias .....	465
Evaluación de la ablación quirúrgica mediante un láser Nd:YAG en aplicaciones dermatológicas sobre tejido porcino .....	469
Metodología de diseño y fabricación de guías personalizadas mediante impresión 3D para cirugía de columna. Patología: Escoliosis idiopática en adolescentes.....	473
Evaluación mediante simulación de nuevos enfoques en la terapia de resincronización cardiaca	477
Utilización de dispositivos de realidad mixta en cirugía maxilofacial .....	481

Viernes, 25 de noviembre de 2022, a las 10:00

---

### Sesión de pósteres III

**Moderador:** Verónica Barroso García

Simulador clínico para el entrenamiento de la técnica de transferencia embrionaria .....	486
Algoritmo y herramienta para la segmentación y volumetría esplénica sobre imágenes tomográficas .....	490
Discriminación Histológica mediante Propiedades Ópticas de Fase de Muestras Tisulares de Colon .....	494
Efecto del plomo en la generación de arritmias. Estudio de simulación en un modelo 3D de aurícula humana .....	498
Realidad Virtual para la rehabilitación motora de miembro inferior .....	502
Desarrollo de un encaje transfemoral para prótesis de miembro inferior asequible mediante el uso de tecnologías disruptivas .....	506
Design and implementation of cases for serious games to train soft skills for surgical teams ..	510
Uso de la inteligencia artificial para la formación en cirugía mínimamente invasiva: Revisión sistemática .....	514
Knowledge elicitation of pedagogical needs for medical learning based on 3D models .....	517
Modelos de predicción de la evaluación neuropsicológica a partir de cuestionarios sobre factores percibidos de salud y estilo de vida .....	521
Modelos de predicción del pronóstico neurológico de pacientes con hemorragia subaracnoidea (HSA) con ventilación asistida .....	525
Towards early detection of melanoma using deep learning and multimodal data .....	529

# Imágenes Biomédicas

---

Miércoles, 23 de noviembre de 2022

# Breast cancer classification based on hyperspectral imaging

G. Bueno<sup>1</sup>, N. Vállez<sup>1</sup>, L. González<sup>2</sup>

<sup>1</sup> VISILAB, E.T.S. Ingeniería Industrial, Ciudad Real, Spain {gloria.bueno, noelia.vallez}@uclm.es

<sup>2</sup> Servicio de Patología Digital, Hospital Universitario de Ciudad Real, Spain lucia.gonzalez@uclm.es

## Abstract

*This work aims to classify unstained breast tissue biopsies into two class tumour and non-tumour. For this purpose, deep learning techniques and multispectral images (MSI) acquired at different wavelengths, from 425 nm to 700 nm and a spectral step of 4 nm, are used. The work carried out includes an analysis to determine whether the different spectral bands provide additional information to that obtained with images acquired with ordinary cameras. As a result of the work, the MSI classification model which provides the closest hit rate to classification models using stained images is obtained. In addition, several digital staining methods have been implemented by applying deep learning techniques. Virtual staining will be performed after classification to visually corroborate the results, allowing for further analysis at cellular and tissue level.*

## 1. Introduction

Biopsy is a procedure consisting in the removal of tissue from the area to be analysed, which is subjected to preparative processes (e.g. staining or cutting) so that it can be interpreted by an expert. This process is essential for the diagnosis of some diseases such as cancer. Interpretation may be affected by the subjectivity of the expert, the complexity of the image and even the preparation process.

In recent years, several advances have been made in the automatic detection of tumours in multispectral imaging (MSI) and hyperspectral imaging (HSI), using different methods such as self-organised maps, decision trees, support vector machine (SVM) or Neural Networks, highlighting convolutional neural networks (CNN) [1-4]. There are studies suggesting that this type of images can improve upon the results obtained with RGB images in biomedical applications [5, 6], specifically in biopsy analysis. MSI or HSI contain information at different wavelengths within the electromagnetic spectrum.

Because most approaches to the above-mentioned detection problem focus on the classification of stained tissue, in this work we address the development of an automatic classification system on different types of tissue using multispectral images of unstained samples.

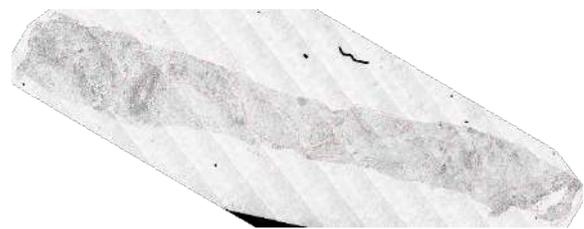
In addition, a study of the images acquired at different wavelengths is carried out to determine whether it is necessary to use all available wavelengths to achieve the best result.

Within neural networks, recent research support the use of Generative Antagonistic Networks (GANs) in the medicine field [7, 8]. Therefore, the application of different GANs is proposed in order to generate virtual staining of

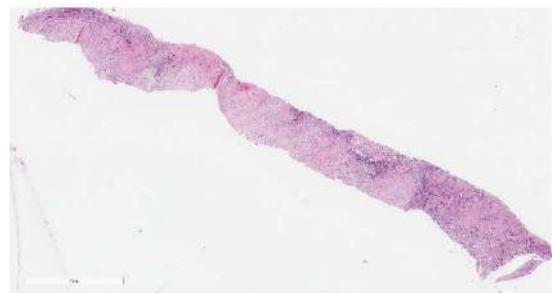
the images of unstained samples, assessing the effectiveness of this technology in this particular use.

## 2. Materials

The biopsies used in this work were obtained from the University Hospital in Ciudad Real, where the complete slides of a total of 13 pairs of whole slide images (WSI) stained breast biopsy specimens (haematoxylin-eosin) and non-stained were prepared and digitalised, classified as malignant (8 WSI), normal (4 WSI) and benign (1 WSI). Figures 1 and 2 show a complete WSI (Whole Slide Image) capture of tissue with and without staining. MSI were captured using the hyperspectral microscope IMA VIS hyperspectral microscope manufactured by Photon etc. The microscope is equipped with three objectives of 10x, 20x and 40x, a motorised stage in the X, Y and Z axes and a programmable spectral filter with a resolution of 1 nm and with a range of 400-700 nm.



*Figura 1. Complete unstained histological specimen*



*Figura 2. Complete histological specimen*

The regions of interest (ROI) were selected from these samples with the help of a specialist. The tissue found within these ROIs was classified into 5 classes: adipose tissue, stroma, non-tumour cellularity, tumour and tumour stroma. The samples were captured with a 20x objective in the spectral range 425-700 nm and a spectral step of 4 nm.

For the classification, the WSIs were subdivided into 255x255 pixel images, thus making a total of 13794 multispectral image clippings of unstained samples and 4631 image clippings of stained samples.

For the virtual staining, RGB images synthesised from the MSIs of the unstained samples were used, as well as the RGB images of the stained samples. The results obtained were used in the wavelength study to represent the MSIs in 3 channels. For this data set, the WSIs were subdivided into images of 1024x1024 pixels, yielding a total of 1509 images of unstained samples and 1520 of stained samples.

The datasets used for classification and the one used in the virtual staining with the pix2pix implementations were divided into training, validation, and test subsets with a proportion of 80%, 13% and 7%, respectively. The remaining implementations were divided into training (80%) and test (20%) subsets.

### 3. Methods

#### 3.1. Classification and wavelength study

Most multispectral image oriented CNNs are aimed at semantic segmentation applications and therefore their input sizes are very small. It was decided to use a fast-training architecture in the stained database and to add a number of layers in its input to convert the N-channel multispectral image, treated as a cube of pixels, to a three-channel RGB synthetic image.

Three implementations were made based on the GoogLeNet, Resnet-101 and Xception architectures, which were trained with the following parameters: 30 epochs, mini-batch size of 30, initial learning rate of 0.003, learning rate reduction factor of 0.1 and learning rate reduction period of 13 epochs.

An analysis of the wavelengths composing the MSI was carried out to test if it was possible to extract useful information for the classification of the MSI to improve the trained model with images of stained samples. This process was used for the virtual staining phase, as a 3-channel representation of the images was necessary.

This study consisted in eliminating the information of each wavelength in the classification and see how the absence of this information affected the hit rate. A wavelength is understood to be more informative when its absence causes a noticeable worsening in the accuracy of the classification with respect to the model trained with all the wavelengths. The same experiment was later carried out with RGB combinations, i.e., combinations consisting of a total of three channels, of which one belongs to the red, one to the green and one to the blue spectrum.

#### 3.2. Dimensionality Reduction and Virtual Staining

Because neural network architectures are implemented to deal with RGB images, it was necessary to adapt the input for the MSI. Three methodologies were proposed for this, convolutional layers, principal component analysis (PCA) and an Autoencoder architecture. The best results were obtained with the convolutional layers and all results are given with this methodology.

For the implementation of virtual staining, three GAN implementations were used, CycleGAN [9, 10], pix2pix [11] and CUT [12]. For the training of the pix2pix network, the database with the paired images was used, i.e., the input

image was composed of the unstained sample image and the corresponding stained image. In contrast, CycleGAN and CUT were trained using the database with unpaired images, although further training was also performed by forcing the network to perform the training with the paired data. The training parameters for the different implementations were: initial learning rate of 0.0002, batch sizes of 1, 200, 400 and 600 epochs for training and rescaling of images to 256x256x3 pixels.

To initiate the experimentation, a model of each implementation was trained for the virtual images staining. Subsequently, with the intention of capturing the characteristics of each tissue and improving the performance of these networks, a model was trained for each type of tissue, with a total of 5 models per implementation.

The pix2pix implementation requires image pairs for training. Because the sample preparation processes (cutting, staining, etc.) cause changes in the tissue structure, it was necessary to perform a registration process in order to correctly align the image pairs. In total, the number of images forming this dataset is 1509 images.

## 4. Results

#### 4.1. Classification and wavelength study

The results shown in Table 1 were obtained from the first classification models, measured as accuracy in terms of True Positive, True Negative, False Positive and False Negative.

The Xception model trained with MSIs outperformed the results of the model trained with stained samples. The ResNet model trained with MSI matched the results of the model trained with stained samples.

ResNet-101 is a residual neural network which is deeper than GoogLeNet and skips connections to jump to specific layers, which are called skip connections or shortcuts. This architecture was proposed to address the problem of the so-called vanishing gradient. This occurs when the network is extended by adding layers to solve complex machine vision problems, resulting in performance saturation. These connections provide a shortcut between layers for gradient stepping, improving efficiency and allowing a 101-layer network.

On the other hand, Xception is a neural network composed of 36 convolutional layers with linear residual connections, which means that it includes shortcuts except in the first and last module. The hypothesis proposed is that the correlation between channels and the spatial correlation can be completely decoupled, therefore, it is better not to map them together. For this reason, it is based on modified depth-separable convolutional layers. This hypothesis is a version of the one proposed for the Inception architecture. The differences between the separable convolutional layers and those used in this architecture are the order of operations and the absence of non-linearity after the first operation.

Model	Multispectral	Stained
GoogLeNet	85.5%	93.6%
ResNet-101	93.8%	95.4%
Xception	<b>95.1%</b>	93.9%

**Tabla 1.** Results of the classification models trained with the 69 wavelengths and the RGB images of the stained samples

The following conclusions were drawn from the wavelength study:

- The models using the 3, 6 and 9 channels, considered most informative, were implemented. The hit rates obtained were lower than the models trained with all wavelengths. Most of the most informative channels belonged to the red spectrum, so it was concluded that the use of these channels together caused duplicities and noise to the model, which worsened its performance.
- The results improved with the use of combinations of 3 channels considered to be more locally informative within each of the RGB spectrum, i.e., one channel from the red, one from the green and one from the blue spectrum.
- It was determined that the channel combination with the highest accuracy in the classification models, and therefore considered to be the most informative combination, is the one consisting of the wavelengths 645 nm, 517 nm and 457 nm. The GoogLeNet architecture achieved a rate of 91.9%, 94.2% achieved by ResNet-101 and 95.1% achieved by Xception.

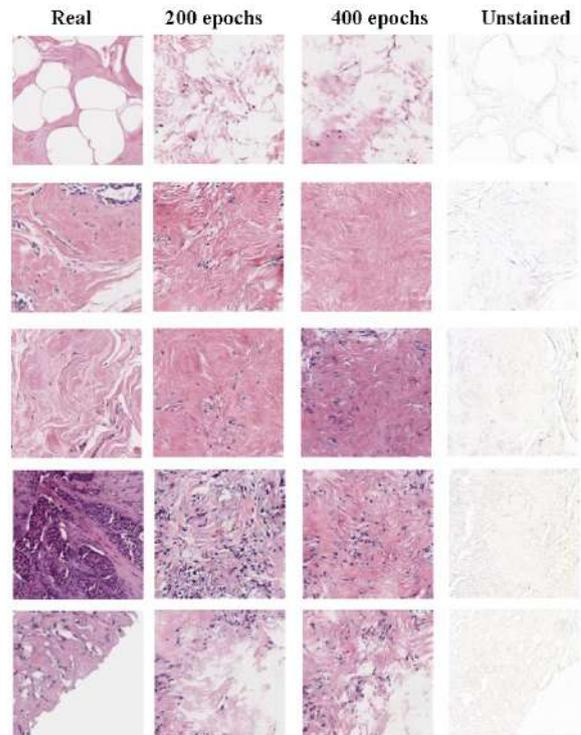
#### 4.2. Virtual Staining

An example of the results of the **pix2pix** implementation for virtual staining is presented in Figure 3, where each row corresponds to a tissue type: adipose, non-tumour, stroma, tumour and tumour stroma. The images generated by the **pix2pix** implementations show distortions with respect to the original images and even fail to differentiate correctly between tissue and background. The increase of epochs did not bring any improvement. For these reasons, the **pix2pix** implementations are discarded as valid options for virtual staining of multispectral sample images in this work.

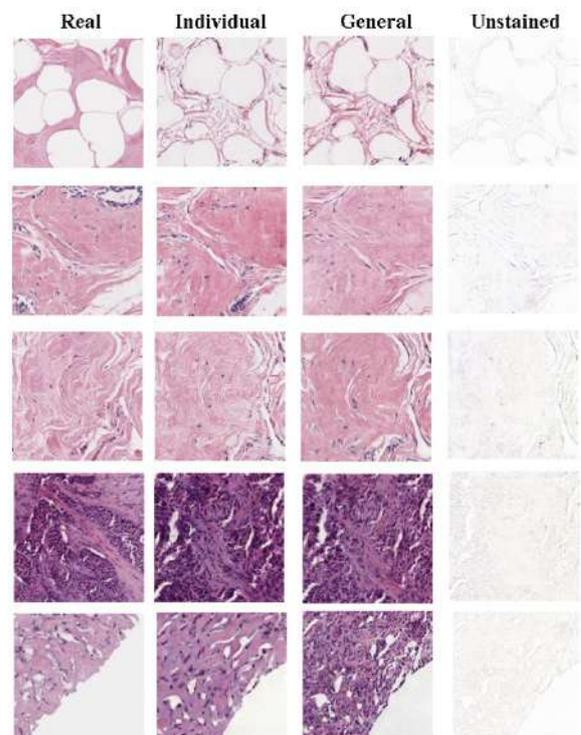
The **CycleGAN** models trained for virtual staining of all tissues are able to properly detect the image structure by applying a colouring which is very similar to the original image of stained samples. In addition, a **CycleGAN** model was trained for each tissue type, obtaining a total of 5 models. Figure 4 shows an example of the results of each model according to the tissue type in comparison with the images generated by the general model, all of them using 200 epochs.

Figure 5 shows the virtual staining generated by the models with the **CUT** implementation. As with **CycleGAN**, models were implemented for specific staining of each type of tissue. In most cases, it can be observed that it is able to respect the original structure of the tissue, but it generates colouring errors causing staining flaws. **CUT** models trained to stain adipose tissue, stroma and tumour

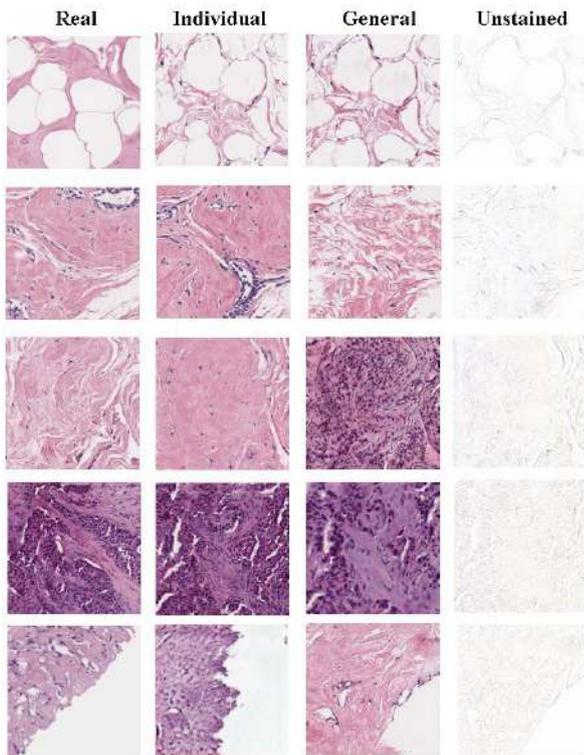
cellularity are able to apply a very realistic and consistent staining with respect to the unstained sample image. For the staining of the non-tumour cellularity and tumour stroma (row 2 and 5), the model suffers from difficulties in detecting shapes and to apply an appropriate staining. In the case of non-tumour cellularity, the increase in epochs during training is sufficient to overcome this problem.



**Figura 3.** Example of virtual staining with the **pix2pix** implementation



**Figura 4.** Example of virtual staining with the **CycleGAN** implementations trained with 200 epochs.



**Figura 5.** Example of virtual staining results with the CUT models for each tissue type at 600 epochs

## 5. Conclusions

This work demonstrates that useful information can be extracted from multispectral images of unstained samples which allows for automatic classification, potentially avoiding the staining process. The wavelength study shows that there are bands which are more useful for the model, allowing the image to be represented with only the 3 most representative channels without loss of information. These bands are 645 nm (channel 56 of the blue spectrum), 517 nm (channel 24 of the green spectrum) and 457 nm (channel 9 of the red spectrum).

In addition, it is shown that the use of all the wavelengths produces noise and duplicities that lead to a deterioration in the accuracy of the model.

The model based on Xception architecture, trained with multispectral images represented by the three most informative wavelengths above mentioned, obtained the best result (95.1%), outperforming the Xception model trained with images of stained samples and almost matching the best value of the stained images (95.4% with ResNet-101).

In the virtual staining, it is shown that the technologies used have great potential in this field. Although the results obtained are not qualitatively precise enough, the colour distance to the real samples is being analysed using different metrics. As an example, for future improvements, the number of samples could be increased in order to increase and improve the available labelled data sets.

Thus, the proposed methodology achieves an accuracy of 95.1% in distinguishing cancerous tissue from non-cancerous tissue compared to 95.4% obtained by

traditional methods of sample staining using biomarkers, which are considered good results for cancer screening. As biomarker staining is not required, the proposed methodology can reduce costs in hospitals, speed up the diagnostic process and increase its accuracy.

## Acknowledgements

This work has been funded by the HIPERDEEP project granted by Junta de Comunidades de Castilla-La Mancha (reference SBPLY/19/180501/000273). The authors would like to thank Elena M<sup>a</sup> Ruiz Izquierdo for her assistance with software development.

## References

- [1] Ma L., et al. Automatic detection of head and neck squamous cell carcinoma on histologic slides using hyperspectral microscopic imaging. *Journal of Biomedical Optics*, Vol. 27(4) pp 1–21, 2022.
- [2] Quintana L., et al. In the use of artificial intelligence and hyperspectral imaging in digital pathology for breast cancer cell identification. In Tomaszewski J. E., Ward A. D. and M.D., R. M. L., editors, *Medical Imaging 2022: Digital and Computational Pathology*, International Society for Optics and Photonics, SPIE, Vol. 12039, pp 73 – 82, 2022.
- [3] Ortega S., et al. Hyperspectral superpixelwise glioblastoma tumor detection in histological samples. *Applied Sciences*, Vol. 10 (13), 2020.
- [4] Kunthoth S. and Al-Maadeed S. Multispectral biopsy image based colorectal tumor grader. In: Valdés Hernández M., González-Castro V. (eds) *Medical Image Understanding and Analysis (MIUA)*. Communications in Computer and Information Science, Vol 723 pp 330–341, 2017.
- [5] Ortega S., et al. Hyperspectral superpixelwise glioblastoma tumor detection in histological samples. *Applied Sciences*, Vol. 10 (13), 2020.
- [6] Ortega S., et al. Hyperspectral and multispectral imaging in digital and computational pathology: a systematic review [invited]. *Biomedical Optics Express*, Vol. 11(6) pp 3195–3233, 2020.
- [7] Rivenson Y., et al. Phasestain: the digital staining of label-free quantitative phase microscopy images using deep learning. *Light: Science & Applications*, Vol. 8(1), 2019.
- [8] Xu Z., et al. Gan-based virtual re-staining: A promising solution for whole slide image analysis. *Computer Vision and Pattern Recognition*, 2019.
- [9] Zhu J.-Y., Park T., Isola P. and Efros A. A. Unpaired image-to-image translation using cycle-consistent adversarial networks. *IEEE International Conference on Computer Vision (ICCV)*, 2017.
- [10] Bayramoglu N., Kaakinen M., Eklund L. and Heikkila J. Towards virtual H&E staining of hyperspectral lung histology images using conditional generative adversarial networks. *Proceedings IEEE International Conference on Computer Vision Workshops (ICCVW)*, pp 64–71, 2017.
- [11] Isola P., Zhu J.-Y., Zhou T. and Efros A. Image-to-image translation with conditional adversarial networks. *Proceedings IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp 5967–5976, 2017.
- [12] Park T., Efros A. A., Zhang R. and Zhu J.-Y. Contrastive learning for unpaired image-to-image translation. *Computer Vision and Pattern Recognition*, 2020.

# Metodologías de cuantificación de imágenes [<sup>18</sup>F]FDG PET/CT para la detección de infiltración medular en pacientes con linfoma folicular

J. I. Velasco<sup>1</sup>, E. Milara<sup>1</sup>, P. Sarandeses<sup>2,3</sup>, A.P. Seiffert<sup>1</sup>, A. Gómez-Grande<sup>2,3</sup>, E.J. Gómez<sup>1,4</sup>, P Sánchez-González<sup>1,4</sup>

<sup>1</sup> Grupo de Bioingeniería y Telemedicina, ETSI Telecomunicación, Centro de Tecnología Biomédica, Universidad Politécnica de Madrid, Madrid, España; juan.velasco@alumnos.upm.es, {eva.milara.hernando, ap.seiffert, enriquejavier.gomez, p.sanchez}@upm.es

<sup>2</sup> Servicio de Medicina Nuclear, Hospital Universitario 12 de Octubre, Madrid, España; {mariadelpilar.sarandeses, adolfo.gomez}@salud.madrid.org

<sup>3</sup> Facultad de Medicina, Universidad Complutense de Madrid, Madrid, España

<sup>4</sup> Centro de Investigación Biomédica en Red de Bioingeniería, Biomateriales y Nanomedicina, Instituto de Salud Carlos III, Madrid, España

## Resumen

*El linfoma folicular es uno de los subtipos de linfoma de mayor incidencia a nivel mundial. Uno de sus síntomas más habituales es la infiltración de la médula ósea, presente en el 40-70% de los casos. Actualmente, la evaluación de esta infiltración se realiza mediante biopsia medular combinada con la valoración visual de las imágenes de tomografía por emisión de positrones. Sin embargo, es una valoración dependiente del observador y de su experiencia en el campo. En este trabajo se propone una prueba de concepto de una metodología de cuantificación de estas imágenes y detección de la infiltración medular. Para ello, se realiza una segmentación de la médula ósea de la cual se extraen un total de 30 características de radiómica. En primer lugar, se realizan un análisis U de Mann-Whitney y se calculan los coeficientes de correlación de Spearman para evaluar la relación entre las características extraídas de las imágenes y la infiltración medular. Posteriormente, se implementan 6 modelos de Machine Learning para la diferenciación entre los casos con infiltración medular positiva y negativa, alcanzando valores de entorno al 70-80% de sensibilidad y especificidad para los modelos árbol de decisión y random forest. Se puede concluir que la metodología propuesta para la cuantificación basada en el análisis textural puede ser usada en un sistema de ayuda a la decisión clínica de utilidad en la práctica actual de evaluación de infiltración medular.*

## 1. Introducción

Según la Organización Mundial de la Salud, el linfoma folicular (LF) es uno de los subtipos de linfoma con mayor incidencia (35%) de entre los Linfomas de No-Hodgkin, abarcando el 70% de los linfomas indolentes a nivel mundial [1]. El LF deriva de la ‘congelación’ del centro germinal de las células B dentro del folículo linfoide secundario que impide la reproducción y diferenciación de las células B [2]. Los pacientes con LF suelen presentar linfadenopatía asintomática e indolora. Sin embargo, algunos pueden experimentar molestias abdominales como dolor, saciedad precoz e hinchazón. Además, la enfermedad suele generalizarse incluyendo la afectación de múltiples ganglios linfáticos, el hígado y el bazo, siendo la infiltración más predominante la de la médula ósea (IM,

infiltración medular), presente en el 40-70% de los pacientes [1], [3]. Clínicamente, la IM puede producir anemia, trombocitopenia y neutropenia, por ello una detección adecuada es clave para la estadificación del LF, la predicción de su pronóstico y las consecuencias terapéuticas [2].

En la práctica clínica actual, la IM se evalúa mediante el análisis de una biopsia medular por trépano tomada a partir de la cresta iliaca. Sin embargo, la pequeña muestra que se toma puede no ser representativa del estado real de la médula, además de ser un método invasivo y con riesgo de complicaciones. Por este motivo, las imágenes [<sup>18</sup>F]FDG PET/CT están ganando importancia en el diagnóstico de la IM [4]. Para detectar su presencia normalmente se realiza una valoración visual por parte de un experto en medicina nuclear, evaluando la captación de la médula ósea, considerándose positiva cuando hay captación focal [5].

Con el objetivo de reducir la arbitrariedad de las evaluaciones visuales, se suelen utilizar métricas derivadas del valor de captación normalizado (SUV, *Standardize Uptake Value*). Sin embargo, estas métricas pueden estar afectadas por múltiples factores como el tiempo entre la inyección del radiotrazador y la adquisición de la imagen, actividad residual, características tecnológicas del escáner, etc. [6]. Por ello, se están desarrollando nuevas métricas de cuantificación de las imágenes [<sup>18</sup>F]FDG PET basadas en el análisis de radiómica. A pesar de que este tipo de análisis ha sido implementado en la evaluación de IM en otras patologías como leucemia [7], no se han realizado estudios para su evaluación en LF.

El objetivo principal de este estudio es el desarrollo de metodologías de cuantificación de imágenes [<sup>18</sup>F]FDG PET/CT para la evaluación de la IM en pacientes con LF. Para ello, se implementan múltiples análisis univariantes y modelos de *Machine Learning* basados en las características radiómicas extraídas de las imágenes [<sup>18</sup>F]FDG PET sobre la región de la médula ósea segmentada en imágenes CT.

## 2. Metodología

Para el presente estudio, se utiliza un conjunto de 41 imágenes [ $^{18}\text{F}$ ]FDG PET/CT de pacientes con LF recogidas en el Servicio de Medicina Nuclear del Hospital Universitario 12 de Octubre de Madrid. Para su posterior análisis, se cuenta con la valoración visual de las imágenes diferenciando entre infiltración medular positiva (IM+) y negativa (IM-) realizadas por expertos en Medicina Nuclear, obteniendo un total de 12 pacientes con IM+ y 29 con IM-.

Para la segmentación de la médula ósea se utiliza una herramienta desarrollada en el Grupo de Bioingeniería y Telemedicina (GBT) de la UPM e implementada en MATLAB (The MathWorks, Inc., Natick, Massachusetts, EEUU) [8]. Para la extracción de las características se emplea la librería de Vallieres llamada *Radiomics* [9], [10], desarrollada para MATLAB, y para el análisis de estas características se utiliza la herramienta SPSS *Statistics* (IBM Corp., Armonk, NY) [11] para el análisis univariante y el entorno de lenguaje de desarrollo *Python 3.9* (Python Software Foundation, EEUU) [12] para la implementación de los modelos de *Machine Learning*.

### 2.1. Segmentación de la médula ósea

En primer lugar, mediante el uso de una aplicación desarrollada en el GBT, se realiza una segmentación automática y posterior edición manual de la médula ósea del esqueleto axial. Esta segmentación incluye los huesos presentes en el protocolo de adquisición de imágenes para la evaluación de LF, es decir, desde mitad de fémur hasta el cráneo. Sin embargo, el cráneo se descarta para evitar artefactos causados por el movimiento. De esta forma, la segmentación final está formada por la columna cervicodorsolumbar, pelvis, parrilla costal con esternón, clavículas y escápulas junto con los húmeros y fémures. Esta segmentación se realiza sobre la imagen CT para obtener la región anatómica de la médula ósea de forma precisa y, posteriormente, se extrae la misma región de la imagen [ $^{18}\text{F}$ ]FDG PET para analizar cuantitativamente la captación de la médula.

### 2.2. Extracción de características

Con la máscara de la médula ósea sobre la imagen [ $^{18}\text{F}$ ]FDG PET, se extraen un total de 30 características. En concreto, se extraen 3 características de histograma o de primer orden, 9 características de segundo orden correspondientes a la matriz *Gray Level Co-occurrence Matrix* (GLCM) y 18 características de alto orden, perteneciendo 13 de ellas a la matriz *Gray Level Run Length Matrix* (GLRLM) y 5 a *Neighborhood Gray Tone Difference Matrix* (NGTDM).

Una vez extraídas las características texturales, se realiza un proceso de sobremuestreo para crear características texturales de nuevos casos a partir de los casos ya existente y equilibrar el número de casos IM+ e IM-, aumentando así la representatividad de la clase minoritaria, en este caso, IM+. La función utilizada es *Safe-Level-Synthetic Minority Over-Sampling Technique* (*Safe-Level-SMOTE*) [13] implementada en Python.

### 2.3. Análisis univariante

Las diferencias entre los grupos IM+ e IM- se evalúan mediante la prueba U de Mann-Whitney. Para la diferenciación de los grupos se cuenta con la valoración visual de las imágenes [ $^{18}\text{F}$ ]FDG PET realizada por los especialistas en Medicina Nuclear. Además, se calculan los coeficientes de correlación de Spearman ( $\rho$ ) y el  $p$ -valor de la correlación para evaluar la relación entre la magnitud de las características texturales y la presencia de IM. En ambos casos, un  $p$ -valor  $< 0,05$  se considera estadísticamente significativo, y se considera que la correlación es moderada cuando los coeficientes de correlación  $\rho$  alcanzan valores superiores a 0,5 o inferiores a -0,5, fuerte cuando superan a 0,7 o son inferiores a -0,7 y débil para el resto de los valores.

### 2.4. Modelos de *Machine Learning*

Partiendo de todas las características texturales extraídas a partir de las imágenes junto con la valoración visual de la presencia de IM, se realizan 6 modelos de *Machine Learning*: árbol de decisión; bosque aleatorio (*random forest*); máquinas de vectores de soporte (SVM, *Support Vector Machines*) con tres *kernels* distintos que son lineal, polinomial y función de base radial (FBR);  $k$ -vecino más cercano (K-NN, *K-Nearest Neighbours*), perceptrón multicapa (PMC) y algoritmo Naive-Bayes. Para el test se utiliza la función de validación cruzada con 5 *folds*.

Para la validación de los modelos, las métricas Exactitud (Ex), Precisión (Pr), Sensibilidad (S) y Especificidad (Es) son calculadas, considerando aceptables los valores superiores a 0,7 y excelentes los superiores a 0,9. Además, se obtiene la curva ROC (*Receiver Operating Characteristic*) con su correspondiente AUC (*Area Under the Curve*), considerando aceptables los valores inferiores a 0,3 o superiores a 0,7 y excelentes los inferiores a 0,1 y superiores a 0,9. Estos modelos son entrenados y evaluados con la base de datos original y con la sobremuestreada.

## 3. Resultados y discusión

### 3.1. Análisis univariante

Los resultados del análisis univariante aplicado a la base de datos original se muestran en la Tabla 1, incluyendo exclusivamente las características con  $p$ -valor de Mann-Whitney y de significancia de la correlación de Spearman  $< 0,05$ . Solo 2 de las 30 características obtienen resultados significativos en la diferenciación de IM+ e IM-, siendo ambas pertenecientes al grupo de características de histograma o de primer orden. Aunque se obtienen  $p$ -valores significativos para el cálculo de la correlación, el valor del coeficiente  $\rho$  para ambos casos indica una correlación débil.

	Mann-Whitney ( $p$ -valor)	Spearman ( $\rho$ )
Asimetría	0,014	0,3851
Kurtosis	0,034	0,3353

**Tabla 1.** Resultados del análisis univariante para la base de datos original.

Para el caso de la base de datos sobremuestreada, los resultados de los análisis de Mann-Whitney y correlación de Spearman se muestran en la Tabla 2. De nuevo, solo se muestran las características con  $p$ -valor de Mann-Whitney y de significancia de la correlación de Spearman  $<0,05$ . Al conseguir equilibrar el número de casos positivos y negativos, se observa la capacidad del análisis textural de las imágenes [ $^{18}\text{F}$ ]FDG PET en la diferenciación de los grupos IM+ e IM-, ya que 12 de las 30 obtienen diferencias estadísticamente significativas. Se puede destacar que las variables pertenecientes a las matrices GLCM y GLRLM obtienen  $p$ -valores inferiores a 0,009, considerablemente mejores que los de las características de histograma, mientras que los valores de  $\rho$  son similares en todos los casos, sin alcanzar ninguna de las variables valores asociados a correlaciones moderadas o fuertes.

	Mann-Whitney ( $p$ -valor)	Spearman ( $\rho$ )
Asimetría	0,014	0,385
Kurtosis	0,034	0,335
GLCM_Suma media	0,004	-0,380
GLCM_Autocorrelación	0,006	-0,364
GLRLM_LGRE	0,006	0,368
GLRLM_HGRE	0,007	-0,357
GLRLM_SRLGE	0,008	0,351
GLRLM_SRHGE	0,007	-0,355
GLRLM_LRLGE	0,001	0,440
GLRLM_GLV	0,006	-0,361
NGTDM_Ocupación	0,029	0,289
NGTDM_Fuerza	0,016	0,318

**Tabla 2.** Resultados significativos del análisis univariante para la base de datos sobremuestreada. LGRE: Low Gray Level Run Emphasis; HGRE: High Gray Level Run Emphasis; SRLGE: Short Run Low Gray Level Emphasis; SRHGE: Short Run High Gray Level Emphasis; LRLGE: Long Run Low Gray Level Emphasis; GLV: Gray Level Variance.

Además, se puede destacar que, aunque haya variables con  $p$ -valores inferiores a los de las características de histograma tras el sobremuestreado, Asimetría y Kurtosis son las únicas variables que consiguen resultados significativos para ambas bases de datos, destacando así su capacidad para diferenciar entre IM+ e IM-. Los resultados obtenidos coinciden con los de Parvez *et al.* [14], que realiza un análisis de las características de radiómica extraídos en los sitios de tumor representativos en pacientes con linfoma de no-Hodgkin, obteniendo una correlación significativa ( $p$ -valor = 0,035) entre la característica Kurtosis y la supervivencia global de estos pacientes.

### 3.2. Modelos de Machine Learning

Los resultados de los modelos basados en todas las características texturales obtenidas a partir de la base de datos original se muestran en la Tabla 3. Como se puede observar, no hay ningún modelo que obtenga valores aceptables para todas las métricas. De hecho, todos los

modelos obtienen valores aceptables e incluso excelentes de especificidad, y los que obtienen valores aceptables de AUC son  $<0,3$ , lo que quiere decir que predicen correctamente IM-. Mientras tanto, los valores de precisión y sensibilidad son considerablemente bajos, ya que predice los casos de IM+ como IM- para la mayoría de los sujetos. Partiendo de que la base de datos original está formada por 12 pacientes con IM+ y 29 con IM-, los modelos sufren un sobreentrenamiento para la clase negativa y no cuentan con suficiente información sobre la positiva y, por ello, varios de los modelos predicen todos los casos como IM-, como pasa con SVM polinomial, SVM FBR y KNN, alcanzando valores de sensibilidad de 0.

	Ex	Pr	S	Es	AUC
Á. de decisión	0,61	0,33	0,33	<b>0,72</b>	0,53
<i>Random Forest</i>	0,59	0,17	0,22	<b>0,76</b>	0,52
SVM Lineal	<b>0,76</b>	0,33	0,67	<b>0,93*</b>	<b>0,27</b>
SVM Polinomial	<b>0,71</b>	0	0	<b>1*</b>	<b>0,25</b>
SVM FBR	<b>0,71</b>	0	0	<b>1*</b>	0,32
K-NN	0,66	0	0	<b>0,93*</b>	0,33
PMC	0,63	0,25	0,33	<b>0,79</b>	0,50
Naive-Bayes	0,66	0,5	0,43	<b>0,72</b>	0,67

**Tabla 3.** Resultados de los modelos de Machine Learning basados en las características obtenidas de la base de datos original. Se muestra en negrita las métricas con valores aceptables ( $>0,7$  y, solo en AUC,  $<0,3$ ) y con \* las que obtienen valores excelentes ( $>0,9$  y, solo en AUC,  $<0,1$ ).

Los resultados de los modelos basados en la base de datos sobremuestreada se muestran en Tabla 4. Se puede observar que, aunque los valores de especificidad se ven reducidos en comparación con los obtenidos para la base de datos original, la mayoría de los modelos mejoran sus métricas al entrenarse con el mismo número de casos positivos que negativos. En este caso, el modelo de árbol de decisión y *random forest* consiguen valores aceptables para todas las métricas, obteniendo valores de AUC  $> 0,7$ , lo que indica que ambos modelos podrían ser buenos predictores en la diferenciación entre IM+ e IM-.

	Ex	Pr	S	Es	AUC
Á. de decisión	<b>0,78</b>	<b>0,83</b>	<b>0,75</b>	<b>0,72</b>	<b>0,78</b>
<i>Random Forest</i>	<b>0,76</b>	<b>0,76</b>	<b>0,76</b>	<b>0,76</b>	<b>0,83</b>
SVM Lineal	0,62	0,62	0,62	0,62	0,34
SVM Polinomial	0,28	0,24	0,26	0,31	<b>0,21</b>
SVM FBR	0,43	0,45	0,43	0,41	<b>0,19</b>
K-NN	0,43	0,38	0,42	0,48	0,45
PMC	0,41	0,45	0,42	0,38	0,36
Naive-Bayes	<b>0,71</b>	<b>0,76</b>	0,69	0,66	<b>0,77</b>

**Tabla 4.** Resultados de los modelos de Machine Learning basados en las características obtenidas de la base de datos sobremuestreada. Se muestra en negrita las métricas con valores aceptables ( $>0,7$  y, solo en AUC,  $<0,3$ ) y con \* las que obtienen valores excelentes.

La caracterización de estos modelos como los mejores, concretamente el modelo de *random forest* va en línea con el estudio de Li *et al.* [7], que para el mismo modelo basado en características de radiómica extraídas de la máscara del esqueleto, obtiene valores de sensibilidad, especificidad y precisión de entorno al 90% en la diferenciación de IM+ e IM- en pacientes con leucemia. Así, los estudios previos, junto con los resultados obtenidos, denotan la capacidad de los modelos para caracterizar la IM en LF, especialmente para el modelo de *random forest*. Sin embargo, sería necesario contar con más casos de IM+ para un entrenamiento adecuado y, obtener así, modelos más robustos. Además, los resultados se encuentran limitados al tipo de segmentación semiautomático utilizado, pudiendo mejorarse los resultados con segmentaciones de mayor precisión, automáticas basadas en aprendizaje profundo y/o completamente manuales realizadas por expertos.

#### 4. Conclusiones

La cuantificación de la presencia de IM es de vital importancia en la definición de la terapia de los pacientes con LF gracias a su carácter pronóstico. En el presente estudio se realiza un análisis de las características de radiómica extraídas de la máscara de la médula ósea. En primer lugar, se analizan las características individualmente y, posteriormente, se desarrollan modelos de *Machine Learning* basados en características texturales para la diferenciación entre IM+ e IM-. Se concluye que la metodología propuesta puede resultar de gran utilidad como sistema de ayuda a la decisión clínica.

#### Agradecimientos

Este trabajo está parcialmente financiado por una beca predoctoral (ayuda del Programa Propio de I+D+i 2020) de la Universidad Politécnica de Madrid.

#### Referencias

[1] A. Freedman and E. Jacobsen, “Follicular lymphoma: 2020 update on diagnosis and management,” *Am. J. Hematol.*, vol. 95, no. 3, pp. 316–327, Mar. 2020, doi: 10.1002/AJH.25696.

[2] A. Carbone *et al.*, “Follicular lymphoma,” *Nat. Rev. Dis. Prim.*, vol. 5, no. 1, Dec. 2019, doi: 10.1038/S41572-019-0132-X.

[3] H. Rajnai *et al.*, “Impact of the reactive microenvironment on the bone marrow involvement of follicular lymphoma,” *Histopathology*, vol. 60, no. 6B, May 2012, doi: 10.1111/J.1365-2559.2012.04187.X.

[4] T. C. Kwee, J. M. H. De Klerk, and R. A. J. Nieuvelstein, “Imaging of Bone Marrow Involvement in Lymphoma: State of the Art and Future Directions,” *Rev. TheScientificWorldJOURNAL*, vol. 11, pp. 391–402, 2011, doi: 10.1100/tsw.2011.40.

[5] S. F. Barrington *et al.*, “Role of imaging in the staging and response assessment of lymphoma: Consensus of the international conference on malignant lymphomas imaging working group,” *J. Clin. Oncol.*, vol. 32, no. 27, pp. 3048–3058, 2014, doi: 10.1200/JCO.2013.53.5229.

[6] L. Kostakoglu and S. Chauvie, “Metabolic Tumor

Volume Metrics in Lymphoma,” *Semin. Nucl. Med.*, vol. 48, no. 1, pp. 50–66, Jan. 2018, doi: 10.1053/J.SEMNUCLMED.2017.09.005.

[7] H. Li *et al.*, “<sup>18</sup>F-FDG PET/CT radiomic analysis with machine learning for identifying bone marrow involvement in the patients with suspected relapsed acute leukemia,” *Theranostics*, vol. 9, no. 16, pp. 4730–4739, 2019, doi: 10.7150/thno.33841.

[8] “MathWorks,” 2022. <https://es.mathworks.com> (accessed Jul. 13, 2022).

[9] M. Vallières, “Radiomics,” 2019. <https://github.com/mvallieres/radiomics> (accessed Jul. 18, 2022).

[10] M. Vallières, C. R. Freeman, S. R. Skamene, and I. El Naqa, “A radiomics model from joint FDG-PET and MRI texture features for the prediction of lung metastases in soft-tissue sarcomas of the extremities,” *Phys. Med. Biol.*, vol. 60, no. 14, pp. 5471–5496, 2015, doi: 10.1088/0031-9155/60/14/5471.

[11] “SPSS Statistics,” *IBM*, 2022. <https://www.ibm.com/products/spss-statistics> (accessed Jul. 18, 2022).

[12] Python.org, “Python Release Python 3.9.1,” 2020. <https://www.python.org/downloads/release/python-391/> (accessed Jun. 08, 2022).

[13] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, “SMOTE: Synthetic Minority Over-sampling Technique Nitesh,” *J. Artif. Intell. Res.*, vol. 16, pp. 321–357, 2002, doi: 10.1613/jair.953.

[14] A. Parvez, N. Tau, D. Hussey, M. Maganti, and U. Metser, “<sup>18</sup>F-FDG PET/CT metabolic tumor parameters and radiomics features in aggressive non-Hodgkin’s lymphoma as predictors of treatment outcome and survival,” *Ann. Nucl. Med.*, vol. 32, no. 6, pp. 410–416, 2018, doi: 10.1007/s12149-018-1260-1.

# Corrección automática del contraste en imagen radiográfica mediante aprendizaje profundo

D. Sanderson<sup>1,2</sup>, M. Desco<sup>1,2,3,4</sup>, C. F. Del Cerro<sup>1,2</sup>, M. I. García-Real<sup>5</sup>, M. J. Ruíz-Fernández<sup>5</sup>, J. García-Blas<sup>2,6</sup>, M. Abella<sup>1,2,3</sup>

<sup>1</sup> Departamento de Bioingeniería e Ingeniería Aeroespacial, Universidad Carlos III de Madrid, Leganés, España, {mabella@ing.uc3m.es}

<sup>2</sup> Laboratorio de Imagen Médica, Instituto de Investigación Sanitaria Gregorio Marañón, Madrid, España,

<sup>3</sup> Centro Nacional de Investigaciones Cardiovasculares Carlos III (CNIC), Madrid, España

<sup>4</sup> Centro de Investigación Biomédica en Red de Salud Mental (CIBERSAM), Madrid, España

<sup>5</sup> Departamento de Medicina y Cirugía Animal, Universidad Complutense, Madrid, España

<sup>6</sup> Departamento de Informática, Universidad Carlos III de Madrid, Leganés, España

## Resumen

*La imagen radiológica plana es una herramienta de gran utilidad diagnóstica debido a su bajo coste, interpretabilidad y facilidad de manejo. Sin embargo, presenta limitaciones debido a la escasa diferencia en atenuación entre tejidos blandos y a la reducción en contraste causado por la dispersión de los rayos X. El uso de herramientas de procesamiento de imagen permite mejorar la calidad de la imagen a la vez que se reduce la dosis de radiación ionizante recibida por el paciente. Los algoritmos convencionales dependen de muchos parámetros que requieren ajuste heurístico, suponen un alto coste computacional y se ven influenciados por el tipo de estudio. En este trabajo se propone un algoritmo de mejora de imagen rápido y automático basado en aprendizaje profundo. La evaluación preliminar, comparando con métodos de procesamiento tradicionales validados por un radiólogo, muestra la viabilidad del método propuesto.*

## 1. Introducción

La radiografía plana es la técnica de imagen más utilizada [1,2] debido a su facilidad de uso, rapidez y bajo coste. Sin embargo, a pesar de su valor clínico, ofrece poco contraste en tejido blando [3]. Se estima que el 93,7% de la dosis de radiación media utilizada para usos médicos se debe al diagnóstico con rayos X [4], por lo que es necesario minimizar la dosis de estos exámenes alcanzando un compromiso con el nivel de ruido resultante en la imagen. Por ello, es indispensable el uso de algoritmos de procesado que permitan mejorar tanto el contraste como la relación señal-ruido de la imagen sin perder resolución espacial (ver Figura 1).



**Figura 1.** Radiografía de tórax de gato en datos crudos (izquierda) y procesada mediante algoritmos convencionales (derecha).

Para mejorar el contraste, se han propuesto métodos basados en la ecualización del histograma, siendo la ecualización adaptativa del histograma (CLAHE) uno de los más utilizados [5-7]. Una de sus principales limitaciones es el potenciamiento del ruido [8]. Para reducir el nivel de ruido, sin perder detalle en la imagen, es común utilizar la pirámide laplaciana o la transformada wavelet [9]. Sin embargo, ambos métodos tienen un alto coste computacional proporcional al número de niveles de escala utilizados, dificultando el procesamiento en tiempo real. En general, los métodos de procesamiento de imagen convencionales dependen de numerosos parámetros que requieren ajuste heurístico específico para el tipo de estudio e implican un alto coste computacional [10].

En los últimos años se han aplicado técnicas de aprendizaje profundo a la mejora de imagen radiológica. Frejlichowski et al. combinaron una RNC (Red Neuronal Convolutiva) con imágenes residuales e interpolación bicúbica para mejorar la calidad de la imagen, aunque en sus resultados se puede apreciar un emborronamiento de la imagen [11]. Krouma et al. utilizaron una red neuronal para aplicar OFAs (Ordenes Fraccionarios Adaptativos), pero los resultados presentan ruido [12].

Un método de aprendizaje profundo que integre tanto la mejora de contraste como la eliminación de ruido permitiría realizar un procesamiento plenamente automático en tiempo real aplicable a la práctica clínica y sin necesidad de ajuste heurístico de parámetros. Este trabajo propone el uso de una red U-NET basada en ResNet-34, evaluando dos funciones de coste, Índice de Similitud Estructura MultiEscala (MS SSIM) [13] y Perceptual Loss (PL) [14], comparando los resultados con los obtenidos con los algoritmos de procesado más usados en imagen radiológica veterinaria.

## 2. Materiales y métodos

### 2.1. Generación de la base de datos

Se han utilizado 24 imágenes de gato, 34 imágenes de perro y 2 de tortuga, todas de distintas partes anatómicas y adquiridas en el Hospital Clínico Veterinario de la

Universidad Complutense de Madrid. Para acelerar la convergencia del entrenamiento y evitar variaciones bruscas de los gradientes de la primera capa de la red, se han normalizado las imágenes de entrada a un rango de intensidades comprendido entre 0 y 1. Del total de imágenes, se han elegido de forma aleatoria 48 para el conjunto de entrenamiento y 12 para el de validación, de las cuales 4 imágenes son de perro, 6 de gato, y 1 de tortuga. Debido al limitado número de imágenes y al carácter preliminar del trabajo, no hemos utilizado un conjunto de datos de “test”.

Lás imágenes de referencia con las que se ha entrenado la red se han obtenido con un método de procesamiento basado en algoritmos convencionales y validado por un radiólogo experto. El primer paso es la segmentación del fondo. Se aplica un umbral de intensidad seleccionado de forma heurística evitando usar algoritmos de segmentación automática para minimizar el tiempo de ejecución. A continuación, se realiza una mejora de contraste mediante el algoritmo CLAHE, que aplica una ecualización de histograma por parches de 500×500 píxeles obtenidos sin solape. Para eliminar los artefactos resultantes de aplicar la ecualización a nivel local, se aplica interpolación bilineal a cada parche utilizando los histogramas de los parches circundantes. Finalmente se realiza un realce de bordes con limitación de ruido mediante una pirámide laplaciana de 3 niveles, construida de la siguiente forma:

$$I_{enhanced} = \sum_{i=1}^n w_i (I_i - (I_{i+1})^{up})$$

$$I_i = (I_{i-1} * g(\sigma))_{down}$$

donde  $w_i$  es el peso que se le aplica al nivel  $i$  de la pirámide (4,4,1),  $I$  es la imagen,  $g(\sigma)$  es un kernel gaussiano con desviación estándar ( $\sigma$ ) igual a 1, y  $up$  y  $down$  son interpolaciones de la imagen hacia mayores o menores resoluciones, respectivamente. Para limitar el ruido se aplica un umbral igual al 5% del máximo a cada una de las imágenes de bordes de la pirámide.

Este método basado en algoritmos convencionales se ha implementado íntegramente en lenguaje C partiendo de un prototipo basado en Matlab. Dada su naturaleza, el algoritmo se ejecuta en CPU en su totalidad. Para reducir el tiempo de ejecución, se ha desarrollado una implementación multi hilo basada en OpenMP. El tiempo de ejecución del algoritmo convencional es de 2 segundos para imágenes de 4000×4000 píxeles por lo que no se ha considerado acelerar la implementación en dispositivos tipo GPU.

## 2.2. Red Convolutiva

Se ha utilizado una red U-NET [15] basada en la arquitectura Resnet-34 preentrenada con la base de datos conocida como *ImageNet* [16]. Para poder aplicar la red a imágenes radiológicas en niveles de gris, se han sumado los pesos de los tres canales de color de la primera capa convolutiva. Para normalizar la salida de la red y evitar la aparición de picos de intensidades en la imagen resultante, se ha utilizado un sigmoide con un rango dinámico de [0,1].

## 2.3. Entrenamiento y validación de la red

Se han implementado dos modelos diferentes, entrenando la red U-NET con las funciones de coste MS-SSIM, que identifica información estructural de la imagen, y PL, que evalúa tanto el contenido como el aspecto de la imagen. La evaluación se ha hecho con una tarjeta gráfica NVIDIA GeForce RTX 3090 con 24 GB de memoria.

Debido a la limitación de memoria, el tamaño máximo de matriz de entrada para el entrenamiento es de 1024×1024 píxeles. Dado que las radiografías tienen un tamaño de matriz de entre 1288×950 y 4188×4188 píxeles, se dividen en parches obtenidos mediante un recorte aleatorio de la imagen, manteniendo la resolución original. Estos parches no tienen información del contraste general, lo que limita la capacidad de la red para extrapolar sus resultados a imágenes completas. Por ello, se ha incluido un paso de entrenamiento con imágenes completas submuestreadas a un tamaño de 1024×1024 píxeles. Evaluamos tres posibles escenarios para determinar la mejor estrategia de entrenamiento:

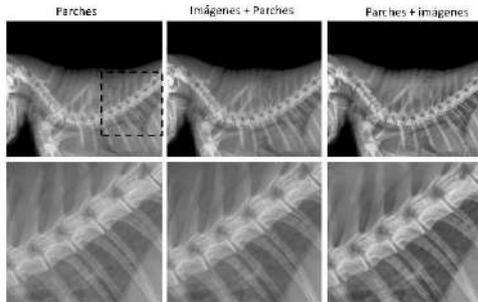
- 1- Entrenamiento con parches de resolución original.
- 2- Entrenamiento con imágenes originales de menor resolución y reentrenamiento con parches de resolución original.
- 3- Entrenamiento con parches de resolución original y reentrenamiento de la red con imágenes originales de menor resolución.

Cada fase del entrenamiento ha tenido una duración de 100 épocas. Para el entrenamiento se ha utilizado un optimizador Adam [17], ya que es uno de los optimizadores más robustos y permite sortear la mayoría de los mínimos locales. Para acelerar el entrenamiento de la red, se ha aplicado un programador de tasa de aprendizaje basado en tasas de aprendizaje cíclicas [18]. Se ha entrenado con un decaimiento de peso de  $10^{-3}$  y una tasa inicial de aprendizaje de 0,052 seleccionada mediante la prueba de Leslie N. Smith [18]. Para compensar el reducido número de imágenes disponibles se ha aplicado un aumento de datos al conjunto de entrenamiento consistente en rotaciones y volteos aplicados a la imagen de entrada con una probabilidad de 0.5. La implementación se ha hecho utilizando la API de Pytorch Fastai 2.7.5 [19].

Para evaluar la red en imágenes completas, se ha aplicado un relleno en forma de espejo ajustando el tamaño de las imágenes al múltiplo de 32 más próximo a la dimensión original, para evitar el redondeo en la reducción de tamaño en el codificador. Se ha seleccionado un tamaño máximo de 3904×3904 debido a la limitación de memoria de la tarjeta gráfica, lo cual ha implicado un pequeño recorte de bordes de 7 imágenes que no ha implicado una pérdida de información útil. Para validar los resultados se han obtenido perfiles y se han utilizado las métricas de Error Cuadrático Medio (MSE), Error Medio Absoluto (MAE), MS-MSSIM y PL. Los valores han sido normalizados con el resultado de aplicar las métricas a la imagen original y a la imagen referencia.

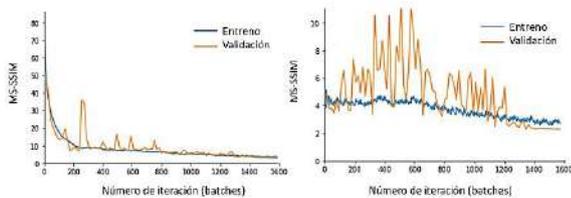
### 3. Resultados

La Figura 2 muestra el resultado de las tres estrategias de entrenamiento en una imagen de tórax de gato, donde se puede observar una mayor recuperación de contraste tanto en tejido blando como en hueso con la tercera estrategia.



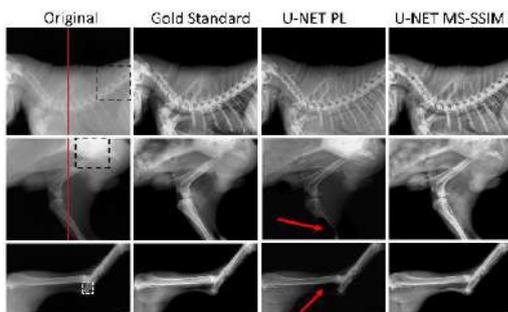
**Figura 2.** Resultados del modelo con MS-SSIM (arriba) en una imagen de tórax de gato, con sus respectivos aumentos (abajo), para un entrenamiento de 40 épocas con las 3 estrategias de izquierda a derecha.

La Figura 3 muestra oscilaciones en la función de coste para el conjunto de validación hasta la época número 60 (1200 batches), momento en el que la red completa su aprendizaje para imágenes completas. Por tanto, esta ha sido la estrategia de entrenamiento que se ha seleccionado para el resto de los experimentos.



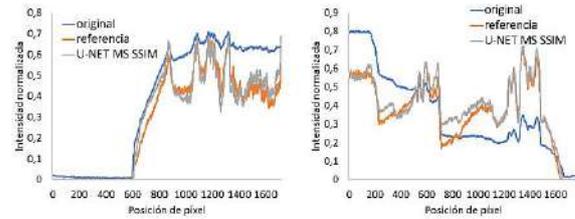
**Figura 3.** Evolución de la función de coste MS-SSIM para los conjuntos de entreno y validación para las dos fases del entrenamiento: con parches de alta resolución (izquierda) y con imágenes completas de baja resolución (derecha).

La Figura 4 muestra los resultados de los dos modelos propuestos. El uso de MS-SSIM es capaz de recuperar un mayor contraste y de corregir saturaciones de imagen y recuperar niveles de contraste equivalentes a los del resto de la imagen (segunda fila de la Figura 4). El uso de PL tiene un realce de contraste menor e introduce artefactos de imagen, principalmente en las zonas de bajo contraste situadas en los bordes entre el sujeto y el fondo (ver artefacto en el hueso indicado por las flechas en la segunda y tercera fila de la Figura 4).



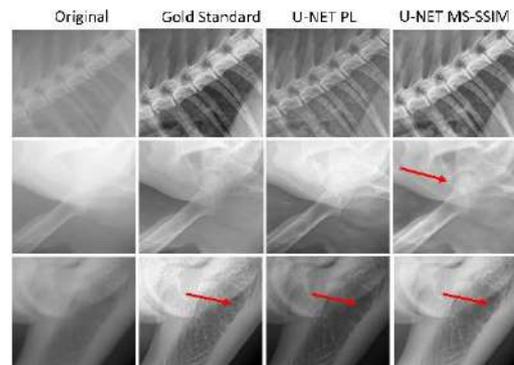
**Figura 4.** Imagen de tórax de gato (primera fila), y de las zonas proximal y distal de extremidad posterior de perro (segunda y tercera fila) con los dos modelos.

Los perfiles de la Figura 5 muestran cómo el método propuesto consigue recobrar los valores de intensidad.



**Figura 5.** Perfiles de tórax de gato (izquierda) y pata de perro (derecha) indicados en la Figura 4 con una línea de color.

Como limitación del método propuesto, se puede apreciar una ligera pérdida de resolución espacial (ver flechas en la tercera fila de la Figura 6).



**Figura 6.** Detalles de las regiones indicadas por los recuadros de la Figura 4.

Los resultados de las métricas de evaluación mostrados en la Tabla 1 confirman las observaciones visuales y demuestran que la red entrenada con MS-SSIM obtiene mejores resultados. El tiempo medio de ejecución del modelo es de 0.143 segundos.

**Tabla 1.** Valores normalizados de las métricas de evaluación (en negrita se resaltan los mejores resultados).

Función de coste	Conjunto de datos	MAE	MSE	MS-SSIM	PL
PL	Entrenamiento	0,76	0,68	0,85	0,05
	Validación	0,99	1,10	0,93	<b>0,04</b>
MS-SSIM	Entrenamiento	0,42	0,27	0,36	0,05
	Validación	<b>0,47</b>	<b>0,34</b>	<b>0,32</b>	<b>0,04</b>

### 4. Discusión

Este trabajo propone el uso de la red convolucional U-NET como un método de mejora de imagen radiológica automático, con dos funciones de coste diferentes. Debido al límite de memoria del acelerador GPU utilizado, hemos dividido las imágenes en parches manteniendo la resolución original. Sin embargo, al entrenar sólo con parches de alta resolución, la red no es capaz de recuperar el contraste en imágenes de tórax o abdomen, debido a la heterogeneidad de hueso y tejido blando y debido al reducido campo de visión de los parches. El incluir un paso de reentrenamiento con imágenes completas de menor resolución, permite que la red se ajuste a mayores heterogeneidades de contraste y que mejore los resultados en imagen completa.

Los mejores resultados se obtienen para el modelo con la función de coste MS-SSIM. Esto puede deberse a que evalúa tanto el contraste y la estructura como la luminancia, lo cual actúa como regularizador para evitar descompensaciones de intensidad en la imagen. Sin embargo, para permitir su introducción en la clínica, es necesario mejorar la resolución de las imágenes de salida. La función de coste PL ha mostrado un peor realce de contraste, con un carácter más artificial y con pérdidas de resolución. Esto podría deberse a que la PL incorpora un factor de regularización para reducir la aparición de ruido, tal y como se explica en [13], lo cual puede dificultar la recuperación de los bordes en caso de una excesiva regularización. Como trabajo futuro, se explorarán distintas formas de mejorar la resolución de la imagen. Utilizaremos valores menores del factor de regularización, y combinaremos los modelos con otras funciones de coste como MixGE [20] o Wavelets [21]. Además, se evaluará el reentrenamiento de la red U-NET mediante una Red Discriminadora en un formato GAN.

Los resultados preliminares muestran la viabilidad del método propuesto para el procesamiento automático de imagen en la práctica clínica veterinaria en tiempo real, válido para todo tipo de estudios y liberando al radiólogo de la necesidad de optimizar parámetros de forma heurística.

## Agradecimientos

Este trabajo ha sido apoyado por el Ministerio de Ciencia e Innovación, Agencia Estatal de Investigación: PDC2021-121656-I00 (MULTIRAD), financiado por MCIN/AEI/10.13039/501100011033 y por la Unión Europea 'NextGenerationEU'/PRTR; PID2019-110369RB-I00/AEI/10.13039/501100011033 (RADHOR). También ha sido financiada por la Comunidad de Madrid: Convenio plurianual con la UC3M en la línea de 'Fomento de la Investigación de Jóvenes Doctores' (DEEPCM-CM-UC3M), y en el contexto del V PRICIT (Programa Regional de Investigación e Innovación Tecnológica); S2017/BMD-3867 RENIM-CM, cofinanciado por el Fondo Estructural y de Inversión Europeo. Y también financiado parcialmente por CRUE Universidades, CSIC y Banco Santander (Fondo Supera Covid19), proyecto RADCOV19. El CNIC cuenta con el apoyo del Instituto de Salud Carlos III, el Ministerio de Ciencia e Innovación y la Fundación Pro CNIC.

## Referencias

- [1] Bhargavan, M., & Sunshine, J. H. Utilization of radiology services in the United States: levels and trends in modalities, regions, and populations. *Radiology*, vol 234, sup 3, 2005.
- [2] Smith-Bindman, R., Miglioretti, D. L., & Larson, E. B. Rising use of diagnostic medical imaging in a large integrated health system. *Health affairs*, vol 27, sup 6, 2008, pp 1491-1502.
- [3] Qiu, J., Harold Li, H., Zhang, T., Ma, F., & Yang, D. Automatic x-ray image contrast enhancement based on parameter auto-optimization. *Journal of applied clinical medical physics*, vol 18, sup 6, 2017, pp 218-223.
- [4] Consejo de Seguridad Nuclear. Dosis de radiación. 2010, SDB-04.07
- [5] Koonsanit, K., Thongvigitmanee, S., Pongnapang, N., & Thajchayapong, P. Image enhancement on digital x-ray images using N-CLAHE. *10th Biomedical engineering international conference (BMEICON)*, 2017, pp. 1-4.
- [6] Saenpaen, J., Arwathananukul, S., & Aunsri, N. (2018, July). A comparison of image enhancement methods for lumbar spine X-ray image. *15th International Conference on Electrical Engineering/Electronics*, 2018, pp. 798-801.
- [7] Qiu, J., Harold Li, H., Zhang, T., Ma, F., & Yang, D. Automatic x-ray image contrast enhancement based on parameter auto-optimization. *Journal of applied clinical medical physics*, vol 18, sup 6, 2017, pp 218-223.
- [8] Lidong, H., Wei, Z., Jun, W., & Zebin, S. Combination of contrast limited adaptive histogram equalisation and discrete wavelet transform for image enhancement. *IET Image Processing*, vol 9, sup 10, 2015, 908-915.
- [9] Frejlichowski, D., & Wanat, R. Application of the Laplacian pyramid decomposition to the enhancement of digital dental radiographic images for the automatic person identification. *In International Conference Image Analysis*, 2010 pp. 151-160.
- [10] Ackar, H., Abd Almisreb, A., & Saleh, M. A. A review on image enhancement techniques. *Southeast Europe Journal of Soft Computing*, 2019, vol 8 sup 1.
- [11] Sun, Y., Li, L., Cong, P., Wang, Z., & Guo, X. Enhancement of digital radiography image quality using a convolutional neural network. *Journal of X-ray Science and Technology*, 2017, vol 25
- [12] Krouma, H., Ferdi, Y., & Taleb-Ahmedx, A. Neural adaptive fractional order differential based algorithm for medical image enhancement. *In 2018 International Conference on Signal, Image, Vision and their Applications (SIVA)*, 2018, pp. 1-6
- [13] Wang, Z., Simoncelli, E. P., & Bovik, A. C. Multiscale structural similarity for image quality assessment. *In The Thirty-Seventh Asilomar Conference on Signals, Systems & Computers*, 2003, pp. 1398-1402.
- [14] Johnson, J., Alahi, A., & Fei-Fei, L. Perceptual losses for real-time style transfer and super-resolution. *In European conference on computer vision*, 2016, pp. 694-711.
- [15] Ronneberger, O., Fischer, P., & Brox, T. U-net: Convolutional networks for biomedical image segmentation. *In International Conference on Medical image computing*, 2015, pp. 234-241.
- [16] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, et al., "Imagenet large scale visual recognition challenge," *International journal of computer vision*, vol. 115, pp. 211-252, 2015.
- [17] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014
- [18] Smith, L. N. Cyclical learning rates for training neural networks. *IEEE winter conference on applications of computer vision (WACV)* 2017, pp. 464-472.
- [19] Howard, J., & Gugger, S. Fastai: a layered API for deep learning. *Information*, vol 11, sup 2, 2020.
- [20] Lu, Z., & Chen, Y. Single image super-resolution based on a modified U-net with mixed gradient loss. *Signal, image and video processing*, 2021, pp 1-9.
- [21] H. -H. Yang, C. -H. H. Yang, Y. -C. James Tsai, "Y-Net: Multi-scale feature aggregation network with wavelet structure similarity loss function for single image dehazing", *2020 IEEE International Conference on Acoustics (ICASSP)*, 2020.

# Mecanismo de atención y *deep learning* en la evaluación automática de la calidad de las retinografías

R. Romero-Oraá<sup>1,2</sup>, M. García<sup>1,2</sup>, M. I. López Gálvez<sup>1,2,3,4</sup>, M. Herrero<sup>1</sup>, R. Hornero Sánchez<sup>1,2,5</sup>

<sup>1</sup> Grupo de Ingeniería Biomédica, Universidad de Valladolid, Valladolid, España, roberto.romero@gib.tel.uva.es, {margar,robhor}@tel.uva.es

<sup>2</sup> Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), España

<sup>3</sup> IOBA (Instituto de Oftalmobiología Aplicada), Universidad de Valladolid, España

<sup>4</sup> Hospital Clínico Universitario de Valladolid, Valladolid, España, maribel@ioba.med.uva.es

<sup>5</sup> IMUVA, Instituto de Investigación en Matemáticas de la Universidad de Valladolid, Valladolid, España

## Resumen

*La evaluación de la calidad de las retinografías es una etapa fundamental en los sistemas automáticos de cribado de distintas patologías oculares, como son la retinopatía diabética, el glaucoma, las cataratas o la degeneración macular asociada a la edad. El objetivo de este trabajo fue desarrollar un método basado en deep learning y un mecanismo de atención para clasificar las retinografías en dos clases según su nivel de calidad. Se empleó la arquitectura InceptionResNetV2 como modelo base en combinación con transfer learning, fine-tuning y data augmentation. El método se entrenó con 2348 retinografías de una base de datos privada y se evaluó con 425 imágenes de la misma base de datos, alcanzando 96.00% de precisión. También se utilizó la base de datos DRIMDB, compuesta por 194 imágenes, para validar el método, obteniendo una precisión del 99.48%. Finalmente, se estudió la interpretación del método, en el contexto de Explainable Artificial Intelligence, obteniendo mapas de atención.*

## 1. Introducción

En el campo de la oftalmología, la retinografía o imagen de fondo de ojo es la modalidad de imagen retiniana más establecida. Su uso se extiende al diagnóstico precoz y mejorado de múltiples patologías oculares, como son la retinopatía diabética, el glaucoma, las cataratas o la degeneración macular asociada a la edad [1]. Este diagnóstico se lleva a cabo analizando los cambios morfológicos del fondo de ojo. Sin embargo, la inspección manual de estas imágenes requiere mucho tiempo, esfuerzo y costes [2]. La escasez de oftalmólogos y el gran número de pacientes evidencian la necesidad de sistemas automáticos de cribado para asistir a los médicos [1]. De esta manera, sería posible aliviar su carga de trabajo y los costes asociados ofreciendo un diagnóstico rápido y objetivo [3].

Sin embargo, la fiabilidad del diagnóstico depende completamente de la calidad de las retinografías [1]. Si la imagen de entrada no presenta una calidad adecuada, podrían darse diagnósticos erróneos [3]. La realidad es que, en entornos clínicos, un elevado porcentaje de las imágenes carece de suficiente calidad para su análisis [4]. La presencia de polvo en la cámara, los parpadeos o la oclusión por pestañas son algunos de los factores que producen distorsiones en la imagen tales como iluminación

desigual, bajo contraste, desenfoque o destellos de luz. Por este motivo, la evaluación de la calidad de la imagen es una etapa esencial en los sistemas automáticos de cribado y ha recibido mucha atención por parte de la comunidad científica en los últimos años [1].

En este contexto, pueden encontrarse en la literatura varios métodos para determinar de forma automática la calidad de las retinografías. Algunas propuestas se basan en extraer características globales de la imagen [5], [6] mientras que otras se basan en un análisis estructural para identificar puntos de referencia [3], [7]. En los últimos años, las arquitecturas *deep learning*, como las redes neuronales convolucionales (CNN), se han empleado en multitud de problemas de visión artificial, incluida la evaluación de la calidad de las retinografías [1]. En este sentido, se han explorado distintas arquitecturas CNN: Alexnet, VGG19, ResNet18, ResNet50, DenseNet121 y Inception v3 [2], [4], [8]–[10]. En uno de nuestros trabajos previos comprobamos que la arquitectura InceptionResNetV2, junto con *transfer learning* y *fine-tuning*, ofrece resultados muy prometedores para nuestro propósito [11]. Asimismo, los mecanismos de atención han supuesto un avance significativo en las arquitecturas *deep learning* [12]. En un intento de imitar a los humanos, estos mecanismos permiten encontrar las regiones más significativas de la imagen y optimizar la tarea de clasificación centrando la atención del modelo en los píxeles relevantes. Además de esta optimización, los mecanismos de atención proporcionan un mapa de atención que permite interpretar los resultados obtenidos y explicar el comportamiento del modelo, contribuyendo así en el campo de *Explainable Artificial Intelligence* (XAI) [12].

En este trabajo partimos de la hipótesis de que los mecanismos de atención pueden mejorar la evaluación automática de la calidad de las retinografías. El objetivo fue desarrollar un modelo basado en *deep learning* y un mecanismo de atención para clasificar las retinografías en dos clases: calidad adecuada y calidad inadecuada. Se utilizó la arquitectura InceptionResNetV2 como modelo base en combinación con *transfer learning*, *fine-tuning* y *data augmentation*. Hasta donde sabemos, esta solución nunca ha sido explorada. Otra contribución del trabajo son los mapas de atención obtenidos, que permiten señalar las

regiones concretas de la imagen donde la calidad es insuficiente. Para desarrollar el método se utilizó una base de datos privada y se validó posteriormente con la base de datos pública DRIMDB [13].

## 2. Bases de datos de retinografías

En este estudio se utilizó una base de datos privada para la configuración del modelo. Estaba formada por 2773 retinografías proporcionadas por el Instituto de Oftalmobiología Aplicada (IOBA) de la Universidad de Valladolid y el Hospital Clínico Universitario de Valladolid. Estas imágenes se capturaron con el retinógrafo no midriático Topcon TRC-NW400, con un campo de visión de 45°, a 1956×1934 píxeles de resolución y en formato JPEG de 24 bits. Un especialista determinó, para cada una, si tenían o no calidad suficiente para su análisis. Se dividieron en dos conjuntos. El conjunto de entrenamiento (2348 imágenes, 85%) contenía 1778 imágenes de calidad adecuada y 570 imágenes de calidad inadecuada. El conjunto de test (425 imágenes, 15%) contenía 314 imágenes de calidad adecuada y 111 imágenes de calidad inadecuada. Asimismo, se empleó la base de datos pública DRIMDB [13] (194 imágenes de 760×570 píxeles de resolución) para validar la robustez del modelo. Esta base de datos está formada por 125 imágenes de calidad adecuada y 69 imágenes de calidad inadecuada.

## 3. Métodos

En esta sección se describe, en primer lugar, la etapa de preprocesado inicial. A continuación, se explican las técnicas *data augmentation*, *transfer learning* y *fine-tuning*. Después se detalla la arquitectura CNN utilizada, que incluye el mecanismo de atención propuesto. Finalmente, se especifica la fase de entrenamiento.

### 3.1. Preprocesado

En esta etapa, se redimensionaron, en primer lugar, todas las imágenes a una resolución de 512×512 píxeles [2]. Por tanto, la resolución original de las imágenes es irrelevante. Después, las imágenes se normalizaron en el intervalo [-1,1] para facilitar el aprendizaje de la red [2].

### 3.2. Data augmentation

Las redes neuronales profundas requieren un elevado número de datos de entrenamiento [2]. Para aumentar el número de imágenes con las que entrenar la CNN de este trabajo, se empleó la técnica de aumentación de datos en tiempo real. Esta técnica permite generar nuevas imágenes

aleatorias exclusivas para cada época de entrenamiento [2]. De esta manera, el modelo se entrenó con un conjunto diferente de 2348 imágenes para cada época, generadas a partir del conjunto de entrenamiento original. Estas retinografías sintéticas se obtuvieron aplicando transformaciones simples [2]:

- Rotaciones aleatorias en el rango [-20, +20] grados.
- Desplazamientos verticales y horizontales hasta un máximo de un 7% de la anchura de la imagen.
- Volteos verticales y horizontales aleatorios.
- Escalado entre 0.85 y 1.15.

Disponer de una mayor base de datos para entrenar ofrece mejores resultados y aumenta la capacidad de generalización del método [2].

### 3.3. Transfer learning y fine-tuning

*Transfer learning* consiste en resolver un problema de *machine learning* en un dominio de interés particular con el conocimiento aprendido de los datos de entrenamiento de otro dominio de interés [14]. De esta manera, no es necesario disponer de un conjunto de imágenes tan amplio como si se tuviera que entrenar la red desde cero. Para aplicar esta técnica, se utilizó un modelo base preentrenado con las imágenes del proyecto ImageNet [15]. Esta base de datos se compone de más de 14 millones de imágenes pertenecientes a más de 20000 clases distintas.

A continuación, se aplicó la técnica *fine-tuning*, que consiste en reentrenar la red preentrenada utilizando un conjunto de imágenes de entrenamiento adaptado al problema bajo estudio. De esta manera, los pesos se adaptan mejor a cada caso particular. En nuestro caso, se entrenaron todas las capas de nuestra CNN.

### 3.4. Arquitectura CNN

Las CNN son redes neuronales compuestas por capas convolucionales y, opcionalmente, capas de *pooling* y *fully-connected*, entre otras. Son capaces de extraer de forma optimizada características representativas de un gran conjunto de imágenes en crudo [2], [4]. En este trabajo se desarrolló una arquitectura CNN compuesta por 3 bloques: un modelo base, un mecanismo de atención y un conjunto de capas *fully-connected* a la salida de la red. La Figura 1 muestra el esquema de esta arquitectura. Como modelo base se empleó la arquitectura InceptionResNetV2, que actúa de extractor de características. Esta arquitectura híbrida es una combinación de Inception [16] y ResNet

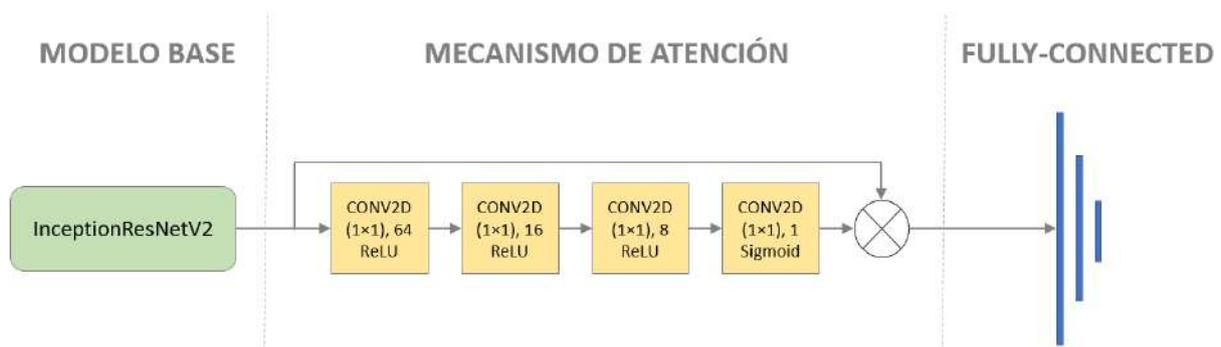


Figura 1. Esquema de la arquitectura CNN propuesta.

[17] y ha demostrado un gran rendimiento para evaluar la calidad de las retinografías [11]. El mecanismo de atención toma la matriz de características obtenida por el modelo base y se compone de 4 capas convolucionales 1x1 en serie, de 64, 16, 8 y 1 filtros cada una. Las 3 primeras capas tenían la función de activación ReLu y la última capa la función de activación sigmoide, de manera que su salida produce un mapa de atención cuyos píxeles representan el grado de relevancia en el rango [0-1] [12]. Para aplicar este mapa de atención a las características extraídas por el modelo base se utilizó la multiplicación elemento a elemento [12]. A continuación, se aplicó la operación *global average pooling* y, después, la técnica *dropout* para evitar el sobreentrenamiento [1]. Finalmente, la arquitectura CNN se cerró con 3 capas *fully-connected* de 1024, 512 y 1 neuronas cada una. Las dos primeras contaban con una función de activación ReLU [2]. Para la última, se empleó una función de activación sigmoide, que proporciona una salida entre 0 y 1 y puede interpretarse como una probabilidad a posteriori en nuestro problema de clasificación binaria [11].

### 3.5. Entrenamiento

Se entrenó el modelo CNN propuesto durante 100 épocas [10]. Se aplicó la entropía cruzada binaria como función de pérdidas y el descenso de gradiente estocástico como algoritmo de optimización [2], [4], [10]. La tasa de aprendizaje se fijó en 0.001 y el valor de *momentum* en 0.9 [2], [18]. Para evitar sobreentrenamiento en épocas avanzadas, se redujo la tasa de aprendizaje en un factor de 10 cada vez que el error de validación alcanzase un mínimo y se mantuviese constante [2], [18]. Además, se estableció un tamaño de *batch* de 16 imágenes [2].

## 4. Resultados

El entrenamiento de la CNN se llevó a cabo con las 2348 imágenes del conjunto de entrenamiento de la base de datos privada. Se utilizó un Workstation con procesador Intel Xeon CPU E5-1620 v4 @ 3.5GHz × 8, 32GB de RAM y dos tarjetas gráficas NVIDIA TITAN X (Pascal).

El método se evaluó sobre el conjunto de test de la base de datos privada y el total de imágenes de la base de datos pública DRIMDB [13]. La Tabla 1 recoge los resultados en términos de precisión (PR), sensibilidad (SE), especificidad (ES), F1-score y área bajo la curva ROC (AUC). Asimismo, en la siguiente sección se discuten algunos de los mapas de atención obtenidos.

## 5. Discusión

En este trabajo se ha propuesto un nuevo método basado en *deep learning* para evaluar la calidad de las retinografías. Como principal contribución, hemos incluido un mecanismo de atención como parte de la arquitectura. Este tipo de mecanismos se ha utilizado

Base de datos	PR (%)	SE (%)	ES (%)	F1	AUC
Privada	96.00	98.09	90.09	0.9224	0.9438
DRIMDB	99.48	99.20	100.00	0.9928	0.9960

**Tabla 1.** Resultados sobre el conjunto de test de la base de datos privada y sobre la base de datos pública DRIMDB.

previamente en el análisis de retinografías pero, hasta donde sabemos, no se ha utilizado en el contexto de calidad de la imagen. El método se ha desarrollado utilizando exclusivamente retinografías de una base de datos privada. Sin embargo, se ha obtenido una alta precisión con imágenes distintas procedentes de una base de datos pública, lo que demuestra su robustez. Los resultados de la Tabla 1 indican que el método propuesto ofrece un gran rendimiento para evaluar la calidad de las retinografías.

Estos resultados son similares a los obtenidos en estudios previos, como se muestra en la Tabla 2. Al considerar la base de datos privada, el método propuesto (PR=96.24%; SE=98.09%; ES=90.09%) supera nuestro antiguo trabajo en términos de precisión (95.29%) y sensibilidad (96.82%) pero no de especificidad (91.00%). Al comparar nuestros resultados para la base de datos DRIMDB con los obtenidos en otros estudios para la misma base de datos, observamos que nuestro método iguala nuestro antiguo trabajo (PR=99.48%; SE=99.20%; ES:100%) y supera al resto de trabajos de la literatura [2], [6], [18]–[20]. En otros trabajos se han utilizado bases de datos distintas, lo que influye sobre los resultados obtenidos y dificulta la comparación directa entre estudios. Destaca el trabajo de Saha et ál. [2], donde se obtuvo un 100% de precisión. Sin embargo, descartaron para el estudio las imágenes de calidad ambigua, que son las más propensas a error.

Otra contribución importante, en el contexto de XAI, son los mapas de atención generados, que se validaron de forma cualitativa. Estos mapas representan la relevancia de cada zona de la imagen para la toma de decisión del modelo. La Figura 2 muestra algunos ejemplos de estos mapas. En el primer ejemplo se aprecia que el modelo identifica correctamente la mancha negra. En el segundo ejemplo, se puede observar que el modelo encuentra relevante el borde superior, que carece de contraste.

Este estudio también presenta algunas limitaciones. En primer lugar, se ha llevado a cabo la validación del método con un número relativamente pequeño de imágenes. En el futuro sería deseable ampliar el número de imágenes de test. En segundo lugar, la sensibilidad y la especificidad han resultado desbalanceadas. Sería deseable aplicar estrategias como la función de pérdidas focal o la ponderación de clases para balancear esas métricas. Por último, el mecanismo de atención aplicado puede verse

Estudio	Base de datos	PR (%)	SE (%)	ES (%)
Yu et al. 2017 [8]	EyePACS	95.42	96.63	93.10
Saha et al. 2018 [4]	EyePACS	100	100	100
Zago et al. [2]	DRIMDB	98.55	97.10	100
Shao et al. [6]	DRIMDB	89.10	89.30	94.10
Chalakkal et al. [18]	DRIMDB	97.77	97.60	97.80
Karlsson et al. [19]	DRIMDB	98.10	99.30	95.80
Abramovich et al. [20]	DRIMDB	99.00	98.40	100
Romero-Oraá et al. [11]	Privada	95.29	96.82	91.00
Romero-Oraá et al. [11]	DRIMDB	99.48	99.20	100
<b>Método propuesto</b>	<b>Privada</b>	<b>96.00</b>	<b>98.09</b>	<b>90.09</b>
<b>Método propuesto</b>	<b>DRIMDB</b>	<b>99.48</b>	<b>99.20</b>	<b>100</b>

**Tabla 2.** Comparación de los resultados con otros estudios.

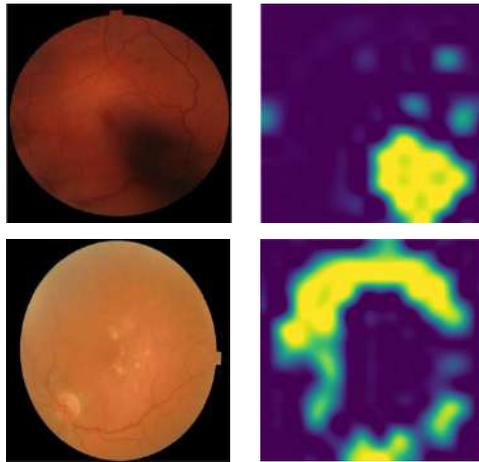


Figura 2. Ejemplos de algunos mapas de atención generados.

como una primera aproximación. En líneas futuras se estudiarán mecanismos de atención más avanzados y se analizarán los mapas generados de forma cuantitativa.

## 6. Conclusiones

En este trabajo se ha propuesto un método basado en *deep learning* y un mecanismo de atención para evaluar de la calidad de retinografías con un gran rendimiento. Los resultados obtenidos indican que método propuesto podría ser de utilidad en un entorno clínico como primera etapa en los sistemas de ayuda al diagnóstico de patologías como la retinopatía diabética, el glaucoma, las cataratas y la degeneración macular asociada a la edad.

## Agradecimientos

Esta investigación se ha desarrollado en el marco de las ayudas PID2020-115468RB-I00 y PGC2018-098214-A-I00 financiadas por el 'Ministerio de Ciencia e Innovación/Agencia Estatal de Investigación/10.13039/501100011033/' y el Fondo Europeo de Desarrollo Regional (FEDER). Una forma de hacer Europa; y por el 'CIBER en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN)' a través del 'Instituto de Salud Carlos III' cofinanciado con fondos FEDER. R. Romero-Oraá disfruta de un contrato predoctoral cofinanciado por el Fondo Social Europeo, el Programa Operativo de Castilla y León, y la Junta de Castilla y León, a través de la Consejería de Educación.

## Referencias

- [1] A. Raj, A. K. Tiwari, and M. G. Martini, "Fundus image quality assessment: survey, challenges, and future scope," *IET Image Process.*, vol. 13, no. 8, pp. 1211–1224, Jun. 2019.
- [2] G. T. Zago, R. V. Andreão, B. Dorizzi, and E. O. Teatini Salles, "Retinal image quality assessment using deep learning," *Comput. Biol. Med.*, vol. 103, pp. 64–70, 2018.
- [3] J. Lin, L. Yu, Q. Weng, and X. Zheng, "Retinal image quality assessment for diabetic retinopathy screening: A survey," *Multimed. Tools Appl.*, 2019.
- [4] S. K. Saha, B. Fernando, J. Cuadros, D. Xiao, and Y. Kanagasigam, "Automated Quality Assessment of Colour Fundus Images for Diabetic Retinopathy Screening in Telemedicine," *J. Digit. Imaging*, vol. 31, no. 6, pp. 869–878, Dec. 2018.
- [5] J. M. Pires Dias, C. M. Oliveira, and L. A. Da Silva Cruz, "Retinal image quality assessment using generic image quality indicators," *Inf. Fusion*, vol. 19, no. 1, pp. 73–90, 2014.
- [6] F. Shao, Y. Yang, Q. Jiang, G. Jiang, and Y. S. Ho, "Automated Quality Assessment of Fundus Images via Analysis of Illumination, Naturalness and Structure," *IEEE Access*, vol. 6, pp. 806–817, Nov. 2017.
- [7] R. A. Welikala *et al.*, "Automated retinal image quality assessment on the UK Biobank dataset for epidemiological studies," *Comput. Biol. Med.*, vol. 71, pp. 67–76, 2016.
- [8] F. Yu, J. Sun, A. Li, J. Cheng, C. Wan, and J. Liu, "Image quality classification for DR screening using deep learning," in *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2017, pp. 664–667.
- [9] R. Tennakoon, D. Mahapatra, P. Roy, S. Sedai, and R. Garnavi, "Image Quality Classification for DR Screening Using Convolutional Neural Networks," in *Proceedings of the Ophthalmic Medical Image Analysis Third International Workshop*, 2016, pp. 113–120.
- [10] A. S. Coyner *et al.*, "Deep Learning for Image Quality Assessment of Fundus Images in Retinopathy of Prematurity," *AMIA ... Annu. Symp. proceedings. AMIA Symp.*, vol. 2018, pp. 1224–1232, 2018.
- [11] R. Romero-Oraá, M. García, J. Oraá-Pérez, M. I. López, and R. Hornero, "Transfer learning para evaluar de forma automática la calidad en imágenes de fondo de ojo," in *XXXVII Congreso Anual de la Sociedad Española de Ingeniería Biomédica (CASEIB 2019)*, 2019, pp. 175–178.
- [12] M. H. Guo *et al.*, "Attention mechanisms in computer vision: A survey," *Computational Visual Media*, vol. 8, no. 3. Springer, pp. 331–368, 15-Mar-2022.
- [13] U. Sevik, "DRIMDB (Diabetic Retinopathy Images Database) Database for Quality Testing of Retinal Images." 2014.
- [14] S. J. Pan and Q. Yang, "A Survey on Transfer Learning," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345–1359, Oct. 2010.
- [15] Jia Deng, Wei Dong, R. Socher, Li-Jia Li, Kai Li, and Li Fei-Fei, "ImageNet: A large-scale hierarchical image database," in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 2009, pp. 248–255.
- [16] C. Szegedy *et al.*, "Going deeper with convolutions," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2015, vol. 07-12-June, pp. 1–9.
- [17] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2016, vol. 2016-Decem, pp. 770–778.
- [18] R. J. Chalakkal, W. H. Abdulla, and S. S. Thulaseedharan, "Quality and content analysis of fundus images using deep learning," *Comput. Biol. Med.*, 2019.
- [19] R. A. Karlsson, B. A. Jonsson, S. H. Hardarson, O. B. Olafsdottir, G. H. Halldorsson, and E. Stefansson, "Automatic fundus image quality assessment on a continuous scale," *Comput. Biol. Med.*, vol. 129, Feb. 2021.
- [20] O. Abramovich, H. Pizem, J. Van Eijgen, I. Stalmans, E. Blumenthal, and J. A. Behar, "FundusQ-Net: a Regression Quality Assessment Deep Learning Algorithm for Fundus Images Quality Grading," *ArXiv Prepr.*, May 2022.

# Uso de la Termografía para el Análisis y Seguimiento de los Hemangiomas Infantiles

J. A. Leñero-Bardallo<sup>1</sup>, J. A. Pérez-Carrasco<sup>2,\*</sup>, C. Serrano<sup>2</sup>, B. Acha<sup>2</sup>, J.-O. Álvarez<sup>3</sup>, J. Bernabéu-Wittel<sup>3</sup>

<sup>1</sup> Departamento de Electrónica y Electromagnetismo, Universidad de Sevilla, Sevilla, España, jlenero@us.es.

<sup>2</sup> Dpto de Teoría de la Señal y Comunicaciones, Universidad de Sevilla, Sevilla España, {jperez2, bacha, cserrano}@us.es

<sup>3</sup> Unidad de Dermatología. Hosp. Univ. Virgen del Rocío, Sevilla, España, {juan.ortiz.alvarez94, jbernabeuw}@gmail.com

## Resumen

Los hemangiomas infantiles se encuentran entre los tumores más comunes en los bebés. Para una correcta evolución es fundamental determinar su extensión, volumen y ubicación. Sin embargo, esto puede representar un desafío, ya que los hemangiomas pueden presentar patrones irregulares, estar cubiertos por cabello y su profundidad no se puede estimar fácilmente. Habitualmente, su diagnóstico se basa en estimaciones de área, textura y color. En este trabajo presentamos una técnica alternativa de diagnóstico, la termografía, que es una herramienta útil e inmediata para un examen preciso de los hemangiomas. Los hemangiomas alteran localmente la temperatura corporal, debido a que el flujo sanguíneo en su interior es mayor debido a la proliferación de células endoteliales de los vasos sanguíneos. Por tanto, midiendo las variaciones de temperatura en la piel, se puede estimar la extensión del hemangioma. Se analizaron 30 pacientes con hemangiomas infantiles con una cámara térmica durante varias sesiones en dos fases. A raíz de los resultados obtenidos concluimos que un aumento de la temperatura en el interior del tumor durante el seguimiento era indicativo de una evolución no deseable.

## 1. Motivación

Los hemangiomas infantiles (HI) son tumores benignos producidos por la proliferación de células endoteliales de los vasos sanguíneos, con una alta incidencia en niños menores de un año [1,2]. Suelen tratarse con propranolol oral. La evaluación de la extensión y el volumen del HI no es sencilla. Pueden presentar patrones irregulares, pueden estar cubiertos por cabello o no ser accesibles directamente a través de la inspección visual. Léauté-Labrèze y col., monitorearon la evolución del HI mediante fotografías estandarizadas y que fueron evaluadas por diferentes expertos [1].

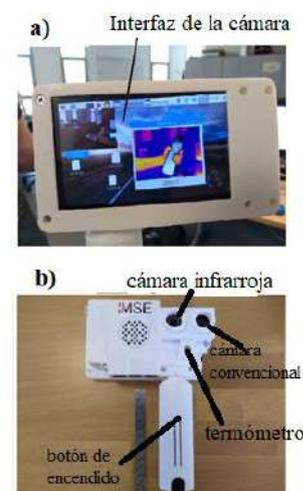
En la práctica clínica habitual, los médicos controlan la evolución del HI mediante la inspección visual de fotografías previas que pueden no mostrar los límites reales del tumor. En 2017, Strumila y col., publicaron un estudio que presenta la termografía como una herramienta para prever la evolución correcta del hemangioma sin tratamiento [3]. Sin embargo, hasta donde sabemos, no existen procedimientos automatizados para medir la evolución de la HI durante el tratamiento. En este artículo, la termografía se presenta como una herramienta

útil para monitorear la evolución de la HI (temperatura y extensión) antes y después del tratamiento.

Durante la última década, las cámaras de infrarrojos han bajado de precio y se han vuelto asequibles y fiables para el diagnóstico clínico. Hoy en día, estos dispositivos tienen una demanda muy alta porque se han empleado masivamente en entornos públicos para detectar personas que puedan estar afectadas por el COVID-19. Las cámaras térmicas pueden medir con precisión las variaciones espaciales de la temperatura corporal [4,5] dentro de hemangiomas [6,7], malformaciones vasculares [8-10], y otras entidades clínicas [4]. Sin embargo, hasta donde sabemos, ningún estudio ha demostrado cómo las cámaras infrarrojas pueden monitorear la temperatura y la extensión del HI durante el tratamiento. La medición de dichos parámetros correlacionados con la actividad del hemangioma puede optimizar la dosis de medicamento administrado.

## 2. Método

Para el seguimiento de los hemangiomas en los pacientes infantiles se diseñó y construyó un sistema portátil con dos sensores de imagen, una pantalla táctil y un termómetro [10], como se muestra en la Figura 1.



**Figura 1.** Sistema portátil utilizado. Implementación detallada del sistema. (a) Vista frontal. (b) Vista posterior. Tiene dos cámaras que funcionan simultáneamente. Uno opera en el espectro visible y la otra en la banda infrarroja.

Uno de los sensores es una cámara digital estándar que opera en el espectro de luz visible. El segundo sensor es una cámara termográfica que detecta radiación en la banda de infrarrojos de onda larga (LWIR). La Figura 2 muestra como se utiliza el dispositivo para estudiar a un paciente.



**Figura 2.** Ilustración de como el sistema es usado para estudiar a un paciente. El hemangioma es analizado sin tocar al paciente. La distancia entre el dispositivo y la lesión es fija.

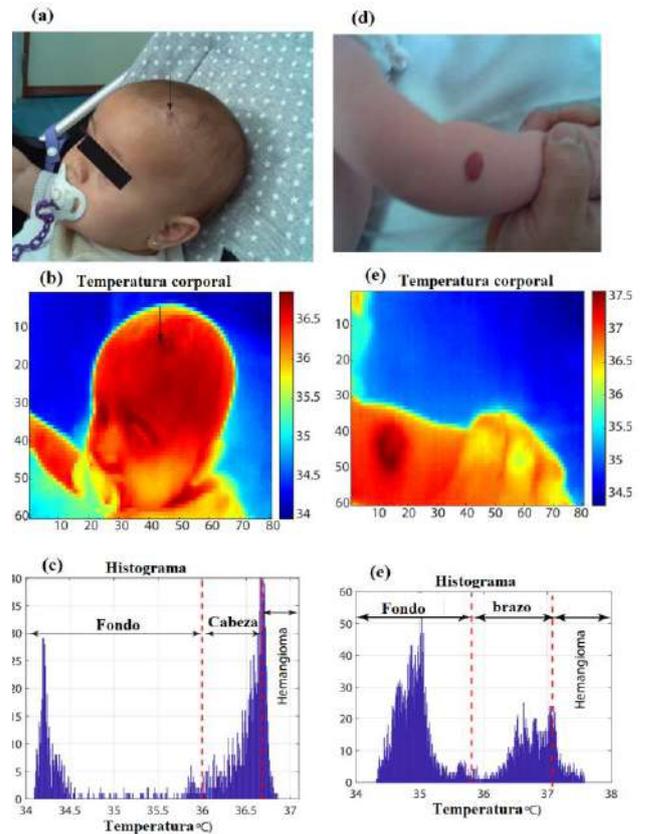
Para este trabajo, 30 pacientes con al menos un hemangioma infantil fueron examinados mediante termografía en su primera consulta de dermatología. El seguimiento de estos pacientes ha sido de dos años. Para todos ellos, se tomaron dos imágenes diferentes (imágenes clínicas e infrarrojas) que fueron adquiridas simultáneamente con un protocolo específico: a una distancia fija de 30 cm y en las mismas condiciones de iluminación y temperatura ambiente. En cada sesión, la cámara infrarroja fue calibrada con un termómetro para traducir los niveles de radiación en la banda LWIR a temperaturas absolutas. Todas las lesiones habían sido diagnosticadas previamente como HI.

Para todos los pacientes, se calculó la variación de temperatura entre el centro del hemangioma y su piel sana circundante en la consulta de dermatología de la primera visita. Grandes variaciones de temperatura entre el HI y la piel sana son indicativas de un flujo sanguíneo alto indeseable dentro del tumor [9].

En una segunda fase del estudio se analizó la evolución de los hemangiomas infantiles en 9 de estos 30 pacientes. El criterio de inclusión para este segundo estudio fue seleccionar pacientes que hubieran sido examinados con la cámara IR tres veces en nueve meses en la consulta de dermatología. Este segundo estudio tuvo como objetivo determinar si la evolución del hemangioma administrando diferentes medicamentos se correlacionaba con el análisis térmico proporcionado por la cámara. Para ello, las imágenes infrarrojas y clínicas fueron adquiridas simultáneamente en cada sesión. Se calculó la variación de temperatura entre el centro del HI y la piel sana circundante para cada sesión. Finalmente, se analizó la evolución de esta variación de temperatura a lo largo de las tres sesiones para cada paciente. Por lo tanto, con los

datos de temperatura y extensión del hemangioma, la dosis de medicamento administrada puede ser ajustada más convenientemente según aumente o disminuya la temperatura y/o extensión de la lesión.

La Figura 3 muestra dos casos de hemangioma utilizando el sistema de diagnóstico descrito. La imagen térmica muestra con mayor temperatura la zona afectada. La tercera imagen muestra el histograma de la imagen térmica.



**Figura 3.** Estudio de la temperatura de hemangioma infantil en dos casos diferentes. a) y d) Instantáneas del hemangioma adquiridas con una cámara digital estándar. b) y e) Imágenes termográficas del hemangioma que muestran valores absolutos de temperatura. c) y f) Histogramas de las imágenes térmicas.

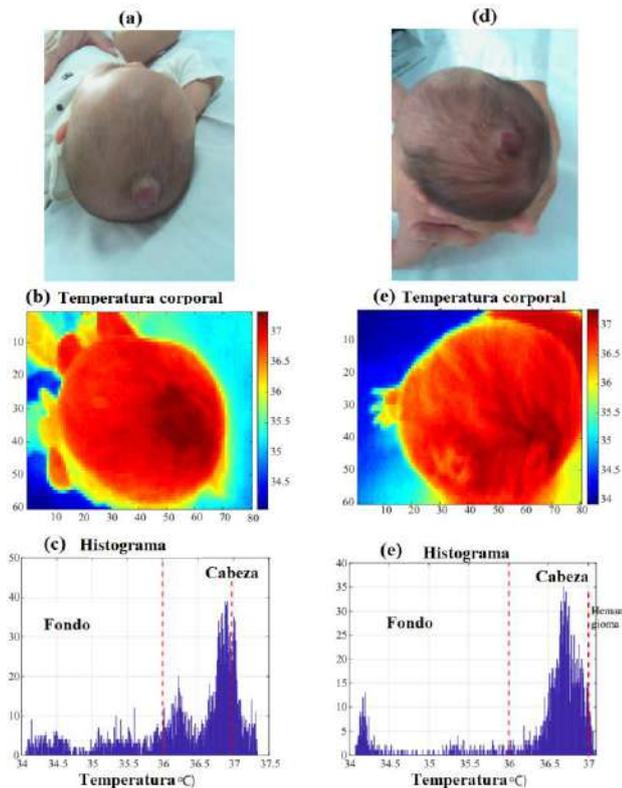
Las Figuras 3 (a, d) muestran el tumor en el espectro visible. Las Figuras 3 (b, e) muestran la misma imagen en la banda LWIR. Finalmente, las Figuras 3 (c, f) muestran los histogramas de las imágenes infrarrojas. Cada bin del histograma representa el número de píxeles con un determinado valor de temperatura.

### 3. Resultados Experimentales

La temperatura y la extensión de los hemangiomas infantiles en los diferentes pacientes fueron analizados en las distintas sesiones de tratamiento.

La Figura 4 ilustra la evolución de un HI localizado en la cabeza antes (fila superior) y después del tratamiento (fila inferior) en un plazo de cuatro meses. Como se puede apreciar en la figura, se observa una reducción del área de la región afectada. Además de la reducción de área se

observa también una reducción de la temperatura entre la región afectada y la no afectada.



**Figura 4.** Estudio de la temperatura de hemangioma infantil en la cabeza en un plazo de cuatro meses de tratamiento con propranolol oral. a) y d) Instantáneas del hemangioma adquiridas con una cámara digital estándar. b) y e) Imágenes termográficas del mismo hemangioma que muestran valores absolutos de temperatura. c) y f) Histogramas de las imágenes térmicas.

Mediante procesamiento de imagen se pudo determinar que la extensión del HI disminuyó tras la administración de propranolol oral. Por lo tanto, el área es una medida indicativa de la involución del HI. La extensión del HI se puede determinar incluso si está parcial o totalmente cubierta por cabello.

Los histogramas fueron útiles para identificar el número de píxeles afectados por el tumor (extensión) en diferentes regiones de la imagen. Las regiones corporales no afectadas por el HI y el fondo de la imagen proporcionaron temperaturas más bajas. Los modos del histograma indican la temperatura promedio de cada región (afectada por HI o no). Si la evolución del tratamiento es correcta, la diferencia de temperatura entre los diferentes modos del histograma debe minimizarse.

Por lo tanto, se concluye que una involución del hemangioma HI está asociada con decrementos de área y temperatura.

La Tabla 1 resume los datos experimentales obtenidos tras analizar el HI durante dos años en dos fases. Las dos columnas de la Tabla 1 muestran los resultados del estudio de HI durante la primera visita clínica de los

pacientes en la primera fase. En pacientes con múltiples hemangiomas solo se consideró el de mayor tamaño. El 87% de hemangiomas indujeron variaciones de temperatura con respecto a la piel circundante. La variación de temperatura en este análisis inicial tiene una gran variabilidad porque depende fuertemente del subtipo de hemangioma y de la fase de proliferación (el tiempo que transcurre entre la aparición de la HI y la primera inspección clínica). La variación de temperatura promedio fue de 0.2 °C. El intervalo de confianza del 90% de la variación de temperatura fue de [0.16, 0.25] °C.

**Primera fase del Estudio**

Nº pacientes	30
Nº de hemangiomas (solo primera visita)	30
Nº de imágenes con variación de Temperatura	26
Incremento de Temperatura promedio (°C)	0.20
Intervalo de confianza del 90% de variación de Temperatura	[0.16-0.25]
Porcentaje de hemangiomas con variaciones de Temperatura	87 %

**Tabla 1.** Resultados primera fase del estudio

**Segunda fase del Estudio**

Nº pacientes	9
Nº de hemangiomas (varias visitas)	17
Incremento de Temperatura positivo	2
Incremento de Temperatura negativo	14
Correlación con la decisión del dermatólogo	17/17
Variación de la temperatura media del tratamiento del hemangioma (°C)	0.19

**Tabla 2.** Resultados segunda fase del estudio

En la Tabla 2 se muestran los resultados de la segunda fase del trabajo de investigación. En esta fase, de entre los 30 pacientes en estudio, nueve de ellos con HI de alto o moderado riesgo fueron seleccionados para someterse a diferentes tratamientos (propranolol oral, nadolol oral, timolol tópico, o simplemente observar y esperar la evolución del HI) y a ser monitorizados con nuestra cámara térmica durante nueve meses. Algunos de ellos tenían más de un hemangioma y un paciente presentó hemangiomatosis cutánea múltiple. En total, fueron 17 HI en estudio en esta segunda etapa. Se monitorizaron sus temperaturas de HI durante tres sesiones espaciadas tres meses en el tiempo.

La tabla muestra los resultados que resumen la respuesta del tratamiento y la evolución de la temperatura. En la primera sesión, la variación de temperatura promedio de HI fue de 0,38 °C. En la última sesión, la variación media fue de 0,27 °C. La evolución de la temperatura del HI mostró una variación negativa de la temperatura a lo largo de las tres sesiones, acompañada de una reducción del área del HI durante el tratamiento en 7 de los 9 pacientes. Dos pacientes mostraron una variación de temperatura positiva después de la interrupción del tratamiento. Esto es concordante con un nuevo crecimiento del HI después de detener el tratamiento.

En la quinta fila de la tabla 2 se analiza la correlación entre la variación de temperatura y la decisión de tratamiento del dermatólogo. Consideramos que existe una correlación entre ambos si: i) existe un aumento relativo de temperatura en el hemangioma a lo largo del tiempo y el dermatólogo diagnostica una respuesta negativa al tratamiento; ii) existe un descenso relativo de la temperatura en el hemangioma y el dermatólogo diagnostica una respuesta positiva al tratamiento.

#### 4. Conclusiones

En conclusión, las cámaras infrarrojas son herramientas útiles para monitorear la actividad y extensión de la temperatura HI durante su tratamiento. Los resultados experimentales realizados con 30 pacientes en su primera consulta de dermatología muestran que existe un aumento de temperatura en el HI en comparación con la piel sana. Así fue en el 87% de los 30 casos. Por otro lado, el estudio realizado con 9 pacientes durante nueve meses muestra una clara correlación entre la evolución de HI y su variación y extensión de temperatura. La termografía puede coexistir con otras técnicas auxiliares de diagnóstico, como las basadas en análisis de ultrasonido Doppler o resonancia magnética. El análisis termográfico es inmediato, no invasivo y proporciona una estimación de la extensión del hemangioma cuando los tumores están parcialmente cubiertos por cabello o no tienen un color o textura uniforme. La rápida difusión y el desarrollo de las cámaras infrarrojas hacen que la tecnología sea fácilmente accesible para los médicos.

En el futuro cercano se pretende obtener imágenes de mayor resolución y obtener medidas de área de superficie afectada, además de análisis más cuantitativos utilizando parámetros obtenidos del análisis del histograma.

#### Agradecimientos

Este trabajo ha sido realizado gracias a la financiación obtenida mediante los siguientes proyectos e instituciones: Proyectos de Generación de Conocimiento 2021 del Ministerio de Ciencia e Innovación (PID2021-127871OB-I00) y proyecto FEDER-US Fondo Europeo de Desarrollo Regional (FEDER) y Consejería de Economía, Conocimiento, Empresas y Universidad de la Junta de Andalucía (US-1381640), Ayudas a Actividades de Transferencia de Conocimiento entre los Agentes del Sistema Andaluz del Conocimiento y el Tejido Productivo (AT21\_00096), Proyectos de I+D+i de

Entidades públicas – Convocatoria 2020 (P20\_01206) (VERSO).

#### Referencias

- [1] Léauté-Labrèze C. et al. A randomized, controlled trial of oral propranolol in infantile hemangioma. *The New England Journal of Medicine*, vol 372, 2015, pp 735-746. <https://www.doi.org/10.1056/NEJMoa1404710>.
- [2] Wassef M, Blei F, Adams D, et al. Vascular anomalies classification: Recommendations from the international society for the study of vascular anomalies. *Pediatrics*, vol 136, sup 1, 2015, pp 203-14 <https://doi.org/10.1542/peds.2014-3673>.
- [3] Strumila A. et al. Infantile hemangioma: Predicting proliferation by infrared thermography. *Medicina*, vol 53, sup 2, 2017, pp 85-89. <https://doi.org/10.1016/j.medici.2017.04.002>.
- [4] Saxena AK, Willital GH. Infrared thermography: Experience from a decade of pediatric imaging. *Eur J Pediatr*, vol 167, 2008, pp 757-64. <https://www.doi.org/10.1007/s00431-007-0583-z>.
- [5] Samantha LS, Indermeet K, Iltefat H et al. Emerging imaging technologies in dermatology: Part I: Basic principles. *Journal of the American Academy of Dermatology*, vol 80, sup 4, 2019, pp 1114-20. <https://doi.org/10.1016/j.jaad.2018.11.042>.
- [6] Mohammed JA, Balma-Mena A, Chakkittakandiyil A et al. Infrared thermography to assess proliferation and involution of infantile hemangiomas: A prospective cohort study. *JAMA Dermatol*, vol 150, sup 9, 2014, pp 964-9. <https://www.doi.org/10.1001/jamadermatol.2014.112>.
- [7] García-Romero MT, Chakkittakandiyil A, Pope E. The role of infrared thermography in evaluation of proliferative infantile hemangiomas. Results of a pilot. *Int J Dermatol*, vol 53, sup 3, 2014, pp 216-7. <https://doi.org/10.1111/ijd.12045>.
- [8] Hardwicke JT, Titley OG. Thermographic assessment of a vascular malformation of the hand: A new imaging modality. *J Clin Imaging Sci*, vol 6, 2016, pp 6-9. <https://doi.org/10.1111/ced.14346>.
- [9] Leñero-Bardallo, JA, Serrano C, Acha B, Pérez-Carrasco JA, Bernabéu-Wittel J. Thermography for the differential diagnosis of vascular malformations. *Clinical and Experimental Dermatology*, vol 46, sup 2, 2021, pp 314-318. <https://doi.org/10.1111/ced.14346>.
- [10] Leñero-Bardallo, JA, Serrano C, Acha B, Pérez-Carrasco JA, Bernabéu-Wittel J. Thermography for the differential diagnosis of vascular malformations, Mendeley Data, V1, 2019, <http://dx.doi.org/10.17632/y82fj6296z.1>. Último acceso Agosto 2022.

# Análisis textural de imágenes [<sup>18</sup>F]FDG PET/CT para la predicción de enfermedad mínima residual en pacientes con sintomatología de mieloma múltiple

E. Milara<sup>1</sup>, R. Cenarro<sup>1</sup>, A. Gómez-Grande<sup>2,3</sup>, A.P. Seiffert<sup>1</sup>, R. Alonso<sup>3,4</sup>, E.J. Gómez<sup>1,5</sup>,  
P. Sánchez-González<sup>1,5</sup>

<sup>1</sup> Grupo de Bioingeniería y Telemedicina, ETSI Telecomunicación, Centro de Tecnología Biomédica, Universidad Politécnica de Madrid, Madrid, España; {eva.milara.hernando, ap.seiffert, enriquejavier.gomez, p.sanchez}@upm.es, r.cenarro@alumnos.upm.es

<sup>2</sup> Servicio de Medicina Nuclear, Hospital Universitario 12 de Octubre, Madrid, España; adolfo.gomez@salud.madrid.org

<sup>3</sup> Facultad de Medicina, Universidad Complutense de Madrid, Madrid, España; rafaelalberto.alonso@salud.madrid.org

<sup>4</sup> Servicio de Hematología y Hemoterapia e Instituto de Investigación Sanitaria (i+12), Hospital Universitario 12 de Octubre, CNIO, CIBERONIC, Madrid, España;

<sup>5</sup> Centro de Investigación Biomédica en Red de Bioingeniería, Biomateriales y Nanomedicina, Instituto de Salud Carlos III, Madrid, España

## Resumen

*La Enfermedad Residual Medible (ERM) es una afectación en la cual parte de la carga tumoral causada por el mieloma múltiple (MM) persiste tras el tratamiento. Esta afectación se evalúa mediante una combinación de biopsia medular e imágenes de tomografía por emisión de positrones. Sin embargo, en algunos casos, sigue siendo de difícil valoración. En el presente estudio, se presenta una metodología de cuantificación de las imágenes tomadas en el momento de diagnóstico del MM con el objetivo de predecir la aparición de ERM. Para ello, se realiza una segmentación de la médula ósea de la cual se extraen un total de 32 características de imagen. A partir de estas características, se realiza una prueba U de Mann-Whitney para evaluar individualmente la capacidad de las variables para predecir ERM y 7 modelos de Machine Learning basados en el conjunto de características alcanzando valores de exactitud de 0,90. Se concluye que el análisis de radiómica resulta de gran utilidad en el contexto clínico y que puede apoyar a la toma de decisiones clínicas.*

## 1. Introducción

El mieloma múltiple (MM) es una neoplasia maligna caracterizada por el crecimiento incontrolado de células plasmáticas monoclonales en la médula ósea que representa el 9% de todos los cánceres hematológicos. Los síntomas más habituales de esta enfermedad son anemia, lesiones óseas, hipercalcemia, insuficiencia renal, fatiga y dolor [1], [2]. Tras el tratamiento, algunos pacientes no alcanzan la respuesta completa manteniendo carga tumoral residual condición nombrada Enfermedad Residual Medible (ERM) por el grupo de Trabajo Internacional sobre Mieloma (IMWG, *International Myeloma Working Group*) [3], [4]. El IMWG define ERM como un biomarcador pronóstico de peores resultados de supervivencia. Actualmente, el IMWG [3] recomienda combinar la biopsia medular con las imágenes [<sup>18</sup>F]FDG PET/CT tanto para el diagnóstico de MM como para la detección de ERM. Para medir la carga tumoral en la biopsia se utiliza la Citometría de Flujo Multiparamétrica

(CFM), mientras que para la evaluación de las imágenes [<sup>18</sup>F]FDG PET se realiza una valoración visual por parte de un experto en Medicina Nuclear [5], [6]. A pesar de que existen los criterios estandarizados IMPeTUs [7] para la evaluación visual de las imágenes, que permiten valorar la extensión de la enfermedad ósea y la presencia de enfermedad extramedular, resultan insuficientes en la evaluación de la afectación de la médula ósea. Por este motivo, surge la necesidad de aplicar técnicas de cuantificación en la evaluación de las imágenes.

Aunque las métricas relacionadas con el valor de captación normalizado (SUV, *Standardize Uptake Value*) son las más utilizadas para la cuantificación de imágenes [<sup>18</sup>F]FDG PET en contextos oncológicos, se descartan en los criterios IMPeTUs [7] para la evaluación de MM por estar habitualmente afectadas por procesos inflamatorios, ruido o métodos de reconstrucción. Como medida complementaria a las métricas SUV, el análisis de radiómica ha sido estudiado como un indicador pronóstico de supervivencia en pacientes recién diagnosticados con MM [8]. Además, estas características extraídas de las imágenes se han utilizado para crear modelos para el diagnóstico de MM difuso [9]. Sin embargo, no se han definido métricas o modelos basados en las imágenes [<sup>18</sup>F]FDG PET/CT de diagnóstico para predecir la ERM.

El objetivo del presente estudio es realizar un análisis de radiómica de las imágenes de diagnóstico de [<sup>18</sup>F]FDG PET/CT de pacientes con MM para la predicción de ERM. Para ello, se analiza la capacidad discriminativa entre ERM+/- de las características teniendo en cuenta tanto la evaluación de la valoración visual (PET+/PET-) como la de la biopsia (CFM+/CFM-).

## 2. Metodología

En este estudio, se utiliza una cohorte de 27 imágenes [<sup>18</sup>F]FDG PET/CT del momento de diagnóstico de MM recogidas en el Servicio de Medicina Nuclear del Hospital

Universitario 12 de Octubre de Madrid. Para su posterior análisis, se cuenta con la valoración visual de las imágenes (7 PET+ y 20 PET-) realizadas por expertos en Medicina Nuclear y la evaluación por estudio CFM de la biopsia medular (6 CFM+ y 21 CFM-), siendo ambas evaluaciones correspondientes al estudio de ERM postratamiento.

Para la segmentación de la médula ósea se utiliza una herramienta desarrollada en el Grupo de Bioingeniería y Telemedicina (GBT) de la Universidad Politécnica de Madrid e implementada en MATLAB (The MathWorks, Inc., Natick, MA, EEUU) [10]. Para la extracción de las características se emplea la librería de Vallieres llamada *Radiomics* [11], [12], desarrollado para MATLAB. Para su análisis se utilizan la herramienta SPSS *Statistics* (IBM Corp., Armonk, NY, EEUU) [13] y el software Orange Data Mining (Bled, Slovenia) [14].

### 2.1. Segmentación de la médula ósea

En primer lugar, mediante el uso de una aplicación previamente desarrollada en el GBT, se realiza una segmentación automática y posterior edición manual de la médula ósea del esqueleto axial incluyendo la columna cervico-dorsolumbar, pelvis, parrilla costal con esternón, clavículas y escápulas junto con los húmeros y fémures. Se descartan tibias, peronés, cubitos, radios y cráneo para evitar artefactos causados por el movimiento. Esta segmentación se realiza sobre la imagen CT para obtener la región anatómica de forma precisa y, posteriormente, se extrae la misma región de la imagen [<sup>18</sup>F]FDG PET para analizar cuantitativamente la captación de la médula.

### 2.2. Extracción de características

Para la máscara de la médula ósea en la imagen [<sup>18</sup>F]FDG PET, se extraen un total de 32 características: 3 características de histograma o primer orden, 9 características de segundo orden correspondientes a la *Gray Level Co-occurrence Matrix* (GLCM), 18 de alto orden, pertenecientes 13 de ellas a la *Gray Level Run Length Matrix* (GLRLM) y 5 a la *Neighborhood Gray Tone Difference Matrix* (NGTDM),  $SUV_{medio}$  y  $SUV_{max}$ . Posteriormente, se realiza un proceso de sobremuestreo para equilibrar el número de casos positivos y negativos tanto en PET como en CFM, y aumentar así la representatividad de la clase positiva. La función utilizada es *Safe-Level-Synthetic Minority Over-Sampling Technique* (*Safe-Level-SMOTE*) [15].

### 2.3. Análisis Univariante

Las diferencias entre los grupos ERM+/- se evalúan mediante la prueba U de Mann-Whitney. Para ello, se toma como variable de clasificación la valoración visual de la presencia de ERM realizada por los especialistas en Medicina Nuclear (PET+/-) y el análisis de la biopsia (CFM+/-) independientemente, y como base de datos la original y la sobremuestreada. Un  $p$ -valor  $< 0,05$  se considera estadísticamente significativo en la diferenciación de los grupos.

### 2.4. Modelos de Machine Learning

Partiendo de las características texturales extraídas a partir de las imágenes [<sup>18</sup>F]FDG PET se realizan 7 modelos de

*Machine Learning* para la diferenciación de los pacientes entre los grupos PET+/- y CFM+/- . Los modelos de *Machine Learning* implementados son: árbol de decisión; máquinas de vectores de soporte (SVM, *Support Vector Machines*) con tres *kernels* distintos que son lineal, polinomial y función de base radial (FBR); bosque aleatorio (*random forest*); redes neuronales; el algoritmo Naive-Bayes; regresión logística y k-vecino más cercano (K-NN, *K.Nearest Neighbours*).

Para la validación de los modelos, se utiliza la función de validación cruzada con 5 *folds* y se calculan las métricas exactitud (Ex), precisión (Pr), sensibilidad (S) y especificidad (Es), considerando aceptables los valores superiores a 0,7 y excelentes los superiores a 0,9. Además, se obtiene el valor de AUC (*Area Under the Curve*) correspondiente a la curva ROC (*Receiver Operating Characteristic*), considerando aceptables los valores inferiores a 0,3 o superiores a 0,7 y excelentes los valores inferiores a 0,1 y superiores a 0,9. Estos modelos son entrenados y evaluados tanto con la base de datos original como con la base de datos sobremuestreada.

## 3. Resultados y discusión

### 3.1. Análisis Univariante

Los resultados del análisis de la prueba U de Mann-Whitney se muestran en la Tabla 1 incluyendo exclusivamente aquellas variables con  $p$ -valor  $< 0,05$  para alguna de las bases de datos. Para el caso del análisis tomando como variable de clasificación PET+/-, se puede observar que para la base de datos original se obtienen 8 variables con  $p$ -valor  $< 0,05$ , destacando la variable GLRLM\_LRLGE ( $p$ -valor = 0,006) y en general la matriz GLCM o variables de segundo orden (con 5 de las 8 variables con  $p$ -valor  $< 0,05$ ) igualando e incluso mejorando los resultados del  $SUV_{max}$  ( $p$ -valor = 0,036). De la misma manera, en la base de datos sobremuestreada, se puede observar como las variables que obtienen diferencias estadísticamente significativas en el análisis anterior mejoran sus resultados, alcanzando  $p$ -valores inferiores a 0,005. Además, se incluyen 8 nuevas variables con  $p$ -valores  $< 0,05$ , obteniendo un total de 16 variables con diferencias entre los grupos PET+ y PET-. Para el caso de los grupos CFM+/-, los resultados para la base de datos original no obtienen ninguna variable con  $p$ -valor  $< 0,05$ . Sin embargo, aumentado el equilibrio de los grupos en la base de datos sobremuestreada, se obtienen 19 variables que muestran diferencias significativas entre los grupos, alcanzando algunas variables valores inferiores a 0,001.

Por un lado, se puede destacar la utilidad de la cuantificación de las imágenes [<sup>18</sup>F]FDG PET a partir de las características de radiómica en la diferenciación de los pacientes PET+/-, ya que, incluso para la base de datos original con un número reducido de casos positivos, existen variables con diferencias significativas entre los grupos. Por el contrario, se puede determinar que el mismo análisis no consigue diferenciar los grupos CFM+/CFM-, ya que solo obtiene diferencias significativas con el aumento de los datos, generando un posible sobreentrenamiento con los datos creados en el sobremuestreo.

		Mann-Whitney ( <i>p</i> -valor)			
		PET		CFM	
		O	S	O	S
GLCM	Contraste	<b>0,026</b>	<b>0,004</b>	0,140	<b>&lt;0,001</b>
	Homogeneidad	<b>0,026</b>	<b>0,005</b>	0,140	<b>&lt;0,001</b>
	Suma Media	<b>0,031</b>	<b>0,002</b>	0,376	<b>0,002</b>
	Dissimilitud	<b>0,026</b>	<b>0,005</b>	0,140	<b>&lt;0,001</b>
	Autocorrelación	<b>0,036</b>	<b>0,003</b>	0,408	<b>0,002</b>
GLRLM	LRE	0,072	<b>0,038</b>	0,140	<b>0,003</b>
	RP	0,072	<b>0,038</b>	0,140	<b>0,002</b>
	LGRE	<b>0,036</b>	<b>0,004</b>	0,289	<b>0,002</b>
	HGRE	0,063	<b>0,015</b>	0,512	<b>0,003</b>
	SRLGE	0,198	0,060	0,887	<b>0,028</b>
	SRHGE	0,145	0,072	0,441	<b>0,004</b>
	LRLGE	<b>0,006</b>	<b>0,001</b>	0,110	<b>0,001</b>
	GLV	0,145	0,081	0,512	<b>0,043</b>
NGTDM	Grosor	1,000	0,738	0,125	<b>0,001</b>
	Ocupación	0,130	<b>0,018</b>	0,712	0,122
	Complejidad	0,055	<b>0,008</b>	0,140	<b>&lt;0,001</b>
	Fuerza	0,092	<b>0,007</b>	0,670	<b>0,010</b>
H	Varianza	0,288	0,091	0,289	<b>0,008</b>
	Asimetría	0,935	0,947	0,550	<b>0,038</b>
SUV	SUV <sub>medio</sub>	0,198	<b>0,007</b>	0,977	0,204
	SUV <sub>max</sub>	<b>0,036</b>	<b>0,004</b>	0,289	<b>0,001</b>

**Tabla 1.** Resultados de Mann-Whitney clasificación PET+/- y CFM+/-, para la base de datos original (O) y sobremuestreada (S). Se muestran en negrita los p-valor<0,05. LRE: Long Run Emphasis; RP: Run Percentage; LGRE: Low Gray Level Run Emphasis; HGRE: High Gray Level Run Emphasis; SRLGE: Short Run Low Gray Level Emphasis; SRHGE: Short Run High Gray Level Emphasis; LRLGE: Long Run Low Gray Level Emphasis; GLV: Gray Level Variance.

### 3.2. Modelos de Machine Learning

Los resultados de los modelos de *Machine Learning* basados en las características de radiómica extraídas de las imágenes se muestran en la Tabla 2. Para los modelos entrenados para la diferenciación de PET+/- en la base de datos original, se puede observar que la especificidad de la mayoría de ellos alcanza valores aceptables e incluso excelentes, mientras que las métricas precisión y sensibilidad son destacablemente bajas. Por este motivo, se puede considerar que los modelos se están sobreentrenando con casos PET- y apenas entrenan con PET+, generando un alto grado de falsos negativos. Al aumentar el número de casos positivos en la base de datos sobremuestreada, se observa la reducción de la especificidad a consecuencia de una mayor representación de los casos positivos en el entrenamiento, observable en los valores que alcanzan el resto de las métricas. De hecho, 6 de los modelos alcanzan valores aceptables para todas sus métricas, destacando los modelos de SVM con *kernel* Polinomial y redes neuronales, ambos con sensibilidades de 1 y 0,95 y especificidades de 0,7 y 0,75, respectivamente.

Para los modelos basados en la diferenciación de CFM+/-, se puede observar resultados similares a los obtenidos para PET+/-, con altos valores de especificidad y muy bajos de precisión y sensibilidad para la base de datos original, y mejora general de las métricas para sobremuestreada gracias al aumento de los casos positivos en el entrenamiento. Para este último caso, destaca la métrica AUC, alcanzando para 4 de los 9 modelos valores excelentes. Además, es destacable que, al igual que para PET+/-, los modelos que obtienen mejores métricas son redes neuronales y SVM, aunque en este caso con *kernel* lineal con sensibilidades de 0,95 y 0,9 y sensibilidades de 0,86 y 0,81 respectivamente.

Modelo	BBDD Original					BBDD Sobremuestreada					
	AUC	Ex	Pr	S	Es	AUC	Ex	Pr	S	Es	
PET	Á. decisión	<b>0,71</b>	<b>0,74</b>	0,50	0,29	<b>0,90*</b>	<b>0,80</b>	<b>0,75</b>	<b>0,81</b>	0,65	<b>0,85</b>
	SVM FBR	0,43	<b>0,74</b>	0,00	0,00	<b>1,00*</b>	<b>0,91*</b>	<b>0,78</b>	<b>0,74</b>	<b>0,85</b>	<b>0,70</b>
	SVM Polinomial	0,38	0,59	0,17	0,14	<b>0,75</b>	<b>0,89</b>	<b>0,85</b>	<b>0,77</b>	<b>1,00*</b>	<b>0,70</b>
	SVM Lineal	<b>0,25</b>	0,56	0,22	0,29	0,65	<b>0,80</b>	<b>0,80</b>	<b>0,77</b>	<b>0,85</b>	<b>0,75</b>
	Random Forest	0,58	0,63	0,00	0,00	<b>0,85</b>	<b>0,78</b>	0,68	0,67	<b>0,70</b>	0,65
	Redes neuronales	0,63	0,63	0,20	0,14	<b>0,80</b>	<b>0,88</b>	<b>0,85</b>	<b>0,79</b>	<b>0,95*</b>	<b>0,75</b>
	Naive Bayes	<b>0,70</b>	0,63	0,41	<b>1,00*</b>	0,50	<b>0,79</b>	<b>0,78</b>	<b>0,74</b>	<b>0,85</b>	<b>0,70</b>
	Regresión logística	0,60	0,59	0,00	0,00	<b>0,80</b>	<b>0,81</b>	<b>0,80</b>	<b>0,77</b>	<b>0,85</b>	<b>0,75</b>
	K-NN	0,40	0,63	0,00	0,00	<b>0,85</b>	0,62	0,68	0,61	<b>0,95*</b>	0,40
CFM	Á. decisión	0,61	<b>0,74</b>	0,40	0,33	<b>0,86</b>	<b>0,82</b>	<b>0,74</b>	<b>0,73</b>	<b>0,76</b>	<b>0,71</b>
	SVM FBR	<b>0,25</b>	<b>0,70</b>	0,00	0,00	<b>0,90*</b>	<b>0,92*</b>	<b>0,83</b>	<b>0,85</b>	<b>0,81</b>	<b>0,86</b>
	SVM Polinomial	<b>0,28</b>	<b>0,74</b>	0,33	0,17	<b>0,90*</b>	<b>0,93*</b>	<b>0,81</b>	<b>0,78</b>	<b>0,86</b>	<b>0,76</b>
	SVM Lineal	<b>0,23</b>	0,59	0,00	0,00	<b>0,76</b>	<b>0,92*</b>	<b>0,86</b>	<b>0,83</b>	<b>0,90*</b>	<b>0,81</b>
	Random Forest	0,31	0,67	0,00	0,00	<b>0,86</b>	<b>0,83</b>	<b>0,79</b>	<b>0,75</b>	<b>0,86</b>	<b>0,71</b>
	Redes neuronales	0,40	0,63	0,00	0,00	<b>0,81</b>	<b>0,95*</b>	<b>0,90</b>	<b>0,87</b>	<b>0,95*</b>	<b>0,86</b>
	Naive Bayes	<b>0,70</b>	0,22	0,17	0,67	0,10	<b>0,81</b>	<b>0,83</b>	<b>0,85</b>	<b>0,81</b>	<b>0,86</b>
	Regresión logística	0,57	<b>0,74</b>	0,33	0,17	<b>0,90*</b>	<b>0,83</b>	0,69	0,65	<b>0,81</b>	0,57
	K-NN	0,52	<b>0,78</b>	0,00	0,00	<b>1,00*</b>	<b>0,74</b>	0,62	0,59	<b>0,81</b>	0,43

**Tabla 2.** Métricas de los modelos de Machine Learning basados en PET y CFM tanto con la base de datos original como con la sobremuestreada. Se muestra en negritas los valores aceptables (>0,7 y, solo en AUC, <0,3) y con \* los excelentes (>0,9 y, solo en AUC, <0,1).

A pesar de que el estudio de Mesguich *et al.* [9], seleccionara el algoritmo de *random forest* basado en las características texturales como el predilecto en la diferenciación de MM difuso, se puede comprobar con los resultados del siguiente estudio que, para la predicción de ERM, los modelos de redes neuronales y SVM obtiene mejores resultados. Sin embargo, cabe destacar que el estudio nombrado utiliza tanto imágenes de [<sup>18</sup>F]FDG PET/CT como de resonancia magnética para la extracción de características y que el contexto clínico es distinto, aunque ambos intentan predecir patrones difusos de la médula ósea. Además, cabe destacar que los modelos con métricas aceptables surgen de las bases de datos sobremuestreadas, por lo que deberían utilizarse estos modelos con una cohorte original más equilibrada, es decir, con mayor representatividad de la clase positiva, ya que esta mejora puede deberse al sobreajuste de los modelos y no al aumento de la representatividad e la clase positiva. Por último, los resultados se encuentran limitados tanto por el número de pacientes disponibles en la cohorte como por el tipo de segmentación semiautomático utilizado, pudiendo mejorarse con segmentaciones de mayor precisión, automáticas basadas en aprendizaje profundo y/o manuales realizadas por expertos.

#### 4. Conclusiones

En este trabajo, se desarrolla una metodología de cuantificación de las imágenes [<sup>18</sup>F]FDG PET/CT de pacientes diagnosticados con MM para la predicción de la evolución a ERM. Las características de radiómica extraídas son capaces de discriminar entre los grupos PET+/- y CFM+/- . Por lo tanto, se demuestra su posible aplicabilidad a la práctica clínica actual como un método de ayuda a la decisión terapéutica, gracias a su capacidad de predecir si el paciente tiene más probabilidad de desarrollar ERM postratamiento en función de la imagen adquirida en el momento del diagnóstico.

#### Agradecimientos

Este trabajo está parcialmente financiado por una beca predoctoral (ayuda del Programa Propio de I+D+i 2020) de la Universidad Politécnica de Madrid.

#### Referencias

[1] K. Brigle and B. Rogers, “Pathobiology and Diagnosis of Multiple Myeloma,” *Semin. Oncol. Nurs.*, vol. 33, no. 3, pp. 225–236, 2017, doi: 10.1016/j.soncn.2017.05.012.

[2] H. Sung *et al.*, “Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries,” *CA. Cancer J. Clin.*, vol. 71, no. 3, pp. 209–249, 2021, doi: 10.3322/caac.21660.

[3] S. Kumar *et al.*, “International Myeloma Working Group consensus criteria for response and minimal residual disease assessment in multiple myeloma,” *Lancet Oncol.*, vol. 17, no. 8, pp. e328–e346, 2016, doi: 10.1016/S1470-2045(16)30206-6.

[4] N. C. Munshi *et al.*, “Association of minimal residual disease with superior survival outcomes in patients with multiple myeloma: A meta-analysis,” *JAMA Oncol.*, vol. 3, no. 1, pp. 28–35, 2017, doi:

10.1001/jamaoncol.2016.3160.

[5] J. Hillengass *et al.*, “Whole-body computed tomography versus conventional skeletal survey in patients with multiple myeloma: A study of the International Myeloma Working Group,” *Blood Cancer J.*, vol. 7, no. 8, 2017, doi: 10.1038/bcj.2017.78.

[6] P. Moreau *et al.*, “Prospective evaluation of magnetic resonance imaging and [<sup>18</sup>F]fluorodeoxyglucose positron emission tomography-computed tomography at diagnosis and before maintenance therapy in symptomatic patients with multiple myeloma included in the IFM/DFCI 2009 trial,” *J. Clin. Oncol.*, vol. 35, no. 25, pp. 2911–2918, 2017, doi: 10.1200/JCO.2017.72.2975.

[7] C. Nanni *et al.*, “Image interpretation criteria for FDG PET/CT in multiple myeloma: a new proposal from an Italian expert panel. IMPeTUs (Italian Myeloma criteria for PET USE),” *Eur. J. Nucl. Med. Mol. Imaging*, vol. 43, no. 3, pp. 414–421, 2016, doi: 10.1007/s00259-015-3200-9.

[8] B. Jamet *et al.*, “Random survival forest to predict transplant-eligible newly diagnosed multiple myeloma outcome including FDG-PET radiomics: a combined analysis of two independent prospective European trials,” *Eur. J. Nucl. Med. Mol. Imaging*, vol. 48, pp. 1005–1015, 2021, doi: 10.1007/s00259-020-05049-6.

[9] C. Mesguich *et al.*, “Improved 18-FDG PET/CT diagnosis of multiple myeloma diffuse disease by radiomics analysis,” *Nucl. Med. Commun.*, vol. 42, no. 10, pp. 1135–1143, 2021, doi: 10.1097/MNM.0000000000001437.

[10] “MathWorks,” 2022. <https://es.mathworks.com> (accessed Jul. 13, 2022).

[11] M. Vallières, “Radiomics,” 2019. <https://github.com/mvallieres/radiomics> (accessed Jul. 18, 2022).

[12] M. Vallières, C. R. Freeman, S. R. Skamene, and I. El Naqa, “A radiomics model from joint FDG-PET and MRI texture features for the prediction of lung metastases in soft-tissue sarcomas of the extremities,” *Phys. Med. Biol.*, vol. 60, no. 14, pp. 5471–5496, 2015, doi: 10.1088/0031-9155/60/14/5471.

[13] “SPSS Statistics,” *IBM*, 2022. <https://www.ibm.com/products/spss-statistics> (accessed Jul. 18, 2022).

[14] “Orange Data Mining,” 2022. <https://orangedatamining.com/> (accessed Jul. 19, 2022).

[15] C. Bunkhumpornpat, K. Sinapiromsaran, and C. Lursinsap, “Safe-level-smote: Safe-level-synthetic minority over-sampling technique for handling the class imbalanced problem,” in *Advances in Knowledge Discovery and Data Mining. PAKDD 2009.*, T. Theeramunkong, B. Kijssirikul, N. Cercone, and T. Ho, Eds. Springer, Berlin, Heidelberg, 2009, pp. 475–482. doi: 10.1007/978-3-642-01307-2\_43.

# Señales Biomédicas I

---

Miércoles, 23 de noviembre de 2022

# Análisis de los cambios en la conectividad funcional tras un entrenamiento cognitivo mediante Neurofeedback

D. Marcos-Martínez<sup>1</sup>, A. Martín-Fernández<sup>1,2</sup>, S. Pérez-Velasco<sup>1</sup>, E. Santamaría-Vázquez<sup>1,3</sup>, V. Martínez-Cagigal<sup>1,3</sup>, S. Moreno-Calderón<sup>1</sup>, R. Hornero<sup>1,3</sup>

<sup>1</sup> Grupo de Ingeniería Biomédica (GIB), E.T.S Ingenieros de Telecomunicación, Universidad de Valladolid, Valladolid, España, diego.marcos@gib.tel.uva.es

<sup>2</sup> Escuela Politécnica Superior, Universidad Carlos III de Madrid, Madrid, España

<sup>3</sup> Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), España

## Resumen

*El Neurofeedback (NF) es una de las principales aplicaciones de los sistemas Brain-Computer Interfaces (BCI). Esta técnica busca inducir cambios en la actividad cerebral del usuario a través del aprendizaje de la modulación voluntaria de los ritmos neuronales. En este trabajo se presenta un análisis exploratorio de los efectos del NF desde una perspectiva novedosa en este campo de investigación: la conectividad funcional. Para ello, se empleó las señales de electroencefalograma (EEG) registradas durante un estudio de NF que constó de 6 sesiones de entrenamiento y dos de evaluación. En el estudio se dividió a los sujetos en un grupo de entrenamiento (GE), que entrenó su actividad theta (4 – 8 Hz), y un grupo de placebo (GP). Los posibles efectos del NF se analizaron empleando Phase Lag Index (PLI) en las bandas theta (4 – 8 Hz), alpha (8 – 13 Hz) y beta (13 – 19 Hz) durante la realización de la prueba de evaluación cognitiva “N-back” durante las sesiones de evaluación inicial y final. El análisis de conectividad entre sensores mostró una mayor tendencia al aumento de la conectividad en el GE con relación al GP, en especial en la banda theta. Por otro lado, se observó un incremento significativo del valor de conectividad global para la banda theta únicamente en el GE. Estos resultados podrían sugerir la efectividad del entrenamiento NF en theta realizado por el GE, así como la utilidad de las métricas de conectividad funcional para caracterizar los efectos del NF en la actividad cerebral.*

## 1. Introducción

Los sistemas *Brain-Computer Interfaces* (BCI) permiten controlar dispositivos externos mediante la actividad cerebral del usuario [1]. Para la adquisición de dicha actividad, el método más extendido en el campo de BCI es la señal de electroencefalograma (EEG) debido a su carácter no invasivo, su portabilidad y su bajo coste. Entre las distintas aplicaciones de los sistemas BCI se encuentra el entrenamiento mediante *Neurofeedback* (NF). En esta técnica se presenta al usuario una realimentación en tiempo real de una determinada característica de su actividad cerebral (p.ej. la potencia de una determinada banda de frecuencia) [2]. De esta manera, por medio del condicionamiento operante, se espera que la realimentación pueda ayudar al usuario a encontrar la estrategia mental que le permita adquirir un control voluntario sobre dicha característica de su actividad cerebral. Una vez probada la posibilidad de inducir neuroplasticidad mediante NF [2], es necesario investigar si estos cambios pueden resultar beneficiosos para aliviar

los síntomas de determinadas patologías relacionadas con desórdenes neuronales. En este sentido, se ha planteado el NF como posible tratamiento para el TDAH [2] o la epilepsia [3], como terapia de rehabilitación neuronal para pacientes de ictus [4], o como método para prevenir el deterioro cognitivo asociado a la edad [5].

A pesar del gran interés suscitado por el NF, se ha señalado la necesidad de ampliar las técnicas empleadas para estudiar los cambios en la actividad cerebral inducidos por el NF [6]. La mayoría de estas técnicas están basadas en el análisis espectral, lo cual podría no ser suficiente para determinar el impacto general del NF. En este sentido, se ha propuesto que los métodos de análisis no lineal [5] y los métodos de conectividad funcional [6] podrían aportar una perspectiva complementaria a los métodos tradicionales. En concreto, las métricas de conectividad funcional reflejan la dependencia estadística entre la actividad registrada en distintas regiones cerebrales, permitiendo obtener así una visión global de las interacciones entre dichas regiones [7]. Estas métricas se han mostrado útiles en la caracterización de la dinámica cerebral, tanto en personas sanas como en personas con patologías de origen neuronal [8].

La hipótesis en la que se basa este trabajo es que las métricas de conectividad funcional son útiles a la hora de analizar los posibles cambios en la actividad cerebral inducidos por el NF. Para evaluar dicha hipótesis, se ha empleado la métrica *Phase Lag Index* (PLI) [7] en registros de EEG obtenidos durante un estudio de NF con sujetos jóvenes sanos. El objetivo de este trabajo es, por tanto, explorar los posibles cambios en la conectividad funcional derivados de un entrenamiento de NF en un grupo experimental y un grupo placebo.

## 2. Materiales y métodos

### 2.1. Estudio de NF

El análisis de los cambios en la conectividad funcional se realizó sobre una base de datos obtenida durante un estudio de NF. En él participaron 19 sujetos jóvenes sanos (media de edad de  $25.05 \pm 4.18$  años, 11 mujeres). El estudio tuvo como objetivo el incremento de la actividad cerebral en la banda theta ( $\theta$ , 4-8 Hz) en la región frontal-medial (fm- $\theta$ ). Esta actividad ha sido relacionada con el desarrollo de tareas que implican a la memoria de trabajo, y se ha

propuesto que su entrenamiento puede repercutir positivamente en dicha función cognitiva [9].

El protocolo experimental se compuso de 8 sesiones de una hora de duración distribuidas a lo largo de dos semanas. En la primera y la última sesión se realizó una evaluación de la memoria de trabajo de los sujetos empleando 3 versiones informatizadas de los tests *N-Back*, *Digit span* y *Corsi block-tapping* [10]. Se realizó un registro del EEG de los sujetos mientras realizaron estas aplicaciones. Durante las otras seis sesiones se realizó el entrenamiento de NF. A lo largo del entrenamiento se emplearon tres aplicaciones diferentes para presentar una realimentación en tiempo real de la actividad del sujeto registrada en el electrodo Fz cada 250 ms. Cada sesión de NF se compuso de 2 registros de dos minutos del EEG del sujeto en estado de reposo (al inicio y al final de la sesión), y 6 bloques de 5 minutos cada uno de NF, con descansos entre cada bloque. Tanto las aplicaciones de evaluación cognitiva como el entorno de NF han sido desarrollados por el Grupo de Ingeniería Biomédica de la Universidad de Valladolid [10].

Con el objetivo de identificar posibles efectos no asociados al uso del NF, se dividió a los sujetos en dos grupos. Un grupo de entrenamiento (GE, 11 sujetos, 7 mujeres), a los cuales se les dio realimentación de su actividad  $fm-\theta$  durante las sesiones de NF, y un grupo de placebo (GP, 8 sujetos, 4 mujeres) a los cuales se les dio realimentación de su actividad variando el intervalo de la banda de frecuencia en cada sesión. Esto permitió que los sujetos de GP pudieran experimentar sensación de control durante la sesión. Las bandas empleadas para la realimentación del GP fueron: 16-18 Hz, 14-16 Hz, 18-20 Hz, 20-22 Hz, 12-14 Hz y 22-24 Hz.

Para el registro del EEG, tanto en las sesiones de evaluación como en las de entrenamiento, se empleó un amplificador g.USBamp (Guger Technologies OG, Graz, Austria) con 16 electrodos activos (F7, F3, Fz, F4, F8, FCz, C3, Cz, C4, CPz, P3, Pz, P4, PO7, POz, PO8), situados en un gorro elástico según el estándar 10-10. Para la tierra se empleó el canal AFz, y se situó una referencia común en el lóbulo de la oreja derecha. La señal se adquirió a una frecuencia de muestreo de 256 Hz. El procesado de la señal en tiempo real constó de un filtro frecuencial de paso banda de segundo orden tipo Butterworth y un filtro espacial tipo Laplaciano. Para evitar la influencia de pestañeos y de ruido muscular, la realimentación presentada fue nula siempre que se superó determinado umbral de potencia en las bandas 0.5-3.5 Hz y 35-40 Hz.

## 2.2. Análisis de conectividad funcional

Para llevar a cabo el análisis de conectividad funcional se emplearon los registros de EEG obtenidos durante la realización de la prueba *N-back* en la primera y última sesión. En esta prueba el sujeto debía prestar atención a los elementos que se presentaban por pantalla (cuadrados en una matriz de 3x3). Si el elemento presentado coincidía con el *N*-elemento anterior (siendo *N* un parámetro a definir: *N* = 1 representa el elemento anterior al presentado, *N* = 2 el anterior del anterior, etc.), el sujeto debía pulsar una tecla determinada. La prueba se realizó asignando un valor *N* = 3. En ambas sesiones se realizaron 6 bloques de *N-back*,

con 30 estímulos por bloque, de los cuales 12 requerían que el sujeto reaccionara. Se seleccionaron los registros de EEG de esta prueba ya que la menor probabilidad de observar un estímulo *target* que uno *non-target* genera una respuesta en el EEG que permite caracterizar el procesamiento cognitivo del estímulo.

De manera previa al análisis de conectividad funcional, se siguió el siguiente preprocesado *offline* de la señal. En primer lugar, se realizó un filtrado espectral, empleando un filtro *notch* en 50 Hz y un filtro FIR pasa banda entre 1 y 40 Hz. A continuación, la señal fue referenciada al promedio de la actividad común de todos los electrodos (CAR) con el objetivo de minimizar los efectos los movimientos oculares sacádicos [8]. Después, la señal fue dividida en épocas de 1.5 s de duración, tomando para el análisis únicamente los estímulos *target* que fueron respondidos correctamente por los sujetos. El intervalo se definió desde los 500 ms previos a la presentación del estímulo *target* a los 1000 ms posteriores. Sobre estas épocas se aplicó un método de rechazo automático de artefactos, descartando aquellas épocas cuya amplitud excediera un umbral basado en la desviación estándar de la señal [8]. Por último, las épocas sin artefactos fueron divididas en dos intervalos con el objetivo de caracterizar los cambios en la conectividad funcional durante la respuesta al estímulo [8]. Por un lado, se definió un intervalo de *baseline* desde 300 ms previos al estímulo hasta la presentación del estímulo. Por otro lado, se definió un intervalo de respuesta al estímulo, desde 150 a 450 ms posteriores al estímulo.

Los registros de EEG fueron analizados empleando la métrica PLI [7]. Esta métrica da cuenta del grado de grado de acoplamiento temporal entre dos señales a partir de la distribución de sus diferencias de fase instantánea:

$$PLI = \left| n^{-1} \sum_{k=1}^{k=n} \text{sgn}(\text{Im}[e^{i(\Delta\phi_k)t_k}]) \right|, \quad (1)$$

siendo  $\Delta\phi_k$  la diferencia de fases instantáneas de dos señales. PLI, a diferencia de otras métricas de conectividad funcional, es más robusta a la influencia de efectos de conducción de volumen [7].

Para el uso de PLI es necesario un ancho de banda de la señal analizada sea estrecho. Por ello, el análisis se llevó a cabo filtrando las épocas resultantes del preprocesado en las bandas theta ( $\theta$ , 4-8 Hz), alpha ( $\alpha$ , 8-13 Hz) y beta-1 ( $\beta$ , 13-19 Hz). La extracción de la fase instantánea de la señal se llevó a cabo mediante la transformada de Hilbert. Para poder estudiar los cambios dinámicos de conectividad funcional entre el pre-estímulo y el post-estímulo, se realizó una substracción simple de la matriz de adyacencia del periodo *baseline* ( $PLI_{BAS}^{ij}$ ) del de respuesta al estímulo ( $PLI_{RES}^{ij}$ ).

$$\Delta PLI^{ij} = \langle PLI_{RES}^{ij} - PLI_{BAS}^{ij} \rangle_{ep}, \quad (2)$$

siendo *ij* cada una de las posibles conexiones entre electrodos a estudiar y  $\langle \rangle_{ep}$  el promediado entre épocas. Se obtuvo así una única matriz de adyacencia  $\Delta PLI$  por cada banda de frecuencia y sujeto. Por otro lado, también se estudió los cambios en la conectividad funcional a nivel

global [8]. Para ello, se obtuvo un valor de conectividad global  $\Delta PLI_{Global}$  promediando los pesos de las conexiones entre todos los electrodos. Después, esto se promedió entre todas las épocas, obteniendo un único valor de conectividad global por banda de frecuencia y sujeto.

$$\Delta PLI_{Global} = \langle \langle PLI_{RES}^{ij} \rangle_{con} - \langle PLI_{BAS}^{ij} \rangle_{con} \rangle_{ep}, \quad (3)$$

donde  $\langle \rangle_{con}$  denota el promediado entre todas las posibles conexiones entre electrodos. Los valores negativos o positivos, tanto de  $\Delta PLI_{Global}$  como  $\Delta PLI$ , indican un decremento o incremento del *baseline* a la respuesta, respectivamente.

La influencia del NF se estudió analizando los cambios de estas métricas tras las 6 sesiones en ambos grupos (GE y GP). Para evaluar la significancia estadística de los posibles cambios se empleó el test de los rangos con signo de Wilcoxon ( $\alpha=0.05$ ) debido a que los datos no satisfacían la condición de normalidad.

### 3. Resultados

#### 3.1. Análisis de la conectividad entre electrodos

En el análisis de conectividad funcional se comparó el cambio en los valores de la matriz de adyacencia entre la evaluación inicial y la final ( $\Delta PLI_{post} - \Delta PLI_{pre}$ ). Esto se hizo para cada una de las bandas estudiadas ( $\theta$ ,  $\alpha$  y  $\beta$ ). En la Figura 1 se muestra el promediado global para los sujetos de GE y GP, para cada banda y cada sesión de evaluación.

Para  $\theta$  se observó un aumento generalizado de los valores en el GE, en especial entre Fz y los canales de las regiones central y posterior. Sin embargo, tras aplicar una corrección para evitar errores de tipo I (FDR-BH), no se observó significancia estadística en este incremento ( $p$ -valores  $> 0.05$ ). Por su parte, en el GP no se observó este incremento especial en las conexiones de Fz, ni se observó cambio significativo en sus conexiones ( $p$ -valores  $> 0.05$ ).

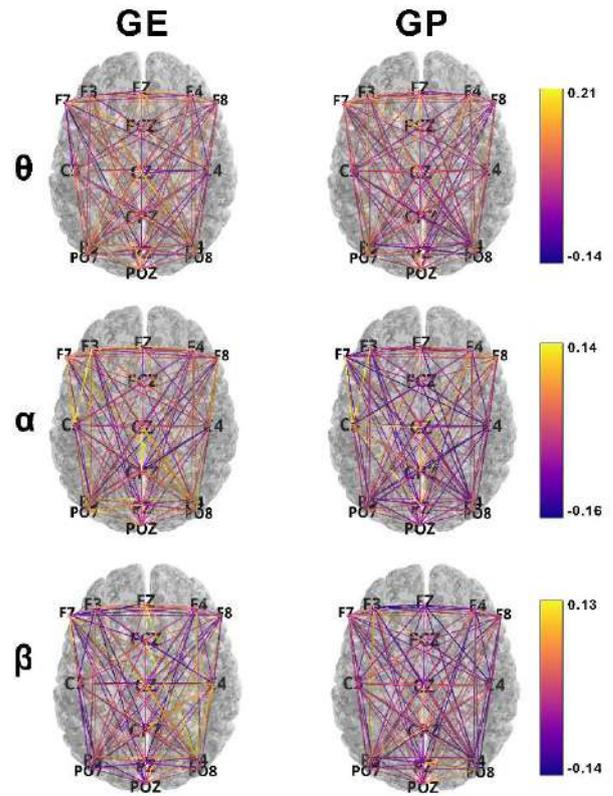
El análisis de  $\alpha$  mostró en el GE un incremento de la conectividad entre los canales frontales, así como entre los canales de la región parietal. Asimismo, se observó un decremento de la conectividad entre estas dos regiones. Ninguno de los cambios en la conectividad resultó ser estadísticamente significativo. Por otro lado, en el GP se observó un mayor decremento de las conexiones, especialmente entre los canales frontales y los canales P3 y P4. Ninguno de estos cambios en la conectividad resultó significativo.

Por último, para  $\beta$  en el GE se observó un aumento de la conectividad entre los canales frontales y P4. Por otro lado, este análisis también mostró una disminución de la conectividad entre los canales frontales y los parietales. Para el GP se observó un decremento generalizado de la conectividad, en especial entre los canales frontales y los parietales. No se encontró cambios estadísticamente significativos en ninguno de los dos grupos de sujetos.

#### 3.2. Análisis de la conectividad global

En la Tabla 1 se presenta un resumen de los  $p$ -valores obtenidos en el análisis de los cambios en  $\Delta PLI_{Global}$  para las 3 bandas de frecuencia estudiadas. Únicamente se

observó un incremento significativo de la conectividad global en la banda  $\theta$  para el GE. Esto no fue observado en el GP. No se observó ningún cambio significativo más para ninguno de los dos grupos de sujetos en el análisis de las otras dos bandas restantes.



**Figura 1.** Mapas de los cambios conectividad ( $\Delta PLI_{post} - \Delta PLI_{pre}$ ) promediados por grupo y para cada banda de frecuencia estudiada.

Banda	GE	GP
$\theta$	$p = 0.0175$	$p > 0.05$
$\alpha$	$p > 0.05$	$p > 0.05$
$\beta$	$p > 0.05$	$p > 0.05$

**Tabla 1.** Resultados de la prueba de rangos con signo de Wilcoxon para  $\Delta PLI_{Global}$ . Los  $p$ -valores han sido corregidos (FDR-BH).

### 4. Discusión

El objetivo de este trabajo ha sido explorar los cambios en el EEG inducidos por el NF desde la perspectiva de la conectividad funcional. Para ello, se han analizado los posibles cambios en la modulación del PLI durante el desarrollo de una prueba de evaluación cognitiva, tanto las relaciones entre electrodos como la conectividad global.

A pesar de que el análisis de los cambios de conectividad entre electrodos no arrojó valores significativos, se observó una mayor tendencia al incremento de los valores de conectividad en el GE que en el GP. Esto fue remarcable en la banda  $\theta$ , en la que predominó el aumento de la conectividad entre la región frontal y la parietal. Este resultado es de especial interés ya que estudios previos reportaron una relación entre un aumento de la

sincronización frontal-parietal y la realización de tareas cognitivas que implican el uso de la memoria de trabajo [11]. En este sentido, este resultado podría sugerir que el entrenamiento de NF planteado para mejorar la memoria de trabajo a través de fm- $\theta$  ha resultado efectivo [11]. Por otro lado, esta tendencia al incremento de la conectividad entre electrodos también se observó en el análisis de las bandas  $\alpha$  y  $\beta$  en el GE. A pesar de no haber encontrado diferencias estadísticamente significativas, las diferencias observadas con respecto a los cambios en el GP podrían deberse a la acción del NF. Si bien estas dos bandas de frecuencia no han sido entrenadas directamente durante el NF, estudios previos reportaron cambios en la potencia de otras bandas no entrenadas, argumentando una influencia global del NF en la dinámica cerebral [5].

En relación con los cambios en la conectividad global, el análisis reveló un incremento significativo en la modulación del PLI en la evaluación final con respecto a la inicial únicamente en el GE y en la banda  $\theta$ . Puesto que la métrica de conectividad global permite describir la interacción de todos los electrodos con un único valor, este resultado podría sugerir un efecto específico del NF en la sincronización de distintas regiones cerebrales en el rango de frecuencia de la banda  $\theta$ .

Sin embargo, nuestro trabajo no se encuentra exento de limitaciones. En primer lugar, sería deseable incrementar la población de estudio, tanto en el GE como en el GP. Esto aportaría una mayor potencia estadística a los resultados obtenidos. En segundo lugar, sería interesante estudiar los cambios en la conectividad funcional a lo largo de las sesiones de NF. Esto podría aportar más información sobre los efectos del NF en la actividad cerebral del usuario. Por último, el uso de la teoría de redes ha resultado relevante a la hora de caracterizar la actividad cerebral de diversas patologías neuronales [12]. Sin embargo, esta metodología no está aún extendida entre los estudios de NF, por lo que podría ser valioso y novedoso tenerlo en cuenta en futuros análisis.

## 5. Conclusiones

En este trabajo se han explorado los efectos del NF en la actividad cerebral desde la perspectiva de la conectividad funcional. Esta metodología, novedosa en el campo del NF, ha mostrado ser efectiva en la caracterización de los cambios en el EEG inducidos por el NF. En concreto, el análisis de los cambios en la modulación del PLI, tanto a nivel sensor como a nivel global, mostraron un incremento de la conectividad en el rango de frecuencia entrenado (4–8 Hz) únicamente en el GE. Estos resultados sugieren la efectividad del entrenamiento de NF planteado. Asimismo, enfatizan la necesidad de emplear esta metodología para obtener un mayor conocimiento de la influencia del NF en la dinámica cerebral.

## Agradecimientos

Este estudio ha sido financiado por los proyectos PID2020-115468RB-I00, RTC2019-007350-1 y PDC2021-120775-I00 del Ministerio de Ciencia e Innovación/Agencia Estatal de Investigación/10.13039/501100011033/ y FEDER Una forma de hacer Europa; por el proyecto I+D+i ‘Análisis y

correlación entre la epigenética y la actividad cerebral para evaluar el riesgo de migraña crónica y episódica en mujeres’ (‘Programa de Cooperación Interreg V-A España-Portugal POCTEP 2014-2020’) y FEDER; y por ‘Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN)’ a través de ‘Instituto de Salud Carlos III’. D. Marcos-Martínez, E. Santamaría-Vázquez y S. Pérez-Velasco son beneficiarios de una ayuda PIF de la Consejería de Educación de la Junta de Castilla y León.

## Referencias

- [1] J. Wolpaw *et al.*, *Brain-Computer Interfaces: Principles and Practice*. Oxford University Press, 2012.
- [2] R. Sitaram *et al.*, “Closed-loop brain training: The science of neurofeedback,” *Nat. Rev. Neurosci.*, vol. 18, no. 2, pp. 86–100, 2017.
- [3] M. B. Stermán, “Biofeedback in the treatment of epilepsy,” *Cleve. Clin. J. Med.*, vol. 77, no. SUPPL. 3, 2010.
- [4] M. A. Cervera *et al.*, “Brain-computer interfaces for post-stroke motor rehabilitation: a meta-analysis,” *Ann. Clin. Transl. Neurol.*, vol. 5, no. 5, pp. 651–663, 2018.
- [5] D. Marcos-Martínez *et al.*, “Neurofeedback training based on motor imagery strategies increases EEG complexity in elderly population,” *Entropy*, vol. 23, no. 12, pp. 1–19, 2021.
- [6] J. Gonzalez-astudillo *et al.*, “Network-based brain – computer interfaces: principles and applications Network-based brain – computer interfaces: principles and applications,” *J. Neural Eng.*, vol. 18, 2021.
- [7] C. J. Stam *et al.*, “Phase lag index: Assessment of functional connectivity from multi channel EEG and MEG with diminished bias from common sources,” *Hum. Brain Mapp.*, vol. 28, no. 11, pp. 1178–1193, 2007.
- [8] A. Bachiller *et al.*, “A comparative study of event-related coupling patterns during an auditory oddball task in schizophrenia,” *J. Neural Eng.*, vol. 12, no. 1, 2015.
- [9] S. Enriquez-Geppert *et al.*, “Modulation of frontal-midline theta by neurofeedback,” *Biol. Psychol.*, vol. 95, no. 1, pp. 59–69, 2014.
- [10] D. Marcos-Martínez *et al.*, “ITACA: Un nuevo sistema de entrenamiento cognitivo mediante Neurofeedback,” in *XIII Simposio CEA de Bioingeniería*, pp. 239–242, 2022.
- [11] B. Tóth *et al.*, “Frontal midline theta connectivity is related to efficiency of WM maintenance and is affected by aging,” *Neurobiol. Learn. Mem.*, vol. 114, pp. 58–69, 2014.
- [12] J. Gomez-Pilar *et al.*, “Altered predictive capability of the brain network EEG model in schizophrenia during cognition,” *Schizophr. Res.*, vol. 201, pp. 120–129, 2018.

# Metodología Robusta Basada en los Fundamentos del Machine Learning Para la Clasificación de Señales Biomédicas. Aplicación a 3 Desafíos de la Ingeniería Biomédica Moderna

S. Ruipérez Campillo<sup>1,2,3</sup>, F.S. Ramón Castells<sup>1</sup>, J. Millet Roig<sup>1</sup>

<sup>1</sup> Instituto ITACA, Universitat Politècnica de València, Valencia, España

<sup>2</sup> Escuela de Medicina, Universidad de Stanford, Palo Alto, CA, Estados Unidos

<sup>3</sup> D-ITET, Instituto Federal de Tecnología Suizo (ETH), Zúrich, Suiza {sruiperez@ethz.ch}

## Resumen

La clasificación no supervisada de señales médicas o secuencias biológicas es uno de los problemas de la bioingeniería moderna que no está resuelto de una forma general y extrapolable. En este estudio proponemos una metodología robusta y tratamos al detalle sus fundamentos matemáticos para representar estas señales en espacios prehilbertianos mediante transformadas espaciales, y utilizamos técnicas basadas en los fundamentos del machine learning y la teoría de la señal para clasificarlas. Para ilustrar la capacidad del método, lo aplicamos a tres desafíos de la bioingeniería moderna. Los resultados del primero muestran la capacidad de hacer un seguimiento a nivel neuronal de señales cerebrales. El segundo muestra una exactitud del 96% para clasificar proteínas de acuerdo a su función. El tercero permite clasificar señales sintéticas de flúter auricular en un espacio muy reducido. Por tanto, se propone un método robusto, controlado y generalizable para clasificar señales y secuencias biológicas.

## 1. Motivación

Las técnicas de inteligencia artificial son tendencia en ingeniería y aplicaciones médicas. Entre ellas, los métodos no supervisados están tomando un papel importante dada su flexibilidad e interpretabilidad [1]. Sin embargo, es común que se apliquen algoritmos sin conocer los fundamentos matemáticos que lo sustentan, o diseñar una metodología de forma adecuada, generalizable y replicable. Se propone pues un método robusto basado en la descripción matemática de la intersección de los fundamentos del *machine learning*, el álgebra en espacios prehilbertianos y la teoría de la señal para representar cualquier señal médica o secuencia biológica que pertenezca a un dominio concreto (espacio, tiempo, frecuencia, ...). Con este método, muchos problemas de la bioingeniería moderna se pueden resolver desde una nueva perspectiva. En concreto lo aplicamos a tres escenarios:

Seguimiento de las activaciones neuronales: aunque se han desarrollado muchas técnicas para detectar las activaciones neuronales, no existe un método para hacer un seguimiento de la actividad de cada neurona en sesiones grabadas en distintos momentos temporales [2].

Clasificación de proteínas de acuerdo con su función: Los métodos para caracterizar la estructura y función de proteínas son costosos en recursos y tiempo. Se han aplicado técnicas de *deep learning* (DL) para clasificar estas secuencias de acuerdo a su función, pero no existe ningún método preciso hasta el momento.

El diagnóstico de flúter auricular (FLA) de forma no invasiva solo ha sido tratado utilizando asunciones

fisiológicas que todavía no han sido demostradas. La clasificación de señales cardiacas es, en general un problema que no se resuelve con técnicas provistas de precisión y explicabilidad.

### 1.1. Materiales

Las señales neuronales han sido adquiridas en 17 sesiones, detectando 92 unidades en todas ellas, en un ratón mBY03. El dispositivo consiste en 64 electrodos flexibles en el neocórtex (1.5 mm). El software JRCLUST© se utiliza para la detección de activaciones neuronales.

Las secuencias de proteínas provienen de la base de datos SwissProt en formato FASTA. Las anotaciones provienen de la base de datos GOA (*Gene Ontology Annotation*) [7]. Estas secuencias se cruzan con Uniprot y se filtran. La base de datos utilizada se conforma de 9427 proteínas. Para el modelo utilizamos una muestra de 250 GPCR (receptor acoplada a proteínas G), 150 nucleoporinas, 150 histonas y 18 interleucinas. El desequilibrio permite probar la robustez de los métodos en casos reales.

Para el caso de FLA, se dispone de 8.000 bucles de vectorcardiogramas de 8 tipos de flúter (1000 por tipo) basándonos en una técnica ya publicada [3] que se basa en las zonas de conducción lenta [4], habiendo 4 diferentes regiones y 2 direcciones de conducción para cada una.

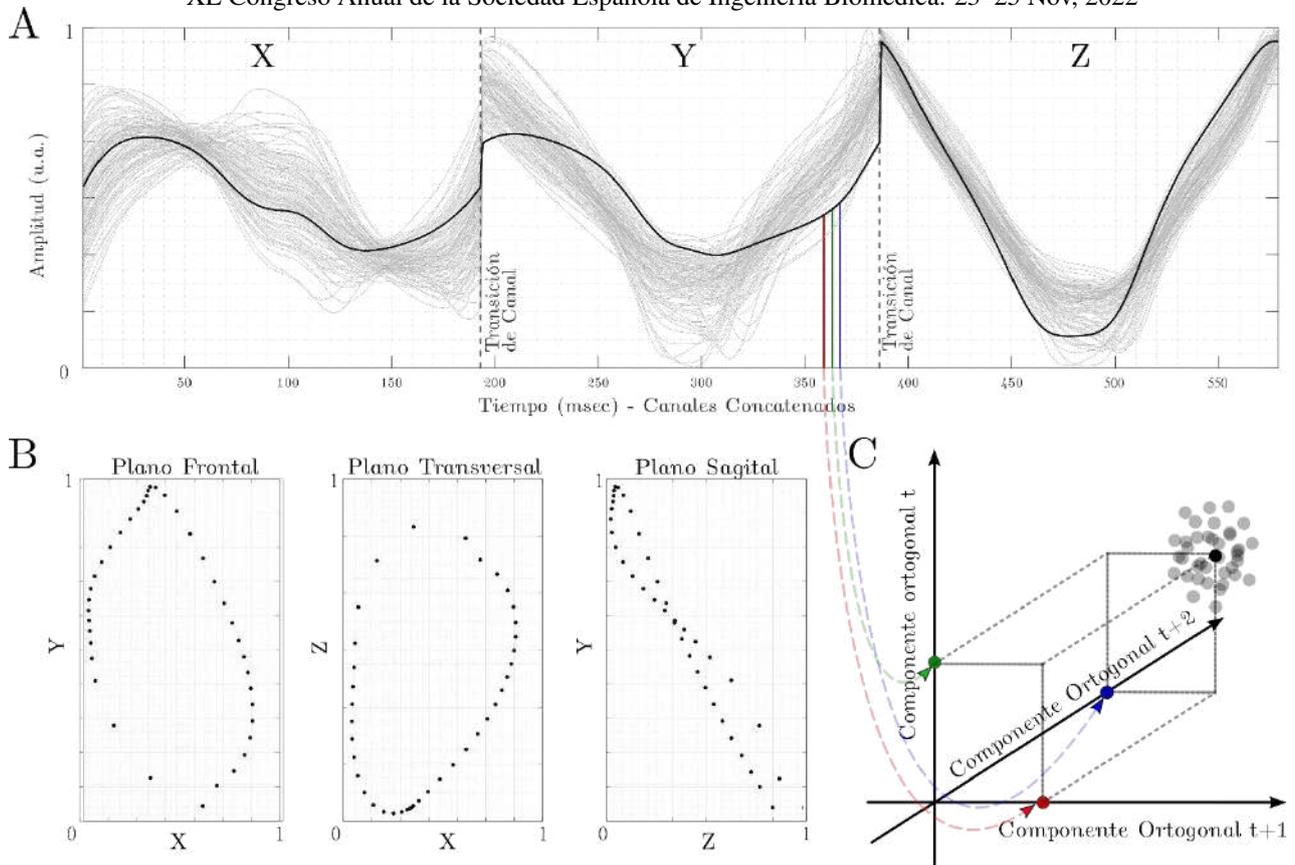
### 1.2. Representación General en Espacios Vectoriales

Definase  $\mathcal{C}$  como la representación de una distribución en el contexto de los *espacios prehilbertiano* – evitando por tanto la necesidad de generalizar la metodología a *espacios de Banach* u otros; y  $\mathcal{F}$  como el campo  $\mathcal{F} = \mathbb{R}$  ó  $\mathbb{C}$ .

Cualquier señal adquirida en varios canales puede representarse en función de un dominio como la concatenación de las ondas en los distintos canales como un vector en  $\mathcal{F}$ . Cada muestra del dominio en esta concatenación se traduce como la magnitud de éste en cada una de las componentes de este espacio dotado de producto escalar. Este vector es simplemente una generalización de  $\mathcal{F}^n$  de  $n$  tuplas reales (o complejas), por tanto análogo a la señal en particular de cada paciente (o identificador) – véase la figura 1.

### 1.3. Proyección de las Señales

Con ánimo de trabajar en un entorno eficiente a nivel computacional que se pueda generalizar a cualquier análisis, independiente de su complejidad o al tamaño de la base de datos, se propone un algoritmo de reducción de dimensionalidad, basado en descomposición de valores



**Figura 1. A:** Concatenación de los 3 canales de las señales de una clase de FLA. Cada una de las muestras provee la magnitud de cada una de las componentes del espacio vectorial (véase C) de  $n$  dimensiones análogo (véanse los ejemplos azul, rojo y verde para una representación con  $n=3$  de la correspondiente sección de la señal). Un caso en particular se marca en línea gruesa para ejemplificar, con la señal original representada en B.

singulares (SVD) y análisis de componentes principales (PCA). El subespacio vectorial resultante comprende el hospedaje de la proyección de los vectores originales en un espacio prehilbertiano de menor dimensionalidad. En particular, se cumple que para un espacio original  $\mathcal{S}$ , todas las *secuencias de Cauchy* de las muestras de la señal (neuronal, cardíaca, o representación de la proteína) en dicho espacio convergen a un elemento del propio espacio, además de que  $\forall \lambda_i \in \mathbb{R}^+$ , existe un producto escalar que podemos entender como una función  $\mathcal{S} \times \mathcal{S} \rightarrow \mathbb{F}$  de modo que  $(v, w) \mapsto \langle v, w \rangle, \forall v, w \in \mathcal{S}$ .

Con este fin, podemos definir  $\mathcal{S}$  como el espacio  $m$ -dimensional original donde nuestras muestras se representan en primer lugar, y  $\mathcal{R}$  como un subespacio de  $\mathcal{S}$  sobre  $\mathcal{F}^p$ . La única condición a tener en cuenta es que  $\forall s \in \mathcal{S}$  y  $f \in \mathcal{F}$ , se cumple que  $f \cdot s \in \mathcal{R}$  y  $\forall s, s' \in \mathcal{S}, s + s' \in \mathcal{S}$ . Por tanto, para el subconjunto  $V = \{v_1, v_2, \dots, v_N\} \subset \mathcal{S}$  sobre  $\mathcal{F}$ , las bio-senales o secuencias componen un espacio generado por la expansión de  $V$  tal que

$$\text{gen}(v_1, v_2, v_3, \dots, v_N) := \left\{ \sum_{\kappa} \xi_{\kappa} v_{\kappa} : \xi_{\kappa} \in \mathcal{F} \right\}$$

Ahora, para reducir las dimensiones y transformar de un espacio original a otro, representamos las señales en  $\mathcal{S}$  como  $y_1, y_2, \dots, y_N \in \mathcal{F}^m$ , con  $N$  el número de pacientes del espacio  $m$ -dimensional, y restamos la media ( $\sum_{\kappa} y_{\kappa} = 0$ ). La proyección al espacio  $P$ -dimensional supone  $\mathcal{F}^m \mapsto \mathcal{F}^P$  con el objetivo de minimizar:

$$\min_{\tilde{y}} \left( \sum_{\kappa} \|y_{\kappa}\|_2 - \sum_{\kappa} \|\tilde{y}_{\kappa}\|_2 \right)$$

La matriz de proyección se puede definir por tanto como  $\Theta : \Theta^H \Theta = I_P, \mathbb{F}^m \rightarrow \mathbb{F}^P : y \mapsto \tilde{y} = \Theta \Theta^H y$ .

#### 1.4. Midiendo Distancias

De las variadas métricas que se han propuesto en la literatura, escogemos la *distancia de Mahalanobis* ( $d_m$ ) que considera la distribución de las muestras en un espacio de múltiples variables, ponderando la cercanía mediante la inversa de la matriz de covarianza ( $Q := K_{\zeta\zeta}^{-1}$ ) de una distribución -naturalmente no singular. En otras palabras, ¿cuántas desviaciones estándar está el punto  $\zeta \in \mathcal{F}^m$  de la media  $\mu \in \mathcal{F}^m$  de la distribución  $\mathcal{C}$ ? Definimos por tanto:

$$d_M = \sqrt{(\zeta - \mu)^T Q (\zeta - \mu)}$$

Nótese que la elección de la métrica no es intuición (sobre otras como la *distancia de Chebyshev*, *Minkowsky* en un espacio normalizado,...) sino que se hicieron pruebas en modelos controlados con distribuciones Gaussianas con densidad:

$$f_{\zeta}(\zeta_1, \dots, \zeta_M) = \frac{e^{-\frac{1}{2}(\zeta - \mu)^T Q (\zeta - \mu)}}{\sqrt{(2\pi)^M |K_{\zeta\zeta}|}}$$

para el caso no degenerado, siendo la generalización a casos degenerados sencillamente la analogía utilizando la pseudoinversa de Moore-Penrose:  $Q := K_{\zeta\zeta}^{\dagger}$ .

#### 1.5. Agrupamiento Jerárquico

Para llevar a cabo las clasificaciones vectoriales utilizando las métricas mencionadas utilizamos métodos aglomerativos de forma estratificada. Proponemos el

método de Ward dado que es el más eficiente en nuestros modelos, implementándolo de la siguiente manera:

$$d_w(u, v) = \sqrt{\frac{|v| + |r|}{C} d_w(u, v)^2 + \frac{|v| + |l|}{C} d_w(u, l)^2 - \frac{|v|}{C} d_w(r, l)^2}$$

donde combinando  $r$  y  $l$  formamos  $u$ , con  $v$  libre, y representando la cardinalidad con  $|\cdot|$  ( $C = |v| + |r| + |l|$ ).

### 1.6. Especificaciones para cada Desafío

Para detectar el número de clases en el *desafío primero* y *segundo* se utiliza la métrica de *Calinski-Harabasz* y la ilustración del diagrama de coste. Técnicas estadísticas clásicas se proponen para evaluar la exactitud de las clasificaciones en base a la anotación manual de expertos en el campo. Además, para el *desafío segundo* se precisa de un preprocesado para representar las proteínas como vectores en el espacio original. Se utiliza una adaptación del modelo de redes neuronales *DeepGo* y el modelo *Go* [5], junto con técnicas de codificación *one-hot*.

En el *desafío tercero* se define la métrica  $C$  para clasificar las simulaciones de acuerdo a su dirección y velocidad, que supone la caracterización del tipo de FLA.

## 2. Resultados

### 2.1. Desafío Primero: Identificación de Neuronas

El algoritmo es capaz de identificar las distintas neuronas para cada una de las adquisiciones en todas las grabaciones con una precisión superior al 99% y un  $p$ -valor < 0.01 basados en las métricas derivadas de la *distancia de Mahalanobis*, tomando la anotación de expertos y la detección de activaciones del software JRCLUST© - véase la figura 2.

Con el propósito de identificar la actividad neuronal a lo largo del tiempo en distintas sesiones, el *método de Ward* permite el seguimiento de cada neurona a lo largo del tiempo, y la identificación de nueva actividad, o la pérdida de actividad de las mismas, con precisión absoluta contrastada con la anotación de expertos. Véase la figura 2.

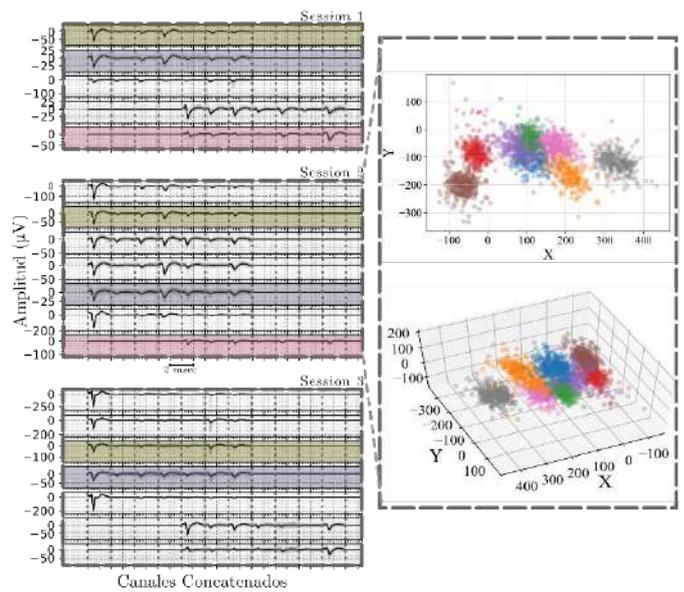
### 2.2. Desafío Segundo: Clasificación de Proteínas

El resultado de la clasificación de proteínas de acuerdo con su función se refleja en la tabla 1, destacando la precisión.

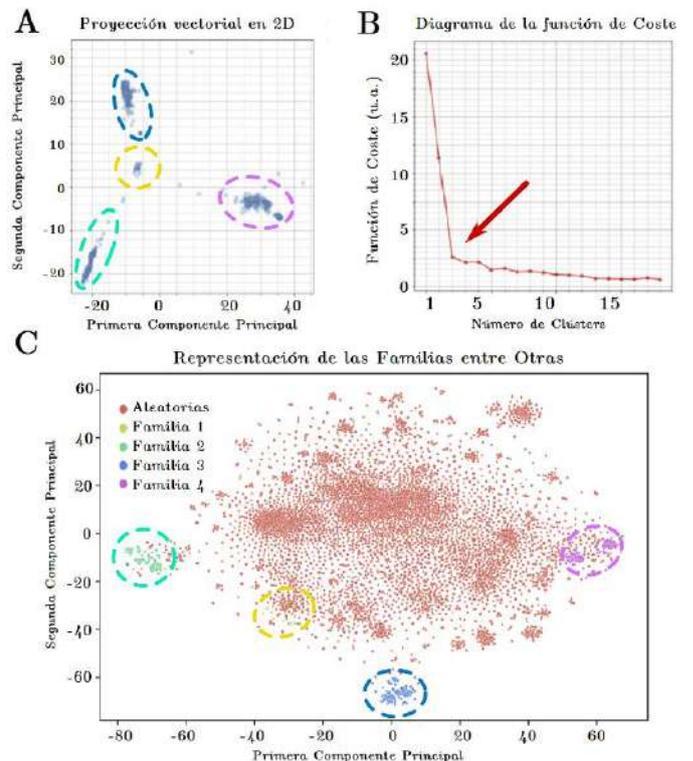
	Precisión	Exhaustividad	Punt. F1	Exactitud
General	N.A.	N.A.	N.A.	0.96
Micro	0.96	0.96	0.96	N.A.
Macro	0.85	0.97	0.87	N.A.
Ponderada	0.98	0.96	0.97	N.A.

**Tabla 1.** Estadística de la clasificación de las familias de proteínas basándose en su función. N.A. significa 'No Aplica'.

El resultado de la agrupación no supervisada y las técnicas de *machine learning* para la clasificación de las proteínas se resume en la figura 3. El panel A muestra la proyección en un espacio bidimensional. El panel B muestra la selección no supervisada del número ideal de familias, a partir de la función de coste – ilustración de la métrica de *Calinski-Harabasz*. El panel C muestra, en una proyección bidimensional, 9000 proteínas de cientos de familias escogidas de forma aleatoria junto con las cuatro familias seleccionadas.



**Figura 2.** *Izquierda:* Representación de tres sesiones y tres ejemplos de las neuronas identificadas y su seguimiento. La neurona roja desaparece, dado que el algoritmo detecta que ninguna actividad se corresponde con ese identificador. Las neuronas verde y azul se detectan en las tres sesiones. *Derecha:* Para la sesión 2, la representación en 2 y 3 dimensiones de la actividad neuronal con un color para cada neurona –  $X$  y  $Y$  representan componentes principales en unidades arbitrarias.



**Figura 3.** Resultados de la clasificación de proteínas de acuerdo a su función: **A:** Proyección 2D; **B:** Función de coste en base al número de clústers; **C:** Representación de las familias seleccionadas entre miles de proteínas aleatorias. Familia 1 = 18 interleukinas; Familia 2 = 250 GPCRs, Familia 3 = 250 nucleoporinas; Familia 4 = 250 histonas.

### 2.3. Desafío Tercero: Diagnóstico de FLA

El resultado del parámetro C se evalúa en distintos subespacios n-dimensionales con una n pequeña. Los resultados de las agrupaciones de vectores se evalúan sobre clústers con la misma velocidad y dirección de conducción, con solo uno de ellos, y con distintos valores para ambos parámetros. Los resultados se resumen en las tablas 2 y 3.

Dimensiones	Grupos	C
Varianza del espacio = 0.996		
n = 6	= Velocidad, = Dirección	0.35
	= Velocidad, ≠ Dirección	0.93
	≠ Velocidad, ≠ Dirección	0.04
Varianza del espacio = 0.880		
n = 3	= Velocity, = Direction	0.47
	= Velocidad, ≠ Dirección	0.94
	≠ Velocidad, ≠ Dirección	0.02
Varianza del espacio = 0.820		
n = 2	= Velocity, = Direction	0.94
	= Velocidad, ≠ Dirección	0.93
	≠ Velocidad, ≠ Dirección	0.02

**Tabla 2.** Análisis de la métrica de la distancia y varianza.

Dimensiones	Parámetro	Valor
# Grupos = 8		
n = 3	Macro-Prec = Recall = F1 Sc.	0.874
	Total TP	6998
	Total FP = Total FN	1002
# Grupos = 4		
n = 3	Macro-Prec = Recall = F1 Sc.	0.999
	Total TP	7994
	Total FP = Total FN	6
# Grupos = 4		
n = 2	Macro-Prec = Recall = F1 Sc.	0.997
	Total TP	7980
	Total FP = Total FN	20

**Tabla 3.** Análisis estadístico de la clasificación de AFL.

### 3. Discusión y Conclusiones

Los resultados del *desafío primero* ilustran una alta capacidad, basada en la anotación de expertos, de seguir la actividad neuronal a lo largo de distintas sesiones, grabadas en distintos momentos temporales (entre 1 y 10 días entre las mismas). Además, el modelo demuestra ser robusto al identificar el momento en el que una neurona deja de expresar activaciones, detectando la pérdida de la misma, mientras que mantiene el seguimiento de aquellas cuyas activaciones cambian a lo largo de las sesiones. La utilidad de estos resultados es inmediata, pues permite evaluar el efecto de nuevos electrodos implantados en el tejido cortical, uno de los grandes desafíos del desarrollo de dispositivos implantados en el cerebro [6]. Además, permite evaluar el cambio de la actividad neuronal en unidades específicas, lo que es de gran utilidad para evaluar el efecto de estímulos, tratamientos, y enfermedades.

En el *segundo caso* podemos observar una ilustración perfecta de los métodos utilizados: Las agrupaciones son claras por sí mismas (fig. 3A), la identificación de los grupos de forma no supervisada es indudable (fig. 3B) y la capacidad de distinguir las familias en un grupo de miles de otras proteínas (ruido para el modelo) es clara (fig. 3C). Estos resultados ofrecen una nueva visión para el desarrollo de productos farmacéuticos basados en la función de las proteínas (cuyos métodos tradicionales son costosos y toman mucho tiempo), para la medicina

personalizada y el tratamiento de enfermedades basándose en la genética específica del individuo en particular. Por ejemplo, si introducimos una proteína desconocida en el algoritmo, y la función se podrá identificar midiendo la distancia a los distintos grupos del *pre-entrenado* espacio vectorial. Además, la generalización de la metodología a secuencias abre infinitas posibilidades de aplicación al campo de la genética.

Finalmente, el *desafío tercero* provee al artículo de una prueba controlada del comportamiento de las diferentes partes de la metodología. Aunque se entienden las limitaciones de un modelo sintético, la aplicación a casos clínicos es el siguiente paso. Sin embargo, la gran capacidad de discernir entre los casos sintéticos de este modelo basado en señales fisiológicas y previamente publicado [3] permite demostrar la capacidad de clasificación y el interesante resultado de la alta varianza en espacios muy reducidos, permitiendo discernir entre tipos de AFL en tan solo 3 o 6 dimensiones – remuestrear una señal de 500 muestras a 3 suena inviable, mientras que esta reducción análoga en el modelo planteado es posible. Esto permite vislumbrar el diagnóstico no invasivo de señales cardíacas, con un gran impacto en la salud de los pacientes.

En resumen, la generalidad del método con pilares matemáticos y aplicación a señales clínicas y simulaciones controladas proveen a esta técnica de una potente aplicabilidad, explicabilidad, replicabilidad y precisión.

### Agradecimientos

*Este proyecto ha sido financiado por el Ministerio de Ciencia e Innovación del Gobierno de España (PID2019-109547RB), el Instituto de Salud Carlos III (CIBERCV CB16/11/00486) y la Generalitat Valenciana (PROMETEOII/2014/037). Gracias también a la Fundación Rafael del Pino y La Caixa por el apoyo proporcionado para la formación del primer autor del artículo. Se agradece también la colaboración del grupo de F. Yanik del Instituto de Neuroinformática del ETH Zürich y el grupo de J.L Merino de la Unidad de Arritmias y Electrofisiología Robotizada del Hospital Universitario la Paz.*

### Referencias

- [1] M. Cikes, S. Sanchez-Martinez, et al., “Machine learning-based phenogrouping in heart failure to identify responders to cardiac resynchronization therapy,” *European journal of heart failure*, vol. 21, no. 1, pp. 74–85, 2019.
- [2] N. Steinmetz et al., “Neuropixels 2.0: A miniaturized high-density probe stable, long-term brain recordings,” *Science*, vol. 372, no. 6539, 2021.
- [3] S. RUIPÉREZ-CAMPILLO, et al. “Non-invasive characterisation of macroreentrant atrial tachycardia types from a vectorcardiographic approach with the slow conduction region as a cornerstone”. *Computer methods and programs in biomedicine*, 200, p.105932, 2021.
- [4] S. RUIPÉREZ-CAMPILLO, et al. “Slow Conduction Regions as a Valuable Vectorcardiographic Parameter for the Non-Invasive Identification of Atrial Flutter Types.” In *2020 Computing in Cardiology*, pp. 1-4. IEEE, 2020.
- [5] M. Kulmanov, M, et al. “DeepGO: predicting protein functions from sequence and interactions using a deep ontology-aware classifier”. *Bioinformatics*, 34(4), pp.660-668, 2018
- [6] M. Gulino, et al. Tissue response to neural implants: the use of model systems toward new design solutions of implantable microelectrodes. *Frontiers in neuroscience*, 13, p.689, 2019
- [7] Huntley, Rachael P., et al. "The GOA database: gene ontology annotation updates for 2015." *Nucleic acids research* 43.D1 (2015): D1057-D1063.

# Analysis of surface electromyographic parameters for the assessment of muscle fatigue during moderate exercises

I. Junquera Godoy<sup>1</sup>, J. Gomis-Tena<sup>1</sup>,  
J.L. Martínez-de-Juan<sup>1</sup>, J. Saiz<sup>1</sup>, G. Prats-Boluda<sup>1</sup>

<sup>1</sup> Centro de Investigación e Innovación en Bioingeniería, Universitat Politècnica de València, 46022 Valencia, Spain

{ijungod<sup>1</sup>, jgomiste<sup>1</sup>, jlmartinez<sup>1</sup>, jsaiz<sup>1</sup>, gprats<sup>1</sup>} [@ci2b.upv.es](mailto:ci2b.upv.es)

## Summary

*Muscle fatigue is a neuromuscular phenomenon which prevents the muscles to generate the required force. Although there are multiple studies about detecting this condition with surface electromyography, this method has still some limitations. The objective of this study was to determine the viability of some biomarkers as early indicators of fatigue during moderate exercise. Eight healthy volunteers performed 2 dynamic exercises involving the flexo-extension of the knee lifting 2 and 4kg attached to the ankle. A bipolar electromyographic signal was recorded simultaneously from: Rectus Femoris, Vastus Lateralis and Biceps Femoris. Muscle activations were extracted by supervised automatic segmentation and their root mean square, bandwidth, entropy and spectral moment were calculated. The maximum of the normalized coefficient of the linear cross-correlation between each muscle pair were also computed as well as the slope throughout the execution of the exercise repetitions of all above mentioned parameters. Results from 2kg exercise did not reveal clear signs of fatigue. By contrast those from 4kg showed an increase of the root mean square and spectral moment ratio, and decrease of bandwidth, sample entropy and linear cross correlation coefficient, pointing out fatigue. In this study, it has been confirmed the viability of the cross-correlation technique as a possible new biomarker related to the first signs of muscle fatigue.*

## 1. Introduction

Muscle fatigue is a complex neuromuscular condition developed by a chain of structural, energetic and metabolic variations due to lack of oxygen and other necessary nutritive substances in combination with changes in the performance of the nervous system. These changes make it impossible to the muscle to generate the required force. This phenomenon can appear in healthy and pathological subjects and is one of the most common symptoms in some neurological disorders such as Parkinson or multiple sclerosis [1, 2].

There are several methods to estimate muscle fatigue such as measuring the lactate concentration in a muscle taking blood samples, being an invasive procedure. Another method is timing a person doing certain task until their surrender. This last method is known as the ‘mechanical manifestation of muscle fatigue’. One of the main problems of this method is that there is no insight to the physiological manifestations of the muscle during the activity and detects the fatigue only after it has already occurred. Moreover, it cannot be distinguished the origin of fatigue between physical or psychological factors [2].

Surface electromyography (sEMG) is a non-invasive method that can be used to assess local muscle condition in real time. Thus, EMG could target the fatigue of specific muscles and can provide early biomarkers of fatigue. The parameters most frequently used to determine muscle condition from EMG are their Root Mean Square (RMS), mean frequency or median frequency [1,2]. During the contraction of the muscles, the increase of RMS and the shift of the spectrum to lower frequencies are indicators of muscle fatigue [3]. However, despite the multitude of studies aiming the detection of fatigue, and the advances in new signal processing algorithms, it is still very complex to analyse these biosignals due to their variations, their lack of repeatability and their sensitivity to electrode location. On the other hand, there is great controversy about the performance of EMG-derived biomarkers for the early detection of fatigue [1,2,4].

The aim of this study is to analyse and compare traditional and novelty s-EMG parameters, during a non-extenuating protocol to assess the viability of these biomarkers as possible indicators of mild fatigue.

## 2. Materials and methods

The sEMG signals were recorded using an electromyograph designed and manufactured by our research group. This is a portable device, with three acquisition channels with a sampling rate of 1000Hz, bandwidth between 8.75 and 500Hz, and a gain of 24V/V. It is based on the microcontroller PIC24FJ128GC006 and the analog front-end ADS1299-4. The communication is wireless using Bluetooth technology, specifically a RN-42 module.

Eight healthy subjects, 2 women and 6 men, with an age of 33 ±13 years and with no history of neurological or neuromuscular disorders participated in this study. Three bipolar sEMG signals were recorded placing two Ag – AgCl disc type disposable electrodes with an inter electrode distance of 20mm over the muscle belly of Rectus Femoris (RF), Vastus Lateralis (VL) and Biceps Femoris (BF). Previously skin was cleaned with alcohol and shaved when needed.

The subjects warmed up for 3 minutes walking on a treadmill. Then, 2 dynamic exercises of flexo-extension of the knee were executed with 2 minutes of rest between exercises, lifting 2kg attached to their ankle and doing 15

repetitions with each leg, the same protocol was repeated for 4kg. Muscle activations from the 3 bipolar sEMG recorded signals (one from each muscle) were automatically segmented [5] and 15 muscle activation signals (one for each repetition) were extracted from each muscle and exercise.

Afterwards, a set of parameters from sEMG activations was obtained: bandwidth - difference in frequency between 5% and 95% of the total power in the spectrum, RMS [6], sample entropy [7], Spectral Moment Ratio (SMR) [8] and the maximum of the normalized coefficient of the linear cross-correlation (MCC) between each pair of muscles [9] and the slope of these parameters over the 15 muscle activations was analysed calculating the slope for each muscle.

### 3. Results and discussion

In Table 1 it is displayed the mean of the sEMG features for the first and last 5 muscle activations for the exercises lifting 2 and 4kg.

Mean $\pm$ std values	Activ.	RF	VL	BF
Bandwidth [Hz] 2kg	[1-5]	317.9 $\pm$ 15.3	285.2 $\pm$ 17.6	365.8 $\pm$ 20.9
	[10-15]	308.3 $\pm$ 9.4	284.0 $\pm$ 15.4	373.1 $\pm$ 12.0
Bandwidth [Hz] 4kg	[1-5]	317.4 $\pm$ 11.1	285.2 $\pm$ 14.2	369.2 $\pm$ 13.2
	[10-15]	307.8 $\pm$ 8.8	281.6 $\pm$ 13.9	363.2 $\pm$ 15.2
RMS [ $\mu$ V] 2kg	[1-5]	49.1 $\pm$ 7.0	68.2 $\pm$ 12.1	27.0 $\pm$ 3.9
	[10-15]	52.7 $\pm$ 6.1	74.0 $\pm$ 8.5	28.3 $\pm$ 2.5
RMS [ $\mu$ V] 4kg	[1-5]	58.4 $\pm$ 9.7	70.3 $\pm$ 10.8	28.5 $\pm$ 3.7
	[10-15]	66.2 $\pm$ 6.8	81.3 $\pm$ 9.7	30.9 $\pm$ 3.3
Entropy 2kg	[1-5]	1.74 $\pm$ 0.16	1.11 $\pm$ 0.15	1.37 $\pm$ 0.13
	[10-15]	1.82 $\pm$ 0.12	1.06 $\pm$ 0.12	1.33 $\pm$ 0.11
Entropy 4kg	[1-5]	1.86 $\pm$ 0.14	1.12 $\pm$ 0.13	1.38 $\pm$ 0.12
	[10-15]	1.85 $\pm$ 0.13	1.07 $\pm$ 0.11	1.32 $\pm$ 0.13
SMR 2kg	[1-5]	-30.4 $\pm$ 0.15	-29.8 $\pm$ 0.23	-30.9 $\pm$ 0.33
	[10-15]	-30.5 $\pm$ 0.10	-29.9 $\pm$ 0.12	-31.0 $\pm$ 0.16
SMR 4kg	[1-5]	-30.6 $\pm$ 0.17	-29.9 $\pm$ 0.14	-31.0 $\pm$ 0.27
	[10-15]	-30.6 $\pm$ 0.10	-29.9 $\pm$ 0.13	-31.1 $\pm$ 0.22

**Table 1.** Mean and std values of the traditional parameters for RF (Rectus Femoris), VL (Vastus Lateralis) and BF (Biceps Femoris) muscles. during the first (1-5) and last (10-15) 5 muscle activations of the exercise.

An increase of RMS and decrease in bandwidth and sample entropy, between the beginning and the end of the exercises was observed for all muscles pointing out possible muscular fatigue. SMR did not reveal an observable difference between the mean value at the beginning and the end of the exercise. However, a more in-depth analysis of its trend of all the parameters will be carried out.

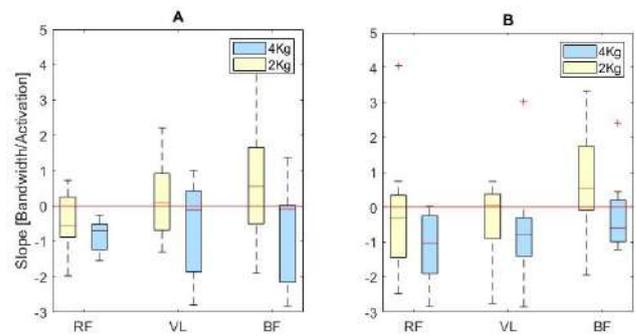
Table 2. represent the MCC mean and standard deviation of the first and last 5 muscle activations between each pair of muscles for the exercises lifting 2 and 4kg. In this case, there was a decrease of the correlation during the exercise.

Mean $\pm$ std values	Activ.	RF-VL	RF-BF	VL-BF
MCC 2kg	[1-5]	0.17 $\pm$ 0.04	0.13 $\pm$ 0.03	0.13 $\pm$ 0.03
	[10-15]	0.15 $\pm$ 0.03	0.12 $\pm$ 0.03	0.11 $\pm$ 0.03
MCC 4kg	[1-5]	0.17 $\pm$ 0.04	0.13 $\pm$ 0.03	0.13 $\pm$ 0.03
	[10-15]	0.15 $\pm$ 0.03	0.12 $\pm$ 0.03	0.11 $\pm$ 0.03

**Table 2.** Mean and std of the cross-correlation value between each pair of muscles during the first (1-5) and last (10-15) 5 muscle activations of each exercise.

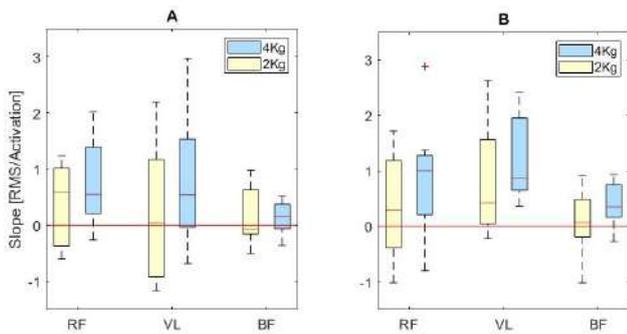
Then, for a more in-depth analysis of the evolution of the parameters, the calculation of their slope was carried out.

Fig. 1. shows the tendencies of the bandwidth evolution throughout the exercises. The slope of every muscle for the 4kg exercise is negative, this means, there was a reduction of the bandwidth throughout the exercise, which, according to the literature, is a sign of muscle fatigue [10]. It has also been noticeable that the agonist muscles exhibit less positive and more negative slopes than the antagonist muscle. This could imply that the agonist muscles point to a higher fatigue than the antagonist muscle.



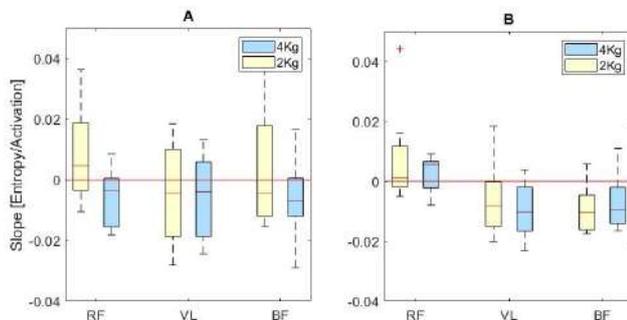
**Fig. 1.** Box & Whisker plots of the bandwidth slopes through the 15 repetitions of the dynamic exercise with 2kg (yellow) and 4kg (blue) computed for each muscle. (A) Left leg, (B) Right leg.

The results display an overall increase of the RMS (positive slope) in all exercises and muscles of both legs (Fig. 2). In this case, according to existing bibliography, the increase of the RMS values indicates muscle fatigue, although this parameter appears controversial due to difficulties in normalisation of the RMS [2,3]. RMS values for 4kg also has a higher slope comparing with 2kg, as occurs in the preview parameter, due to a higher effort and the outcome of a quicker fatigue state. Analysing the differences between agonists and antagonist muscles, the antagonist muscle exhibits less steep slope than the agonists.



**Fig. 2.** Box & Whisker plots of the RMS slopes through the 15 repetitions of the exercise with 2kg (yellow) and 4kg (blue). (A) Left leg, (B) Right leg.

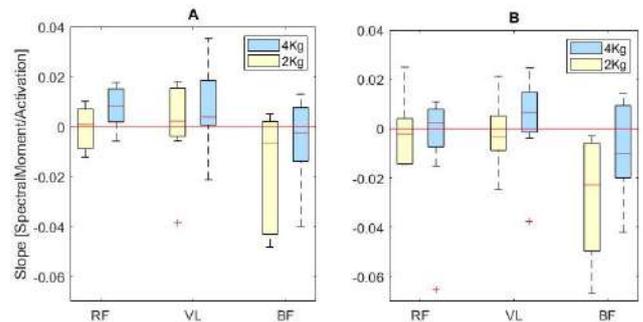
Inspecting the slope for the sample entropy during the execution of the exercises, in Fig. 3., it is appreciable generally negative slopes except for the Rectus Femoris muscle in the right leg pointing out muscle fatigue as stated in the literature [11]. The probable cause of this effect is that the Rectus Femoris is the most trained muscle of the leg, and because all subjects are dexterous, and therefore less prone to fatigue. Moreover, the results of this parameter in this study are quite controversial due to various reasons. Firstly, there is no observable difference between the exercise with 2kg and 4kg in most of the cases. Secondly, the agonist muscles use to have similar behaviour in contraposition to the antagonist muscle, which is something that cannot be observed in this instance. According to bibliography, sample entropy is a parameter with low susceptibility to fatigue detection [12].



**Fig. 3.** Box & Whisker plots of the sample entropy slopes through the 15 repetitions of the exercise with 2kg (yellow) and 4kg (blue). (A) Left leg, (B) Right leg.

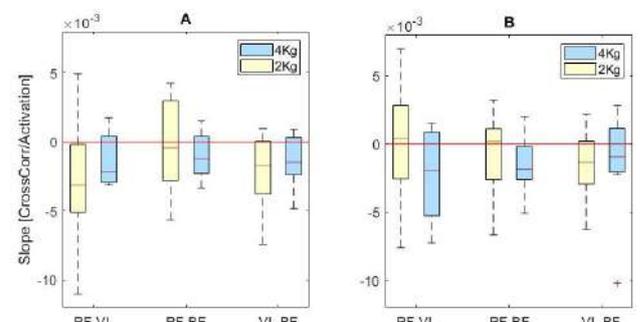
Regarding the slope of the SMR (see Fig. 4) results show an evident displacement from more negative slopes with 2kg exercises to a more positive slopes with 4kg exercises in all muscles, according with the literature [8]. This increase of the slope is more blatant for agonist muscles (as the slopes reach positive values) than the antagonist muscle (the slopes become less negative, but not yet positive). Agonist muscles of the right leg are showing negative tendencies for 2kg exercises, which could be since the subjects are right-handed, their right leg is more trained leg, which may require more effort to show muscle fatigue. Although it doesn't show a clear tendency for mild fatigue, as the 2kg exercises show both positive and negative slopes, this parameter could lead to

a possible discerning factor of mild muscle fatigue, as there are observable differences between 2kg and 4kg exercises.



**Fig. 4.** Box & Whisker plots of the SMR slopes through the 15 repetitions of the exercise with 2kg (yellow) and 4kg (blue). (A) Left leg, (B) Right leg.

The next step for the assessment of mild fatigue has been the analysis of muscle interactions. One of the undercurrent physiological behaviours in muscle fatigue is the slowing of conduction velocity of muscle fibers [13]. Cross-correlation technique is used in experiments to estimate the average muscle fiber conduction velocity during isometric exercises [13, 14]. Fig. 5. shows the box and whisker plots of the slopes of the MCC during the exercise. The results reveal mostly negative slopes for all exercises and pair of muscles. This increase of the delay between signals, leading to a decrease of the correlation may be caused by the reduction of the conduction velocity in the fibers due to muscle fatigue. This muscle impairment could also be associated to the first signs of muscle fatigue.



**Fig. 5.** Box & Whisker plots of the Cross-correlation slopes through the 15 repetitions of the exercise with 2kg (yellow) and 4kg (blue) for each muscle. (A) Left leg, (B) Right leg.

In this study, the final objective has been the determination of mild muscle fatigue during the realisation of a dynamic exercise in healthy subjects. Bandwidth, SMR, RMS and sample entropy parameters show signs of fatigue in line with the state of the art [2,3,8,10-12], due to a specific tendency during the exercise, especially in agonist muscles lifting 4kg. These parameters did not show a specific tendency with the 2kg exercise, which could mean two possibilities, or they are not sensible enough to mild muscle fatigue, or due to the variability in the population, some did not achieve that state. From the rest of the parameters analysed in this

study, MCC stands out because the reduction of the value during the exercise means a desynchronization of muscles leading to a mild fatigue. As these results are in accordance with the traditional parameters (temporal, spectral mainly), then, this is a possible feature for the determination of not only muscle fatigue, but other muscular injuries too, regarding muscular interaction and reduction of conduction velocity in muscle fibers.

The principal limitation of this study is the reduced database. For further investigation of these parameters, including a sensibility study, is imperative to increase it. Also, another handicap observed was the lack of a gold standard for the determination of the start of the fatigue state. This, correlated to the variability of the training status of the subjects and the exercise had low limited number of repetitions, then it was very difficult to determinate a state of muscular fatigue. Future studies will be done with higher number of participants and including an extenuating exercise ending with the surrender of the subject.

#### 4. Conclusion

The purpose of this study has been to analyse the behaviour of traditional and novel parameters during a non-extenuating task with the objective of determine new biomarkers for muscular fatigue. The bandwidth, SMR, RMS and sample entropy did not reveal muscle fatigue during the 2kg dynamic exercise. However, the dynamic exercise with 4kg has given the expected results in the tendency of the traditional parameters. Between all these parameters, the SMR exhibited great performance in the early detection of fatigue, then bandwidth and RMS and lastly, sample entropy. The relationship between the reduction of conduction velocity of fibers during muscle fatigue and the muscle impairment seen in MCC values, as well as its accordance with the results of bandwidth, SMR, RMS reveals a potential biomarker for the first signs of muscle fatigue. Further investigations and sensitivity studies with a larger recording database will be done to verify present study.

#### Acknowledgements

This study was supported by the Agència Valenciana de la Innovació project 'iSARC-GENETICS, Plataforma Inteligente para la Detección Temprana, Ayuda al Diagnóstico y Seguimiento de la Sarcopenia'. (INNEST/2021/365).

#### References

- [1] Karthick, P. A., Ghosh, D. M., and Ramakrishnan, S. (2018). Surface electromyography-based muscle fatigue detection using high-resolution time-frequency methods and machine learning algorithms. *Computer Methods and Programs in Biomedicine*, 154, 45–56. <https://doi.org/10.1016/j.cmpb.2017.10.024>
- [2] Cifrek, M., Medved, V., Tonković, S., and Ostojić, S. (2009). Surface EMG based muscle fatigue evaluation in Biomechanics. *Clinical Biomechanics*, 24(4), 327–340. <https://doi.org/10.1016/j.clinbiomech.2009.01.010>
- [3] Gerdle, B., Larsson, B., and Karlsson, S. (2000). Criterion validation of surface EMG variables as fatigue indicators using peak torque. *Journal of Electromyography and Kinesiology*, 10(4), 225–232. [https://doi.org/10.1016/s1050-6411\(00\)00011-0](https://doi.org/10.1016/s1050-6411(00)00011-0)
- [4] Farina, D., Gazzoni, M., and Merletti, R. (2003). Assessment of low back muscle fatigue by surface EMG signal analysis: Methodological aspects. *Journal of Electromyography and Kinesiology*, 13(4), 319–332. [https://doi.org/10.1016/s1050-6411\(03\)00040-3](https://doi.org/10.1016/s1050-6411(03)00040-3)
- [5] Iranzo, S., Piedrabuena, A., García Torres, F., Martínez de Juan, J. L., Prats Boluda, G., Sanchis, M., and Belda Lois, J. M. (2022). Assessment of a passive lumbar exoskeleton in material manual handling tasks under laboratory conditions. <https://doi.org/10.20944/preprints202204.0024.v1>
- [6] IEEE standard on transitions, pulses, and related waveforms. (2003). <https://doi.org/10.1109/ieeestd.2003.94394>
- [7] Richman, J. S., and Moorman, J. R. (2000). Physiological time-series analysis using approximate entropy and sample entropy. *American journal of physiology. Heart and circulatory physiology*, 278(6), H2039–H2049. <https://doi.org/10.1152/ajpheart.2000.278.6.H2039>
- [8] Dimitrov, G., Arabadzhiev, T., Mileva, K., Bowtell, J., Crichton, N., and Dimitrova, N. (2006). Muscle fatigue during dynamic contractions assessed by new spectral indices. *Medicine and Science in Sports and Exercise*, 38(11), 1971–1979. <https://doi.org/10.1249/01.mss.0000233794.31659.6d>
- [9] Buck, John R., Michael M. Daniel, and Andrew C. Singer. *Computer Explorations in Signals and Systems Using MATLAB®*. 2nd Edition. Upper Saddle River, NJ: Prentice Hall, 2002.
- [10] Karthick, P. A., and Ramakrishnan, S. (2016). Muscle fatigue analysis using surface EMG signals and time–frequency based medium-to-low Band Power Ratio. *Electronics Letters*, 52(3), 185–186. <https://doi.org/10.1049/el.2015.3460>
- [11] González-Izal, M., Malanda, A., Gorostiaga, E., and Izquierdo, M. (2012). Electromyographic models to assess muscle fatigue. *Journal of Electromyography and Kinesiology*, 22(4), 501–512. <https://doi.org/10.1016/j.jelekin.2012.02.019>
- [12] Kahl, L., and Hofmann, U. G. (2016). Comparison of algorithms to quantify muscle fatigue in upper limb muscles based on SEMG signals. *Medical Engineering and Physics*, 38(11), 1260–1269. <https://doi.org/10.1016/j.medengphy.2016.09.009>
- [13] Arendt-Nielsen, L., Mills, K. R., and Forster, A. (1989). Changes in muscle fiber conduction velocity, mean power frequency, and mean EMG voltage during prolonged submaximal contractions. *Muscle and Nerve*, 12(6), 493–497. <https://doi.org/10.1002/mus.880120610>
- [14] Marco, G., Alberto, B., and Taian, V. (2017). Surface EMG and muscle fatigue: Multi-channel approaches to the study of myoelectric manifestations of muscle fatigue. *Physiological Measurement*, 38(5). <https://doi.org/10.1088/1361-6579/aa60b9>

# Estimación de la severidad de apnea del sueño pediátrica a partir de la señal de flujo aéreo usando redes neuronales convolucionales

V. Barroso García<sup>1,2</sup>, F. Vaquerizo Villar<sup>1,2</sup>, G.C. Gutiérrez Tobal<sup>1,2</sup>, D. Álvarez González<sup>2,3</sup>, F. del Campo<sup>2,3</sup>, L. Kheirandish Gozal<sup>4</sup>, D. Gozal<sup>4</sup>, R. Hornero Sánchez<sup>1,2</sup>

<sup>1</sup> Grupo de Ingeniería Biomédica, Universidad de Valladolid, Valladolid, España, veronica.barroso@gib.tel.uva.es

<sup>2</sup> Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), España

<sup>3</sup> Servicio de Neumología, Hospital Universitario Río Hortega, Valladolid, España, fsas@telefonica.net

<sup>4</sup> Department of Child Health, The University of Missouri School of Medicine, Columbia, Missouri, EEUU, gozald@health.missouri.edu

## Resumen

*El objetivo de este estudio fue evaluar la utilidad de un modelo de deep learning basado en redes neuronales convolucionales (convolutional neural networks, CNN) para estimar automáticamente la severidad de la apnea obstructiva del sueño (AOS) infantil a partir de la señal de flujo aéreo monocal (FA). Para ello, se analizaron 3158 registros pediátricos de FA: 1639 usados para entrenar el modelo (entrenamiento), 759 para optimizar sus hiperparámetros (validación) y 760 para evaluar su capacidad diagnóstica (test). Todos los registros fueron divididos en segmentos de 5 minutos. Para cada uno de estos segmentos, el modelo CNN estimó su correspondiente índice de apnea-hipopnea (IAH). Después, se obtuvo el IAH de cada sujeto pediátrico como el promedio de los valores estimados por la CNN para los segmentos contenidos en la señal de FA. El modelo propuesto alcanzó una precisión global del 53.42% y un Cohen's kappa de 0.31 en test con cuatro clases. Respecto a la evaluación del rendimiento diagnóstico por grupos de severidad, la CNN obtuvo una precisión del 76.18% para 1 evento/h, 81.45% para 5 eventos/h y 90.66% para 10 eventos/h en test. Además, también destaca el alto ratio de verosimilitud positiva alcanzado para 10 eventos/h (22.16). Así, nuestra propuesta mejoró la capacidad diagnóstica obtenida por el FA mediante métodos convencionales de feature engineering. Estos resultados sugieren que la información extraída de la señal de FA mediante técnicas de deep learning puede ser una potente herramienta para ayudar a diagnosticar la AOS pediátrica, especialmente en los niños más severamente afectados.*

## 1. Introducción

La apnea obstructiva del sueño (AOS) es un trastorno crónico caracterizado por presentar recurrentes ceses respiratorios durante el sueño [1]. Dependiendo del tipo de evento, estos ceses pueden ser totales (eventos de apnea) o parciales (eventos de hipopnea) [1]. La presencia de estos eventos conducen a anomalías en el intercambio de gases, hipoxia, micro despertares y, por tanto, a un sueño fragmentado y no reparador [2]. Consecuentemente, los niños afectados por AOS pueden sufrir graves consecuencias para su salud y desarrollo, como alteraciones neuroconductuales, cardiovasculares, metabólicas y endocrinas [2], [3].

Esta enfermedad afecta hasta al 5.7% de la población infantil [3]. Sin embargo, a pesar de su alta prevalencia y de las graves consecuencias que conlleva, se estima que el 90% de los niños con AOS aún no cuentan con un diagnóstico objetivo [4]. Dicho diagnóstico se lleva a cabo mediante la polisomnografía nocturna (PSG), una prueba que consiste en registrar simultáneamente numerosas señales neurofisiológicas y cardiorrespiratorias mientras el niño duerme [5]. Posteriormente, los especialistas médicos inspeccionan visualmente todos registros y derivan el índice de apnea-hipopnea (IAH), el parámetro clínico usado para determinar la presencia y severidad de AOS [2]. Aunque la PSG es la técnica de referencia para diagnosticar la enfermedad, esta técnica es compleja, costosa, incómoda para los niños y de disponibilidad muy limitada [6]. Por ello, es necesario desarrollar pruebas diagnósticas más sencillas.

Una alternativa es el análisis automático de un conjunto reducido de registros cardiorrespiratorios, como la señal de flujo aéreo (FA). Esta señal es capaz de reflejar la actividad respiratoria del niño mientras duerme, incluidos los ceses respiratorios asociados con la AOS [1]. Además, puede ser fácilmente adquirida en la casa del paciente mediante un dispositivo de monitorización portátil con termistor incorporado. Así, diversos estudios ya han mostrado la utilidad del análisis automático del FA para diagnosticar la AOS pediátrica [7]–[11]. Sin embargo, estos estudios están basados en técnicas de *feature engineering*, por lo que presentan algunas limitaciones [12]. Estos enfoques requieren una caracterización exhaustiva de la señal, por lo que debe conocerse *a priori* la idoneidad de las características que van a ser extraídas. Además, proporcionan un nivel de abstracción bajo, lo que podría limitar su capacidad para detectar patrones complejos. Consecuentemente, podría haber información relevante del FA que pasase desapercibida en estos análisis. Por el contrario, las técnicas de *deep learning* permiten identificar automáticamente las características útiles y encontrar los patrones existentes a partir de la señal en crudo [12]. En este sentido, las redes neuronales convolucionales (*convolutional neural networks, CNN*)

son la técnica de *deep learning* más ampliamente utilizada dada su idoneidad para procesar matrices multidimensionales [13]. En el contexto de la AOS infantil, sólo dos estudios recientes de nuestro grupo de investigación han propuesto modelos CNN para estimar el IAH a partir de la señal de oximetría sola y combinada con FA [14], [15]. Sin embargo, y hasta donde sabemos, este es el primer estudio que evalúa individualmente el FA pediátrico mediante técnicas de *deep learning*.

De acuerdo con las consideraciones anteriores, el presente estudio parte de la hipótesis de que una arquitectura CNN podría identificar automáticamente la información relevante que está asociada con la presencia de eventos apneicos en FA. Por tanto, nuestro objetivo es evaluar la utilidad de un modelo CNN para estimar el IAH y determinar la severidad de AOS pediátrica a partir del FA.

## 2. Sujetos y señales

Nuestro estudio ha empleado tres bases de datos de FA pediátrico: (i) the *Childhood Adenotonsillectomy Trial* (CHAT), una base de datos pública compuesta por 1639 registros y obtenida de la web <https://sleepdata.org>, (ii) una base de datos privada procedente del *Comer Children's Hospital* de la Universidad de Chicago (UofC) que incluía 974 registros, y (iii) una base de datos privada procedente del *LeBonheur Children's Hospital* de la Universidad de Tennessee (UofT), formada por 545 registros pediátricos. El estudio fue aprobado por el Comité Ético de los centros participantes y los tutores legales de los niños dieron su consentimiento informado.

Todos los niños involucrados fueron diagnosticados por especialistas médicos siguiendo las reglas de la *American Academy of Sleep Medicine* [1]. Para determinar la severidad de cada sujeto se usaron los umbrales de IAH 1, 5 y 10 eventos/hora (e/h) [2]. La Tabla 1 muestra los datos demográficos y clínicos de los sujetos bajo estudio. Cabe señalar que sólo los registros de la base de datos CHAT incluían marcajes de eventos apneicos. Consecuentemente, ésta fue usada para entrenar el modelo de *deep learning*. Respecto a UofC y UofT, éstas fueron distribuidas aleatoriamente en dos grupos: validación (50%) y test (50%), para optimizar los hiperparámetros y evaluar la capacidad diagnóstica del modelo propuesto, respectivamente. Se encontraron diferencias estadísticamente significativas ( $p$ -valor  $< 0.01$  al aplicar el test de Mann-Whitney  $U$  y la corrección de Bonferroni) en edad, sexo, índice de masa corporal (IMC) e IAH entre entrenamiento-validación y entrenamiento-test, pero no entre validación-test.

Las señales de FA fueron obtenidas durante la PSG. Estas se registraron utilizando un termistor oronasal a frecuencias de muestreo que oscilaban entre 20 y 512 Hz.

## 3. Métodos

### 3.1. Preprocesado y segmentación

Para que todas las señales tuvieran la misma frecuencia de muestro, estas fueron remuestreadas a 10 Hz, frecuencia que ha demostrado ser adecuada para caracterizar la señal de FA y reducir el coste computacional [14]. Después, se

	Entrenamiento (CHAT)	Validación (UofC y UofT)	Test (UofC y UofT)
<b>Sujetos (n)</b>	1639 (100.0%)	759 (24.0%)	760 (24.1%)
<b>Edad (años)</b>	7.0 [2.0]	6.0 [6.9]	6.00 [6.2]
<b>Varones (n)</b>	777 (47.4%)	437 (57.6%)	455 (59.9%)
<b>IMC (kg/m<sup>2</sup>)</b>	17.3 [6.2]	18.4 [8.0]	18.4 [7.9]
<b>IAH (e/h)</b>	2.5 [4.8]	3.1 [6.7]	3.3 [7.3]
<b>No AOS (n)</b>	355 (21.7%)	173 (22.8%)	174 (22.9%)
<b>AOS Leve (n)</b>	814 (49.6%)	303 (39.9%)	302 (39.8%)
<b>AOS Moderado (n)</b>	252 (15.4%)	127 (16.7%)	128 (16.8%)
<b>AOS Severo (n)</b>	218 (13.3%)	156 (20.6%)	156 (20.5%)

IMC: índice de masa corporal; IAH: índice apnea-hipopnea. Datos presentados como número de sujetos (%) o mediana [rango intercuartil].

Tabla 1. Datos demográficos y clínicos de los sujetos.

sometieron a un proceso de estandarización para normalizar los niveles de referencia de FA entre los diferentes sujetos pediátricos. Finalmente, los registros fueron divididos en segmentos de 5 minutos como en estudios previos [14]. En la etapa de entrenamiento y validación, esta segmentación se realizó con un solapamiento del 50% para incrementar el número de segmentos disponibles.

### 3.2. Arquitectura CNN

Como puede observarse en la Figura 1, la arquitectura CNN que proponemos está compuesta de tres módulos. El primero de ellos corresponde al módulo de entrada, que recibe los segmentos de 5 minutos de FA.

El siguiente es el módulo CNN, el cual está constituido por  $\lambda$  bloques convolucionales. Cada uno de estos bloques está formado por cinco capas consecutivas: convolucional, *batch normalization*, activación, *pooling* y *dropout* [12], [13]. La capa convolucional es la encargada de extraer los mapas de características mediante filtros de convolución [12], [13]. En nuestro estudio, cada capa convolucional tenía 64 filtros de tamaño =  $5 \times 1$ , paso = 1 y *padding* = 0 [15]. Después, se aplica un *batch normalization* para normalizar los mapas de características [12]. En la siguiente capa, se usa una función no lineal para determinar cuáles de esos mapas de características se activan [12], [13]. Así, la función de activación empleada en este estudio fue la *Rectified Linear Unit* (ReLU) [14], [15]. Respecto al *pooling*, éste permite reducir la dimensionalidad fusionando patrones similares [12]. En concreto, nosotros utilizamos un *max pooling* con factor = 2, el más utilizado en las CNN [14], [15]. Cada bloque convolucional finaliza con una capa de *dropout*. Ésta elimina aleatoriamente activaciones con probabilidad  $p_{drop}$  para reducir el sobreajuste durante el entrenamiento [12], [13].

Por último, el tercer módulo es el de salida. Este se compone de una capa de *flattening*, cuyo objetivo es transformar los mapas de característica en vectores unidimensionales. También incluye una capa de activación lineal, la cual es utilizada para obtener el IAH estimado de cada segmento de FA de entrada.

### 3.3. Entrenamiento y optimización de la CNN

El proceso de entrenamiento de la CNN comienza con la inicialización de los pesos y sesgos de la red. Para ello, utilizamos el inicializador He-normal [15]. Después, los

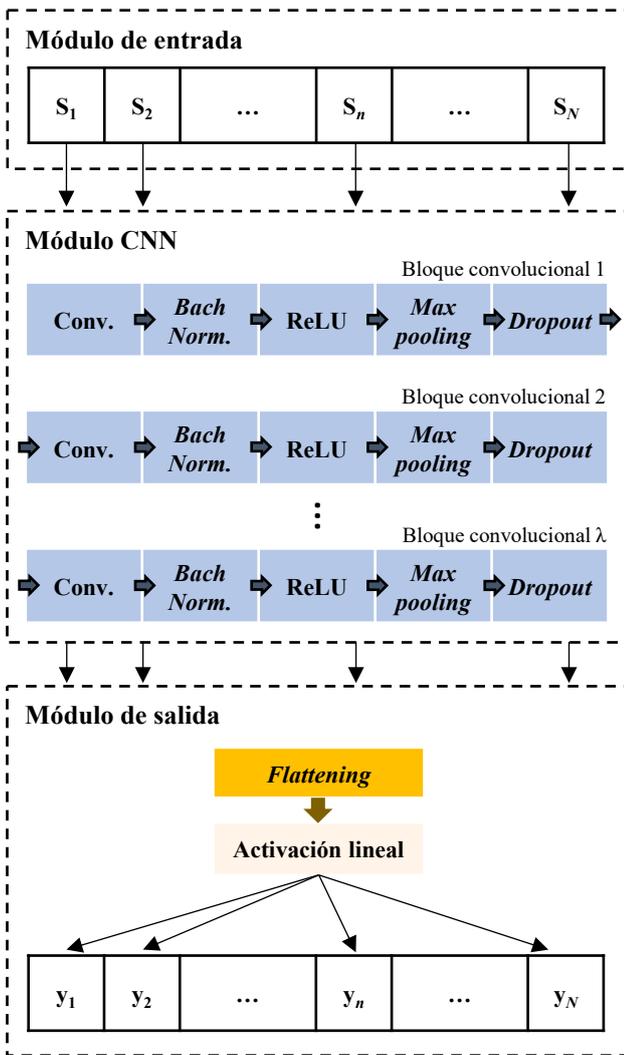


Figura 1. Arquitectura CNN propuesta

pesos y sesgos son actualizados en cada iteración mediante la estimación del momento adaptativo (algoritmo de Adam) [12]. En este estudio, el algoritmo de Adam utilizó una tasa de aprendizaje inicial de 0.001 y la función de error *Huber loss* [15]. Esta función contiene el hiperparámetro delta ( $\delta$ ), el cual debe ser ajustado para controlar el efecto de los valores atípicos e incrementar la robustez [16]. Así, la función *Huber loss* fue calculada en el conjunto de validación para monitorear la convergencia de la CNN durante su entrenamiento. En este sentido, la tasa de aprendizaje se redujo en un factor de 2 durante el entrenamiento cuando la pérdida en el conjunto de validación no mejoró en las últimas 10 épocas [14], [15]. Además, se aplicó un criterio de parada temprana de 30 épocas sin mejoras para evitar el sobreajuste.

### 3.4. Estimación del IAH

Para obtener el IAH de cada sujeto pediátrico, se han utilizado los valores reportados por el módulo de salida de la CNN. En primer lugar, se calcula el promedio de los valores de IAH estimados para los  $N$  segmentos de 5 minutos contenidos en la señal de FA correspondiente [15]:

$$\bar{y} = \frac{\sum_{n=1}^N y_n}{N}, \quad (1)$$

donde  $y$  es la salida de la red obtenida para el segmento  $n$ .

Después, el IAH de cada sujeto es calculado mediante un modelo regresión lineal para corregir la posible subestimación causada por la CNN [15]:

$$\text{IAH} = (\beta \cdot \bar{y}) + \varepsilon, \quad (2)$$

siendo  $\beta$  la pendiente y  $\varepsilon$  el sesgo del modelo. Dichos coeficientes fueron ajustados en el conjunto de validación.

### 3.5. Análisis estadístico

El rendimiento diagnóstico del modelo CNN se evaluó globalmente mediante la matriz de confusión (4 clases), la precisión global y el Cohen's kappa ( $\kappa$ ). También se evaluó por grados de severidad mediante las siguientes métricas: sensibilidad (S), especificidad (E), precisión (P), valor predictivo positivo y negativo (VPP y VPN) y razón de verosimilitud positiva y negativa (LR+ y LR-).

## 4. Resultados

### 4.1. Optimización de la CNN

Los hiperparámetros a optimizar fueron el número de bloques convolucionales  $\lambda$ , la probabilidad  $p_{drop}$  y el  $\delta$  de la función *Huber loss*. El valor de estos hiperparámetros fue variando para configurar distintos modelos CNN que fueron entrenados en el grupo de entrenamiento. Para cada uno de estos modelos, el índice kappa fue obtenido en el grupo de validación. Así, aquella configuración que obtuvo el mayor kappa fue establecida como óptima. En este estudio, los valores óptimos fueron  $\lambda = 7$ ,  $p_{drop} = 0.075$  y  $\delta = 5.5$ , con un kappa de 0.29 en el grupo de validación.

### 4.2. Rendimiento diagnóstico de la CNN

La Tabla 2 muestra la matriz de confusión obtenida por el modelo CNN en el grupo de test. Éste alcanzó una precisión global del 53.42% (406/760 sujetos correctamente clasificados) y un kappa de 0.31 en test con cuatro clases. A partir de la matriz de confusión, se obtuvo el rendimiento diagnóstico alcanzado por el modelo CNN para 1, 5 y 10 e/h en el grupo de test. Los resultados obtenidos son mostrados en la Tabla 3. Como puede observarse, el modelo CNN alcanzó una alta precisión diagnóstica en 5 y 10 e/h (81.45% y 90.66%, respectivamente), así como un alto LR+ en 10 e/h (22.16).

		CNN			
		No AOS	Leve	Moderado	Severo
PSC	No AOS	11	155	8	0
	Leve	13	246	38	5
	Moderado	3	66	46	13
	Severo	2	19	32	103

Tabla 2. Matriz de confusión obtenida por el modelo CNN en el grupo de test.

IAH	S (%)	E (%)	P (%)	VPP (%)	VPN (%)	LR+	LR-
1 e/h	96.93	6.32	76.18	77.70	37.93	1.03	0.49
5 e/h	68.31	89.29	81.45	79.18	82.52	6.38	0.35
10 e/h	66.03	97.02	90.66	85.12	91.71	22.16	0.35

Tabla 3. Rendimiento diagnóstico obtenido para 1, 5 y 10 e/h por el modelo CNN en el grupo de test.

## 5. Discusión y conclusiones

En este trabajo se ha evaluado la utilidad de analizar la señal de FA mediante un modelo de *deep learning* basado en CNN para simplificar el diagnóstico de la AOS infantil. Así, el modelo propuesto logró estimar el IAH de los sujetos pediátricos con una alta precisión diagnóstica (76.18%, 81.45% y 90.66% para 1, 5 y 10 e/h, respectivamente), especialmente en los casos con AOS moderada y severa. Además, destaca el alto LR+ alcanzado por la CNN en 10 e/h (22.16). Ya que un valor superior a 10 en esta métrica es un robusto indicador para establecer la presencia de una enfermedad [17], la metodología propuesta podría ser una útil herramienta para determinar automáticamente la presencia de AOS severa en niños.

En los últimos años, diversos estudios de nuestro grupo de investigación se han centrado en el análisis automático del FA como ayuda al diagnóstico de la AOS pediátrica [7]–[11], [14]. La mayoría de estos estudios se basaron en la caracterización de la señal de FA mediante técnicas convencionales de *feature engineering*, con las cuales alcanzaron precisiones que oscilan entre el 73.61% – 81.10% para 1 e/h, 57.14% – 62.82% para 5 e/h y 70.47% – 80.60% para 10 e/h [8]–[11]. En el presente trabajo, alcanzamos precisiones dentro de este rango para 1e/h, así como notablemente superiores para 5 y 10 e/h. Otro estudio analizó conjuntamente el FA y la señal de oximetría mediante CNN [14]. Éste alcanzó altas precisiones en CHAT (84.64%, 93.46% y 94.44% para 1, 5 y 10 e/h, respectivamente), que se redujeron notablemente cuando la propuesta se evaluó en UofC (84.10%, 84.10% y 90.26% para 1, 5 y 10 e/h, respectivamente). Analizando un único canal, nuestro modelo obtuvo una precisión ligeramente inferior en 1 y 5 e/h, y superior en 10 e/h. Además, también lo superó en S en 1 e/h y en E, VPP y LR+ en 5 y 10 e/h.

Este estudio presenta algunas limitaciones que deben ser consideradas. En primer lugar, el modelo CNN fue entrenado únicamente con la base de datos CHAT, ya que los registros de UofC y UofT no contenían anotaciones sobre el comienzo y finalización de los eventos apnéicos. Además, se observaron diferencias entre estas bases de datos, particularmente entre CHAT y UofC y entre CHAT y UofT, lo que podría haber afectado a la capacidad de generalización de nuestro enfoque. En este sentido, el modelo propuesto podría ser validado en otras bases de datos, e incluso con registros domiciliarios de FA, lo que aumentaría la solidez de nuestros resultados. También se podría estimar el IAH mediante la aplicación y/o combinación de otros métodos de *deep learning*, como las redes *long short-term memory* (LSTM), lo que constituye una interesante línea de investigación futura. Finalmente, sería interesante utilizar técnicas de *explainable artificial intelligence* (XAI). Esto nos permitirían identificar y comprender qué características del FA extrae el modelo CNN para derivar su predicción.

En conclusión, el modelo CNN mostró un alto rendimiento diagnóstico estimando el IAH pediátrico. La mejora obtenida respecto a las técnicas convencionales pone de manifiesto la utilidad de los métodos de *deep learning*, y en particular de las arquitecturas CNN, para incrementar la

capacidad diagnóstica del FA. Estos resultados sugieren que la información extraída de la señal de FA mediante técnicas de *deep learning* puede ser una potente herramienta para ayudar a diagnosticar la AOS infantil, especialmente en los niños más severamente afectados.

## Agradecimientos

Este estudio ha sido financiado por el Ministerio de Ciencia e Innovación - Agencia Estatal de Investigación y ERDF (PID2020-115468RB-I00 y PDC2021-120775-I00), por el CIBER-BBN (CB19/01/00012), por la Sociedad Española de Neumología y Cirugía Torácica (SEPAR 649/2018) y por la Sociedad Española de Sueño (Beca de Investigación SES 2019). D. Álvarez es beneficiario de una ayuda Ramón y Cajal (RYC2019-028566-I) del Ministerio de Ciencia e Innovación - Agencia Estatal de Investigación, cofinanciada por el Fondo Social Europeo.

## Referencias

- [1] R. B. Berry *et al.*, “Rules for Scoring Respiratory Events in Sleep: Update of the 2007 AASM Manual for the Scoring of Sleep and Associated Events,” *J. Clin. Sleep Med.*, vol. 08, no. 05, pp. 597–619, 2012.
- [2] M. Luz Alonso-Álvarez *et al.*, “Consensus document on sleep apnea-hypopnea syndrome in children,” *Arch. Bronconeumol.*, vol. 47, no. SUPPL. 5, pp. 2–18, 2011.
- [3] C. L. Marcus *et al.*, “Diagnosis and management of childhood obstructive sleep apnea syndrome,” *Pediatrics*, vol. 130, no. 3, 2012.
- [4] L. Kheirandish-Gozal, “What Is ‘Abnormal’ in Pediatric Sleep?,” *Respir. Care*, vol. 55, no. 10, pp. 1366–1376, 2010.
- [5] C. Jon, “Polysomnography in Children,” in *Pediatric Otolaryngology for the Clinician*, Totowa, NJ: Humana Press, 2009, pp. 35–47.
- [6] P. J. Ryan *et al.*, “Validation of British Thoracic Society guidelines for the diagnosis of the sleep apnoea/hypopnoea syndrome: can polysomnography be avoided?,” *Thorax*, vol. 50, no. 9, pp. 972–975, 1995.
- [7] G. C. Gutiérrez-Tobal *et al.*, “Diagnosis of pediatric obstructive sleep apnea: Preliminary findings using automatic analysis of airflow and oximetry recordings obtained at patients’ home,” *Biomed. Signal Process. Control*, vol. 18, pp. 401–407, 2015.
- [8] V. Barroso-García *et al.*, “Usefulness of recurrence plots from airflow recordings to aid in paediatric sleep apnoea diagnosis,” *Comput. Methods Programs Biomed.*, vol. 183, pp. 105083, 2020.
- [9] V. Barroso-García *et al.*, “Bispectral analysis of overnight airflow to improve the pediatric sleep apnea diagnosis,” *Comput. Biol. Med.*, vol. 129, pp. 104167, 2021.
- [10] V. Barroso-García *et al.*, “Wavelet Analysis of Overnight Airflow to Detect Obstructive Sleep Apnea in Children,” *Sensors*, vol. 21, no. 4, pp. 1491, 2021.
- [11] J. Jiménez-García *et al.*, “Assessment of Airflow and Oximetry Signals to Detect Pediatric Sleep Apnea-Hypopnea Syndrome Using AdaBoost,” *Entropy*, vol. 22, no. 6, pp. 670, 2020.
- [12] I. Goodfellow *et al.*, *Deep learning*. MIT press, 2016.
- [13] Y. Lecun *et al.*, “Deep learning,” *Nat.*, vol. 521, no. 7553, pp. 436–444, 2015.
- [14] J. Jiménez-García *et al.*, “A 2D convolutional neural network to detect sleep apnea in children using airflow and oximetry,” *Comput. Biol. Med.*, vol. 147, pp. 105784, 2022.
- [15] F. Vaquerizo-Villar *et al.*, “A Convolutional Neural Network Architecture to Enhance Oximetry Ability to Diagnose Pediatric Obstructive Sleep Apnea,” *IEEE J. Biomed. Heal. Informatics*, vol. 25, no. 8, pp. 2906–2916, 2021.
- [16] P. J. Huber, “Robust Estimation of a Location Parameter,” *Ann. Math. Stat.*, vol. 35, no. 1, pp. 73–101, 1964.
- [17] J. J. Deeks and D. G. Altman, “Diagnostic tests 4: likelihood ratios,” *BMJ*, vol. 329, pp. 168–169, 2004.

# Medición automática del “jitter” en registros electromiográficos patológicos

A. Malanda Trigueros<sup>1</sup>, D. Stashuk<sup>2</sup>, O. Garnés Camarena<sup>3</sup>, C. Valle Morgado<sup>1</sup>, J. Navallas Irujo<sup>1</sup>, J. Rodríguez Falces<sup>1</sup>

<sup>1</sup> Dpto. Ingeniería Eléctrica y Electrónica, Universidad Pública de Navarra, Pamplona, España. [malanda@unavarra.es](mailto:malanda@unavarra.es)

<sup>2</sup> System Design Engineering Department, University of Waterloo, Waterloo (ON), Canada.

<sup>3</sup> Servicio de Rehabilitación y Neurofisiología. Hospital Universitario de la Fundación Jiménez Díaz, Madrid, España.

## Resumen

La dispersión en los instantes de disparo de los potenciales generados en las fibras de una unidad motora (“jitter”) se ve aumentada en procesos patológicos o degenerativos que afectan a la unión neuromuscular o a las motoneuronas. Tradicionalmente el “jitter” se ha medido con electrodos de fibra simple, que son caros y al ser reutilizables presentan potencial riesgo de contagio, por lo que gradualmente se va imponiendo la utilización de electrodos de fibra concéntrica desechables. El “jitter” es más difícil de medir en estos electrodos, pues al ser de mayor tamaño, se recogen los potenciales eléctricos de un número mayor de fibras y éstos se solapan más entre sí, dificultando su análisis. En 2019 los autores presentaron un método automático para medir el “jitter”, probando su capacidad y prestaciones con señales simuladas. En el presente trabajo se evalúa el funcionamiento del método automático utilizando señales reales de músculos con posibles patologías neurógenas y de la unión neuromuscular. Las mediciones obtenidas fueron comparadas con las halladas según procedimientos manuales estándar. La media y desviación estándar de las diferencias entre las medidas manuales y automáticas fueron 0.55 y 5.28  $\mu\text{s}$ , respectivamente. Aunque se precisan todavía análisis más exhaustivos, las estrechas diferencias entre las dos medidas sostienen la utilidad clínica del método automático.

## 1. Introducción

La unidad motora (UM), compuesta por una motoneurona y las fibras musculares a las que ésta inerva, es la unidad estructural y funcional de la contracción voluntaria. Al ser activada, un tren de impulsos eléctricos recorre su motoneurona, causando el disparo de trenes de potenciales en cada una de sus fibras. Un electrodo colocado en las proximidades de una UM registrará la suma de los potenciales de las distintas fibras, lo que se llama el potencial de unidad motora (PUM). Los potenciales de las distintas fibras de la UM se generan de forma sincrónica, aunque con una pequeña dispersión temporal en el inicio de los mismos. A esta dispersión se le conoce como “jitter neurofisiológico” o simplemente “jitter”. Distintas condiciones fisiopatológicas que afectan al funcionamiento de la unión neuromuscular pueden ocasionar un aumento del “jitter”. Es el caso de las enfermedades de la unión neuromuscular, como la *miastenia gravis* [1] o el *síndrome de Lambert-Eaton*, de los procesos de denervación-reinervación en enfermedades neurógenas como la *esclerosis lateral amiotrófica* [2] o de la vejez [3].

Para medir el “jitter” se han utilizado tradicionalmente electrodos de fibra simple, que tienen una parte activa muy

pequeña (de unos 0.005 mm<sup>2</sup>) [4]. La técnica consiste en manipular la aguja del electrodo buscando PUMs en los que aparezcan suficientemente separados los potenciales de únicamente dos fibras (potenciales de fibra simple o PFS) (Fig. 1). Se medirá el intervalo entre ambos potenciales en las sucesivas descargas de la fibra (Fig. 1.B), y el “jitter” se computará como la desviación estándar de los intervalos medidos [1]. Otra medida alternativa es la media de las diferencias consecutivas (“mean of consecutive differences” o MCD) de los intervalos entre potenciales. Esta medida es la preferida en la práctica clínica por su capacidad de eliminar sesgos o tendencias en los intervalos del tren de potenciales, que pueden producirse por el movimiento involuntario de la aguja en el registro EMG o por otros factores difíciles de controlar. La medición del “jitter” con electrodos de fibra simple es una técnica difícil y laboriosa. Además, debido a su elevado precio los electrodos son reutilizables, lo cual conlleva un potencial riesgo de infección. El electrodo de aguja concéntrico es el que se utiliza en la mayor parte de las exploraciones EMG clínicas actuales. Este electrodo tiene una superficie activa más grande (del orden de 0.070 mm<sup>2</sup>) [4] y resulta más sencillo de usar, más barato y además es desechable. Por todo ello cada vez es más generalizado su uso en la práctica clínica para realizar mediciones de “jitter” [3][4].

La señal de EMG registrada en estas exploraciones es filtrada con un filtro paso alto [4] o un filtro de aceleración [5]. Estos filtros eliminan las bajas frecuencias de la señal, debidas en buena parte a los potenciales de las fibras alejadas del electrodo, y estrechan los picos de los potenciales del PUM, facilitando su separación. También con este tipo de electrodos, la tarea de estimar el “jitter” resulta complicada y requiere tiempo y habilidad del

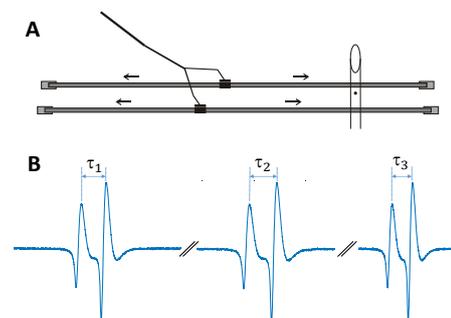


Fig. 1. Registro (A) y dispersión temporal (B) de un PUM de dos fibras

especialista para obtener señales que contengan al menos dos PFS reconocibles y disjuntos.

En 2017 Abdelmaseeh et al., publicaron una técnica automática para la determinación del “jitter” a partir de registros de aguja concéntrica [6]. Esta técnica es computacionalmente compleja e incluye decisiones basadas en ejemplos tomados de señales simuladas; por todo ello resulta difícil de reproducir y, a día de hoy, no se ha empleado para procesar señales patológicas reales.

Más recientemente, se propuso una nueva técnica llamada “near fiber segment jitter” [7] para evaluar la inestabilidad de la UM proporcionando una medida promedio ponderada de la fluctuación temporal asociada a las fibras cercanas que contribuye significativamente al PUM analizado. Esta técnica se probó con datos simulados y también con datos de EMG de tres sujetos (control, neurogénico y miopático), obteniendo una correlación consistente entre las medidas de “jitter” y los cambios fisiológicos incluidos en la simulación o presumiblemente existentes en los músculos de los sujetos. Sin embargo, no se presentó ninguna evaluación directa de la precisión de estas mediciones.

En el Congreso CASEIB de 2019 los autores presentaron un nuevo método automático para estimar el “jitter” a partir de trenes de PUMs tomados con ambos tipos de electrodos [8]. Este algoritmo fue probado con un conjunto de señales sintéticas basadas en modelos de generación de PFS, obteniéndose medidas de “jitter” de gran precisión. El objetivo del estudio que aquí se presenta ha sido el de evaluar el funcionamiento de aquel método automático utilizando señales patológicas reales. A continuación, se muestra el método empleado para la evaluación, los resultados obtenidos y las conclusiones del estudio.

## 2. Material y métodos

### 2.1. Medición automática del “jitter”

En la Figura 2.A aparecen superpuestos varios potenciales de un tren de PUMs compuestos por dos PFS. Para calcular el “jitter” se alinean todos los potenciales por el primer pico positivo y luego se mide la desviación estándar de los instantes temporales en los que aparece el segundo pico positivo. En general, dado un punto de alineamiento se puede medir el “jitter” en cualquier instante temporal de la señal, aunque no se corresponda con un pico (Fig. 2.B). Esto se hará hallando el desplazamiento óptimo (mediante el máximo de correlación cruzada o mínimo de la distancia euclídea) entre segmentos de señal centrados en este instante temporal correspondientes a distintos potenciales del tren y computando la desviación estándar de esos

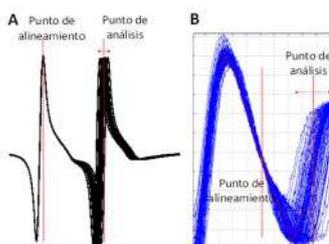


Fig. 2. Medición del “jitter”

desplazamientos.

Dependiendo de la diferencia temporal entre los dos PFS, se obtienen mediciones de “jitter” que podrán diferir apreciablemente del “jitter” real. Esto es debido a la influencia que tiene cada potencial individual sobre la forma de onda del potencial vecino. Sólo cuando dos potenciales están separados al menos 0.5 ms, el “jitter” medido se ajusta estrechamente al “jitter” real. Cuando la separación es inferior, habrá solapamiento entre los potenciales y el “jitter” medido no será fiable [9]. Se precisa, por tanto, encontrar aquellas partes del PUM cuyo origen proceda mayormente de un solo PFS (separado al menos 0.5 ms de los otros PFS que componen el PUM).

El método presentado en [8], se basa en el concepto de “jitter de vecindad”, que es el “jitter” obtenido en las proximidades de un instante temporal dado, después de que el conjunto de potenciales del tren de PUMs haya sido alineado en ese instante temporal. Un “jitter de vecindad” bajo indica que en el instante elegido el potencial procede de un solo PFS y no de la suma de varios PFS; por tanto, este instante es válido o bien para alinear los PUMs del tren, o bien para medir su “jitter” una vez alineados. El “jitter de vecindad” se analiza en todas las muestras temporales del tren de PUMs obteniéndose una curva de “jitter de vecindad” (Fig. 3). Los intervalos temporales situados bajo un umbral determinado serán considerados válidos para alinear o para medir “jitter”. Cada intervalo se caracteriza por su duración y el jitter de vecindad promedio. La fiabilidad para medir adecuadamente el “jitter” puede establecerse como función de estos dos parámetros.

En un par de intervalos válidos, se puede tomar uno para el alineamiento y el otro para el análisis del “jitter”. El valor del “jitter” se calcula como el promedio ponderado entre los puntos que constituyen estos intervalos. Los pesos serán inversamente proporcionales a los valores de “jitter de vecindad” de las muestras consideradas. Si hay  $N$  intervalos válidos, entonces habrá  $\binom{N}{2} = \frac{N \cdot (N-1)}{2}$  parejas distintas de intervalos, pero sólo  $N-1$  de ellos serán independientes. De todos los pares posibles, se eligen los

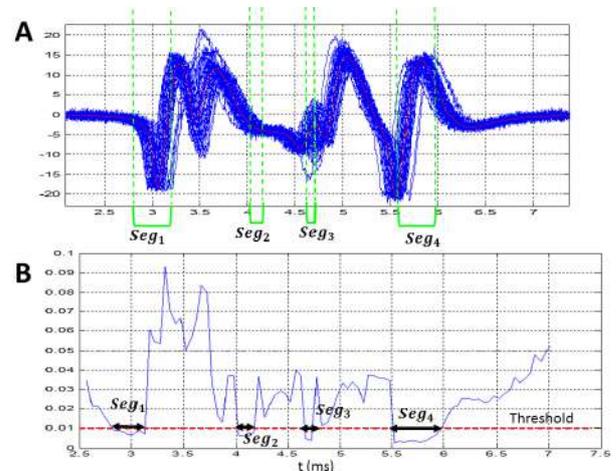


Fig. 3. Tren de PUMs (A) y su curva de “jitter de vecindad” (B). Se muestran los intervalos válidos (Seg1-Seg4).

N-1 pares independientes que tengan mayor fiabilidad promedio, de los cuales saldrán N-1 medidas de “jitter”.

## 2.2. Pacientes y registros

Se tomaron 132 señales de EMG de 20 segundos de duración de los músculos de ocho pacientes con síntomas de enfermedades neurógenas o de la unión neuromuscular registradas durante los exámenes clínicos de rutina en el Departamento de Neurofisiología Clínica del Hospital Universitario Fundación Jiménez Díaz, Madrid. Se utilizó un sistema de EMG Keypoint (v.3.22), un filtro paso banda de 0.5-10 kHz ó de 1-10 kHz y electrodos de aguja concéntricos faciales (área de captación de 0,019 mm<sup>2</sup>). El estudio se llevó a cabo de acuerdo con la Declaración de Helsinki para Experimentos con Humanos y fue aprobado por el Comité de Ética del mencionado Hospital.

## 2.3. Pruebas de validación

De cada señal de EMG el sistema Keypoint extrae un tren de PUMs utilizando un determinado umbral de amplitud. Aunque este sistema registra digitalmente las señales de EMG crudas, los MUP extraídos no se pueden exportar digitalmente. Por lo tanto, para cada señal EMG sin procesar, se realizó una captura de pantalla de los MUPs extraídos de Keypoint (Fig. 4.A). Por otro lado, la señal EMG cruda se descompuso usando el software DQEMG [5], extrayéndose varios trenes de PUMs (Fig. 4.B-D). De estos trenes, se seleccionó para el análisis de “jitter” el que más coincidía con el tren de MUPs extraído de Keypoint, según la comparación visual con el tren de potenciales observado en la captura de pantalla correspondiente. Es el caso del PUM 1 (Fig. 4.B), en el ejemplo de la Fig. 4. Tras descartar algunas señales en las que no se pudo encontrar una coincidencia adecuada, al final de este proceso se obtuvieron 115 trenes de PUMs.

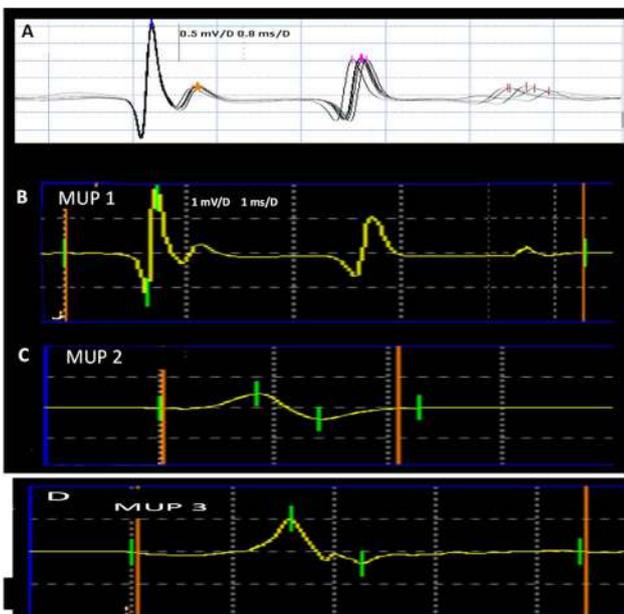


Fig. 4. Captura de pantalla de los PUMs extraídos por el equipo Keypoint (A). Tres plantillas de PUMs extraídas por DQEMG de la misma señal EMG (B-D). La primera de estas plantillas (B) representa el tren de PUMs extraído de DQEMG que se corresponde al tren extraídos por el equipo y mostrado en A.

Con ellos se empleó una interfaz gráfica construida ad-hoc [10] para descartar manualmente potenciales corrompidos por el ruido o inapropiados, obteniéndose 96 trenes de PUMs ya listos para ser analizados.

Se utilizaron programas propios desarrollados en el entorno Matlab para obtener el “jitter” con el método automático, y la mencionada interfaz gráfica para realizar medidas manuales siguiendo un procedimiento análogo al llevado a cabo en la práctica clínica para la medición del “jitter” a partir de los picos del potencial [11]. Se obtuvieron 102 medidas de “jitter” del método automático y 84 de la evaluación manual (usando la interfaz). Entre los 84 pares de picos detectados en el método manual, 82 resultaron cercanos en el tiempo a pares de intervalos válidos determinados por el método automático. Utilizando exclusivamente estos pares de picos coincidentes se realizó un análisis comparativo de los valores MCD obtenidos por los métodos manual y automático.

## 3. Resultados

En la Fig. 5.A se muestran diagramas de caja de las medidas de MCD obtenidos por los métodos manual y automático para las medidas coincidentes y de la diferencia de valores MCD en estas medidas. En la Fig. 5.B se muestra el histograma de estas diferencias. Las medianas y desviaciones estándar medidas en  $\mu\text{s}$  para las medidas de “jitter” manual y automático, y para la diferencia entre las dos medidas fueron, respectivamente: (29.03, 15.17), (26.95, 14.40) y (0.555, 5.28). Los percentiles 5, 25, 75 y 95 de las diferencias fueron: -2.62, -0.35, 2.53 y 7.56  $\mu\text{s}$ , respectivamente. Por último, se contabilizaron 6 “outliers”.

En la Fig. 6 se muestran dos ejemplos de trenes de PUMs diferentes tomados del conjunto estudiado, junto con sus curvas de “jitter de vecindad”. Se dan las medidas de MCD obtenidas por los métodos manual y automático. En el primer ejemplo (Fig. 6.A, C, E) los dos picos detectados manualmente están cerca de los dos intervalos obtenidos por el método automático; por lo tanto, es un caso coincidente. Los valores de MCD obtenidos por los dos métodos (37,6 y 35,5  $\mu\text{s}$ , respectivamente) son muy cercanos. El segundo ejemplo (Fig. 6.B, D, F) es un tren de PUMs polifásicos. Por el método manual se detectaron tres picos y se extrajeron dos mediciones de MCD. Asimismo, el método automático determinó tres intervalos válidos y

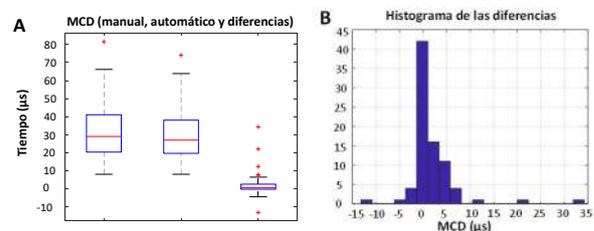


Fig. 5. Diagramas de caja (“boxplots”) de valores de MCD obtenidos por los métodos manual y automático para las medidas coincidentes (ver texto) y de la diferencia de valores MCD en estas medidas. Estos valores atípicos se trazan individualmente (A). Histograma de la diferencia de valores MCD en las medidas coincidentes (B).

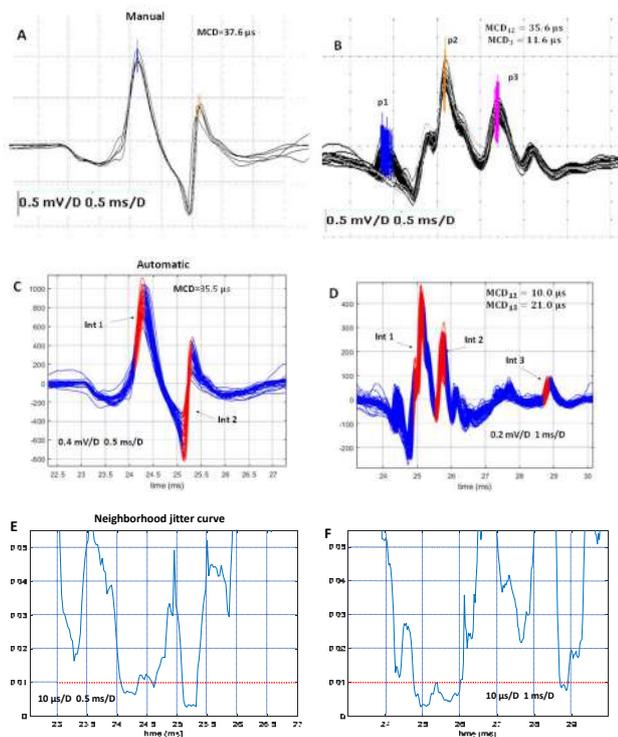


Fig. 6. Ejemplos de trenes de PUMs analizados por el método manual, con los picos negativos detectados marcados y los valores MCD obtenidos (A, B). Los mismos trenes de PUMs analizados por el método automático, con intervalos válidos resaltados y valores MCD (C, D). Curvas correspondientes de “jitter” de vecindad (E, F).

extrajo dos medidas MCD, pero sólo una de ellas coincidió con una de las medidas manuales (MCD=11.6  $\mu$ s tomado de p1 y p3 en el método manual, y MCD=10  $\mu$ s tomado de los intervalos 1 y 2 en el método automático). Las otras dos mediciones de MCD no se pueden comparar ya que no corresponden a pares de picos e intervalos cercanos en el tiempo.

#### 4. Conclusiones y líneas futuras

- 1- Se ha probado un algoritmo automático de estimación de “jitter” para trenes de PUMs registrados con electrodos concéntricos con un banco de señales de EMG de sujetos con patologías neurógenas y de la unión neuromuscular
- 2- Los resultados de las medidas automáticas se ajustaron estrechamente a las medidas manuales (mediana y desviación estándar de las diferencias de 0.555 y 5.28  $\mu$ s, respectivamente).
- 3- El algoritmo automático produce mayor número de estimaciones de “jitter” válidas que el método manual. (De los 96 trenes de PUMs válidos para el análisis, del método automático se obtuvieron 102 medidas válidas de “jitter”, mientras que del método manual sólo se obtuvieron 84).
- 4- La utilidad clínica práctica del algoritmo propuesto es considerable, ya que permite obtener medidas automáticas fiables y precisas del “jitter” utilizando electrodos concéntricos, eliminando así los inconvenientes del uso de electrodos fibra simple:

dificultad práctica, elevado precio y riesgo de infección.

- 5- Como subproducto de la técnica, los intervalos válidos detectados podrían servir para validar la posición de picos o instantes de análisis marcados por los expertos en exámenes de EMG de fibra única o concéntrica, mejorando la fiabilidad y el rendimiento de los procedimientos manuales y semiautomáticos.
- 6- Se deben realizar pruebas con mayores y más diversos conjuntos de señales EMG para comprender mejor el comportamiento, ventajas y deficiencias del método automático.

**Agradecimientos:** Este trabajo ha sido financiado por el Ministerio de Ciencia e Innovación, como parte del Proyecto PID2019-109062RB-I00.

#### Referencias

- [1] D.B. Sanders, J.A. Kouyoumdjian, E. Stålberg. “Single fiber electromyography and measuring jitter with concentric needle electrodes”. *Muscle Nerve*. 2022 Aug;66(2):118-130.
- [2] De Carvalho M, Swash M. Lower motor unit dysfunction in ALS. *Clinical Neurophysiology* 127 (2016) 2670–2681.
- [3] Hourigan ML, McKinnon NB, Johnson M, Rice CL, Stashuk DW, Doherty TJ. Increased motor unit potential shape variability across consecutive motor unit discharges in the tibialis anterior and vastus medialis muscles of healthy older subjects. *Clin Neurophysiol*. 2015 Dec; 126(12):2381-9.
- [4] Stålberg E. Jitter analysis with concentric needle electrodes. *Ann. N.Y. Acad. Sci.* 1274 (2012) 77–85.
- [5] Stashuk DW. Detecting single fiber contributions to motor unit action potentials. *Muscle Nerve* 1999; 22:218–29.
- [6] Abdelmaseeh M, Stashuk DW. Motor unit potential jitter: a new measure of neuromuscular transmission instability. *IEEE Trans Neural Syst Rehabil Eng*. 2017 Jul;25(7):1018-1025.
- [7] Piasecki M, Garnés-Camarena O, Stashuk DW. “Near-fiber electromyography”. *Neurophysiol*. 2021 May;132(5): 1089-1104.
- [8] A. Malanda Trigueros, D. Stashuk, J. Navallas Irujo, J. Rodríguez Falces, “Medición automática del “jitter” en registros electromiográficos”. Congreso Anual de la Sociedad Española de Ingeniería Biomédica (CASEIB). Santander, Noviembre de 2019.
- [9] Navallas J. Simulation in relation to concentric needle jitter. *International Congress of Clinical Neurophysiology*, Berlin. 2014.
- [10] C. Valle, A. Malanda, D. Stashuk, O. Garnés, “Interfaz para el estudio del “jitter” neuromuscular”. Congreso Anual de la Sociedad Española de Ingeniería Biomédica (CASEIB). Valladolid, Noviembre de 2021.
- [11] D.B. Sanders, J.A. Kouyoumdjian, E. Stålberg. “Single fiber electromyography and measuring jitter with concentric needle electrodes”. *Muscle Nerve*. 2022 Aug;66(2):118-130.

# Impacto de la expresión del gen *MAPT* en las relaciones inter-banda de la actividad eléctrica cerebral

A. Maturana-Candelas<sup>1,2</sup>, J. Poza<sup>1,2,3</sup>, R. Hornero<sup>1,2,3</sup>, V. Gutiérrez-de Pablo<sup>1,2</sup>, C. Gómez<sup>1,2</sup>

<sup>1</sup> Grupo de Ingeniería Biomédica, Universidad de Valladolid, Valladolid, España, aaron.maturana@gib.tel.uva.es

<sup>2</sup> Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), España

<sup>3</sup> IMUVA, Instituto de Investigación en Matemáticas de la Universidad de Valladolid, Valladolid, España

## Resumen

*La enfermedad de Alzheimer (EA) a menudo se relaciona con alteraciones en la actividad eléctrica cerebral. Estas pueden manifestarse como interacciones neurodinámicas entre diferentes bandas de frecuencia en regiones cerebrales específicas. Aunque la EA es una enfermedad multicausal, la implicación de factores bioquímicos en estas alteraciones no es frecuentemente estudiada. Es por ello que la expresión de genes que pueden contribuir a estas perturbaciones es de gran interés. El gen *MAPT* (microtubule associated protein tau) codifica la estructura de la proteína tau, la cual está íntimamente relacionada con la EA. El objetivo de este trabajo es estudiar el impacto de la expresión de este gen en las interacciones inter-banda de la actividad neuronal en la EA. Para ello, se calculó la información mutua (IM) inter-banda a partir de datos de electroencefalograma de 155 sujetos divididos entre pacientes con EA y sujetos de control. A su vez, éstos se subdividieron entre portadores de alelos de riesgo y protectores de *MAPT*. La IM entre la banda zeta y las restantes bandas de frecuencia reveló las diferencias en disrupción más significativas entre genotipos. Este fenómeno podría deberse a las diferencias estructurales de tau, provocadas por diferentes expresiones de *MAPT*, que estarían provocando alteraciones cerebrales incluso en estados preclínicos.*

## 1. Motivación

La enfermedad de Alzheimer (EA) es un trastorno neurodegenerativo caracterizado por cambios bioquímicos y anatómicos en el cerebro [1]. Estas alteraciones están asociadas a daño del tejido del sistema nervioso central que culmina en deterioro cognitivo [1]. La disrupción de la integridad estructural cerebral implica variaciones en su funcionalidad, manifestándose en la actividad eléctrica cerebral [2]. En las últimas décadas se han estudiado ampliamente la naturaleza de estas alteraciones y se han desarrollado numerosas técnicas de adquisición y procesado de actividad eléctrica para tal fin. Uno de los métodos más comúnmente empleados para obtener datos funcionales cerebrales es el electroencefalograma (EEG).

El EEG registra información sobre las fluctuaciones de potencial eléctrico en la superficie del cuero cabelludo. Las técnicas de adquisición de EEG pueden realizar un seguimiento de la naturaleza rápida y transitoria de la dinámica neuroeléctrica, gracias a su alta resolución temporal [3]. Además, cuentan con ventajas adicionales, como bajo coste, portabilidad y no invasividad. Estudios previos han demostrado la utilidad de los análisis de señales de EEG en el estudio de las alteraciones de la actividad cerebral provocadas por la EA [3]. Además, el

EEG se ha empleado para llevar a cabo estudios de caracterización de la funcionalidad eléctrica para comprobar el impacto de diversos biomarcadores en la actividad cerebral [4, 5].

Uno de los biomarcadores establecidos como estándar a la hora de diagnosticar EA es la presencia de la proteína tau hiperfosforilada en el tejido nervioso [6]. Tau es una proteína enlazada al citoesqueleto axonal de la neurona, confiriendo estabilidad estructural, además de formar parte de procesos de transporte molecular [6]. Durante la evolución de la EA se observa un desprendimiento de esta proteína a causa de la hiperfosforilación [6]. Este fenómeno resulta en agregados de tau intracelulares llamados ovillos neurofibrilares [6]. Estas estructuras han sido previamente asociadas a efectos neurotóxicos y se ha sugerido como una de las principales causas de atrofia [6]. Por otro lado, se ha demostrado que pequeñas variaciones en la estructura de tau causadas por diferentes genotipos pueden tener un impacto en el desarrollo de la neurodegeneración [7]. Es por ello que el estudio de las características genéticas vinculadas a tau es crucial para comprender sus implicaciones en la actividad neural.

La proteína tau humana está codificada por el gen *MAPT* (microtubule associated protein tau), que abarca alrededor de 150 kb de secuencia de nucleótidos en el cromosoma 17q21.3 [8]. Entre las modificaciones genéticas, la más abundante se manifiesta en forma de polimorfismos de nucleótido único (SNP, *Single Nucleotide Polymorphism*) [9]. Los SNP son cambios de base en loci genéticos específicos. Se han realizado múltiples estudios de asociación de todo el genoma para determinar las asociaciones estadísticas entre genotipos concretos y la neurodegeneración [10]. En este sentido, diferentes alelos del gen *MAPT* se han relacionado con la EA [11], lo que indica que diferentes alelos de *MAPT* pueden conferir riesgos o efectos protectores contra los trastornos neurológicos. Previamente, se ha establecido el SNP rs8070723-G como un factor de riesgo significativo en el desarrollo de la EA [12].

Los cambios fisiológicos del cerebro derivados de diferentes especies de tau pueden tener repercusiones en el desarrollo de enfermedades neurodegenerativas. Debido a que la EA se considera un “síndrome de desconexión”, ello podría reflejarse en anomalías en la comunicación de ciertos subsistemas cerebrales con otros. Se ha sugerido anteriormente que estas sinergias pueden manifestarse en interacciones de la actividad cerebral entre diferentes

bandas de frecuencia [13]. Por este motivo, comprobar en qué medida se alteran estas interacciones en la EA ayudaría a obtener información sobre la perturbación de la funcionalidad cerebral. Así pues, el objetivo de este trabajo es estimar la información que comparte la actividad cerebral en cada banda de frecuencia y las demás. De este modo, será posible cuantificar la disrupción funcional causada por la EA en función de la expresión génica.

## 2. Materiales y métodos

### 2.1. Sujetos

Ciento cincuenta y cinco sujetos de edad avanzada (mínimo, 65 años) participaron en este estudio y fueron clasificados como sujetos de control (45) y pacientes con EA (110). La evaluación del estado cognitivo se realizó mediante el test *Mini-Mental State Examination* (MMSE). Cada participante, familiar o representante legal brindó su consentimiento informado, según las recomendaciones del Código de Ética de la Asociación Médica Mundial. Este estudio cumplió con la Declaración de Helsinki y su protocolo fue aprobado por el Comité de Ética de la Universidad de Oporto (Oporto, Portugal, nº 38/CEUP/2018).

### 2.2. Análisis genético

Se recolectó una muestra de saliva de cada sujeto mediante un kit DNA Genotek Oragene DNA (OG-500). Las muestras biológicas se genotipificaron utilizando Thermo Fisher Scientific Axiom™ Spain Biobank Arrays en el Centro Nacional de Genotipado de España (CeGEN, Santiago de Compostela, España), y el *variant calling* del SNP se realizó con Affymetrix Power Tools. Este protocolo fue seguido por análisis individuales y de marcadores, de acuerdo con la guía de mejores prácticas de Affymetrix. El SNP rs8070723 fue calificado como válido en el análisis de control de calidad. Los datos genéticos y demográficos se muestran en la tabla 1.

### 2.3. Adquisición y pre-procesado de EEG

Se adquirieron cinco minutos de EEG para cada sujeto, con una frecuencia de muestreo de 500 Hz y referenciado promedio. El dispositivo de adquisición fue un sistema EEG Nihon Kohden Neurofax JE-921A de 19 canales. Las posiciones de los electrodos se establecieron en Fp1, Fp2, F3, F4, F7, F8, T3, T4, T5, T6, C3, C4, P3, P4, O1, O2, Fz, Cz y Pz del sistema internacional 10–20. Los registros de EEG se obtuvieron en condiciones de reposo en un ambiente relajado y sin ruido con los ojos cerrados. Se

monitorizó el estado de cada sujeto para minimizar la somnolencia. Los datos fueron almacenados en formato ASCII para luego aplicar una etapa de pre-procesado a las señales EEG: (i) sustracción de la media; (ii) filtrado pasabanda con ventana Hamming de 1-30 Hz; (iii) eliminación de ruido causado por la impedancia de los electrodos y señales no cerebrales mediante análisis de componentes independientes; (iv) segmentación de la señal en épocas de 5 segundos; y (v) inspección visual de la señal y rechazo de épocas contaminadas con artefactos. Finalmente, el EEG fue filtrado en las bandas de frecuencia clásicas: delta (1-4 Hz), zeta (4-8 Hz), alfa (8-13 Hz) y beta (13-30 Hz). El procesado y pre-procesado del EEG se realizaron con MATLAB® (versión R2021b, Mathworks, Natick, MA).

### 2.4. Análisis de EEG

Con el fin de estimar en qué medida la información contenida en la actividad eléctrica es compartida entre bandas de frecuencia, se utilizó la información mutua (IM). La IM permite cuantificar la información compartida entre dos series temporales [14]. Esta medida cuenta con la ventaja sobre el cálculo de correlación simple de ser sensible a interacciones tanto lineales como no lineales [14]. La IM entre dos series temporales  $X = \{x_i\}$  e  $Y = \{y_i\}$  se define como [14]:

$$IM_{XY} = \sum_{x_i, y_j} P_{XY}(x_i, y_j) \log_2 \frac{P_{XY}(x_i, y_j)}{P_X(x_i)P_Y(y_j)},$$

siendo  $P_X(x_i)$  y  $P_Y(y_j)$  las probabilidades de que una medición aislada de la serie se encuentre en el  $i$ -ésimo elemento de los histogramas  $P_X(x)$  y  $P_Y(y)$ , y siendo  $P_{XY}(x, y)$  la densidad de probabilidad conjunta de  $X$  e  $Y$ . Si las series temporales  $X$  e  $Y$  son completamente independientes una de otra, la IM es cero [14].

La IM se calculó a partir de cada época de EEG en cada canal consigo misma en diferentes bandas de frecuencia. Por ejemplo, se calculó la IM entre la primera época en el primer canal filtrada en delta y esa misma época filtrada en zeta. Este proceso se iteró para el resto de las bandas de frecuencia, reflejando así las interacciones inter-banda (p. ej. alfa con respecto a delta, zeta y beta). En este trabajo, únicamente se estudió la IM de cada canal consigo mismo debido al elevado coste computacional. A continuación, las interacciones de cada banda con el resto fueron también promediadas, obteniéndose un vector de 19 valores (uno por canal) para cada sujeto y banda “dominante” (delta vs. resto, zeta vs. resto, etc.), clasificados a su vez en portadores del alelo de riesgo o protector. Finalmente, se promediaron los valores de IM por épocas y por sujetos.

SNP rs8070723		Sujetos		Sexo (H:M)		Edad (media ± DT) (años)		MMSE (media ± DT)	
Base	Efecto	C	EA	C	EA	C	EA	C	EA
A	Riesgo	23	62	12:11	34:28	77.7±6.2	81.1±7.0	28.7±1.2	15.7±7.5
G	Protector	22	48	11:11	24:24	81.8±7.8	79.8±7.6	29.0±1.0	14.4±8.4

**Tabla 1.** Datos genéticos y demográficos. A, adenina; G, guanina; DT, desviación típica; H, hombre; M, mujer; MMSE, puntuación del Mini-Mental State Examination; C, controles; EA, pacientes con enfermedad de Alzheimer.

### 3. Resultados

La IM del EEG en cada banda de frecuencia fue calculada con respecto a las demás y se promedió entre ellas. Para reflejar las disparidades que pueden revelar los grupos genéticos, se ha propuesto una visualización de los resultados similar a la empleada por Yu *et al.* [15] para expresar diferencias de actividad eléctrica en regiones cerebrales específicas. Para ello, se comparan los valores de controles y los valores de las diferencias entre controles y pacientes. A éstos últimos se les denomina valores de disrupción. De esta manera, quedan expuestas las tendencias provocadas por la enfermedad en un parámetro concreto y cuán disruptivos pueden ser sus efectos en el mismo. En nuestro caso, se ha empleado este método usando la IM. La banda de frecuencia que reveló las diferencias más significativas entre genotipos en ambos ejes fue la banda zeta ( $p$ -valores  $< 0.001$ , test  $U$  de Mann-Whitney), cuyos resultados son mostrados en la figura 1. Se puede observar que los valores de IM asociados a los sujetos control con alelo de riesgo son más elevados y dispares que los del alelo protector. Por otro lado, los valores de disrupción de IM (*i.e.* eje  $y$ ) están más agrupados y próximos a cero en el genotipo protector que en el de riesgo. Además, existe una dependencia lineal mayor y negativa en los portadores del alelo de riesgo ( $\rho = -0.931$ ,  $p$ -valor  $< 0.001$ , correlación de Pearson) que en los del alelo protector ( $\rho = 0.404$ ,  $p$ -valor  $< 0.086$ , correlación de Pearson), siendo ésta última no significativa (para un umbral de significación de  $\alpha = 0.05$ ).

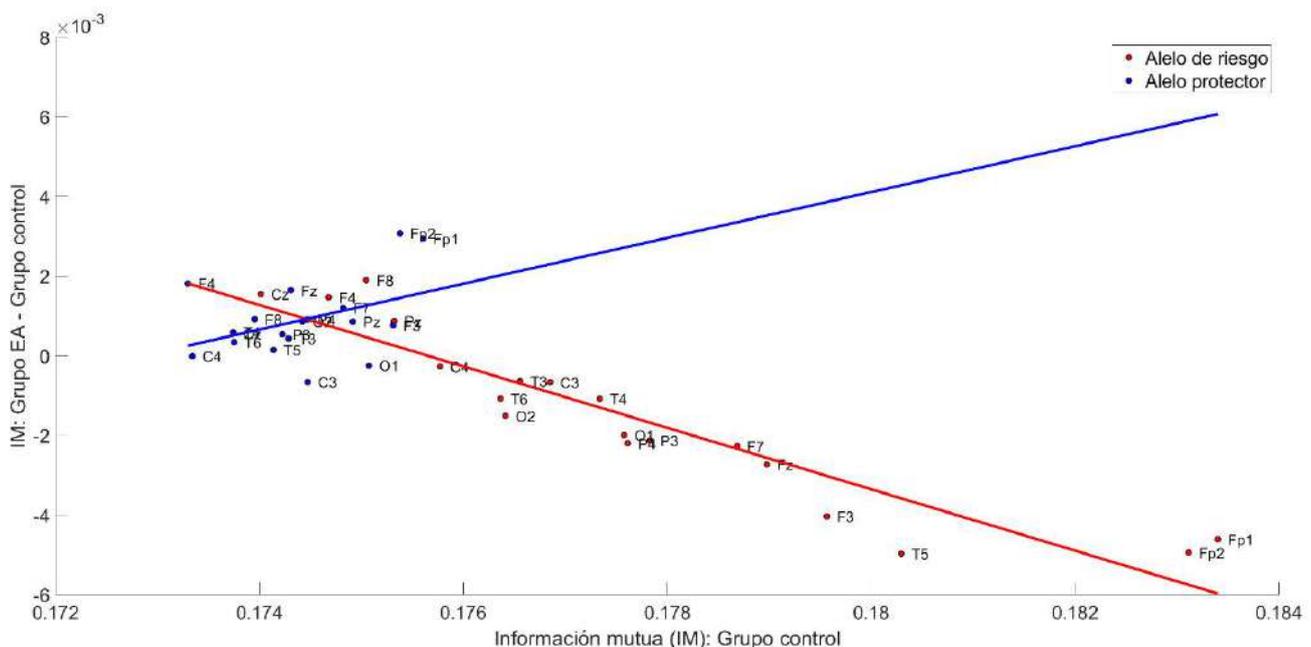
### 4. Discusión

En este trabajo se obtuvieron los valores de IM que reflejan las interacciones inter-banda de cada canal. Dicho de otra manera, se cuantificó la cantidad de información en una

banda que es transmitida o manifestada en el resto de las bandas de frecuencia en la misma región. La banda zeta reveló las mayores diferencias entre genotipos.

Se pueden comprobar diferencias sustanciales entre los resultados obtenidos en los portadores del alelo de riesgo y protector. Estas variaciones parecen deberse no tanto a los pacientes con EA, sino a la naturaleza de la actividad eléctrica de los propios controles. Uno de los aspectos clave es la elevada diversidad de los valores de IM en sujetos control portadores del alelo de riesgo. Esto podría ser debido a una actividad eléctrica aberrante en diferentes regiones cerebrales. Es sabido que el desarrollo de enfermedades neurodegenerativas está asociado a la pérdida de conectividad funcional [3]. Así pues, se podría sugerir que los sujetos control con alelo de riesgo podrían estar manifestando signos más cercanos a cuadros neurodegenerativos que los sujetos con alelo protector, a pesar de no haber desarrollado todavía ningún deterioro cognitivo.

Por otro lado, los valores más elevados de IM en los controles con alelo de riesgo indican un aumento de información compartida entre el EEG en la banda zeta y el resto. Esto podría ser un indicador de la pérdida de especialización de los subsistemas que trabajan en dicha banda de frecuencia en la misma región cerebral. También podría ser consecuencia de una ralentización del espectro del EEG. A pesar de que el algoritmo de IM es inmune a disminuciones de potencia, sí podría afectar que la actividad de bandas de frecuencia más elevadas influya en los intervalos asignados a zeta. Se conoce bien la ralentización del EEG como un signo de neurodegeneración [3]. Aludiendo al punto anterior, esta observación también podría estar reflejando estadios fisiológicos más próximos a la neurodegeneración.



**Figura 1.** Diagrama de dispersión entre los valores de IM en controles y los valores de disrupción de IM en cada canal de registro. Cada valor representa el promedio de la IM entre el EEG la banda zeta y el resto de bandas de frecuencia.

Por último, el hecho de que los valores de disrupción estén más próximos a cero en el caso de los portadores del alelo protector sugiere un efecto amortiguador, manteniendo las propiedades de la actividad eléctrica en la EA más próximas a estadios preclínicos.

Las diferencias observadas entre grupos genéticos podrían deberse a múltiples causas relacionadas con la naturaleza estructural de tau. Para empezar, podría ser una cuestión de cantidad; se ha sugerido que los distintos alelos *MAPT* están asociados con diversas concentraciones de tau en sujetos amiloides positivos [16]. Otra razón podría ser un plegamiento anormal de tau; un estudio previo sugirió que las alteraciones en el plegamiento de tau pueden conducir a agregaciones de proteínas más elevadas causadas por una mayor interacción tau-tau [17]. Por ello, dado que la conformación espacial de la proteína puede influir en su agregación, una isoforma de tau cuya configuración afecte a su plegamiento podría conducir a mayores concentraciones de esta molécula y desarrollar respuestas citotóxicas e inmunológicas más intensas. Este plegamiento también podría repercutir en la propagación de tau por el tejido nervioso, lo cual podría tener un impacto en sus efectos citotóxicos. Anteriormente, se ha reportado una asociación entre diferentes configuraciones estructurales de tau y distintos patrones de propagación en el cerebro [18].

A pesar de que los resultados son relevantes, es necesario mencionar ciertas limitaciones de este estudio. El más evidente es la inviabilidad de estudiar la IM entre diferentes canales debido al elevado coste computacional asociado. Realizar este cálculo permitiría comprobar interacciones entre sistemas cerebrales alejados entre sí que trabajen en diferentes bandas de frecuencia sincrónicamente. Por otro lado, aunque la actividad cerebral a nivel de sensor ofrece información sobre los procesos de activación local, la aplicación de un algoritmo de localización de fuentes podría ser útil para afinar la interpretación fisiológica de los resultados. Por último, muchos otros SNPs han sido documentados previamente en *MAPT*, siendo el rs8070723 sólo uno de ellos. Ampliar este estudio exploratorio a diferentes variaciones de este gen sería crucial para sugerir un vínculo causal entre las alteraciones inter-banda y las propiedades de tau.

## 5. Conclusión

A la vista de los resultados obtenidos, se puede sugerir que la expresión del gen *MAPT* tiene un impacto en las interacciones inter-banda. Estas alteraciones podrían estar ligadas a diferentes configuraciones estructurales de la proteína tau. Además, la banda zeta es más sensible al genotipo que las demás, puesto que las interacciones de esta con el resto de bandas son las más significativas. Esto podría arrojar pistas sobre procesos cerebrales concretos que puedan verse más alterados por causas bioquímicas.

## Agradecimientos

Esta investigación ha sido financiada por la ‘Comisión Europea’ y el ‘Fondo Europeo de Desarrollo Regional’ (FEDER) a través del proyecto ‘Análisis y correlación

entre la epigenética y la actividad cerebral para evaluar el riesgo de migraña crónica y episódica en mujeres’ (‘Programa de Cooperación Transfronteriza Interreg V-A España-Portugal 2014-2020’), por el ‘Ministerio de Ciencia e Innovación - Agencia Estatal de Investigación’ y FEDER a través del proyecto PGC2018-098214-A-I00, y por el ‘CIBER en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN)’ a través del ‘Instituto de Salud Carlos III’ cofinanciado con fondos FEDER.

## Referencias

- [1] Alzheimer's Association. 2022 Alzheimer's disease facts and figures, 2022.
- [2] D'Amelio, M., & Rossini, P. M. Brain excitability and connectivity of neuronal assemblies in Alzheimer's disease: from animal models to human findings. *Prog. Neurobiol.*, vol 99, sup 1, 2012, pp 42-60.
- [3] Vecchio, F., *et al.* Resting state cortical EEG rhythms in Alzheimer's disease: toward EEG markers for clinical applications: a review. *Suppl. Clin. Neurophysiol.*, vol 62, 2013, pp 223-236.
- [4] Smailovic, U., *et al.* Quantitative EEG power and synchronization correlate with Alzheimer's disease CSF biomarkers. *Neurobiol. Aging*, vol 63, 2018, pp 88-95.
- [5] Smailovic, U., *et al.* Synaptic molecular and neurophysiological markers are independent predictors of progression in Alzheimer's disease. *J. Alzheimer's Dis.*, vol 83, sup 1, 2021, pp 355-366.
- [6] Chi, H., *et al.* Tauopathy. In *Cognitive Disorders*. IntechOpen, 2018.
- [7] Chornenky, Y., *et al.* Tau and TDP-43 proteinopathies: kindred pathologic cascades and genetic pleiotropy. *Lab. Invest.*, vol 99, sup 7, 2019, pp 993-1007.
- [8] Pittman, A. M., *et al.* Untangling the tau gene association with neurodegenerative disorders. *Hum. Mol. Genet.*, vol 15, sup 2, 2006, pp 188-195.
- [9] Collins, F. S., *et al.* Variations on a theme: cataloging human DNA sequence variation. *Science*, vol 278, sup 5343, 1997, pp 1580-1581.
- [10] Tábuas-Pereira, M., *et al.* (2020). Alzheimer's disease genetics: Review of Novel Loci associated with disease. *Curr. Genet. Med. Rep.*, vol 8, sup 1, 2020, pp 1-16.
- [11] Gerrish, A., *et al.* The role of variation at AβPP, PSEN1, PSEN2, and MAPT in late onset Alzheimer's disease. *J. Alzheimer's Dis.*, vol 28, sup 2, 2012, pp 377-387.
- [12] Allen, M., *et al.* Association of MAPT haplotypes with Alzheimer's disease risk and MAPT brain gene expression levels. *Alzheimer's Res. Ther.*, vol 6, sup 4, 2014, pp 1-14.
- [13] Morillas-Romero, *et al.* Spontaneous EEG theta/beta ratio and delta-beta coupling in relation to attentional network functioning and self-reported attentional control. *Cogn. Affect. Behav. Neurosci.*, vol 15, sup 3, 2015, pp 598-606.
- [14] Jeong, J., *et al.* (2001). Mutual information analysis of the EEG in patients with Alzheimer's disease. *Clin. Neurophysiol.*, vol 112, sup 5, 2001, pp 827-835.
- [15] Yu, M., *et al.* Selective impairment of hippocampus and posterior hub areas in Alzheimer's disease: an MEG-based multiplex network study. *Brain*, vol 140, sup 5, 2017, pp 1466-1485.
- [16] Kauwe, J. S., *et al.* Variation in MAPT is associated with cerebrospinal fluid tau levels in the presence of amyloid-beta deposition. *Proc. Natl. Acad. Sci.*, vol 105, sup 23, 2008, pp 8050-8054.
- [17] Rudenko, L. K., *et al.* Intraneuronal tau misfolding induced by extracellular amyloid-β oligomers. *J. Alzheimer's Dis.*, vol 71, sup 4, 2019, pp 1125-1138.
- [18] Hallinan, G. I., *et al.* Tau misfolding efficiently propagates between individual intact hippocampal neurons. *J. Neurosci.*, vol 39, sup 48, 2018, pp 9623-9632.

# Telemedicina y E-Salud

---

Miércoles, 23 de noviembre de 2022

# Implementación y evaluación de una plataforma tecnológica para estrategias nutricionales personalizadas en personas mayores

J.M. Iniesta<sup>1</sup>, P. Chausa<sup>1</sup>, A. García-Rudolph<sup>2</sup>, A. García-Molina<sup>2</sup>, C. Galarregui<sup>3</sup>, S. Navas-Carretero<sup>3,4,5</sup>, I. Abete<sup>3,4,5</sup>, M.A. Zulet<sup>3,4,5</sup>, J.A. Martínez<sup>3,7</sup>, C.J. González-Navarro<sup>6</sup>, H. Marcos-Pasero<sup>7#</sup>, E. Aguilar-Aguilar<sup>7</sup>, G. Reglero<sup>7,8</sup>, A. Ramirez de Molina<sup>9</sup>, V. Loria-Kohen<sup>7§</sup>, M.E.Hernando<sup>1,10</sup>, Enrique J. Gómez<sup>1,10</sup>

<sup>1</sup> Grupo de Bioingeniería y Telemedicina, ETSI Telecomunicación, Centro de Tecnología Biomédica, Universidad Politécnica de Madrid, Madrid, España, [jm.iniesta/paloma.chausa/mariaelena.hernando/enriquejavier.gomez@upm.es](mailto:jm.iniesta/paloma.chausa/mariaelena.hernando/enriquejavier.gomez@upm.es)

<sup>2</sup> Department of Research and Innovation, Institut Guttmann, Institut Universitari De Neurorehabilitació Adscrit a La UAB, Barcelona, Spain; Bellaterra (Cerdanyola Del Vallès), Universitat Autònoma De Barcelona, España; Fundació Institut d'Investigació En Ciències De La Salut Germans Trias I Pujol, Barcelona, España. [agarcia/agarciam@guttmann.com](mailto:agarcia/agarciam@guttmann.com)

<sup>3</sup> Department of Nutrition, Food Sciences and Physiology and Center for Nutrition Research, Faculty of Pharmacy and Nutrition, University of Navarra, 31008 Pamplona, España. [cgalarregui@alumni.unav.es](mailto:cgalarregui@alumni.unav.es); [snavas/iabetego/mazulet/jalfmtz@unav.es](mailto:snavas/iabetego/mazulet/jalfmtz@unav.es)

<sup>4</sup> Navarra Institute for Health Research (IdiSNA), 31008 Pamplona, España

<sup>5</sup> CIBER de la Fisiopatología de la Obesidad y Nutrición, Instituto de Salud Carlos III

<sup>6</sup> Center for Nutrition Research, Faculty of Pharmacy and Nutrition, University of Navarra, 31008 Pamplona, Spain. [cgnavarro@unav.es](mailto:cgnavarro@unav.es)

<sup>7</sup> Nutrition and Clinical Trials Unit, GENYAL Platform IMDEA-Food Institute, CEI UAM+CSIC, 28049, Madrid, Spain. [elena.aguilar/guillermo.reglero@imdea.org](mailto:elena.aguilar/guillermo.reglero@imdea.org)

<sup>8</sup> Department of Production and Characterization of Novel Foods, Institute of Food Science Research (CIAL) CEI UAM + CSIC, 28049, Madrid, España.

<sup>9</sup> Molecular Oncology and Nutritional Genomics of Cancer, IMDEA-Food Institute, CEI UAM+CSIC, 28049, Madrid, España. [ana.ramirez@imdea.org](mailto:ana.ramirez@imdea.org)

<sup>10</sup> CIBER de Bioingeniería, Biomateriales y Nanomedicina, Instituto de Salud Carlos III

<sup>#</sup> Current address: Bioactivity and Nutritional Immunology Group (BIOINUT), Faculty of Health Sciences, Universidad Internacional de Valencia—VIU, Pintor Sorolla 21, 46002 Valencia, Spain. [hmarcos@universidadviu.com](mailto:hmarcos@universidadviu.com)

<sup>§</sup> Current address: Departamento de Nutrición y Ciencia de los Alimentos, Faculty of Pharmacy, Universidad Complutense de Madrid, 28040 Madrid, Spain. [vloria@ucm.es](mailto:vloria@ucm.es)

## Resumen

*El estilo de vida de los países desarrollados ha orientado a la población hacia una alimentación desequilibrada acompañada de niveles de sedentarismo elevados, lo que ha provocado una mayor prevalencia de síndrome metabólico y alteraciones asociadas. En las últimas décadas se han desarrollado estrategias nutricionales para concienciar a la población sobre la importancia de una alimentación y estilo de vida saludable. Sin embargo, cada vez son más evidentes las amplias diferencias que se dan entre individuos en la respuesta a la alimentación dependen de las características de cada uno. En este sentido se hace necesaria la adaptación del tratamiento nutricional a cada condición individual. Este trabajo de investigación ha definido, desarrollado y evaluado una plataforma tecnológica para dar soporte a estrategias nutricionales personalizadas en personas mayores. La plataforma provee a los profesionales de la salud con un sistema de ayuda a la decisión para la personalización de la dieta y herramientas de gestión para facilitar las intervenciones nutricionales, y a las personas mayores con herramientas que facilitan la recogida de datos, mecanismos para promover la adherencia, información rigurosa y el acceso*

*a un sistema para valorar y entrenar sus funciones cognitivas. La plataforma ha sido evaluada mediante una intervención de 3 meses de duración (n=127) y ha demostrado ser usable (SUS=76), funcional, robusta y confiable (tiempo de respuesta medio = 113ms; hasta 200 usuarios simultáneos).*

## 1. Introducción

Evidencias científicas recientes demuestran la influencia del estilo de vida en el riesgo de desarrollar síndrome metabólico, el cual se caracteriza por la agrupación de varias anomalías metabólicas observadas con frecuencia en la práctica clínica: obesidad abdominal, dislipemia, hiperinsulinemia, glucosa en ayunas alterada e hipertensión arterial [1]. Además, el envejecimiento de la población está produciendo un aumento de la prevalencia de enfermedades crónicas no transmisibles, entre ellas la obesidad y las comorbilidades metabólicas asociadas, como la diabetes mellitus y las enfermedades cardiovasculares, que son causas muy importantes de

mortalidad y discapacidad [2]. El estilo de vida, y en particular la alimentación, son factores modificables con especial importancia en el desarrollo del síndrome metabólico, la evolución del envejecimiento y las patologías asociadas [1,2].

La nutrición de precisión se presenta como un prometedor enfoque terapéutico para el desarrollo de pautas dietéticas en grupos específicos, adaptándose a los requerimientos nutricionales y metabólicos de cada individuo, ya que tiene en cuenta el perfil genético y la marca epigenética de cada persona, así como su edad, sexo, actividad física, alimentación, factores ambientales o estado fisiopatológico. A pesar de los beneficios potenciales, la aplicación de este enfoque aún está limitada a unas pocas personas debido a las dificultades encontradas en la gestión de datos y la prestación de servicios. Además, la eficacia de las prescripciones dietéticas se ve limitada por la falta de adherencia. Esto suele deberse a la complejidad de la prescripción misma y/o a la falta de compromiso del individuo

En este contexto, la Ingeniería Biomédica y las Tecnologías de la Información y las Comunicaciones (TIC) pueden resultar de gran ayuda gracias a la incorporación de herramientas para una gestión eficaz de las estrategias nutricionales. Estas herramientas podrían ayudar a generar pautas de alimentación adecuadas a cada usuario, proporcionando información precisa sobre su estado de salud y potenciando la comunicación con los profesionales que intervienen en el proceso. Esto se puede realizar de forma personalizada y dinámica, llevando al usuario a ser parte activa del proceso y en consecuencia, mejorando la autogestión y la adherencia [3].

El proyecto Nutriprecisión tiene como objetivo desarrollar y evaluar estrategias para la mejora de la calidad de vida de personas mayores (>50 años) basadas en la nutrición de precisión. Para ello ha comparado la efectividad de dos estrategias dietéticas: una estrategia control basada en las recomendaciones actuales de la Sociedad Española de Nutrición Comunitaria y una estrategia experimental, compuesta por una dieta personalizada, alimentos funcionales y soportada por herramientas TIC.

Este artículo describe las herramientas tecnológicas desarrolladas para dar soporte a la intervención y el resultado de evaluación de las mismas.

## 2. Materiales y métodos

El estudio del proyecto Nutriprecisión se planteó como un estudio multicentro prospectivo paralelo (grupo control y experimental) y randomizado de 12 semanas de duración donde se analizó la efectividad de dos estrategias nutricionales (estrategia control basada en recomendaciones generales actuales vs estrategia nutricional de precisión) sobre el estado de salud global de población entre 50 y 80 años con alteraciones metabólicas. A lo largo de las 12 semanas de estudio, se establecieron un total de 5 visitas: 1) información del estudio y screening; 2) día 0: inicio de la intervención; 3) día 28: sesión grupal (grupo control) / visita de seguimiento (grupo experimental); 4) día 56: sesión grupal (grupo control) /

visita de seguimiento (grupo experimental) y 5) día 84: finalización de la intervención. En el estudio participaron 127 personas: 65 en el grupo control y 64 en el grupo experimental.

Para la personalización de la dieta, el proyecto Nutriprecisión definió un marcador de salud basado en una serie de determinantes antropométricos y clínicos: edad, sexo, peso, altura, IMC, circunferencia de la cintura, hemoglobina glicosilada, colesterol total, HDL, LDL, triglicéridos, ácido úrico, presión arterial sistólica y diastólica, salud gastrointestinal y función cognitiva. El procesamiento de estas medidas y escalas permiten obtener el estado (0 valores normales; 1 valores de riesgo; 2 valores patológicos) de condiciones comunes en personas mayores: hipertensión, diabetes, obesidad, función cognitiva, función digestiva y síndrome metabólico [4]. El marcador de salud final resulta de la concatenación de los marcadores de cada condición.

El desarrollo de las herramientas TIC se ha realizado mediante metodologías ágiles de desarrollo de software, que permiten que los requisitos y las soluciones evolucionen con el tiempo en base a las necesidades de los usuarios. Se ha trabajado por funcionalidades, de manera que, en cada iteración, se ha incrementado el valor de una parte o funcionalidad del software.

El desarrollo de las aplicaciones se ha basado en el patrón Modelo-Vista-Controlador mediante tecnologías Web (framework Ionic, HTML, CSS, TypeScript, Angular). Para la implementación de los módulos y componentes del “backend” se ha utilizado el servidor NodeJS y para almacenar los datos, el servidor de base de datos MySQL. Por último, se han utilizado diferentes tecnologías para la conexión con diversos servicios en la nube: Firebase para la autenticación, IBM Watson Studio para el algoritmo del marcador de salud, plataforma eCRF REDCap y Google Spreadsheets para la gestión de agenda de citas.

La evaluación de las herramientas TIC que dan soporte a la estrategia nutricional de precisión se ha centrado en el análisis de la usabilidad, funcionalidad, robustez y confiabilidad. La usabilidad ha sido evaluada mediante la Escala de Usabilidad del Sistema (SUS) [5], un cuestionario de 10 ítems con 5 opciones de respuesta, desde Totalmente en desacuerdo a Totalmente de acuerdo. El SUS se puede usar en muestras pequeñas de población con resultados confiables y la escala es fácilmente administrable a los usuarios. El resultado obtenido varía entre 0 y 100, siendo 0 indicativo de muy baja usabilidad y 100 usabilidad excelente. La evaluación de la funcionalidad se ha realizado mediante pruebas de sistema en laboratorio en las cuales se analizaba que el flujo de acciones disponibles en las aplicaciones cumpliera con cada uno de los requisitos funcionales identificados. La robustez y la confiabilidad, se han medido mediante pruebas de carga y estrés en laboratorio haciendo uso de la herramienta Apache JMeter (<https://jmeter.apache.org/>) con diferentes volúmenes de estímulo a aplicar sobre el sistema considerando parámetros como la cantidad de usuarios, cantidad de iteraciones, etc. Las pruebas de carga se configuraron con 100 usuarios simultáneos, haciendo transacciones hacia la plataforma TIC e iterando las

pruebas 10 veces. Las pruebas de estrés buscan romper el sistema. Se va aumentando el número de usuarios que se agregan a la aplicación y se ejecutan las mismas pruebas de cargas realizadas en el punto anterior hasta que se rompe. Con esta prueba se determina la solidez de la aplicación en momentos de carga extrema y si en un caso eventual podría soportar una carga superior a la esperada.

### 3. Resultados

#### 3.1. Arquitectura de la plataforma de intervención

La arquitectura de la plataforma tecnológica que da soporte a la estrategia nutricional de precisión está compuesta de los siguientes módulos (Figura 1).

- **Aplicación móvil del voluntario.** Permite el seguimiento de la estrategia experimental (agenda de eventos y visitas, avisos), el registro de datos (dieta, actividad física, tolerancia y percepción sensorial de los productos funcionales, calidad de vida, sintomatología gastrointestinal, etc.) a través de cuestionarios agrupados en los cuadernos de registro de datos de cada visita e incorpora mecanismos para promover la adherencia a la dieta (mensajes motivacionales) así como información rigurosa a partir de fuentes bibliográficas fiables.
- **Aplicación web para el profesional.** Permite el registro de datos durante las visitas del estudio incluyendo el acceso a cuestionarios digitales implementados en la plataforma REDCap de la Universidad Politécnica de Madrid. Además, los profesionales pueden realizar la gestión de visitas, asignando fecha, hora, voluntario y profesional a cargo. Estos datos actualizarán la agenda y se generarán los avisos de visitas a los voluntarios de forma automática. Este módulo integra a su vez el sistema de ayuda a la decisión para personalización de la dieta, descrito a continuación.
- **Sistema de ayuda a la decisión para personalización de la dieta.** Se ha implementado un sistema de ayuda a la decisión basado en árboles de decisión. A partir de las medidas antropométricas y clínicas citadas

anteriormente, el sistema calcula el marcador de salud final y en base a ese marcador determina la dieta más adecuada de entre las disponibles.

- **Herramienta para la valoración y la estimulación de las funciones cognitivas.** Plataforma web que proporciona ejercicios agrupados en tres fases: fase PRE, con 12 ejercicios de valoración cognitiva; fase de ENTRENAMIENTO, la cual consiste en la propuesta de 32 ejercicios diarios con el objetivo de entrenar las funciones cognitivas durante un tiempo concreto; y la fase POST, igual que la PRE, con el objetivo de poder comparar la evolución del usuario
- **Servidor web, base de datos y API (Application Programming Interface) de conexión con servicios.** Su función principal es la de recibir todos los datos registrados por los voluntarios en la aplicación móvil, procesarlos y almacenarlos en una base de datos, utilizando una API específica para la comunicación de los datos. Por otro lado, a partir de la agenda de visitas administrada por los profesionales, el servidor se encarga de enviar los avisos, recordatorios y recomendaciones a los voluntarios. Los datos de las diferentes dimensiones de interés (dieta, actividad física, consumo de productos, etc.) se exportan y quedan disponibles online para el acceso por parte de los profesionales.

Tanto para la autenticación y login de la aplicación móvil como para el envío de notificaciones al móvil, se han utilizado servicios de Firebase. El servidor Nutriprecisión administra estos servicios y mecanismos, por un lado, para crear las cuentas de los voluntarios y validar el acceso en la aplicación, y por otro para poder utilizar un canal de comunicación con la aplicación móvil para el envío de mensajes y recomendaciones personalizadas. Otro servicio de la nube que se utiliza son las hojas de cálculo de Google. La agenda de visitas y los datos exportados de los voluntarios son almacenados en hojas de cálculo online, disponibles para la gestión y valoración por parte de los profesionales que controlan los ensayos del proyecto.

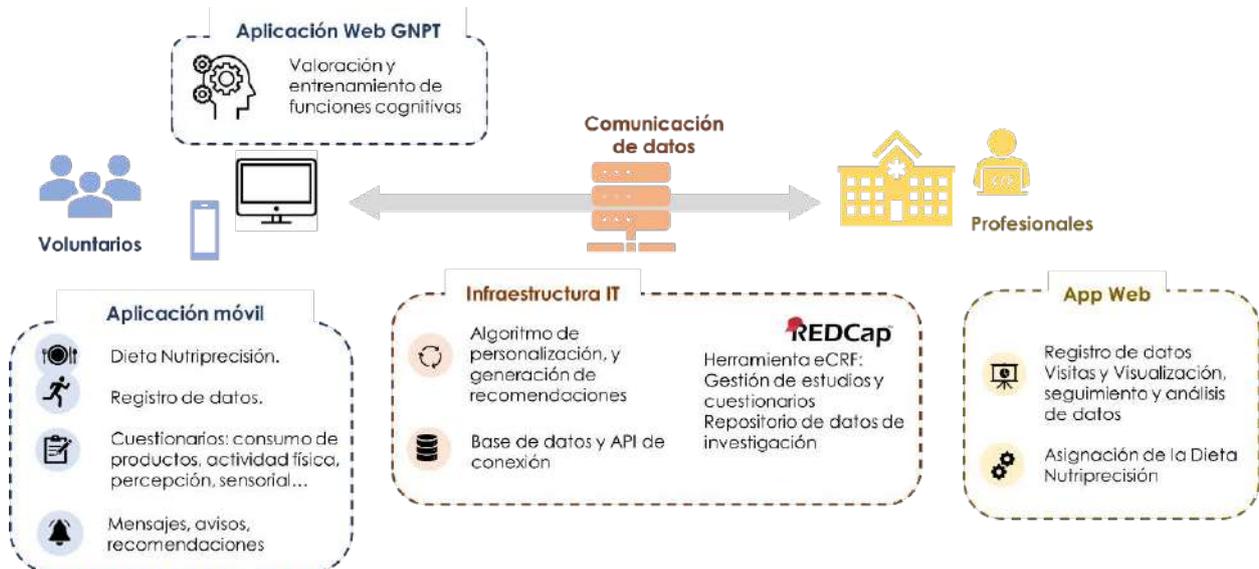


Figura 1 - Arquitectura plataforma Nutriprecisión

### 3.2. Interfaz de usuario

La Figura 2, muestra la interfaz de usuario de la aplicación móvil del voluntario.

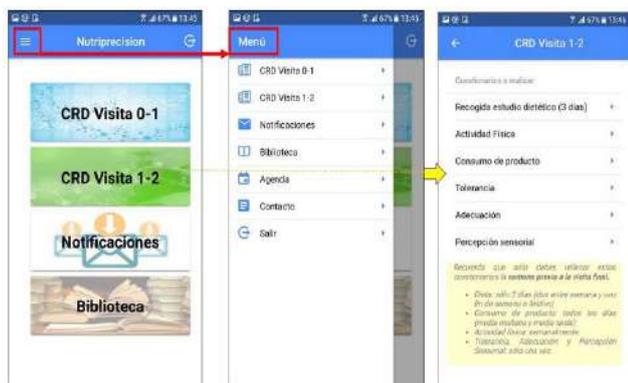


Figura 2 – Interfaz de usuario de la aplicación móvil del voluntario

### 3.3. Evaluación técnica

La aplicación móvil fue utilizada por las personas del grupo experimental (n=64) y la aplicación web para profesionales por los 5 profesionales que llevaron a cabo la intervención. La Tabla 1 muestra el resultado de la evaluación de la usabilidad de las herramientas TIC obtenido mediante el cuestionario SUS.

	n total	n	SUS (0-100)
Aplicación móvil para voluntarios	64	40	76,2
Aplicación web para profesionales	5	5	76

Tabla 1. Resultado evaluación usabilidad

Realizando un análisis detallado de cada una de las preguntas del cuestionario SUS, observamos que tanto los voluntarios como los profesionales han valorado de forma positiva todos los aspectos de usabilidad, destacando la facilidad de uso y una curva de aprendizaje corta en el uso de las aplicaciones.

Con respecto a los criterios de robustez y confiabilidad, vemos como las pruebas de carga realizadas no colapsaron el sistema y no se generaron errores en las transacciones, obteniéndose un tiempo de respuesta medio de 113ms. Por otro lado, las pruebas de estrés mostraron que al llegar a los 200 usuarios haciendo peticiones simultáneas al sistema, el tiempo de respuesta se incrementaba de forma considerable, resultando en 2.6s y comenzaba a detectarse un porcentaje de error del 2.65%.

## 4. Conclusiones

Este trabajo de investigación describe el diseño, desarrollo y evaluación de una plataforma tecnológica que permite dar soporte a la gestión de estrategias nutricionales personalizadas en personas mayores.

La plataforma ha sido evaluada mediante una intervención de 3 meses de duración, en la que se ha comparado la efectividad de 2 estrategias nutricionales diferentes: una

estrategia control (n=65) y una experimental (n=64), compuesta por una dieta personalizada, alimentos funcionales y soportada por herramientas TIC. Los resultados clínicos de esta intervención están pendientes de publicación. Todas las herramientas desarrolladas obtuvieron resultados notables de usabilidad (utilidad, facilidad de aprendizaje, rendimiento y satisfacción) y las pruebas de laboratorio demostraron que la plataforma es funcional, robusta y confiable. El tiempo de respuesta medio, 113ms, se puede considerar más que aceptable ya que hasta 200ms la sensación del usuario es de respuesta instantánea del sistema. La plataforma permite hasta 200 usuarios simultáneos accediendo y transfiriendo datos, con lo que podría dar soporte a estudios clínicos con mayor número de usuarios.

A nivel funcional, la plataforma ayuda a los profesionales de la salud a realizar una gestión eficaz de la intervención facilitando el manejo de la agenda, la visualización de datos y la personalización de la dieta. A su vez, facilita la recogida de datos por parte de las personas mayores, proporciona información rigurosa, recomendaciones y mensajes motivacionales y les dota de herramientas para entrenar sus funciones cognitivas. Todos estos factores favorecen la implicación de las personas mayores en la intervención, que se convierten en un elemento activo en la gestión de su salud en un entorno personalizado y dinámico.

## Agradecimientos

Este trabajo ha sido financiado por el Centro para el Desarrollo Tecnológico Industrial a través del “Programa Estratégico CIEN” [IDI-20160734], cofinanciado por el Fondo Europeo de Desarrollo Regional. En concreto, el trabajo contó con el apoyo de las empresas AMC Innova Juice and Drinks S.L., Iberfruta Muerza, S.A., Hijo de José Martínez Somalo, S.L., Grupo ICA, S.L., Congelados de Navarra, S.A., Galletas Gullón, S.A., Europastry, S.L.

## Referencias

- [1] Garralda-Del-Villar, M., Carlos-Chillerón, S., Diaz-Gutierrez, J., Ruiz-Canela, M., Gea, A., Martínez-González, M. A., ... & Fernández-Montero, A. (2018). Healthy lifestyle and incidence of metabolic syndrome in the SUN cohort. *Nutrients*, 11(1), 65.
- [2] Ageing and health. <https://www.who.int/news-room/factsheets/detail/ageing-and-health> (Consultada: julio 2022).
- [3] Lee JA, Choi M, Lee SA, Jiang N. Effective behavioral intervention strategies using mobile health applications for chronic disease management: a systematic review. *BMC Med Inform Decis Mak*. 2018;18(1):12. Published 2018 Feb 20. doi:10.1186/s12911-018-0591-0
- [4] «Estrategias de guía e ingredientes dietéticos de precisión para enfermedades crónicas en población pre-sénior y sénior,» *Anales del Sistema Sanitario de Navarra*, vol. 41, nº 2, pp. 227-243, 2018. <https://scielo.isciii.es/pdf/asisna/v41n2/1137-6627-asisna-41-02-227.pdf>
- [5] Brooke, J. (1996). SUS-A quick and dirty usability scale. *Usability evaluation in industry*, 189(194), 4-7.

# Percepción de la ciberseguridad en usuarios sobre soluciones de salud digital con dispositivos HIoT

M. I. Varona Torralvo<sup>1</sup>, A. M. Medrano Gil<sup>1</sup>, M. F. Cabrera-Umpiérrez<sup>1</sup>

<sup>1</sup> Life Supporting Technologies, Departamento de Tecnología Fotónica y Bioingeniería, Universidad Politécnica de Madrid, Madrid, España, {ivarona, a.medrano, chiqui}@lst.tfo.upm.es

## Resumen

En los últimos años, se han incrementado las tecnologías en el campo de la salud con la aparición de nuevos dispositivos Healthcare IoT. No obstante, el desarrollo e intrusión de estos dispositivos en los sistemas, ha incrementado el riesgo de aparición de nuevas vulnerabilidades que afecten a la seguridad. Bajo este marco, se presenta una metodología para la creación de una encuesta para la valoración de la experiencia del usuario, que refleje la importancia de la ciberseguridad en este ámbito, permitiendo el análisis y evaluación de su percepción en un grupo de individuos. Para el procedimiento de creación y refinamiento se ha contado con un grupo de expertos con experiencia en el campo. Finalmente, la encuesta se ha ejecutado obteniendo 112 respuestas de usuarios en un rango de edades comprendido entre los 18-67 años.

## 1. Introducción

Actualmente, la evolución y vanguardia en el área de la salud ha llevado a la inclusión de sensores médicos conectados, para brindar un enfoque personalizado en la atención médica [1]. Este tipo de dispositivos inteligentes es lo que se conoce como *Healthcare Internet of Things* o *Internet of Medical Things* (HIoT/IoMT), siendo la combinación del Internet of Things junto con sensores biomédicos portables. De esta forma se crean nuevas oportunidades en los servicios de salud; brindando accesibilidad, disponibilidad o contenido personalizado [2]. Aunque son un gran número las ventajas que ofrece, existen varios desafíos que deben abordarse para crear dispositivos adaptables, apropiados, energéticamente eficientes, y seguros [3].

Los dispositivos desarrollados hasta el momento presentan una serie de riesgos y vulnerabilidades de seguridad, debido a una serie de limitaciones. Principalmente, la creación de los equipos sin la seguridad por diseño que compromete la conexión, información y datos manejados por los equipos, así como, las escasas regulaciones y normativas legales en esta área suponen un alto riesgo [4]. Todo esto, puede llevar a la exposición de la información médica personal de un individuo, atacando su privacidad. Además, la falta de formación en el campo de los profesionales sanitarios incrementa las posibles brechas [5]. No obstante, la aplicación de medidas de seguridad y la preservación de la privacidad es la tendencia en el campo, mejorando protocolos bajo los requisitos canónicos de la ciberseguridad, conocidos como la triada CIA [6] (**Confidencialidad**, garantía de la no divulgación de la información a las entidades no autorizadas; **Integridad**,

inmutabilidad y exactitud de los datos a lo largo de su ciclo de vida; **Disponibilidad**, accesibilidad a la información por entidades autorizadas cuando se requiera).

En este artículo, se presenta el diseño y análisis de una encuesta para la valoración de la experiencia del usuario sobre la concienciación y la percepción de la ciberseguridad en la salud.

## 2. Metodología

El desarrollo de la metodología se ha enmarcado dentro de tres pilares esenciales para su proposición: 1) conocer la situación actual de la ciberseguridad dentro de la bioingeniería; 2) la importancia de su aplicación y 3) el valor de su aplicación en las tecnologías médicas. La siguiente sección describe la metodología implementada para definir y establecer la encuesta propuesta.

### 2.1. Diseño de la encuesta

Se ha realizado una búsqueda de estudios similares en esta área, encontrando cuestionarios que plantean un análisis de la ciberseguridad en otros ámbitos de salud atendiendo a conceptos como la triada CIA, privacidad, etc. [7]. No obstante, ninguna investigación aborda la ciberseguridad en dispositivos HIoT y se ha generado una encuesta propia. Para llevar a cabo la conformación de la encuesta, se establece un procedimiento de 4 pasos como muestra la *Figura 1*, que permite definir y obtener las cuestiones atendiendo a conceptos específicos sobre seguridad. Este proceso, tiene el objetivo de concretar de manera clara las preguntas de la encuesta, enmarcadas dentro del área del HIoT, y adaptadas para cualquier usuario. Las cuestiones de la encuesta se dividen en dos secciones; una primera de carácter demográfico para conocer las características de la muestra, ya que la encuesta es de carácter anónimo. Y, una segunda sección con las cuestiones enmarcadas en la concienciación de la ciberseguridad en salud. A través de estos conceptos principales, se establecen las cuestiones teniendo en cuenta que su valoración se realizará mediante la Escala de Likert con valores de 1 al 5.



*Figura 1. Metodología para la conformación de la encuesta.*

La selección de esta escala se debe a su aplicación universal, su cobertura en el rango de respuesta y la comparabilidad de la respuesta entre las preguntas [8]. El formulario está enfocado a cualquier tipo de usuario, independientemente de los conocimientos técnicos que presente, por lo que las preguntas se construyen y plantean de manera sencilla y con aclaraciones (definiciones o ejemplos) sobre las mismas, resolviendo las posibles dudas que puedan surgir respecto a los conceptos más técnicos. Para la construcción y validación de la encuesta se ha realizado un proceso de refinamiento iterativo con expertos sobre las cuestiones, para focalizar el objetivo de las preguntas. Donde, una vez conformado una primera versión del cuestionario, dos investigadores expertos en formularios para la valoración de la experiencia del usuario revisan y comprueban que hay exactitud, entendimiento y que el valor a obtener de las mismas queda reflejado y conciso. Por otro lado, las opciones de respuesta (de 1 a 5) se han discutido en paralelo a la creación de las preguntas para obtener una linealidad conceptual en la escala de cada una de ellas. Las iteraciones necesarias en el refinamiento de las cuestiones han sido de un total de 3. Tras esta verificación se conforma el cuestionario.

## 2.2. Proceso de difusión

Para la difusión y acceso a la encuesta se emplea la herramienta Google Forms que permite generar y compartir formularios, al mismo tiempo que almacena las respuestas y permite su análisis. El proceso de difusión se realiza mediante correo electrónico, y redes sociales compartiendo el enlace de la encuesta.

## 3. Resultados

Gracias a la metodología expuesta en la sección anterior, se conforma la encuesta para la valoración de la experiencia del usuario, obteniendo como resultado las respuestas que permite la evaluación preliminar.

### 3.1. Encuesta

Gracias a las distintas iteraciones realizadas con los expertos, se concluyen un total de 12 cuestiones, repartidas entre las dos secciones que componen la encuesta, de la siguiente manera: 4 cuestiones demográficas y 8 cuestiones para la concienciación de la ciberseguridad (Enlace a la encuesta: <https://forms.gle/4dfGtc6YoF9xh5op8>).

Las 4 primeras cuestiones tienen como objetivo poder realizar un estudio de la muestra, pudiendo detectar sesgos o incluso similitudes entre las respuestas a través del género, edad, nivel de estudios que presenta el usuario y rama a la que se asocia el mismo, ya que la anonimidad de la encuesta es un factor clave. Adicionalmente, en la presentación al cuestionario se presenta un descargo de responsabilidad que el usuario debe aceptar para acceder a las cuestiones, reflejando que entiende y acepta el carácter anónimo, y que el objetivo de sus respuestas tiene un fin meramente informativo. Las cuestiones finales presentadas en la sección demográfica pueden observarse en la *Tabla 1*. Por otro lado, la sección de la concienciación de la ciberseguridad en salud contiene las 8 cuestiones específicas para observar la percepción de los usuarios, (ver *Tabla 2*), junto con la Escala de Likert. La primera

cuestión abarca el tema de la **privacidad** de los datos médicos, la segunda cuestión indica una comparativa sobre el **contenido** de los datos médicos y datos bancarios. En el caso de la tercera pregunta, se introduce el empleo de dispositivos *Healthcare IoT* en la vida de los usuarios y la **facilidad de su uso**. Para la cuarta y quinta cuestión se sigue la línea de la tercera pregunta, pero en esta ocasión sobre los datos que capturan los dispositivos y la **confidencialidad** y **soberanía** de los mismos. La sexta pregunta hace reflexionar al usuario acerca de la **transmisión** de la información del dispositivo HIoT, frente a la información contenida por su historial médico. Para la séptima pregunta, se plantea al usuario el dilema de las **notificaciones telemáticas** sobre citas o resultados de salud y la confianza que le suscitan. Por último, la octava pregunta expone al usuario el caso del **robo** de su información médica, debido a un ataque al sistema de salud. Igualmente, las dimensiones positiva y negativa (1 y 5 o 5 y 1, dependiendo de la formulación) de la Escala de Likert también se determinan en la *Tabla 2*, habiendo sido también quintaesenciadas junto a los expertos.

### 3.2. Análisis de las respuestas

En total la encuesta ha recibido 112 respuestas. En primer lugar, observando las características de la muestra, se ha obtenido que el porcentaje entre géneros que han respondido es equitativo (50,9% masculino y 49,1% femenino). Si se analizan las edades de los usuarios destacan 3 rangos de edad de 18-25 años (35,7%), de 25-35 años (29,5%) y de 35-45 (19,6%). Sin embargo, todas las franjas de edad establecidas presentan al menos un representante. En cuanto al nivel de estudios, el 45,5% presentan estudios de máster, seguido de grado universitario con el 25%, y el doctorado con un 11%. Aunque, al igual que en el caso anterior, se presenta al menos un usuario de cada nivel de estudios definido. Además, en cuanto a las ramas de estudio, la gran mayoría de usuarios, el 62,5% pertenecen al área de Ingeniería y Arquitectura, seguido de usuarios especializados en Ciencias de la Salud con un 17,9%. La predominancia de estas dos áreas se debe a que un grupo de investigación de una universidad y el departamento de oncología de un hospital público madrileño, han sido objetivo principal del proceso de difusión. En segundo lugar, en cuanto a las cuestiones de concienciación en salud, para observar y comparar las elecciones de la escala entra las cuestiones se genera la gráfica que se observa en la *Figura 2*. Las 8 cuestiones están reflejadas desde la parte inferior de la gráfica en orden ascendente, reflejando el número de usuarios que han escogido cada valor de la escala. Igualmente, para obtener consonancia en la representación de las respuestas, se ha normalizado invirtiendo aquellas que se encontraban planteadas de manera inversa. De los resultados obtenidos sobre las distintas preguntas, hay que resaltar la privacidad de los datos médicos (cuestión 1), donde los valores escogidos por las dos terceras partes, han sido el 4 y el 5 indicando su alto percepción privada. En paralelo, la pregunta 2, aunque las respuestas están algo más repartidas que en el caso anterior, el 50% de los usuarios consideran (valor 5) que los datos médicos presentan información de iguales características que los datos bancarios, y un cruce revelaría información sensible.

**Concienciación de la Ciberseguridad en Salud**

Cuestiones	Opciones							
1. ¿Con qué genero se identifica?	Femenino		Masculino				Otro	
2. ¿En qué franja de edad se encuentra?	18-25	25-35	35-45	45-55	55-65	< 65		
3. ¿Cuál es su nivel de estudios?	Sin estudios	Graduado escolar	Enseñanza Obl.	Bach. / COU	Ciclo formativo	Grado Univ.	Máster	Doctorado
4. ¿A qué rama de estudios?	Sin rama		Arte y humanidades	Ciencias de la Salud		Ciencias Sociales y Jurídicas		Ingeniería y Arquitectura

Tabla 1. Cuestiones resultantes en la sección demográfica. Obl= Obligatoria, Bach= Bachillerato y Univ.= Universitario.

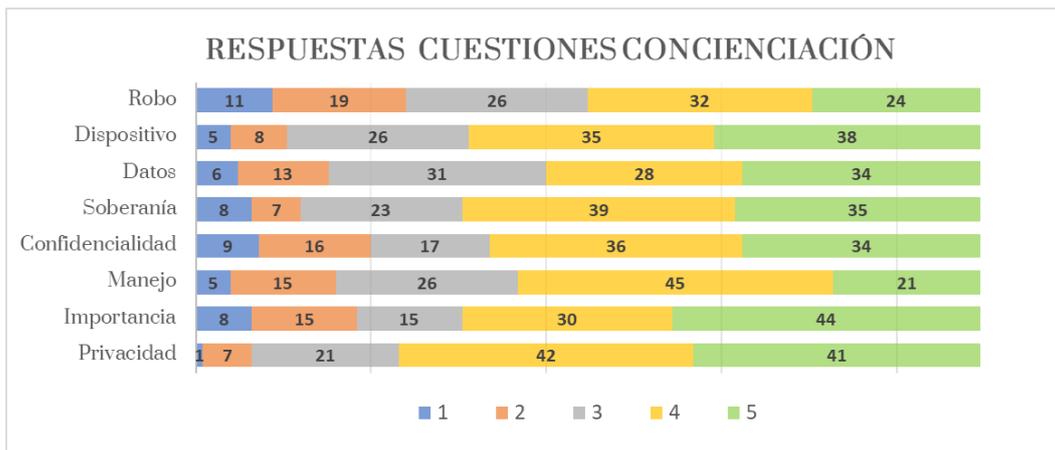
**Concienciación de la Ciberseguridad en Salud**

Cuestiones	1	5
1. ¿Considera que sus datos médicos son de carácter privado?	Todo el mundo debería poder acceder a ellos	Son de conocimiento exclusivo por el propietario de los datos
2. ¿Considera que sus datos médicos contienen información de igual importancia que sus datos bancarios?	Los datos bancarios y los datos médicos no guardan ninguna relación	La obtención de uno de ellos podría conllevar la exposición de los otros
3. Teniendo en cuenta desde la instalación de aplicaciones y uso del dispositivo, hasta la toma de medidas y visualización de los resultados ¿Considera sencillo el manejo de dispositivos inteligentes que permiten la medida de parámetros vitales?	El proceso requiere a una persona con conocimientos técnicos amplios	El proceso es fácil y no provoca ninguna confusión
4. Sabiendo que los datos captados por el dispositivo inteligente se almacenan en una nube del fabricante del equipo ¿Cómo de confidenciales considera sus datos almacenados?	Los datos solo deben ser revelados a su propietario	Los datos pueden ser conocidos tanto por el fabricante, propietario y cualquier otra persona que quiera acceder
5. ¿Está de acuerdo en que el fabricante puede tomar alguna decisión sobre los datos, debido al consentimiento que se otorga al aceptar el uso del dispositivo?	El uso de los datos es exclusivo del usuario	El fabricante al almacenarlos en su sistema tiene el derecho de emplearlos además del usuario
6. ¿Considera que los datos que se transmiten desde su casa a través de un dispositivo inteligente	La información entre ambos no presenta ninguna relación	En el caso de interceptar el envío, un individuo podría obtener

contienen la misma información que su historial médico?		información tanto personal como médica de la persona
7. ¿Cómo de seguro se siente a la hora de recibir citas o resultados médicos a través de su dispositivo móvil?	No existe diferencia entre la forma tradicional y la telemática	La forma telemática no me transmite nada de confianza, y no suelo creerme nada notificado por ella
8. ¿Cómo de peligroso calificaría el robo o filtración de sus datos médicos por un ataque/hackeo al sistema de salud?	No afectaría en nada a mi persona	Podría suponer la pérdida de mi trabajo, extorsión o pago de una cuantía económica

Tabla 2. Cuestiones resultantes en la sección de concienciación.

Para el caso de la facilidad de uso de los dispositivos HIoT (cuestión 3), la tendencia cambia habiendo sido escogidos en mayor proporción los valores centrales 3 y 4 reflejando la no trivialidad de su manejo. Las preguntas 4 y 5 presentan similitud entre los valores de respuesta seleccionados. Ambas cuestiones conciernen el tema de los datos generados y el fabricante del dispositivo inteligente. Los usuarios han escogido las opciones 1 y 2 de manera más habitual que el resto de las opciones, indicando, por un lado; que no consideran que sus datos sean confidenciales bajo el almacenamiento del fabricante, y por otro; que la soberanía de estos debería ser exclusiva de quien lo genera independientemente de la propiedad de la Nube (valores 4 y 5 con la normalización). Seguidamente, se pregunta al usuario (cuestión 6) a cerca de la información transmitida desde un dispositivo inteligente de salud en comparación con los datos que incluye su historial, los valores más predominantes van del 3 al 5, indicando que se consideren un valor similar o igual. La pregunta 7, asocia las notificaciones de información en salud con la validez y seguridad en la transmisión de estas. La tendencia en este caso vuelve a invertirse siendo los valores 1 y 2 los más escogidos (valores 4 y 5 con la normalización de las respuestas), indicando que ambas formas son similares y que se sienten seguros a la hora de ser comunicados a través de la vía telemática. Finalmente, la cuestión 8 hace valorar al usuario la posibilidad que sus datos médicos sean revelados debido a un ataque al sistema de salud. En este caso, el valor más escogido es el 4 donde el usuario considera que supondría unas consecuencias graves para él, pero no críticas.



**Figura 2.** Gráfico de barras apiladas normalizado con las 112 respuestas a las cuestiones según la Escala de Likert. De arriba a abajo las cuestiones 1 a 8 valoradas.

#### 4. Conclusiones

En la actualidad, el sector de la Salud presenta un amplio abanico de dispositivos médicos conectados a las redes de Internet. Sin embargo, la mayoría de estos equipos son obsoletos con sistemas operativos desactualizados, presentando riesgos y vulnerabilidades que pueden llegar a afectar al paciente. El principal objetivo de este artículo es componer un formulario para conocer la percepción de los usuarios acerca de la ciberseguridad en los ambientes de salud, específicamente con el empleo de dispositivos HIoT.

La necesidad del estudio y análisis era evidente, desde el punto de vista de revolución tecnológica surgida en los últimos años en salud, y la falta de incorporación en el diseño y definición de la seguridad en los dispositivos médicos y entornos intra o extrahospitalarios donde se reflejan las carencias.

Por ello, la encuesta para la valoración de la experiencia del usuario ha permitido determinar; por un lado, que el estudio demográfico de la muestra no presenta sesgo de género, ya que el número de mujeres y hombres es equitativo; aunque sí que existe predominancia de los usuarios comprendidos entre los 18 y los 35 años que podría deberse al método de difusión. También predominan los estudios universitarios asociados a la rama de Ingenierías y Arquitectura. Por otro lado, las cuestiones valoradas, para la concienciación de la ciberseguridad en salud, han reflejado que existe una percepción alta, de entorno el 70% de los usuarios, de conocimiento acerca de la vulnerabilidad de los datos médicos y el riesgo que puede suponer su robo o filtración. Igualmente, los resultados del formulario (alrededor del 80%) reflejan que los usuarios están abiertos a nuevas soluciones que aporten un mayor grado de seguridad, ya que la información que se transmite desde los dispositivos inteligentes de salud desde sus casas es similar a la que se contiene dentro de su historial médico. Donde, unido a la hipótesis de que los datos que se almacenan en las nubes de los fabricantes de los dispositivos no son confidenciales escogido por un 65% de los usuarios, la refuerzan. No obstante, aunque la encuesta se ha ejecutado sobre un primer grupo de usuarios, los resultados obtenidos reflejan las tendencias de concienciación sobre la población en esta materia.

#### Agradecimientos

Agradecer al grupo de investigación *Life Supporting Technologies*, por la oportunidad de combinar mis dos áreas de estudio en esta investigación.

#### Referencias

- [1] M. Haghi Kashani, M. Madanipour, M. Nikravan, P. Asghari, and E. Mahdipour, "A systematic review of IoT in healthcare: Applications, techniques, and trends," *Journal of Network and Computer Applications*, vol. 192, p. 103164, Oct. 2021, doi: 10.1016/J.JNCA.2021.103164.
- [2] B. Farahani, F. Firouzi, and K. Chakrabarty, "Healthcare IoT," *Intelligent Internet of Things*, pp. 515–545, 2020, doi: 10.1007/978-3-030-30367-9\_11.
- [3] F. Nausheen and S. H. Begum, "Healthcare IoT: Benefits, vulnerabilities and solutions," *Proceedings of the 2nd International Conference on Inventive Systems and Control, ICISC 2018*, pp. 517–522, Jun. 2018, doi: 10.1109/ICISC.2018.8399126.
- [4] R. de Michele and M. Furini, "IoT Healthcare: Benefits, Issues and Challenges," 2019, doi: 10.1145/3342428.
- [5] P. A. H. Williams and V. McCauley, "Always connected: The security challenges of the healthcare Internet of Things," *2016 IEEE 3rd World Forum on Internet of Things, WF-IoT 2016*, pp. 30–35, Feb. 2017, doi: 10.1109/WF-IOT.2016.7845455.
- [6] S. Nasiri, F. Sadoughi, M. H. Tadayon, and A. Dehnad, "Security Requirements of Internet of Things-Based Healthcare System: a Survey Study," *Acta Informatica Medica*, vol. 27, no. 4, p. 253, 2019, doi: 10.5455/AIM.2019.27.253-258.
- [7] F. Gioulekas *et al.*, "A Cybersecurity Culture Survey Targeting Healthcare Critical Infrastructures.," *Healthcare (Basel)*, vol. 10, no. 2, Feb. 2022, doi: 10.3390/healthcare10020327.
- [8] L. Ángela Guadalupe Canto de Gante, W. Elim Sosa González, J. Bautista Ortega, I. Judith Escobar Castillo, and A. Santillán Fernández, "Escala de Likert: Una alternativa para elaborar e interpretar un instrumento de percepción social," *Alta Tecnología y Sociedad*, vol. 38, no. 1, 2020.

# Herramienta de valoración de aplicaciones móviles de eSalud: programadores, pacientes y profesionales de salud

L. Franco Terriza<sup>1</sup>, M. J. Esteban Gata<sup>1</sup>, M. Elena<sup>1</sup>

<sup>1</sup> Departamento de Ingeniería Electrónica, Universidad de Sevilla, España, marelen@us.es

## Resumen

*El mercado de aplicaciones móviles de salud ha crecido considerablemente en los últimos años, buscando soluciones personalizadas y centradas en las necesidades de los pacientes. Sin embargo, en la actualidad existen pocas herramientas de ayuda al diagnóstico que faciliten elegir la mejor opción disponible. Para solucionar este problema vamos a desarrollar una herramienta que integre tres puntos de vista: programadores, pacientes y profesionales de la salud. La metodología seguida para desarrollar la herramienta consiste en la elaboración de un cuestionario para cada uno de los grupos, generación de ejecutables y diseño de un cuadro de mando interactivo, más una valoración parcial mediante la selección de algunas aplicaciones ya existentes. Se analizarán los resultados obtenidos del uso de la herramienta para demostrar su potencial. En la discusión se expone la importancia y utilidad de valorar las aplicaciones con una herramienta que cubra los tres puntos de vista, así como limitaciones y margen de mejora.*

## 1. Introducción

Debido a los avances en medicina y al aumento del volumen de pacientes que presentan alguna patología en las últimas décadas, ha sido necesario apoyarse en la tecnología para proporcionar soluciones que se adapten tanto a las necesidades de los pacientes como de los profesionales de salud. Así surge el concepto de mHealth o salud móvil, consistente en el uso de las aplicaciones médicas de salud como herramientas de diagnóstico, prevención y tratamiento de patologías [1]. Entre otras ventajas permiten que el profesional realice un seguimiento médico de un gran volumen de pacientes sin necesidad de consultas diarias presenciales, o que los propios pacientes puedan recibir una atención médica personalizada desde su lugar de residencia.

Sin embargo, las aplicaciones de salud móvil no siempre aseguran que el paciente se adhiera a su tratamiento o lo haga de forma correcta. Algunos de los problemas que pueden darse es que la aplicación no cubra las necesidades de su público objetivo, no sea clara o incluso que resulte poco llamativa, lo que puede conllevar a que el paciente abandone el tratamiento. A veces también puede suceder que el paciente o profesional de salud no elija la aplicación que mejor cubra sus necesidades entre las que estén disponibles en el mercado, ya que desconocen qué criterio seguir para realizar esta elección. Como solución, se recurre a diferentes parámetros y escalas que dan la posibilidad a pacientes y profesionales de valorar qué aplicaciones son las que mejor se ajustan a sus necesidades, al mismo tiempo que también resultan útiles para que los

desarrolladores valoren qué aspectos deben abordar para cubrir las necesidades de los usuarios.

Los métodos de validación de aplicaciones móviles más conocidos dentro del entorno español son, con diferencia, la adaptación a nuestro idioma del MARS [1] y el iSYScore [2]. Más allá de estos dos no se han encontrado métodos validados y reconocidos que puedan utilizarse en la actualidad para la valoración de aplicaciones móviles de salud. Sin embargo, ambos cuestionarios están enfocados a un grupo de usuarios específico, como pueden ser los propios usuarios o los desarrolladores de las aplicaciones. Esto conlleva que la evaluación de aplicaciones solo va a resultar útil desde el punto de vista de un perfil determinado.

Debido a lo anterior, se plantea la necesidad de crear una única herramienta que integre los puntos de vista de todos los perfiles involucrados en el proceso de selección de una aplicación móvil de salud: el paciente, que valorará su experiencia con la aplicación; el profesional de salud, que valorará cómo influye la aplicación en el tratamiento de sus pacientes; y el programador, que valorará en qué grado su implementación es capaz de satisfacer las necesidades de los otros dos grupos. Además, es importante que los cuestionarios tengan una extensión adecuada y los resultados sean fácilmente interpretables para facilitar el uso de la herramienta. Por este motivo, se implementará un cuadro de mando interactivo para la representación de los resultados recogidos.

### 1.1. Antecedentes

La primera versión de la herramienta consistía en un único cuestionario dirigido a los programadores, de ahora en adelante desarrolladores de aplicaciones móviles, entendiendo como desarrollador a cualquier persona con conocimiento técnico del campo de aplicaciones móviles de salud, como los miembros de las instituciones responsables de las mismas. La elaboración del cuestionario tuvo en cuenta la opinión de los usuarios finales, los pacientes, para conseguir así una herramienta centrada en el usuario. Por medio de un cuestionario preliminar valoraron diez aspectos relacionados con las aplicaciones móviles de salud. Cada uno de estos aspectos conformaban una categoría, a las que se les asignó un número de preguntas proporcional a la puntuación obtenida, definiendo así las 22 preguntas [4] que conformaban el cuestionario. Posteriormente, a través de Python® se creó una aplicación capaz de almacenar las respuestas para representarlas en un cuadro de mando

creado con Business Intelligence. Finalmente, se llevó a cabo una primera validación usando seis aplicaciones móviles de salud, obteniendo resultados visuales fácilmente interpretables.

## 1.2. Objetivo

El objetivo de esta segunda versión de la herramienta es ampliar la funcionalidad de la anterior mediante el desarrollo de dos cuestionarios adicionales: el cuestionario de pacientes y el cuestionario de profesionales de salud. Al igual que para el cuestionario de desarrolladores, se van a obtener los ejecutables correspondientes y los resultados se van a representar de forma visual y conjunta de forma gráfica, mediante la herramienta de Business Intelligence que mejor se aplique para este propósito.

Una vez que la herramienta esté disponible para ser utilizada, vamos a comprobar su utilidad mediante una valoración parcial de algunas aplicaciones disponibles en el mercado. Además de valorar la utilidad de la herramienta, también se extraerán conclusiones sobre qué características debería presentar una aplicación en fase de desarrollo para que consiga destacar frente a otras soluciones existentes.

## 2. Metodología

La herramienta ha sido desarrollada en cuatro etapas, que se detallarán a continuación:

- Definición de las preguntas que conforman los cuestionarios.
- Implementación de los cuestionarios para su uso y distribución.
- Desarrollo de un cuadro de mando interactivo para la representación visual de los resultados.
- Aplicación de la herramienta a un campo de estudio.

### 2.1. Definición de cuestionarios

La extracción de las preguntas enfocadas a pacientes y a profesionales de salud se ha llevado a cabo mediante una búsqueda bibliográfica. En ella se han revisado los aspectos y parámetros relativos al desarrollo de una aplicación móvil, que pueden ser definidos en términos de usabilidad, calidad, objetividad o subjetividad, entre otras [5-8].

Como resultado de la revisión llevada a cabo, se ha conseguido plantear un listado de 50 preguntas que evalúan diferentes aspectos de las aplicaciones móviles de salud. Estas preguntas se han diferenciado, en primer lugar, en función de si respondían a evaluar la opinión de un paciente, de un profesional de salud o de ambos. En segundo lugar, se han clasificado en las mismas diez categorías que el cuestionario de desarrolladores previamente validado. Finalmente se ha realizado una última selección entre esta lista de preguntas, eliminando o combinando las preguntas redundantes, priorizando las más importantes y asegurando que todas las categorías tuviesen asignada, al menos, una pregunta. También se ha revisado el lenguaje para asegurar que todas las preguntas sean entendibles para el público al que van dirigidas.

El cuestionario final para los pacientes consta de 18 preguntas mientras el de los profesionales de salud consta de 15. En la Figura 2 se ha representado el peso que cada categoría tiene en cada cuestionario.

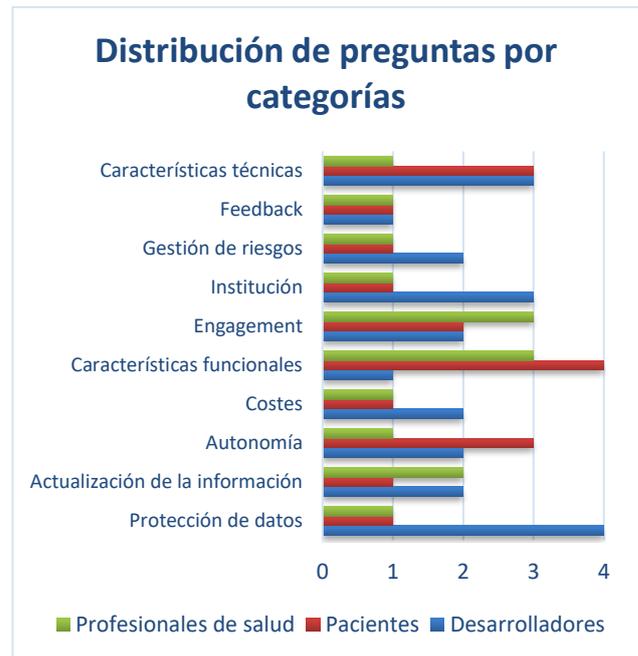


Figura 1. Distribución de preguntas por categorías para cada una de las tres escalas.

### 2.2. Implementación de los cuestionarios

La implementación de los dos nuevos cuestionarios se ha realizado con Python®, generando scripts similares al primer cuestionario de desarrolladores incluyendo algunas actualizaciones, como pantalla de instrucciones y de finalización, entrada de texto de nombres de usuario y aplicación, y otros aspectos relativos a la visualización de las ventanas de preguntas. En total, se han generado dos ejecutables de 18 y 15 preguntas puntuables entre 1 y 5 puntos, siguiendo la escala de Likert. Estos dos cuestionarios complementan el primer ejecutable de 22 preguntas dirigido a desarrolladores.

Cuando un usuario responde a alguno de los tres cuestionarios sus respuestas se almacenan en un Excel generado de forma automática, el cual contiene los siguientes campos:

- Tipo de cuestionario (desarrollador, paciente o profesional de salud)
- Iniciales del nombre del encuestado
- Nombre de la aplicación
- Campo de la aplicación
- Puntuación de cada pregunta
- Categoría de cada pregunta

### 2.3. Cuadro de mando

El cuadro de mando conforma la parte visual e interactiva de la herramienta que facilitará la toma de decisiones a través de la representación de la información, que en este caso serían las respuestas a los cuestionarios. Para implementarlo se puede hacer uso de diversas herramientas

de Business Intelligence, entre las cuales hemos elegido Google Data Studio ® por ser online y gratuita.

Las gráficas y funciones que se estiman oportunas para evaluar los resultados obtenidos de los cuestionarios son las siguientes:

- Gráfica de puntuación media de cada aplicación valorada, en %
- Gráfica de puntuación media de cada aplicación desglosadas por categorías, en %
- Gráfica de puntuación media de cada aplicación distribuida por ramas, según el campo
- Gráfica de puntuación media de cada aplicación distribuida por ramas, según el tipo de cuestionario
- Botones de filtrado por tipo de cuestionario, nombre de la aplicación, campo y categoría

### 2.4. Aplicación de la herramienta

Completadas las fases anteriores, necesitamos disponer de las valoraciones de algunas aplicaciones para poder realizar la valoración visual, y así comprobar y validar el uso de la herramienta. Vamos a abordar esta validación desde el punto de vista de su desarrollo, de forma que, mediante la selección de varias aplicaciones destacadas en el mercado, no necesariamente de la misma área, podamos observar en qué categorías destaca cada una de ellas. Así, podemos conocer a qué aspectos prestar especial atención en el desarrollo de una aplicación nueva.

Hemos supuesto que los desarrolladores están trabajando en una aplicación de salud móvil para pacientes con patologías respiratorias. Las funcionalidades principales de esta aplicación son los ejercicios personalizados, el seguimiento de la evolución del paciente, cuestionarios de salud e integración de dispositivos externos. En base a esto, se ha llevado a cabo una búsqueda de aplicaciones que dispongan de alguna de dichas funcionalidades, tanto del mismo para la misma patología como para otras. Otros criterios de selección han sido que las aplicaciones sean gratuitas, accesibles para el público general, destinadas a los pacientes y disponibles tanto en la App Store como en PlayStore. En la Tabla 1 se muestra la lista de aplicaciones entre las que se ha realizado la selección final.

Aplicación	Patología	a	b	c	d
EPOCuidate	EPOC	x	x	x	
GOLD EPOC	EPOC			x	
Paseos EPOC	EPOC	x	x		
Ser+ contra COVID19	COVID			x	
Hailie	EPOC		x		x
Social Diabetes	Diabetes		x	x	x
MediSafe	Medicación		x		
Fissios	Fisioterapia	x			
MySugr	Diabetes	x	x		x
YoTeCuido	Alzhéimer	x			

Tabla 1. Selección de aplicaciones. a) Ejercicios/retos, b) Seguimiento, c) Cuestionarios, d) Dispositivos.

La valoración parcial ha sido llevada a cabo por las tres investigadoras desde el punto de vista del desarrollo, utilizando las cuatro aplicaciones seleccionadas: Hailie, Social Diabetes, Fissios y MySugr.

### 3. Resultados

Los resultados que se presentan son las gráficas obtenidas de la aplicación de la herramienta a una selección de aplicaciones móviles. Por un lado, esto sirve para mostrar el resultado final de la herramienta. Por otro, se va a comprobar su funcionamiento y utilidad a la hora de interpretar los datos y tomar decisiones.

Tras la valoración parcial de las aplicaciones móviles seleccionadas, se han obtenido un 81,82% para Social Diabetes, posicionada en primer lugar. Le siguen Hailie con un 80,91%, Fissios con un 79,09% y MySugr con un 78,18%. En la Figura 2 se han representado estos porcentajes de forma visual, mientras en la Figura 3 se muestra el desglose por categorías.

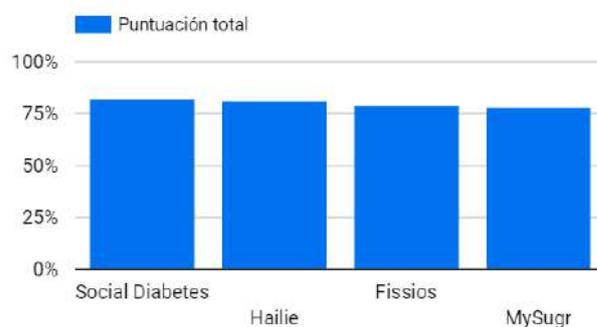


Figura 2. Ampliación de la segunda gráfica: gráfico de barras de los porcentajes obtenidos por cada aplicación.

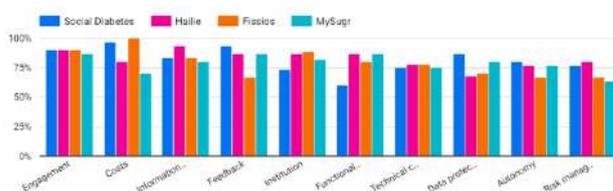


Figura 3. Ampliación de la tercera gráfica: gráfico de barras de los porcentajes obtenidos por cada categoría y aplicación.

La distribución de los puntos por categorías varía según la aplicación, aunque destacan categorías como “Engagement” o “Actualización de la información”, en las que todas las aplicaciones han superado el 75% de la puntuación. Por el lado contrario, en categorías como “Protección de datos”, “Autonomía” o “Gestión de riesgos” solo dos de las cuatro aplicaciones consiguen superar ese 75%.

### 4. Discusión

El desarrollo de los dos nuevos cuestionarios supone la posibilidad de poder evaluar cualquier aplicación móvil de salud desde tres puntos de vista (desarrolladores, pacientes y profesionales de salud). La integración de las respuestas, almacenadas por medio de tres scripts, en un único cuadro de mando permite conseguir una valoración visual, interactiva e intuitiva de las mismas en todos los planos, lo

cual podrá facilitar la toma de decisiones y mejorar la usabilidad de las propias aplicaciones.

El cuestionario inicial, de desarrolladores, constaba de preguntas técnicas distribuidas por categorías, respondiendo a las necesidades de los pacientes. Sin embargo, el peso de cada categoría ha cambiado en cada uno de los cuestionarios ya que cada grupo será capaz de aportar más información respecto a un aspecto determinado. Por ejemplo, los programadores sabrán más de aspectos técnicos, los pacientes conocerán la experiencia de uso y los profesionales de salud la calidad del contenido y el impacto de la aplicación en el tratamiento. Se puede observar que los cuestionarios se relacionan entre sí como se observa en la Figura 4, complementando la información aportada por los otros dos.

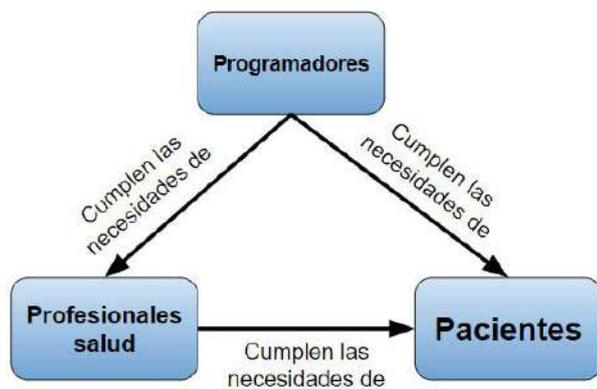


Figura 4. Relación entre los tres cuestionarios.

El cuadro de mando se ha podido implementar en Google Data Studio ® de forma muy sencilla y la visualización y filtrado de datos permite observar claramente qué aplicaciones destacan. A pesar de esto, las dos últimas gráficas no han sido implementables tal como se esperaban debido a las limitaciones de la herramienta de Business Intelligence elegida, y se han sustituido por gráficas de barras con las que se puede realizar una interpretación similar de los datos.

Por otro lado, la valoración de las aplicaciones móviles elegidas ha sido llevada a cabo de forma parcial, ya que solo han participado desarrolladores. En posteriores valoraciones, esto nos sirve para hacer una primera criba y disminuir el número de aplicaciones a considerar. Por tanto, dividir el proceso en partes que ayuden a tomar decisiones parciales puede simplificar el proceso e influir positivamente en los resultados.

El resultado de la valoración parcial muestra que todas las aplicaciones superan el 75% de la puntuación y hay menos de un 4% de diferencia entre la primera y la última aplicación de la lista. En el caso de desarrollar una aplicación nueva, independientemente de la patología que se trate, las puntuaciones obtenidas por categoría pueden resultar orientativas para determinar a qué aspectos de dicha aplicación se les debe prestar especial atención durante su desarrollo.

Aunque la herramienta ya puede ser empleada para valorar aplicaciones móviles, quedan pendientes algunas mejoras

para facilitar su uso. Entre ellas, se implementarán mejoras en el cuadro de mando que permitan sacarle más provecho a los datos, se buscará una forma de poner los cuestionarios a disposición de los usuarios de forma online y se buscará que las respuestas se almacenen automáticamente en una plataforma tipo nube. Como paso final, se realizará una valoración más amplia y completa de aplicaciones integrando los tres puntos de vista: Programadores, Pacientes y Profesionales de la salud.

## Referencias

- [1] Vaghefi I, Tulu B. The Continued Use of Mobile Health Apps: Insights From a Longitudinal Study. *JMIR Mhealth Uhealth* 2019;7(8):e12983
- [2] Stoyanov S. R, Hides L, Kavanagh D. J, et al. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. *JMIR Mhealth and Uhealth*. 2015 Mar;3(1):e27
- [3] Grau I, et al. Método de valoración de aplicaciones móviles de salud en español: el índice iSYScore. *SEMERGEN-Medicina de Familia*, 2016, vol. 42, no 8, p. 575-583.
- [4] Página web Distintivo AppSaludable, Agencia de Calidad Sanitaria de Andalucía <http://www.calidadappsalud.com/listado-completo-recomendaciones-app-salud/> (Consultada: Agosto 2022)
- [5] Ramírez L, Guillen E; Cifuentes Y. 37. Estrategia de validación para aplicaciones móviles de salud. *Desarrollo e Innovación en Ingeniería*, 2016, p. 503.
- [6] Paniagua L. A, Bedoya R. D, Mera C. Un método para la evaluación de la accesibilidad y la usabilidad en aplicaciones móviles. *Tecnológicas*. 2020;23:99-117.
- [7] Dolado Martín, C., Berlanga-Fernández, S., Fabrellas i Padrès, N., & Galimany Masclans, J. (2017). Uso de aplicaciones móviles de salud en usuarios de Atención Primaria. *ROL. Revista Española de Enfermería*, 2017, vol. 40, num. 2, p. 16-21.
- [8] González, B. S., Alcazar, V. R., Álvarez-Rementería, J. F., López, V. S., Vázquez, M. V., Usagre, M. H., & Olivera, A. T. (2019). Uso y valoración de aplicaciones móviles de salud por los profesionales sanitarios de Andalucía para la mejora de la continuidad asistencial y la comunicación. *International Journal of Integrated Care*, 19(4).

# Entrenamiento de la memoria declarativa episódica mediante serious games en personas mayores con deterioro cognitivo leve

J.F. Ortega Morán<sup>1</sup>, J.B. Pagador<sup>1</sup>, A. Romero García<sup>2</sup>, V. Gilete Preciado<sup>3</sup>, M. Guerrero Pérez<sup>2</sup>, L.F. Sánchez-Peralta<sup>1</sup>, F.M. Sánchez Margallo<sup>1</sup>

<sup>1</sup> Centro de Cirugía de Mínima Invasión Jesús Usón, Cáceres, España, {jfortega,jbpagador,lfsanchez,msanchez}@ccmijesususon.com

<sup>2</sup> Servicios del Fresno S.L., Almendralejo, España, {merida,mguerrero}@serviciosdelfresno.es

<sup>3</sup> Neuropsicólogo, Cáceres, España, vgp.nps@gmail.com

## Resumen

*Según envejece la población, aparecen problemas cognitivos en las personas, como el deterioro cognitivo leve. El uso de serious games permite realizar actividades para trabajar funciones cognitivas de una manera útil y entretenida para los mayores. En el proyecto EuroAGE se ha implementado un serious game de hacer la compra para estimular la memoria declarativa episódica, cuya eficacia durante un programa de intervención con personas con deterioro cognitivo leve se ha evaluado en este trabajo. Para ello, 4 personas (grupo de estudio) utilizaron el juego durante 10 semanas, mientras 2 personas (grupo control) realizaron fichas tradicionales de memoria. El progreso de dicha función cognitiva se ha monitorizado en ambos grupos con un test pre-post de memoria verbal intencional declarativa, cuya principal conclusión indica que los resultados de memoria inmediata son mejores en el grupo de estudio, llegando a recordar más palabras y a controlar la evolución de su rendimiento hasta un 50% con respecto al grupo control, gracias a un sistema de pistas del juego que ejercita la memoria facilitada. Los aspectos del juego relacionados con la estimulación de la memoria diferida son adecuados, puesto que el grupo de estudio ha obtenido mejores resultados de rendimiento. Las métricas facilitadas por el serious game de tiempo, aciertos, fallos e intentos mejoraban con las semanas, sugiriendo que el juego no supone una barrera tecnológica para los mayores y resulta de interés para su uso. Como trabajo futuro, la muestra tan reducida de este estudio hace necesario corroborar los resultados con más participantes.*

## 1. Introducción

Debido al envejecimiento de la población y al aumento de la esperanza de vida, un número cada vez mayor de individuos tienen algún grado de deterioro cognitivo. La demencia es un síndrome que puede estar causado por una gran variedad de enfermedades, que incluirían prácticamente todas las lesiones del cerebro de causa degenerativa, tóxica, metabólica, traumática, infecciosa, tumoral o vascular [1]. Este síndrome adquirido afecta en la actualidad a unas 640.000 personas en España y alrededor de 46 millones de personas en todo el mundo [2]. El Deterioro Cognitivo Leve (DCL) es una fase de transición entre un estado cognitivamente sano y la demencia [3]. Aunque los tratamientos suelen ser farmacológicos, son también muy importantes los tratamientos no farmacológicos, como entrenamiento y rehabilitación cognitiva en Actividades de la Vida Diaria (AVD) [4]. En este sentido, las herramientas TICs, y más

en concreto los serious games, son especialmente útiles, ya que son juegos cuyo objetivo principal es dar un valor terapéutico y de evaluación a la acción lúdica de los juegos, coincidiendo con los objetivos trabajados en la intervención de pacientes con demencia [5].

La memoria declarativa episódica es la función neurocognitiva que se ve afectada más tempranamente en la enfermedad de Alzheimer, por lo que la estimulación de la memoria declarativa episódica en personas con DCL afectadas por el deterioro mnésico es muy importante. En el proyecto EuroAGE se ha desarrollado un serious game basado en la AVD de hacer la compra en un hipermercado con el objetivo de estimular la memoria declarativa intencional verbal a nivel de los tres procesos mnésicos de codificación o fijación, consolidación o almacenamiento, y recuerdo o recuperación de la información [6]. El juego consta por tanto de tres fases: 1) fase de aprendizaje, donde el usuario debe memorizar productos de una lista de la compra; 2) fase de interferencia, donde el usuario debe preparar el dinero en el monedero para pagar la compra; y 3) fase de recuerdo, donde al usuario se le presenta un mapa de un hipermercado y tiene que rellenar una cesta de la compra con los productos de la lista de la compra, entre productos distractores.

Aunque el juego fue validado por usuarios finales como intuitivo, entretenido y útil para su aplicación en la vida diaria, es necesario evaluar los efectos sobre la función cognitiva de la memoria declarativa episódica de las personas mayores con DCL mediante una intervención durante un período de tiempo para probar la eficacia de este serious game, siendo éste el objetivo del trabajo.

## 2. Métodos

El criterio de inclusión de los participantes en el programa de intervención fue de personas con Escala de Deterioro Global (GDS) [7] de valor 3, que se corresponde con el DCL. Por lo tanto, previo al comienzo del estudio, los participantes pasaron el test MEC de Lobo [8] para valorar su estado cognitivo general.

Los participantes se dividieron en un grupo de estudio y un grupo control, consistiendo el protocolo de intervención en la realización de 2 sesiones por semana con una duración de 30 minutos por sesión. Durante cada sesión, los participantes del grupo de estudio jugaron al

serious game con la tablet, mientras que los participantes del grupo control completaron una ficha con los ejercicios tradicionales de memoria que normalmente se utilizan para deterioro cognitivo, y que han sido extraídas del manual de Tárraga y Rovira [9]. Los participantes del grupo de estudio realizaron antes de empezar el estudio unas pruebas de entrenamiento para familiarizarse con el uso de la tablet y el juego.

Antes y después del periodo de la intervención, los participantes pasaron un test pre- y post-estudio respectivamente con el fin de evaluar el progreso del deterioro cognitivo. Este test es el denominado ENM.dem (Exploración Neuropsicológica Mínima) [10], centrándonos en el test de memoria verbal intencional declarativa. Esta prueba evalúa el proceso básico de aprendizaje en lo referente a la fijación de la memoria inmediata y la consolidación en la memoria diferida. El test consiste en el siguiente protocolo: 1) Se le presentan al participante una serie de 12 palabras o imágenes y tiene que leerlas (o decir qué imágenes son) en voz alta; 2) se realiza una tarea de interferencia consistente en restar de 3 en 3 a partir de 40; 3) se realiza el primer ensayo de memoria inmediata libre (RL1) en el que el participante debe recordar el mayor número de palabras o imágenes que le fueron presentadas; 4) se realiza el primer ensayo de memoria inmediata facilitada (RF1) donde se le da pistas al paciente para que recuerde las palabras o imágenes que no ha recordado espontáneamente; 5) se realiza de nuevo la tarea de interferencia (contar desde el 40 hacia atrás de 3 en 3); 6) se realiza el segundo ensayo, similar al anterior, de memoria inmediata libre (RL2) y facilitada (RF2); 7) finalmente se realiza un último ensayo de memoria diferida libre (RLD) y facilitada (RFD) 20 minutos después de finalizar el ensayo 2, realizando previamente varias pruebas (evocación categorial, Poppelreuter, abstracción, etc.). La puntuación es el número de palabras que el participante ha recordado en cada ensayo. Por otra parte, se analizó la evolución de las métricas (intentos, aciertos, fallos, tiempo,...) que ofrece el serious game con el paso de las semanas.

Este estudio fue realizado con la aprobación de la comisión de bioética y bioseguridad de la Universidad de Extremadura (Nº Registro: 61//2021). Tanto los participantes (o familiares) del grupo de estudio como del grupo control firmaron un consentimiento informado para participar en el estudio en el que se les ha informado sobre el objeto del estudio, el procedimiento y duración, los resultados, los riesgos, los beneficios, los costes, la confidencialidad de los datos y su derecho de participar libremente o no en la investigación, así como de retirar su consentimiento en cualquier momento.

### 3. Resultados

Los participantes que iniciaron el estudio fueron 10 personas (8 mujeres, 2 hombres, edad promedio 82,7 años) en el grupo de estudio y 5 personas (4 mujeres, 1 hombre, edad promedio 79,4 años) en el grupo control, todos ellos con DCL y en situación de dependencia de grado I. Sin embargo, al final de las 10 semanas que duró el estudio, solo 4 personas (3 mujeres, 1 hombre, edad

promedio 84,25 años) del grupo de estudio y 2 personas (1 hombre, 1 mujer, edad promedio 77,5 años) del grupo de control finalizaron el programa de intervención.

#### 3.1. Test ENM pre-post

Los resultados del test ENM pre- y post-estudio se muestran en la Tabla 1.

Test	Grupo estudio		Grupo control	
	Pre	Post	Pre	Post
RL1	5,8	4,5	5,5	4,5
RF1	4,8	4,5	6,0	3,5
RL2	6,0	4,3	5,0	5,0
RF2	5,5	5,0	6,0	5,0
RLD	5,5	4,5	5,0	5,0
RFD	6,0	4,8	6,0	4,0

Tabla 1. Resultados del test ENM pre-post

#### 3.2. Métricas del juego

Este apartado muestra los resultados de las métricas que proporciona el serious game a lo largo de las 10 semanas del protocolo de intervención del grupo de estudio.

En la fase 1 de aprendizaje, el tiempo que emplean los participantes para finalizar la fase va disminuyendo con el paso de las semanas (Figura 1), siendo el tiempo promedio de 265,43 segundos. Las primeras semanas eran todos aciertos (promedio 7,82), sin embargo las últimas semanas hubo algunos fallos (promedio 0,11). Eso sí, los fallos con pistas disminuían con el paso de las semanas (promedio 0,22).

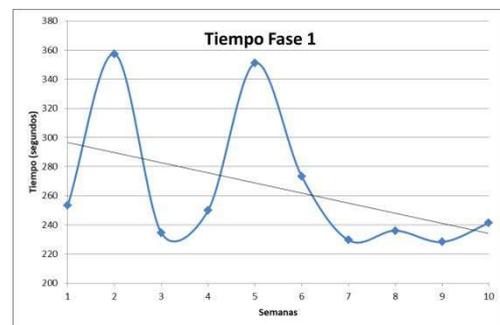


Figura 1. Tiempo promedio (segundos) en la fase 1 del grupo de estudio y línea de tendencia lineal de los datos

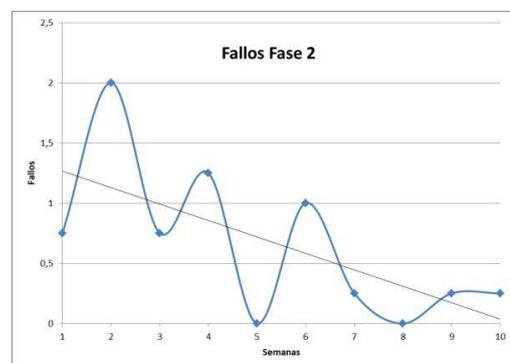
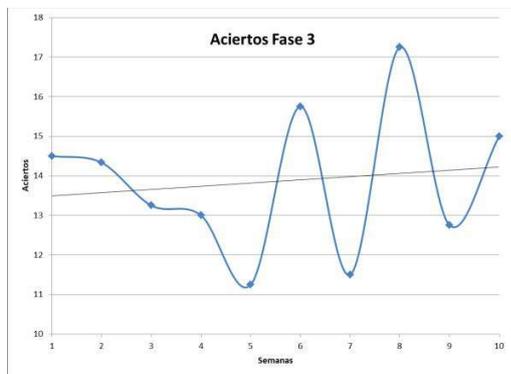


Figura 2. Fallos promedio en la fase 2 del grupo de estudio y línea de tendencia lineal de los datos

En la fase 2 de interferencia, con el paso de las semanas empleaban menos tiempo en hacer la prueba (promedio 144,23 segundos), realizaban cada vez más intentos (promedio 3,29) y cometían cada vez menos fallos (promedio 0,65) (Figura 2).

En la fase 3 de recuerdo, el número de aciertos aumentaba (Figura 3) (promedio 13,86) y el número de fallos disminuía (promedio 2,01) con el paso de las semanas. Los intentos (promedio 1,98) y el tiempo empleado (promedio 165,89 segundos) también aumentaban.



**Figura 3.** Aciertos promedio en la fase 3 del grupo de estudio y línea de tendencia lineal de los datos

En cuando al tiempo (promedio 574,79 segundos) y los fallos (promedio 2,22) del total del serious game, disminuían con las semanas.

#### 4. Discusión y conclusiones

El objetivo de este trabajo era evaluar los efectos sobre la función cognitiva de la memoria declarativa episódica de las personas mayores con DCL mediante una intervención durante un periodo de tiempo con el serious game desarrollado. En ese sentido, tras analizar los resultados del test ENM pre- y post-estudio, la evolución tanto del grupo control como del grupo de estudio es muy similar en cuanto a la memoria inmediata y la memoria diferida. La puntuación en todos los casos y grupos es menor en el post- que en el pre-, que es lo normal por el deterioro cognitivo que sufren las personas mayores con el tiempo. Al igual que en estudios previos [11], se produce una disminución de la capacidad de la memoria episódica verbal en los índices de recuerdo y reconocimiento de listas de palabras inmediatas y retardadas evaluados. La tasa de disminución de la capacidad de memoria episódica verbal no es estática, y aumenta con el avance de la edad [11], por eso la evolución del grupo control disminuye a menor ritmo que el grupo de estudio puesto que tienen una edad promedio de casi 7 años menos.

Sin embargo, cabe destacar que los resultados de memoria inmediata son mejores en el grupo de estudio que en el grupo control, llegando a recordar más palabras (75% vs 66,67% en el ensayo 1; 76,25% vs 75% en el total). A ello ha contribuido principalmente el test de memoria facilitada, puesto que ambos grupos consiguieron recordar entre 36-40% de las palabras en los distintos ensayos de memoria libre. Esto puede ser debido a que en el serious game, durante la primera fase de aprendizaje, hay una sub-fase en la que una vez organizada la lista de la

compra, el usuario debe señalar el producto de cada categoría que cumple una característica específica que lo diferencia del resto de los productos de la lista de la misma categoría. Como dicha sub-fase de refuerzo no se realiza en el grupo control, es posible que el grupo de estudio haya conseguido recordar más palabras tras el test de memoria facilitada, ya que ha ejercitado la fijación de la información mediante la utilización de pistas jugando al serious game, y eso permite optimizar el recuerdo incrementando la retención y transferencia [12]. De hecho, si nos fijamos en las métricas proporcionadas por el serious game, los fallos con pistas disminuían con el paso de las semanas, por lo tanto hace pensar que el sistema de pistas es adecuado.

Por otro lado, cabe destacar que los efectos positivos del serious game en la memoria inmediata también se ven reflejados por la evolución del rendimiento de la misma. En el grupo de estudio, se ha pasado a recordar 10,6 palabras en el pre- a 9 palabras en el post-estudio en el ensayo 1, lo que significa una evolución (negativa por el deterioro cognitivo) del -15,09%. Sin embargo, en el grupo control se ha pasado de recordar 11,5 palabras a 8, lo que significa una evolución del -30,43%. En el total de ensayos de memoria inmediata, la evolución del grupo de estudio es del -17,19%, por el -20% del grupo control. Por lo tanto, nuestro estudio muestra que mediante el programa de intervención realizado con el serious game se ha conseguido controlar la evolución del rendimiento de la memoria inmediata hasta un 50% con respecto a los participantes del programa de intervención tradicional.

En las enfermedades de deterioro cognitivo, antesala de demencia tipo Alzheimer, los problemas de memoria derivan de alteraciones en el almacenamiento de la información, puesto que las personas con estas enfermedades muestran un marcado descenso en el recuerdo diferido en tareas de memoria verbal [13]. La media establecida en el recuerdo diferido libre corresponde a 8 puntos y se considera que existe bajo rendimiento cuando la puntuación total se sitúa por debajo de 5 puntos [14]. Según los resultados, el grupo de estudio consigue una puntuación total de 9,3 puntos, mientras que el grupo control consigue un 9. Por lo tanto, ambos grupos superan ampliamente el umbral de bajo rendimiento, obteniendo el grupo de estudio mejores resultados de rendimiento de la memoria diferida total.

La principal barrera identificada por las personas mayores para no utilizar la tecnología es su falta de interés en ella. El uso de la tecnología aumenta a medida que las personas mayores se dan cuenta de lo útil que puede ser en su calidad de vida [15]. Los profesionales reconocen la existencia de barreras en el uso de las nuevas tecnologías, como son los sistemas de interacción con los dispositivos, debido a las dificultades que plantean las interfaces de usuario en cuanto a menús y tamaños [16]. Sin embargo, en nuestro estudio las métricas proporcionadas por el serious game durante las 10 semanas del programa de intervención muestran que el tiempo para completar las fases disminuía con el paso de las semanas, realizaban cada vez más intentos, y cometían más aciertos y menos fallos. Todo esto hace pensar que el serious game

diseñado causa interés en las personas mayores, superando la barrera tecnológica que les permita usarlo en su día a día de manera útil mejorando su calidad de vida.

La alta tasa de abandonos de participantes (60% en ambos grupos) limita la potencia del estudio. Solo dos de ellos abandonaron por motivos de salud, así que es necesario analizar los factores que influyen en la tasa de abandonos para futuros estudios. Todos los participantes tienen similar estado cognitivo (GDS 3) y nivel cultural o socioeconómico, por lo tanto, la falta de adherencia se puede deber al escaso interés de la persona mayor para participar en el programa, la desmotivación del terapeuta para incentivar la participación del paciente, o una planificación intensiva de las sesiones del protocolo.

Como conclusión, los resultados del presente estudio sugieren que se ha diseñado un serious game que permite controlar la evolución del rendimiento de la memoria declarativa episódica inmediata y mejorar el rendimiento de memoria diferida de las personas mayores con deterioro cognitivo leve tras realizar un programa de intervención. Esto se alinea con la tendencia de que los serious games mejoran el desempeño de la memoria de los participantes [17], y no parece estar influido por la edad, ya que los participantes del grupo de estudio eran mayores que el grupo control y obtienen mejores resultados. Sin embargo, el número tan reducido de participantes sugiere que tomemos los resultados con cautela, al igual que indican estudios similares [18]. Por ello, estos resultados deberían ser corroborados mediante un programa de intervención más largo, al menos 6 meses de duración, y una muestra más amplia de participantes.

## Agradecimientos

El proyecto EUROAGE está co-financiado por el Fondo Europeo de Desarrollo Regional (FEDER), Programa INTERREG V-A España-Portugal (POCTEP) 2014-2020.

GR21201, financiado por la Consejería de Economía, Ciencia y Agenda Digital de la Junta de Extremadura, y cofinanciado por la Unión Europea (FEDER "Una manera de hacer Europa").

## Referencias

[1] Subirana Mirete J, Crusat Basté M, Cullell Gómez N, et al. Capítulo |18| - Demencias y enfermedad de Alzheimer. En: Bruna O, Roig T, Puyuelo M, et al. editores. *Rehabilitación Neuropsicológica*. Elsevier España, 2011, pp 289-e54.

[2] Sousa L, Sequeira C, Ferré-Grau C, et al. Manual del cuidador familiar. (Es)tar con la demencia. Programa de capacitación para cuidadores familiares de personas con demencia que residen en el domicilio. 1th. ed. Tarragona: Publicacions URV, 2019.

[3] Mooldijk SS, Yaqub A, Wolters FJ, et al. Life expectancy with and without dementia in persons with mild cognitive impairment in the community. *Journal of the American Geriatrics Society*, vol 70, sup 2, 2022, pp 481-489.

[4] Jiménez Palomares M. Estudio sobre efectos del programa SADL (Stimulation Activities of Daily Living) de terapia ocupacional sobre el nivel de autonomía en las actividades de la vida diaria de personas con demencia institucionalizados, 2014.

[5] Bennett S, Laver K, Voigt-Radloff S, et al. Occupational therapy for people with dementia and their family carers provided at home: a systematic review and meta-analysis. *BMJ open*, vol 9, sup 11, 2019, pp e026308.

[6] Ortega Morán JF, et al. Serious game for cognitive stimulation of older people with Mild Cognitive Impairment: Design and pilot study. *JMIR Serious Games (submitted)*. doi:10.2196/41437.

[7] Reisberg B, Ferris SH, De Leon MJ, et al. Global deterioration scale (GDS). *Psychopharmacol Bull*, vol 24, sup 4, 1988, pp 661-3.

[8] Lobo A, Escobar V, Ezquerro J, et al. "El Mini-Examen Cognoscitivo" (Un test sencillo, práctico, para detectar alteraciones intelectuales en pacientes psiquiátricos). *Revista de psiquiatría y Psicología Médica*, 1980.

[9] Tàrraga L, Rovira MB. Ejercicios prácticos de estimulación cognitiva para enfermos de Alzheimer en fase leve. Glosa (Eds.), 2004.

[10] Castellanos Pinedo F, Cid Gala M, Duque San Juan P, et al. Plan integral de atención sociosanitaria al deterioro cognitivo en Extremadura (PIDEX). Primera parte: Demencias degenerativas. Consejería de Sanidad y Consumo. Consejería de Bienestar Social. Publicaciones Junta de Extremadura, 2007.

[11] Constantinidou F, Zaganas I, Papastefanakis E, et al. Age-related decline in verbal learning is moderated by demographic factors, working memory capacity, and presence of amnesic mild cognitive impairment. *Journal of the International Neuropsychological Society*, vol 20, sup 8, 2014, pp 822-835.

[12] Ríos, E. Diseño de juego serio para promover la motivación y el aprendizaje de las matemáticas. En "Manuel Prieto, Silvia Pech y Joel Angulo. Tecnología, Innovación y Práctica Educativa", 2020, pp 61.

[13] Ultra-Cucarella J, Pérez-Elvira R, Duque P. Beneficios de la codificación profunda en la enfermedad de Alzheimer. Análisis del rendimiento en una tarea de memoria mediante el Item Specific Deficit Approach. *Neurología*, vol 29, sup 5, 2014, pp 286-293.

[14] Estévez F, Webster F, Piedra MJ. Cerebellum: Not Just Motor Function. About A Case. *Revista Ecuatoriana de Neurología*, vol 29, sup 3, 2020, pp 107-113.

[15] Vaportzis E, Clausen MG, Gow AJ. Older adults perceptions of technology and barriers to interacting with tablet computers: A focus group study. *Frontiers in Psychology*, vol 8, sup OCT, 2017, pp 1–11.

[16] Valenzuela T, Okubo Y, Woodbury A, et al. Adherence to Technology-Based Exercise Programs in Older Adults: A Systematic Review. *J Geriatr Phys Ther.*, vol 41, sup 1, 2018, pp 49-61.

[17] Chalodgeridis A, Tsiatsos T. Implementation and Evaluation of a Serious Game for Working Memory Enhancement. *Applied Sciences*, vol 10, sup 24, 2020, pp 9128.

[18] Abd-Alrazaq A, Alhuwail D, Al-Jafar E, et al. The effectiveness of serious games in improving memory among older adults with cognitive impairment: systematic review and meta-analysis. *JMIR serious games*, vol 10, sup 3, 2022, pp e35202.

# Plataforma mHealth ALIBIRD®: Descripción y resultados preliminares de su evaluación en la asistencia clínica de pacientes oncológicos

J.M. Iniesta-Chamorro<sup>1</sup>, M. Sereno Moyano<sup>2,3</sup>, B. Garrido Rubiales<sup>2,3</sup>, M. Villarino Sanz<sup>2,3</sup>, I. Espinosa-Salinas<sup>4</sup>, J. Moreno-Rubio<sup>2,3</sup>, S. Falagán<sup>2</sup>, M.V. Moreno-Arribas<sup>5</sup>, A. Ramírez<sup>4</sup>, E. Casado<sup>2,3</sup>, M.E. Hernando<sup>1,6</sup>, E.J. Gómez<sup>1,6</sup>, Consorcio ALIBIRD2020-CM

<sup>1</sup> Biomedical Engineering and Telemedicine Centre, ETSIT, Center for Biomedical Technology, Universidad Politécnica de Madrid, Spain, {jm.iniesta, mariaelena.hernando, enriquejavier.gomez}@upm.es

<sup>2</sup> Medical Oncology, Infanta Sofía University Hospital. FIIB HUIS HHEN, Madrid, Spain, {beatriz.garrido.externo, maria.sereno, marta.villarino, sandra.falagan, enrique.casado}@salud.madrid.org

<sup>3</sup> Clinical Oncology Group, IMDEA Food Institute, CEI UAM + CSIC, Madrid, Spain, {juan.moreno}@imdea.org

<sup>4</sup> Molecular Oncology Group, IMDEA Food Institute, CEI UAM + CSIC, Madrid, Spain, {mariaisabel.espinosa, ana.ramirez}@imdea.org

<sup>5</sup> Institute of Food Science Research (CIAL), CSIC-UAM, Campus de Cantoblanco, Madrid, Spain, {victoria.moreno}@csic.es

<sup>6</sup> Centro de Investigación Biomédica en Red de Bioingeniería, Biomateriales y Nanomedicina, CIBER, Madrid, España

## Resumen

*En el marco del consorcio ALIBIRD2020-CM se ha diseñado y desarrollado una plataforma mHealth que mediante el registro de Patient Reported Outcomes (PROs) y el análisis de la nutrigenética y la microbiota, genera recomendaciones personalizadas, fomentando el empoderamiento del paciente y apoyando la toma de decisiones clínicas durante el tratamiento oncológico. El objetivo de este artículo es describir el proceso y los resultados preliminares de la evaluación de la usabilidad y la aceptación de la plataforma mHealth ALIBIRD® que se está llevando a cabo en un estudio clínico con 20 pacientes con cáncer de pulmón avanzado en el Hospital Universitario Infanta Sofía (HUIS). Los resultados preliminares reflejan un alto nivel de puntuación en la usabilidad, la aceptación y la satisfacción, respondiendo a las necesidades de los usuarios (pacientes y profesionales sanitarios). Las métricas de uso e interacciones de los pacientes con la aplicación móvil se relacionan con las dimensiones clínicas y funcionalidades mejor valoradas: control, apoyo y recomendaciones sobre variables del estilo de vida, gestión de síntomas y contacto telemático con el servicio clínico. Los datos registrados y su procesamiento han facilitado a los profesionales realizar diferentes acciones clínico-terapéuticas: detección de progresiones, adelantos de citas, ajustes terapéuticos y derivaciones a especialidades. Los resultados preliminares muestran que la plataforma es una solución de salud digital útil, tanto para el empoderamiento y apoyo del paciente oncológico, como para los profesionales clínicos a la hora de llevar a cabo un seguimiento más estrecho de la enfermedad, apoyando y agilizando la toma de decisiones.*

## 1. Introducción

El cáncer representa uno de los problemas sanitarios más importantes, ya que constituye una de las principales causas de mortalidad a nivel mundial. Atendiendo a los datos proporcionados por el Centro Internacional de Investigaciones sobre el Cáncer (International Agency for Research on Cancer, IARC), el número de muertes relacionadas con cáncer en el año 2020 fue de 9,9M aproximadamente, siendo el responsable del mayor número de fallecimientos el cáncer de pulmón. La OMS estima que alrededor de un tercio de las muertes por cáncer son debidas

a cinco factores de riesgo evitables: el tabaco, el alcohol, las infecciones, el sedentarismo y las dietas inadecuadas [1,2].

La medicina personalizada representa la adaptación del tratamiento médico a las características clínico-genéticas del paciente. A este respecto, la nutrigenética permite alcanzar una nutrición de precisión a través del análisis de marcadores genéticos que indican los efectos de ciertos alimentos sobre el organismo y la predisposición a padecer determinadas enfermedades [3]. En el ámbito de la oncología, el análisis del perfil nutrigenético del paciente permitiría diseñar un plan de nutrición y estilo de vida personalizado, potenciando los efectos positivos del tratamiento oncológico. En la misma línea, el estudio de la microbiota intestinal está adquiriendo una gran relevancia, ya que, aunque todavía no puede establecerse una clara relación de causalidad, parece existir una asociación entre la alteración de la composición y funcionalidad normal de la misma, estado conocido como “disbiosis”, y el desarrollo de algunas patologías, como la obesidad o diversos tipos de cáncer [4].

La eHealth (salud electrónica) y la mHealth (salud móvil) utilizan las Tecnologías de la Información y la Comunicación (TIC) y de la Ingeniería Biomédica en el campo de la salud para apoyar procesos como la prevención, diagnóstico, tratamiento y seguimiento de pacientes.

La evidencia clínica ha demostrado como las deficiencias en el control de los síntomas afectan a la calidad de vida, la adherencia al tratamiento, y el pronóstico del paciente oncológico. La revisión de diferentes intervenciones con aplicaciones móviles en este ámbito demuestra como el control estrecho de los síntomas permite detectar recaídas de forma precoz, lo que se traduce en un incremento de supervivencia [5,6]. Por ello, el desarrollo de nuevas herramientas que permitan el autorregistro y una monitorización más frecuente, más allá de los datos recogidos en las visitas médicas rutinarias, serían de gran utilidad al permitir la adopción de un enfoque de seguimiento continuo en la atención clínica.

El programa ALIBIRD2020CM, formado por un consorcio multidisciplinar, tiene entre sus objetivos el diseñar y validar estrategias de nutrición de precisión, dirigidos a enfermos de cáncer a través de una plataforma mHealth [7]. La hipótesis del programa de investigación es que un abordaje basado en la medicina de precisión y la monitorización personalizada del paciente con tecnologías mHealth, pulseras de actividad y otros datos reportados por el propio paciente (*Patient Reported Outcomes Measures*, PROMs), potenciarán el empoderamiento de este y mejorarán la calidad asistencial.

En este trabajo se presenta el análisis preliminar de la aplicabilidad de la plataforma ALIBIRD en el manejo y seguimiento de pacientes con cáncer a partir de los resultados disponibles de la evaluación clínica que se está llevando a cabo en el Hospital Universitario Infanta Sofía, y se aborda

## 2. Materiales y métodos

### 2.1. Plataforma mHealth ALIBIRD®

La plataforma mHealth ALIBIRD® está compuesta por 3 componentes principales (Tabla 1): 1) aplicación móvil para el paciente oncológico, 2) aplicación Web para los profesionales sanitarios, y 3) infraestructura tecnológica que soporta el sistema [8].

App móvil (paciente)	Aplicación Web (profesional)
Registro de PROs y PROMs <ul style="list-style-type: none"> <li>- Síntomas</li> <li>- Alimentación, actividad física y sueño (pulsera de actividad)</li> <li>- Estado de ánimo</li> <li>- Cuestionarios validados</li> </ul>	<ul style="list-style-type: none"> <li>- Seguimiento de pacientes</li> <li>- Monitorización de variables clínicas (PROMs)</li> <li>- Visualización y atención de alertas por síntomas, dieta o reducción importante del peso</li> <li>- Acceso a reportes semanales</li> <li>- Apoyo personalizado al paciente y contacto por mensajería</li> </ul>
Apoyo en la autogestión <ul style="list-style-type: none"> <li>- Recomendaciones personalizadas</li> <li>- Informes y artículos</li> <li>- Mensajería</li> </ul>	
<b>Infraestructura IT</b> (Servidores, bases de datos y servicios: eCRF REDCap [9], Google Services, Fitbit, etc.)	
<ul style="list-style-type: none"> <li>- Procesamiento y almacenamiento de datos</li> <li>- Generación de recomendaciones personalizadas</li> <li>- Envío de notificaciones</li> </ul>	

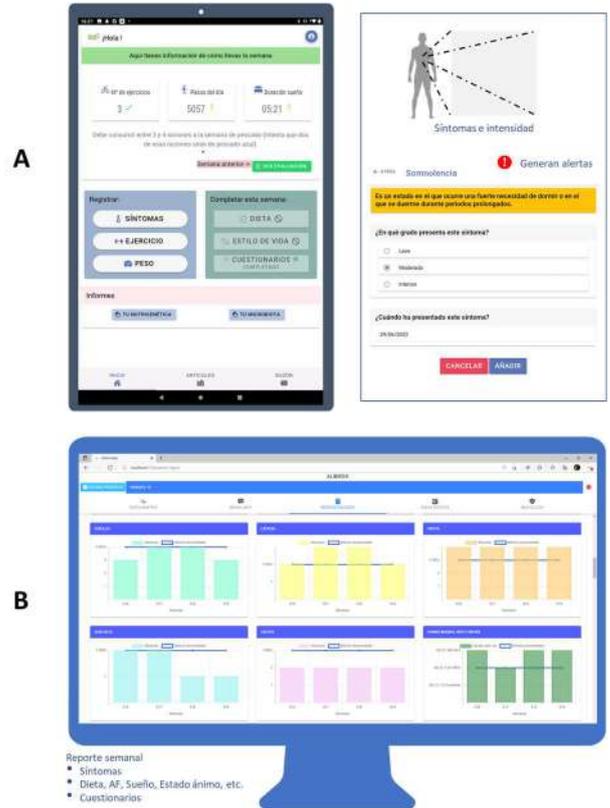
**Tabla 1.** Características y funcionalidades de la plataforma

En las fases de conceptualización, definición y diseño de las aplicaciones, se utilizó el Diseño Centrado en el Usuario o UCD, incluyendo estrategias de codiseño o diseño participativo. Para ello, desarrollaron diferentes actividades, como entrevistas individuales, sesiones del tipo “focus group” o grupos de enfoque en las que participaron pacientes, sanitarios e investigadores del consorcio.

La aplicación móvil para el paciente oncológico incluye escenarios para el registro y visualización de datos, conexión con pulseras de actividad física, y recepción de recomendaciones personalizadas basadas en la nutrigenética y microbiota del paciente (Figura 1-A).

Toda la información es procesada automáticamente y queda disponible en una aplicación Web para que los profesionales sanitarios puedan acceder y monitorizar los datos de los

pacientes para apoyar la toma de decisiones de los profesionales sanitarios implicados (Figura 1-B).



**Figura 1.** A) app móvil paciente, B) app Web profesional

### 2.2. Estudio clínico

El ensayo ha sido aprobado por el Comité de Ética de la Investigación con Medicamentos (CEIm) IdiPAZ (código PI-4735). La muestra de estudio la componen 20 pacientes con cáncer de pulmón no microcítico (CPNM) avanzado, en tratamiento activo. Entre los objetivos del estudio están: 1) evaluar la experiencia de los participantes como usuarios de la aplicación; 2) evaluar el impacto clínico, analizando la detección precoz de progresiones clínico-radiológicas y las recomendaciones nutricionales de precisión; y, 3) evaluar la posible asociación pronóstico y respuesta al tratamiento teniendo en cuenta factores nutricionales, genéticos, metabólicos y la microbiota.

El periodo de intervención y seguimiento es de 6-7 meses y se realizan 10 visitas, en las que se registran y visualizan las variables relativas a las dimensiones clínicas de interés a través de las aplicaciones del paciente y el profesional (ver Tabla 2). Además, cada visita incluye una consulta oncológica, nutricional y de soporte general (gestión clínica y apoyo técnico). En la consulta nutricional, se recogen datos antropométricos y se revisan los datos capturados por el paciente con la aplicación móvil y la pulsera de actividad física, en base a los cuales se dan recomendaciones nutricionales personalizadas. La evaluación nutricional se complementa con la realización de una ecografía muscular en las visitas 1, 5 y 10 que permite analizar la evolución del músculo durante el periodo de intervención. En la visita oncológica, el clínico, a través de la aplicación Web, monitoriza los PROMs registrados por el paciente lo que agiliza la toma de decisiones clínicas.

Nº visita	Actividades y variables	Instrumentos
V0	Reclutamiento ✓ Criterios I/E y firma CI y HIP	Documentos
V1	Inicio de la intervención ✓ Muestras: saliva (nutrigenómica) y heces (microbiota basal) ✓ Antropometría y eco muscular ✓ Datos basales ✓ Entrega Tablet y pulsera	-App móvil paciente y app Web profesional -Cuestionarios (REDCap)
V2-9	Seguimiento ✓ Antropometría y eco muscular (V5) ✓ Registro y visualización PROMs	-App móvil paciente y app Web profesional -Cuestionarios (REDCap)
V10	Fin de la intervención ✓ Antropometría y eco muscular ✓ Muestra heces (microbiota) ✓ Registro y visualización PROMs ✓ Valoración final usabilidad	-App móvil paciente y app Web profesional -Cuestionarios (REDCap)

**Tabla 2.** Protocolo del estudio

A través del análisis de las muestras de saliva y heces se generan dos informes:

- Nutrigenética: agrupa biomarcadores relacionados con el metabolismo, inflamación e inmunidad, biorritmos y respuesta a compuestos.
- Microbiota: analiza la presencia de potenciales grupos bacterianos, su actividad metabólica, perfil microbiano y relación con efectos positivos y negativos en la salud.

En la etapa de intervención, el sistema procesa los datos registrados y se genera un reporte de evaluación semanal, accesible desde la aplicación móvil, con recomendaciones personalizadas y moduladas por la nutrigenética.

Una vez finalizado el ensayo se procederá al estudio de los datos para llevar a cabo un análisis descriptivo y comparativo que permita determinar cambios significativos, tendencias y correlaciones entre las variables.

La evaluación de la plataforma se centra en la usabilidad de las aplicaciones, satisfacción y experiencia de los usuarios. Para evaluar la usabilidad se utiliza la Escala de Usabilidad del Sistema (SUS) [10], cuestionario de 10 ítems con 5 opciones de respuesta. El resultado varía entre 0 y 100, siendo 0 indicativo de muy baja usabilidad y 100 excelente. Adicionalmente, para valorar la aceptabilidad se utilizan dos preguntas como mecanismo de experiencia reportada por el paciente (Patient Reported Experience, PREM).

### 3. Resultados

Se ha completado el reclutamiento de los 20 voluntarios pacientes con cáncer de pulmón avanzado y en tratamiento activo, 12 mujeres (60%) y 8 hombres (40%), con una media de edad de 61 años (43-78).

Los resultados aquí presentados son preliminares y se corresponden con los datos recogidos y analizados en un momento en el que todos los pacientes han superado la visita 7 y el 65% han finalizado el estudio, es decir, 13 pacientes entre los que diferenciamos 6 abandonos prematuros por “exitus” y 7 pacientes que han completado todas las visitas.

Algunas cifras de los datos capturados a través de la app móvil del paciente y procesados por el sistema se muestran

en la tabla siguiente (Tabla 3). De los más de 500 síntomas registrados, el 9,66% eran graves y generaron 54 alertas atendidas por los profesionales sanitarios en una media de 1,6 días, evitando 54 visitas a urgencias.

SÍNTOMAS: 559 registros			
Leve	Moderado	Intenso	Nº de alertas
62,25 %	28,98 %	9,66%	54

ACTIVIDAD FÍSICA: 2736 registros			
Nº total de sesiones	Nº de sesiones por semana $\bar{X}$ (SD)	Nº máximo de sesiones por semana	Pasos diarios $\bar{X}$ (SD)
1254	3 (2,72)	20	5470 (2798,91)

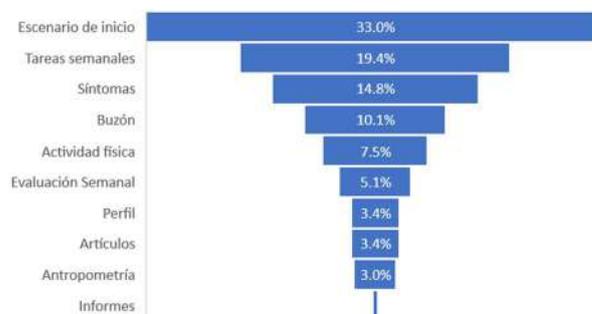
SUEÑO: 2338 registros		
Duración media $\bar{X}$ (SD)	Duración máxima	Duración mínima
7,7h (1,09)	12,31h	1,78h

CUESTIONARIOS				
Semanales	Asociados a las visitas			
Nutrición y estilo de vida	Actividad física, Toxicidad, Calidad de Vida	Adherencia dieta mediterránea	Sueño, ansiedad y depresión, adherencia tto.	Sexualidad
326	121	20	35	49

**Tabla 3.** Datos capturados (PROMs) e indicadores

En la figura siguiente (Figura 3) se muestran datos asociados al uso de los diferentes escenarios por parte de los pacientes.



**Figura 3.** Distribución del uso de los escenarios de la app móvil

Los 4 escenarios más utilizados son el de inicio de la aplicación (33%); la sección de tareas semanales (19,4%) que incluye acciones relacionadas con la dieta y otras variables del estilo de vida (sueño, estado de ánimo, energía y ritmo gastrointestinal); el escenario de los síntomas (14,8%), la sección del buzón (10,1%) que incluye notificaciones automáticas y la mensajería médico-paciente.

Respecto a la evaluación de la usabilidad y la aceptación por parte de los 7 pacientes que han finalizado el estudio, la Figura 5 muestra la media de los resultados de las preguntas del cuestionario SUS. La puntuación total obtenida es de 92,9 puntos, que se corresponde con una calificación “A”, el grado más alto de aceptación. Todos los pacientes se han mostrado satisfechos o muy satisfechos con el servicio proporcionado por la plataforma y han declarado que recomendarían la aplicación.

En lo relativo a la utilidad de la plataforma para la toma de decisiones clínicas y terapéuticas, los PROMs registrados facilitaron la detección precoz de 4 progresiones clínico-radiológicas, fueron decisivos para realizar 13 derivaciones a otras especialidades como Digestivo, Endocrinología-Nutrición o Dermatología, y permitieron realizar 26 ajustes terapéuticos (analgésia, antibioterapia, etc.) así como anticipar 6 citas. Hasta la fecha, a través de la plataforma se han intercambiado un total de 70 mensajes entre pacientes y profesionales sanitarios atendidos en una media de 1,6 días.

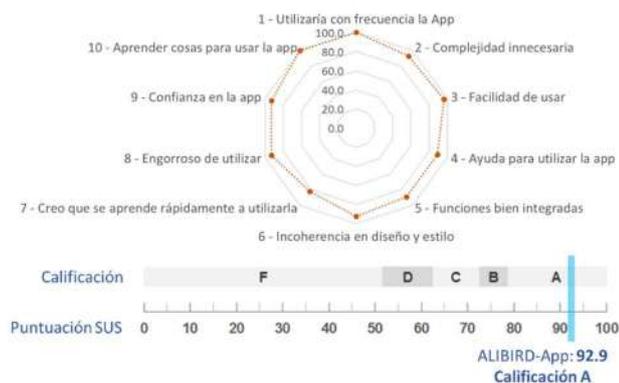


Figura 5. Puntuación SUS app móvil

Los análisis nutrigenéticos han permitido la detección de 5 pacientes intolerantes a la lactosa y 4 pacientes con tolerancia intermedia al gluten. En cuanto a los análisis de la microbiota intestinal, los resultados preliminares nos indican que el 75% de los pacientes presentan disbiosis.

#### 4. Discusión y conclusiones

Factores como la adherencia al tratamiento, la monitorización de síntomas, y la adopción de hábitos de vida saludables son factores determinantes en el pronóstico del paciente oncológico. En este sentido, el aspecto innovador y diferenciador de la plataforma mHealth ALIBIRD® reside en que unifica, en una sola herramienta digital, los beneficios derivados de un registro amplio de PROMs con el análisis de la nutrigenética y la microbiota del paciente lo que contribuye al desarrollo de recomendaciones personalizadas. A través de este enfoque multidisciplinar, se busca que el clínico disponga de información para analizar y evaluar de manera global el estado y la evolución del paciente, y su vez apoyar y agilizar la toma de decisiones clínicas.

La evaluación preliminar de la interacción de los usuarios con la plataforma indica que el “Escenario de inicio” ha sido el más utilizado, lo que tiene sentido ya que su interfaz cuenta con un resumen de datos asociados al estilo de vida (actividad física y el sueño), y es donde se concentra el acceso al resto de escenarios y funcionalidades. Otro de los escenarios más utilizados es “Tareas semanales”, a través del cual los pacientes debían cumplimentar una serie de registros sobre dieta, estado de ánimo, y cuestionarios validados vinculados a las visitas del estudio. Destaca también el uso de “Síntomas” y “Buzón”, el primero de ellos por representar una variable clave en la autogestión y control de la enfermedad, y el segundo, tanto para visualizar las notificaciones/recomendaciones automáticas de la aplicación como para el envío de mensajes al servicio clínico

y correspondiente resolución de dudas. Por su parte, las acciones clínico-terapéuticas de profesionales a través de la aplicación Web y su tiempo de respuesta, refuerzan el propósito y utilidad de la implementación de este tipo de soluciones en la práctica clínica.

El estudio actual presenta algunas limitaciones, como el tamaño de la muestra o la ausencia de una rama control, y además aún no ha finalizado. Aun siendo cautos en las conclusiones, los hallazgos obtenidos hasta el momento relacionan la plataforma mHealth ALIBIRD® con un impacto clínico favorable, elevada satisfacción y se configura como una potente herramienta basada en la evidencia científica y con un enfoque integral en la atención del paciente oncológico. Por ello creemos que su implantación en la práctica clínica sería beneficiosa y apoyaría la mejora del pronóstico del tratamiento, la calidad de vida del paciente, la optimización de recursos asistenciales y la generación de conocimiento.

#### Agradecimientos

El consorcio ALIBIRD2020-CM (S2018/BAA-4343): “Fórmulas terapéuticas de nutrición de precisión para el cáncer” ha sido financiado por la DG de Investigación e Innovación de la Comunidad de Madrid (DGII) y cofinanciado con Fondos Estructurales de la UE.

Al proyecto COVITECH-CM. REACT-UE a través del Fondo Europeo de Desarrollo Regional (FEDER), el Fondo Social Europeo (FSE) y el Fondo de Ayuda Europea para las Personas Más Desfavorecidas (FEAD).

#### Referencias

- [1] International Agency for Research on Cancer. World Cancer Report: Cancer Research for Cancer Prevention, (2020), eBook (ISBN: 978-92-832-0448-0).
- [2] Sociedad Española de Oncología Médica (SEOM). Las cifras del cáncer en España 2021, (2022). [https://seom.org/images/Cifras\\_del\\_cancer\\_en\\_España\\_2021.pdf](https://seom.org/images/Cifras_del_cancer_en_España_2021.pdf)
- [3] Corella, D. et al. Papel de las ómicas en la nutrición de precisión: fortalezas y debilidades. *Nutrición Hospitalaria*, 35(spe4), 10-18, (2018), doi: 10.20960/nh.2119
- [4] Manor, O. et al. Health and disease markers correlate with gut microbiome composition across thousands of people. *Nat Commun* 11, 5206, (2020), doi: 10.1038/s41467-020-18871-1
- [5] Copeland, C. et al. M-Health Application Interface Design for Symptom Checking. *Proc. of the 10th Intl. Conf. on e-Health 2018*, Madrid, Spain, July 17-19, pp. 210-214, (2018).
- [6] Ciani O. et al. Lung Cancer App (LuCApp) study protocol: a randomised controlled trial to evaluate a mobile supportive care app for patients with metastatic lung cancer. *BMJ Open* 9(2):e025483, (2019). doi: 10.1136/bmjopen-2018-025483.
- [7] Página web del Programa ALIBIRD2020-CM, “Fórmulas terapéuticas de nutrición de precisión para el cáncer”. <https://alibird.org/2020-CM>
- [8] Iniesta-Chamorro J.M. et al. Plataforma mHealth ALIBIRD para el apoyo en el tratamiento y seguimiento de pacientes con cáncer. *Actas del XXXVIII Congreso Anual de la Sociedad Española de Ingeniería Biomédica (CASEIB)*. 25 – 27 Nov, Valladolid, pp 223, (2020) (ISBN: 978-84-09-25491-0).
- [9] Harris P.A. et al. The REDCap consortium: Building an international community of software partners. *J Biomed Inform.* (2019) doi: 10.1016/j.jbi.2019.103208
- [10] Brooke, J. SUS-A quick and dirty usability scale. *Usability evaluation in industry*, 189(194), 4-7, (1996).

# Sistema de ayuda a la prescripción farmacológica

G. Balastegui García<sup>1</sup>, D. Ruiz Fernández<sup>2</sup>, M. Lillo Crespo<sup>3</sup>, F. Gomis Jimeno<sup>4</sup>

<sup>1</sup> Escuela Politécnica Superior, Universidad de Alicante, Alicante, España, guillermo.balastegui@ua.es

<sup>2</sup> Escuela Politécnica Superior, Universidad de Alicante, Alicante, España, druiz@gcloud.ua.es

<sup>3</sup> Dpto. Enfermería, Universidad de Alicante, Alicante, España, manuelliglocrespo@gmail.com

<sup>4</sup> Enfermero de hospitalización, Clínica HLA Vistahermosa, Alicante, España, fgomisji@gmail.com

## Resumen

*Los sistemas expertos y la inteligencia artificial en su conjunto cobran cada día mayor relevancia dentro del mundo de la medicina y la atención médica. Su uso dentro del proceso de prescripción de medicamentos puede ser de gran ayuda para optimizar los procesos asistenciales y mejorar la eficiencia de la atención farmacológica. Mediante el uso de los datos de la historia clínica de cada paciente y las pautas de prescripción de cada medicamento y las variables médicas de este durante la consulta, se puede crear una herramienta que ahorre a los profesionales sanitarios, en el caso de este proyecto, los enfermeros, a prescribir medicinas, mejorando la seguridad del paciente y los tiempos de atención sanitaria. El proyecto llevado a cabo consiste en una herramienta accesible desde cualquier plataforma para la prescripción farmacológica y de formación para el enfermero que le otorga un sistema en el que apoyarse para realizar prescripciones.*

## 1. Introducción

Tras la publicación del Real Decreto 1302/2018, del 22 de octubre, que regula la indicación, uso y autorización de dispensación de medicamentos y productos sanitarios por parte de enfermeros, los profesionales de la rama de enfermería cobraron aún una mayor importancia dentro del cuidado farmacéutico de los pacientes, siendo un personaje clave en este tipo de cuidados de la población y aportando un efecto muy positivo al Sistema de Salud Español gracias a su nuevo papel de prescripción[1].

Los profesionales de enfermería cuentan con una gran formación de tipo universitario que les aporta unas nociones básicas sobre la prescripción de medicamentos, pero que no son suficientes para desarrollar esta labor asegurando la máxima seguridad para el paciente [1].

En un contexto en el que todos los servicios cada vez están más digitalizados y la inteligencia artificial mejora a pasos agigantados, surge la idea de desarrollar una aplicación web multiplataforma que proporcione una herramienta fiable a los enfermeros y enfermeras con el objetivo de ayudarles con la prescripción de medicamentos a pacientes y mejorar su formación y conocimientos en cuanto a esta labor.

El Real Decreto mencionado previamente establece que los enfermeros deben ser titulares de una acreditación emitida por la comunidad autónoma en la que trabajan (o bien tener un año de experiencia laboral de un año), siendo estas comunidades autónomas las que proporcionan las pautas

de prescripción de los medicamentos que habilitan a prescribir, definiendo concretamente los posibles casos y procesos que puede atender el enfermero[2].

Con las directrices marcadas por este Real Decreto junto con las indicaciones y la reorientación directa del Profesor Manuel Lillo y el enfermero Francisco Gomis se define la creación de una aplicación con dos pilares, un sistema de formación basado en pruebas de tipo test y un sistema de apoyo a la decisión clínica (CDSS). Estos pilares se sostienen en las pautas de prescripción clínicas proporcionadas por la Junta de Andalucía[3] [4] [5], que en la fecha de elaboración de este artículo son únicas en España. Estas pautas tratan la prescripción de medicamentos para el tratamiento de la Diabetes Mellitus tipo 2, enfermedades de riesgo vascular (hipercolesterolemia e hipertensión) y personas en riesgo de coagulación.

El CDSS empleará la información de los pacientes y de los medicamentos que se pueden prescribir en España. Los datos sobre los pacientes se obtienen mediante su registro en la base de datos de la aplicación y la información relativa a los medicamentos se recibe mediante el consumo de la API REST CIMA[6]. Esta API proporciona acceso al registro de medicamentos del Ministerio de Sanidad y de la Agencia Española de Medicamentos y Productos Sanitarios (AEMPS).

El sistema experto en el que se basa el CDSS se fundamenta en tres pilares, la base de hechos, la base de conocimientos y el motor de inferencia:

- La base de hechos se forma mediante los datos de historia clínica del paciente, los datos recogidos sobre este en la consulta y la información obtenida de la API REST CIMA.
- La base de conocimientos está integrada por las pautas de prescripción clínicas de Andalucía.
- El motor de inferencia es el encargado de comparar los datos de la base de hechos con las normas de la base de conocimientos para obtener una respuesta que será emitida al usuario.



**Ilustración 1 - Esquema de funcionamiento del CDSS**

### 1.1. Proyecto Demophac

Este proyecto de desarrollo de un CDSS se ubica dentro del marco del proyecto Demophac, que busca unificar los criterios de cuidado farmacéutico en Europa[7], el proyecto Demophac obtuvo como conclusiones que el cuidado farmacéutico es de vital importancia para la salud del paciente y que es necesario que existan unos criterios unificados, otorgando relevancia a los enfermeros dentro de este, dando más cabida a aplicaciones como la que se ha desarrollado en el proyecto que se trata en este documento.

### 1.2. Otros sistemas expertos

En la actualidad se pueden encontrar otros sistemas expertos para el cuidado de la salud, no solo de las personas, también de los animales. Por ejemplo, desde hace varios años, ante el aumento de la población equina de China, se comenzó a emplear un CDSS basado en un sistema experto para el tratamiento de hasta 91 enfermedades de caballos, mostrando buenos resultados y potencial para su empleo en la práctica veterinaria[8]. En cuanto al uso con pacientes humanos, se han encontrado aplicaciones similares a la creada en este proyecto, como es el caso de un sistema experto para el prediagnóstico de hipertensión, DM 2 y síndrome metabólico[9], sin embargo, este solo es empleado para el prediagnóstico de estas enfermedades y no para su tratamiento, aunque los resultados de este son muy parejos con los de los doctores.

## 2. Partes de la aplicación

Para desarrollar la aplicación primero hubo que definir la información que necesita el sistema experto para poder emitir prescripciones. Esta información gira entorno al paciente y los medicamentos que toma y puede tomar.

### 2.1. Información de historia clínica

Los pacientes deben ser registrados en la aplicación mediante un formulario que permite introducir datos sobre sus tratamientos actuales, las alergias que pueda tener, las patologías que tiene o ha tenido, si se encuentra en un estado de gestación o amamantando a un hijo y la información habitual como su nombre, su fecha de nacimiento y su peso.

Toda esta información permite que el sistema experto pueda comprobar la siguiente información:

- Incompatibilidades entre medicamentos.
- Incompatibilidades del medicamento con el estado del paciente, ya sea por una enfermedad o por gestación o lactancia.
- Riesgos por alergias.
- Cálculo de nuevas dosis en función del peso.

Una vez registrado, el formulario siempre estará disponible para poder realizar cambios sobre esta información ya que puede sufrir modificaciones con el tiempo. Estos formularios están creados para usarse fuera de la consulta, que su uso acarrea un tiempo del que no se dispone en las consultas que realizan los enfermeros.

### 2.2. Información sobre los medicamentos

Para completar la base de hechos de la aplicación se requiere los datos de los medicamentos que el CDSS va a prescribir. Estos datos deben estar actualizados y contener el menor número de errores posibles. El Ministerio de Sanidad y la AEMPS proporcionan un buscador de medicamentos en su web[10] y como se ha mencionado previamente una API que puede ser consumida y devuelve la información de los medicamentos buscados en formato JSON. La información que recoge la aplicación de la API es la siguiente:

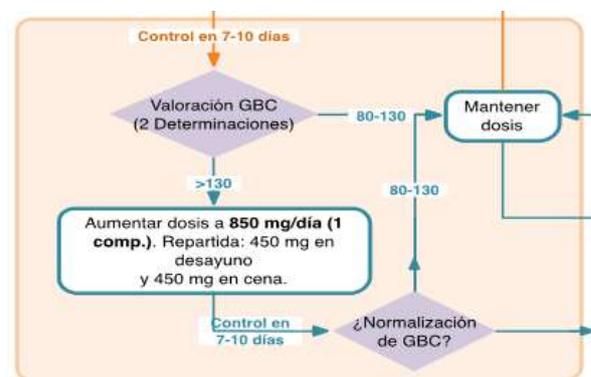
- Nombre del medicamento.
- Principios activos.
- Forma farmacéutica simplificada.
- Dosis.
- Imagen de la caja y forma farmacéutica.
- Vía de administración.
- Número de registro.
- Ficha técnica.

Toda esta información permite que el motor de inferencia compare los datos de los medicamentos con los requerimientos del paciente y las pautas que marca la base de conocimientos. La documentación de la API se puede encontrar en las referencias [6].

### 2.3. Base de conocimientos

La base de conocimientos se encuentra incrustada en el código, y define las normas que sigue el sistema experto para recomendar unos medicamentos u otros en función del estado del paciente y sus variables médicas. Esta parte del sistema experto está implementada mediante funciones correspondientes a cada uno de los principios activos que aparecen en las pautas de prescripción empleadas.

Las pautas de prescripción empleadas están representadas en los documentos en forma de diagramas de flujo que avanzan en función de los datos del paciente como su glucemia capilar, su colesterol, su hemoglobina glucosilada, su INR y otros.



**Ilustración 2 - Ejemplo de pauta de prescripción (diabetes)**

Los principios activos (agrupados por enfermedad) para los que se ha creado una función en la base de conocimientos son los siguientes:

- Diabetes [3]:
  - o Metformina.
  - o Gliclazida, glipizida y glimepirida.
  - o Insulina.
- Anticoagulación [4]:
  - o Acenocumarol.
  - o Warfarina.
- Hipercolesterolemia [5]:
  - o Simvastatina.
- Hipertensión [5]:
  - o Enalapril y Ramipril.
  - o Clortalidona.
  - o Amlodipino.

Estas funciones comprueban en primer lugar que el paciente no tenga ninguna alergia a alguno de los principios activos, ninguna incompatibilidad con sus patologías y sus tratamientos actuales y en caso de que el medicamento suponga algún peligro para la gestación o la lactancia comprueba si el paciente se encuentra en estas situaciones. Solo si el medicamento ha pasado estos filtros pasará entonces a actualizar la dosis que debe tomar el paciente, modificar el tratamiento con el correspondiente cálculo de la nueva dosis o en caso de que sea necesario devolver al paciente a la atención de su médico.

#### 2.4. Motor de inferencia

El motor de inferencia está implementado mediante una función que hace pasar los datos del paciente y sus variables médicas por las funciones de los medicamentos en el orden de prioridad en el que se deben prescribir hasta que se llega a una salida. Cuando una función genera un tratamiento, el motor de inferencia envía la respuesta al usuario y en caso de que no se llegue a ningún posible tratamiento porque ninguno es compatible, esto se notifica al enfermero, que es el que debe tomar siempre la decisión final sobre la prescripción.

Este componente se encarga también de discernir la enfermedad con la que se está trabajando y de emplear un conjunto de funciones u otro.

#### 2.5. Sistema de formación

El sistema de formación para los enfermeros se ha pensado para que se realicen pruebas de forma periódica cada mes. Estas pruebas son creadas y publicadas por un administrador de forma que el enfermero puede realizarlas en cualquier momento durante el ciclo correspondiente. Están divididas en función de la enfermedad sobre la que se esté tratando en ese momento, es decir, cada enfermero debe realizar tres pruebas al mes, una sobre el tratamiento para diabetes, otra para el tratamiento de ACOs y otra para pacientes en riesgo vascular.

Estas pruebas ofrecen un informe al final para el enfermero, permitiendo que conozca en qué conocimientos falla, de esta forma puede reforzarlos en caso de que sea necesario. Además, el administrador puede comprobar los

resultados de los evaluados para tomar las decisiones que considere necesarias sobre su formación.



Ilustración 3 - Informe de prueba

### 3. Implementación

La aplicación está desarrollada sobre una arquitectura cliente-servidor. El servidor está implementado mediante una API REST, que es la encargada de comunicarse con la base de datos y de emplear el CDSS.

La aplicación se ha creado mediante el conjunto de HTML, CSS y JavaScript, empleando también este último para implementar el sistema experto. La base de datos está implementada mediante MySQL y XAMPP.

La API REST emplea se ejecuta empleando el entorno NodeJS y la aplicación tiene una capa de seguridad que emplea JWT.

La aplicación ofrece un sistema de gestión de usuarios, pacientes y sus citas, un sistema de gestión para la formación, y el CDSS usado en las citas con los pacientes.

Al tratarse de una aplicación que debe estar disponible en el entorno de trabajo del sanitario, esta se ha diseñado con “responsive design” de forma que pueda ser empleada en cualquier dispositivo, ya sea móvil, tableta u ordenador.

### 4. Resultados

El CDSS obtenido es capaz de emitir prescripciones en función de los parámetros mencionados y se ha implementado de una forma práctica y usable. Este da recomendaciones sobre la dosis que debe tomar el paciente y la duración del nuevo tratamiento.



Ilustración 4 - Interfaz del CDSS

Además, incluye recomendaciones para el enfermero, para que no olvide dar indicaciones importantes al paciente.

Tras haber mostrado los resultados al enfermero Francisco Gomis y al Profesor Manuel Lillo, la conclusión es que la aplicación se puede implementar en un entorno clínico.

## 5. Conclusiones y líneas de trabajo

Una vez finalizada la primera versión de la aplicación de seguridad en la prescripción y tras haber analizado esta, las conclusiones que se pueden extraer diversas conclusiones.

Los CDSS son herramientas que pueden favorecer enormemente a los profesionales sanitarios, aunque estos pueden tener algunos efectos negativos. Los usuarios pueden perder tiempo usando esta aplicación si no está bien diseñada, tiempo propio a la atención del paciente. También puede darse el caso de que al emplear esta aplicación reiteradamente, los profesionales se vuelvan complacientes y olviden los conocimientos y habilidades que aporta el CDSS al no tener que usarlos por su propia cuenta, por lo que el sistema de formación es importante.

Los sistemas de apoyo a la decisión clínica ayudan a mejorar la seguridad del paciente y la velocidad de la atención médica, el hecho de que sean capaces de emplear estas pautas de prescripción y comprobar todas las incompatibilidades del paciente en pocos segundos permiten que el enfermero tenga más tiempo que dedicar al paciente y un refuerzo que aporte más seguridad en su trabajo.

La aplicación todavía no ha sido lanzada a un entorno real, por lo que no se pueden extraer unas conclusiones con más profundidad.

En cuanto a las posibles mejoras de la aplicación, estas pueden ir en varias direcciones:

- Integración con los sistemas de historia clínica del hospital para disponer de todos los datos del paciente.
- Creación de una red neuronal que sustituya al sistema experto.
- Mejora del sistema experto, haciendo que este pueda ser actualizado por el enfermero.
- Mejoras en el rendimiento y la interfaz.
- Añadir nuevas funciones para el paciente, mejorando así su adherencia al tratamiento.
- Emplear la aplicación para extracción de datos sobre la eficacia de los tratamientos.
- Mejoras en el sistema de formación, sumando más modalidades e integrando los contenidos en la app

Los CDSS ofrecen muchas posibilidades a explotar y su uso no está generalizado, por lo que se postulan como un gran campo de trabajo que puede ayudar a optimizar la atención farmacológica y médica de los pacientes.

Los profesionales del campo de la enfermería se mostraron muy satisfechos con los resultados de este proyecto y admitieron que una herramienta de este tipo puede ser de gran ayuda en su día a día.

## Agradecimientos

Este proyecto ha sido realizado gracias a la ayuda y tutorización por parte de los tutores Daniel Ruiz y Manuel Lillo y el enfermero Francisco Gomis.

## Referencias

- [1] M. Lillo-Crespo, J. Riquelme-Galindo, E. D. Baetselier, B. V. Rompaey, y T. Dilles, «Understanding pharmaceutical care and nurse prescribing in Spain: A grounded theory approach through healthcare professionals' views and expectations», *PLoS ONE*, vol. 17, n.º 1, p. e0260445, ene. 2022, doi: 10.1371/journal.pone.0260445.
- [2] Ministerio de Sanidad, Consumo y Bienestar Social, *Real Decreto 1302/2018, de 22 de octubre, por el que se modifica el Real Decreto 954/2015, de 23 de octubre, por el que se regula la indicación, uso y autorización de dispensación de medicamentos y productos sanitarios de uso humano por parte de los enfermeros*, vol. BOE-A-2018-14474. 2018, pp. 102636-102643. Accedido: 9 de abril de 2022. [En línea]. Disponible en: <https://www.boe.es/eli/es/rd/2018/10/22/1302>
- [3] Aguilar Diosdado, Manuel, Molina López, Teresa, Bellón Saameño, Juan Ángel, Morilla Herrera, Juan Carlos, Carrascosa García, M<sup>a</sup> Isabel, y Pajares Bernaldo de Quirós, Ignacio, «Seguimiento Protocolizado del Tratamiento Farmacológico Individualizado en DIABETES». Junta de Andalucía - Consejería de Salud y Familias, 2019. Accedido: 14 de mayo de 2022. [En línea]. Disponible en: [https://www.sspa.juntadeandalucia.es/servicioandaluzdesalud/sites/default/files/sincfiles/wsas-media-mediafile\\_sasdocumento/2019/Seguimiento%20Prot\\_Tto\\_%20Farmac\\_%20DIABETES\\_Nueva.pdf](https://www.sspa.juntadeandalucia.es/servicioandaluzdesalud/sites/default/files/sincfiles/wsas-media-mediafile_sasdocumento/2019/Seguimiento%20Prot_Tto_%20Farmac_%20DIABETES_Nueva.pdf)
- [4] Aranda Escribano Manuela, Bel Peña Nieves, Llamas del Castillo M<sup>a</sup> Dolores, Guarino Núñez Manolo, y Pablo Guzmán Puello, «Seguimiento Protocolizado del Tratamiento Farmacológico Individualizado en Personas en Tratamiento con Anticoagulación oral». Junta de Andalucía - Consejería de Salud y Familias, 2019. Accedido: 14 de mayo de 2022. [En línea]. Disponible en: [https://www.sspa.juntadeandalucia.es/servicioandaluzdesalud/sites/default/files/sincfiles/wsas-media-mediafile\\_sasdocumento/2019/Seguimiento%20farmacologico%20de%20personas%20con\\_ACO\\_2019\\_3\\_edicion.pdf](https://www.sspa.juntadeandalucia.es/servicioandaluzdesalud/sites/default/files/sincfiles/wsas-media-mediafile_sasdocumento/2019/Seguimiento%20farmacologico%20de%20personas%20con_ACO_2019_3_edicion.pdf)
- [5] Aguilar Diosdado, Manuel, Bellón Saameño, Juan Ángel, Morilla Herrera, Juan Carlos, y Pajares Bernaldo de Quirós, Ignacio, «Seguimiento Protocolizado del Tratamiento Farmacológico Individualizado en RIESGO VASCULAR: Hipolipemiantes Antihipertensivos». Junta de Andalucía - Consejería de Salud y Familias, 2019. [En línea]. Disponible en: [https://www.sspa.juntadeandalucia.es/servicioandaluzdesalud/sites/default/files/sincfiles/wsas-media-mediafile\\_sasdocumento/2019/SeguimientoProtocolizado%20RV.pdf](https://www.sspa.juntadeandalucia.es/servicioandaluzdesalud/sites/default/files/sincfiles/wsas-media-mediafile_sasdocumento/2019/SeguimientoProtocolizado%20RV.pdf)
- [6] «CIMA-REST-API\_1\_19.pdf». Accedido: 18 de mayo de 2022. [En línea]. Disponible en: [https://sede.aemps.gob.es/docs/CIMA-REST-API\\_1\\_19.pdf](https://sede.aemps.gob.es/docs/CIMA-REST-API_1_19.pdf)
- [7] «Erasmus+ DeMoPhaC», *nuphac*. <https://www.nuphac.eu/erasmus-demophac> (acceso 14/5/2022).
- [8] H. Qin, J. Xiao, X. Gao, y H. Wang, «Horse-Expert: An aided expert system for diagnosing horse diseases», *Pol J Vet Sci*, vol. 19, n.º 4, pp. 907-915, dic. 2016, doi: 10.1515/pjvs-2016-0112.
- [9] C. Urrea y A. Mignogna, «Development of an expert system for pre-diagnosis of hypertension, diabetes mellitus type 2 and metabolic syndrome», *Health Informatics J*, vol. 26, n.º 4, pp. 2776-2791, dic. 2020, doi: 10.1177/1460458220937095.
- [10] :: «CIMA :: Centro de información de medicamentos». <https://cima.aemps.es/cima/publico/home.html#quees> (accedido 9 de abril de 2022).

# Pósteres I

---

Miércoles, 23 de noviembre de 2022

# Implementación de Sistema de Cribado para la Audición y la Visión de Neonatos

E. Velarde-Reyes<sup>1</sup>, J.C. Santos-Ceballos<sup>1</sup>, J.G. Pérez-Blanco<sup>1</sup>, Y. Pantoja-Gómez<sup>1</sup>, A. Torres-Fortuny<sup>2</sup>, R. Cabal-Rodríguez<sup>2</sup>, B. Ramirez-Hernández<sup>3</sup>, E. Eimil-Suarez<sup>2</sup>, Ángel Regueiro-Gómez<sup>4</sup>

<sup>1</sup> Departamento de Electrónica, Centro de Neurociencias de Cuba, La Habana, Cuba, [velarde1980@gmail.com](mailto:velarde1980@gmail.com), [jsantosceballos@gmail.com](mailto:jsantosceballos@gmail.com), [jorge.perez@cneuro.cu](mailto:jorge.perez@cneuro.cu), [yosvani.pantoja@cneuro.cu](mailto:yosvani.pantoja@cneuro.cu)

<sup>2</sup> Grupo de Audiología, Centro de Neurociencias de Cuba, La Habana, Cuba, [fortuny@cneuro.cu](mailto:fortuny@cneuro.cu), [ramon.cabal@cneuro.cu](mailto:ramon.cabal@cneuro.cu), [eddy@cneuro.cu](mailto:eddy@cneuro.cu)

<sup>3</sup> Departamento de Neurofisiología, Hospital Ginecobstétrico “Ramón González Coro”, La Habana, Cuba, [belinda.ramirez@infomed.sld.cu](mailto:belinda.ramirez@infomed.sld.cu)

<sup>4</sup> Departamento de Bioingeniería, Universidad Tecnológica de La Habana (CUJAE), La Habana, Cuba, [regueiro@automatica.cujae.edu.cu](mailto:regueiro@automatica.cujae.edu.cu)

## Resumen

*La hipoacusia infantil constituye un importante problema sanitario, con una incidencia no menor de 6 por cada 1000 nacidos vivos. El padecimiento de alguno de los diferentes tipos de hipoacusia puede implicar graves trastornos en el desarrollo, aprendizaje y rendimiento escolar del niño. La prevalencia de las pérdidas visuales también es alta, con incidencia de 1 por cada 20 niños en edades pre-escolares. El diagnóstico y tratamiento precoz de la hipoacusia y de las pérdidas visuales son esenciales para la adquisición del lenguaje, el desarrollo motor, psicológico, cognitivo y social del menor. Para el diagnóstico precoz se requiere la implementación de programas de cribado neonatal que utilicen las tecnologías adecuadas. El objetivo de este trabajo es presentar la implementación del Sistema INFANTIX NA6.2, un sistema capaz de realizar pruebas de cribado auditivo y visual a neonatos. Se describen los diferentes módulos del sistema, describiendo su implementación técnica y métodos de procesamiento empleados. Finalmente se describe el proceso de validación, que incluyó ensayos técnicos y una validación clínica preliminar que brindó resultados satisfactorios.*

## 1. Introducción

La hipoacusia infantil constituye un importante problema sanitario. La Organización Mundial de la Salud indica que la incidencia de hipoacusia, en cualquiera de sus grados y en todo el mundo, es no menor de 6 por cada 1000 nacidos vivos [1]. El padecimiento de alguno de los diferentes tipos de hipoacusia puede implicar un retraso en la adquisición del lenguaje oral o la ausencia total de este, con consecuencias sobre el desarrollo del pensamiento lógico, el desarrollo de la memoria y la adquisición de la capacidad de lectura, influyendo sobre el aprendizaje y el rendimiento escolar posterior del niño. El diagnóstico tardío o fallido de la hipoacusia constituye un grave problema, pues se considera que existe un período de tiempo óptimo y no recuperable para comenzar el tratamiento o rehabilitación: antes de los 6 meses de edad [2].

La prevalencia de las pérdidas visuales es también importante. En los Estados Unidos de América, según la organización Prevent Blindness [3], es de 1 por cada 20 niños en edades pre-escolares y 1 de cada 4 niños en edades

escolares. La Academia Americana de Oftalmología (AAO) sugiere que el pesquisaje visual debe ser realizado al nacer, a los 6 meses, a los 2 años, y después cada 2 años [4]. Según registros epidemiológicos, del 2 al 4 % de los niños norteamericanos desarrollan estrabismo y/o ambliopía, por lo que su detección y tratamiento precoces son esenciales para evitar pérdidas visuales permanentes. En Cuba existen estudios donde se registraron resultados similares [5].

La detección temprana de las pérdidas visuales es importante pues el desarrollo de la agudeza visual (maduración neural y cambios conformacionales en retina) permite también el desarrollo del sistema motor extraocular indispensable en la visión. Los niños pequeños con deterioro de la visión grave de inicio temprano pueden sufrir retrasos en el desarrollo motor, lingüístico, emocional, social y cognitivo, con consecuencias para toda la vida. Los niños en edad escolar con deterioro de la visión también pueden presentar niveles más bajos de rendimiento académico.

La detección temprana de pérdidas auditivas y visuales impone la correcta implantación de programas de cribado neonatal. Para la implementación efectiva de cualquier programa de cribado auditivo y/o visual en neonatos, se requieren tecnologías adecuadas.

El objetivo de este trabajo es presentar la implementación de un sistema capaz de realizar todas las técnicas de cribado auditivo, como son las Emisiones Otoacústicas (EOA), los Potenciales Evocados Auditivos de Tallo Cerebral (PEATC) automáticos y los Potenciales Evocados Auditivos de Estado Estable (PEAEE) con posibilidades de estimulación por vía aérea y ósea de forma simultánea. Para el cribado visual se implementará la técnica de Potenciales Evocados Visuales (PEV) a Goggle de manera automática, como completamiento de las técnicas de cribado visual ya estandarizadas. Se destaca que las funciones de cribado auditivo con estimulación simultánea aérea y ósea, y de cribado visual mediante PEV a Goggle automáticos son novedosas en los equipos de cribado

neonatal. El sistema que se implementó se denominó INFANTIX NA6.2.

## 2. Diseño e implementación del Sistema INFANTIX NA6.2

El Sistema INFANTIX NA6.2 se concibió como un sistema modular, con módulos para los diferentes tipos de pruebas (ver Figura 1). De esta manera el sistema podría distribuirse como un sistema de cribado auditivo basado en la técnica de EOA, con la posibilidad de agregarse un módulo que implementara las diferentes técnicas de Potenciales Evocados Auditivos (PEA) y Visuales (PEV). La distribución del sistema, solo con el módulo de PEA y/o con el módulo de PEV, también sería posible.

La modularidad hace posible dividir el sistema en varios equipos individuales, con posibilidad de expansión hacia otras aplicaciones. Esto aporta ventajas a los usuarios, pues de esta forma podrán adquirir un equipo de cribado a un precio económico, al cual luego podrían agregar módulos según sus posibilidades o necesidades. La modularidad también aporta ventajas al equipo de desarrollo del sistema, pues siempre será más sencillo desarrollar módulos físicos independientes que aporten nuevas aplicaciones, que desarrollar un equipo o sistema completo. También será posible modificar los módulos de manera individual, ante futuras mejoras o por la obsolescencia de componentes y partes.



**Figura 1.** Sistema de Cribado Neonatal INFANTIX NA6.2. A: Módulo de Control, B: Módulo de Registro y Estímulo de Potenciales Evocados Auditivos y Visuales, C: Módulo de Emisiones Otoacústicas, D: Transductor Aéreo, E: Transductor Aéreo-Óseo and F: Estimulador Goggle.

A continuación se describen los principales módulos del sistema.

### 2.1. Módulo de Control

La función principal del Módulo de Control (ver Figura 1-A) es la de controlar el resto del sistema. Este debe ser capaz de registrar en una pantalla táctil los datos del usuario, graficar bioseñales y aplicar métodos de procesamiento digital de señales. Además, debe enviar y recibir comandos al resto de los módulos, así como administrar la energía del sistema. Este módulo se conecta al resto mediante una interfaz de comunicación serie. Se podrá conectar a un Ordenador (PC) mediante interfaz USB, para actualizar el sistema o descargar los registros de las pruebas realizadas. Esta interfaz USB también será un puerto de recarga de las baterías internas, mediante el uso de un Adaptador CA/CD.

Para realizar todas las funciones anteriores se seleccionó un microcontrolador STM32f429ZIT6 con otros componentes electrónicos asociados. Este microcontrolador dispone de un CPU Cortex-M4 con núcleo Floating Point Unit (FPU) y soporte para OS. El firmware de este módulo fue desarrollado en C++, sobre el Sistema Operativo en Tiempo Real (RTOS) Keil RTX. La implementación del firmware sobre un RTOS brinda las facilidades de la ejecución multitarea con restricciones de

tiempo que garantizan el funcionamiento del sistema en tiempo real.

### 2.2. Módulo de Registro y Estímulo de Potenciales Evocados Auditivos y Visuales

La función de este módulo (ver Figura 1-B) es la de registrar las bioseñales eléctricas del cerebro del paciente de forma sincronizada con un estímulo auditivo o visual que el mismo módulo debe generar. El módulo debe amplificar y digitalizar esas bioseñales, para luego enviarlas hacia el Módulo de Control.

El firmware del Módulo de Registro y Estímulo se ejecutará sobre el microcontrolador STM32F405RGT6. Este microcontrolador, al igual que el del Módulo de Control, también dispone de un CPU Cortex-M4 con núcleo Floating Point Unit (FPU) y soporte para OS. Su firmware fue desarrollado usando el mismo sistema operativo.

Para la estimulación auditiva y visual, se requiere la conexión de transductores auditivos aéreos (ver Figura 1-D), aéreos-óseos (ver Figura 1-E) o de un dispositivo basado en LED rojo conocido como Goggle (ver Figura 1-F), especialmente diseñado para recién nacidos.

En el Módulo de Registro y Estímulo se ejecutan algoritmos para la detección automática de PEATC y PEV.

Estos algoritmos están basados en la comparación de la señal promediada, luego de un proceso de promediación sincrónica de segmentos de señal de 10 ms de duración, con plantillas de la señal de potencial evocado y la medición de su energía [6]. Los segmentos de señal se enumeran y se dividen en grupos de segmentos pares e impares, para realizar una promediación parcial de estos segmentos que se conoce como hemipromedio. Para reforzar los algoritmos se agregan medidas que verifican la calidad de la señal registrada, como son la medición de la correlación entre los hemipromedios pares e impares, y el cálculo del ruido residual.

Para la detección de los PEAE se implementó el algoritmo basado en el cálculo de la T2 de Hotelling (T2H) implementado para el cribado auditivo sobre el equipo AUDIX 5, evaluado por Mijares-Nodarse y colaboradores [7, 8]. Debe destacarse que en el Sistema INFANTIX NA6.2 se replicaron los mismos valores de frecuencia de muestreo y frecuencias de corte de registro del equipo AUDIX 5.

### 2.3. Módulo de Emisiones Otoacústicas

La función de este módulo (ver Figura 1-C) es la de registrar las emisiones otoacústicas generadas por la cóclea del paciente, luego de una adecuada estimulación auditiva. Este módulo también debe digitalizar esas bioseñales y enviarlas hacia el Módulo de Control. Las modalidades de registro de emisiones otoacústicas que este módulo implementa son las de Producto de Distorsión (EOAPD) y las Transientes (EOAT).

El módulo cuenta con una sonda donde se encuentran ubicados un micrófono de alta sensibilidad para el registro y dos altavoces para la estimulación. Se requiere de dos altavoces para poder implementar tanto la técnica de EOAT, como la de EOAPD, ruido. Para mejorar la relación señal-ruido, la señal del micrófono es amplificada mediante un amplificador integrado dentro de la sonda. Antes de iniciar cada prueba, siempre se evalúa de forma automática, el estado de colocación de la sonda (afuera, adentro u ocluido).

## 3. Validación

La validación del Sistema INFANTIX NA6.2 se dividió en dos etapas: Ensayos Técnicos y Validación Clínica.

En los Ensayos Técnicos se evaluó por un laboratorio acreditado para estos fines y con resultados satisfactorios, el cumplimiento de la normativa general de seguridad UNE-EN 60601-1:2008 (Equipos electromédicos. Parte 1: Requisitos generales para la seguridad básica y funcionamiento esencial). También fue evaluado el cumplimiento de las diferentes normativas particulares, como es el caso de la IEC 60601-2-40:2016 (requerimientos particulares para la seguridad básica y funcionamiento esencial de los equipos de respuesta evocada y electromiografía) y la EN 60645-6:2010 (instrumentos de medición de emisiones Otoacústicas).

En las próximas secciones se expone el proceso de validación preliminar de los diferentes módulos del Sistema INFANTIX NA6.2. Debe destacarse que en todas

las pruebas realizadas se contó con el Consentimiento Informado de los sujetos, o de sus padres o tutores. En todo momento se veló por el cumplimiento de los principios contenidos en la Declaración de Helsinki [9].

### 3.1. Validación preliminar de la función de cribado utilizando Potenciales Evocados Auditivos

Se realizó una validación preliminar del Sistema INFANTIX NA6.2, en su función de cribado auditivo mediante PEATC automático, en 20 neonatos nacidos en el Hospital Ginecobiológico “Ramón González Coro” de La Habana, en los meses de enero a abril del 2019. Como *gold standard* se utilizaron las pruebas de PEATC realizadas con el Electroaudiómetro AUDIX 5, Modelo ND001A, del fabricante IC Neuronics SL.

Los estímulos utilizados en las pruebas de PEATC-a fueron clics de 100  $\mu$ s de ancho. La intensidad de los estímulos fue de 30 dB HL y la frecuencia de presentación de 71 Hz. Estos estímulos fueron previamente calibrados usando un Sonómetro Brüel & Kjær (Modelo 2250), acoplado a una oreja artificial Modelo 4152 con micrófono Brüel & Kjær (Modelo 4144). Todas las calibraciones fueron realizadas según la norma ISO 389-6:2007.

Con la muestra estudiada (40 oídos, 11 hembras y 9 varones), se obtuvieron valores de Especificidad/Sensibilidad de 100/100 %, con tiempos de prueba por oído de entre 10 y 38 segundos, para un tiempo promedio de 24 segundos. Debe señalarse que para simular una condición de hipoacusia, a 7 de los neonatos del estudio seleccionados aleatoriamente, les fue realizada la prueba sin aplicar estimulación auditiva.

La validación de los algoritmos y métodos de cribado auditivo mediante PEAE fue realizada por Mijares-Nodarse y colaboradores. Esta validación se realizó con el equipo Audix 5 y se estudiaron 80 bebés con factores de riesgo de pérdidas auditivas, con resultados de Especificidad/Sensibilidad de 97.7/100 %, con tiempos de prueba por oído de entre 84 y 480 segundos, para un tiempo promedio de 318 segundos [8].

### 3.2. Validación preliminar de la función de cribado utilizando Emisiones Otoacústicas

Se desarrolló un estudio con un total de 140 sujetos voluntarios, de ambos sexos, con edades entre 18 y 72 años. Se seleccionó un grupo con este rango de edades para la validación preliminar teniendo en cuenta criterios éticos, puesto que estos sujetos no suelen pertenecer a grupos vulnerables. Debe tenerse en cuenta además, que en estudios anteriores se observó que las señales de emisiones otoacústicas de los neonatos e infantes, tienen mayor amplitud a las altas frecuencias que las señales emisiones otoacústicas de adultos [10]. Por tanto, los resultados que se obtengan en una validación con un grupo de neonatos, siempre deberían ser mejores a los obtenidos con adultos. Siendo esta una hipótesis a demostrar en futuras evaluaciones.

En el estudio se evaluó el Coeficiente de Kappa de Cohen, del resultado de la prueba de cribado (Pasó o Referir) del Sistema INFANTIX NA6.2, comparándola con el

resultado del Sistema Titan del fabricante Interacoustics, con su módulo de EOA con referencia *Preamplifier*.

Se realizaron registros de las EOAT en ambos oídos de cada paciente, alcanzándose un total de 280 pruebas. Durante cada uno de los registros se aseguró que el ruido ambiente no superara los 50 dBA.

Como resultado del estudio se obtuvo un Coeficiente de Kappa de 0.68, equivalente a un grado de acuerdo Sustancial, entre las respuestas de ambos sistemas [11].

### 3.3. Validación de la función de cribado utilizando Potenciales Evocados Visuales

Para la validación preliminar de la función de cribado del Sistema INFANTIX NA6.2, empleando Potenciales Evocados Visuales, se utilizó el Módulo de Registro y Estimulo de Potenciales Evocados Auditivos y Visuales, con el Goggle a led rojo monocular.

El estudio se realizó con 20 sujetos sanos (10 hembras y 10 varones), con más de 28 semanas de edad gestacional y hasta 3 meses de nacidos.

Se utilizaron como *gold standard* los resultados de las pruebas de Potenciales Evocados Visuales realizadas previamente con el equipo de Electromiografía y Potenciales Evocados NEURONICA 5, Modelo N6.5, del fabricante IC Neuronic SL.

El estudio fue realizado en un local oscuro, con la mínima incidencia de luz ambiente. Con el fin de dilatar las pupilas y facilitar la prueba, a los niños se les aplicó una gota del midriático Tropicamida al 1%, en cada ojo, 15 minutos antes del estudio.

En los estudios realizados se empleó como estímulo un destello luminoso con una frecuencia de 1 Hz. El estímulo se presentó en cada ojo por separado. Para minimizar las molestias a los sujetos, el tiempo de prueba para cada ojo, fue menor de 2 min en toda la muestra.

Todos los sujetos estudiados resultaron sanos, con latencias de P1 correspondientes a su edad. Con el Sistema INFANTIX NA6.2 se evaluó además de manera automática la diferencia de amplitudes interoculares intrasujeto, obteniéndose resultados positivos en todos los casos.

## 4. Conclusiones y Futuros Trabajos

En este trabajo ha sido presentada la implementación de un sistema de cribado para la audición y la visión de neonatos, el cual fue nombrado Sistema INFANTIX NA6.2.

De acuerdo a los resultados obtenidos en las pruebas técnicas y validaciones preliminares de los diferentes módulos, el sistema se impone como una valiosa herramienta para el cribado de neonatal.

En el momento de la escritura de este trabajo (septiembre 2022), se encuentra en ejecución el Ensayo Clínico titulado “Evaluación de Rendimiento Clínico del Sistema de Cribado Neonatal NEURONIC / INFANTIX A6.2 en un Hospital Materno”. Este ensayo se encuentra inscrito en el Registro Público Cubano de Ensayos Clínicos (RPCEC) con el código RPCEC00000405 y evaluará el rendimiento

clínico de los diferentes módulos del sistema con una muestra de al menos 300 recién nacidos.

## Agradecimientos

Se expresa un especial agradecimiento a todas las madres y padres que consintieron la participación de sus hijos recién nacidos en las validaciones preliminares del Sistema INFANTIX NA6.2.

## Financiamiento

Este proyecto fue financiado por el Fondo Financiero de Ciencia e Innovación (FONCI) del Ministerio de Ciencia, Tecnología y Medio Ambiente (CITMA) de la República de Cuba.

## Referencias

- [1] B. Olusanya, S. Wirz, and L. Luxon. (2008, June 19th, 2019). Community-based infant hearing screening for early detection of permanent hearing loss in Lagos, Nigeria: a cross-sectional study. *Bulletin of the World Health Organization*. Available: <http://www.who.int/bulletin/volumes/86/12/07-050005/en/>
- [2] World Health Organization, "Childhood hearing loss: strategies for prevention and care," 2016.
- [3] Prevent Blindness. (2010, December 2021). *Prevent Blindness America: The Sight Saving People*. Available: [https://preventblindness.org/wp-content/uploads/2011/12/FS14\\_SightSavers.pdf](https://preventblindness.org/wp-content/uploads/2011/12/FS14_SightSavers.pdf)
- [4] American Academy of Ophthalmology. (2020, December 2021). *Exámenes oftalmológicos de detección para niños*. Available: <https://www.aaopt.org/salud-ocular/consejos/tamizaje-oftalmologico-para-ninos>.
- [5] O. Capetillo Biart, I. Triana Casado, Z. d. I. C. Martínez Legón, S. Roche Caso, and A. Broche Hernández, "Frecuencia de la ambliopía en escolares," *Revista cubana de pediatría*, vol. 83, pp. 372-381, 2011.
- [6] I. M. Cabana-Pérez, E. Velarde-Reyes, A. Torres-Fortuny, E. Eimil-Suarez, and A. García-Giró, "Automatic ABR detection at near-threshold intensities combining template-based approach and energy analysis," in *VII Latin American Congress on Biomedical Engineering CLAIB 2016*, Bucaramanga, Santander, Colombia., 2016.
- [7] E. M. Nodarse, D. H. Alonso, J. G. Vázquez, E. S. Febles, M. C. P. Abalo, L. M. Alarcón, *et al.*, "Cribado auditivo neonatal con potenciales evocados auditivos de estado estable a múltiples frecuencias," *Acta Otorrinolaringológica Española*, vol. 62, pp. 87-94, 2011.
- [8] E. Mijares, L. Báez, L. Cabrera, M. C. Pérez-Abalo, and A. Torres-Fortuny, "Cribado auditivo mediante potenciales evocados auditivos de estado estable obtenidos por estimulación simultánea de la vía aérea y la ósea," *Acta Otorrinolaringológica Española*, vol. 66, pp. 8-15, 2015.
- [9] A. M. Mundial, "Declaración de Helsinki de la AMM-Principios éticos para las investigaciones médicas en seres humanos," 2019.
- [10] M. C. Stamate, C. I. Bondor, and M. Cosgarea, "Differences in otoacoustic emissions in infants and adults with normal hearing," *Human and Veterinary Medicine*, vol. 7, pp. 126-134, 2015.
- [11] V. Abaira, "El índice kappa," *Semergen-Medicina de Familia*, vol. 27, pp. 247-249, 2001.

# Fabricación de fibras de alta resistencia obtenidas a partir de glándulas sericígenas de *Samia cynthia ricini*. Caracterización y comparación con *Bombyx mori* y potenciales aplicaciones en Ingeniería Tisular

M.J. Candel<sup>1,2</sup>, A. Pagan<sup>2</sup>, J. Pérez-Rigueiro<sup>3</sup>, M.L. Arnal<sup>1,4</sup>, J. L. Cenis<sup>2</sup>, S. D. Aznar<sup>2</sup>

<sup>1</sup> Escuela Superior de Ingeniería, Ciencia y Tecnología, Universidad Internacional de Valencia (VIU), 46002 Valencia, España [mjcandelc@gmail.com](mailto:mjcandelc@gmail.com) <sup>2</sup> Departamento de Biotecnología, Genómica y Mejora Vegetal, Instituto Murciano de Investigación y Desarrollo Agrario y Ambiental (IMIDA), 30150, La Alberca, Murcia, España. [sdac1@um.es](mailto:sdac1@um.es) [anapagan@um.es](mailto:anapagan@um.es) [josel.cenis@carm.es](mailto:josel.cenis@carm.es) <sup>3</sup> Centro de Tecnología Biomédica, Universidad Politécnica de Madrid, Pozuelo de Alarcón, 28223, Madrid, España. [jose.perez@ctb.upm.es](mailto:jose.perez@ctb.upm.es) <sup>4</sup> Dpto. de Ciencia de los Materiales Universidad Simón Bolívar, Caracas, 89000, Venezuela, [marnal@usb.ve](mailto:marnal@usb.ve)

## Resumen

*El uso de la fibroína de la seda como biomaterial, dadas sus excelentes propiedades mecánicas, de biocompatibilidad y biodegradabilidad, lo erige como un material con gran potencial para el desarrollo de dispositivos médicos, con el fin de fabricar armazones o andamios (scaffolds) para el crecimiento de células y su diferenciación en tejidos u órganos implantables.*

*En este trabajo, los autores presentan la obtención y caracterización de las fibras derivadas de las glándulas sericígenas de *Samia cynthia ricini* o gusano eri. Las glándulas se sumergieron en disoluciones acuosas de distintas concentraciones de ácido acético y se sometieron a tracción para fabricación de las fibras. Los materiales producidos se caracterizaron, en términos de morfología, propiedades mecánicas, cristalinidad y contenido de estructuras secundarias, comparándolos a su vez, con los producidos por el procedimiento estándar publicado para *Bombyx mori* [1]*

*Los resultados obtenidos muestran que las fibras de *B. mori* presentan un mayor grado de cristalinidad que las de *S. c. ricini*, lo que se refleja en valores más altos de módulo elástico y valores más bajos de deformación a la rotura. Además, se observó una disminución en los valores de elongación de las fibras de *S. c. ricini* a medida que aumentaba la concentración de ácido acético empleada para su obtención.*

## 1. Introducción y objetivos

La fibroína de seda (FS) es un material textil antiguo y de lujo que no se limita a la sericultura, sino a lo largo de los años, se ha expandido ampliamente en la investigación de biomateriales y en las industrias biomédicas de alta tecnología, aportando un valor adicional a las aplicaciones presentes y futuras. También, se ha utilizado la FS como sedales para pesca, dada su gran resistencia a tracción

(hasta la aparición del nylon que, prácticamente, lo sustituyó) y como material biomédico; como ejemplos, el hilo de sutura para ligadura de heridas o las telas de seda para el tratamiento de afecciones dermatológicas.

Una de las razones principales para la intensa investigación de los últimos treinta años de la fibroína de seda (FS) como biomaterial es su potencial uso en dispositivos biomédicos debido a sus excelentes propiedades mecánicas, su óptima biocompatibilidad y las amplias y versátiles formas que pueden adaptar sus diversas estructuras entre ellas cuerdas *films*, fibras, membranas, hidrogeles, nanopartículas y materiales electrohilados [2].

La FS de *S. c. ricini* presenta una estructura primaria compuesta por 100 repeticiones de dominios alternos de poli-(L)-alanina (PA) y glicina, mientras que la FS *B. Mori* se compone principalmente de residuos de glicina, alanina y serina, con un mayor contenido en glicina que en el caso de la fibroína de *S. c. ricini* [3]. Estas diferencias en la composición repercuten en las propiedades mecánicas de las sedas nativas y los biomateriales de FS derivados de ellas, así como en su funcionalidad [4]. Existen algunos indicios de posibles ventajas en términos de adhesión, proliferación celular y comportamiento mecánico en fibras y películas de fibroína producidas a partir de gusano eri comparándola con las que utilizan FS *B. mori* [5],[6].

Los objetivos de este trabajo son:

- El diseño de una metodología de fabricación de fibras de alta resistencia procedentes de *S.c. ricini*; así como la obtención de fibras de alta resistencia empleando diferentes concentraciones de ácido acético.
- La caracterización físico-química y comparación de las fibras procedentes de *S.c. ricini*. y *B. mori* mediante microscopía óptica, espectroscopía infrarroja y

ensayos tensiles a fin de conocer su morfología, grado de cristalinidad y desempeño mecánico

## 2. Material y Métodos

### 2.1. Crianza del gusano *Samia cynthia ricini* y fabricación de fibras de alta resistencia

Los gusanos *S.c. ricini* se alimentaron con *Ricinus communis*, en IMIDA (Murcia), durante 35 días, hasta el final del quinto estado larvario, momento en el que comenzarían a hilar el capullo. Se seleccionaron los ejemplares en base a su peso (aproximadamente 7 gr) para la obtención de las glándulas sericígenas.

Los ejemplares seleccionados, se anestesiaron bajando su temperatura corporal a 4°C y se extrajeron, por disección, las glándulas sericígenas. Se fabricaron las fibras o hijuelas mediante un proceso de inmersión de la glándula en una disolución acuosa de ácido acético de concentración variable (2, 10 y 20 % v-v) durante 2 minutos. Después de un lavado con agua destilada, el material se sometió a tracción manual. Posteriormente, se procedió al lavado, secado y cortado de las fibras obtenidas (algunas de más de 3 metros) en trozos de 30 cm de longitud. El material de *B. mori* fue aportado por IMIDA, obtenidos por un procedimiento similar [1]. También se caracterizaron las sedas nativas (SN) de ambas especies.

### 2.2. Evaluación de la morfología de la fibra

Se utilizó un microscopio óptico de campo claro, Nikon eclipse 50i con cámara Nikon DS-F1, se fotografiaron las fibras de *S.c. ricini* y *B. mori* con aumentos 5x, 10x y 20x y se pudieron procesar y fotografiar con detalle con el software NIS-Elements F3.0.

### 2.3. Caracterización de fibras mediante espectroscopía infrarrojo transformada de Fourier de reflexión total atenuada (ATR-FTIR)

Se empleó ATR-FTIR para analizar las posibles diferencias en la superestructura de las fibras de *S.c. ricini* y de *B. mori*. Cada espectro se obtuvo en un espectrómetro Nicolet iS5, equipado con un accesorio iD5 ATR (Thermo Scientific, EE. UU.), controlado con el software OMNIC (Ver. 9.3.30), midiendo en modo de absorbancia con una resolución de 4 cm<sup>-1</sup>, un rango espectral de 4000–550 cm<sup>-1</sup> y 64 barridos.

#### 2.3.1 Cálculo del Índice de cristalinidad.

El índice de cristalinidad se calculó con el cociente de los valores de absorbancia de las láminas β (estructuras cristalinas) y la fracción amorfa [7]. El estudio detallado de ciertas bandas de carácter conformacional permite estudiar la cristalinidad de las fibras mediante su espectro IR.

El análisis fue enfocado en la región amida I (calculado en el rango de los valores de medida de absorbancia entre 1625 a 1650 cm<sup>-1</sup>) y la región amida II (entre los valores de absorbancia de 1515 a 1538 cm<sup>-1</sup>), para establecer el predominio de estructuras cristalinas insolubles sobre estructuras amorfas hidrosolubles y menos organizadas [7]. Se determinaron las intensidades de las señales asociadas a la fracción de cristalinidad (láminas β) en 1625 cm<sup>-1</sup> (en el

eje X del espectro) y de la fracción amorfa en 1650 cm<sup>-1</sup> (*Random coil* o cadenas conformación de ovillo al azar) [8].

#### 2.3.2 Procesado de los espectros por deconvolución

En este trabajo se realizó el procesado de datos por deconvolución con el fin de investigar más profundamente el contenido de estructuras secundarias de la FS. El análisis fue finalmente enfocado en la región amida I (los valores de medida de absorbancia entre 1703 y 1594 cm<sup>-1</sup>). La deconvolución se realizó, mediante la función resolución de picos, utilizando 9 funciones Gaussianas para el ajuste del pico amida I. Con una sensibilidad media, un FWHH de 8, un objetivo de ruido de 0,1 y la selección la opción de la segunda derivada se consideraron como condiciones iniciales del ajuste. Con el fin de medir las áreas relativas de los componentes estructurales del pico amida I, se ajustaron las curvas de los espectros de la deconvolución. Las posiciones (en cm<sup>-1</sup>) de la banda de los máximos en los espectros de la deconvolución se hicieron para corresponder a la frecuencia de los mínimos en la segunda derivada de los espectros. Los puntos mínimos de la segunda derivada se utilizaron para determinar el número y la posición de las funciones gaussianas utilizadas para el ajuste. Finalmente, los espectros de amida I de la deconvolución se normalizaron en área, y las áreas relativas de las bandas simples se utilizaron para calcular la fracción (%) de los elementos estructurales secundarios, como se muestra en la Figura 1. Las asignaciones de bandas se basaron en los datos resumidos en estudios publicados [9], [10]. Las áreas de los gaussianos se utilizaron para estimar el contenido de las diferentes estructuras secundarias.

#### 2.4. Evaluación de propiedades mecánicas mediante ensayos de tracción.

Las pruebas de tracción se realizaron utilizando una máquina de marco de prueba electromecánica universal, Shimadzu EZ-SX 200N. Las propiedades mecánicas de las probetas, de diámetro variable y longitud entre mordazas de 20 mm, se registraron con una velocidad de cruceta de 0,1 mm/s<sup>-1</sup> y una celda de carga de 200 N, en condiciones ambientales procesados con el software TRAPEZIUM (Versión 1.4.0). El módulo elástico (GPa), la resistencia a la tracción (MPa) y la deformación a la rotura (%) se determinaron utilizando las curvas de esfuerzo-deformación. El módulo de elasticidad se calculó en la parte elástica lineal de las curvas.

#### 2.5. Análisis estadístico.

Con el software IBM SPSS Statistics, se detectaron si había o no diferencias significativas en los datos obtenidos en la caracterización mecánica, índices de cristalinidad y deconvolución. Se comprobó que los datos de los tratamientos sí cumplían los requisitos de normalidad (Kolmogórov-Smirnov,  $p > 0.05$ ) y homogeneidad de varianzas (Levene,  $p > 0.05$ ). Posteriormente, se compararon globalmente todos los tratamientos con el test paramétrico ANOVA ( $p < 0.05$ ). La comparativa por tratamientos dos a dos, se llevó a cabo mediante el test de Bonferroni ( $p < 0.05$ ).

### 3. Resultados y Discusión

A continuación, se presentan los resultados más relevantes tras una extensa caracterización de fibras de FS obtenidos a partir del gusano de seda eri y de *B. mori*. Las fibras de *S. c. ricini* se denominarán, en adelante, (según la disolución de ácido acético utilizada en su fabricación) Scr FS Ac 2%, Scr FS Ac 10% o Scr FS Ac 20%. Las fibras de *B. mori*, Bm FS Ac 2% y las sedas nativas corresponderán a Scr SN y Bm SN respectivamente.

#### 3.1. Crianza del gusano *Samia cynthia ricini* y fabricación de fibras de alta resistencia

Obtenidas las fibras de *S.c. ricini* y *B. mori*, se estudiaron sus dimensiones (Tabla 1). Existe una importante diferencia tanto entre los diámetros, como en su longitud. Las fibras de *S.c. ricini* llegaron a medir 3,5 m (variando y dependiendo del ejemplar de estudio y la manipulación) frente a las de *B. mori* que tienen una longitud de 50-60 cm [1]. El diámetro de las fibras Bm F S AC 2% es mayor que los obtenidos de *S.c. ricini* en cualquier condición.

	Scr FS Ac 2%	Scr FS Ac 10%	Scr FS Ac 20%	Bm FS Ac 2%
Diámetro (µm)	133 ± 6	156 ± 3	128 ± 21	248 ± 46

Tabla 1. Valores medios de los diámetros de las fibras en seco (valores expresados como valor medio ± desviación estándar)

#### 3.2. Evaluación de la morfología de las fibras

Se observaron las fibras de *S.c. ricini* y *B. mori*, con el microscopio óptico obteniendo las imágenes de la Figura 1, donde se aprecia un cierto incremento de la rugosidad con el contenido en ácido acético.



Figura 1. Imágenes de las fibras en seco de los distintos tratamientos observadas con el microscopio óptico a microescalas 200 y 100 µm

#### 3.3. Caracterización de fibras mediante espectroscopía infrarrojo transformada de Fourier de reflexión total atenuada (ATR-FTIR)

##### 3.3.1 Cálculo del Índice de cristalinidad.

Los resultados obtenidos en el cálculo del índice de cristalinidad (Tabla 2) muestran que la hijuela de *B. mori* es mucho más cristalina que *S.c. ricini* en cualquiera de sus tratamientos. Curiosamente, es más cristalina la seda de *S.c. ricini* que sus propias hijuelas. Al contrario de lo que

sucede con la seda de *B. mori*, que resulta menos cristalina que su hijuela.

	Índice de Cristalinidad	
	Amida I	Amida II
Scr FS Ac 2%	1.079 ± 0.108	1.038 ± 0.040
Scr FS Ac 10%	1.100 ± 0.047	1.043 ± 0.023
Scr FS Ac 20%	1.125 ± 0.054	1.065 ± 0.052
Bm FS Ac 2%	1.333 ± 0.089	1.122 ± 0.059
Scr SN	1.203 ± 0.041	1.032 ± 0.043
Bm SN	0.947 ± 0.025	1.011 ± 0.007

Tabla 2. Resultados de los índices de cristalinidad

Respecto al análisis estadístico, ANOVA detecta que el pico amida I muestra diferencias significativas en cuanto al índice de cristalinidad (ANOVA,  $p < 0.05$ ) pero no el pico amida II (ANOVA,  $p > 0.05$ ). La hijuela de *B. mori* es distinta a todas las de *S.c. ricini* en cuanto a su grado de cristalinidad siendo un valor significativamente superior a las de *S.c. ricini* (Bonferroni,  $p < 0.05$ ). La hijuela de *B. mori* tiene el mayor valor de índice de cristalinidad, incluso que su propia seda. La cristalinidad de la seda de *S.c. ricini* es significativamente mayor que de la seda *B. mori* (Bonferroni,  $p < 0.05$ ).

##### 3.3.2 Procesado de los espectros por deconvolución

Los resultados de proceso de deconvolución (Figura 2) muestra la distribución de las estructuras secundarias en todas las muestras.

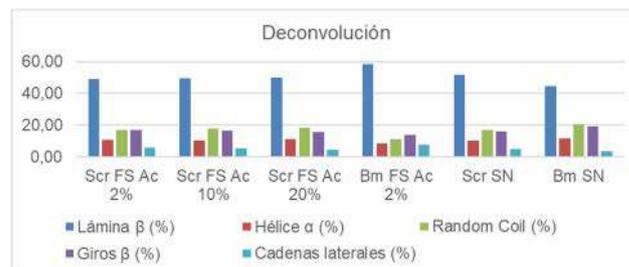


Figura 2. Distribución de estructuras secundarias por tratamientos

La comparación estadística dos a dos entre todos los grupos de estudio reveló diferencias en el contenido en láminas β de la hijuela de *B. mori* siendo significativamente mayor a las hijuelas de *S.c. ricini* (Bonferroni,  $p < 0.05$ ), como sucede con el índice de cristalinidad por lo que estos resultados entran en concordancia con los mostrados en la Tabla 2. El contenido en *random coil* o estructuras aleatorias es significativamente menor en el caso de las hijuelas de *B. mori* al compararlo con las obtenidas con las de *S.c. ricini* (Bonferroni,  $p < 0.05$ ). Comparando la seda de *B. mori* y *S.c. ricini*, el contenido de láminas β en la seda de *B. mori* es menor que en *S.c. ricini*, por lo que su seda tiene menor cristalinidad, como también corroboran los datos obtenidos en el índice de cristalinidad. Los valores de contenido en láminas β de las hijuelas de *S.c. ricini* son iguales entre ellas independientemente del tratamiento. En el apartado 3.4. se expondrán las consecuencias en sus propiedades mecánicas.

### 3.4. Evaluación de propiedades mecánicas mediante ensayos de tracción.

Los resultados (Tabla 3) de los valores de tensión máxima antes de rotura de la fibra de *S.c. ricini* con tratamiento con ácido acético alcanzan valores medios de hasta 388 MPa, aunque no se aprecia una tendencia clara puesto que los valores en todos los tratamientos entran dentro del intervalo dado por la desviación estándar. La misma falta de tendencia se observa en los valores del Módulo de Young, con valores de entre 8,7-9 GPa. Respecto al valor de deformación máxima, el mayor valor se alcanza con el tratamiento al 2% de ácido acético, la elongación disminuye con el aumento de la concentración de ácido acético. La fibra de *B. mori*, es más rígida y menos deformable que las de *S.c. ricini*.

	Tensión máxima (MPa)	Máxima deformación (%)	Módulo elástico (GPa)
Scr FS Ac 2%	388 ± 12	65.0 ± 2.5	8.7 ± 1.0
Scr FS Ac 10%	387 ± 39	39.2 ± 3.6	8.7 ± 0.3
Scr FS Ac 20%	305 ± 64	28.0 ± 0.6	9.0 ± 1.5
Bm FS Ac 2%	306 ± 78	7.9 ± 3.3	11.6 ± 0.9

**Tabla 3.** Valores medios de las propiedades mecánicas de las fibras en seco

Respecto a la estadística, ANOVA detecta globalmente diferencias en máxima deformación y en módulo elástico para algún tratamiento (ANOVA,  $p < 0.05$ ), sin embargo, los valores de tensión máxima son estadísticamente equivalentes en todas las hijuelas (ANOVA,  $p > 0.05$ ). Se revelaron diferencias en los valores de máxima deformación entre todos los grupos de estudio (*Bonferroni*,  $p < 0.05$ ), siendo significativamente mayor en las hijuelas de *S.c. ricini* con acético al 2% respecto al resto de tratamientos. En el módulo elástico, los valores de la hijuela de *B. mori* son distintos a las hijuelas producidas de *S.c. ricini* empleando acético al 2% y al 10% (*Bonferroni*,  $p < 0.05$ ) e iguales a las hijuelas de *S.c. ricini* fabricada con acético al 20% (*Bonferroni*,  $p > 0.05$ ).

## 4. Conclusiones

La síntesis y caracterización de las fibras permite concluir que:

-. La comparación de las propiedades mecánicas se tiene que la hijuela de *S.c. ricini* es considerablemente más elongable que la de *B. mori* con el mismo tratamiento de acidificación al 2% dado que su módulo de *Young* es significativamente menor. Esta tendencia se mantiene para todos los contenidos de ácido acético evaluados

-. Tras el estudio por espectroscopía infrarroja de la estructura química y grado de cristalinidad de ambas fibras y procesado de sus espectros por deconvolución, se puede concluir que las hijuelas de *B. mori* poseen el índice de cristalinidad mayor y mayor contenido en láminas  $\beta$  que las de *S.c. ricini*. La seda de *S.c. ricini* es más cristalina que sus hijuelas, al contrario de lo que sucede con la seda *B. mori*.

## Agradecimientos

Agradecer a IMIDA por la realización de los experimentos, discusión de resultados y publicaciones derivadas [4][6].

## Referencias

- [1] J. L. Cenis, R. Madurga, S. Aznar-Cervantes, A. Lozano-Pérez, N. Mari-Buyé, L. Meseguer-Olmo, G. Plaza, G. Guinea, M. Elices, F. Del Pozo Guerrero, J. Pérez-Rigueiro, «Mechanical behaviour and formation process of silkworm silk gut», *Soft Matter*, vol. 11, n.o 46, pp. 8981-8991, sep. 2015, doi: 10.1039/c5sm01877c.
- [2] D. N. Rockwood, R. C. Preda, T. Yücel, X. Wang, M. L. Lovett, y D. L. Kaplan, «Materials fabrication from *Bombyx mori* silk fibroin», *Nat Protoc*, vol. 6, n.o 10, pp. 1612-1631, oct. 2011, doi: 10.1038/nprot.2011.379.
- [3] Y. Nakazawa, «Tightly winding structure of sequential model peptide for repeated helical region in *Samia cynthia ricini* silk fibroin studied with solid-state NMR», *Protein Science*, vol. 12, n.o 4, pp. 666-671, abr. 2003, doi: 10.1110/ps.0239203.
- [4] S. D. Aznar-Cervantes, A. Pagán, M. J. Candel, J. Pérez-Rigueiro, y J. L. Cenis, «Silkworm Gut Fibres from Silk Glands of *Samia cynthia ricini*. Potential Use as a Scaffold in Tissue Engineering», *Int J Mol Sci*, vol. 23, n.o 7, abr. 2022, doi: 10.3390/IJMS23073888.
- [5] K. Mai-Ngam, K. Boonkitpattarakul, J. Jaipaw, y B. Mai-Ngam, «Evaluation of the properties of silk fibroin films from the non-mulberry silkworm *Samia cynthia ricini* for biomaterial design», *J Biomater Sci Polym Ed*, vol. 22, n.o 15, pp. 2001-2022, 2011, doi: 10.1163/092050610X530964.
- [6] S. D. Aznar-Cervantes, J.L. Cenis, P. Lozano-Picazo, A.L. Bruno, A. Pagán, Y. Ruiz-León, M.J. Candel, D. González-Nieto, F.J. Rojo, M. Elices, G.V. Guinea J. Pérez-Rigueiro., «Unexpected high toughness of *Samia cynthia ricini* silk gut», *Soft Matter*, 2022, doi: 10.1039/d2sm00340f.
- [7] S. D. Wang, Y. Z. Zhang, G. B. Yin, H. W. Wang, y Z. H. Dong, «Fabrication of a composite vascular scaffold using electrospinning technology», *Materials Science and Engineering C*, vol. 30, n.o 5, pp. 670-676, jun 2010, doi: 10.1016/j.msec.2010.02.021.
- [8] W. C. Jao, M. C. Yang, C. H. Lin, y C. C. Hsu, «Fabrication and characterization of electrospun silk fibroin/TiO<sub>2</sub> nanofibrous mats for wound dressings», *Polym Adv Technol*, vol. 23, n.o 7, pp. 1066-1076, jul. 2012, doi: 10.1002/pat.2014.
- [9] R. Lacalle, «Development of a new biomimetic approach for the spinning of regenerated silkworm silk fibers», Tesis, Universidad P.Madrid, Madrid, España, 2015.
- [10] R. Madurga, A. Gañán-Calvo, G. R. Plaza, G. Guinea, M. Elices, y J. Pérez-Rigueiro, «Straining flow spinning: Production of regenerated silk fibers under a wide range of mild coagulating chemistries», *Green Chemistry*, vol. 19, n.o 14, pp. 3380-3389, 2017, doi: 10.1039/c7gc01254c.

# Simulación paralela multidispositivo de secuencias de imagen por resonancia magnética

I. Fernández-Arias<sup>1</sup>, P. Villacorta-Aylagas<sup>1</sup>, C. Castillo-Passi<sup>2</sup>, P. Irarrazaval<sup>2</sup>,  
F. Simmross-Wattenberg<sup>1</sup>, M. Rodríguez-Cayetano<sup>1</sup>, C. Alberola-López<sup>1</sup>

<sup>1</sup> Laboratorio de Procesado de Imagen. ETSI Telecomunicación. Universidad de Valladolid.

<sup>2</sup> Instituto de Ingeniería Biológica y Médica. Pontificia Universidad Católica de Chile.

## Abstract

*La simulación del fenómeno de resonancia magnética (RM) es de gran interés tanto en la investigación científica como en la formación de personal técnico. En la literatura hay documentados varios simuladores de RM, pero muchos de ellos no son suficientemente realistas, mientras que otros solamente funcionan en CPUs o en GPUs de un único fabricante. Uno de estos simuladores, KomaMRI.jl, simula de forma fidedigna el fenómeno de RM y funciona en CPU y GPU (estas últimas de un único fabricante), pero no se ha diseñado con especial énfasis en la rapidez de la simulación. La conjunción de velocidad y realismo es crucial para simuladores interactivos y para investigación en secuencias de RM. En este trabajo se presenta un simulador de RM que funciona sobre dispositivos de cualquier tipo y fabricante, y cuyo rendimiento supera al mencionado simulador KomaMRI.jl, a la vez que mantiene su grado de realismo.*

## 1. Introducción

La imagen de resonancia magnética (MRI) es una de las modalidades de imagen médica más potentes entre las disponibles hoy en día. Su alto contraste en los llamados “tejidos blandos” y su gran versatilidad para la obtención de imagen anatómica y funcional la convierten en la técnica preferida en muchos ámbitos clínicos.

Desde una perspectiva educativa, en cambio, manejar un escáner de MRI no es tarea fácil, puesto que las adquisiciones dependen fuertemente de un gran número de parámetros que han de ser calibrados oportunamente para obtener una calidad de la imagen óptima y, en especial, para evitar la aparición de artefactos. Los parámetros a calibrar dependen de las secuencias de pulsos magnéticos que se programan en el escáner, existiendo una amplia variedad de secuencias y siendo, en la actualidad, un campo abierto de investigación que requiere de un uso intensivo de escáneres reales. Sin embargo, las demandas del ámbito clínico y los altos costes suponen que los escáneres sean utilizados constantemente, limitando las posibilidades de programar sesiones de aprendizaje para los técnicos en formación o para la investigación de nuevas secuencias de pulsos. Consecuentemente, el desarrollo de simuladores *in silico* constituye una alternativa realista. En este artículo nos centra-

remos en la paralelización de un simulador ya descrito, con el objetivo de maximizar sus prestaciones y posibilitar su uso universal, es decir, en cualquier dispositivo de cualquier fabricante.

## 2. Estado del Arte

La complejidad de la MRI ha hecho que la presencia de la simulación en este campo haya resultado indispensable desde sus primeros días [1]. Los objetivos perseguidos en el desarrollo de los simuladores de MRI difieren, pudiendo encontrar simuladores con fines educativos o de investigación en materia de desarrollo de nuevas secuencias de pulsos. Centrando el análisis en los simuladores con fines no educativos y, por tanto, con motores de simulación basados en el principio físico subyacente, se pueden destacar los siguientes [2]: JEMRIS [3], que hace uso de las ecuaciones de Bloch y se centra en el desarrollo de secuencias; MRILab [4], que simula la evolución del vector de magnetización a través de la ecuación de Bloch y que fue creado para el desarrollo de nuevas técnicas de MRI; SpinBench [5], que permite simular secuencias de pulsos diseñadas a través de su interfaz gráfica, pero que no genera imágenes MRI; PSUdoMRI [6], que emplea la simulación realista de los campos magnéticos para la obtención de la señal y el ruido de la resonancia magnética, y POS-SUM [7], diseñado para la investigación de artefactos que pueden aparecer en las imágenes MRI.

De forma adicional a los simuladores mencionados, destaca la existencia del simulador KomaMRI.jl [8], desarrollado en la Pontificia Universidad Católica de Chile, en el que se ha empleado el lenguaje de programación Julia [9] y que cuenta con un motor de simulación realista. Este simulador ofrece resultados muy competitivos con respecto a los simuladores que se acaban de mencionar, y, actualmente, es capaz de realizar operaciones paralelas. Sin embargo, dispone únicamente de soporte para dispositivos GPU de un único fabricante.

En el campo de la educación, se han descrito simuladores que emulan el proceso llevado a cabo por un técnico en su labor diaria (ver [10] y sus referencias) y se ha descrito también la inclusión de un tutor inteligente [11]. Los si-

muladores educativos, sin embargo, suelen poner el énfasis en la interactividad, por lo que los motores de simulación suelen ser rápidos y simples, lo cual impide desarrollar con ellos nuevas secuencias de pulsos y, por lo tanto, limita su utilidad como herramienta de investigación.

Recientemente se ha desarrollado el *framework* OpenCLIPER [12], que facilita el desarrollo de aplicaciones en el lenguaje OpenCL [13]. El *framework* está diseñado como un conjunto de clases de C++ centradas en proporcionar al programador diversas funcionalidades como la gestión del dispositivo de computación, la carga y guardado de datos y la gestión de los algoritmos. El hecho de emplear OpenCL hace que el código desarrollado cumpla el paradigma WORA (*write once, run anywhere*), de forma que el código es multidispositivo (CPU, GPU, FPGA, etc.).

En este artículo describimos la metodología llevada a cabo para escribir el núcleo del simulador [8] en OpenCL, haciendo uso de OpenCLIPER, de forma que las ventajas de dicho simulador sean reforzadas con la potencia de OpenCLIPER y, así, los simuladores basados en principios físicos puedan ser empleados indistintamente, tanto con finalidades docentes como con finalidades investigadoras, y sobre cualquier dispositivo, con independencia del fabricante de que se trate. Para ello, describiremos los fundamentos del simulador (sección 3); a continuación describiremos cómo se ha llevado a cabo su desarrollo en OpenCL/OpenCLIPER (sección 4), y, finalmente, llevaremos a cabo un análisis comparativo de prestaciones de ambos simuladores sobre un fantoma que sirve como banco de pruebas (sección 5).

### 3. Diseño del Simulador

El simulador descrito en [8] hace uso de los siguientes conceptos:

- Grad es una estructura formada por una amplitud ( $A$ ) y una duración ( $T$ ), que permite, a partir de la concatenación de bloques, construir las formas de ondas de los gradientes.
- RF es una estructura formada por la amplitud ( $A$ ) del pulso de RF ( $B_{1,x} + jB_{1,y}$ ) y una duración ( $T$ ).
- ADC es una estructura formada por un número de muestras ( $N$ ) y una duración ( $T$ ), que permite determinar los instantes temporales de lectura de la señal de MRI.
- Sequence es una estructura formada por las formas de onda de los gradientes, los pulsos de RF y los tiempos de adquisición de datos. Por tanto, integra tres *arrays* de bloques Grad, RF y ADC.
- Phantom representa el objeto virtual sobre el que se aplica la simulación del fenómeno MRI. Se trata de una estructura que incluye las posiciones espaciales iniciales de los *spins* ( $x, y, z$ ), la densidad de los protones  $\rho$ , las constantes de tiempo  $T1$  y  $T2$ , entre otros, mayoritariamente definidos como vectores.
- Mag corresponde con el vector de magnetización y, como tal, es una estructura formada por su componente transversal ( $M_{xy}$ ) y longitudinal ( $M_z$ ).

- Spinor es una estructura empleada durante la simulación para aplicar una rotación sobre el vector de magnetización y evitar el uso de matrices para tal fin.

Estos conceptos son los elementos que se emplean para resolver las ecuaciones de Bloch mediante una aproximación por eventos discretos. No entraremos en los detalles en este artículo, sino en cómo se ha llevado a cabo esta operativa en OpenCL, haciendo uso del *framework* OpenCLIPER.

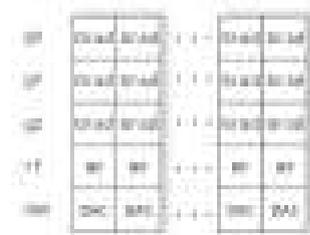
### 4. Desarrollo del Simulador en OpenCL/OpenCLIPER

El simulador KomaMRI.jl [8] está implementado en Julia, un lenguaje de alto nivel que abstrae la disposición de los datos en memoria, lo cual permite desarrollar código rápidamente, pero puede influir negativamente en su rendimiento. Nuestro simulador, en cambio, se ha diseñado con el objetivo de alcanzar el máximo rendimiento posible, minimizando las transferencias *host*-dispositivo y los movimientos de datos durante las operaciones, y maximizando la cohesión de los datos para favorecer el uso eficiente de las cachés. El simulador se incorporará al *framework* OpenCLIPER y se publicará con licencia GNU, igual que el resto del *framework*.

En ese sentido, el simulador desarrollado va a aprovechar las facilidades ofrecidas por el *framework* OpenCLIPER en cuanto a estructuras de datos y gestión de algoritmos, definiendo, a partir de las clases predefinidas por el mismo [12], nuevas clases destinadas a estructuras de datos y al proceso de simulación, las cuales aparecen representadas en la *Figura 1*.

Las principales clases son:

- Grad, RF y DAC son clases que se definen de forma inmediata a partir de sus correspondientes estructuras en el simulador KomaMRI.jl (Grad, RF y ADC, respectivamente).
- Sequence es la clase correspondiente a las secuencias del simulador. Un objeto Sequence se constituye a partir de tres punteros a vectores de objetos Grad, cada uno en una dirección ( $x, y, z$ ), un puntero a un vector de objetos RF y un puntero a un vector de objetos DAC. Una representación aproximada de un objeto de la clase Sequence correspondería con la matriz mostrada en la *Figura 2*.



*Figura 2.* Representación objeto Sequence.

- Phantom es la clase que representa el objeto escaneado cuya imagen se desea reconstruir. Es una clase derivada de la clase Data de OpenCLIPER y, como tal, consta de un único NDAarray con 11 filas, que co-

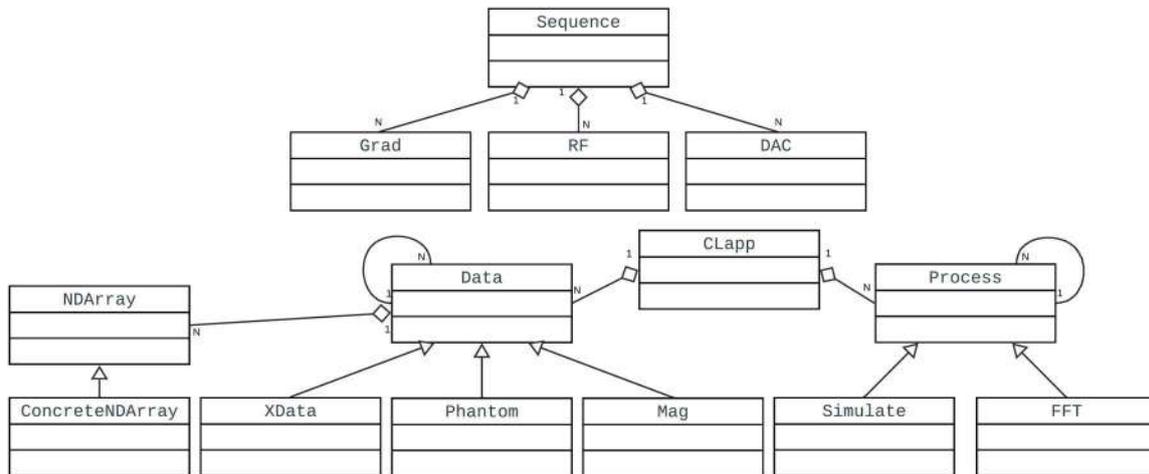


Figura 1. Diagrama de clases del simulador en OpenCLIPER.

responden a cada una de las componentes del objeto Phantom, tal como se muestra en la Figura 3.

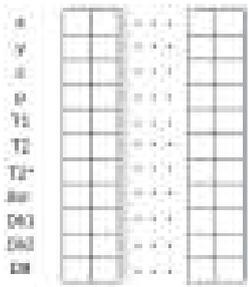


Figura 3. Representación objeto Phantom.

- Mag es la clase que corresponde al vector de magnetización. Deriva también de la clase Data y consta de un único NDAarray de 2 filas correspondientes a la componente transversal ( $M_{xy}$ ) y longitudinal ( $M_z$ ) del vector de magnetización.

En cuanto a la simulación, se define la clase Simulate, que deriva de la clase abstracta Process de OpenCLIPER. El proceso Simulate requiere esencialmente, para que la simulación pueda tener lugar, de un objeto Phantom, que se establece como objeto Data de entrada, y de un objeto Sequence.

Para favorecer el rendimiento, en OpenCLIPER la inicialización y ejecución son fases separadas, de forma que se puedan lanzar múltiples ejecuciones con una única inicialización allá donde sea posible. En la inicialización se efectúan las operaciones relativas al cálculo de los instantes de tiempo de la simulación. El número de instantes de tiempo resulta imprescindible para la obtención en esta misma fase, de todos aquellos objetos de la clase XData de OpenCLIPER imprescindibles para el posterior procesado a través de kernels.

En la fase de ejecución, el vector correspondiente a los instantes temporales se divide en un número de particiones, de forma que el procesado se realiza por intervalos de tiempo. Las operaciones a realizar en cada intervalo dependen de si el pulso de RF se encuentra activo, en cuyo caso se está an-

te *forced precession*, o, en caso contrario, ante *free precession*. En ambos casos, las operaciones se efectúan a través de distintos kernels, de forma análoga a las operaciones realizadas por el simulador KomaMRI.jl [8]. Los kernels implementados en nuestro simulador efectúan un procesamiento en paralelo de todas las estructuras de datos anteriormente descritas, sin necesidad de reorganización de los datos ni transferencias *host*-dispositivo, maximizando así el rendimiento.

La señal de MRI únicamente se lee en los intervalos de *free precession*. Posteriormente, la señal es interpolada en los instantes de tiempo determinados por los bloques DAC, obteniendo así las posiciones correctas en el espacio K. Finalmente, se efectúa una Transformada Inversa de Fourier para la obtención de la imagen final reconstruida, para lo cual se hace uso del proceso FFT, que deriva de la clase Process de OpenCLIPER.

## 5. Resultados y discusión

Los resultados se evalúan a través de la comparación del rendimiento del simulador desarrollado en OpenCL/OpenCLIPER y del simulador KomaMRI.jl [8]. Para ello, se ha efectuado el proceso de simulación en ambos simuladores mediante una secuencia común y un mismo fantoma, cuya imagen reconstruida se muestra en la Figura 4. El fantoma está constituido por 12714 spins, que dan lugar a una primera señal que es posteriormente interpolada en 10201 instantes de tiempo. Esta señal se estructura en una matriz de dimensiones  $101 \times 101$ , que en el simulador desarrollado en OpenCLIPER se redimensiona mediante la introducción de ceros a una matriz de  $128 \times 128$ , para facilitar el empleo del proceso FFT.

Las medidas de rendimiento se efectúan en términos de tiempos de ejecución, a partir de 100 ejecuciones, utilizando como parámetros la media y la varianza. Se emplean tres dispositivos distintos, ejecutando en los dos primeros ambos simuladores y en el tercero únicamente el desarrollado en OpenCLIPER (puesto que KomaMRI.jl no está actualmente preparado para funcionar en una GPU AMD).

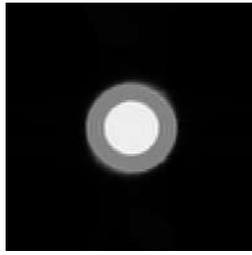


Figura 4. Fantoma reconstruido.

En la *Tabla 1* se muestran los tiempos, medidos en segundos, para cada dispositivo.

		Media	Varianza
Nvidia Quadro RTX 6000	OpenCLIPER	0.789284	0.159327
	KomaMRI.jl	4.634991	0.059775
Intel Core i7-4790	OpenCLIPER	2.545896	0.080647163
	KomaMRI.jl	7.024032	0.086608
AMD Radeon RX 5700 XT	OpenCLIPER	2.701813	0.000027

**Tabla 1.** Tiempos de ejecución de los simuladores en distintos dispositivos (en segundos).

De acuerdo con la *Tabla 1*, se aprecia la importante ganancia en términos de tiempos de ejecución que ha supuesto la implementación del simulador en OpenCL/OpenCLIPER, así como la posibilidad de extender los dispositivos en que puede ejecutarse. Los valores reducidos de la varianza en todos los casos permiten determinar la fiabilidad de los resultados obtenidos.

## 6. Conclusiones y Líneas Futuras

Se ha presentado la implementación de un simulador MRI sobre el *framework* OpenCLIPER, empleando el simulador KomaMRI.jl como base respecto al proceso de simulación.

El trabajo desarrollado ha dado lugar a unos resultados que superan al simulador de partida en tiempos de ejecución, que mantiene su mismo grado de realismo y que funciona sobre dispositivos de cualquier fabricante, tanto CPU como GPU. Adicionalmente, se pone de manifiesto la facilidad que proporciona OpenCLIPER respecto a la ejecución del simulador en cualquier dispositivo de cualquier vendedor.

Los resultados obtenidos son favorables e introducen, además, la posibilidad de futuras mejoras. En particular, la implementación actual no hace uso de memoria local, cuyo uso junto con otras técnicas de optimización podría resultar en una reducción de los tiempos de ejecución aún mayor.

## 7. Agradecimientos

Este trabajo ha sido financiado por el Ministerio de Ciencia e Innovación a través del proyecto PID2020-115339RB-I00 y por el Ministerio de Educación y Formación Profesional a través de la Beca SIA 998142.

## Referencias

- [1] J. Bittoun, J. Taquin, y M. Sauzade, “A computer algorithm for the simulation of any nuclear magnetic resonance (NMR) imaging method,” *Magnetic Resonance Imaging*, vol. 2, pp. 113–120, 1984.
- [2] R. de Luis-García, D. Treceño-Fernández, J. Calabía-del Campo, F. J. de Paz-Fernández, y C. Alberola-López, “Magnetic resonance imaging simulator: From a proof of concept to a useful tool in the education of radiographers,” in *Technological Adoption and Trends in Health Sciences Teaching, Learning, and Practice*. IGI Global, 2022, pp. 66–93.
- [3] P. D. Stoecker T, Vahedipour K y S. N, “High-performance computing MRI simulations,” *Magnetic Resonance in Medicine*, vol. 64, no. 1, pp. 186–193, 2010.
- [4] F. Liu, J. V. Velikina, W. R. Block, R. Kijowski, y A. A. Samsonov, “Fast realistic MRI simulations based on generalized multi-pool exchange tissue model,” *IEEE Transactions on Medical Imaging*, vol. 36, no. 2, pp. 527–537, 2017.
- [5] W. Overal y J. Pauly, “An extensible, graphical environment for pulse sequence design and simulation,” *Int. Soc. Magn. Reson. Med.*, 2007.
- [6] Z. Cao, C. Sica, S. Oh, J. McGarrity, T. Horan, B. Park, y C. Collins, “An MRI simulator for effects of realistic field distributions and pulse sequences, including SAR and noise correlation array coils,” *Proceedings of ISMRM*, p. 1456, 2010.
- [7] I. Drobnjak, G. S. Pell, y M. Jenkinson, “Simulating the effects of time-varying magnetic fields with a realistic simulated scanner,” *Magnetic Resonance Imaging*, vol. 28, no. 7, pp. 1014–1021, 2010.
- [8] C. Castillo-Passi y P. Irrazaval. (2021, Sep.) cncastillo/komamri.jl. Disponible en: <https://doi.org/10.5281/zenodo.6627503>
- [9] J. Bezanson, A. Edelman, S. Karpinski, y V. B. Shah, “Julia: A fresh approach to numerical computing,” *SIAM review*, vol. 59, no. 1, pp. 65–98, 2017.
- [10] D. Treceño-Fernández, J. Calabía-del Campo, M. L. Bote-Lorenzo, E. Gómez-Sánchez, R. Luis-García, y C. Alberola-López, “A web-based educational magnetic resonance simulator: Design, implementation and testing,” *Journal of Medical Systems*, vol. 44, 2019.
- [11] D. Treceño-Fernández, J. Calabía-del Campo, M. L. Bote-Lorenzo, E. Gómez-Sánchez, R. Luis-García, y C. Alberola-López, “Integration of an intelligent tutoring system in a magnetic resonance simulator for education: Technical feasibility and user experience,” *Computer Methods and Programs in Biomedicine*, vol. 195, 2020.
- [12] F. Simmross-Wattenberg, M. Rodríguez-Cayetano, E. Moya-Sáez, M. Martín-Fernández, y C. Alberola-López, “OpenCLIPER: An OpenCL-based C++ framework for overhead-reduced medical image processing and reconstruction on heterogeneous devices,” *IEEE journal of biomedical and health informatics*, vol. 23, pp. 1702–1709, 2019.
- [13] J. E. Stone, D. Gohara, y G. Shi, “OpenCL: A parallel programming standard for heterogeneous computing systems,” *Computing in science & engineering*, vol. 12, no. 3, p. 66, 2010.

# Análisis hemodinámico en dos localizaciones de la Aorta

J.P. Tello Portillo<sup>1</sup>, J. Vélez Díaz<sup>1</sup>, A. Jutinico Alarcón<sup>2</sup>, A. Cadena Bonfanti<sup>3</sup>

<sup>1</sup> Departamento Ing. Eléctrica y Electrónica, Universidad del Norte, Barranquilla, Colombia, [telloj@uninorte.edu.co](mailto:telloj@uninorte.edu.co), [jcvelez@uninorte.edu.co](mailto:jcvelez@uninorte.edu.co)

<sup>2</sup> Facultad de Ingeniería, Universidad Distrital Francisco José de Caldas, Bogotá, Colombia, [aljutinicoa@udistrital.edu.co](mailto:aljutinicoa@udistrital.edu.co)

<sup>3</sup> Departamento de Cardiología, Clínica de la Costa, Barranquilla Colombia, [acadenabon@gmail.com](mailto:acadenabon@gmail.com)

## Resumen

*Se presenta la simulación hemodinámica en dos secciones transversales de la aorta torácica de dos pacientes, uno de ellos con endoprótesis. El objetivo es analizar la influencia de la endoprótesis en el comportamiento del flujo sanguíneo y predecir posibles complicaciones clínicas. Se analizan las magnitudes de velocidad y presión en diferentes instantes de tiempo durante un ciclo cardíaco y en un número de posiciones de las secciones transversales (ST1 y ST2). Los resultados muestran dispersiones notorias de la velocidad y pequeñas dispersiones de la presión de una posición a otra. A pesar de la dispersión, la velocidad y presión en la ST1 de ambos pacientes siguen un comportamiento acorde a los principios de la mecánica de fluidos, dada su forma geométrica y localización en la aorta. En la ST2, la dispersión en ambos pacientes no es predecible, debido a los vórtices generados por la bifurcación del fluido y más aún, en el paciente II, por la coincidencia de la ST2 con el borde proximal de la endoprótesis, que presenta una geometría irregular, producto de la reparación endovascular del aneurisma, que una vez en contacto con el endotelio, se reacomoda a la pared y toma su forma. Como consecuencia, se presentan alteraciones en los valores de velocidad y cambios de hasta un 10% en los diferenciales de presión. Estas alteraciones del flujo multiplicadas por la frecuencia cardíaca pueden generar cambios en la anatomía y posibles desplazamientos que pueden conducir a una pérdida de continuidad prótesis-endotelio, creando complicaciones clínicas en el paciente.*

## 1. Introducción

Las enfermedades cardiovasculares (ECV) son la principal causa de muerte en el mundo. Se estima que 17.9 millones de personas murieron por ECV en 2019, lo que representa el 32% de todas las muertes mundiales [1].

Según la Organización Panamericana de la Salud - OPS, las ECV, son la causa principal de muerte en las Américas, cobrando 2 millones de vidas cada año. En América Latina y el Caribe, la hipertensión es una condición muy común que conduce tanto a la enfermedad cardíaca como al accidente cerebrovascular [2].

Hay muchos factores de riesgo bien conocidos que se correlacionan con la incidencia de ECV, entre algunos de ellos están: el estilo de vida sedentario, la dislipidemia, la diabetes, la predisposición genética y el tabaquismo. Sin embargo, los estudios epidemiológicos han demostrado que el mayor riesgo de todos es la edad avanzada [3].

En los grandes vasos, la aorta torácica por su anatomía compleja está sujeta a las ECV típicas que a menudo se manifiestan en formas de aterosclerosis, aneurisma y disección [4]. La aterosclerosis es una ECV en la que se va

depositando material graso y colesterol en las arterias, haciendo que su superficie interna sea irregular y más estrecha, dificultando el paso de la sangre [5].

Por otro lado, el desarrollo de la dinámica de fluidos computacional (CFD) en los últimos años ha permitido el uso de simulaciones computacionales 3D para investigar la hemodinámica específica del paciente y comprender el origen y el desarrollo de diferentes ECV; principalmente por su capacidad de cuantificar variables no medibles in vivo, su reproducibilidad y su papel como herramienta de diagnóstico y tratamiento para diferentes ECV. A diferencia de los estudios in vitro, la CFD es más conveniente, ya que permite modificar parámetros del modelo (condiciones iniciales y de frontera, entre otras), y dada la complejidad del flujo sanguíneo en geometrías reales, la combinación entre técnicas de simulación de alta resolución y datos de mediciones basadas en imágenes, ha demostrado ser una herramienta confiable para el modelado realista del flujo sanguíneo arterial [4].

Los primeros análisis de CFD basados en imágenes fueron motivados por el objetivo de estudiar el papel de las fuerzas hemodinámicas en el desarrollo de la aterosclerosis en las arterias carótida y coronaria, cuantificar la hemodinámica en aneurismas aórticos abdominales y planificar protocolos quirúrgicos específicos para diferentes pacientes [6].

En este trabajo se presenta un estudio del comportamiento del flujo sanguíneo en la aorta torácica para dos pacientes con ECV y la influencia en los patrones de flujo debida a la prótesis endovascular presente en uno de ellos. Se describe el proceso metodológico, comenzando con la adquisición de las imágenes de tomografía (TAC), continuando con el procesamiento para finalmente obtener un modelo 3D que se parametriza y simula usando Comsol Multiphysics. Finalmente se muestra el análisis de los resultados de la simulación, para ciertas posiciones dentro de dos secciones transversales previamente seleccionadas.

## 2. Metodología

### 2.1. Adquisición de imágenes

Las imágenes DICOM fueron adquiridas usando un escáner TAC de la Clínica de la Costa en Barranquilla - Col., y siguiendo un mecanismo para la sincronización de la captura con el latido del corazón de cada paciente, que consiste en la disminución del tiempo de adquisición y/o en la disminución de su frecuencia cardíaca. El propósito fue obtener imágenes susceptibles de evaluación y aptas para generar un modelo 3D de la región de estudio.

Las imágenes adquiridas, con resolución de  $512 \times 512$ , tamaño de píxel de  $0.625\text{mm}$  y espesor de  $0.6\text{mm}$ , fueron anonimizadas para proteger la identidad del paciente.

## 2.2. Procesamiento y reconstrucción del modelo 3D

Dado que, la geometría vascular afecta el comportamiento del fluido [7-10], la etapa de procesamiento junto con la de adquisición (sincronización con el latido del corazón), es clave a la hora de conseguir una geometría 3D que coincida anatómicamente con la estructura real de la aorta, de tal manera, que los valores de la simulación vislumbren lo que sucede realmente con el flujo sanguíneo en la aorta.

Mediante técnicas clásicas de procesamiento de imágenes (filtrado y segmentación), parametrizadas en el software 3DSlicer, se obtuvo la región de interés-ROI, y a partir de ella, la reconstrucción 3D de la geometría de la aorta. Otras estructuras que rodean la ROI son removidas usando el mismo software, obteniéndose una geometría limpia.

En la Figura 1 se muestran las estructuras reconstruidas de los pacientes, la ubicación y la geometría de las dos secciones transversales. La ubicación de la ST1 se elige antes de la bifurcación del fluido con el fin de verificar, que mientras el conducto aórtico no presente alteraciones morfológicas, tampoco el fluido responderá a manifestaciones hemodinámicas impredecibles. La ubicación de la ST2 se hace con la finalidad de conocer los patrones de flujo después de la bifurcación del fluido, y específicamente, para conocer las diferencias y la posible influencia de la ST2 del Paciente II, que coincide con el borde proximal de la endoprótesis.

## 2.3. Condiciones iniciales y de contorno

La sangre se asumió como un fluido homogéneo, incompresible y newtoniano, con densidad constante de  $1060\text{kg/m}^3$  y viscosidad dinámica de  $0.0035\text{Pa} \times \text{s}$ . Las paredes de la aorta se consideraron como rígidas y siguiendo la condición de no deslizamiento (no-slip).

Las ondas de velocidad y presión fueron generadas, de manera similar a los modelos expuestos por otros autores [9, 11-15], y con un ciclo cardíaco de duración de  $0.8\text{s}$ .

## 2.4. Modelado numérico

La solución de las ecuaciones de Navier Stokes que gobiernan el modelo del fluido, se llevó a cabo mediante el software Comsol Multiphysics. El proceso de mallado quedó conformado por triángulos de diferentes tamaños que se ajustan a la forma y cambios geométricos de la aorta en casi toda la estructura, y de rectángulos en los bordes (paredes, entrada y salida) de la aorta.

## 3. Resultados y discusión

Se realizó la simulación del flujo sanguíneo en la aorta torácica de dos pacientes de edad avanzada, uno de ellos con prótesis endovascular. Se tomó medidas en diferentes posiciones de 2 secciones transversales (ST1 y ST2) de la aorta torácica. Se simuló 4 ciclos cardíacos con un tiempo de paso de  $0.0032\text{s}$ . El modelo logró estabilizarse al final de primer ciclo cardíaco, sin embargo, las medidas se tomaron en el tercer ciclo, donde la variación máxima de las magnitudes de velocidad no superó el  $0.5\%$ .

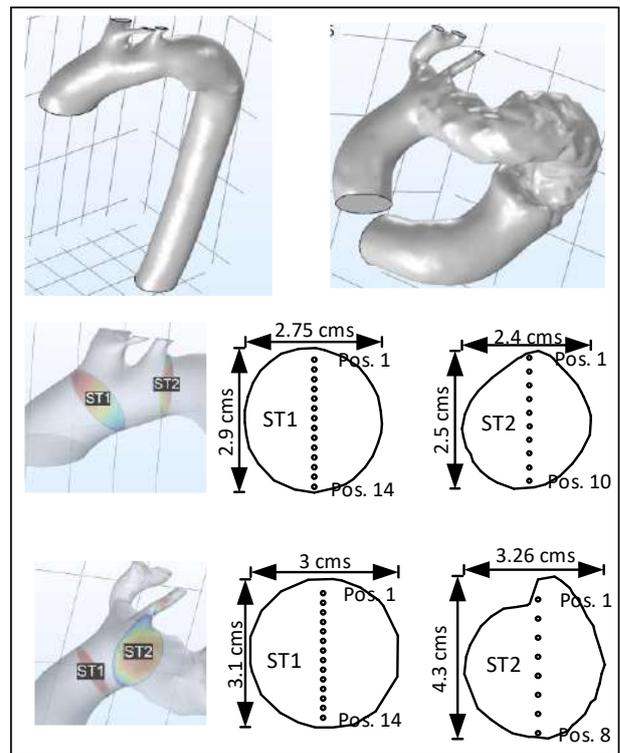


Figura 1. Geometría 3D y localización de las secciones transversales dentro de la aorta

Se analizó el comportamiento de las magnitudes de la velocidad y la presión mediante diagramas de box-plots, donde se observa, como el flujo se modifica de acuerdo a la localización y geometría del conducto aórtico.

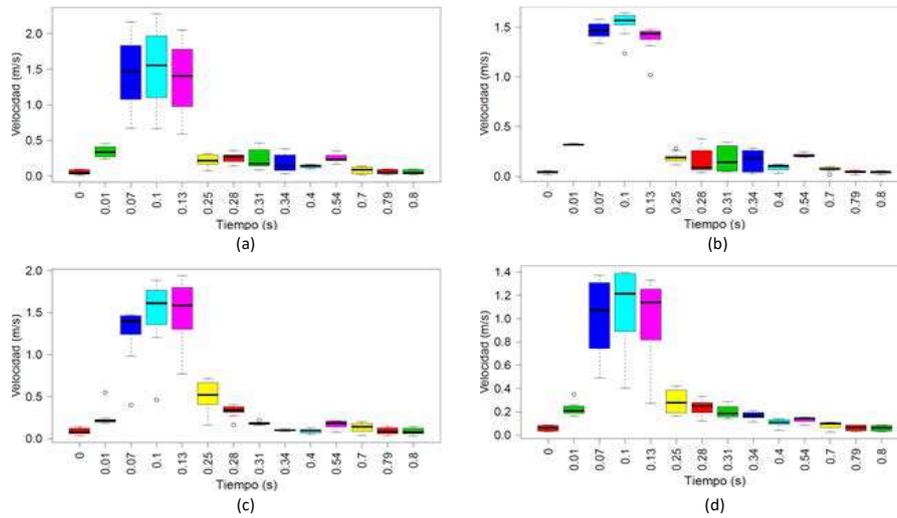
### 3.1. Comportamiento de la velocidad

En la ST1 del Paciente I (Figura 2.(a)) y en la fase intermedia de la sístole ( $0.07\text{s}$  a  $0.13\text{s}$ ) las magnitudes de la velocidad se dispersan ampliamente, siendo mayor donde la velocidad es ( $0.1\text{s}$ ).

A pesar de la dispersión se mantiene una distribución casi que simétrica, con una media muy cercana a la mediana. Luego la dispersión disminuye rápidamente, y en el intervalo de tiempo de  $0.25\text{s}$  a  $0.34\text{s}$ , se presenta otra ligera dispersión, causada por la posición donde se toman las medidas y en mayor parte, por la presencia de un flujo retrógrado debido al cambio de fase en el ciclo cardíaco.

A diferencia del Paciente I, en la ST1 del Paciente II (Figura 2.(b)), las magnitudes de velocidad en la fase intermedia de la sístole presentan una dispersión menor, conservando una distribución simétrica con pequeñas diferencias en el rango entre cuartiles. Una dispersión más notable se da en los instantes de tiempo posteriores, comenzando en  $0.25\text{s}$  y finalizando en  $0.4\text{s}$ . Aquí, no hay uniformidad en la distribución de los datos, causadas nuevamente por la presencia de un flujo retrógrado que altera el comportamiento del fluido.

En la ST2 del Paciente I (Figura 2.(c)), la dispersión de la velocidad empieza a ser perceptible a partir de los  $0.07\text{s}$  y continúa creciendo de manera casi que lineal, hasta alcanzar su valor máximo en el instante de tiempo de  $0.13\text{s}$ , donde la velocidad también es máxima.



**Figura 2.** Comportamiento de las magnitudes de la velocidad durante un ciclo cardíaco. (a). ST1 – Paciente I, (b). ST1 – Paciente II, (c). ST2 – Paciente I y (d). ST2 – Paciente II.

Los valores atípicos de los diagramas de box-plots, son producto de las medidas tomadas cerca de la bifurcación del fluido (arteria subclavia), y en los bordes de la ST2. En 0.25s la velocidad cae abruptamente y la dispersión disminuye, sin dejar de ser notoria. Esto se debe nuevamente al comienzo de un flujo retrogrado.

En la ST2 del Paciente II (Figura 2.(d)) y en la fase intermedia de la sístole, se presenta una alta dispersión de las magnitudes de la velocidad, similares a los observados en la ST1 del paciente I. La diferencia está, en que la distribución de los valores entre cuartiles, cambian en cada uno de los instantes de tiempo. Este comportamiento se da, no solo por la cercanía de la ST2 a la bifurcación, sino también por su geometría irregular (Figura 1), adoptada de la reparación endovascular del aneurisma. Finalmente, en 0.25s, el comportamiento del flujo es muy similar al presentado en la ST2 del Paciente I.

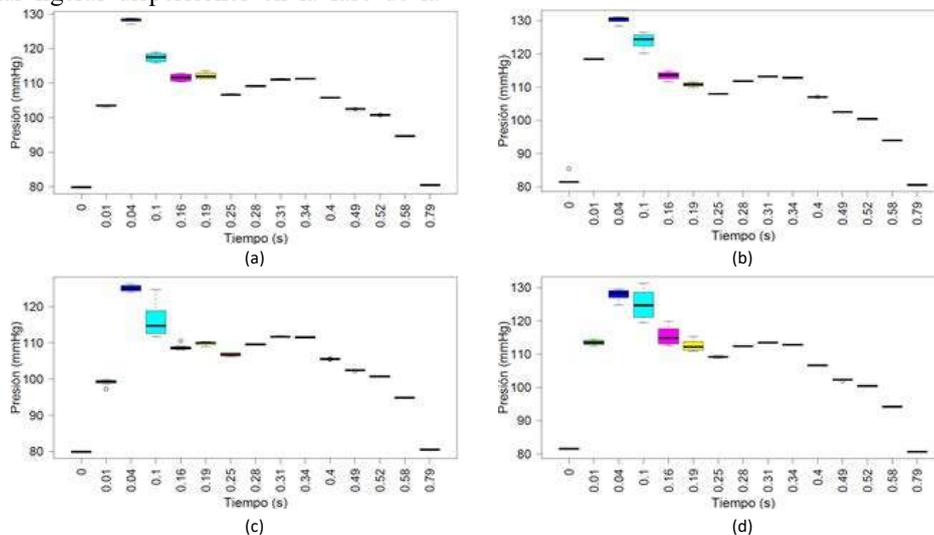
### 3.2. Comportamiento de la presión

En cuanto a la presión y en la ST1 de ambos pacientes, los diagramas de box-plots de las Figuras 3.(a) y 3.(b), permiten visualizar ligeras dispersiones en la fase de la

sístole, específicamente en los instantes de tiempo de 0.1s, 0.16s y 0.19s, siendo un poco más notorias en el instante de tiempo donde la velocidad es máxima (0.1s). Sin embargo, no hay cambios importantes en los diferenciales de presión que sean motivo de análisis.

En la ST2 del Paciente I (Figura 3.(c)), se presenta una marcada dispersión de las magnitudes de la presión, en el instante de tiempo de 0.1s, que está estrechamente relacionada con las alteraciones del fluido, provocadas por la cercanía de la ST2 con la bifurcación de la sangre hacia la arteria subclavia izquierda tal como se muestra en la Figura 1.

Además, los valores máximos de las magnitudes de presión se dan en el instante de tiempo de 0.04s, tiempo en el que aún no se ha alcanzado la velocidad máxima (0.1s). A pesar de eso el diferencial de presión es bajo (2%), con una presión máxima de 126.3mmHg. En el instante siguiente donde la velocidad es máxima, se exhibe un diferencial de presión del 10.6% con solo dos valores de presión (124mmHg y 121mmHg) que están por encima de la presión base (120mmHg).



**Figura 3.** Comportamiento de las magnitudes de la presión durante un ciclo cardíaco. (a). ST1 – Paciente I, (b). ST1 – Paciente II, (c). ST2 – Paciente I y (d). ST2 – Paciente II.

Este diferencial no representa un motivo de análisis clínico ya que se da cerca de la bifurcación del fluido, donde naturalmente el flujo sufre alteraciones que también se ven reflejadas en variaciones de presión, además porque en el siguiente instante de tiempo de 0.16s, nuevamente el diferencial de presión cae al 2%.

Así mismo, en la ST2 del paciente II (Figura 3.(d)), la dispersión, que ahora extiende a casi toda la fase de la sístole, no solo está relacionada con la bifurcación del fluido, sino aún más, por la coincidencia de la ST con la parte proximal de la endoprótesis, que como ya se mencionó, su geometría irregular, dada por el sellamiento prótesis-endotelio, hace que el flujo modifique su patrón reflejándose en diferenciales de presión de hasta un 10% de sus valores nominales, que además pueden transmitirse a los extremos de la endoprótesis. Específicamente hay una amplia y desplegada de diferenciales de presión en la fase de la sístole: 4% en 0.04s, 10% en 0.1s, 6.1% en 0.16s y 3.6% en 0.19s, donde los valores de magnitud máxima y mínima son respectivamente 131.3mmHg y 119.5mmHg que ocurren en el instante donde la velocidad es máxima.

En los demás instantes de tiempo, se exhibe mayor regularidad que es más evidente en la fase de la diástole. En ella, las magnitudes de la presión en los dos pacientes, se mantienen siguiendo un patrón, muy cercano al comúnmente reportado por otros autores en geometrías de aortas similares.

Finalmente, las alteraciones del fluido se maximizan en la ST2 del Paciente II, debidas a la prótesis. En esta región, se exhiben fenómenos reológicos que desencadenan en la formación de vórtices y en la aparición de diferenciales de presión que, sumados a los continuos latidos del corazón, pueden conducir a un desprendimiento parcial de la endoprótesis, produciendo daños irreparables. De ahí que la simulación se convierte en una herramienta de gran utilidad en la asistencia médica, con propósitos de toma oportuna y/o anticipada de decisiones en los diagnósticos que redundan en la calidad de vida y salud de los pacientes.

#### 4. Conclusión

La simulación mostró diferencias hemodinámicas en las STs de los dos pacientes estudiados. Sin embargo, el comportamiento del fluido en las ST1 y ST2 del Paciente I responde a valores esperados acorde a la forma geométrica y localización dentro de la aorta torácica e idénticos a los reportados en trabajos anteriores. En ST1 del Paciente II, ocurre algo similar. En cambio, para la ST2, que además coincide con los bordes de la endoprótesis, el flujo exhibe un comportamiento con cambios abruptos en las magnitudes de la velocidad y diferenciales de presión que, al transmitirse a los extremos de la endoprótesis, provocan fuerzas que pueden deformar la anatomía, impidiendo el anclaje correcto y un posterior desplazamiento, que impactaría negativamente la salud del paciente.

#### Referencias

[1] Página web de la World Health Organization - WHO, Cardiovascular diseases (CVDs), 2021. [https://www.who.int/health-topics/cardiovascular-diseases#tab=tab\\_1](https://www.who.int/health-topics/cardiovascular-diseases#tab=tab_1) (consultada Mayo 2022).

[2] Página web de la Organización Panamericana de la Salud – OPS, Las enfermedades del corazón siguen siendo la principal causa de muerte en las Américas, 2021. <https://www.paho.org/es/noticias/29-9-2021-enfermedades-corazon-siguen-siendo-principal-causa-muerte-americas> (consultada Mayo 2022).

[3] Lakatta EG, Levy D, Arterial and cardiac aging: major shareholders in cardiovascular disease enterprises: Part I: aging arteries: a “set up” for vascular disease. *Circulation*, vol. 107, no 1, 2003, pp 139-146.

[4] Caballero A, Laín S, A review on computational fluid dynamics modelling in human thoracic aorta, *Cardiovascular Engineering and Technology*, vol. 4, no. 2, 2013, pp 103–130.

[5] Mendis S, Puska P, Norrving B, & World Health Organization. *Global atlas on cardiovascular disease prevention and control*, World Health Organization, 2011.

[6] Taylor CA, Steinman DA. Image-based modeling of blood flow and vessel wall dynamics: applications, methods and future directions. *Annals of biomedical engineering*, vol. 38, no 3, 2010, pp 1188-1203.

[7] Podyma M, Zbicinski I, Estimation of the effect of system geometry on blood flow hemodynamics in arteries, *Inzynieria Chemiczna i Procesowa*, vol. 25, no 3, 2004, pp 1527-1532.

[8] Morris, L., et al. 3-D numerical simulation of blood flow through models of the human aorta, *Journal of biomechanical engineering*, vol. 127, no 5, 2005, pp 767-775.

[9] Tang, AYS, Fan Y, Cheng SWK, Chow KW, Biomechanical factors influencing type B thoracic aortic dissection: computational fluid dynamics study, *Engineering Applications of Computational Fluid Mechanics*, vol. 6, no 4, 2012, pp 622-632.

[10] Dabagh M, Vasava P, Jalali P, Effects of severity and location of stenosis on the hemodynamics in human aorta and its branches, *Medical & biological engineering & computing*, vol. 53, no 5, 2015, pp 463-476.

[11] Ku DN, Blood flow in arteries. *Annual review of fluid mechanics*, vol. 29, no 1, 1997, pp 399-434.

[12] Olufsen, MS, Peskin CS, Kim WY, Pedersen EM, Nadim A, Larsen J, Numerical simulation and experimental validation of blood flow in arteries with structured-tree outflow conditions, *Annals of biomedical engineering*, vol. 28, no 11, 2000, pp 1281-1299.

[13] Leung JH, Wright AR, Cheshire N, Crane J, Thom, SA, Hughes AD, Xu Y, Fluid structure interaction of patient specific abdominal aortic aneurysms: a comparison with solid stress models. *Biomedical engineering online*, vol. 5, no 1, 2006, pp 1-15.

[14] Fung GS, Lam SK, Cheng, SW, Chow KW, On stent-graft models in thoracic aortic endovascular repair: a computational investigation of the hemodynamic factors, *Computers in biology and medicine*, vol. 38, no 4, 2008, pp 484-489.

[15] Li J, Tian X, Wang Z, Xiong J, Fan Y, Deng X, Liu X, Influence of endoleak positions on the pressure shielding ability of stent-graft after endovascular aneurysm repair (EVAR) of abdominal aortic aneurysm (AAA), *BioMedical Engineering OnLine*, vol. 15, no 2, 2016, pp 421-429.

# Influencia del tipo de muerte del donante en la supervivencia del trasplante renal: un estudio centrado en la población andaluza

A. Talaminos Barroso<sup>1</sup>, J. Reina Tosina<sup>1,2</sup>, L. M. Roa Romero<sup>1,2</sup>, J. Calvillo Arbizu<sup>1,3</sup>, M. A. Pérez Valdivia<sup>4</sup>, R. Medina López<sup>4</sup>, P. Castro de la Nuez<sup>5</sup>

<sup>1</sup> Grupo de Ingeniería Biomédica, Universidad de Sevilla, Sevilla, España {atalaminos, jreina, lroa, jcalvillo}@us.es

<sup>2</sup> Departamento de Teoría de la Señal y Comunicaciones, Universidad de Sevilla, Sevilla, España

<sup>3</sup> Departamento de Ingeniería Telemática, Universidad de Sevilla, Sevilla, España

<sup>4</sup> Unidad de Urología y Nefrología, Hospital Universitario Virgen del Rocío, Sevilla, España, {miguel.perez, rantonio.medina}.sspa@juntadeandalucia.es}

<sup>5</sup> Coordinación Autonómica de Trasplante de Andalucía, Sevilla, España, pablo.castro.sspa@juntadeandalucia.es

## Resumen

*El trasplante renal es el tratamiento de elección para pacientes con enfermedad renal terminal. La mayoría de los injertos proceden de donantes fallecidos por muerte encefálica. Sin embargo, en los últimos años se están incrementando los casos de donantes en asistolia o muerte circulatoria. El objetivo de este trabajo es analizar la supervivencia del trasplante renal para donantes por muerte encefálica y por muerte circulatoria en Andalucía. Para ello se ha realizado un análisis de supervivencia mediante el estimador de Kaplan-Meier junto a la técnica de significación estadística log-rank. Los resultados pusieron de manifiesto que el tipo de muerte no tiene un impacto significativo en la supervivencia del injerto considerando los primeros días después del trasplante, pero sí a partir del día 30 y hasta un periodo de dos años ( $p < 0.005$ ). A partir de esta fecha, el impacto del tipo de muerte en la supervivencia del injerto no es significativo en nuestros datos.*

## 1. Introducción

La prevalencia de la enfermedad renal crónica terminal (ERC) ha aumentado en la última década. El envejecimiento de la población y la diabetes son los dos principales factores que han motivado este aumento. La morbimortalidad asociada a la ERC la convierte en un problema de salud pública mundial. Un informe reciente [1] estima que el 10% de la población mundial mayor de 30 años se encuentra afectada por ERC y en algunos países como Estados Unidos se estima que esta incidencia es aún mayor [2].

El trasplante continúa siendo el tratamiento de elección para esta enfermedad, si bien la oferta existente de donantes no logra suplir la demanda creciente. Por otro lado, el perfil del donante también ha cambiado, siendo cada vez más mayor y con un número creciente de comorbilidades, lo que puede provocar una alta tasa de descartes. Uno de los principales avances en el ámbito del trasplante ha sido la creciente utilización de injertos procedentes de donantes en muerte circulatoria controlada (cDCD), también conocida como asistolia tipo III según la clasificación de Maastricht. La muerte encefálica (DBD) ha sido la principal causa de donación en Europa [3] desde el inicio de los programas de trasplante. Sin embargo, este tipo de donación se ha reducido en los últimos años debido

a factores epidemiológicos, como son el descenso de la siniestralidad vial y los accidentes laborales, así como la mejora en el tratamiento asociado a los factores de riesgo cerebrovasculares. En consecuencia, los injertos cDCD vienen incrementándose en los últimos años. Estas donaciones se caracterizan por una elevada complejidad logística y el poco tiempo que se dispone durante la toma de decisiones.

En este trabajo se presenta un estudio del impacto del tipo de muerte del donante (DBD o cDCD) en la probabilidad de supervivencia mediante análisis estadísticos y la aplicación del estimador de Kaplan-Meier en datos referentes a trasplantes renales de cadáveres en la población andaluza en el periodo 2006-2019. Con una población equivalente a algunos países europeos, en la comunidad autónoma de Andalucía no existen hasta la fecha estudios que comparen los resultados en la supervivencia del injerto de donantes cDCD respecto a DBD. Dado el aumento reciente de casos cDCD, este tipo de análisis puede ayudar a la toma de decisiones durante el proceso de pre-trasplante.

## 2. Métodos y materiales

### 2.1. Datos disponibles

Se dispone de una base de datos con datos provenientes del Sistema de Información de la Coordinación Autonómica de Trasplantes de Andalucía (SICATA). Los datos se refieren a trasplantes de un único riñón de donantes con edades comprendidas entre 3 y 90 años, y receptores mayores de 18 años. El número total de casos es de 4907, de los cuales 535 (12%) son casos de fracaso. El número de donaciones por DBD es 4176, mientras que por cDCD es 731. El número de variables disponibles es 73, de las cuales 22 están asociadas al donante, 44 al receptor y 7 al proceso de trasplante. En las tablas I, II y III se presentan las variables más relevantes del donante, receptor y proceso de trasplante, respectivamente, separando por tipo de muerte. Para las variables edad, peso, talla, creatinina, anticuerpos citotóxicos, tiempo en diálisis, KDPI y tiempo de isquemia frío se presenta el valor medio y, entre paréntesis, la desviación estándar. Para el resto de variables se representa

el número de casos y, entre paréntesis, el porcentaje respecto al total de casos para cada tipo de muerte.

Característica del donante	DBD	cDCD	p
Edad del donante	53.92 (15.57)	56.52 (11.13)	<0.05
Género (%hombres)	2471 (59.17%)	502 (68.67%)	<0.05
Peso (kg)	77.92 (15.4)	80.34 (15.47)	<0.05
Talla (cm)	168.45 (10)	169.6 (8.79)	<0.05
Diabetes	297 (7.11%)	37 (5.06%)	0.04
Hepatitis B	390 (9.34%)	83 (11.35%)	0.22
Hepatitis C	12 (0.29%)	5 (0.68%)	0.87
Citomegalovirus	528 (12.64%)	80 (10.94%)	0.01
Hipertensión	1285 (30.77%)	206 (28.18%)	0.16
Creatinina (mg/dl)	0.86 (0.4)	0.69 (0.31)	<0.05
Anticuerpos citotóxicos	10.45 (25.09)	13.73 (29.75)	<0.05

Tabla 1. Características principales de los donantes

Característica del receptor	DBD	cDCD	p
Edad del donante	53.08 (13.02)	55.42 (12.15)	<0.05
Género (%hombres)	2637 (63.15%)	457 (62.52%)	0.74
Peso (kg)	72.61 (14.06)	73.73 (15.06)	0.14
Talla (cm)	164.53 (12.35)	166.13 (11.48)	0.01
Diabetes	541 (12.95%)	167 (22.85%)	<0.05
Hepatitis B	51 (1.22%)	8 (1.09%)	0.75
Hepatitis C	127 (3.04%)	14 (1.92%)	0.08
Citomegalovirus	343 (8.21%)	41 (5.61%)	<0.05
Tiempo en diálisis (días)	856.04 (817.06)	681.57 (726.24)	<0.05

Tabla 2. Características principales de los receptores

Característica del injerto	DBD	cDCD	p
KDPI	48.49 (28.99)	57.84 (26.37)	<0.05
Tiempo de isquemia fría (minutos)	970.24 (306.08)	762.98 (328.72)	<0.05

Tabla 3. Características principales del proceso de trasplante

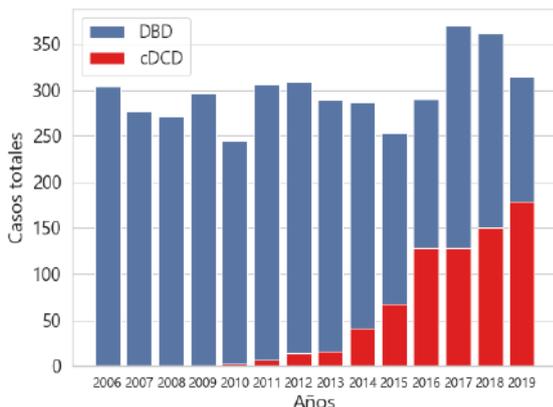


Figura 1. Casos totales DBD y cDCD durante todo el tiempo de seguimiento

En la Figura 1 se presentan los casos totales de trasplantes renales de la población considerada para injertos DBD (azul) y cDCD (rojo) desde que comenzó el estudio hasta los últimos datos disponibles, donde en el eje de abscisas se representa el año en el que se realizó el trasplante y en ordenadas el número de casos. Con respecto a los casos DBD, el número de trasplantes ha permanecido constante durante todo el periodo del estudio

## 2.2. Metodología

La metodología general seguida en este trabajo para el análisis de probabilidad de supervivencia censurada al fracaso del trasplante a partir de los datos disponibles ha sido la siguiente:

1. Adecuación de los datos disponibles para su posterior análisis, incluyendo el tratamiento de valores ausentes y conversión de variables categóricas mediante la técnica *One Hot Encoding* [4].
2. Análisis de datos, tomando las variables a priori más significativas del donante, receptor y proceso de trasplante, considerando con especial interés la variable tipo de muerte del donante (DBD o cDCD).
3. Análisis univariante mediante el estimador de Kaplan-Meier [5] y la técnica de significación estadística *log-rank* [6], agrupando a los casos en distintos grupos poblacionales por edad y diferentes periodos temporales tras la realización del trasplante.
4. Representación gráfica y discusión de los resultados.

## 3. Resultados

### 3.1. Análisis de los datos

En la Figura 2 se presentan las curvas de supervivencia obtenidas mediante el estimador de Kaplan-Meier considerando los dos tipos de muerte. En general, puede observarse que la probabilidad de supervivencia es más baja en cDCD que en DBD, pero ambos presentan un pronunciado descenso en la primera etapa post-trasplante. A partir del décimo año, aproximadamente, no se dispone de datos para injertos cDCD, debido a que este tipo de trasplantes no comenzaron a realizarse hasta 2010.

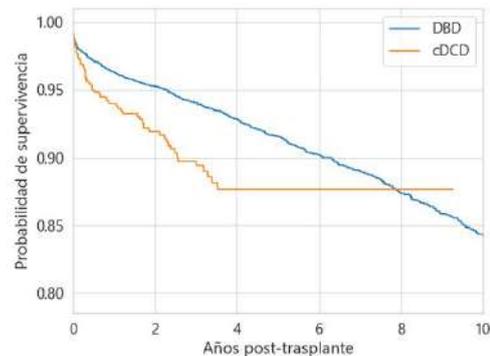


Figura 2. Probabilidad de supervivencia del injerto para DBD y cDCD

Los estudios centrados en el análisis de supervivencia incorporan generalmente la edad del donante [7], con independencia del tipo de estudio, dado que esta variable está considerada como la de mayor influencia en la supervivencia del trasplante renal [8]. En la Figura 3 se presentan los efectos de la edad del donante sobre la supervivencia del injerto en años, para DBD y cDCD. Los casos de fracasos DBD están distribuidos en función de la edad y ocurren en un rango más amplio de tiempo post-trasplante que los cDCD.

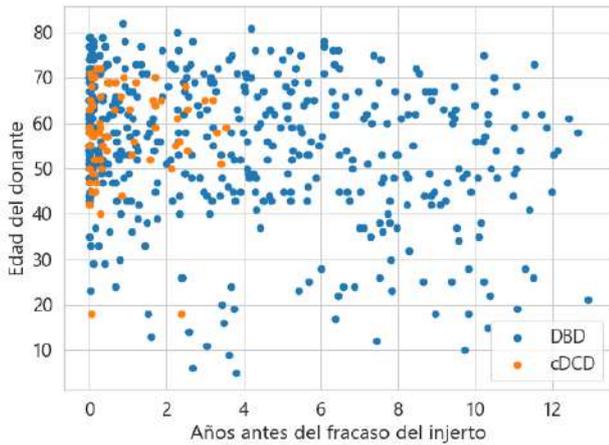


Figura 3. Años de supervivencia de los casos de fracaso para DBD y cDCD considerando la edad del donante

Un resultado complementario al anterior incluye los casos de fracaso y no fracaso para DBD y cDCD (Figuras 4 y 5). En el eje de abscisas se representa el número de años de seguimiento para los casos de no fracaso, mientras que para los casos de fracaso el eje de abscisas representa el número de años anteriores al evento de fracaso. En particular, para los casos de no fracaso el número de días representa el tiempo hasta el que se dispone de datos. Ignorando los casos de fracaso en los primeros días, se confirma también que los casos de fracaso para DBD (Figura 4), comparados con cDCD (Figura 5), se distribuyen en un mayor rango de edades y periodos de seguimiento, mientras que los fracasos de cDCD están localizados fundamentalmente en mayores de 45 años y un periodo de seguimiento inferior a 4 años.

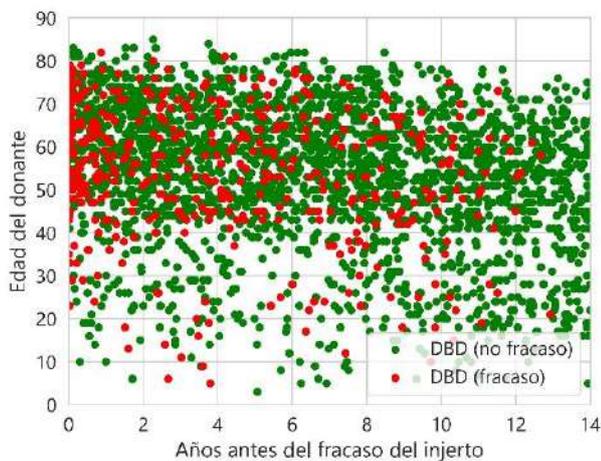


Figura 4. Evolución temporal con la misma escala de los casos de fracaso y no fracaso del injerto para DBD

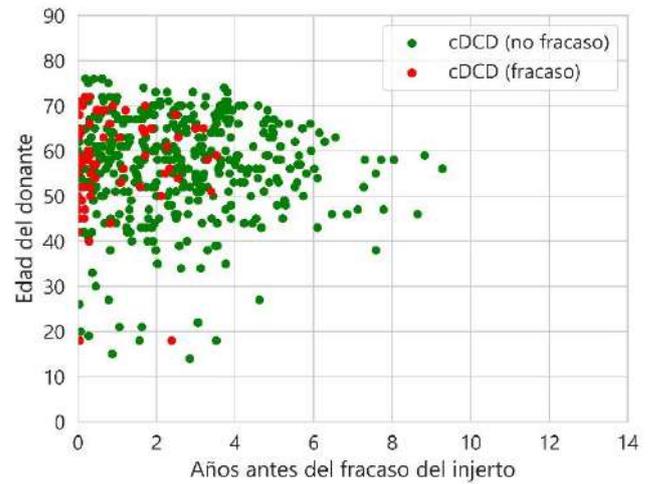


Figura 5. Evolución temporal con la misma escala de los casos de fracaso y no fracaso del injerto para cDCD

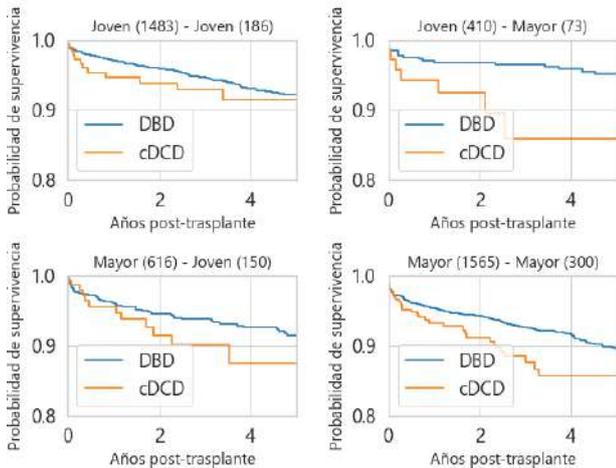
Atendiendo a las diferencias de comportamiento de los resultados mostrados en las figuras anteriores, un estudio por búsqueda exploratoria utilizando la técnica del log-rank permite identificar diferentes tramos temporales con diferente significación estadística en relación a la probabilidad de supervivencia. En particular, los periodos de tiempo identificados son: menos de 15 días, de 15 a 30 días, de 30 a 730 días y más allá de 730 días. La distribución de los casos de fracaso en función de estos grupos a distintas escalas temporales se presenta en la Tabla 3. El 79% de los casos de fracaso de donante en asistolia controlada se produce antes de 730 días, mientras que para DBD es solamente del 40%. Por otra parte, el número de casos de fracaso totales durante los 30 primeros días es significativo, aunque son superiores para los casos con donantes en cDCD.

	DBD	cDCD
0 ≤ días < 5	36 (8%)	7 (11%)
5 ≤ días < 30	30 (6%)	8 (12%)
30 ≤ días < 730	124 (26%)	36 (56%)
días ≥ 730	281 (60%)	13 (21%)
<b>Total</b>	<b>471</b>	<b>64</b>

Tabla 4. Resumen de los comportamientos observados en las Figuras 4 y 5 de casos de fracaso para cDCD y DBD

El análisis de las probabilidades de supervivencia entre jóvenes/mayores donantes y jóvenes/mayores receptores para DBD y cDCD se presenta en la Figura 6, considerando un periodo de tiempo de 5 años. El número entre paréntesis en el título de cada gráfica representa el número de casos para cada grupo poblacional. En otros trabajos similares [7], se considera a donante/receptor joven con una edad de menos de 45 años y, a partir de esa edad, donante/receptor mayor. Sin embargo, en este trabajo se ha optado por tomar como edad umbral 55 años para obtener grupos poblacionales más homogéneos. En general, en los dos tipos de muerte se observa que la probabilidad de supervivencia es más baja en donantes y receptores jóvenes. Las diferencias más significativas se alcanzan en

la combinación joven–mayor y mayor–joven para DBD, donde la probabilidad es más alta para el primero.



**Figura 6.** Probabilidad de supervivencia del injerto para DBD y cDCD post-trasplante para diferentes combinaciones de donante-receptor por grupo de edad

Algunos trabajos similares de otros autores no han encontrado diferencias significativas de la influencia de los dos tipos de muerte sobre la supervivencia del injerto renal [9]. Sin embargo, estos estudios manejan bases de datos de trasplantes con antigüedad de varias décadas y están centrados en la población británica y americana, con importantes particularidades económicas, culturales, sociopolíticas y de gestión de la sanidad. Un trabajo presentado recientemente [26] puso de manifiesto la menor probabilidad de supervivencia de cDCD en comparación a injertos DBD, incluso a largo plazo. El trabajo recomienda considerar cuidadosamente la utilización de injertos cDCD y esperar, si es posible, injertos DBD o de donantes vivos. Otro trabajo similar reciente [10] identifica un mayor riesgo de retraso de la función del injerto para los trasplantes de riñón DCD y recomienda considerar cuidadosamente otros factores de riesgo en la asignación de órganos, como el tiempo de isquemia caliente funcional o la edad del donante.

#### 4. Conclusiones

En este trabajo se ha presentado un análisis de supervivencia de la influencia del tipo de muerte (DBD y cDCD) sobre la supervivencia del trasplante renal. Hasta el conocimiento de los autores, es el primer análisis de supervivencia aplicado a datos de trasplantes en Andalucía centrado en la influencia del tipo de muerte del donante sobre la supervivencia renal.

Los resultados ponen de manifiesto que el impacto del tipo de muerte sobre la probabilidad de supervivencia del injerto renal es desigual a lo largo de todo el tiempo de seguimiento. En general, la probabilidad de supervivencia es más baja durante todo el tiempo del estudio para trasplantes cDCD que DBD. La influencia del tipo de muerte comienza a ser importante a partir del día 5 y aumentando a partir del día 30 hay una diferencia más significativa hasta los 2 años. Transcurrido ese periodo de tiempo, la dependencia de la probabilidad de supervivencia con el tipo de muerte es poco significativa y escasamente

concluyente dado el limitado número de casos de fracaso cDCD en la base de datos analizada.

El progresivo incremento de injertos cDCD debido a las necesidades de órganos y las diferencias de comportamiento expuestas en este trabajo refuerzan la conveniencia de un adecuado y cuidadoso emparejamiento donante-receptor para maximizar la supervivencia, así como la justificación social y económica de la realización de trasplantes de donantes en parada cardíaca controlada.

#### Agradecimientos

Este trabajo ha sido financiado en parte por la Fundación Mutua Madrileña, bajo el proyecto VÍA-RENAL y parcialmente por la Consejería de Economía y Conocimiento, Junta de Andalucía, a través del proyecto P18-TPJ-3074.

#### Referencias

- [1] J. Sundström et al. Prevalence, outcomes, and cost of chronic kidney disease in a contemporary population of 2.4 million patients from 11 countries: The CaReMe CKD study. *The Lancet Regional Health - Europe*, 2022, p. 100438.
- [2] C. P. Kovesdy. Epidemiology of chronic kidney disease: an update 2022. *Kidney International Supplements*, vol. 12, n° 1, 2022, pp. 7-11.
- [3] D. M. Summers et al. Kidney donation after circulatory death (DCD): state of the art. *Kidney International*, vol. 88, n° 2, 2015, pp. 241-249.
- [4] P. Cerda y G. Varoquaux. Encoding High-Cardinality String Categorical Variables. *IEEE Transactions on Knowledge and Data Engineering*, vol. 34, n° 3, 2022, pp. 1164-1176.
- [5] P. Schober y T. R. Vetter. Kaplan-Meier Curves, Log-Rank Tests, and Cox Regression for Time-to-Event Data. *Anesthesia & Analgesia*, vol. 132, n° 4, 2021, pp. 969-970.
- [6] M. F. Danzer, J. Feld, A. Faldum, y R. Schmidt. Reference curve sampling variability in one-sample log-rank tests. *PLOS ONE*, vol. 17, n° 7, 2022, p. e0271094.
- [7] K. Sasaki et al. Conditional probability of graft survival in liver transplantation using donation after circulatory death grafts – a retrospective study. *Transplant International*, vol. 34, n° 8, 2021, pp. 1433-1443.
- [8] T. Hiramitsu et al. Adult Living-Donor Kidney Transplantation, Donor Age, and Donor-Recipient Age. *Kidney International Reports*, vol. 6, n° 12, 2021 pp. 3026-3034.
- [9] D. M. Summers et al. Analysis of factors that affect outcome after transplantation of kidneys donated after cardiac death in the UK: a cohort study. *The Lancet*, vol. 376, n° 9749, 2010, pp. 1303-1311.
- [10] A. Elmer, M.-L. Rohrer, C. Benden, N. Krügel, F. Beyeler, y F. F. Immer. Organ donation after circulatory death as compared with organ donation after brain death in Switzerland – an observational study. *Swiss Medical Weekly*, n° 7, 2022.

# Realidad Virtual Inmersiva para rehabilitación del miembro superior en patología neurológica

A. de los Reyes Guzmán<sup>1</sup>, V. Herrera Tirado<sup>2</sup>, R. Perales Gómez<sup>3</sup>, S. Ceruelo Abajo<sup>4</sup>, J. Albusac<sup>2</sup>, C. González Morcillo<sup>2</sup>

<sup>1</sup> Unidad de Biomecánica y Ayudas Técnicas, Hospital Nacional de Paraplégicos, Toledo, España [adlos@sescam.jccm.es](mailto:adlos@sescam.jccm.es)

<sup>2</sup> Departamento de Tecnologías y Sistemas de Información, Universidad de Castilla La-mancha, Ciudad Real, España [{Vanesa.Herrera, JavierAlonso.Albusac, Carlos.Gonzalez}@uclm.es](mailto:{Vanesa.Herrera, JavierAlonso.Albusac, Carlos.Gonzalez}@uclm.es)

<sup>3</sup> Unidad de Terapia Ocupacional, Hospital Nacional de Paraplégicos, Toledo, España [rpgomez@sescam.jccm.es](mailto:rpgomez@sescam.jccm.es)

<sup>4</sup> Servicio de Rehabilitación, Hospital Nacional de Paraplégicos, Toledo, España [scruelo@sescam.jccm.es](mailto:scruelo@sescam.jccm.es)

## Resumen

*Actualmente, hay un interés creciente en el uso de tecnología en el campo de la Neurorehabilitación con el objetivo de cuantificar y generar conocimiento acerca de los trastornos sensorimotores tras enfermedades neurológicas, entendiendo que la tecnología posee un elevado potencial para su uso como herramientas terapéuticas. Además, las enfermedades neurológicas producen en la mayor parte de los casos trastornos discapacitantes que perduran durante toda la vida del paciente afectando a su calidad de vida y participación social. En este contexto, se entiende que el proceso rehabilitador de los trastornos motores se debe extender más allá de la fase de ingreso hospitalario del paciente, siendo necesario involucrar tecnología de bajo coste en el proceso de rehabilitación con el objetivo de contar con soluciones tecnológicas que se puedan aproximar al domicilio del paciente. Así, existen soluciones que mediante el uso de tecnología de bajo coste, cuentan con aplicaciones virtuales diseñadas y desarrolladas para la rehabilitación de la destreza manipulativa de los miembros superiores en pacientes con afecciones neurológicas. Además, actualmente existe una tendencia a incorporar metodologías inmersivas en los programas terapéuticos que reciben los pacientes, que ya han sido incorporadas por algunos centros clínicos aunque la evidencia científica aún es escasa. Así, el objetivo principal de esta contribución es presentar el proyecto REHAB-IMMERSIVE en el que se pretende desarrollar una plataforma para la rehabilitación de los aspectos motores de los miembros superiores basada en Realidad Virtual inmersiva, diseñada específicamente para el tratamiento de pacientes neurológicos, concretamente pacientes con lesión medular cervical.*

## 1. Introducción

Actualmente, hay un interés creciente en el uso de tecnología en el campo de la Neurorehabilitación con el objetivo de cuantificar y generar conocimiento acerca de los trastornos sensorimotores tras enfermedades neurológicas, entendiendo que la tecnología posee un elevado potencial para su uso como herramientas terapéuticas [1]. Sin embargo, estas tecnologías, tanto la robótica como la basada en redes de sensores, no poseen actualmente una verdadera implantación en la práctica clínica diaria por varios motivos. Uno de ellos, de elevada importancia, es el elevado coste que poseen, que reduce

su ámbito de aplicación únicamente al de los centros de rehabilitación especializados.

Por otro lado, las enfermedades neurológicas producen en la mayor parte de los casos trastornos discapacitantes que perduran durante toda la vida del paciente afectando a su calidad de vida, nivel de independencia en ejecución de actividades de la vida diaria y participación social. Un ejemplo, es la lesión medular cervical, que carece de un proceso curativo completo y en la que la fuerza residual de músculos parcialmente paralizados es un importante determinante para poseer independencia y funcionalidad en los miembros superiores [2]. En este contexto, se entiende que el proceso rehabilitador se debe extender más allá de la fase de ingreso hospitalario del paciente, siendo necesario involucrar tecnología de bajo coste en el proceso de rehabilitación con el objetivo de contar con soluciones tecnológicas que se puedan aproximar al domicilio del paciente. La tecnología unida al uso de aplicaciones virtuales permitiría la ejecución de ejercicios terapéuticos de la intensidad y número de repeticiones necesarias para inducir la neuroplasticidad, proporcionando las fuentes de feedback necesarias como para motivar y asegurar la adherencia del paciente a las tareas [3]. Además, existe una tendencia al diseño y desarrollo de entornos virtuales inmersivos para rehabilitación motora. Sin embargo, antes de adoptar estos entornos con un propósito rehabilitador, es necesario determinar la equivalencia cinemática entre los movimientos realizados en el entorno virtual con aquéllos realizados en entornos físicos reales. Este aspecto es especialmente importante en los movimientos de alcance y agarre, que dependen de la información acerca del objeto a manipular con el objetivo de planificar las trayectorias de alcance y el tipo de agarre. Este estudio se realizó en personas sanas utilizando head-mounted display junto con guantes como Cybergrasp y Cyberglove y un equipo de fotogrametría para analizar la cinemática [4].

Actualmente, esta modalidad inmersiva se está poniendo en práctica en algunos centros clínicos para el tratamiento de enfermedades neurológicas como Daño Cerebral Adquirido, con el objetivo de trabajar aspectos

relacionados con la atención, percepción y control motor dentro de entornos realistas que simulan situaciones y actividades del día a día [5]. Sin embargo, pese a la realización de estudios en muestras de pacientes, es todavía mucha la investigación necesaria, en términos de efectividad, para poder utilizar estas terapias consistentemente con un propósito rehabilitador, y en términos de viabilidad en cuanto a la duración óptima de estas terapias.

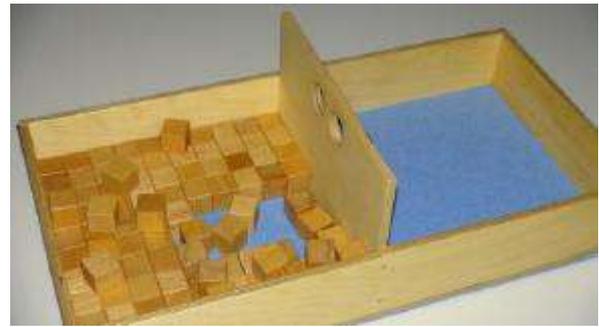
El objetivo de esta contribución es presentar el proyecto de investigación REHAB-IMMERSIVE que tiene como propósito desarrollar una plataforma basada en aplicaciones virtuales inmersivas para rehabilitación de los aspectos motores de los miembros superiores. Dada la complejidad del desarrollo tecnológico y software de la nueva propuesta, ésta se realiza en coordinación con la Escuela Superior de Informática de la Universidad de Castilla-La Mancha. El proyecto parte de los resultados de un proyecto anterior, que consiste en un software basado en aplicaciones virtuales manipuladas por medio de Leap Motion Controller (LMC) para la rehabilitación de los aspectos motores de los miembros superiores en pacientes con afecciones neurológicas [6]. Cada aplicación virtual posee un objetivo terapéutico en cuanto a los aspectos motores del miembro superior y, más concretamente la mano [7]. La novedad radica en crear una plataforma en la que coexistan ambas modalidades terapéuticas, la actual de escritorio y la inmersiva, de la que se hará en el ámbito del proyecto un estudio de su funcionalidad y usabilidad en pacientes con lesión medular cervical. La idea es que los pacientes comiencen con la terapia no inmersiva y poco a poco, y de forma progresiva en función de su tolerancia, se introduzcan en las sesiones experimentales pequeñas dosis de la modalidad inmersiva.

## 2. Métodos

El proyecto se lleva a cabo mediante la participación colaborativa y coordinada de un equipo multidisciplinar de trabajo formado por médicos, terapeutas ocupacionales e ingenieros pertenecientes a los dos centros.

Una metodología basada en Diseño Centrado en el Usuario se aplicará de forma transversal a lo largo de todo el proyecto. A partir de los requisitos clínicos y necesidades en el paciente definidas por el personal clínico y por los propios pacientes se definirán y desarrollarán distintos entornos virtuales inmersivos. La implementación de esta metodología permite llevar a cabo un proceso de diseño y desarrollo iterativo en función de los resultados de las pruebas realizadas con pacientes.

El contenido de los entornos virtuales estará orientado a lograr tareas funcionales completas que requieren la realización de gestos y movimientos del miembro superior y, en concreto de la mano, que son necesarios rehabilitar en el paciente para una mejora de la destreza manipulativa. Para ello se ha comenzado por el desarrollo de una versión inmersiva del test clínico Box and Block [8].



*Figura 1. Test clínico Box and Block*

### 2.1. Desarrollo de la versión inmersiva del test Box and Block

El desarrollo de la aplicación inmersiva Box And Block se ha realizado con Unity [9] e integrado en las gafas de realidad virtual Oculus Quest 2 [10]. Para su implementación se ha utilizado la reciente SDK experimental (Oculus Interaction SDK Experimental) [11], publicada en febrero de 2022 y que incorpora mejoras en el tracking de las manos y facilita una interacción más robusta con diferentes elementos. Para el diseño del juego se ha tenido en cuenta las recomendaciones y pautas marcadas por los terapeutas: tamaño de los bloques, altura y profundidad en la que se encuentra la caja, mano con la que se efectúa el ejercicio, tipo de agarre y almacenamiento de cinemática entre otros. Algunos de los elementos descritos anteriormente son configurables desde el menú inicial.



*Figura 2. Menú inicial para adaptar la terapia al paciente*

### 2.2. Intervención

El proyecto se encuentra en plena fase de desarrollo y la intervención se limita a las pruebas realizadas en 6 personas sanas y dos pacientes con afectación de la función del miembro superior tras una lesión medular. El primero de ellos, varón de 21 años, tiene una lesión medular de nivel metamérico C5 y severidad ASIA C, con un Índice Motor en el miembro superior dominante de 10 sobre 25 puntos. El segundo es un varón de 15 años

con una lesión medular de nivel metamérico C5 ASIA B, con Índice Motor de 6 sobre 25 puntos en el miembro superior dominante. Ambos poseen afectación en la función y fuerza muscular de los miembros superiores y ausencia de función de agarre. Ambas lesiones son de etiología traumática.

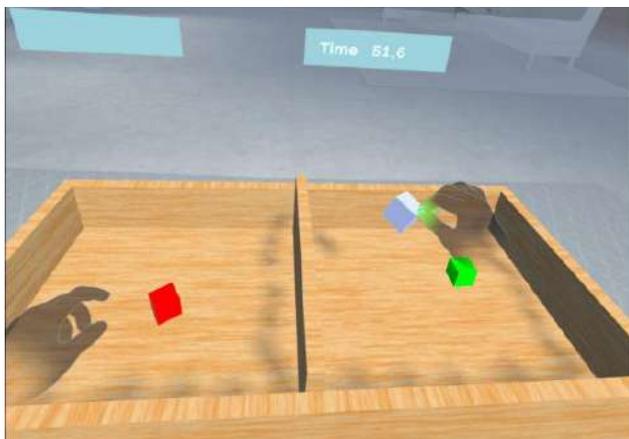
Dado que el proyecto se encuentra en una fase muy incipiente de su desarrollo, en esta primera fase el objetivo fundamental es comprobar la calidad del tracking de las manos por medio del dispositivo utilizado, así como recoger la opinión de los pacientes en cuanto a determinados criterios de diseño.

La aplicación se ejecuta en un ordenador portátil Intel®Core i7-10750 CPU 2.6 GHz, 32GB de RAM y sistema operativo Windows 10 Pro. La tarjeta gráfica es NVIDIA® Quadro T2000 con Max-Q Design.

### 3. Resultados

En todos los participantes, sanos y pacientes, el tracking de las manos mediante Oculus Quest 2 fue muy bueno y realista.

Todos los participantes sanos pudieron realizar el test Box and Block en su versión inmersiva. Aunque los participantes tuvieron que realizar movimientos compensatorios del tronco en algunos momentos para poder llevar a cabo la tarea, el desempeño como versión inicial de prueba fue bueno logrando pasar al otro lado de la partición una media de 76.50 (10.60) cubos. En la Figura 3 se ve un participante sano manipulando la realidad virtual inmersiva.



**Figura 3.** Participante sano ejecutando el test Box and Block en su versión inmersiva mediante Oculus Quest 2

Los dos pacientes con lesión medular que participaron pudieron alcanzar los cubos dentro del entorno virtual inmersivo, pero no pudieron manipularlos debido a la ausencia de función de agarre. Esto indica la necesidad de introducir adaptaciones del entorno inmersivo según las necesidades de los pacientes.

### 4. Discusión

Es necesario tener en cuenta que los dispositivos de Realidad Virtual se han diseñado y desarrollado primeramente con un fin lúdico y de ocio. Sin embargo, esta disciplina, que posee un potencial tremendo en el

tratamiento de pacientes neurológicos, ha de ser adaptada para su correcta aplicación sobre los pacientes. El primer hándicap al que nos enfrentamos es en relación al tracking de las manos. En este caso, no podemos depender del uso de los mandos del dispositivo para interactuar con los entornos virtuales.

Como primera aproximación puesto que se trata actualmente de una versión beta, hemos podido comprobar el tracking tan realista que el dispositivo hace de las manos reconociendo gestos muy sutiles en las personas sanas. En ambos pacientes con lesión medular, el dispositivo ha realizado correctamente el tracking de las manos y no se ha perdido la visualización en todas las pruebas realizadas.

En cuanto al desempeño dentro de la tarea, el Box and Block es un test de destreza manipulativa que consiste en pasar el mayor número de cubos al otro lado de la partición durante 1 minuto. En la versión actual, es necesario realizar una prensión con todos los dedos para coger los cubos en el escenario inmersivo. En estas pruebas iniciales, las personas sanas pudieron ejecutar la tarea sin mayor dificultad. Los dos pacientes que participaron pudieron visualizar sus manos en el entorno virtual inmersivo y alcanzar los cubos que aparecían en la parte derecha de la partición. Sin embargo, no podían cogerlos debido a la función alterada de la mano que presentan y no poder realizar un agarre completo.

Las pruebas iniciales han sido satisfactorias y la opinión del paciente nos lleva a introducir cambios y mejoras que permitan su desempeño de forma óptima en versiones posteriores. Algunas de ellas son las siguientes:

- Altura de la caja configurable, según la posición óptima de las manos que es aquella en la que el codo se encuentra flexionado 90°.
- Reducción de las dimensiones de la caja y de los cubos para adecuarlo al test real, evitando así la realización de movimientos compensatorios del tronco. Los cubos deben tener un tamaño de 2.5 cm, sin embargo, las unidades en Unity no mantienen la misma métrica, por lo que se ha ido refinando el tamaño en función de las pruebas. El entorno mostrado en la Figura 3 muestra este cambio ya implementado.
- Configuración adaptativa de la capacidad de agarre. El agarre máximo que pueda realizar cada paciente ha de ser lo suficiente para poder coger y manipular los objetos. Asimismo la opción de autogrip ha de estar habilitada en caso de que el paciente tenga capacidad nula para abrir y cerrar la mano.

En cuanto a experiencias previas sobre este test desarrollado en formato inmersivo, se tiene evidencia de desarrollos por parte de otros autores para su aplicación en la enfermedad de ictus [12] y Parkinson [13]. Actualmente no disponemos de resultados propios para poder contrastar con estos estudios, dado que nuestro proyecto se encuentra en una etapa muy incipiente de su desarrollo, por lo que se incluirá en posteriores ediciones de este congreso.

## 5. Conclusión

En esta contribución se ha presentado el proyecto REHAB-IMMERSIVE. Debido a que la fase de desarrollo se encuentra en una situación muy incipiente, no se han podido presentar resultados concluyentes. Además, se desarrollarán nuevos escenarios inmersivos dotados de un sentido terapéutico. Por tanto, en sucesivas ediciones se podrá presentar las experiencias en población sana y en pacientes neurológicos.

La aplicación de una metodología de Diseño Centrado en el Usuario es la única vía para lograr resultados satisfactorios verdaderamente adaptados a las necesidades del paciente.

## Agradecimientos

Proyecto PID2020-117361RB-C22 financiado por MCIN/AEI/ 10.13039/501100011033.

## Referencias

- [1] Shirota C, Balasubramanian S, Melendez-Calderon A. Technology-aided assessments of sensorimotor function: current use, barriers and future directions in the view of different stakeholders. *Journal of neuroengineering and rehabilitation*, 16(1), 2019, 53.
- [2] Yozbatiran N, Francisco GE. Robot-assisted Therapy for the Upper Limb after Cervical Spinal Cord Injury. *Physical Medicine and Rehabilitation Clinics*, 2019, 367–384. <https://doi.org/10.1016/j.pmr.2018.12.008>
- [3] Dimbwadyo-Terrer I, Gil-Agudo A, Segura-Fragoso A, de los Reyes-Guzmán A, Trincado-Alonso F, Piazza S, Polonio-López B. Effectiveness of the virtual reality system Toyra on upper limb function in people with tetraplegia: a pilot randomized clinical trial. *BioMed research international*, 2016.
- [4] Magdalon EC, Michaelsen SM, Quevedo AA, Levin MF. Comparison of grasping movements made by healthy subjects in a 3-dimensional immersive virtual versus physical environment. *Acta psychologica*, 138(1), 2011, 126-134.
- [5] Fundación Lescer. <https://www.centrolescer.org/centro-lescer/la-realidad-virtual-al-servicio-de-la-neurorehabilitacion/> (Consultada: Junio 2019). Bronzino JD. The Biomedical Engineering Handbook. Vol 1. CRC Press, 1999 (ISBN: 084930461X).
- [6] de Los Reyes-Guzmán A, Fernández García L, Alvarez-Rodríguez M, Lozano-Berrio V, Domingo-García AM, Ceruelo-Abajo S. Realidad virtual de bajo coste. Una nueva aplicación para rehabilitación motora de los miembros superiores en patología neurológica: Estudio piloto [Low-cost virtual reality. A new application for upper extremity motor rehabilitation in neurological pathology: Pilot study]. *Rehabilitacion (Madr)*. 2022;56(3):173-181. doi:10.1016/j.rh.2021.07.001
- [7] de Los Reyes-Guzmán A, Lozano-Berrio V, Alvarez-Rodríguez M, et al. RehabHand: Oriented-tasks serious games for upper limb rehabilitation by using Leap Motion Controller and target population in spinal cord injury. *NeuroRehabilitation*. 2021;48(3):365-373. doi:10.3233/NRE-201598
- [8] Mathiowetz V, Volland G, Kashman N, Weber K. Adult norms for the Boz and Block Test of manual dexterity. *American Journal of Occupational Therapy*, 39(6), 1985, 386-391.
- [9] Unity. <https://unity.com/> (Consultada: Julio 2022)
- [10] Oculus Quest 2. <https://store.facebook.com/es/es/quest/products/quest-2/> (Consultada: Junio 2022).
- [11] Oculus Interaction SDK Experimental. <https://developer.oculus.com/documentation/unity/unity-isdk-interaction-sdk-overview/> (Consultada: Junio 2022).
- [12] Everard G, Otmane-Tolba Y, Rosselli Z, et al. Concurrent validity of an immersive virtual reality version of the Box and Block Test to assess manual dexterity among patients with stroke. *J Neuroeng Rehabil*. 2022;19(1):7. Published 2022 Jan 22. doi:10.1186/s12984-022-00981-0.
- [13] Oña ED, Jardón A, Cuesta-Gómez A, Sánchez-Herrera-Baeza P, Cano-de-la-Cuerda R, Balaguer C. Validity of a Fully-Immersive VR-Based Version of the Box and Blocks Test for Upper Limb Function Assessment in Parkinson's Disease. *Sensors (Basel)*. 2020;20(10):2773. Published 2020 May 13. doi:10.3390/s20102773.
- [14] Cikajlo I, Peterlin Potisk K. Advantages of using 3D virtual reality based training in persons with Parkinson's disease: a parallel study. *J Neuroeng Rehabil*. 2019;16(1):119. Published 2019 Oct 17. doi:10.1186/s12984-019-0601-1.

# A review on the use of sEMG for the analysis of motor cues in Parkinson's disease: The AI4HealthyAging project

M. Algaba-Vidoy<sup>1</sup>, J.A. Gómez-García<sup>1</sup>, S. Ruiz-Díez<sup>1</sup>, A. Torres-Pardo<sup>1</sup>, F. Oliveira-Barroso<sup>1</sup>,  
D. Torricelli<sup>1</sup>, J.C. Moreno<sup>1</sup>

<sup>1</sup> Neural Rehabilitation Group, Cajal Institute, Spanish National Research Council (CSIC), 28006. Madrid, Spain.  
{marina.algaba, jorge.gomez.garcia, sara.ruiz, adriana.torres, filipe.barroso}@cajal.csic.es;  
{diego.torricelli, jc.moreno}@csic.es

## Abstract

*Parkinson's Disease (PD) is the second most common neurodegenerative disorder. It severely impairs the quality of life of patients, due to the motor and non-motor underlying alterations that it carries. Early diagnosis is not possible, and the current diagnostic procedures depend on the clinical subjective analysis of motor cues in accordance with certain rating scales. To help in the process of a better identification and assessment of PD, motor biomarkers extracted from surface electromyography (sEMG) along with computational methods for the automatic analysis of the disorder have started to gain attention. In this context and seeking to contribute to the body of knowledge of PD analysis using automatic methods, the project AI4HealthyAging was envisioned. The objective of this paper is to introduce AI4HealthyAging, specifically the work package related to Parkinson, and to present a critical literature review on the use of computational methods for the analysis of PD using sEMG. The literature review reveals that more research is needed in the use of machine learning methodologies in PD analysis, as well as in the design of novel biomarkers that describe the disorder and its signs. It also evidences the lack of databases that promote research in the topic and the lack of longitudinal studies.*

## 1. Introduction

After Alzheimer's disease, PD is the most prevalent neurodegenerative disorder (1). It is mainly produced by disruptions of the basal ganglia, specifically by the death of dopaminergic neurons in the substantia nigra pars compacta projecting to the striatum. PD is a multisystemic condition present in both non-motor and motor domains (2,3). Non-motor symptoms include, among others, mental health disorders, cognitive impairment, sleep disorders and other sensory disturbances. Likewise, the main motor symptoms of PD are resting tremor, bradykinesia, rigidity, postural instability and gait disturbance (3).

Although non-motor alterations are of significant interest, the analysis of motor cues is still the guiding principle for the identification of disorder. Indeed, diagnosis is fundamentally clinical, and it is based on the subjective assessment of the motor signs that are manifested in patients, in accordance with clinical scales such as the UPDRS (4). However, these motor signs show up when more than 70% of dopaminergic neurons have already died

(5). To make things worse, current treatments [Levodopa/Carbidopa and Deep Brain Stimulation (DBS)] are only symptomatic (i.e., they don't stop the progression of the disease), and they can produce side effects such as dyskinesias in the case of long-term dopaminergic treatment. In the meantime, because of these motor alterations, PD patients' quality of life, mobility and individual independence are reduced, making them more susceptible to falls and injuries, significantly affecting their own quality of life and that of their caregivers (6).

For the abovementioned motives, plenty of effort has been dedicated to the discovery of new biomarkers that help to enlarge the body of knowledge about the pathophysiology of PD. This is crucial to develop new therapeutic strategies, to assess the degree to which a patient has been affected by the disorder and to reach an earlier diagnosis of the pathology. The advancement in the field has been partially due to the use of novel computational methods, including those based on machine learning (ML), that have served to discover by automatic means methodologies to analyze PD motor alterations. In this context, the AI4HealthyAging project was envisioned, and particularly, the work package dedicated to PD (WP-PD). The objectives of the WP are on the development of novel biomarkers that, by means of ML and computational methods, permit an automatic and objective analysis of the disease. It is expected that the resulting models and outcomes become useful in clinical settings, as supporting tools for the diagnosis and assessment of PD. Although WP-PD in AI4HealthyAging considers different modalities for the evaluation and analysis of motor and non-motor cues in PD patients, this paper will focus on sEMG due to its non-invasiveness. sEMG has been shown useful assessing human motor control, allowing to study myoelectric activity, along its corresponding muscle synergies, during dynamic tasks like gait (7). The objectives of this paper are: i) to introduce the AI4HealthyAging (specifically the WP dedicated to PD) and the scientific objectives that it pursues, and ii) to present a critical analysis on the state of the art on the use of sEMG for the automatic analysis of PD. With respect to the latter, attention has been put into revising schemes that also consider the use of ML within this context.

This paper is organized as follows: section 2 introduces the AI4HealthyAging project, explaining its main objectives,

organization, but then describing WP-PD in AI4HealthyAging, devoted to the analysis and discovery of novel biomarkers for PD diagnosis and assessment. Section 3 reviews the literature and describes the approaches dedicated to the analysis of PD with sEMG, revising as well, those works that have employed ML for data modelling. Finally, section 4 introduces the discussions and conclusions of the paper.

## 2. AI4HealthyAging

AI4HealthyAging, “*Inteligencia artificial distribuida para el diagnóstico y tratamiento temprano de enfermedades con gran prevalencia en el envejecimiento*”, is a multicentric project, running until 31st December 2024, and funded by the Spanish ministry of Economy and Digital transformation within the 2021 MisionesIA call. This in turn is within the Recovery, Transformation and Resilience Plan, Next Generation EU plan, for the recovery and economic growth after the COVID-19 crisis. AI4HealthyAging embodies a consortium of 15 partners including, among others, Cajal Institute of the Consejo Superior de Investigaciones Científicas (CSIC), Universidad Politécnica de Madrid, Universitat Politècnica de Catalunya, Barcelona Supercomputing Center, Cruz Roja Española, Universitat de València, etc.

The aim of AI4HealthyAging is to promote the advance in the study of disorders related with aging, including PD, Alzheimer, sleep disorders, etc. Particularly for this paper, the focus is on the work package 5 (WP5), related to PD: WP-PD. This work is led and carried out by the Neural Rehabilitation Group of the Cajal Institute at CSIC.

The main objective of WP-PD in AI4HealthyAging, is to design systems capable of automatically analyzing, through ML and signal processing, the motor (fine and grain) and non-motor cues of patients affected by PD. The modalities considered in WP-PD in AI4HealthyAging, include non-motor (EEG); fine-grained motor (speech, handwriting and oculometrics); and gross motor (gait using inertial measurement units, sEMG and optoelectronics) cues. For the purposes of this paper, only sEMG is considered because of the innovative computational opportunities to model and assess motor control.

The first specific objective of the project includes the design of systems that automatically distinguish between PD and healthy subjects accurately and reliably through ML and computational methods. Another specific objective is the assessment of the degree on which a certain subject is affected by PD, in accordance with novel clinical scales designed for the purposes of the project, or clinical scales typically used in the clinical diagnosis of PD (UPDRS, H&Y, etc.). The third specific aim is obtaining objective, but also explainable, biomarkers for them to be useful in clinical settings. These explainable biomarkers are designed to correlate to motor or non-motor signs of parkinsonian patients. WP-PD in AI4HealthyAging also aims to register a longitudinal database of 50 patients with PD and 50 healthy controls. These registers will be recorded during three different periods over three years, one at basal time, other 12 months after, and the last one

24 months after that initial record. For the gross motor activities, this database will include recordings of overground and slope walking.

The techniques that are often considered in the analysis of sEMG include those that quantify and characterize the spectral content of the signal, along those that describe the magnitude and coordination of the signal (muscle synergies). However, before the design of any automatic classification/assessment system, it is crucial to examine the state of the literature to find out the type of signal processing techniques, the characteristics, the computational methods, the open databases that are available, and in general the analysis that others have been carried in the context of the analysis of PD with sEMG.

## 3. State of the art

Effective sEMG signal characterization is critical in the diagnosis of neuromuscular disorders (8) because it allows, for instance, to find alterations in gait parameters in a non-invasive manner. In this respect, muscle synergies have been widely used to characterize PD and describe some of the principal motor alterations of the disease. Muscle synergies describe motor alterations, as they provide direct insight into how the central nervous system regulates motor activity. They unveil the brain ability to coordinate a group of muscles, known as motor modules, to work as a single unit in a specific task, and to minimize neural processing (6,9,10). Regarding muscle synergies, it has been proven that there is a decrease in neuromuscular complexity in PD patients associated with decreased gait performance, meaning the existence of fewer number of motor modules and different activation gait profiles while muscle weighting vector were unchanged (11). Moreover, PD patients showed fewer motor modules for balance control compared to healthy subjects (12), reduced synergies when stabilizing the center of mass during balance perturbations (10,13) and disinhibition of sensorimotor feedback in antagonist muscles in center of mass stabilization (14). During the quantification of the effects of PD medication, sEMG is also useful, existing evidence that dopaminergic treatment does not influence synergies during balance perturbations (10,12).

Regarding gait, a higher variability in the control of lower limb muscles in PD individuals was found with sEMG during sub-phases of the gait (15). This fact could explain the significant increase of atypical gait cycles (16) and alterations in normal gait parameters (17). Similarly, in (18,19), while evaluating improvements of muscle activation in PD after hydrotherapy, the authors identified changes in timing and amplitude of leg muscles during gait cycles in comparison with matched healthy subjects. Not only gait and muscle synergies seem to be affected in PD. Indeed, frequential analysis of sEMG features, revealed that the mean frequency content was different among muscles, and between activations of the same muscles during gait (1). Moreover, significant differences in neuromuscular control were found in PD. They were related to muscle force and joint torques: lower magnitude of force and higher amount of variability (20). A reduction

of the ability to inhibit antagonist muscles (21) and a greater proximal and decreased distal activity of lower limb muscles was also reported (3). Those last two effects seemed to be reduced with Levodopa, which increase distal activity and DBS that increase both distal and proximal (3).

Literature also describes the use of novel computational approaches to analyze PD, employing, for instance, ML and artificial intelligence methodologies. In this regard, ML approaches have been explored recently with the objective of classifying subjects in categories. It was found that, compared to traditional ML, [Support Vector Machine (SVM), K-Nearest Neighbour (KNN), Naïve Bayes (NB), etc.], Deep Learning (DL) techniques tend to provide better performance at the expenses of needing significantly larger amounts of data. Typically, DL systems involve the use of sophisticated multi-layered neural systems to perform an automated diagnosis of PD. In this respect, literature report that gait analysis seems to perform better with DL hybrid models (22) compared to conventional ML approaches. In (23), a highly accurate (91.9%), automatic classification system of normal and abnormal sEMG patterns during gait, was reported with Convolution Neural Network (CNN) but not with conventional ML models such as SVM (67.6%) or KNN (48.7%). Authors in (24) also reported a system to classify between controls and PD subjects, using DT and NB, but obtaining moderate results (DT = 88.38%, NB = 54.07%). Most recently, the use of pre-trained DL models with Transfer Learning (TL) illustrated that this approach to PD classification could lead to classification rates up to 99% (25). This strategy has been used efficiently as an automated approach to diagnose PD from sEMG signals, also tackling one of the main drawbacks of DL: the need of huge amount of data for learning.

#### 4. Discussions and conclusions

This paper introduces the WP-PD in AI4HealthyAging project, which is aimed at using ML methodologies and a myriad of modalities for the discovery of novel biomarkers that enhance the knowledge on PD and help on diagnostic and assessment of the disorder in clinical settings. A revision of the state of the art regarding the use of sEMG for the characterization of the gait has been presented, paying special attention to the use of ML for the automatic analysis of parkinsonian gait.

With respect to the literature review, it was observed that sEMG has the potentiality of characterizing motor activity during gait of PD patients. This has been shown by the multiple works presented in the literature: 14 works have employed sEMG to characterize PD and they all succeeded at extracting possible biomarkers to distinguish PD subjects from healthy controls. As previously stated, to perform gait analysis through sEMG, different features (time, frequency, time-frequency and muscle synergies) can be extracted from each specific gait phase. Nevertheless, there is a lack of consensus into which parameters are better. The extraction of muscle synergies seems to be the most common approach. This is because literature has revealed that muscle synergies have great potentiality in detecting gait abnormalities, such as altered

PD gait patterns. However, more work is needed to discover new potential biomarkers to analyze the disease and to characterize disease progression by means of muscle synergies, as just 5 works employ this technique, or try with other parameters such as time or frequency features.

Despite the volume of works considering sEMG for PD detection, the literature also evidenced that it is necessary to design novel biomarkers that help to characterize PD gait. Indeed, most of the works are dedicated to the classification of PD vs healthy controls, but almost none has studied the assessment of the degree to which patients have been affected by the disorder. In the same line, literature reveals a lack of research at early and prodromal stages, which would be of significant relevance for better diagnosis and treatment of the disorder. Furthermore, the biomarkers described on the literature have not been directly associated to specific motor signs of the pathology. For example, it is known that gait disturbance is characterized by slow shuffling steps, asymmetry and high stride-to-stride variability (3), but these features are too general and could be related to any other motor manifestations as well. Therefore, an analysis that relates features and motor signs of PD (i.e., bradykinesia, stiffness, freezing of gait, etc.) would be of interest to distinguish different stages of the disease and assist in diagnostic and assessment tasks.

With respect to the use of ML techniques, the literature reveals the use of this type of methodologies in the analysis of abnormal gait patterns. Currently, most of the works use classical classification algorithms for this purpose, but some others (3 studies) employed DL models with promising results. Another of the biggest issues in using DL is the lack of the data volumes that are needed to make DL algorithms work properly. To solve this matter, authors have started to use transfer learning, employing databases coming from other contexts to pretrain models. Then, the pretrained models are adjusted with the data of PD subjects. Nonetheless, most of the works are incipient and much more research in this direction is required. In the line with the previous issue, literature evidences the lack of openly available databases, as none was found during the review of the literature. Also, no database included a longitudinal study that evaluates the evolution of the disease at different time scales.

As conclusions of the literature review, it can be stated that the use of sEMG and muscle synergies to characterize the muscle activations of PD patients during gait, has great potential, providing good results as evidenced by the literature. Notwithstanding, more research is needed to find novel biomarkers that better correlate to the disorder and describe their signs more accurately, as this has been barely explored before. In the same line, more investigation is required into the use of ML, especially by means of DL algorithms, so that their capabilities are better validated in the context of PD analysis. This, however, requires much more data which is not currently available. There lies one of the aims of the AI4HealthyAging project: to develop a longitudinal database composed by a large cohort of individuals with PD and healthy subjects, and which

investigates a comprehensive set of muscles and extract a range of parameters from sEMG over an extended-time-period in different settings.

## Acknowledgments

This work is partially funded by the Spanish ministry of Economy and Digital transformation within the 2021 MisionesIA call, AI4HealthyAging.

## References

- Romanato M, Strazza A, Piatkowska WJ, Spolaor F, Fioretti S, Volpe D, et al. Characterization of EMG time-frequency content during Parkinson walking: a pilot study. In: 2021 IEEE International Symposium on Medical Measurements and Applications (MeMeA). IEEE; 2021. p. 1–6.
- Parkinson Disease [Internet]. [cited 2022 Sep 4]. Available from: <https://www.who.int/news-room/factsheets/detail/parkinson-disease>
- Islam A, Alcock L, Nazarpour K, Rochester L, Pantall A. Effect of Parkinson's disease and two therapeutic interventions on muscle activity during walking: a systematic review. *NPJ Parkinsons Dis.* 2020 Dec 9;6(1):22.
- Goetz CG, Tilley BC, Shaftman SR, Stebbins GT, Fahn S, Martinez-Martin P, et al. Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): Scale presentation and clinimetric testing results. *Movement Disorders.* 2008 Nov 15;23(15):2129–70.
- Lang AE, Lozano AM. Parkinson's Disease. *New England Journal of Medicine.* 1998 Oct 8;339(15):1044–53.
- Mileti I, Zampogna A, Santuz A, Ascì F, del Prete Z, Arampatzis A, et al. Muscle Synergies in Parkinson's Disease. *Sensors.* 2020 Jun 5;20(11):3209.
- Agostini V, Ghislieri M, Rosati S, Balestra G, Knaflitz M. Surface Electromyography Applied to Gait Analysis: How to Improve Its Impact in Clinics? *Front Neurol.* 2020 Sep 4;11.
- Yousefi J, Hamilton-Wright A. Characterizing EMG data using machine-learning tools. *Comput Biol Med.* 2014 Aug;51:1–13.
- Ballarini R, Ghislieri M, Knaflitz M, Agostini V. An Algorithm for Choosing the Optimal Number of Muscle Synergies during Walking. *Sensors.* 2021 May 11;21(10):3311.
- Freitas SMSF, de Freitas PB, Falaki A, Corson T, Lewis MM, Huang X, et al. Synergic control of action in levodopa-naïve Parkinson's disease patients: II. Multi-muscle synergies stabilizing vertical posture. *Exp Brain Res.* 2020 Dec 17;238(12):2931–45.
- Rodriguez KL, Roemmich RT, Cam B, Fregly BJ, Hass CJ. Persons with Parkinson's disease exhibit decreased neuromuscular complexity during gait. *Clinical Neurophysiology.* 2013 Jul;124(7):1390–7.
- Mileti I, Zampogna A, Taborri J, Martelli F, Rossi S, del Prete Z, et al. Parkinson's disease and Levodopa effects on muscle synergies in postural perturbation. In: 2019 IEEE International Symposium on Medical Measurements and Applications (MeMeA). IEEE; 2019. p. 1–6.
- Falaki A, Huang X, Lewis MM, Latash ML. Impaired synergic control of posture in Parkinson's patients without postural instability. *Gait Posture.* 2016 Feb;44:209–15.
- McKay JL, Lang KC, Bong SM, Hackney ME, Factor SA, Ting LH. Abnormal center of mass feedback responses during balance: A potential biomarker of falls in Parkinson's disease. *PLoS One.* 2021 May 27;16(5):e0252119.
- Keloth SM, Radcliffe PJ, Raghav S, Arjunan SP, Kumar D. Differentiating between Parkinson's disease patients and controls using variability in muscle activation during walking. In: 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC). IEEE; 2020. p. 3158–61.
- Ghislieri M, Agostini V, Rizzi L, Knaflitz M, Lanotte M. Atypical Gait Cycles in Parkinson's Disease. *Sensors.* 2021 Jul 27;21(15):5079.
- Zanardi APJ, da Silva ES, Costa RR, Passos-Monteiro E, dos Santos IO, Kruel LFM, et al. Gait parameters of Parkinson's disease compared with healthy controls: a systematic review and meta-analysis. *Sci Rep.* 2021 Dec 12;11(1):752.
- Spolaor F, Volpe D, Pavan D, Guiotto A, Fichera F, Torresin P, et al. Surface EMG analysis in Parkinson Disease patients before and after underwater gait training. *Gait Posture.* 2017 Sep;57:53–4.
- Volpe D, Spolaor F, Sawacha Z, Guiotto A, Pavan D, Bakdounes L, et al. Muscular activation changes in lower limbs after underwater gait training in Parkinson's disease: A surface emg pilot study. *Gait Posture.* 2020 Jul;80:185–91.
- Romanato M, Volpe D, Guiotto A, Spolaor F, Sartori M, Sawacha Z. Electromyography-informed modeling for estimating muscle activation and force alterations in Parkinson's disease. *Comput Methods Biomech Biomed Engin.* 2022 Jan 2;25(1):14–26.
- Keloth SM, Arjunan SP, Raghav S, Kumar DK. Muscle activation strategies of people with early-stage Parkinson's during walking. *J Neuroeng Rehabil.* 2021 Dec 8;18(1):133.
- Loh HW, Hong W, Ooi CP, Chakraborty S, Barua PD, Deo RC, et al. Application of Deep Learning Models for Automated Identification of Parkinson's Disease: A Review (2011–2021). *Sensors.* 2021 Oct 23;21(21):7034.
- Fricke C, Alizadeh J, Zakhary N, Woost TB, Bogdan M, Classen J. Evaluation of Three Machine Learning Algorithms for the Automatic Classification of EMG Patterns in Gait Disorders. *Front Neurol.* 2021 May 21;12.
- Subba R, Bhoi AK. Feature Extraction and Classification Between Control and Parkinson's Using EMG Signal. In 2020. p. 45–52.
- Rezaee K, Savarkar S, Yu X, Zhang J. A hybrid deep transfer learning-based approach for Parkinson's disease classification in surface electromyography signals. *Biomed Signal Process Control.* 2022 Jan;71:103161.

# Diseño de un dispositivo implantado con transmisión inalámbrica de energía para tratamiento del Glioblastoma mediante TTF

P. Cano Mayo<sup>1</sup>, G. Fernández Moreno<sup>1</sup>, M. Jiménez Carrizosa<sup>1</sup>, R. Ramos Hortal<sup>1</sup>, R. Gil Simoes<sup>2</sup>

<sup>1</sup> Centro de Electrónica Industrial, Universidad Politécnica de Madrid, Madrid, España,

{paula.cmayo, gonzalo.fernandez.moreno}@alumnos.upm.es

{miguel.jimenezcarrizosa, regina.ramos}@upm.es

<sup>2</sup> Servicio de Neurocirugía, Hospital Universitario de la Princesa, Madrid, España

josercardo.gil@salud.madrid.org

## Resumen

*Los glioblastomas son los tumores cerebrales más agresivos del sistema nervioso central y al que la mayoría de los gliomas de grado inferior tienden a evolucionar en su desarrollo. El glioblastoma es más frecuente en adultos y su incidencia es de entre 16-17 casos por cada 100.000 habitantes al año. La edad media de los pacientes con glioblastoma es de 62 años. La tasa relativa de supervivencia a 5 años cuenta con peores resultados que el resto de los tumores cerebrales, con un 9% entre adultos de 45-54 años y tan solo un 6% entre adultos de 55 y 64 años. En 2015, la FDA aprobó un tratamiento denominado “Tumor Treating Fields” que mediante campos eléctricos pulsantes inhiben la reproducción de las células cancerígenas. Actualmente, los dispositivos comerciales se administran en el lecho tumoral de forma no invasiva. Sin embargo, sus efectos terapéuticos se relacionan con la adherencia al tratamiento durante al menos 18 horas diarias, lo que supone un estigma social para estos pacientes al tener que llevar un aparatoso casco que además exige estar afeitado completamente y generando frecuentes complicaciones médicas en la piel. En este trabajo se muestra el diseño de un dispositivo intracraneal implantado que sea autónomo energéticamente mediante transferencia inalámbrica de energía, y que sea capaz de generar dichos campos eléctricos para el tratamiento del glioblastoma.*

## 1. Introducción

El cáncer continúa siendo una de las principales causas de muerte en nuestra sociedad, por ello se necesitan desarrollar nuevas terapias eficaces para su tratamiento. La terapia mediante campos eléctricos pulsantes, “Tumor Treating Fields” (TTF), por su nombre en inglés, ha sido recientemente aprobada por la FDA para el tratamiento en el más agresivo de los tumores cerebrales, el glioblastoma (GBM), debido al incremento en supervivencia y supervivencia libre de enfermedad asociado a este tratamiento [1]. Sus beneficios están basados en el control que ejerce sobre la división celular y por sus efectos antimetabólicos en células con alta tasa de proliferación sin efectos nocivos sobre las células sin alta replicación.

El tratamiento actual se basa en la aplicación sobre el tejido tumoral y peritumoral de campos eléctricos pulsantes de frecuencia intermedia (100-500 kHz) y baja intensidad (1-3 V/cm), mediante un casco con electrodos acoplados que se alimenta mediante una batería externa. Para que sea

efectivo, los pacientes deben llevar el dispositivo entre 18 y 24 horas al día, dado que el factor más importante relacionado con el incremento en supervivencia es la estricta adherencia al tratamiento, alcanzándose, en estos casos, una media de supervivencia global de 28,7 meses desde el diagnóstico y una tasa de supervivencia a 5 años del 29,3%. Sin embargo, el estigma social y el impacto psicológico de llevar puesto un aparatoso dispositivo, además de la reducida esperanza de vida de estos pacientes, y junto con los efectos secundarios locales que produce (irritación cutánea), hace que un porcentaje muy importante de pacientes no cumpla con el tratamiento de una forma adecuada y por lo tanto baje su efectividad [2].

Recientes estudios sugieren que la dosis global de TTFs administrada lo más cercana posible al tumor, pero en el cuero cabelludo, está fuertemente correlacionada con la supervivencia [3]. Además, parece que los TTFs incrementan la respuesta de otros tratamientos como la quimioterapia al ejercer un efecto sensibilizante sobre las células tumorales frente a estas terapias [4].

Los dispositivos comerciales existentes utilizan 2 pares de transductores ortogonales fijados en el cuero cabelludo (afeitado) con las ubicaciones específicas de colocación en función de la localización del tumor [5]. Como alternativa a este tratamiento, se plantea el desarrollo de un dispositivo miniaturizado que pueda ser implantado en un futuro con una apertura craneal menor de 1 cm de diámetro, o bien posicionarlo en el lecho quirúrgico cuando se efectúe cirugía para reseca el tumor, circunstancia que ocurre en una gran mayoría de estos pacientes. Con ello se aseguraría una tasa de adherencia al tratamiento adecuada, disminuyendo entre 10-20 veces los costes actuales, siendo éstos de 25000 dólares/ mes por alquiler del dispositivo. Además, se eliminarían los efectos cutáneos adversos y la carga social/psicológica del enfermo oncológico. Finalmente, como se efectuaría un tratamiento directo sobre las células tumorales sin la atenuación correspondiente al atravesar los tejidos cerebrales y el cráneo, se confía en encontrar resultados más eficaces y eficientes para esta terapia.

En este trabajo se presenta el diseño de un dispositivo que es capaz de generar TTFs, pero que a diferencia de los actuales estaría implantado, por lo que de esta manera el tratamiento se podría aplicar de forma ininterrumpida. Sin embargo, uno de los principales inconvenientes que presentan los dispositivos implantados más convencionales, como por ejemplo los marcapasos o los desfibriladores automáticos implantados (DAI), es la necesidad de reemplazar la batería del dispositivo en la mayoría de los modelos comerciales, lo que implica una nueva cirugía.

Existen numerosos métodos de almacenamiento y transmisión de energía que permiten alimentar los implantes en aplicaciones médicas [6]. De entre ellos, la transmisión inalámbrica de energía (WPT, *Wireless Power Transfer*) se ha postulado como un método prometedor en el ámbito biomédico ya que ofrece importantes beneficios tales como implantes más pequeños, eliminación de cables y conectores, mayor vida útil o mejora de la eficiencia del implante. En los últimos años se han investigado y desarrollado micro-sistemas biomédicos que presentan esta tecnología, como cápsulas endoscópicas, prótesis de retina, marcapasos, corazones artificiales, etc...[6].

Generalmente, se emplea la técnica de transmisión de energía inductiva para alimentar dispositivos implantados de forma inalámbrica a bajas frecuencias, ya que los campos magnéticos no sufren una gran atenuación a través del cuerpo humano en este rango de frecuencias. Además, para los enlaces inductivos, la eficacia de transmisión de energía es muy alta y es una técnica muy utilizada debido a su simplicidad y alta seguridad.

## 2. Diseño y metodología.

### 2.1. Diseño del sistema implantado

En la Figura 1 se presenta el esquema del sistema global que consta de una unidad externa y otra implantada alimentada mediante transmisión inalámbrica de potencia (WPT).

El dispositivo implantado consta de un módulo receptor que capta la energía de forma inalámbrica, la etapa de potencia para controlar dicha energía y almacenarla en una batería recargable biocompatible, y finalmente el módulo de estimulación que es capaz de generar los campos eléctricos pulsantes (TTF).

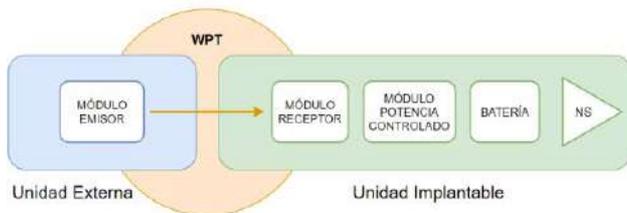


Figura 1. Sistema global

### 2.2. Enlace inductivo

El diagrama de bloques básico de un WPT de acoplamiento inductivo (IC, Inductively Coupled) se presenta en la Figura 2. Normalmente, se coloca un sensor en la unidad implantada y el dispositivo de control fuera, dando lugar a

la posibilidad de comunicación bidireccional. Un factor que afecta significativamente a la eficiencia del enlace es la desalineación de las bobinas, por lo que éstas deberán permanecer correctamente alineadas para que exista una transmisión de energía eficiente y para conseguir una buena comunicación entre el primario y el secundario.

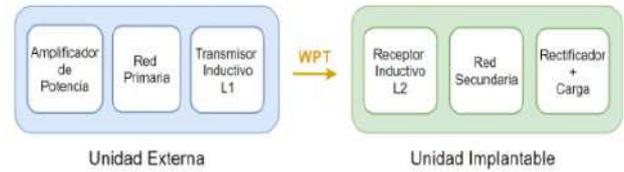


Figura 2. Unidades externa e implantada

Para satisfacer los requisitos de biocompatibilidad del implante, su estructura debe ser compacta y la eficiencia energética lo más alta posible, con el fin de alargar la vida útil de éste. El objetivo principal de la arquitectura es transmitir energía de forma inalámbrica para ello se ha de trabajar en corriente alterna (CA), la cual es generada en el llamado módulo emisor de la Figura 1, y se transforma en corriente continua (CC) a la salida del módulo receptor. Dicha conversión debe tratarse de una transmisión estable y eficiente.

Una vez conseguida dicha conversión se obtiene una alimentación en CC que a través de una etapa de regulación (módulo potencia controlado en la Figura 1) controla la tensión e intensidad con distintos modos de operación para realizar la carga de la batería. Es decir, la etapa de control se realiza entre el módulo de WPT y la batería, y se explicará en la sección 2.3

### 2.3. Módulo de control de potencia

El módulo de control de la potencia para la recarga de la batería viene mostrado en la Figura 3, donde el reductor es el encargado de controlar la potencia a través del BMS, "Battery Management System". El principal desafío que presenta la etapa de control es asegurar la estabilidad de los lazos de control para la función de carga de la batería.

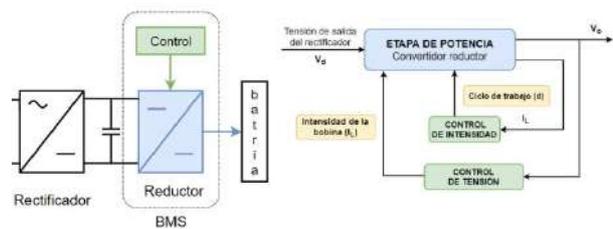


Figura 3. Control de tensión y de intensidad

A partir del modelo promediado del reductor se realiza la etapa de control de tensión e intensidad, la cual consta de:

-**Lazo interno:** encargado del control de la intensidad mediante un regulador PI, el cual actuará sobre el ciclo de trabajo del interruptor del convertidor reductor con el fin de cumplir con la intensidad requerida

-**Lazo externo,** encargado del control de tensión de la batería.

## 2.4. Batería

La batería seleccionada para este trabajo es de la familia de las de Litio-ión recargable con una capacidad de 130 mAh y una tensión nominal de 3.7V tal y como muestra la figura 4.



Figura 4. Batería 130 mAh GMB401524

## 2.5. Generador de TTF

El módulo NS es capaz de generar campos eléctricos de baja intensidad (1-3 V/cm) y alta frecuencia (100-300 kHz) para poder aplicar los TTF. Para ello se ha seleccionado el microcontrolador PIC12F1572 de la familia de microcontroladores PIC de 8 bits que será el encargado de la generación de los TTF.

La elección de este modelo se debe, entre otras especificaciones, a su capacidad para trabajar a baja potencia, su rango de tensión de funcionamiento, su bajo sobrecalentamiento y su comodidad a la hora de la programación y generación del PWM.

El entorno de desarrollo utilizado para la programación del microcontrolador es el MPLAB X Integrated Development Environment (MPLAB-X-IDE). Por otro lado, la plataforma de desarrollo que soporta el microcontrolador es la placa DM160228 con funciones Explorer 8. Dicha placa tiene una configuración de 6 a 100 pines, lo que hace que sea una placa muy completa. Por último, para conseguir la transferencia efectiva del código programado en el entorno de desarrollo al microcontrolador insertado en la placa Explorer 8 se ha utilizado el depurador en circuito MPLAB PICkit 4

## 3. Resultados

### 3.1. Simulación del control del BMS

En este apartado se muestran las simulaciones de la etapa de control de para la recarga de la batería. Dichas simulaciones han sido desarrolladas en Matlab-Simulink. Los parámetros del convertidor reductor vienen reflejados en la Tabla 1:

Tensión de entrada	15 V
Tensión de salida	3.7 V
Frecuencia de conmutación	125 kHz.
Condensador de salida	2,2 $\mu$ F
Bobina	250 $\mu$ H

Tabla 1. Parámetros del convertidor reductor

Con el fin de obtener unas especificaciones dinámicas en las respuestas tanto de la corriente como de la tensión, se definió una sobreoscilación máxima del 5% ( $M_p < 5\%$ ) y un tiempo de establecimiento máximo de 5 ms ( $t_s < 5$  ms). Para

ello se utilizaron dos reguladores PI para el lazo externo ( $R_1$ ) e interno ( $R_2$ ) respectivamente:

$$R_1 = 0.054 \frac{(s+10000)}{s} \quad (1)$$

$$R_2 = 500 \frac{(s+200)}{s} \quad (2)$$

En la Figura 5 se muestra la respuesta dinámica del sistema ante un escenario de cambio en la referencia de la tensión. En la gráfica superior se observa la corriente por la bobina, la cual no supera nunca los 0.1 A, mientras que en la gráfica inferior se observa cómo la tensión en la batería alcanza siempre el valor de referencia y cómo cumple con las especificaciones dinámicas.

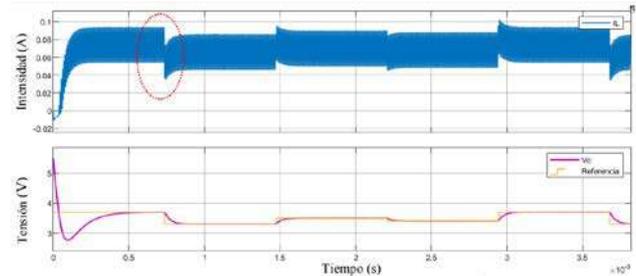


Figura 5. a) Corriente en el reductor b) Tensión en el reductor

En la Figura 6 se muestra el transitorio en el cambio de referencia para  $t=0.74$ s. Se puede comprobar cómo se cumplen con las especificaciones dinámicas y el bajo rizado en la tensión.

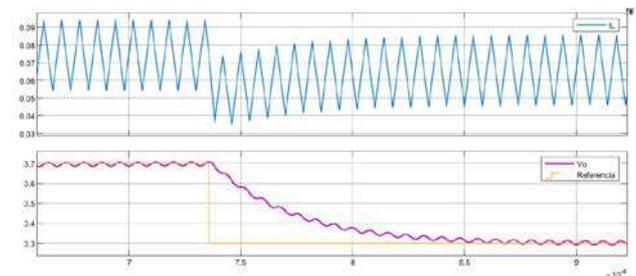


Figura 6. a) Detalle de la corriente en el reductor b) Detalle de la tensión en el reductor

### 3.2. Experimentación del generador de TTF

En esta sección se mostrarán los experimentos para el modulo generador de campos eléctricos.

En la Figura 7 se muestra el microcontrolador PIC12F1572 para la generación de la señal, así como las conexiones a la placa de desarrollo Explorer 8 del depurador de circuito MPLAB PICkit 4.

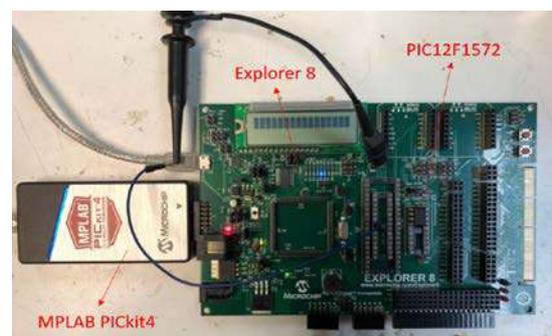


Figura 7. Placa de desarrollo

Para poder hacer la medida del campo eléctrico se ha utilizado el set de sondas isotrópicas del distribuidor R&S (HZ-15 E and H near-field) capaces de medir tanto campo eléctrico como magnético [7]. Entre las características se destaca que tienen un ancho de banda ajustable de 10 MHz y 1 MHz, un conector de salida SMA (para alta frecuencia), capacidad de conectarse a cualquier entrada estándar de un osciloscopio de 1 M $\Omega$  de impedancia y el ajuste automático del offset. La sensorización de corriente utiliza un ancho de banda CC de 10 MHz, con una sensibilidad de 0.25 V/A, una corriente máxima de  $\pm 5$  A, una precisión del  $\pm 15\%$  y una impedancia de inserción de 100 nH en paralelo con 70  $\Omega$ . El ruido es de 25 mA p-p con un ancho de banda de 10 MHz, y de 10 mA p-p con un ancho de banda de 1 MHz.

En la Figura 8 se muestran las medidas de campo eléctrico que genera el PIC12F1572 situándose la sonda a 1 cm del generador. Se puede apreciar cómo se ha generado un campo eléctrico alterno de frecuencia 300 kHz y cuyo módulo es 3 V/cm.

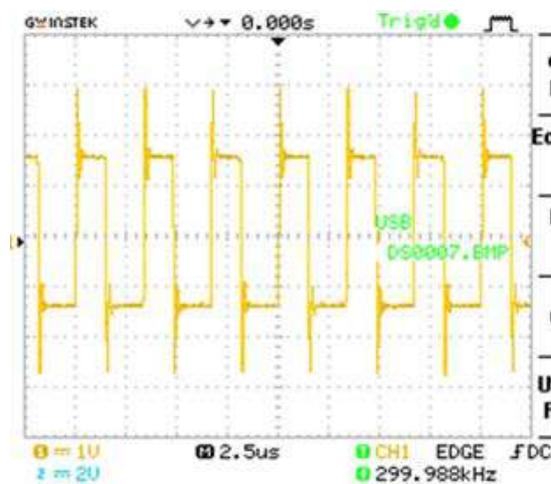


Figura 8. Medida de campo eléctrico en osciloscopio

#### 4. Discusión

Por una parte, las simulaciones muestran cómo la recarga de la batería mediante el BMS cumple con las especificaciones dinámicas impuestas, ya que la corriente por la batería nunca supera los 0.1 A. Por su parte la tensión alcanza siempre el valor de referencia y con un rizado menor al 5 %.

También se destaca que los resultados experimentales del módulo generador de los campos eléctricos confirman que éstos cumplen con las especificaciones eléctricas para poder aplicar los TTF.

#### 5. Conclusiones

En este trabajo se presenta un diseño de un dispositivo implantado miniaturizado que es capaz de generar campos eléctricos variables para aplicar en el tratamiento de los GBM gracias a la terapia TTF y que es energéticamente autónomo debido a la recarga inalámbrica.

En particular, se ha mostrado el diseño de los diferentes módulos del dispositivo: etapa receptora del enlace inductivo, gestión de la recarga de la batería y el módulo

generador de los TTF gracias al microcontrolador PIC12F1572.

Como trabajos futuros se tiene previsto poder verificar los TTF en cultivos celulares, mientras que se va a desarrollar un módulo de comunicación del dispositivo implantado para poder remitir al exterior el estado del mismo.

#### Agradecimientos

Esta actividad de investigación ha sido financiada por el Programa “CEI Grants Program 2022” del Centro de Electrónica Industrial de la Universidad Politécnica de Madrid.

#### Referencias

- Rivera F, Benavides M, Gallego J, Guillen-Ponce C, Lopez-Martin J, Küng M. Tumor treating fields in combination with gemcitabine or gemcitabine plus nab-paclitaxel in pancreatic cancer: Results of the PANOVA phase 2 study. *Pancreatology*. 2019 Jan;19(1):64-72. doi: 10.1016/j.pan.2018.10.004. Epub 2018 Oct 17. PMID: 30396819.
- Rominiyi, O., Vanderlinden, A., Clenton, S.J. et al. “Tumour treating fields therapy for glioblastoma: current advances and future directions” *Br J Cancer* 124, 697–709 (2021). <https://doi.org/10.1038/s41416-020-01136-5>
- Ballo MT, Urman N, Lavy-Shahaf G, Grewal J, Bomzon Z, Toms S. “Correlation of Tumor Treating Fields Dosimetry to Survival Outcomes in Newly Diagnosed Glioblastoma: A Large-Scale Numerical Simulation-Based Analysis of Data from the Phase 3 EF-14 Randomized Trial”. *Int J Radiat Oncol Biol Phys*. 2019 Aug 1;104(5):1106-1113. doi: 10.1016/j.ijrobp.2019.04.008. Epub 2019 Apr 23. PMID: 31026557.
- Kesari S, Ram Z. “Tumor-treating fields plus chemotherapy versus chemotherapy alone for glioblastoma at first recurrence: a post hoc analysis of the EF-14 trial”. *CNS Oncol*. 2017;6:185–193
- Novocure. Optune: instructions for use. 2019; Available from: [https://www.optune.com/Content/pdfs/Optune\\_IFU\\_8.5x11.pdf](https://www.optune.com/Content/pdfs/Optune_IFU_8.5x11.pdf). Última visita 15 Agosto 2022.
- R. V. Taalla, M. S. Arefin, A. Kaynak, and A. Z. Kouzani, “A review on miniaturized ultrasonic wireless power transfer to implantable medical devices,” *IEEE Access*, vol. 7, pp. 2092–2106, 2019, doi: 10.1109/ACCESS.2018.2886780.
- R&S - Cómo funciona la depuración de EMI con un osciloscopio,» [En línea]. Available: [https://www.rohde-schwarz.com/es/productos/test-y-medida/osciloscopios/educational-content/como-funciona-la-depuracion-de-emi\\_254516.html](https://www.rohde-schwarz.com/es/productos/test-y-medida/osciloscopios/educational-content/como-funciona-la-depuracion-de-emi_254516.html). [Último acceso: 02 Septiembre 2022].

# Estudio de las alteraciones en la arquitectura temporal de la actividad neuronal provocadas por la enfermedad de Alzheimer

P. Carretero-Calvo<sup>1</sup>, P. Núñez<sup>2</sup>, V. Rodríguez-González<sup>1,3</sup>, M.A. Tola-Arribas<sup>4</sup>, M. Cano del Pozo<sup>5</sup>, C. Gómez<sup>1,3</sup>, J. Poza<sup>1,3,6</sup>

<sup>1</sup> Grupo de Ingeniería Biomédica, Universidad de Valladolid, Valladolid, España, pablo.carretero.calvo@alumnos.uva.es

<sup>2</sup> Coma Science Group, GIGA-Consciousness, University of Liège, Bélgica

<sup>3</sup> Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), España

<sup>4</sup> Departamento de Neurología, Hospital Universitario Río Hortega, Valladolid, España

<sup>5</sup> Departamento de Neurofisiología Clínica, Hospital Universitario Río Hortega, Valladolid, España

<sup>6</sup> IMUVA, Instituto de Investigación en Matemáticas, Universidad de Valladolid, Valladolid, España

## Resumen

*La enfermedad de Alzheimer (EA) es una de las principales causas de demencia. Cada vez más estudios evidencian que el diagnóstico y caracterización de la EA puede apoyarse en el análisis de señales de electroencefalografía (EEG). Este estudio se ha enfocado en la caracterización de la EA mediante métodos conectómicos, que permiten cuantificar las propiedades dinámicas de la conectividad funcional que presentan las agrupaciones de neuronas. La base de datos utilizada estaba compuesta por 160 sujetos: 43 sujetos de control cognitivamente sanos, 50 pacientes con deterioro cognitivo leve (DCL) por EA y 67 pacientes con EA. A partir de registros EEG obtenidos durante el reposo, se extrajeron patrones de conectividad neuronal recurrentes, denominados meta-estados. En base a la secuenciación temporal de los meta-estados cerebrales, se calcularon dos nuevas medidas conectómicas: la entropía de transición, que caracteriza los patrones de transición entre meta-estados; y la singularidad del meta-estado dominante, que cuantifica el nivel de definición del meta-estado dominante. Los resultados mostraron diversas alteraciones asociadas al DCL y a la EA, en las bandas zeta y alfa, que se resumen en perturbaciones en la comunicación entre regiones cerebrales, pérdida de modularidad, aparición de oscilaciones más aleatorias y pérdida de estabilidad y estructura temporal. Estos hallazgos sugieren que las medidas conectómicas pueden ser una herramienta útil para el apoyo en la caracterización de la EA.*

## 1. Introducción

Para comprender el funcionamiento del cerebro es necesario conocer el comportamiento de los grupos neuronales que lo conforman. Cuando este órgano recibe estímulos, se produce un flujo de campos eléctricos y magnéticos a través de las agrupaciones de neuronas. Cuando dicho estímulo atraviesa una agrupación lo suficientemente grande, el campo eléctrico puede medirse desde fuera del cráneo mediante el uso de electrodos. Esta técnica no invasiva, de bajo coste y alta resolución temporal, se conoce como electroencefalografía (EEG) [1].

La naturaleza de las señales de EEG sufre alteraciones provocadas por diversas enfermedades, entre las que se encuentra la enfermedad de Alzheimer (EA). La EA es una

de las principales causas de demencia y produce cambios paulatinos en la conducta, en la funcionalidad y en las capacidades cognitivas de los pacientes. Al comienzo de la progresión de la EA puede distinguirse una etapa prodrómica conocida como deterioro cognitivo leve (DCL), donde las alteraciones se distinguen del envejecimiento normal sin llegar a considerarse demencia [2].

Hasta el momento, la mayoría de los estudios de EEG que estudian el acoplamiento de la activación funcional entre regiones cerebrales han analizado los patrones de conectividad funcional estática (sFC) [3], siguiendo el enfoque clásico de la conectómica. Sin embargo, en los últimos años, el estudio de la dinámica de las redes neuronales funcionales ha cobrado gran importancia, ya que cada vez más estudios sugieren que la variabilidad del acoplamiento neuronal es fundamental para entender cómo el cerebro procesa y transmite información neuronal [4,5]. Es en este punto donde cobran importancia las técnicas conectómicas o de conectividad funcional dinámica (dFC), que permiten evaluar las propiedades altamente dinámicas de la red cerebral [6].

En un estudio anterior [7] se introdujo un nuevo método de detección de patrones de conectividad funcional recurrentes (denominados meta-estados) en señales EEG para explorar las alteraciones en la arquitectura temporal debidas al DCL y a la demencia por EA. Los resultados obtenidos indicaron que la progresión de la EA afecta a la secuenciación de los meta-estados, probablemente debido a un déficit en la flexibilidad de la red cerebral. En el presente estudio, se pretende avanzar sobre dicho estudio de las alteraciones que la EA provoca en la estructura temporal de la actividad neuronal. La hipótesis de partida asume que el avance de la EA está asociado a una menor capacidad para reconfigurar la red neuronal funcional, incluso en el estado de reposo. Esto se traduce en una paulatina desestructuración de la arquitectura temporal reflejada en la secuenciación de los meta-estados funcionales. El objetivo del estudio consiste en cuantificar estas alteraciones mediante nuevos parámetros

cronectómicos, capaces de caracterizar la evolución temporal de los meta-estados.

## 2. Materiales

### 2.1. Participantes

La base de datos empleada en este estudio estaba formada por 160 sujetos: 43 controles de edad avanzada y cognitivamente sanos, 67 pacientes con DCL debido a EA y 50 enfermos con demencia debida a EA. La Tabla 1 resume los datos clínicos y sociodemográficos más relevantes. El diagnóstico de los pacientes se realizó en base a los criterios NIA-AA (*National Institute on Aging and Alzheimer's Association*). Todos los participantes y familiares fueron informados sobre el protocolo y el propósito del estudio, y dieron su consentimiento informado para ser incluidos. El estudio fue aprobado por el comité ético del Hospital Río Hortega de Valladolid, de acuerdo con el Código Ético de la Asociación Médica Mundial (Declaración de Helsinki).

Los análisis estadísticos indicaron que no había diferencias estadísticamente significativas entre grupos para la edad ( $p$ -valor  $> 0.05$ , test de Kruskal-Wallis) ni para el sexo ( $p$ -valor  $> 0.05$ , test Chi cuadrado). Sí que se hallaron diferencias significativas en el nivel de educación ( $p$ -valor  $< 0.05$ , test Chi cuadrado) y en la puntuación del *Mini-Mental State Examination* (MMSE) ( $p$ -valor  $< 0.05$ , test de Kruskal-Wallis).

Grupo	C	DCL	EA
<i>N</i>	43	67	50
Edad (años) (m[IQR])	76 [74 79]	77 [72 81]	79 [76 82]
Sexo (H:M)	13:30	29:38	23:27
Educación (A:B)	16:27	41:26	35:15
MMSE (m[IQR])	29 [28 30]	27 [26 28]	22 [20 24]

**Tabla 1.** Características clínicas y sociodemográficas de los grupos bajo estudio. H: hombres; M: mujeres. A: educación primaria o inferior; B: educación secundaria o superior. m: mediana; [Q1 Q3]: rango intercuartil.

### 2.2. Registros EEG y preprocesado

Para cada sujeto se adquirieron 5 minutos de actividad EEG en reposo, con los ojos cerrados, en el Departamento de Neurofisiología Clínica del Hospital Universitario Río Hortega de Valladolid. Los registros se realizaron mediante un electroencefalógrafo de 19 canales (XLTEK, Natus Medical), colocados según las especificaciones del sistema internacional 10-20, con una referencia común y una frecuencia de muestreo de 200 Hz.

Los registros fueron preprocesados para minimizar la presencia de ruido de acuerdo a los siguientes pasos [7]: (i) análisis de componentes independientes (ICA) para eliminar las componentes ruidosas; (ii) filtrado paso banda con un filtro FIR (*Finite Impulse Response*) entre 1 y 70

Hz, y filtrado de ranura para eliminar la interferencia de red a 50 Hz; y (iii) selección de los primeros 60 s libres de artefactos mediante inspección visual.

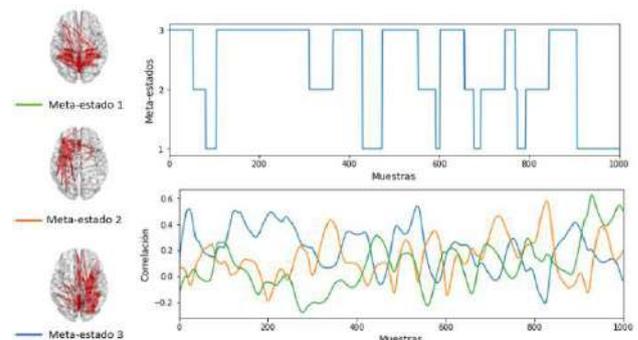
Las épocas de 60 s se proyectaron a nivel de fuentes mediante el algoritmo sLORETA (*standardized low-resolution brain electromagnetic tomography*) y se agruparon en las 68 regiones de interés (ROI) del atlas de Desikan-Killiany [8].

## 3. Métodos

### 3.1. Detección de los meta-estados cerebrales

A continuación, se describe sucintamente la metodología introducida por Núñez *et ál.* [7] para identificar meta-estados cerebrales y generar las secuencias temporales empleadas en el presente trabajo:

- i. Mediante la IAC (*instantaneous amplitude correlation*) [9] se generó un vector de conectividad instantánea para cada una de las muestras de las series temporales a nivel de fuente (68 ROI x 68 ROI x 12000 muestras) en cuatro bandas de frecuencia convencionales (delta: 1-4 Hz; zeta: 4-8 Hz; alfa: 8-13 Hz; y beta-1: 13-19 Hz). Las bandas beta-2 y gamma no pudieron analizarse debido al elevado coste computacional asociado.
- ii. Se calcularon diagramas de recurrencia a partir de los vectores de conectividad instantánea.
- iii. Se aplicó el algoritmo de detección de comunidades, Louvain-GJA, a los diagramas de recurrencia para identificar topologías de red recurrentes (*i.e.*, meta-estados).
- iv. Se alinearon los tres meta-estados identificados para cada sujeto en base a su similitud topológica, calculada mediante la distancia de correlación.
- v. Se calcularon dos secuencias temporales para cuantificar las fluctuaciones de los meta-estados en los 60 s de cada época: (a) la *temporal activation sequence* (TAS), que es una representación discreta del meta-estado dominante en cada instante; y (b) el *instantaneous correlation tensor* (ICT), que es el tensor de la correlación de Spearman de cada meta-estado con la IAC en cada instante. La Figura 1 ilustra los primeros 5 s del TAS y del ICT para un paciente con DCL.



**Figura 1.** Representación del TAS (fila superior) y del ICT (fila inferior) para los primeros 5 s de un paciente con DCL, que tiene 3 meta-estados.

### 3.2. Nuevas métricas cronectómicas

El objetivo principal de este trabajo consistió en caracterizar la arquitectura temporal asociada a la activación dinámica de los meta-estados cerebrales. Para ello, proponemos dos nuevas métricas cronectómicas que permiten: (i) caracterizar la pérdida de estructura temporal; y (ii) el grado de definición del meta-estado dominante.

#### Entropía de transición

La entropía de transición se basa en el TAS. Concretamente, agrupa las muestras temporales de esta serie simbólica en símbolos de longitud  $L$  y analiza la distribución de probabilidades de transición entre ellos. La siguiente ecuación define la medida, considerando que  $X$  es una variable aleatoria discreta que se corresponde con la transición entre símbolos,

$$H(X) = - \sum_{i=1}^M \sum_{j=1}^M p(x_{i \rightarrow j}) \cdot \log[p(x_{i \rightarrow j})]$$

donde  $M$  representa el número total de símbolos y  $p(x_{i \rightarrow j})$  la probabilidad de que se produzca la transición desde el símbolo  $i$  al  $j$ .

#### Singularidad del meta-estado dominante

Esta medida se define en base al ICT. En cada muestra, se selecciona el mayor valor de correlación, que corresponde al meta-estado dominante,  $\rho_1[n]$ . A este valor se le resta el valor de correlación inmediatamente inferior en el mismo instante,  $\rho_2[n]$ , que corresponde a uno de los otros dos meta-estados. La medida se define según la siguiente ecuación,

$$S = \frac{1}{N} \sum_{n=1}^N (\rho_1[n] - \rho_2[n])$$

donde  $N$  representa el número total de muestras del ICT. Un valor de  $S$  alto indicaría que el meta-estado dominante está bien definido con respecto al resto; mientras que un valor bajo de  $S$  estaría relacionado con una secuenciación de topologías de red con distancias de correlación similares para dos meta-estados, indicando un menor grado de dominancia.

### 3.3. Análisis estadísticos

En primer lugar, las métricas cronectómicas introducidas se normalizaron para garantizar que los resultados obtenidos se debían a dFC de origen neural y no a fluctuaciones aleatorias como ruido. Para ello, se generaron 100 secuencias subrogadas mediante AAFT (*amplitude adjusted Fourier transform*), aplicando la misma aleatorización para cada ROI [7]. Las métricas indicadas en el apartado 3.2 se normalizaron dividiendo sus valores por el valor promedio obtenido a partir de los datos subrogados.

Las diferencias entre grupos se evaluaron mediante test no paramétricos. Concretamente, se usó el test de Kruskal-Wallis (KW) para identificar interacciones globales entre los tres grupos y el test  $U$  de Mann-Whitney (MW) para evaluar las diferencias entre pares de grupos. Se aplicó la

corrección FDR (*false discovery rate*) para corregir por el número de bandas de frecuencia y grupos.

## 4. Resultados y discusión

### 4.1. Desestructuración temporal

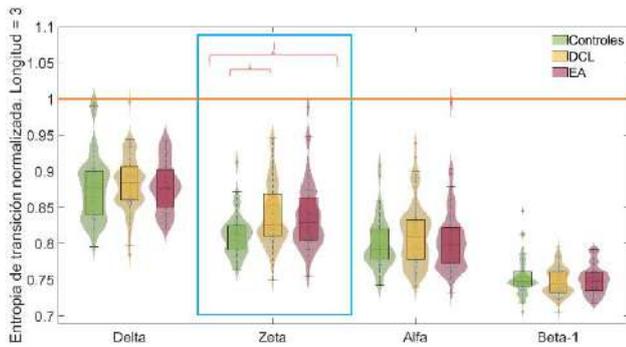
La Figura 2 muestra los resultados obtenidos tras calcular la entropía de transición para una longitud de símbolo de valor 3. Tras analizar diferentes longitudes de símbolo ( $L=1, \dots, 7$ ), se seleccionó el valor  $L=3$  porque era el más sensible a las alteraciones de la EA. Se observaron interacciones globales estadísticamente significativas en la banda zeta ( $p=0.0070$ , test KW, corrección FDR). Al comparar entre pares de grupos en esta banda, se obtuvieron diferencias significativas entre controles y pacientes con DCL ( $p=0.0025$ , test MW, corrección FDR), y entre controles y pacientes con EA ( $p=0.0059$ , test MW, corrección FDR). En ambas situaciones los pacientes con DCL y EA presentaban una entropía más alta que los controles. La presencia de valores de entropía menores en los controles indica una menor variedad de transiciones en comparación con pacientes con DCL o EA, apoyando la hipótesis planteada por [7] sobre la existencia de una estructura temporal subyacente que se va perdiendo progresivamente conforme avanza la EA. Asimismo, se observa que los valores de entropía normalizados son menores que 1, lo cual indica que los patrones observados se deben a dFC de origen neural y apoya la idea anterior de la existencia de una estructura temporal, no enteramente aleatoria, en la actividad neuronal en reposo. Los resultados a su vez apoyan la idea de que la transmisión de información en el cerebro podría darse a diferentes escalas temporales, estando algunas más afectadas por la EA y sugiriendo la utilidad de estudiar con diferentes niveles de granularidad la conectividad instantánea.

### 4.2. Pérdida de definición en la secuenciación de meta-estados

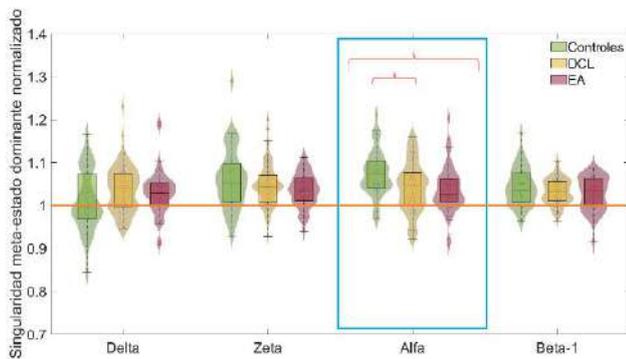
La Figura 3 recoge los resultados obtenidos tras calcular la singularidad del meta-estado dominante. Únicamente se observó una interacción global significativa en la banda alfa ( $p=0.0084$ , test KW, corrección FDR). En esta banda, se obtuvieron diferencias estadísticamente significativas entre controles y pacientes con DCL ( $p=0.0094$ , test MW, corrección FDR), y entre controles y pacientes con EA ( $p=0.0017$ , test MW, corrección FDR). La presencia de valores normalizados más bajos del parámetro en pacientes con DCL y EA, y más cercanos a 1, indica la presencia de meta-estados peor definidos que en el caso de los controles. Estas perturbaciones en la banda alfa podrían estar relacionadas con la reducción en la eficiencia de la comunicación entre las regiones cerebrales de acuerdo con el estudio realizado por Miraglia *et al.* [10]. La reducción en la definición de la secuenciación de meta-estados en los pacientes con DCL y EA se alinea con la pérdida de modularidad en la banda alfa observada por Núñez *et al.* [7], que fue relacionada con una cognición más inestable.

### 4.3. Limitaciones y líneas futuras de investigación

El presente estudio tiene varias limitaciones que motivan futuras investigaciones. En primer lugar, la base de datos analizada está formada por 160 sujetos, lo cual limita la



**Figura 2.** Entropía de transición para  $L=3$  en cada banda de frecuencia. El cuadrado azul indica diferencias significativas globales entre grupos ( $p < 0.05$ , test de Kruskal-Wallis, corrección FDR). Las llaves rojas indican diferencias significativas entre pares de grupos ( $p < 0.05$ , prueba U de Mann-Whitney, corrección FDR).



**Figura 3.** Singularidad del meta-estado dominante en cada banda de frecuencia. El cuadrado azul indica diferencias significativas globales entre grupos ( $p < 0.05$ , test de Kruskal-Wallis, corrección FDR). Las llaves rojas indican diferencias significativas entre pares de grupos ( $p < 0.05$ , prueba U de Mann-Whitney, corrección FDR).

capacidad de generalización de los resultados. Además, se podrían considerar una división en dos subgrupos de los pacientes con DCL, en función de su evolución o no a demencia por EA. Esta información permitiría realizar un estudio más específico sobre las alteraciones de la estructura temporal del EEG asociadas a la progresión a demencia por EA. Finalmente, las bandas de frecuencia consideradas en el estudio no se han extendido a bandas superiores a beta-1 debido al elevado coste computacional que supone la obtención de los meta-estados. Resultaría interesante calcular las medidas en estas bandas para observar las alteraciones que la EA provoca en su estructura temporal.

## 5. Conclusiones

Los resultados de este estudio indican que la progresión de la EA provoca una paulatina desestructuración de la arquitectura temporal subyacente de la actividad neuronal en reposo, la cual no es enteramente aleatoria. Además, el avance de la EA conlleva una reducción significativa en la definición de los meta-estados cerebrales que podría ser debida a una cognición más inestable. En definitiva, el presente estudio constituye un paso más para obtener una caracterización completa de las alteraciones dinámicas de la red neuronal funcional en la EA.

## Agradecimientos

Este estudio ha sido financiado por el ‘Ministerio de Ciencia e Innovación – Agencia Estatal de Investigación’ y el Fondo Europeo de Desarrollo Regional (FEDER) bajo el proyecto PGC2018-098214-A-I00, por la Comisión Europea y el FEDER bajo el proyecto ‘Análisis y correlación entre la epigenética y la actividad cerebral para evaluar el riesgo de migraña crónica y episódica en mujeres’ (‘Programa de Cooperación Interreg V-A España-Portugal POCTEP 2014-2020’), y por el CIBER en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN) a través del Instituto de Salud Carlos III cofinanciado con fondos FEDER. V. Rodríguez-González es beneficiario de una ayuda PIF-UVa de la Universidad de Valladolid.

## Referencias

- [1] Cohen MX. Where does EEG come from and what does it mean?. *Trends in Neuroscience*, vol. 40, no. 4, 2017, pp. 208-18.
- [2] Alzheimer's Association. 2022 Alzheimer's disease facts and figures. *Alzheimer's & Dementia*, vol. 18, no. 4, 2022, pp. 700-89.
- [3] Rossini PM, et al. Early diagnosis of Alzheimer's disease: the role of biomarkers including advanced EEG signal analysis. Report from the IFCN-sponsored panel of experts. *Clinical Neurophysiology*, vol. 131, no. 6, 2020, pp. 1287-310.
- [4] Deco G, Jirsa VK, McIntosh AR. Resting brains never rest: Computational insights into potential cognitive architectures. *Trends in Neuroscience*, vol. 36, no. 5, 2013, pp. 268-74.
- [5] Waschke L, Kloosterman NA, Obleser J, Garrett DD. Behavior needs neural variability. *Neuron*, vol. 109, no. 5, 2021, pp. 751-66.
- [6] Calhoun VD, Miller R, Pearson G, Adali T. The chronnectome: time-varying connectivity networks as the next frontier in fMRI data discovery. *Neuron*, vol. 84, no. 2, 2014, pp. 262-74.
- [7] Núñez P, Poza J, Gómez C, Rodríguez-González V, Hillebrand A, Tewarie P, Tola-Arribas MA, Cano M, Hornero R. Abnormal meta-state activation of dynamic brain networks across the Alzheimer spectrum. *Neuroimage*, vol. 232, 2021, p. 117898.
- [8] Desikan RS, Ségonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, Buckner RL, Dale AM, Maguire RP, Hyman BT, Albert MS, Killiany RJ. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage*, vol. 31, no. 3, 2006, pp. 968-80.
- [9] Tewarie P, Liuzzi L, O'Neill GC, Quinn AJ, Griffa A, Woolrich MW, Stam CJ, Hillebrand A, Brookes MJ. Tracking dynamic brain networks using high temporal resolution MEG measures of functional connectivity. *Neuroimage*, vol. 200, 2019, pp. 38-50.
- [10] Miraglia F, Vecchio F, Rossini PM. Searching for signs of aging and dementia in EEG through network analysis. *Behavioural Brain Research*, vol. 317, 2017, pp. 292-300.

# Technological platforms using bioengineering techniques for studies of the alternative vasculature in neuroblastoma

A. Villasante<sup>1,2</sup>, M.J. Lopez Martinez<sup>1,2,3</sup>, C. Alcon<sup>1</sup>, A. Garcia-Lizarribar<sup>1,3</sup>, J. Samitier<sup>1,2,3</sup>

<sup>1</sup>Institute for Bioengineering of Catalonia, Barcelona, Spain.

<sup>2</sup>Department of Electronic and Biomedical Engineering, University of Barcelona, Barcelona, Spain.

<sup>3</sup>Biomedical Research Networking Center in Bioengineering, Biomaterials, and Nanomedicine, Madrid, Spain.

## Abstract

*Neuroblastoma (NB) is a highly vascularized pediatric tumor originating from undifferentiated neural crest cells in the first years of life. NB displays an alternative vasculature constituted by cancer cells that can transdifferentiate into tumor-derived endothelial cells (TECs), a phenotype associated with drug resistance and tumor relapse. We currently lack treatments against TECs, mainly because of the challenge of developing predictive models that can mimic this phenomenon in vitro. Herein, we deconstructed the NB microenvironment to establish the minimum elements necessary to induce endothelial transdifferentiation and developed two technological platforms using bioengineering techniques. The first strategy to generate TECs in vitro consisted in mimicking the reported stiffness for physiological angiogenesis and induction of angiogenesis in vitro. We developed a biocompatible platform of different stiffnesses. We demonstrated that the most aggressive NB cell phenotypes were mainly transdifferentiated in stiff substrates in the arteries range, while the less aggressive cell lines lacked transdifferentiation capacity. The second approach consisted in mimicking the NB physiological stiffness by fabricating a collagen I/fibrin-based hydrogel designed to support the survival of NB cells. NB cells cultured under multiple combinations of microenvironment factors, but static, did not result in transdifferentiation. However, the inclusion of shear stress into a microfluidic chip device induced conversion into TECs. We concluded that stiffness and shear stress are crucial microenvironment properties to induce the alternative vasculature in NB.*

## 1. Introduction

Neuroblastoma (NB) is a malignant tumor of the neural crest cells that give rise to the sympathetic nervous system. NB is a rare cancer and the most common extracranial solid tumor of childhood diagnosed in the first year of life. It frequently arises in the adrenal medulla, although the tumor's location can be anywhere in the sympathetic nervous system (abdominal origin in 80% of cases)[1]. NB cells *in vitro* are presented as three distinct cellular phenotypic alternatives with different tumorigenic potentials- N, S, and I-type. [2]. N-type (neuroblastic phenotype) cells are highly malignant, while S-type (substrate-adherent) cells are nonmalignant. Finally, the I-type exhibits an intermediate morphology to those of N and S. I-type cell is suggested to represent a cellular intermediate in the N/S interconversion, or a progenitor of N- or S-type cells, capable of bidirectional differentiation [2, 3]. I-type cells have been shown to

represent the most tumorigenic neuroblastoma phenotype. Tumor vascularization is highly implicated in the regulation of growth, survival, and metastasis in multiple solid tumors. Tumor vessels can be formed through physiological-based mechanisms such as angiogenesis (sprouting and intussusceptive) and vasculogenesis [4]. In addition, tumors can apply alternative strategies known as vasculogenic mimicry and endothelial transdifferentiation, which do not involve the host vascular endothelium. During Endothelial transdifferentiation cancer cells acquire endothelial properties, such as expression of CD31, and become tumor-derived endothelial cells (TECs) [4]. NB is a highly angiogenic tumor [5]. Recently was discovered that the alternative vascularization mechanism occurs in NB. NB cells can transdifferentiate into TECs carrying CD31 and MYCN amplification markers[6]. TECs are likely involved in chemoresistance in numerous tumors, including NB[6, 7]. Current antiangiogenic strategies are designed to target the “classical mechanisms” but not the alternative vasculature, which in turn produce clinical failures and adverse outcomes[4].

However, there is a substantial technical limitation regarding predictive models. TECs studies have been hampered by a lack of affordable, non-time consuming experimental models. Nevertheless, tissue-engineered models have successfully mimicked healthy human vasculature in the last decade. Different studies have demonstrated the role of different microenvironment factors such as shear stress [8], substrate stiffness [9, 10], ECM components, dimensionality, and cell culture medium in favoring vasculature formation *in vitro*[11-13].

We aimed to find the minimum number of microenvironment factors (minimally functional unit) necessary to recapitulate neuroblastoma-derived endothelial cells *in vitro*, which consists in developing the most straightforward platform enabling NB endothelial transdifferentiation.

## 2. Materials and Methods

### 2.1. Hydrogel preparation and characterization

FX3 and FX5 formulations are both collagen I/fibrin-based hydrogels. They were fabricated using bovine collagen I (2µg/µL; ThermoFisher Scientific,

#A1064401), fibrinogen from bovine blood plasma (10 $\mu$ g/ $\mu$ L; Sigma #F8630), were cross-linked with Thrombin (Sigma # T4648).

## 2.2. Mechanical properties

Uniaxial compression tests of materials were performed using a Zwick Z0.5 TN instrument (Zwick-Roell, Germany) with a 5 N load cell. Samples were tested at room temperature up to 30% final strain (deformation), using the following parameters: 0.1 mN preload force and 20% min<sup>-1</sup> strain rate. Stress-strain graphs were obtained from load-deformation measurements. Values for the compressive modulus were calculated from the slope of the linear region corresponding to 10–20% strain.

## 2.3. Microfluidic device and perfusion setup

idenTx 3 Chips (AIM Biotech) were assembled in an idenTx 9 Plate. LUC-1 Luer (AIM Biotech) and male Luer integral 1/16 connectors (Cole-Parmer #45518-07) were used to connect Tygon tubes (Tygon AAD04133; inner diameter = 1.27mm, outer diameter = 2.28 mm) to chip inlets and outlets. Tygon tubes were connected to the peristaltic pump tubing (Tygon LMT-55 #070534-04-ND) through a small piece of Tygon tube of 2mm of inner diameter (Tygon # ACSF1S1502-C). The recirculating circuit was achieved by attaching the other extreme of Tygon tubes to a reservoir fabricated with a 50 mL Falcon. The falcon tube was also used as a bubble trap system. The culture medium was pumped using a high precision IPC-N 8 peristaltic pump (Ismatec). The idenTx 9 plate containing the devices was placed inside a humidified incubator at 37°C and 5% CO<sub>2</sub>.

## 2.4. PDMS-based platform

Polydimethylsiloxane (PDMS) Sylgard 184 (Dow Corning) was mixed in a weight ratio of 1:75 and 1:70 (cross-linker: pre-polymer) to get equilibrium modulus values of  $E_{\text{mod}} = 101.38 \pm 6.22$  kPa, and  $E_{\text{mod}} = 140.70 \pm 5.24$  kPa, respectively. PDMS Sylgard 527 was mixed with proportions of 1:1 and 1:2 (A component/B component), corresponding to Equilibrium modulus values of  $E_{\text{mod}} = 3.39 \pm 0.81$  kPa and  $E_{\text{mod}} = 7.09 \pm 1.16$  kPa, respectively.

## 2.5. CD31 staining for flow cytometry analysis

Cells cultured in a six well-plate coated with PDMS of different stiffnesses were trypsinized, collected, pooled together and stained with the primary antibody specific to CD31 (dilution 1:100, DAKO, #M0823) in Hank's Balanced Salt Solution (HBSS) with 2% Fetal Bovine Serum (FBS) for 30 minutes on ice. Cells were washed three times with PBS, centrifuged at 500 g for 5 minutes, and incubated with the secondary antibody goat anti-mouse Alexa Fluor 647 (dilution 1:500, ThermoFisher #A-21235) in HBSS + 2% FBS on ice for 30 minutes. Cells were then washed three times with PBS, finally resuspended in 250 $\mu$ L of PBS, and analyzed on a flow cytometry Gallios instrument (Beckman Coulter, Nyon, Switzerland). Cells incubated without primary antibody were used as negative controls. Data analysis was performed using Flowing Software.

## 2.6. Immunofluorescence (IF)

Cells were fixed with formalin 10% neutral buffered (Sigma #HT5011), permeabilized with PBS containing 0.2% Triton X-100 and blocked during 1 hour with 5% BSA in PBST (PBS+ 0.1% Tween 20). Then, samples were incubated with primary antibodies for MYCN (Abcam, ab198912) and CD31 (DAKO, #M0823) in a humid chamber overnight at 4°C. A day later, the samples were washed with PBST and incubated with the appropriate secondary antibodies (ThermoFisher; Goat anti-rabbit-Alexa Fluor 488 # A-11008; Goat anti-mouse-Alexa Fluor 647 #A-21235) at room temperature, in a humid dark chamber for one hour. Samples were incubated with Hoechst at room temperature and imaged by fluorescence microscopy (Olympus IX81 light microscope).

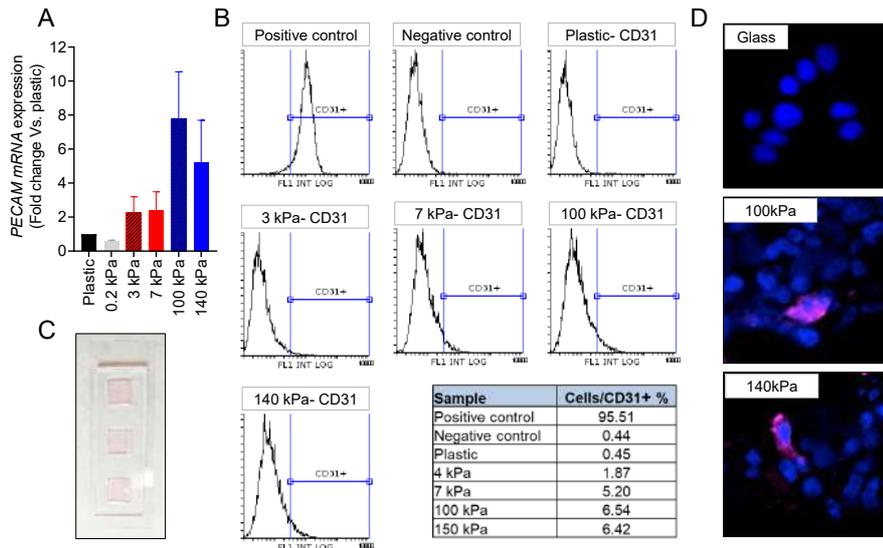
## 2.7. Quantitative Real-Time PCR (qRT-PCR)

Total RNA from cells was obtained using Trizol (Life Technologies), following the manufacturer's instructions. RNA preparations were treated with "Ready-to-go you-prime first-strand beads" (GE Healthcare) to obtain cDNA. Quantitative real-time PCR was performed using PowerUp SYBR Green Master Mix (Applied Biosystems #A25742). mRNA expression levels were quantified applying the  $\Delta$ Ct method,  $\Delta$ Ct = (Ct of the gene of interest - Ct of Actin). PECAM primers were obtained from the PrimerBank database.

## 3. Results

In mammals, the range of venous tissue stiffness is about 3–50 kPa, while arteries are stiff, ranging from about 50–150 kPa [14]. Previous studies using PDMS substrates of varying degrees of stiffness demonstrated the influence of ECM rigidity on arterial-venous differentiation of endothelial progenitor cells (EPCs). Following the same rationale to differentiate healthy EPCs into vascular vessels, we cultured NB cells on venous-like and arterial-like PDMS substrates to induce transdifferentiation into TECs. N-type cells do not express PECAM/CD31 transdifferentiation marker in compliant substrates (**Figure 1. A, B**), nor in highly soft materials such as  $E_{\text{mod}} = 0.2$  kPa (**Figure 1. A**). However, NB cells show significant PECAM mRNA expression levels compared to the plastic and soft controls when cultured for 7 days on arterial-like PDMS substrates (**Figure 1. A**). The presence of a subpopulation of CD31<sup>+</sup> originated in the 100kPa, and 140kPa plates was also observed by flow cytometry (**Figure 1. B**). For further analysis of the transdifferentiated cells, we decided to perform immunofluorescence for CD31. To this end, we fabricated a device with chambers placed on top of a glass slide and coated with 1:70 and 1:75 Sylgard 184 (**Figure 1. C**). As expected, SK-N-BE (2) cells did not express CD31 when cultured on the glass control, but we confirmed the presence of TECs in the 100kPa and 140kPa chamber slides (**Figure 1.D**).

To delve into the capability of NB cells to transdifferentiate on the PDMS platform, we investigate the possibility of obtaining TECs from the other two NB

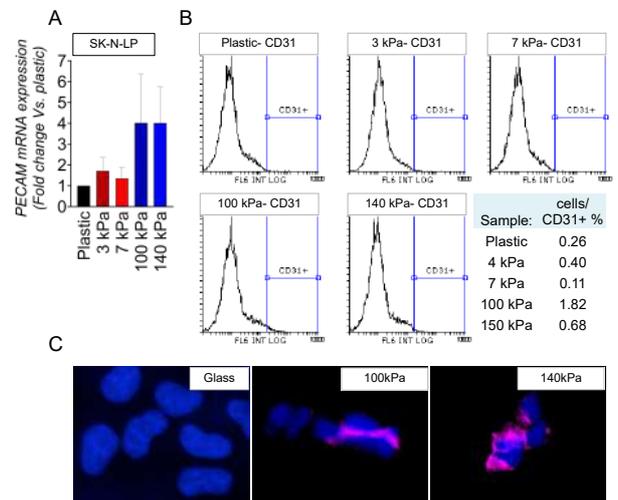


**Figure 1. Endothelial cell transdifferentiation of N-type neuroblastoma on venous-like and artery-like PDMS materials.** (A) mRNA levels of PECAM in SKNBE (2) cells cultured on the indicated substrates stiffness for 7 days. (B) FACS analysis of CD31 endothelial marker in SKNBE (2) cells cultured on the indicated PDMS substrates for 7 days (n=4). HUVEC were used as positive control. Cells incubated without primary antibody were used as negative control. (C) Cell culture chamber slide fabricated with PDMS Sylgard 184. (D) Representative images of CD31 expression (magenta) in SKNBE (2) cells cultured on the indicated substrates; nuclei are stained with Hoechst 33342 (blue) (n=4).

phenotypes. qRT-PCR of PECAM expression and FACS analyses of cell subpopulations expressing CD31 showed that the 100kPa and 140kPa substrates significantly induced endothelial transdifferentiation of the I-type cell phenotype (Figure 2A, B). However, the 100kPa appeared as the most effective to induce the transition. As expected, immunofluorescence images revealed the expression of the CD31 protein in the cell membrane of subpopulations of cells cultured on the arterial-like PDMS material (Figure 2C). Finally, we examined the S-type NB cell line LA1-5s. Unlike the N-type and I-type cells analyzed, the LA1-5s cell line did not express PECAM/CD31 in any stiffness conditions (data not shown).

Then, we proposed recapitulating the native NB stiffness using a collagen I/fibrin-based hydrogel as a second strategy to obtain TECs *in vitro*. Neuroblastoma is a soft tumor. The Young's modulus of NB is of  $E_{mod} = 0.79 \pm 0.20$  kPa [15]. By modifying the concentration of the thrombin enzyme, we obtained two stable formulations in the range of native stiffness named as FX5 ( $E_{mod} = 0.87 \pm 0.28$  kPa) and FX3 ( $E_{mod} = 0.56 \pm 0.30$  kPa). Neither culturing the cells on top of the FX5 biomaterial under static conditions in monoculture or co-culture with HUVEC nor in 3D within the biomaterial supported TEC conversion (data not shown).

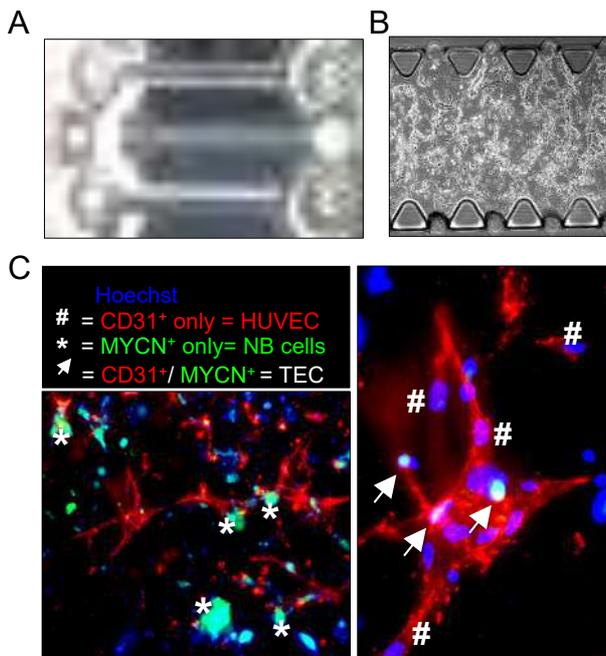
Shear stress is a pivotal microenvironment factor leading to vasculature formation. Thus, we decided to include shear stress in the model system to study the possible effect by using a microfluidic chip device. A co-culture of NB cells and HUVEC were seeded within the hydrogel and cultured in RPMI/EGM mix medium for 7 days under flow perfusion conditions (Figure 3A). A tissue-engineered was observed inside the central chamber at day 7 (Figure 3B). Co-stains of MYCN and CD31 of the tissues demonstrated the presence of cells expressing both markers (Figure 3C), which confirms that NB transdifferentiation using a matrix of physiological stiffness requires shear stress to be effective.



**Figure 2. Endothelial transdifferentiation of I-type neuroblastoma on venous-like and artery-like PDMS materials.** (A) mRNA levels of PECAM in SK-N-LP cells cultured on the indicated substrates stiffness for 7 days. Fold change was calculated by first normalizing to GAPDH levels in the individual samples and then to the corresponding levels in cells cultured in plastic culture dishes. Data are shown as average  $\pm$ SD (n=3). (B) FACS analysis of CD31 endothelial marker in SK-N-LP cells cultured on the indicated PDMS substrates for 7 days (n=4). (C) Representative images of CD31 expression (magenta) in SK-N-LP cells cultured on the indicated substrates; nuclei are stained with Hoechst 33342 (blue) (n=4). Cells were cultured in the chamber slides for 7 days.

#### 4. Conclusions

We had developed a battery of technological platforms to recapitulate *in vitro* the presence of the alternative transdifferentiated neuroblastoma vasculature. These model systems that recapitulate tumor heterogeneity will help defer a specific treatment against the TECs subpopulation to avoid resistance and tumor relapse in pediatric patients. We found that NB cells with aggressive phenotypes can evolve into tumor-derived endothelial cells by mimicking the artery stiffness.



**Figure 3. Analysis of endothelial cell transdifferentiation in a neuroblastoma-on-a-chip.** (A) Image of a butterfly-shaped microfluidic chip with a central chamber and two lateral channels for perfusion. The central chamber is filled with the FX5 hydrogel that crosslinks inside the chamber. (B) Bright-field image of the chip seeded with NB cells and HUVEC within the hydrogel in the central chamber. (C) Representative fluorescence images of neuroblastoma and HUVEC co-cultures double-stained for MYCN (neuroblastoma marker; green) and CD31 (endothelial marker; red) ( $n=3$ ); nuclei are stained with Hoechst 33342 (blue). TECs express both markers MYCN and CD31 (white arrows).

### Acknowledgments

AV is supported by the Spanish Association Against Cancer (AECC), the Association of Families and Friends of Children with Neuroblastoma (NEN), and the Spanish National R+D+ i plan PID2020-117977RA-I00. The IBEC Group is supported by Biomedical Research Networking (CIBER), Spain (RD16/0006/0012; RD16/0011/0022), the Generalitat de Catalunya (2017-SGR-1079).

We are thankful to the staff from Unitat de Microscòpia Electrònica TEM/SEM (Centres Científics i Tecnològics (CCiTUB), Universitat de Barcelona) for the electron microscopy studies. We thank the MicroFabSpace and Microscopy Characterization Facility, Unit 7 of ICTS “NANBIOSIS” from CIBER-BBN at IBEC

### References

[1] Matthay KK, Maris JM, Schleiermacher G, Nakagawara A, Mackall CL, Diller L, et al. Neuroblastoma. *Nat Rev Dis Primers*. 2016;2:16078.  
 [2] Campos Cogo S, Gradowski Farias da Costa do Nascimento T, de Almeida Brehm Pinhatti F, de França Junior N, Santos Rodrigues B, Cavalli LR, et al. An overview of neuroblastoma cell lineage phenotypes and in vitro models. *Exp Biol Med* (Maywood). 2020;245:1637-47.  
 [3] Walton JD, Kattan DR, Thomas SK, Spengler BA, Guo H-F, Biedler JL, et al. Characteristics of Stem Cells from Human

Neuroblastoma Cell Lines and in Tumors. *Neoplasia*. 2004;6:838-45.  
 [4] Krishna Priya S, Nagare RP, Sneha VS, Sidhanth C, Bindhya S, Manasa P, et al. Tumour angiogenesis-Origin of blood vessels. *Int J Cancer*. 2016;139:729-35.  
 [5] Chlenski A, Liu S, Cohn SL. The regulation of angiogenesis in neuroblastoma. *Cancer Lett*. 2003;197:47-52.  
 [6] Pezzolo A, Parodi F, Corrias MV, Cinti R, Gambini C, Pistoia V. Tumor origin of endothelial cells in human neuroblastoma. *J Clin Oncol*. 2007;25:376-83.  
 [7] Villasante A, Sakaguchi K, Kim J, Cheung NK, Nakayama M, Parsa H, et al. Vascularized Tissue-Engineered Model for Studying Drug Resistance in Neuroblastoma. *Theranostics*. 2017;7:4099-117.  
 [8] Ahsan T, Nerem RM. Fluid Shear Stress Promotes an Endothelial-Like Phenotype During the Early Differentiation of Embryonic Stem Cells. *Tissue Engineering Part A*. 2010;16:3547-53.  
 [9] Smith Q, Chan XY, Carmo AM, Trempel M, Saunders M, Gerecht S. Compliant substratum guides endothelial commitment from human pluripotent stem cells. *Science Advances*. 2017;3.  
 [10] Wong L, Kumar A, Gabela-Zuniga B, Chua J, Singh G, Happe CL, et al. Substrate stiffness directs diverging vascular fates. *Acta Biomater*. 2019;96:321-9.  
 [11] Rosa S, Praça C, Pitrez PR, Gouveia PJ, Aranguren XL, Ricotti L, et al. Functional characterization of iPSC-derived arterial- and venous-like endothelial cells. *Sci Rep*. 2019;9.  
 [12] Klein D. iPSCs-based generation of vascular cells: reprogramming approaches and applications. *Cellular and Molecular Life Sciences*. 2017;75:1411-33.  
 [13] Song HHG, Rumma RT, Ozaki CK, Edelman ER, Chen CS. Vascular Tissue Engineering: Progress, Challenges, and Clinical Promise. *Cell Stem Cell*. 2018;22:340-54.  
 [14] Gordon E, Schimmel L, Frye M. The Importance of Mechanical Forces for in vitro Endothelial Cell Biology. *Frontiers in Physiology*. 2020;11.  
 [15] Monferrer E, Martín-Vañó S, Carretero A, García-Lizarribar A, Burgos-Panadero R, Navarro S, et al. A three-dimensional bioprinted model to evaluate the effect of stiffness on neuroblastoma cell cluster dynamics and behavior. *Sci Rep*. 2020;10.

# Diseño de modelos automáticos basados en inteligencia artificial para predicción del Síndrome Metabólico

A. Burgaleta Ledesma<sup>1</sup>, G. Fico<sup>2</sup>, B. Merino<sup>3</sup>, M. F. Cabrera Umpiérrez<sup>4</sup>, M. T. Arredondo Waldmeyer<sup>5</sup>

<sup>1</sup> Life Supporting Technologies (LST), Departamento de Tecnología Fotónica y Bioingeniería, Universidad Politécnica de Madrid (UPM), Madrid, España. [aburgaleta@lst.tfo.upm.es](mailto:aburgaleta@lst.tfo.upm.es), [gfico@lst.tfo.upm.es](mailto:gfico@lst.tfo.upm.es), [bmerino@lst.tfo.upm.es](mailto:bmerino@lst.tfo.upm.es), [Chiqui@lst.tfo.upm.es](mailto:Chiqui@lst.tfo.upm.es), [mta@lst.tfo.upm.es](mailto:mta@lst.tfo.upm.es).

## Resumen

*El Síndrome Metabólico (SM) está conformado por un conjunto de trastornos metabólicos que suponen un factor de riesgo para sufrir diabetes y enfermedades cardiovasculares y cerebrovasculares, patologías que son de las principales causas de muerte a nivel mundial. El objetivo de este proyecto es predecir la presencia de Síndrome Metabólico en un individuo mediante el desarrollo de técnicas de aprendizaje automático capaces de detectar factores de riesgo asociados con la enfermedad.*

*La metodología llevada a cabo ha consistido en una exploración de las bases de datos existentes para seleccionar la más adecuada para este proyecto (NHANES), posteriormente se realizó una limpieza y preprocesamiento de los datos para dar lugar a un análisis estadístico de la muestra. A continuación, se obtuvieron las matrices de correlación de las variables con la presencia del síndrome metabólico entre los individuos para poder descartar aquellas cuya correlación pudiera considerarse despreciable. Finalmente, con la muestra preparada, se entrenan, se optimizan y se validan los modelos de aprendizaje automático.*

*De esta forma, las aportaciones principales de este proyecto son los modelos de aprendizaje automático capaces de predecir el síndrome metabólico de un individuo con unas métricas de precisión, exhaustividad y exactitud superiores al 80% (en base a constantes sencillas de obtener de los pacientes, muchas de ellas a través de un cuestionario de hábitos de vida), así como el conjunto de datos procesado que queda a disposición de futuras investigaciones.*

*Como conclusión principal se obtiene que los factores que más influyen en el Síndrome Metabólico son la edad, padecer un estado prediabético o diabético, los niveles elevados de insulina en ayunas, el exceso de porcentaje de grasa y ser fumador.*

## 1. Introducción

El SM se ha convertido en un problema de salud pública generalizado, especialmente en los países desarrollados. Esta enfermedad se asocia directamente con un incremento de hasta cinco veces en la prevalencia de diabetes tipo dos y de hasta tres veces en enfermedad cardiovascular [1].

El Síndrome Metabólico (SM) es un conjunto de alteraciones metabólicas constituido por la obesidad abdominal, la disminución de las concentraciones de

colesterol unido a las lipoproteínas de alta densidad, la elevación de las concentraciones de triglicéridos, el aumento de la presión arterial y la hiperglucemia [2].

La suma de hipertensión, obesidad, dislipemia y diabetes es un concepto que lleva estando presente mucho tiempo en los entornos clínicos. Se trata de una asociación de problema de salud que pueden aparecer simultáneamente o de manera secuencial en un mismo individuo. Están causados por una combinación de factores genéticos y un estilo de vida poco saludable [3].

Aunque la primera definición del Síndrome Metabólico aparece en 1988 de la mano de Gerald Reaven [3], hoy en día no hay un consenso entre las diferentes organizaciones sanitarias. En el presente trabajo los criterios se eligen conforme a la definición publicada en la declaración científica conjunta sobre el Síndrome Metabólico [4]. Para diagnosticar SM en un sujeto se deben dar como mínimo tres de los cinco requisitos expuestos a continuación:

- Perímetro abdominal mayor o igual a 88cm para mujeres y de 102cm para hombres.
- Triglicéridos con una concentración mayor o igual a 150mg/dL o encontrarse en tratamiento farmacológico para los triglicéridos elevados.
- Colesterol HDL de menos de 40mg/dL para hombres y de 50mg/dL para mujeres o encontrarse en tratamiento farmacológico para el HDL bajo.
- Presión arterial sistólica mayor o igual a 130mmHg, diastólica de más de 85mmHg o ambas, o encontrarse en tratamiento farmacológico antihipertensivo.
- Glucosa en ayunas superior a 100mg/dL o encontrarse en tratamiento farmacológico por niveles elevados de glucosa en sangre.

## 2. Objetivos

Una vez identificada la problemática, la finalidad principal es implementar una serie de modelos predictivos basados en aprendizaje automático que permitan detectar la presencia de SM en un individuo. Para ello, previamente se realiza un análisis estadístico que permita obtener los factores de riesgo principales del SM.

Los objetivos específicos se presentan a continuación:

- Obtención de un conjunto de datos que sea de utilidad para futuras investigaciones sobre el SM, con un total de 26.868 filas y 33 columnas.
- Obtención de un análisis descriptivo del conjunto de datos estratificando a los participantes.
- Obtención de la prevalencia de SM en la muestra recogida en la base de datos.
- Obtención de la correlación de los diferentes atributos con la presencia de SM.
- Obtención de los hiperparámetros ajustados que permitan optimizar el rendimiento de los diferentes modelos de aprendizaje automático desarrollados,
- Validación técnica y de experiencia de usuario. Observación y obtención de la evaluación de profesionales de la salud sobre el potencial interés que puede suponer esta herramienta aplicada a su entorno laboral.

### 3. Materiales y Métodos

#### 3.1. Base de datos NHANES

La base de datos National Health and Nutrition Examination Survey (NHANES) [5] lleva operativa desde el 1959 recopilando, analizando y difundiendo información sobre el estado de salud de los residentes de los Estados Unidos. Cada año cerca de 5.000 individuos participan en sus encuestas (anualmente el conjunto muestral experimenta un aumento, hasta acercarse a los 10.000 participantes en los ciclos más actuales) que incluyen preguntas demográficas, socioeconómicas, dietéticas, de actividad física, hábitos y otros parámetros relacionados con su salud. Además, cuenta con un reconocimiento médico que consiste en mediciones físicas, dentales y fisiológicas, así como pruebas de laboratorio administradas por personal médico altamente capacitado. Los resultados de esta encuesta han sido ampliamente utilizados para determinar la prevalencia de las principales enfermedades y los factores de riesgo de las mismas. La muestra de la encuesta se selecciona para representar a la población estadounidense de todas las edades y etnias, con la finalidad de generar estadísticas lo más fiables posibles.

Se escoge esta base de datos debido a distintos motivos:

- Por la variedad y la cantidad de datos que aporta. Seleccionando los ciclos que comprenden del 2011 al 2018 se extraen un total de 37.606 filas y 42 columnas que tras la limpieza y preprocesado resultan en 26.868 filas y 33 columnas. El grupo muestral es lo suficientemente grande como para realizar un análisis estadístico en condiciones y para poder entrenar los modelos de aprendizaje automático desarrollados.
- El procedimiento que utiliza para la recopilación de datos es un seguimiento anual y exhaustivo de los participantes.
- El hecho de llevar desde 1959 elaborando información sobre el estado de salud de la población estadounidense hace que esté lo suficientemente asentada y que cuente con la experiencia necesaria como para confiar en la precisión de sus datos.

- Esta base de datos ya se ha utilizado previamente para elaborar estudios relacionados con diferentes patologías y problemas de salud. Por el tipo de información que ofrece NHANES, ha contribuido especialmente en el estudio de las siguientes cuestiones: Anemia [6], enfermedades cardiovasculares, diabetes, exposiciones ambientales, enfermedades oculares, pérdidas de audición, enfermedades infecciosas, enfermedades renales, nutrición, obesidad [7], salud bucodental, osteoporosis [8], condición física, historial reproductivo, enfermedades respiratorias y enfermedades de transmisión sexual. Así como del propio síndrome metabólico [9], [10], [11].

Se decide recoger registros de los ciclos comprendidos entre 2011 y 2018. Tras la limpieza y el preprocesado de los mismos se obtienen 26.868 registros de pacientes. De cada registro o paciente se extraen 33 características, seleccionadas en base al análisis estadístico realizado entre las mismas y la presencia de SM, escogiendo aquellas que más correlación presentaran. Estos atributos son de tipo demográfico (etnia, edad, género y nivel educativo), de cuestionario (depresión, calidad de sueño, nivel de actividad física, calidad de la dieta...), antropométrico (peso, altura, perímetro abdominal y presión sanguínea) y de laboratorio (triglicéridos y colesterol en sangre y glucosa e insulina en ayunas).

#### 3.2. Análisis estadístico de la muestra

En este apartado se realiza un análisis estadístico de los datos recopilados con el objetivo de tener una visión de las características generales del grupo muestral de la encuesta.

Para ello se estudia la distribución de los participantes por datos demográficos. Es decir, la división por género, edad y etnia. Se obtiene una muestra bastante homogénea en lo que a género se refiere. De todos los registros el 50.7% representa a mujeres y el 49.3% a hombres. En cuanto a la etnia, hay una mayor participación de personas blancas no hispanas (33%), seguido de no hispanos negros (24%), americanos mexicanos (16%), asiáticos (12%), otros hispanos (11%) y finalizando con la representación de otra raza (5%). Los participantes menores a 12 años se descartan por resultar irrelevantes para el entrenamiento de los modelos de aprendizaje automático, puesto que no se recogen muestras de laboratorio de este grupo de participantes en la base de datos NHANES. Se obtiene una representación equitativa en cuanto a la edad de los encuestados. 22% en los participantes entre 12 y 20 años, 21% en los participantes entre 21 y 25 años, 20% en los participantes entre 26 y 50 años, 19% en los participantes entre 51 y 65 años y finalmente, del 18% en los participantes entre 66 y 80.

Por otro lado, se estudia el porcentaje de sobrepeso y obesidad. Para ello se tiene en cuenta el índice cintura altura [12]. La incidencia de sobrepeso entre los participantes se muestra en la siguiente tabla:

	Sobrepeso	Sobrepeso elevado	Obesidad
Mujeres	12%	13%	54%
Hombres	21%	23%	25%

**Tabla 1.** Incidencia de sobrepeso en la muestra

Finalmente, se realiza un estudio de la prevalencia del Síndrome Metabólico en base a presentar al menos tres de los cinco criterios expuestos en la sección 1 de este documento (perímetro abdominal mayor o igual a 88cm para mujeres y de 102cm para hombres, triglicéridos con una concentración mayor o igual a 150mg/dL, glucosa en ayunas superior a 100mg/dL...). De todos los participantes, el **41% presenta Síndrome Metabólico**. El nivel educativo, el consumo de tabaco, la edad avanzada, el índice cintura altura y el género masculino representan algunos de los factores de riesgos para padecerlo.

### 3.3. Desarrollo de los modelos de aprendizaje automático

Tras el análisis, se procede al desarrollo de los modelos de aprendizaje automático y a su ajuste de hiperparámetros con el objetivo de optimizar su rendimiento. Los modelos implementados son: *Support Vector Machine*, *Decision Tree*, *Random Forest*, *XGBoost* y *Neural Network*.

Tras la optimización de los modelos mediante ajuste de hiperparámetros por validación cruzada. se entrenan con el 80% de la muestra y se utiliza el 20% restante para el evaluar su rendimiento. Para guiar la validación se empleó la métrica F1 además de utilizarse técnicas para evitar el sobre entrenamiento.

## 4. Discusión de los resultados

A pesar de que todos los modelos presentan un rendimiento parecido, siendo el XG Boost el que mejores resultados ofrece. Con una precisión del 83%, una exhaustividad del 85%, una puntuación F1 del 84% y una exactitud del 87%.

	DT	RF	SVC	XGB	NN
<b>Exactitud</b>	0.851	0.863	0.858	<b>0.869</b>	0.865
<b>Precisión</b>	0.795	0.824	0.826	<b>0.828</b>	0.887
<b>Exhaustividad</b>	0.850	0.839	0.820	<b>0.852</b>	0.939
<b>F1</b>	0.821	0.831	0.823	<b>0.840</b>	0.864
<b>AUC ROC</b>	0.929	0.942	0.852	<b>0.94</b>	0.939

**Tabla 2:** Resultados de los modelos

Utilizar técnicas de inteligencia artificial avanzada en un entorno médico para detectar de forma precoz el Síndrome Metabólico podría suponer una herramienta muy útil, especialmente cuando la mayor parte de datos que se deben recopilar del paciente consisten en un formulario de hábitos de vida y unas cuantas mediciones físicas, es decir, que son muy sencillos de obtener. el modelo XGBoost, que ha resultado ser el que mejores resultados obtiene, puede

predecir el SM con una exactitud del 87% y una precisión del 83%. Esto indica que se pueden desarrollar modelos de inteligencia artificial para diagnosticar el Síndrome Metabólico con altos porcentajes de éxito.

Al no existir un criterio unificado para diagnosticar Síndrome Metabólico por las diferentes organizaciones que estudian esta enfermedad [1], se vuelve necesario establecer nuevos mecanismos no convencionales para su diagnóstico. Una buena alternativa es la utilización de modelos de aprendizaje automático. En la literatura no se identifica ningún estudio que haga una comparativa de diferentes modelos para la predicción del SM debido a que la mayoría se limitan a examinar su prevalencia en distintas poblaciones según criterios demográficos.

En cuanto a los atributos que más parecen guardar relación con la presencia de SM en los individuos se destacan el de presión sanguínea elevada, la edad avanzada, el nivel de insulina en sangre en ayunas alto, presencia de diabetes y un porcentaje graso elevado, esta selección de atributos se pueden obtener tanto de la selección interna que realizan el XGBoost, el Decision Tree y el Random Forest, así como mediante el análisis de correlación de la muestra. Cabe destacar que se ha visto que el índice cintura-altura es un marcador más fiable que el índice de masa corporal para evaluar el estado de salud de un sujeto.

A pesar de que los resultados de los modelos de predicción han sido bastante satisfactorios, es conveniente destacar algunas limitaciones.

En primer lugar, es necesario señalar que, en el ámbito de la salud, correlación no implica causalidad. Las patologías médicas suelen tener una causa multifactorial y aislar variables resulta complicado, especialmente en el marco de los hábitos de vida de una persona. Los factores genéticos, ambientales y sociales pueden actuar sinérgicamente para causar enfermedad en un sujeto.

Por otro lado, se debe considerar la recopilación de información de la base de datos. La mayor parte de atributos utilizados para el entrenamiento de los algoritmos son frutos de un cuestionario. Estos datos pueden no ser del todo fiables por el grado de subjetividad que suponen.

## 5. Conclusiones

A pesar de las limitaciones expuestas anteriormente, es posible desarrollar modelos de inteligencia artificial para diagnosticar el Síndrome Metabólico con altos porcentajes de éxito.

La base de datos NHANES se ha utilizado ampliamente en la investigación de factores de riesgo de patologías como la diabetes [13] u obesidad [14]. A pesar de esto, no se identifican estudios en los que se desarrollen y comparen modelos de aprendizaje automático orientados a la predicción del Síndrome Metabólico, cuando es una enfermedad con mayor tasa de incidencia que las mencionadas anteriormente [15].

En este desarrollo se manifiesta que los factores que más influyen en el Síndrome Metabólico son la edad, padecer un estado prediabético o diabético, los niveles elevados de insulina en ayunas, el exceso de porcentaje graso y ser

fumador. Es interesante señalar que se incluye el índice cintura altura, debido a que según diversos estudios parece ser un marcador más fiable que el IMC a la hora de evaluar el estado de sobrepeso de un individuo [16], [17]. En menor grado, la ausencia de actividad física y el índice de pobreza de los sujetos también guarda relación con la presencia del SM, en base a los criterios empleados por los modelos de aprendizaje automático para predecir la presencia de Síndrome Metabólico.

## Referencias

- [1] A. Jover *et al.*, «Prevalencia del síndrome metabólico y de sus componentes en pacientes con síndrome coronario agudo», *Rev Esp Cardiol*, vol. 64, n.º 7, pp. 579-586, jul. 2011, doi: 10.1016/J.RECESP.2011.03.010.
- [2] M. Para Optar Al Grado, «UNIVERSITAT DE LES ILLES BALEARS SÍNDROME METABÓLICO, DIETA Y MARCADORES DE INFLAMACIÓN».
- [3] V. Huggo Córdova-Pluma, G. Castro-Martínez, A. Rubio-Guerra, M. E. Hegewisch, y L. Salle, «Breve crónica de la definición del síndrome metabólico», *Med Int Méx*, vol. 30, pp. 312-328, 2014.
- [4] «Síndrome metabólico. Declaración conjunta, octubre 2009». <https://www.elsevier.es/es-revista-endocrinologia-nutricion-12-pdf-S1575092209735247> (accedido may 05, 2022).
- [5] «NHANES - National Health and Nutrition Examination Survey Homepage». <https://www.cdc.gov/nchs/nhanes/index.htm> (accedido may 05, 2022).
- [6] C. H. H. Le, «The Prevalence of Anemia and Moderate-Severe Anemia in the US Population (NHANES 2003-2012)», *PLoS One*, vol. 11, n.º 11, p. e0166635, nov. 2016, doi: 10.1371/JOURNAL.PONE.0166635.
- [7] Y. Ostchega, J. P. Hughes, A. Terry, T. H. I. Fakhouri, y I. Miller, «Abdominal obesity, body mass index, and hypertension in US adults: NHANES 2007-2010», *Am J Hypertens*, vol. 25, n.º 12, pp. 1271-1278, dic. 2012, doi: 10.1038/AJH.2012.120/2/AJH.1271.F1.JPEG.
- [8] B. Dawson-Hughes *et al.*, «The potential impact of the National Osteoporosis Foundation guidance on treatment eligibility in the USA: an update in NHANES 2005-2008», *Osteoporos Int*, vol. 23, pp. 811-820, 2012, doi: 10.1007/s00198-011-1694-y.
- [9] M. K. Palmer y P. P. Toth, «Young Adult Cancer: Influence of the Obesity Pandemic», *Obesity*, vol. 27, n.º 2, pp. 309-314, 2019, doi: 10.1002/oby.22370.
- [10] A. Mozumdar y G. Liguori, «Persistent Increase of Prevalence of Metabolic Syndrome Among U.S. Adults: NHANES III to NHANES 1999–2006», *Diabetes Care*, vol. 34, n.º 1, pp. 216-219, ene. 2011, doi: 10.2337/DC10-0879.
- [11] J. X. Moore, N. Chaudhary, y T. Akinyemiju, «Metabolic Syndrome Prevalence by Race/Ethnicity and Sex in the United States, National Health and Nutrition Examination Survey, 1988–2012», *Prev Chronic Dis*, vol. 14, n.º 3, 2019, doi: 10.5888/PCD14.160287.
- [12] «Índice cintura-altura - Wikipedia, la enciclopedia libre». [https://es.wikipedia.org/wiki/%C3%8Dndice\\_cintura-altura](https://es.wikipedia.org/wiki/%C3%8Dndice_cintura-altura) (accedido may 08, 2022).
- [13] R. R. Kalyani, C. D. Saudek, F. L. Brancati, y E. Selvin, «Association of Diabetes, Comorbidities, and A1C With Functional Disability in Older Adults Results from the National Health and Nutrition Examination Survey (NHANES), 1999–2006», *Diabetes Care*, vol. 33, n.º 5, pp. 1055-1060, may 2010, doi: 10.2337/DC09-1597.
- [14] J. E. Gangwisch, D. Malaspina, B. Boden-Albala, y S. B. Heymsfield, «Inadequate Sleep as a Risk Factor for Obesity: Analyses of the NHANES I», *Sleep*, vol. 28, n.º 10, pp. 1289-1296, oct. 2005, doi: 10.1093/SLEEP/28.10.1289.
- [15] M. G. Saklayen, «HYPERTENSION AND OBESITY (E REISIN, SECTION EDITOR) The Global Epidemic of the Metabolic Syndrome», 1906, doi: 10.1007/s11906-018-0812-z.
- [16] M. Ashwell, L. Mayhew, J. Richardson, y B. Rickayzen, «Waist-to-Height Ratio Is More Predictive of Years of Life Lost than Body Mass Index», 2014, doi: 10.1371/journal.pone.0103483.
- [17] F. Hadaegh, A. Zabetian, H. Harati, y F. Azizi, «Waist/height ratio as a better predictor of type 2 diabetes compared to body mass index in Tehranian adult men--a 3.6-year prospective study», *Exp Clin Endocrinol Diabetes*, vol. 114, n.º 6, pp. 310-315, jun. 2006, doi: 10.1055/S-2006-924123.

# Transformación OMOP-CDM del repositorio para usos secundarios de la cohorte de la Red Española de Investigación en SIDA (CoRIS)

M. Pascual-Carrasco<sup>1,4</sup>, A. Burgos Llamo<sup>1</sup>, I. Jarrín Vera<sup>2,4</sup>, M. Pedrera-Jiménez<sup>3</sup>, N. García-Barrio<sup>3</sup>, R. Sánchez-de Madariaga<sup>1</sup>, A. Muñoz Carrero<sup>1,4</sup>

<sup>1</sup> Unidad de Investigación en Salud Digital. Instituto de Salud Carlos III, Madrid, España, [mario.pascual@isciii.es](mailto:mario.pascual@isciii.es)

<sup>2</sup> Centro Nacional de Epidemiología, Instituto de Salud Carlos III, Madrid, España

<sup>3</sup> Unidad de Ciencia de Datos, Hospital Universitario 12 de octubre, Madrid, España

<sup>4</sup> CIBER de Enfermedades Infecciosas, CIBERINFEC, Madrid, España

## Resumen

*CoRIS es el registro longitudinal de datos clínico-epidemiológicos de la cohorte española de investigación SIDA. CoRIS es conforme a HICDEP, modelo relacional mono-propósito no alineado con las características que constituyen las bases esenciales en repositorios avanzados para usos secundarios: capacidad de integración flexible de nuevas fuentes de datos, interoperabilidad semántica y diseño multi-propósito. Este aspecto supone una dificultad para la evolución de CoRIS en relación con la reutilización y uso compartido de datos y la replicación de estudios. En el ámbito de los usos secundarios, se posicionan como referencia los modelos comunes de datos (CDM); entre ellos, destaca OMOP-CDM. A partir de CoRIS/HICDEP, se ha implementado una nueva instancia conforme al modelo OMOP-CDM. Se ha evaluado la nueva instancia CoRIS/OMOP frente a CoRIS/HICDEP y se ha demostrado no-inferioridad en relación con: capacidad de cobertura de conceptos; integridad de datos; y capacidad para replicar cohortes complejas. De 410 conceptos clínico-epidemiológicos en CoRIS, se ha alcanzado una cobertura semántica óptima en un 79% y sub-óptima 20,2%. Se ha verificado la integridad de los datos al 100% (identificación y extracción). Se ha replicado de forma precisa las cohortes de tres estudios científicos publicados en dominios combinados relevantes para CoRIS (análisis de enfermedades defintorias SIDA; pautas antirretrovirales; pruebas de laboratorio). CoRIS/OMOP demuestra ser un repositorio no-inferior a CoRIS/HICDEP, adicionalmente aportando a CoRIS capacidad semántica en persistencia y explotación (reutilización y replicabilidad de evidencia), y escalabilidad en volumen e integración de nuevas fuentes y dominios de datos sin modificaciones en su modelo de datos.*

## 1. Introducción

Desde organismos e instituciones internacionales, se promueve la persistencia e integración de los datos en repositorios sujetos a los principios FAIR para compartir, reutilizar, y conformar conjuntos de datos fiables, precisos, normalizados, anonimizados, accesibles, semánticamente interoperables y puestos a disposición de la comunidad científica para la generación y replicación de nueva evidencia, la creación de redes de conocimiento biomédico, dirigidas a maximizar el impacto y los beneficios de la investigación en las necesidades de salud de la sociedad [1].

En el ámbito de la persistencia de los datos para usos secundarios, instituciones de investigación, sanitarias, consorcios científicos, despliegan repositorios y redes de repositorios conformes a modelos de datos comunes (CDM), multi-propósito, que posibilitan el almacenamiento e intercambio de datos, escalado y explotación en entornos confiables, y con alto potencial colaborativo [2,3]. OMOP (Observational Medical Outcomes Partnership), impulsado por OHDSI (Observational Health Data Sciences and Informatics), se sitúa entre los CDM semánticamente interoperables para persistencia y explotación en usos secundarios en registros longitudinales de datos de salud [4].

CoRIS [5] es una cohorte abierta, multicéntrica y prospectiva de ámbito nacional constituida por sujetos reclutados en unidades asistenciales de VIH del Sistema Sanitario Público español, con infección confirmada por el VIH y sin tratamiento antirretroviral en el momento de la inclusión. Desde 2004, CoRIS acumula información longitudinal de datos clínico-epidemiológicos de más de 16 mil pacientes procedentes de 45 hospitales en 13 CCAA, constituyendo un recurso fundamental en la investigación clínico-epidemiológica VIH/ SIDA.

El repositorio de datos CoRIS es conforme a HICDEP (HIV Cohorts Data Exchange Protocol) [6]. HICDEP es un modelo de persistencia e intercambio de datos consensuado entre cohortes europeas de investigación, con desarrollo desde 2011, que responde a una aproximación tradicional de diseño relacional mono-propósito (evolución sujeta a cambios en el propio modelo), y mínimo soporte semántico normalizado (codificaciones propietarias). Estos factores constituyen barreras para avanzar hacia las necesidades en relación con la gestión avanzada de datos para usos secundarios [3]: asimilación de nuevos dominios, conceptos y relaciones, reutilización, uso compartido de datos y capacidad de replicación de evidencia.

## 2. Objetivo

A partir del actual repositorio CoRIS basado en el modelo HICDEP (CoRIS/HICDEP), se plantea llevar a cabo:

- La transformación de la instancia CoRIS/HICDEP a CoRIS/OMOP

- Evaluación comparada CoRIS/OMOP frente a CoRIS/HICDEP en términos de no inferioridad

### 3. Material

La instancia CoRIS/HICDEP facilitada desde CoRIS es la versión de noviembre 2018 (CoRIS/HICDEP v2018) con datos de 15.509 pacientes, sobre ficheros DTA (STATA). Para el soporte de servicios y herramientas se ha habilitado una infraestructura telemática sobre máquinas virtuales (VMWare), S.O. Redhat Linux. Las instancias CoRIS/HICDEP y CoRIS/OMOP han sido desplegadas sobre el gestor de bases de datos relacional PostgreSQL 9.2. La infraestructura OMOP-CDM (v5.3.1) se ha puesto en marcha utilizando las herramientas públicas OHDSI [3]: creación de base de datos (tablas, índices, restricciones, carga de vocabularios); acceso a instancia OMOP: atlas (v2.8) y webAPI (v2.9); examen de calidad de la instancia OMOP: librerías R achilles/achilles heel (v1.6.7) y DQD 1.0; mapeo de conceptos con vocabularios estándar: web Athena y Usagi (v1.4.3), y otras web de terminología médica (SNOMED-CT, RXNORM, ATC, LOINC, principalmente); diseño del proceso ETL: Rabbit-in-a-hat; implementación ETL: java 1.8 (eclipse). Para la explotación de datos de la instancia OMOP se ha empleado SQL (pgAdmin 5) y R v4.0.4 (RStudio 1.4).

El equipo de investigación ha estado compuesto por médicos (dominio sanitario); ciencia e ingeniería de datos y biomédica (dominio de los datos); e informática (dominio de ingeniería de software/ sistemas).

## 4. Métodos

### 4.1. Transformación a CoRIS/OMOP

La metodología desarrollada para la transformación de CoRIS/HICDEP a CoRIS/OMOP consiste, básicamente, en un proceso ETL en 4 fases: identificación de conceptos clínico-epidemiológicos en origen y mapeo a conceptos basados en terminologías estándar; diseño del proceso ETL; implementación del proceso ETL; y evaluación de calidad de la instancia OMOP-CDM.

Para disponer de una infraestructura de persistencia homogénea, inicialmente se llevó a cabo una carga desde CoRIS/HICDEP sobre STATA a CoRIS/HICDEP sobre PostgreSQL, manteniendo la misma estructura de tablas, campos, tipos y codificaciones (extracción CSV y carga PostgreSQL); para ello fue utilizado el framework Pentaho Data Integration (Kettle).

Analizando la documentación CoRIS y la referencia HICDEP [6], se realizó la identificación y catalogación de conceptos clínico-epidemiológicos CoRIS. A partir la lista de conceptos, se llevó a cabo un proceso sistemático de mapeo semántico hacia conceptos estándar OMOP-CDM.

A continuación, se realizó el diseño del proceso ETL CoRIS/HICDEP a OMOP-CDM. Los registros OMOP-CDM destino deben replicar íntegramente los registros CoRIS/HICDEP de origen con transformaciones soportadas por la semántica y los mecanismos implícitos del CDM. Este hecho convierte el diseño ETL en un proceso que requiere diversos ciclos de iteración. Una vez

concluido el diseño, el proceso ETL ha sido implementado mediante un programa java “ad-hoc”.

Finalmente, se llevó a cabo un examen de calidad de la nueva instancia CoRIS/OMOP. Los parámetros de calidad se refieren a la completitud, conformidad y plausibilidad de los registros CoRIS/OMOP [8]. A través un extenso conjunto de sentencias SQL y scripts R (proporcionados desde OHDSI), son analizados, principalmente, aspectos sintácticos y semánticos en la nueva instancia en relación con el CDM y los vocabularios estándar. Los informes de calidad guían el proceso iterativo de depuración y mejora progresiva de la nueva instancia.

### 4.2. Evaluación comparada

Se estableció como hipótesis de partida la no inferioridad en capacidades de persistencia y explotación de la nueva instancia CoRIS/OMOP frente a CoRIS/HICDEP. Para corroborar la hipótesis, se planteó tomar como referencias las metodologías y marcos de trabajo propuestos por Garza [7] (enfoque genérico), Voss [8] y Vaclav [9] (enfoque a OMOP). Garza [2] establece un conjunto de categorías, criterios y métricas para llevar a cabo caracterizaciones globales de capacidades de modelos. Fueron seleccionados los siguientes criterios de capacidad: cobertura de conceptos y datos (“content coverage”); implementación de condicionantes y restricciones (“integrity”); y extracción de datos (“queriability”); se establecieron métricas orientadas a OMOP descritas a continuación.

Para determinar la cobertura de conceptos, el grado de correspondencia entre conceptos originales CoRIS/HICDEP y conceptos mapeados en CoRIS/OMOP fue cuantificado en términos de mapeo óptimo, sub-óptimo y no-mapeado. Para verificar la cobertura de datos y capacidad de extracción de datos, se tomó como patrón de referencia el informe anual emitido por CoRIS (elaborado con CoRIS/HICDEP v2018) en el que son exhaustivamente cuantificados los diferentes tipos de datos y agrupaciones de datos en la instancia: demográficos (género, edad, origen, nivel educativo, etc.); modos de infección; laboratorio (CD4, CD8, carga viral, albúmina, ALT/AST, triglicéridos, HDL, LDL, creatinina, plaquetas); virología (anticuerpos VHB, DNA de VHB, antígenos HBe,VHB, anticuerpos VHC, HCV-RNA, hepatitis delta, HIV-1, HIV western blot, sífilis RPR-VDRL); enfermedades defintorias SIDA; pautas TAR; eventos no-SIDA; y causas de muerte. Se desarrolló un conjunto de 25 sentencias SQL y scripts R para la reproducción del informe CoRIS v2018. Adicionalmente, se realizó el análisis de las sentencias SQL (número de JOIN) y de los planes de ejecución asociados (EXPLAIN ANALYZE: tiempo de ejecución y complejidad del árbol de ejecución).

Finalmente, para evaluar la capacidad de reflejar condicionantes y restricciones CoRIS, se planteó llevar a cabo la replicación de las cohortes de 3 estudios publicados basados en CoRIS/HICDEP v2018. Los trabajos fueron seleccionados por CoRIS en tres dominios combinados relevantes: presentación de eventos defintorios SIDA y mediciones de laboratorio (condition/measurement; Rava et al. PMID: [33882093](https://pubmed.ncbi.nlm.nih.gov/33882093/)); análisis de resultados en pautas de

antirretrovirales específicas (drug/measurement; Suárez et al. PMID: [32690099](#)); y análisis comparado entre pautas antirretrovirales (drug/measurement; Alejos et al. PMID: [32667674](#)).

### 5. Resultados

Se completó la implementación de la instancia CoRIS/OMOP, caracterizada según se muestra en la Tabla 1. Se identificaron 709 conceptos, de los cuáles 410 fueron conceptos relacionados específicamente con el ámbito de la salud (clínicos y epidemiológicos). Se muestra la distribución de conceptos: por vocabulario estándar (SNOMED-CT, RXNORM, ATC, UCUM, LOINC y otras); y por dominio: observación, condición de salud, medicación y tratamiento, medición, procedimiento; y otros. A continuación, se indica el número total de registros CoRIS/HICDEP, y en la nueva instancia CoRIS/OMOP. Finalmente, se indica el número total de registros por dominio en la instancia CoRIS/OMOP.

n° pac.	15.509					
concept.	709	410 *solo salud				
vocab.	snomed	rxnorm	atc	ucum	loinc	otro
	552	70	39	12	10	26
	77,86%	9,87%	5,50%	1,69%	1,41%	3,66%
*salud	253	70	39	12	10	26
	61,71%	17,07%	9,51%	2,93%	2,44%	6,34%
dominio	observ.	condit.	drug	meas.	proc.	otro
	355	158	116	43	7	30
	50,07%	22,28%	16,36%	6,06%	0,99%	4,30%
n° regist.	HICDEP		OMOP-CDM			
	2.981.379		11.646.605			
OMOP						
dominio	observ.	condit.	drug	meas.	proc.	otr.(5)
	8.251.784	66.919	70.252	3.231.531	3.990	22.129
	70,85%	0,57%	0,60%	27,75%	0,03%	0,19%

Tabla 1. Caracterización: vocabularios, conceptos, dominios

La Figura 1 muestra el resultado del examen de calidad de CoRIS/OMOP generada por Data Quality Assessment (DQD), que ofreció un resultado de prácticamente un 100% de tests superados satisfactoriamente (ratio: 2.776/2.781) en completitud, conformidad y plausibilidad. Los 5 errores detectados son consecuencia de inconsistencias acarreadas desde CoRIS/HICDEP; no será posible corregirlos porque no corresponde modificar la instancia origen.

	Verification				Validation				Total			
	Pass	Fail	Total	% Pass	Pass	Fail	Total	% Pass	Pass	Fail	Total	% Pass
Plausibility	1911	0	1911	100%	285	2	287	99%	2196	2	2198	100%
Conformance	325	1	326	100%	54	1	55	98%	379	2	381	99%
Completeness	193	1	194	99%	8	0	8	100%	201	1	202	100%
<b>Total</b>	<b>2429</b>	<b>2</b>	<b>2431</b>	<b>100%</b>	<b>347</b>	<b>3</b>	<b>350</b>	<b>99%</b>	<b>2776</b>	<b>5</b>	<b>2781</b>	<b>100%</b>

Figura 1. Resultados calidad DQD para CoRIS/OMOP

En relación con la cobertura de conceptos, la Tabla 2 muestra los resultados referidos a los conceptos clínico-epidemiológicos CoRIS. Se muestra el grado de cobertura total y por dominio.

La evaluación de la cobertura de datos concluyó con una coincidencia del 100%, es decir, la instancia OMOP-CDM contiene todos los datos de origen (quedando corroborada simultáneamente la capacidad de extracción de datos).

	condit.	drug	meas.	proc.	person	event	Total
óptimo	143	74	33	11	31	32	324
	87,2%	59,2%	97%	78,6%	91,2%	82%	79%
sub- óptimo	21	49	1	2	3	7	83
	12,8%	39,2%	3%	14,3%	8,8%	18%	20,2%
no- mapeado	0	2	0	1	0	0	3
		1,6%		7,2%			0,8%
	164	125	34	14	34	39	410

Tabla 2. Conceptos mapeados (cobertura conceptos)

El resultado del análisis de las sentencias SQL se muestra en la Tabla 3, comparando para ambas instancias (promedio/ desviación estándar) el número de JOIN, y la complejidad de los planes de ejecución asociados.

	HICDEP	OMOP-CDM
n° JOIN	0,6 (±0,58)	1,08 (±1,08)
Plan de ejecución		
Nodos	4,24 (±1,69)	7,4 (±4,95)
Profundidad	3,64 (±1,15)	4,72 (±1,74)
Anchura	1,56 (±0,51)	1,76 (±0,52)
Tiempo ejecución (s)	0,28 (±0,61)	1,81 (±3,7)

Tabla 3. Análisis QUERY: complejidad, tiempos de ejecución

Para finalizar, la Tabla 4 muestra los resultados correspondientes a la replicación de las cohortes de los estudios seleccionados por CoRIS como indicador de capacidad para reflejar condicionantes y restricciones a través de CoRIS/OMOP y corroborando su capacidad para la replicación de evidencia. Para cada estudio se muestra el número de pacientes de la cohorte y sub-cohortes (n), reproducidas en cada instancia, y la diferencia en porcentaje. Para ello, se desarrollaron sentencias SQL y scripts R vinculados a cada estudio, respectivamente.

	HICDEP (n)	OMOP (n)	Dif (%)
Rava et al.	14.876	14.587	1,86
LP	6.636	6.225	1,93
LPAD	3.931	3.536	2,18
Suárez et al.	39	39	0
DRV+TDF	16	16	0
DRV+TAF	23	23	0
Alejos et al.	3.945	3.940	0,12
ABC/3TC/DTG	1.341	1.341	0
CBT/EVG/FTC/TAF	679	679	0
TDF/FTC+DTG	471	471	0
TDF/FTC/EVG/CBT	463	463	0
TDF/FTC/RPV	453	453	0
TDF/FTC+bDRV	329	324	1,5
TDF/FTC+RAL	209	209	0

Tabla 4. Resultado de replicación de cohortes/ sub-cohortes

### 6. Discusión

Los escenarios más habituales en el desarrollo de instancias para usos secundarios son: escenario de extracción que implica la conformación de una nueva instancia a partir de la localización-extracción de datos desde sistemas de información asistenciales; y, escenario de replicación, que plantea la generación de una nueva instancia por transformación de instancias previamente constituidas. La replicación conlleva la necesidad de demostrar objetivamente que se conservan las capacidades de la instancia origen (no-inferioridad). Éste último es el escenario abordado en este trabajo, en línea con otros como Voss [8] y Vaclav [9].

Es necesario señalar que CoRIS/HICDEP es una instancia para usos secundarios lo que, intrínsecamente, conlleva implicaciones como son: agregación/simplificación de conceptos; y alta densidad de conceptos en los registros.

En el primer caso, se generan situaciones de ambigüedad en origen que dificulta mapeo con terminologías estándar y por lo tanto la identificación precisa de conceptos estándar. En ambos casos, se acarrea pérdidas de información irreversibles ya que no existe posibilidad de acceso a las fuentes primarias. No obstante, HICDEP ha resultado suficientemente robusto y se ha alcanzado un mapeo semántico óptimo del 79% de los conceptos, resultado en línea con otras experiencias publicadas [8, 9]. El mapeo sub-óptimo del 20,2% (que implica cierto grado de ambigüedad), se centra esencialmente en el dominio DRUG y viene acarreado desde origen ya que en los tratamientos CoRIS no son registrados parámetros como dosis o vía de administración (simplificación). Los conceptos no mapeados son casi irrelevantes (0,8%) y se corresponde con fármacos no recogidos aún en terminologías estándar.

En segundo lugar, los modelos como HICDEP, plantean para los registros unos constructos extensos que aglutinan conceptos sin vinculación semántica, primando una orientación a explotación en ámbitos de interés concretos. Estas propuestas, minimizan el número de registros en los repositorios (alta densidad de conceptos por registro) y agilizan las consultas para extracción de datos, a costa de una menor flexibilidad (cambios menores implican modificaciones de los constructos). Por el contrario, el CDM OMOP plantea registros simples basados en conceptos únicos vinculados a su dominio natural, determinado por la semántica. Estas aproximaciones, generan más registros (menor densidad de conceptos por registro), haciendo más compleja la extracción de datos, pero son más flexibles para adoptar nuevas variables sin necesidad de cambios en el modelo. La multiplicación de registros se hace evidente en CoRIS/OMOP que cuadruplica el número de registros respecto a CoRIS/HICDEP.

La reproducción del informe CoRIS v2018 muestra que todos los datos están y son accesibles, demostrando la consistencia sintáctica y semántica que proporciona OMOP. Esta consistencia ha sido también corroborada en repositorios de otros ámbitos a través de la replicación de fenotipos consensuados [8, 9]. Finalmente, la fidelidad con la que han sido replicadas las cohortes experimentales de los trabajos analizados implica que es posible generar extracciones complejas de datos en OMOP-CDM y, por lo tanto, replicar evidencia. Tras análisis, las mínimas diferencias en los “n” de las cohortes fueron debidas a sutiles diferencias en el tratamiento aritmético de fechas y periodos temporales efectuados desde STATA, R o SQL.

Es evidente que el acceso SQL a CoRIS/OMOP es más costoso en complejidad de consultas y tiempo requerido que a CoRIS/HICDEP. En promedio, para el principal conjunto de consultas analizado (extracción masiva de datos para la replicación del informe CoRIS), el tiempo ha sido 6 veces superior en OMOP, aunque sobre un conjunto

de consultas no homogéneo, como se observa en las desviaciones estándar. No obstante, se puede afirmar que para consultas que implican menos de 2 operaciones JOIN, el tiempo de ejecución se reduce en OMOP pasa a ser 1,43 veces superior.

## 7. Conclusiones

OHDSI ofrece un amplio conjunto de herramientas abiertas y una activa comunidad profesional y científica que facilita la implementación y explotación sobre OMOP-CDM.

CoRIS/OMOP demuestra ser un repositorio no inferior a la CoRIS/HICDEP en los criterios establecidos (cobertura de conceptos; integridad de datos; replicación de cohortes complejas), poniendo de manifiesto que CoRIS/OMOP permite reproducir la evidencia generada desde CoRIS/HICDEP. Además, CoRIS/OMOP aporta las ventajas inherentes de los repositorios avanzados de datos en investigación biomédica y epidemiológica: interoperabilidad semántica, y escalabilidad en capacidad de integración de nuevos ámbitos y fuentes de datos.

## Agradecimientos

Trabajos realizados en el marco del proyecto AES, Smart-PITeS (PI18CIII/ 00004, PI18/ 01047); y proyecto RIS EPICLIN15\_2019, como contribución a RETIC RIS RD16CIII/0002/0006.

## Referencias

- [1] eHealth: Digital health and care. European health-data space. [https://ec.europa.eu/health/ehealth/dataspace\\_en](https://ec.europa.eu/health/ehealth/dataspace_en) (Consultada: junio 2022)
- [2] Weiskopf NG, Weng C. Methods and dimensions of electronic health record data quality assessment: enabling reuse for clinical research. *J Am Med Inform Assoc.* 2013;20(1):144-151. doi:10.1136/amiainl-2011-000681
- [3] Kent, S., Burn, E., Dawoud, D. et al. Common Problems, Common Data Model Solutions: Evidence Generation for Health Technology Assessment. *PharmacoEconomics*, vol 39, 2021, pp 275–285 (DOI: 10.1007/s40273-020-00981-9)
- [4] Observational Medical Outcomes Partnership (OMOP) CDM. <https://www.ohdsi.org/data-standardization/the-common-data-model/> (Consultada: junio 2022)
- [5] CoRIS: Cohorte de la Red Española de Investigación en SIDA. <https://coris.isciii.es/> (Consultada: junio 2022)
- [6] HICDEP (HIV Cohorts Data Exchange Protocol). <https://hicdep.org> (Consultada: junio 2022)
- [7] Garza Met al. Evaluating common data models for use with a longitudinal community registry. *J Biomed Inform.* 2016 Dec; 64: 333-341. doi: 10.1016/j.jbi.2016.10.016. Epub 2016 Oct 29. PMID: 27989817; PMCID: PMC6810649
- [8] Voss EA et al. Feasibility and utility of applications of the common data model to multiple, disparate observational health databases. *J Am Med Inform Assoc.* 2015 May;22(3):553-64. doi: 10.1093/jamia/ocu023. Epub 2015 Feb 10. PMID: 25670757; PMCID: PMC4457111.
- [9] Vaclav Papez, et al, Transforming and evaluating electronic health record disease phenotyping algorithms using the OMOP common data model: a case study in heart failure, *JAMIA Open*, Volume 4, Issue 3, July 2021, ooab001, <https://doi.org/10.1093/jamiaopen/ooab001>

# Validación preliminar de ITACA: Un entorno novedoso para estudios de Neurofeedback

A. Martín-Fernández<sup>1,2</sup>, D. Marcos-Martínez<sup>2</sup>, V. Martínez-Cagigal<sup>2,3</sup>, S. Pérez-Velasco<sup>2</sup>, R. Hornero<sup>2,3</sup>

<sup>1</sup> E.P.S, Universidad Carlos III de Madrid (UC3M), Madrid, España, 100428762@alumnos.uc3m.es

<sup>2</sup> Grupo de Ingeniería Biomédica (GIB), E.T.S Ingenieros de Telecomunicación, Universidad de Valladolid, Valladolid, España

<sup>3</sup> Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), España

## Resumen

Los sistemas Brain-Computer Interface (BCI) son aquellos que permiten el control de dispositivos externos a partir de la actividad cerebral del usuario. El Neurofeedback (NF) es una de las aplicaciones de dichos sistemas. Esta técnica busca inducir cambios en la plasticidad cerebral mediante la auto regulación voluntaria de la actividad cerebral del usuario. En este contexto, se ha desarrollado ITACA (Interface for Training Against Cognitive Ageing) como una novedosa plataforma para el desarrollo de estudios de NF. Sin embargo, aún es necesario validar la utilidad de este entorno en dichos estudios. Por ello, se ha diseñado y evaluado a través de ITACA un estudio en personas jóvenes y sanas para mejorar la memoria de trabajo mediante un entrenamiento de NF. En la presente comunicación se han analizado los datos obtenidos a partir de una de las pruebas de evaluación, realizadas como parte del estudio. Tres parámetros resultantes de dicha prueba se han analizado: tiempos de reacción ante un estímulo, precisión de acierto y potencia absoluta en la banda de frecuencia relacionada con la memoria de trabajo. Se realizó una comparación de los valores obtenidos para estos parámetros en la realización de la prueba, previa y posteriormente al entrenamiento de NF. Los resultados que se han obtenido sugieren ciertas mejoras en la memoria de trabajo de las personas que participaron en el entrenamiento. Por lo tanto, se sugiere que ITACA podría ser un entorno útil para el desarrollo de entrenamientos de NF.

## 1. Introducción

Los Brain-Computer Interfaces (BCI) son aquellos sistemas que, a través de la adquisición de la señal cerebral, ejecutan respuestas artificiales que reemplazan, reestablecen, suplementan o mejoran las respuestas naturales ejecutadas por el sistema nervioso central (SNC) [1], permitiendo el control de dispositivos externos en tiempo real. La señal empleada por los sistemas BCI se puede registrar mediante diferentes métodos, entre los que destaca el electroencefalograma (EEG) por su carácter no invasivo, su portabilidad y su bajo coste en comparación con otros métodos.

Entre las diversas aplicaciones del campo de BCI, se encuentra el Neurofeedback (NF). En esta técnica se busca la auto regulación voluntaria de la actividad neuronal. Esto se lleva a cabo por medio de condicionamiento operante, presentando *feedback* en forma de estímulos interpretables por el usuario. La hipótesis en la que se basa esta técnica

establece que a través de la regulación voluntaria de la actividad cerebral se puede inducir plasticidad cerebral, que podría resultar beneficiosa para usuarios con trastornos neuropsicológicos [2]. Debido a sus posibles aplicaciones clínicas en el tratamiento de esta clase de patologías, como el TDHA [2], el NF ha generado gran interés entre la comunidad científica. Esto ha dado lugar al desarrollo de plataformas que permiten la implementación de estudios de NF.

ITACA (Interface for Training Against Cognitive Ageing) [3] es una nueva plataforma de NF que nace con la idea de superar las limitaciones presentes en las plataformas existentes. Consta de un módulo de pre-procesado y extracción de características *online* de la señal de EEG, tres aplicaciones de evaluación cognitiva y tres aplicaciones para el entrenamiento mediante NF. A pesar de sus características novedosas, aun es necesario validar su utilidad en el desarrollo de estudios de NF.

Se ha diseñado un estudio de NF para realizar la validación de la plataforma. En este trabajo se presentan los resultados de dicho estudio. Para ello, se han analizado los datos obtenidos durante las dos sesiones de evaluación cognitiva que se realizaron durante el mismo. Este análisis ha permitido hacer una comparación entre la actividad neuronal antes y después de haber recibido el entrenamiento de NF, así como el cambio en el desempeño de los sujetos a la hora de realizar las pruebas de evaluación cognitiva.

## 2. Métodos

Se ha desarrollado un estudio de NF para validar la plataforma con el que se pretendía entrenar la memoria de trabajo en participantes jóvenes y sanos. Esta función cognitiva ha sido relacionada en otros estudios de NF [4] con las oscilaciones en la banda de frecuencia theta ( $\theta$ , 4-8 Hz), en la línea medio-frontal del cerebro. Dentro de esta banda, cada sujeto presenta un pico de potencia máxima ( $f_m\text{-}\theta$ ) durante la realización de tareas que requieren del uso de la memoria de trabajo. Por ello, en este estudio se diseñó una banda de entrenamiento individualizada para cada sujeto. Esta banda comprendía los 2 Hz en el intervalo  $f_m\text{-}\theta \pm 1$  Hz.

El estudio contó con 19 sujetos. Para identificar los efectos específicos del NF y diferenciarlos de otros no asociados, se dividió a los sujetos en dos grupos. 11 de ellos (4 hombres, edad media:  $24,90 \pm 3,67$  años) constituyeron el grupo de entrenamiento (GE), recibiendo *feedback* respecto a su actividad fm- $\theta$  durante el entrenamiento. Los 8 restantes (4 hombres, edad media:  $25,25 \pm 5,06$  años) constituyeron el grupo placebo (GP) y recibían *feedback* de su actividad cerebral en otras frecuencias, variando la banda según la sesión. De este modo podían experimentar cierta sensación de control durante la sesión, pero no estaban entrenando su memoria de trabajo.

Para el registro de EEG se usó un amplificador g.USBamp (Guger Technologies OG, Graz, Austria) con 16 electrodos activos (F7, F3, Fz, F4, F8, FCz, C3, Cz, C4, CPz, P3, Pz, P4, PO7, POz, PO8), localizados en un gorro estándar 10-10. Como tierra se utilizó el canal AFz y como referencia común el lóbulo de la oreja derecha. La señal se adquirió con una frecuencia de muestreo de 256 Hz.

### 2.1. Protocolo de entrenamiento

Tanto para el GE como GP, el estudio constó de 8 sesiones. La primera y la última fueron empleadas para evaluar la memoria de trabajo de los sujetos y las 6 restantes de entrenamiento de NF.

Ambas sesiones de evaluación presentaron la misma estructura: un registro de la actividad cerebral en reposo de 3 minutos, seguido del registro de la señal de EEG durante el uso de tres aplicaciones que implementaban test centrados en evaluar la memoria de trabajo. Dichos test fueron *3 Back*, *Digit Span* y *Corsi* [5]. A partir de la prueba de *3 Back* de la sesión inicial se calculó el pico fm- $\theta$  dentro del rango 4-8 Hz para cada sujeto del GE. Esto se llevó a cabo mediante un análisis de perturbación espectral empleando la Transformada Wavelet continua.

Las 6 sesiones de entrenamiento de NF seguían una estructura común. Todas ellas comenzaban con un registro de 2 minutos de la actividad cerebral en reposo. A continuación, se realizaban 6 bloques de entrenamiento, con 5 intentos para alcanzar un objetivo en términos de regulación cerebral por bloque y con 45 segundos máximo por intento. Para finalizar, se realizaba un registro de 2 minutos en reposo. La diferencia de estas 6 sesiones radicaba en las aplicaciones empleadas para cada una y los parámetros que configuraban esas aplicaciones. Había 3 aplicaciones distintas, todas ellas presentadas a modo de videojuego para el sujeto, y cada una de ellas con un objetivo diferente. En la primera sesión solo se empleaba una de ellas y a medida que el sujeto progresaba en el estudio, las otras dos se iban incorporando paulatinamente en distinto número de bloques y con distintos parámetros. En este sentido se conseguía un aumento de la dificultad progresivo a medida que se avanzaba en el estudio.

La señal era sometida a un pre-procesado *online*, imprescindible para la correcta actualización del *feedback* en tiempo real. Este consistió en un filtrado espacial Laplaciano corto y un filtrado espectral Butterworth paso banda de orden 4 y fase cero. El *feedback* se obtuvo a partir del cálculo de la potencia en fm- $\theta$  en el canal Fz cada 250

ms. Además, si la potencia en las bandas 0,5-3,5 Hz y 35-40 Hz superaba cierto umbral el *feedback* se inhibía, evitando así interferencias de pestañeo y ruido muscular.

### 2.2. Análisis 3 Back

En la prueba *3 Back* se presenta una matriz 3x3 en la que van apareciendo cuadrados en distinta posición. Cuando la posición del estímulo se repite con la de hace 3, el sujeto debe pulsar una tecla determinada. Esta prueba se llevó a cabo 6 veces en cada sesión, en cada una se mostraron 30 elementos de los que un 40% eran objeto de respuesta.

Esta prueba de evaluación se empleó para el análisis del entrenamiento de NF debido a su extendido uso en estudios de memoria de trabajo. Durante la prueba se presentan dos clases de estímulos: objeto de respuesta (*target*) y no objeto de respuesta (*non-target*), lo que permite la caracterización de la respuesta cerebral ante el estímulo. Esto permite estudiar los cambios en la potencia de fm- $\theta$  en el intervalo justo posterior a la presentación del estímulo, periodo en el que tiene lugar su procesamiento cognitivo. Por otro lado, no solo se puede emplear la precisión de aciertos para el análisis, sino que otros parámetros que proporciona la prueba, como el tiempo de respuesta también son útiles.

Para el análisis de la potencia  $\theta$  en Fz durante el *3 Back* se llevó a cabo el siguiente pre-procesado *offline* empleando funciones implementadas en MEDUSA© Kernel [6], un paquete de Python destinado al procesado y análisis de señales cerebrales desarrollado por el Grupo de Ingeniería Biomédica de la Universidad de Valladolid. Primero, se tomó la señal de EEG de cada sujeto y se aplicó un filtro espectral FIR paso banda en un rango entre 0,5-20 Hz. A continuación, se dividió la señal en épocas de modo que solo se emplearan para el análisis los fragmentos de la señal correspondientes a estímulos *target* respondidos correctamente. Estas épocas eran de 1.3 s de duración desde la presentación del estímulo para que se pudiese extraer la información del procesamiento cognitivo ante el estímulo. Luego, se aplicó un método de rechazo automático de artefactos, descartando las épocas cuya amplitud excediera un umbral basado en la desviación estándar de la señal. Posteriormente, se calculó la densidad espectral de potencia (PSD) mediante el periodograma de Welch, con segmentos de 1 s (resolución espectral 1 Hz), ventana "Boxcar" y solapamiento del 90%, luego se promedió la PSD de cada sujeto. Por último, se calculó la potencia absoluta de la banda  $\theta$  para cada sujeto y se analizaron los cambios entre ambas sesiones de evaluación en los GE y GP empleando la prueba de rangos con signo de Wilcoxon, elegida tras comprobar que los datos no seguían una distribución normal.

Para analizar los tiempos de respuesta se consideraron solo las respuestas correctas a estímulos *target*. Se extrajo para cada sujeto el promedio de los tiempos de respuesta de los 6 bloques. Posteriormente, se analizó la significancia estadística de los cambios en los tiempos de respuesta entre las sesiones pre y post para cada grupo. Para ello se empleó una prueba de rangos con signo de Wilcoxon.

De manera similar se analizó la precisión de acierto. Para cada sujeto se calculó el porcentaje de estímulos *target*

respondidos correctamente en cada uno de los 6 bloques. Luego se analizó la significancia estadística de los cambios de precisión de acierto, pre y post, en el GP y el GE empleando la prueba de rangos con signo de Wilcoxon.

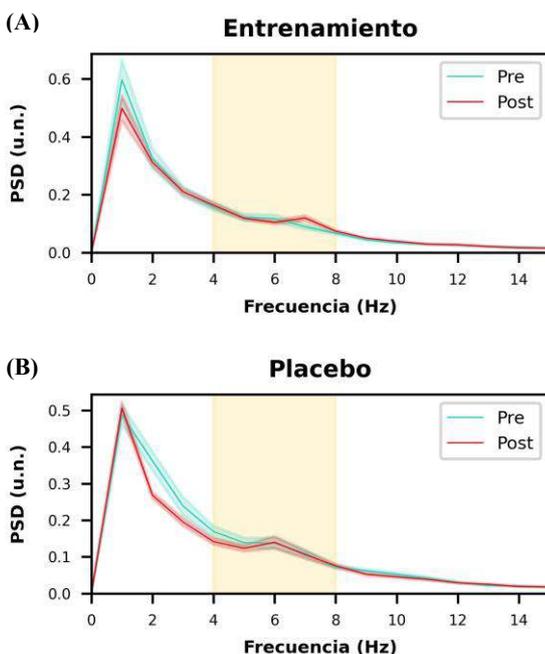
### 3. Resultados

A continuación, se presentan los resultados obtenidos en base al análisis de la prueba 3 Back. En la Tabla 1 se resumen los valores obtenidos para tiempos de respuesta, porcentaje de aciertos y potencia absoluta en  $\theta$  para el canal Fz.

	Placebo		Training	
	Pre	Post	Pre	Post
TR	0,80±0,24	0,81±0,26	0,74±0,27	0,68±0,26
%	73,78±22,91	76,95±25,03	62,37±15,21	75,88±19,56
P( $\theta$ )	0,62±0,35	0,58±0,28	0,55±0,35	0,58±0,19

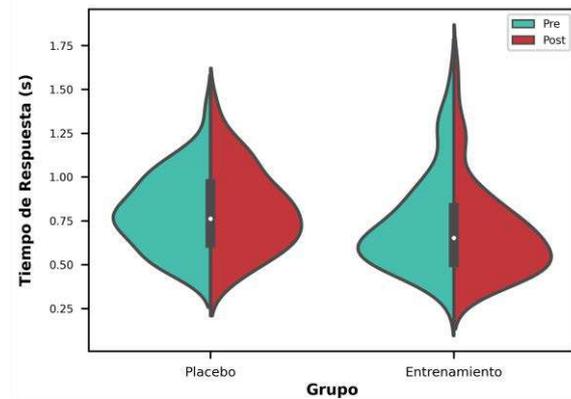
**Tabla 1.** Valores promedio, pre y post, de GE y GP obtenidos en la prueba 3 Back. Tiempos de respuesta (TR) en segundos, porcentaje de aciertos (%) y potencia absoluta en  $\theta$  para el canal Fz (P( $\theta$ )) en unidades normalizadas.

El análisis de potencia absoluta en  $\theta$  para el canal Fz se muestra en la Figura 1. En esta se puede observar que en la PSD del GE aparece un incremento del valor post-NF respecto al pre-NF, entre los 6 y 8 Hz. Dicho incremento no se observa para el GP. Sin embargo, no se encuentran diferencias estadísticamente significativas para ninguno de los grupos ( $p$ -valor > 0,1).



**Figura 1.** Promedio de las PSD, en el canal Fz, correspondientes a las épocas de target respondidos correctamente en la prueba cognitiva 3 Back, en las sesiones pre y post-NF para GE (A) y GP (B).

En el análisis del TR el GE mostró una disminución significativa tras el entrenamiento ( $p$ -valor < 0,01). Por otro lado, para el GP no se observó este cambio en su TR tras las 6 sesiones de entrenamiento ( $p$ -valor > 0,1). Las distribuciones de TR para GE y GP se pueden observar en la Figura 2.



**Figura 2.** Diagrama de violín de la distribución de TR en la prueba de 3 Back, pre y post en GE y GP.

En el análisis de precisión de aciertos el GE mostró un aumento significativo tras las seis sesiones de entrenamiento ( $p$ -valor < 0,001). Por su parte el GP no mostró este aumento en porcentaje de aciertos tras el entrenamiento ( $p$ -valor > 0,1).

### 4. Discusión

En la presente comunicación se ha mostrado una validación preliminar de ITACA [3], una novedosa plataforma para la realización de estudios de NF. Para tal fin, se han empleado los registros de la prueba de evaluación cognitiva 3 Back que realizaron los sujetos del estudio, previa y posteriormente a ser entrenados.

Los resultados obtenidos en el análisis de la prueba 3 Back sugieren una mejora de la memoria de trabajo, función cognitiva evaluada por dicha prueba, únicamente en los sujetos pertenecientes al GE. Esta mejora se ve reflejada en el incremento de aciertos y la disminución del tiempo de respuesta. Puesto que esta mejora no se ha observado en GP, los cambios observados en GE podrían atribuirse al entrenamiento de NF. De hecho, estudios previos en los que se empleó un protocolo de entrenamiento de NF similar, mostraron mejoras de la memoria de trabajo similares a las encontradas en este estudio [7].

Por otro lado, a pesar de no encontrar diferencias significativas en términos de incremento en la potencia absoluta en el rango  $\theta$  para el canal Fz, se ha observado un aumento entre los 6-8 Hz en la PSD de la sesión de evaluación final del GE, Figura 1 (A). Esto podría deberse a que, como ha sido mencionado previamente, cada sujeto perteneciente al GE entrenó una banda de frecuencia, calculada de manera individual en función de su pico  $fm-\theta$ . Para muchos de los sujetos del GE, el rango de frecuencia entrenado fue el comprendido entre 6-8 Hz. Considerando esto, podría explicarse que la diferencia de potencia en la banda  $\theta$  al completo no sea estadísticamente significativa. Otra muestra de que este aumento para el GE pueda ser inducido por el NF es que, en estudios realizados

anteriormente, se ha relacionado el aumento de potencia en  $\theta$  con un mejor rendimiento cognitivo y una mejor memoria de trabajo [8].

Por otro lado, la Figura 1 (B) muestra una disminución de potencia en la banda 2-4 Hz para el GP tras el entrenamiento. Puesto que esta banda está fuera de los rangos entrenados por el GP, se podría asociar este hecho a diferentes componentes asociados al entrenamiento de NF, pero no específicos de este. Por ejemplo, los denominados efectos naturales que pueden ser tanto positivos como negativos e incluyen características propias del sujeto tales como su edad o etapa de desarrollo cognitivo [9]; los efectos relacionados con la repetición de un evento, que incluyen el uso de las mismas aplicaciones varias veces a lo largo de las sesiones, además de la repetición de una misma estructura para todas ellas [9]; los efectos generales no específicos como la interacción entre el sujeto y la persona que conduce la sesión o el entorno en el que esta se desarrolla [9]; y los efectos no específicos relacionados con el NF, como el propio hecho de estar participando en el estudio [9]. La aparición de esta disminución, por tanto, destaca la importancia del uso del GP para reconocer cuales son los efectos específicos del NF en el GE.

Los resultados obtenidos en esta validación preliminar sugieren mejoras en la memoria de trabajo de los sujetos pertenecientes al GE. Por lo tanto, el uso del entorno ITACA [3] como plataforma de entrenamiento de NF podría cumplir con los objetivos propuestos y por tanto ser útil para el desarrollo de estudios de esta clase.

Sin embargo, este estudio no está exento de limitaciones. En primer lugar, es necesario ampliar el número de sujetos para poder así tener una mayor representación de la población de estudio. Además, para la validación solo se consideraron los datos de la prueba cognitiva *3 Back*. Por lo que, sería interesante incorporar los datos de las otras dos pruebas cognitivas realizadas, además de los datos de la propia evolución durante las sesiones de entrenamiento, en un análisis más profundo para obtener unos resultados más completos.

## 5. Conclusión

En este estudio se ha presentado una validación preliminar de la plataforma ITACA [3]. Para ello, se llevó a cabo un estudio de NF a través de las aplicaciones desarrolladas en este entorno. Con el entrenamiento se pretendía mejorar la memoria de trabajo de sujetos jóvenes y sanos, mediante la auto regulación de estos de su potencia en  $\text{fm}-\theta$ .

Se realizaron análisis de los datos obtenidos durante las sesiones de evaluación cognitiva previas y posteriores al entrenamiento de NF. Se consideraron tiempos de reacción a estímulos *target* de la prueba *3 Back*, precisión de aciertos y potencia absoluta en la banda de interés,  $\theta$ , para el canal entrenado, Fz. Se hallaron diferencias significativas para el GE en los tiempos de respuesta y el porcentaje de aciertos, que no se encontraron para el GP. Estos resultados sugieren la utilidad del entorno ITACA [3] como plataforma para el desarrollo de entrenamientos de NF.

## Agradecimientos

Este estudio ha sido financiado por los proyectos PID2020-115468RB-I00, RTC2019-007350-1 y PDC2021-120775-I00 del Ministerio de Ciencia e Innovación/Agencia Estatal de Investigación/10.13039/501100011033/ y FEDER Una forma de hacer Europa; por el proyecto I+D+i ‘Análisis y correlación entre la epigenética y la actividad cerebral para evaluar el riesgo de migraña crónica y episódica en mujeres’ (‘Programa de Cooperación Interreg V-A España-Portugal POCTEP 2014-2020’) y FEDER; y por ‘Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN)’ a través de ‘Instituto de Salud Carlos III’ cofinanciado por FEDER. Diego Marcos-Martínez y Sergio Pérez-Velasco son beneficiarios de una ayuda PIF de la Consejería de Educación de la Junta de Castilla y León.

## Referencias

- [1] J. R. Wolpaw and E. W. Wolpaw. *Brain-Computer Interfaces: Principles and Practice*. Oxford University Press, 2012.
- [2] R. Sitaram *et al.*, “Closed-loop brain training: the science of neurofeedback,” *Nature reviews. Neuroscience*, vol. 18, no. 2, 2017, pp. 86–100.
- [3] D. Marcos Martínez, E. Santamaría Vázquez, V. Martínez Cagigal, S. Pérez Velasco, S. Moreno Calderón, and R. Hornero Sánchez, “ITACA: Un nuevo sistema de entrenamiento cognitivo mediante Neurofeedback,” *Actas de las Jornadas de Robótica, Educación En Automática y Bioingeniería 2022 (Jreb 2022)*, Málaga, 2022, pp. 239–245.
- [4] S. Enriquez-Geppert, R. J. Huster, C. Figge, and C. S. Herrmann, “Self-regulation of frontal-midline theta facilitates memory updating and mental set shifting,” *Frontiers in behavioral neuroscience*, vol. 8, 2014, pp. 420–420.
- [5] A. Baddeley, “Working memory: Looking back and looking forward” *Nat. Rev. Neurosci.*, vol. 4, no. 10, 2003, pp. 829–839.
- [6] E. Santamaría Vázquez, V. Martínez Cagigal, D. Marcos-Martínez, V. Rodríguez-González, S. Pérez-Velasco, S. Moreno-Calderón and R. Hornero, “MEDUSA ©: A novel Python-based software ecosystem to accelerate brain-computer interface and cognitive neuroscience research”, *Comput. Methods Programs Biomed.*, p. (Under Review), 2022.
- [7] T. Brandmeyer and A. Delorme, “Closed-Loop Frontal Midline Theta Neurofeedback: A Novel Approach for Training Focused-Attention Meditation,” *Frontiers in human neuroscience*, vol. 14, 2020, pp. 246–246.
- [8] F. Marlats *et al.*, “SMR/Theta Neurofeedback Training Improves Cognitive Performance and EEG Activity in Elderly With Mild Cognitive Impairment: A Pilot Study,” *Frontiers in Aging Neuroscience*, vol. 12, 2020.
- [9] T. Ros *et al.*, “Consensus on the reporting and experimental design of clinical and cognitive-behavioural neurofeedback studies (CRED-nf checklist),” *Brain*, vol. 143, no. 6, 2020, pp. 1674–1685.

# Bioinstrumentación

---

Miércoles, 23 de noviembre de 2022

# Sistema integral de pupilometría binocular para roedores

Alejandro Lara-Doña<sup>1,2</sup>, Sonia Torres-Sánchez<sup>1,3,4</sup>, Esther Berrocoso<sup>1,3,4</sup>, Daniel Sánchez-Morillo<sup>1,2</sup>

<sup>1</sup> Instituto para la Investigación e Innovación Biomédica de Cádiz (INIBICA), Hospital Universitario Puerta del Mar, Cádiz, España.

<sup>2</sup> Grupo de Investigación de Ingeniería Biomédica y Telemedicina. Departamento de Ingeniería Automática, Electrónica y Arquitectura y Redes de Computadores. Universidad de Cádiz, Puerto Real (Cádiz), España. {alejandro.lara, daniel.morillo}@uca.es

<sup>3</sup> Grupo de Investigación de Neuropsicofarmacología y Psicobiología, Departamento de Psicología, Área de Psicobiología, Universidad de Cádiz, Puerto Real (Cádiz), España, {sonia.torres, esther.berrocoso}@uca.es

<sup>4</sup> CIBER de Salud Mental (CIBERSAM), Instituto de Salud Carlos III, Madrid, España.

## Resumen

*Recientemente el Locus Coeruleus (LC) está siendo objeto de interés por su implicación en los mecanismos de respuesta al dolor y su correlación con la comorbilidad entre trastornos ansio depresivos y el dolor crónico. Las investigaciones sobre este tema presentan evidencias de que cambios en el diámetro de la pupila pueden relacionarse con actividad en el LC. La pupilometría, como una técnica indirecta y no invasiva, permitiría evaluar la actividad del LC cuando se emplea, por ejemplo, estimulación optogenética en esta región. Sin embargo, actualmente esta técnica no se encuentra estandarizada en la investigación preclínica con roedores.*

*En este trabajo se presenta un sistema integral de pupilometría binocular para roedores, incluyendo los componentes físicos y las aplicaciones para la adquisición de imágenes, la estimulación optogenética, y la segmentación automática de las pupilas del animal aplicando técnicas de aprendizaje profundo.*

## 1. Introducción

Existe un creciente interés investigador en el estudio del Locus Coeruleus (LC) como la principal fuente de noradrenalina en el sistema nervioso central y su afectación en patologías neurodegenerativas y trastornos de salud mental. No obstante, la dificultad de estudiar el LC, localizado en la región del puente troncoencefálico, implica un reto tanto en humanos como animales [1,2]. Por ello, la pupilometría se presenta como una técnica indirecta para el estudio de la actividad en el LC. Actualmente, existen evidencias sobre la relación entre las fluctuaciones en el tamaño pupilar y los procesos de atención y procesamiento cognitivo [3,4]. Asimismo, variaciones en el tamaño de la pupila se han correlacionado con cambios en la actividad del LC, que participa en la regulación de dichos procesos [5,6]. Además, el LC está involucrado en los mecanismos de respuesta al dolor y se ha descrito su implicación en la comorbilidad entre dolor crónico y trastornos ansio depresivos [7,8].

Estudios recientes en modelos animales han relacionado la estimulación del LC con la dilatación pupilar, ya sea mediante estimulación eléctrica, técnicas de optogenética o técnicas farmacogenéticas basadas en DREADD (acrónimo en inglés *Designer Receptors Exclusively Activated by Designer Drugs*) [2,9]. La disponibilidad, por tanto, de herramientas que faciliten la monitorización del

tamaño pupilar en modelos animales, permitiendo el registro sincronizado de la respuesta pupilar frente a estímulos de diferente naturaleza, es una prioridad en este ámbito. En este trabajo se presenta un sistema integral de pupilometría para pequeños roedores, hardware y software, que puede ser utilizado para el estudio de la correlación la respuesta pupilar y la actividad del LC cuando se estimula mediante técnicas de optogenética.

## 2. Materiales y métodos

### 2.1. Plataforma para el registro de datos

Para facilitar el registro de datos se diseñó una plataforma para la colocación del animal junto con las cámaras y soportes necesarios [10,11]. La Figura 1 ilustra la plataforma desarrollada en nuestro grupo para el registro binocular.



**Figura 1.** Plataforma desarrollada para la monitorización de los cambios en la pupila de pequeños roedores.

A este montaje se acopló un estimulador de optogenética (Pulser Plus, Prizmatix) que transmite un tren de pulsos TTL a un controlador LED que se encarga de la emisión de los pulsos luminosos necesarios para la fotoactivación/inhibición de la respuesta neuronal en el área cerebral diana. En esta área diana (LC en nuestro caso) se ha administrado previamente un vector viral con una secuencia de un receptor sensible a la longitud de onda de la luz emitida. El controlador LED puede activar o inhibir las células que expresen dichos receptores mediante pulsos luminosos en el LC a través de una fibra óptica implantada en un procedimiento quirúrgico. El dispositivo de estimulación optogenética se conectó vía USB al ordenador encargado de la adquisición, preprocesamiento y registro de los datos experimentales. La comunicación se

realizó mediante puerto serie utilizando el protocolo proporcionado por el fabricante.

## 2.2. Software de control

El método experimental utilizado por los investigadores permite dividir en 2 partes bien diferenciadas el trabajo realizado: a) adquisición de los videos de ambos ojos del sujeto experimental junto con la aplicación de diversos protocolos de estimulación optogenética; b) y análisis de los videos para extraer los datos referentes a la pupila. Este flujo de trabajo es secuencial, es decir, se realizan todos los experimentos planificados y posteriormente se analizan los videos generados. Por ello se han generado dos aplicaciones independientes. En la Figura 2 se muestra el flujo de trabajo y datos entre ambas, así como las entradas/salidas de cada una de ellas.

Ambas aplicaciones se han desarrollado en Python por su capacidad multiplataforma, utilizando Qt para la creación de la interfaz de usuario. Qt, además de permitir la creación de interfaces gráficas, dispone de un mecanismo de comunicación entre componentes y funciones internas que utiliza interrupciones, siendo posible la comunicación entre diferentes funciones e hilos del programa sin llamadas explícitas en el código.

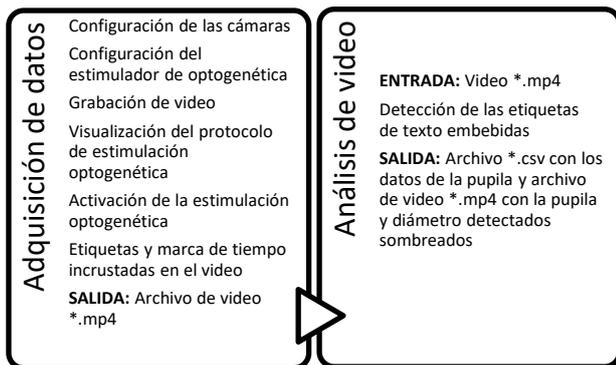


Figura 2. Flujo de trabajo del sistema

### 2.2.1 Adquisición de datos

El objetivo de esta aplicación es el registro y almacenamiento de las imágenes obtenidas por las cámaras para posteriormente medir el cambio del tamaño en la pupila del animal. Esto implica que se ha de asegurar la coherencia temporal entre ambos videos y la estimulación optogenética. Con este fin, se ha optado por una técnica de programación multihilo en la que cada componente del sistema (interfaz de usuario, la comunicación con cada una de las cámaras, el estimulador de optogenética y el almacenamiento de los videos adquiridos) se ejecute en un hilo independiente. En caso de ser necesario, los hilos se pueden ejecutar de forma concurrente en núcleos separados.

Además, para paliar la pérdida de fotogramas, no resuelta de forma óptima por el fabricante lo que origina pérdidas de la coherencia temporal, se ha generado un mecanismo de corrección, duplicando el fotograma anterior al perdido junto con una etiqueta de error para que pueda ser evaluado en el análisis posterior.

La configuración de los parámetros de las cámaras se muestra en la Tabla 1.

Parámetro	Valor
Reloj del píxel	30 MHz
Frecuencia de muestreo	10 fps
Tiempo de exposición	79.085ms
Ganancia	Ajustable por el usuario

Tabla 1. Configuración de las cámaras

### Estimulador de optogenética

La estimulación optogenética se realiza mediante el envío de un tren de pulsos TTL definido por el usuario. Las posibles configuraciones de este tren de pulsos dependerán del experimento a realizar, aunque están limitadas por el fabricante. Se implementaron algunos protocolos útiles para la experimentación en curso.

### 2.2.2 Análisis de video

Esta aplicación funciona como interfaz para la ejecución del algoritmo de detección de la pupila desarrollado anteriormente [12]. El algoritmo de detección de la pupila comprende un proceso de segmentación rápida basado en instancias que emplea la arquitectura SOLOv2 de aprendizaje profundo. A partir de esta segmentación se calcula el diámetro de la pupila como la distancia mayor entre los píxeles de la región seleccionada, las coordenadas del centroide y la circularidad, que junto a las etiquetas detectadas se almacenan en un archivo que permitirá realizar el postratamiento y análisis de los datos en programas externos como MATLAB o Excel. Se ha implementado un mecanismo de lectura de las etiquetas de texto colocadas por el software de registro utilizando la librería easyOCR. De estas etiquetas se extrae la marca de tiempo y el resto del texto para añadirlo como datos al fichero \*.csv que se genera como salida.

## 3. Resultados y Discusión

### 3.1. Aplicación para la adquisición de datos

La interfaz de usuario de la aplicación desarrollada para la adquisición de datos en sesiones experimentales se muestra en la Figura 3. Se pueden visualizar en tiempo real las imágenes obtenidas. En la parte inferior se localizan las pestañas de configuración de las cámaras, el registro en video y la configuración del estimulador de optogenética. La configuración (ganancia de las cámaras, ruta de guardado de los videos, tipo de estimulación, etc.) se guarda automáticamente entre sesiones facilitando la continuación de los experimentos.

#### 3.1.1 Estimulación optogenética

El usuario dispone de 2 modos de configuración del estimulador: a) un modo simple, en el que se envía una señal de una duración, frecuencia y ancho de pulso determinados; b) y un modo avanzado donde es posible controlar el resto de los parámetros para trenes de pulsos asimétricos o retardos entre las foto-estimulaciones (ver Figura 4).

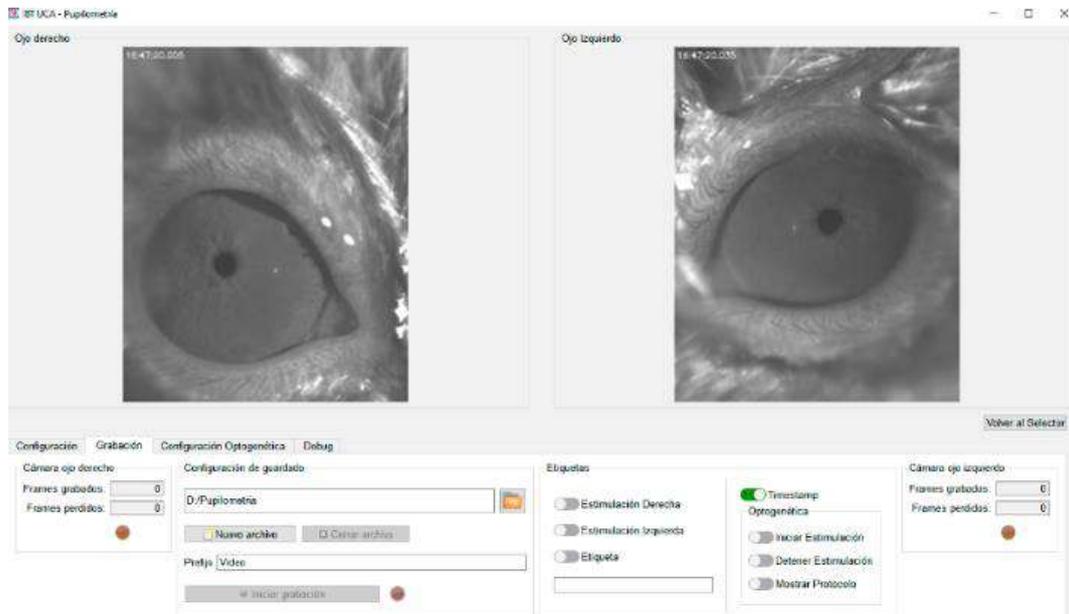


Figura 3. Interfaz de usuario de la aplicación de registro.

Los cambios realizados en estos parámetros se reflejan en tiempo real en la gráfica implementada en la interfaz, permitiendo al usuario visualizar el tren de pulsos que se utilizará en la estimulación.



Figura 4. Configuración de la estimulación optogenética

### 3.1.2 Registro de datos

El usuario puede elegir la ruta de guardado de los videos y el prefijo que estos llevarán, al que se agrega automáticamente la fecha y hora de inicio de la grabación y el ojo a la que esta pertenece (derecho o izquierdo). El usuario ha de iniciar la grabación tras crear los archivos correspondientes para el caso en que no se requiera registrar la sesión de manera continua.

En los extremos de la pantalla se han dispuesto contadores con el número de fotogramas grabados y perdidos de cada cámara, para indicar al usuario si existe algún error en la conexión con alguna de ellas y actuar en consecuencia. Además, se han dispuesto los controles de las etiquetas que serán incrustadas en el video: 3 predefinidas (marca de tiempo, estimulaciones izquierda y derecha) y una personalizada a emplear por usuario durante el transcurso del experimento. También se encuentran los controles del dispositivo de estimulación optogenética: inicio y detención de la estimulación, y detalles del protocolo, que incrusta en el video la información del protocolo

configurado. Esta información se muestra de manera automática durante 1 segundo antes iniciar la estimulación en el sujeto.

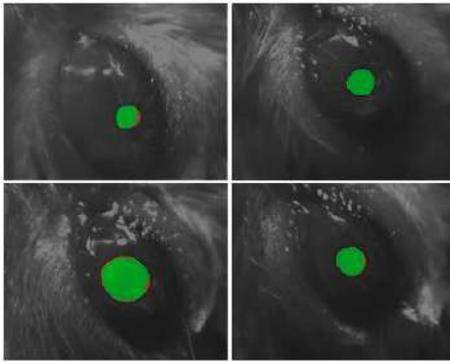
### 3.2. Aplicación para el análisis de videos

En la Figura 5 se pueden observar los elementos de la aplicación desarrollada para el análisis de los registros binoculares de pupila.



Figura 5. Interfaz de la aplicación de análisis

El usuario únicamente ha de indicar el directorio donde se encuentran los archivos de video generados por aplicación de adquisición y, si se desea, detectar las etiquetas incrustadas y/o guardar el video con la región de la pupila detectada superpuesta con el video original. Se ha optimizado la aplicación para que ejecute el análisis de videos en paralelo para aprovechar al máximo el hardware disponible (GPU y CPU). La arquitectura SOLOV2 para la segmentación semántica de la pupila se entrenó y validó con un conjunto de prueba externo. La métrica de intersección calculada sobre la unión fue de 0,93 y el error porcentual absoluto medio fue de 1,89%. Los resultados detallados en [12] confirmaron una alta precisión que hace que el sistema sea adecuado para el seguimiento del tamaño de la pupila en tiempo real, independientemente del tamaño de la pupila, la intensidad de la luz o cualquier característica típica del proceso de grabación en ratones sedados (Figura 6). Además, para evaluar el rendimiento de la aplicación de análisis de los videos se calculó el tiempo empleado en analizar 6 videos (43024 fotogramas en total), utilizando para ello de 1 a 3 hilos, con objeto de analizar la influencia de la paralelización en el rendimiento.



**Figura 6.** Resultado de la segmentación automática de la pupila.

El rendimiento fue evaluado en dos ordenadores equipados con GPUs RTX 2060 y RTX 3070 respectivamente. En la Tabla 2 se puede observar que el tiempo de análisis se reduce en 19% con 2 hilos para ambos PCs.

Nº de hilos	PC 1			PC 2		
	1	2	3	1	2	3
Tiempo de análisis (s)	4218	3434	3388	2811	2279	1969
Reducción (%)	-	19	1	-	19	14
Media uso CPU (%)	55.2	98.5	99.9	53.6	98.7	100
Media uso GPU (CUDA) (%)	61	74.3	78.8	53.2	66.9	77.6

**Tabla 2.** Evaluación del rendimiento de la aplicación de análisis.

#### 4. Conclusiones

En este trabajo se ha presentado una aplicación para el registro binocular de la pupila en pequeños roedores con control de un dispositivo de estimulación optogenética, todo ello unido a la aplicación de análisis de los videos registrados para la segmentación automática de las pupilas empleando una arquitectura basada en aprendizaje profundo para el cálculo de los datos referentes al diámetro, posición y circularidad de la pupila en cada fotograma del video junto con las etiquetas incrustadas.

Estas aplicaciones están siendo utilizadas actualmente en laboratorio para el registro de experimentos en grupos de animales. El siguiente y último paso en el desarrollo del sistema supone la conexión a un hardware de registro de electroencefalografía con el objetivo de correlacionar la actividad en determinadas áreas cerebrales con la dilatación/constricción pupilar.

#### Financiación

Este estudio ha sido financiado conjuntamente por el “Fondo Europeo de Desarrollo Regional” (FEDER)-UE “Una manera de hacer Europa” del “Ministerio de Economía y Competitividad” (MINECO: RTI2018-099778-B-I00); el “Programa Operativo de Andalucía FEDER, Iniciativa Territorial Integrada ITI 2014-2020 Consejería Salud, Junta de Andalucía” (PI-0080-2017); el “Plan Andaluz de Investigación, Desarrollo e Innovación (PAIDI 2020), Consejería de Economía, Conocimiento, Empresas y Universidad, Junta de Andalucía (P20\_00958)”; el “Programa

de Investigación e Innovación de la Unión Europea Horizon 2020 bajo la concesión Marie Skłodowska-Curie N.º 955684”; el “Instituto de Investigación e Innovación en Ciencias Biomédicas de Cádiz (LI19/06IN-CO22; IN-CO09; IN-CO07)”; la “Consejería de Economía, Innovación, Ciencia y Empleo de la Junta de Andalucía” (CTS-510 and TIC-212); el “CIBERSAM”: CIBER-Consortio Centro de Investigación Biomédica en Red- (CB07/09/0033), Instituto de Salud Carlos III, Ministerio de Ciencia e Innovación; y la “Universidad de Cádiz, contrato predoctoral fpuUCA 2019”.

#### Referencias

- Xiang, L. et al. Behavioral Correlates of Activity of Optogenetically Identified Locus Coeruleus Noradrenergic Neurons in Rats Performing T-Maze Tasks. *Sci. Rep.* **2019**, *9*, 1–13.
- Privitera, M. et al. A Complete Pupillometry Toolbox for Real-Time Monitoring of Locus Coeruleus Activity in Rodents. *Nat. Protoc.* **2020**, *15*, 2301–2320.
- de Gee, J.W. et al. Pupil-Linked Phasic Arousal Predicts a Reduction of Choice Bias across Species and Decision Domains. *Elife* **2020**, *9*, 1–25.
- Reimer, J. et al. Pupil Fluctuations Track Fast Switching of Cortical States during Quiet Wakefulness. *Neuron* **2014**, *84*, 355–362.
- Reimer, J. et al. Pupil Fluctuations Track Rapid Changes in Adrenergic and Cholinergic Activity in Cortex. *Nat. Commun.* **2016**, *7*, 1–7.
- Joshi, S.; Li, Y.; Kalwani, R.M.; Gold, J.I. Relationships between Pupil Diameter and Neuronal Activity in the Locus Coeruleus, Colliculi, and Cingulate Cortex. *Neuron* **2016**, *89*, 221–234.
- Llorca-Torralla, M.; Borges, G.; Neto, F.; Mico, J.A.; Berrocoso, E. Noradrenergic Locus Coeruleus Pathways in Pain Modulation. *Neuroscience* **2016**, *338*, 93–113.
- Llorca-Torralla, M. et al. Chemogenetic Silencing of the Locus Coeruleus–Basolateral Amygdala Pathway Abolishes Pain-Induced Anxiety and Enhanced Aversive Learning in Rats. *Biol. Psychiatry* **2019**, *85*, 1021–1035.
- Zerbi, V. et al. Rapid Reconfiguration of the Functional Connectome after Chemogenetic Locus Coeruleus Activation. *Neuron* **2019**, *103*, 702–718.e5, doi:10.1016/j.neuron.2019.05.034.
- Sánchez-Morillo, D.; Lara-Doña, A.; Torres-Sánchez, S.; Berrocoso, E. Plataforma de Pupillometría Lateral En Pequeños Animales 2021.
- Lara-Doña, A.; Torres-Sánchez, S.; Sánchez-Morillo, D.; Berrocoso, E. Sistema Para El Reconocimiento Automático de los cambios en el tamaño de la pupila En Pequeños Roedores. In *XXXVIII Congreso Anual de la Sociedad Española de Ingeniería Biomédica CASEIB 2020*; Sociedad Española de Ingeniería Biomédica, 2020; pp. 455–458 ISBN 9788409254910.
- Lara-Doña, A.; Torres-Sánchez, S.; Priego-Torres, B.; Berrocoso, E.; Sanchez-Morillo, D. Automated Mouse Pupil Size Measurement System to Assess Locus Coeruleus Activity with a Deep Learning-Based Approach. *Sensors* **2021**, *Vol. 21*, Page 7106 **2021**, *21*, 7106.

# Nuevo Método para el Análisis Independiente de la Orientación en Matrices Multielectrodo Equiespaciadas

I. Segarra<sup>1</sup>, S. Ruipérez Campillo<sup>1,2,3</sup>, F. Castells<sup>1</sup>, J. Millet<sup>1</sup>

<sup>1</sup> Instituto ITACA, Universidad Politécnica de Valencia, Valencia, España

{izsegor, saruicam, fcastells, jmillet}@upv.es

<sup>2</sup> Departamento de Medicina, Universidad de Stanford, Palo Alto, Estados Unidos

<sup>3</sup> Departamento de Bioingeniería, Universidad de California, Berkeley, Estados Unidos

## Resumen

El diagnóstico y el tratamiento de las arritmias cardíacas se basan en los registros obtenidos por los catéteres de mapeo electroanatómico con tipología matricial, cuyo potencial no se aprovecha dadas las técnicas de procesamiento y análisis bipolar tradicionales no optimizadas. Esto resulta en la necesidad de un procesamiento más complejo y de exploraciones cardíacas más largas para obtener información precisa sobre el estado del tejido analizado. Este estudio propone una nueva configuración de clique para calcular el electrograma omnipolar en catéteres multielectrodo equiespaciados con el fin de obtener parámetros de interés de forma más robusta y precisa. Se diseñan numerosas simulaciones con parámetros variables para generar un frente de onda plano y emular la propagación de la actividad eléctrica en la superficie del tejido epicárdico. Estas simulan la actividad que capta el catéter, y permiten analizar las diferencias entre el método clásico de análisis omnipolar (clique triangular) y el nuevo método propuesto (clique en cruz). Los resultados muestran que el clique en cruz es más robusto a las variaciones en la dirección de propagación de la onda, y más preciso en la estimación del tiempo de activación local.

## 1. Introducción

La irrupción de los navegadores electroanatómicos cardíacos en los laboratorios de electrofisiología ha supuesto un desarrollo tecnológico sustancial en los últimos años, mejorando la exploración y el diagnóstico de arritmias mediante catéteres reduciendo las tasas de recurrencia, lo que implica importantes beneficios económicos y operativos. Este tipo de procedimiento permite el manejo eficiente de ciertas arritmias como la fibrilación auricular (FA), el aleteo auricular (FIA) o las taquicardias ventriculares (TV), entre otras. El tratamiento más común para esta condición es el procedimiento de ablación, que consiste en la generación de cicatrices en el corazón quemando o congelando la superficie del tejido e interrumpiendo las señales eléctricas responsables de los latidos irregulares [1].

Para identificar y caracterizar la zona que debe ser ablacionada, se utiliza un catéter multielectrodo y un navegador electroanatómico que permite realizar un mapeo 3D de la cavidad cardíaca. La posición del dispositivo multielectrodo es esencial, ya que los mecanismos de perpetuación de las regiones de conducción lenta pueden revelar la región de ablación para diferentes enfermedades, como en el caso del FIA [2, 3]. El análisis de los electrogramas (EGM) es crucial para la toma de decisiones.

En este sentido se buscan variaciones de amplitud, así como la presencia de complejos fraccionados, que pueden reflejar etiologías complejas que deben ser bien definidas [4]. Hasta ahora se han desarrollado varios catéteres flexibles con diferentes formas geométricas para mejorar la exploración y la calidad del mapeo [5]. Aun así, las modalidades de registro más utilizadas siguen siendo los registros unipolares (uEGM) y los bipolares (bEGM). Sin embargo, estas formas tradicionales tienen claras desventajas: los uEGM son sensibles al ruido de baja frecuencia y al campo lejano, y los bEGM dependen en gran medida de la dirección de propagación de la onda, entre otras [6].

El catéter Advisor™ HD Grid de Abbott ha sido recientemente aprobado para uso clínico y consta de 16 electrodos equiespaciados (4 mm de centro a centro) en una matriz de 4x4 con una posición espacial estable para el mapeo del tejido cardíaco de alta densidad. Además, su configuración permite el estudio de registros omnipolares (oEGM), que son independientes de la dirección de propagación de la señal eléctrica [7]. Estas características permitirán una mejor caracterización del sustrato y la determinación de mecanismos-

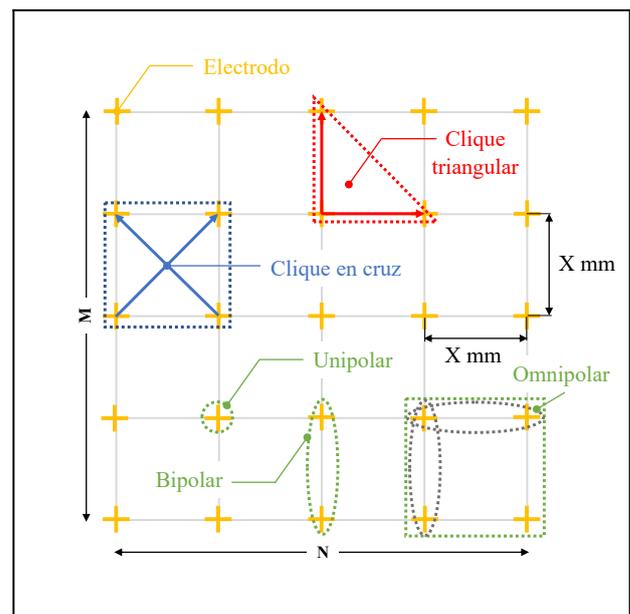


Figura 1. Ilustración de un catéter multielectrodo genérico de tamaño  $M \times N$  con una distancia interelectrodo de  $X$  mm en el eje

vertical y horizontal. En la región superior derecha de la figura se representa el cálculo tradicional del oEGM, es decir, 3 electrodos que conforman un clique triangular. En la parte superior izquierda se ilustra en azul el nuevo método propuesto para calcular el oEGM, utilizando los 4 electrodos disponibles, es decir, todo el clique. Los registros unipolares, bipolares y omnipolares se representan en la parte inferior de la figura en color verde.

El oEGM se calcula utilizando el concepto de *clique*, que se fundamenta en el área formada por 4 electrodos contiguos en forma de cuadrado [10]. El procedimiento habitual para construir un *clique* es elegir 3 de los 4 electrodos para crear un triángulo (véase la Figura 1). No obstante, esta configuración implica limitaciones: entre otras, el centro geométrico resultante no coincide con el centro geométrico del *clique* completo. Además, uno de los electrodos se utiliza dos veces en los cálculos desaprovechando la señal del cuarto electrodo del *clique*.

En este artículo se presenta la hipótesis de que la utilización conjunta de los 4 electrodos (véase la Figura 1) presenta ventajas considerables frente a como se está utilizando en la actualidad, proporcionando un oEGM con mejor amplitud, orientación y un tiempo de activación local (TAL) más preciso, basándose en el hecho de que, utilizando esta configuración, el centro geométrico coincide con el centro del *clique*.

## 2. Materiales y Métodos

### 2.1. Generación del Frente de Onda Plano

Para comprobar la eficacia del método propuesto, se ha diseñado un modelo de propagación plano y homogéneo emulando los datos capturados por un catéter multielectrodo 4x4 (siguiendo el esquema genérico de la Figura 1) en la superficie cardíaca. El modelo de propagación precisa como parámetros de entrada la velocidad, el ángulo, la distancia entre electrodos y una señal unipolar de referencia. La Figura 2 muestra un ejemplo de generación para unos parámetros determinados. Dicha referencia se calcula a partir de las señales unipolares adquiridas por el sistema MapTech®, módulo utilizado para la amplificación, digitalización y procesamiento de señales epicárdicas. Las especificaciones para el análisis son las predeterminadas del software PaceMap®.

	A	B	C	D
1		0,082	0,080	0,079
2	0,080		0,079	0,077
3	0,079	0,077		0,076
4				

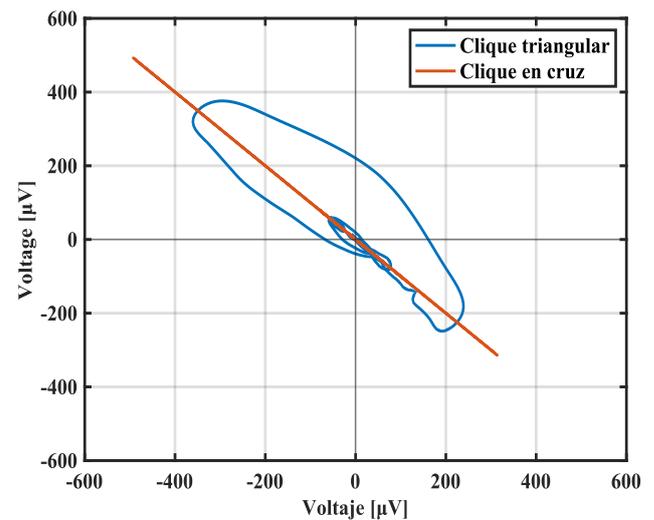
**Figura 2.** Representación de la propagación del frente de onda plano (*s*) siguiendo la disposición del catéter Advisor™ HD Grid para un ángulo de  $3\pi/4$  rad y una velocidad de conducción 0,5 m/s. La escala de grises indica el TAL en segundos en el centro del clique.

Para comprobar el comportamiento de los 4 posibles *cliques* triangulares haciendo uso de uno solo, es necesario realizar un número elevado de simulaciones. Por tanto, se generan 1000 simulaciones con una distancia interelectrodo de 1 mm con una velocidad y un ángulo uniformes generados aleatoriamente en el rango de [0,5-1 m/s] y  $[-\pi, \pi$  rad].

### 2.2. Cálculo de los Bucles del bEGM

El cálculo de la señal bipolar a partir de la unipolar se realiza como indican las flechas de la Figura 1 para el *clique* triangular y en cruz.

Obsérvese en la Figura 3 que el uso de sólo 3 electrodos en el primer caso (triangular) hace que las componentes de la abscisa y la ordenada no estén alineadas, lo que provoca un ensanchamiento del bucle que forman y que sea necesario precisando un tratamiento posterior de la señal [8]. En cambio, en el *clique* en configuración tipo cruz, al coincidir el centro, el único parámetro a considerar es la corrección de  $\pi/4$  rad para alinear los bipolos con el eje de coordenadas.



**Figura 3.** Bucle bEGM obtenido para cada uno de los *cliques* para un ángulo de propagación de  $3\pi/4$  rad y una velocidad de conducción de 0,5 m/s. El efecto de la desincronización del eje en el *clique* triangular es remarkable debido a la anchura del bucle, mientras que en el *clique* en cruz es una línea recta con mayor amplitud.

### 2.3. Estimación del Ángulo de Propagación y Cálculo de los oEGM.

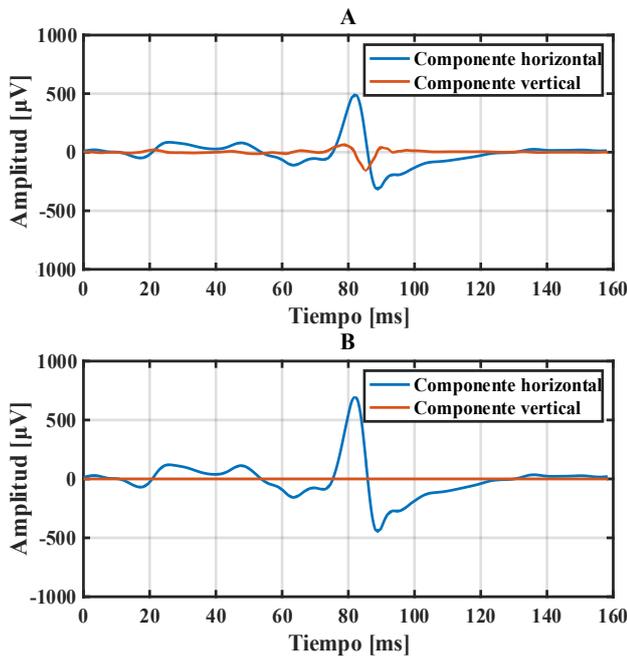
Al transformar el bucle bipolar representado en la Figura 3 de coordenadas cartesianas a polares, se observa que el ángulo de propagación  $\theta$  se corresponde con el ángulo cuya distancia entre la coordenada radial  $\rho$  y el origen es máxima.

Dado que el oEGM se corresponde con el bEGM en la dirección de propagación de la onda, disponer de información de este ángulo permite corregir la dirección

para calcular el oEGM. Analizando las componentes de la abscisa y la ordenada del oEGM, podemos observar que una de ellas es cercana a cero (véase Figura 4). En cambio, la otra tiene una amplitud máxima con respecto a la dirección de propagación.

### 2.4. Estimación del Tiempo de Activación Local

Para corroborar la mejora en la sensibilidad del oEGM, es esencial tener una referencia adecuada con la que comparar el TAL obtenido. La máxima pendiente negativa del uEGM es lo que se aplica comúnmente en la literatura [9]. Para una comparación generalizable y robusta, es necesario tener una referencia para cada una de los *cliques*, ya que no comparten centro. Por lo tanto, el uEGM debe calcularse en el centro del *clique* triangular y en el centro del *clique* en cruz.



**Figura 4. A:** Componentes del oEGM calculados a partir del *clique* triangular. **B:** Mismos componentes del oEGM calculados con el *clique* en cruz. Obsérvese que la amplitud es mayor en el último caso y que la componente perpendicular a la dirección de propagación es despreciable en el método propuesto. Nótese que la corrección angular ha sido aplicada conociendo el ángulo de rotación para evitar sesgos en la detección del ángulo.

Cada referencia se compara con el máximo de la componente horizontal, marcado en azul en la Figura 4, de su correspondiente bEGM. De este modo, es posible determinar el error en la estimación de la TAL.

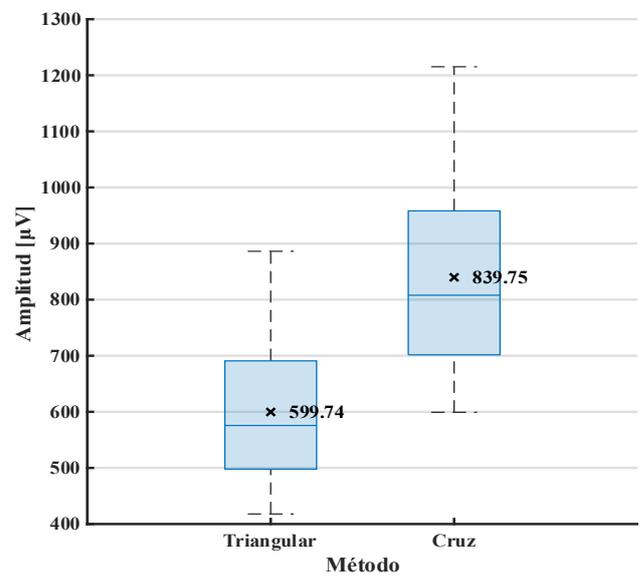
### 3. Resultados

El método tradicional expresa una amplitud media de  $599,74 \pm 117,15 \mu\text{V}$  en la dirección de propagación de la onda, mientras que el nuevo método muestra una amplitud media de  $839,75 \pm 161,18 \mu\text{V}$  (véase la Figura 5). El *clique* en cruz tiene una amplitud media un 40 % mayor que el *clique* triangular. Además, la Figura 6 ilustra cómo la detección es más precisa, disminuyendo el error de detección del TAL de  $210,33 \pm 145,76 \mu\text{s}$  a  $21,51 \pm 29,75 \mu\text{s}$ . Prácticamente un orden de magnitud.

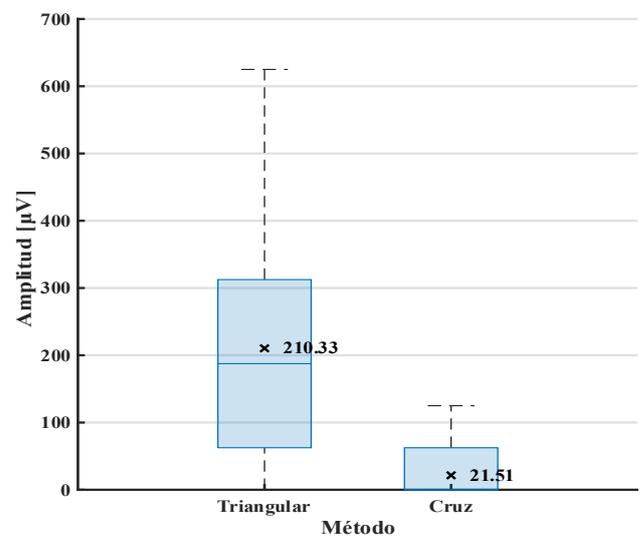
Parámetro	Triangular	Cruz
Amplitud ( $\mu\text{V}$ )	$599,74 \pm 117,15$	$839,75 \pm 161,18$
Error en TAL ( $\mu\text{s}$ )	$210,33 \pm 145,76$	$21,51 \pm 29,75$
Error en TAL ( $\mu\text{s}$ ) *	$217,58 \pm 147,43$	$20,78 \pm 29,68$
Error angular (rad) *	$0,018 \pm 0,14$	$0,013 \pm 0,009$

**Tabla 1.** Resultados de las simulaciones parametrizadas. Los parámetros marcados con \* se han calculado al activar la estimación del ángulo de propagación del bucle bEGM, mientras que los demás se han calculado suponiendo una estimación angular ideal.

Se debe denotar que, al aplicar la detección automática de ángulos, en el caso del método tradicional, la desincronización de los componentes del modelo afecta a los resultados a medida que el bucle bEGM se hace más ancho. Estos resultados se resumen en la Tabla 1.



**Figura 5.** Representación de cajas y bigotes de la amplitud del oEGM, para cada configuración. La mediana se indica con una línea horizontal azul y la media con una x.



**Figura 6.** Ilustración en forma de cajas y bigotes de error absoluto al calcular el TAL. La mediana se indica con una línea horizontal azul y la media con una x.

#### 4. Discusión y Conclusiones

Los catéteres multielectrodo de nueva generación, como el Advisor™ HD Grid, proporcionan nuevas herramientas en el campo de los dispositivos médicos para el diagnóstico de enfermedades cardíacas, la localización y caracterización de anomalías en el tejido cardíaco. No obstante, se necesitan nuevas técnicas de procesamiento de señales para explotar el potencial de estos dispositivos.

El oEGM es una tendencia hoy en día debido al rendimiento que ha demostrado en la determinación de los parámetros necesarios para localizar las regiones que se van a ablacinar, entre otras tareas [6]. El método más habitual en la literatura para el cálculo del oEGM utiliza 3 electrodos, lo que supone una pérdida del potencial del catéter.

Este estudio propone un método para calcular el oEGM de tal forma que se aproveche dicho potencial. Se demuestra además que es posible mejorar el rendimiento en la detección de los parámetros (amplitud y TAL) del oEGM en sistemas electroanatómicos 3D, superando las limitaciones de otras configuraciones de *clique*.

Los resultados de las 1.000 simulaciones con parámetros de entrada variables muestran que, en promedio, la configuración de *clique* propuesto proporciona valores de amplitud y TAL más precisos que el *clique* triangular comúnmente utilizado en la literatura para calcular el oEGM. Esto se debe a que el método propuesto es capaz de detectar siempre la mayor amplitud independientemente de la dirección de propagación. Además, de obtener una estimación más precisa del TAL en la nueva configuración en cruz al compararlo con el TAL de referencia de la señal unipolar.

#### 5. Limitaciones y Direcciones Futuras

Este estudio utiliza una morfología unipolar y una distancia interelectrodo de 1 mm. Además, aunque es una suposición común [10], en el tratamiento de la señal omnipolar, asumimos que la propagación de la onda es localmente plana y una distancia interelectrodo pequeña con respecto a las características electrofisiológicas. Esto implica la necesidad de un estudio exhaustivo para identificar los límites de la distancia entre electrodos de los dispositivos médicos para garantizar el correcto funcionamiento de la tecnología omnipolar. A su vez, las características fisiológicas de cada paciente podrían ser un factor determinante a la hora de localizar estos límites. El trabajo futuro consiste en realizar este análisis en señales clínicas y probar la mejora en los registros obtenidos en el ámbito clínico.

#### Agradecimientos

Este trabajo se ha desarrollado parcialmente en el marco del proyecto PID2019-109547RB-I00 (Programa Nacional de Investigación, Ministerio de Ciencia e Innovación, Gobierno de España), CIBERCV CB16/11/00486 (Instituto de Salud Carlos III) y PROMETEOII/2014/037 (Generalitat Valenciana).

#### Referencias

- [1] Lippi G, Sanchis-Gomar F, Cervellin G. Global epidemiology of atrial fibrillation: An increasing epidemic and public health challenge. *International Journal of Stroke* 2021;16. ISSN 17474949.
- [2] Ruipérez-Campillo S, Castrejón S, Martínez M, Cervigón R, Meste O, Merino JL, Millet J, Castells F. Non-invasive characterisation of macroreentrant atrial tachycardia types from a vectorcardiographic approach with the slow conduction region as a cornerstone. *Computer Methods and Programs in Biomedicine* 2021;200. ISSN 18727565.
- [3] Ruipérez-Campillo S, Castrejón S, Martínez M, Cervigón R, Meste O, Merino JL, Millet J, Castells F. Slow conduction regions as a valuable vectorcardiographic parameter for the non-invasive identification of atrial flutter types. In *2020 Computing in Cardiology. IEEE, 2020; 1–4*.
- [4] Narayan SM, Wright M, Derval N, Jadidi A, Forclaz A, Nault I, Miyazaki S, Sacher F, Bordachar P, Clementy J, Jaïs P, Hasisaguerre M, Hocini M. Classifying fractionated electrograms in human atrial fibrillation using monophasic action potentials and activation mapping: Evidence for localized drivers, rate acceleration, and nonlocal signal etiologies. *Heart Rhythm* 2011;8. ISSN 15475271.
- [5] Maury P, Monteil B, Marty L, Duparc A, Mondoly P, Rollin A. *Three-dimensional mapping in the electrophysiological laboratory, 2018*.
- [6] Magtibay K, MasséS, Asta J, Kusha M, Lai PF, Azam MA, Porta-Sanchez A, Haldar S, Malebranche D, Labos C, Deno DC, Nanthakumar K. Physiological assessment of ventricular myocardial voltage using omnipolar electrograms. *Journal of the American Heart Association* 2017;6. ISSN 20479980.
- [7] Nguyen M, von Alvensleben JC, Runciman M, Collins KK. Advisor high-definition hd grid catheter for mapping accessory pathways in pediatrics. *Heart Rhythm Case Reports* 2020;6. ISSN 22140271.
- [8] Riccio J, Alcaine A, Rocher S, Martinez-Mateu L, Laranjo S, Saiz J, Laguna P, Martínez JP. Characterization of atrial propagation patterns and fibrotic substrate with a modified omnipolar electrogram strategy in multi-electrode arrays. *Frontiers in Physiology* 2021;12. ISSN 1664042X.
- [9] Cantwell CD, Roney CH, Ng FS, Siggers JH, Sherwin SJ, Peters NS. Techniques for automated local activation time annotation and conduction velocity estimation in cardiac mapping. *Computers in Biology and Medicine* 2015;65. ISSN 18790534.
- [10] Deno DC, Balachandran R, Morgan D, Ahmad F, Masse S, Nanthakumar K. Orientation-independent catheter-based characterization of myocardial activation. *IEEE Transactions on Biomedical Engineering* 2017;64. ISSN 15582531.

# Estudio Comparativo con Señales Epicárdicas de las Limitaciones del Omnipolo con Multielectrodos de Alta Densidad

M. Crespo Aguirre<sup>1,2</sup>, S. Ruipérez Campillo<sup>1,3,4</sup>, I. Segarra<sup>1</sup>, A. Guill<sup>1</sup>, A. Tormos<sup>1</sup>, J. Millet<sup>1</sup>, F. Castells<sup>1</sup>

<sup>1</sup>Instituto ITACA, Universitat Politècnica de València, Valencia, España, {saruicam, fcastells, jmillet}@upv.es

<sup>2</sup> Escuela Politécnica Superior, Universidad Carlos III de Madrid, Madrid, España

<sup>3</sup>Escuela de Medicina, Universidad de Stanford, Palo Alto, California, Estados Unidos

<sup>4</sup>Departamento de Ing. de Teleco. e Ing. Electrónica (D-ITET), Instituto Federal Suizo en Zúrich (ETH), Zúrich, Suiza

## Resumen

*Con la aparición de nuevos catéteres multielectrodos se han ampliado las opciones de exploración en un estudio electrofisiológico. En particular el HD Grid de Abbot, con una estructura matricial de 4x4 electrodos separados 4 mm entre ellos, ha dado lugar a un nuevo concepto denominado omnipolo. Este estudio pretende determinar cómo influye la distancia en la determinación de los electrogramas (EGM), realizando un estudio comparativo de distintos métodos de reconstrucción. En efecto, se reconstruyen los EGM a partir de metodologías basadas en configuración triangular y cruzada (recientemente propuesta) a distancias de entre 1 a 4 mm, analizando su efecto sobre el EGM resultante. La base de datos utilizada proviene de una serie experimental de corazón aislado perfundido de conejo de registros multielectrodo epicárdicos de alta densidad (d: 1mm). El enfoque para analizar los resultados se centra en los valores de áreas contenidas en los lazos de los bucles generados por los bipolos, y la relación de componentes de los omnipolos. Se plantea la hipótesis de que a menor distancia inter-electrodo, tanto los electrogramas bipolares como los unipolares son más robustas. Los resultados respaldan esta hipótesis especialmente en cliques reconstruidos mediante configuraciones cruzadas, presentando mejor estimación de la propagación del frente de ondas.*

## 1. Introducción

El estudio local de las propiedades electrofisiológicas es clave en la caracterización del sustrato fibrótico, aspecto estrechamente relacionado con los mecanismos desencadenantes y de perpetuación de arritmias cardíacas [1], siendo su cuantificación y caracterización [2] esencial tanto para el diagnóstico como para el tratamiento de las mismas.

Los navegadores electroanatómicos han demostrado su utilidad para caracterizar el sustrato al proveer cartografías eléctricas del interior de la cavidad cardíaca bajo estudio. Concretamente, el catéter de alta densidad aprobado recientemente (Advisor™ HD Grid mapping catheter -Fa. Abbott, St. Paul, MN, USA-), proporciona una óptima exploración local del tejido miocárdico. Este dispositivo está compuesto por 16 electrodos equiespaciados en una malla de 4x4, proporcionando una mayor resolución sin las limitaciones espaciales de otros catéteres convencionales [3].

La detección de los tiempos de activación local del tejido cardíaco y velocidades de conducción, son entre otros parámetros, cruciales para dicha caracterización del

sustrato fibrótico [4] y el diagnóstico y la caracterización de distintas arritmias [5] [6].

Los electrogramas unipolares (EGMu), registrados a través de diferencias de potencial entre un electrodo de exploración y otro de referencia; y los bipolares (EGMb), obtenidos mediante la diferencia de potencial entre electrodos mínimamente espaciados, siguen siendo los más utilizados.

Recientemente, se ha propuesto una novedosa tecnología de reconstrucción de EGMs, el omnipolo (EGMo) [7], que a diferencia de los métodos empleados anteriormente, es robusta con respecto a interferencias eléctricas de campo lejano y a la interacción del frente de ondas con la orientación de los electrodos que lo registran[8] [9].

La estimación del omnipolo se puede conseguir a partir de varios métodos. El más frecuente se basa en la disposición de 3 registros unipolares en forma de triángulo. Se propone una nueva configuración de *clique*, conjunto de 4 electrodos dispuestos en forma de cuadrado, para la reconstrucción del EGMo. Esta propuesta se basa en una configuración cruzada, utilizando las diagonales de dicho clique para componer los nuevos ejes.

Por tanto se trata de contrastar la robustez del método cruzado para distintas distancias inter electrodo de *cliques* (de 1 a 4 mm) con respecto a las configuraciones triangulares convencionales.

Se hipotetiza que, utilizando la reconstrucción en cruz, cuyo centro coincide con el centro del *clique* se reduce el desfase temporal entre las señales de propagación de onda y se proporciona una medida más precisa de detección del ángulo de propagación de dicho frente. Además, se tiene en cuenta la distancia entre electrodos, pues se asume que para distancias más cortas existe un comportamiento de onda plano y homogéneo y por tanto una estimación más precisa de dicha dirección de activación. En este sentido, se espera que para menores distancias la configuración cruzada sea más rigurosa que las empleadas hasta el momento.

## 2. Materiales

Dada la necesidad de disponer de registros a distintas distancias interelectrodos se optó por utilizar una serie experimental sobre corazones aislados de conejo

perfundidos con la técnica de Langendorff disponible en el grupo de investigación [10]. Los registros empleados, un total de 3, procedentes de un corazón de conejo concreto; se obtuvieron mediante una cuadrícula de 128 electrodos dispuestos a 1mm de distancia entre sí, en contacto con el epicardio del ventrículo a 37°C y con pulsos de estimulación de 4 Hz y 6 Hz mediante el uso de un electrodo bipolar.

Las señales se obtienen mediante un sistema de actividad de mapeado, MapTech© [Waalre, The Netherlands].

### 3. Métodos

#### 3.1 Reconstrucción de EGM bipolares

A partir del multielectrodo de 128 canales unipolares se reconstruyen los bipolos según la propuesta descrita en la figura 1.

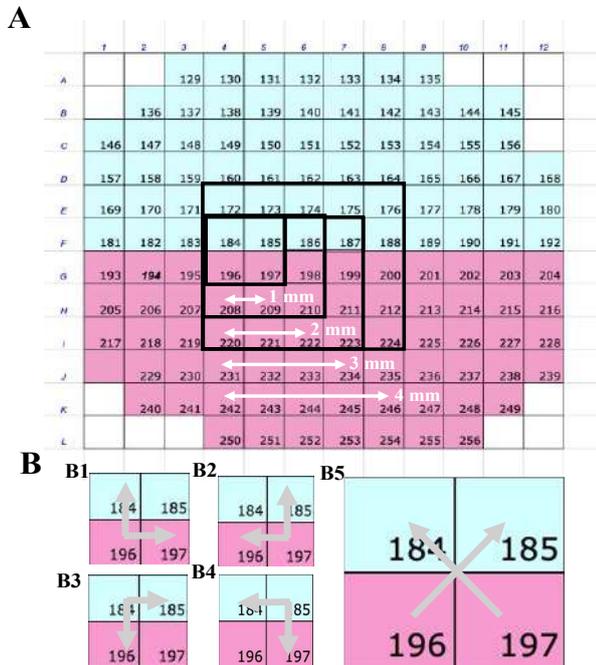


Figura 1. A. Cuadrícula de coordenadas de electrodos espaciados 1 mm. B. Configuraciones de clique. B1-B4. Triangular (1-4) B5. Configuración en cruz.

En el caso del bipolo, viene definido por sus dos componentes:  $\mathbf{b}_x$  y  $\mathbf{b}_y$ . Si tomamos el clique de 1 mm de distancia y siendo  $u$  la señal unipolar registrada por el electrodo en cuestión, definimos las siguientes configuraciones :

- Configuración triangular 1 (fig. 1.B1):  
 $\mathbf{b}_x(t) = u_{197} - u_{196}; \mathbf{b}_y(t) = u_{184} - u_{196}$
- Configuración triangular 2 (fig. 1.B2):  
 $\mathbf{b}_x(t) = u_{196} - u_{197}; \mathbf{b}_y(t) = u_{185} - u_{197}$
- Configuración triangular 3 (fig. 1.B3):  
 $\mathbf{b}_x(t) = u_{185} - u_{184}; \mathbf{b}_y(t) = u_{196} - u_{184}$
- Configuración triangular 4 (fig. 1.B4):  
 $\mathbf{b}_x(t) = u_{184} - u_{185}; \mathbf{b}_y(t) = u_{197} - u_{185}$

- Configuración en cruz (fig. 1.B5):  
 $\mathbf{b}_x(t) = u_{185} - u_{196}; \mathbf{b}_y(t) = u_{184} - u_{197}$

Posterior a la composición de las componentes bipolares, en el caso de la configuración en cruz, es necesaria una corrección en sentido horario de  $\frac{\pi}{4}$  rad de los ejes.

La señal bipolar resultante se expresa en la siguiente ecuación:

$$\mathbf{b}(t) = \begin{bmatrix} \mathbf{b}_x(t) \\ \mathbf{b}_y(t) \end{bmatrix} \cdot \begin{bmatrix} \cos\left(\frac{\pi}{4}\right) & -\sin\left(\frac{\pi}{4}\right) \\ \sin\left(\frac{\pi}{4}\right) & \cos\left(\frac{\pi}{4}\right) \end{bmatrix} \quad [\text{Ec. 1}]$$

#### 3.2 Reconstrucción de EGM Omnipolares

Un bipolo virtual u omnipolo se define como una corrección del EGMb de tal forma que la componente horizontal del mismo, se oriente de forma paralela a la dirección de propagación del frente de ondas, o dirección de activación [11].

Al representar el campo eléctrico, definido por las componentes ortogonales del bipolo, se genera un lazo, cuya área es menor cuanto más precisa es la estimación de la dirección de propagación de onda en el clique. El omnipolo se consigue proyectando las componentes del bipolo en el ángulo de propagación del frente de ondas, estimado mediante dicho lazo. Este ángulo se calcula mediante la ecuación 2.

$$\beta = \arctan\left(\frac{\mathbf{b}_y(t)}{\mathbf{b}_x(t)}\right) \quad [\text{Ec. 2}]$$

donde  $\mathbf{b}_x(t)$  y  $\mathbf{b}_y(t)$  constituyen el vector de módulo máximo generado por el lazo. El omnipolo se reconstruye por tanto mediante la ecuación 3.

$$\mathbf{o}(t) = \mathbf{b}(t) \cdot \begin{bmatrix} \cos(\beta) & -\sin(\beta) \\ \sin(\beta) & \cos(\beta) \end{bmatrix} \quad [\text{Ec. 3}]$$

#### 3.3 Análisis Estadístico

Se propone comparar los resultados obtenidos empleando el método de reconstrucción en cruz con el triangular. En primer lugar, se fija el estudio en la dispersión de áreas descritas por los lazos de los bipolos respecto al mismo clique con distancias interelectrodo de 1 a 4 mm.

Para cuantificar el área, se calcula el sumatorio de áreas delimitadas por los cuadrantes que abarca el lazo del bipolo (véase la figura 2A).

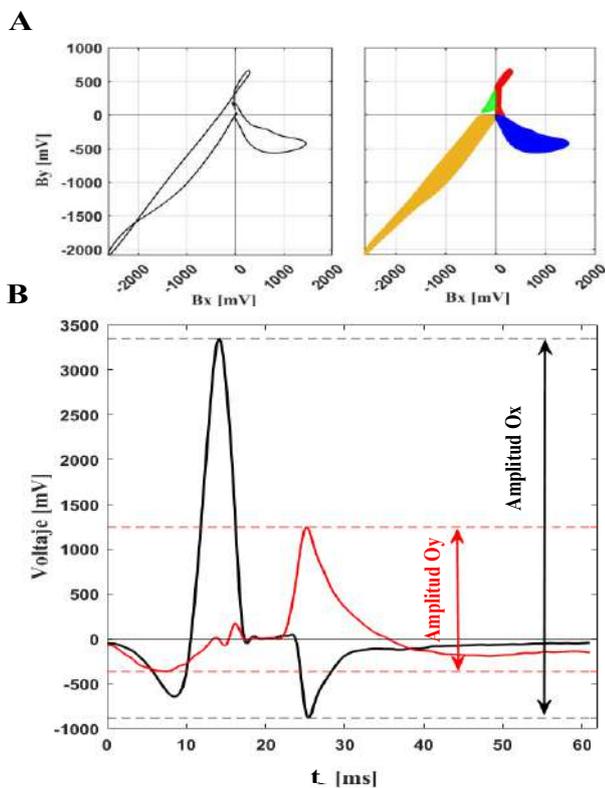
En base a las señales omnipolares, se realiza un estudio de la relación de la componente paralela a dirección de propagación  $\mathbf{o}_x$  y la perpendicular  $\mathbf{o}_y$ . Para ello, se utiliza la amplitud máxima de la componente horizontal, ( $\mathbf{o}_x(t)$ ), y de la vertical ( $\mathbf{o}_y(t)$ ), para calcular la relación entre ambas (Ec. 4).

$$r = \frac{|\mathbf{o}_x(t)_{\max} - \mathbf{o}_x(t)_{\min}|}{|\mathbf{o}_y(t)_{\max} - \mathbf{o}_y(t)_{\min}|} \quad [\text{Ec. 4}]$$

## 4. Resultados

### 4.1 Estudio de los Bipolos

La figura 3 muestra los lazos generados por los bipolos de un registro para la reconstrucción mediante *cliques* triangulares 1 y 2, y en cruz. Se observa que la morfología de los lazos se distorsiona a medida que aumenta la distancia interelectrodo, por el contrario muestran un aspecto más esbelto y alineado en la dirección de propagación cuando  $\leq 2$ mm. Se aprecia una relación directa entre la distorsión de áreas de los lazos y la distancia interelectrodo. La figura 5A ilustra dicha relación. Para distancias de hasta 2 mm, la configuración cruzada genera áreas bipolares en un rango de menor dispersión y para todas las distancias, una media de valores inferiores a los obtenidos por las triangulares.

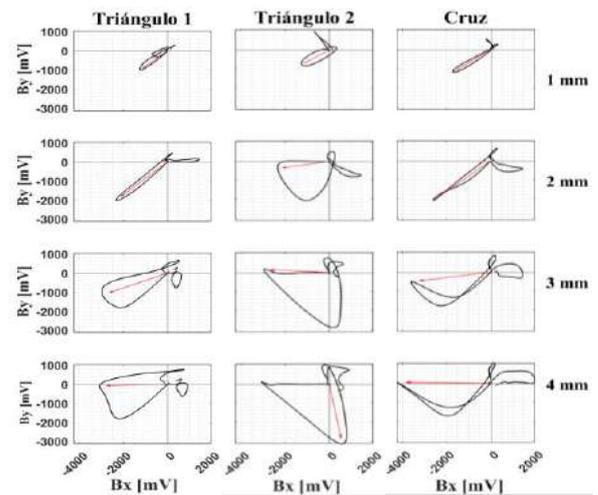


**Figura 2.** Métricas empleadas. **A.** Lazo generado por un bipolo (izquierda) y estimación del área (derecha). **B.** Omnipolo con sus correspondientes componentes representadas para la estimación de la relación.

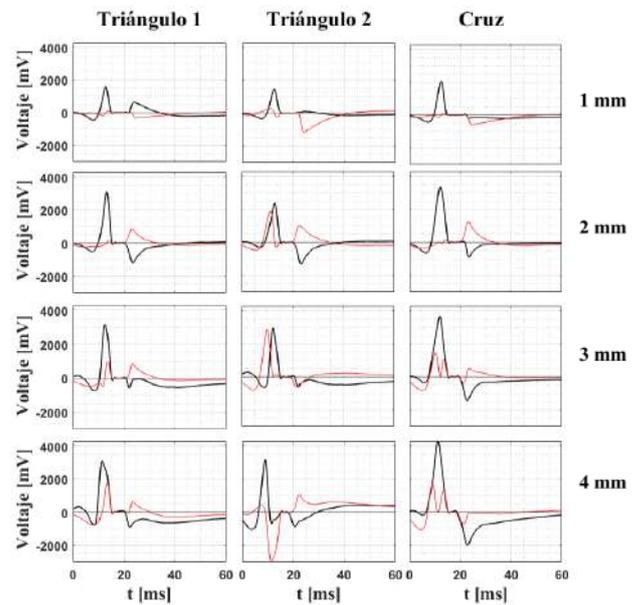
### 4.2 Estudio de los Omnipolos

En contraste con la figura 3, la relación de componentes del omnipolo es inversamente proporcional a la distancia inter-electrodo (véase la figura 5.B).

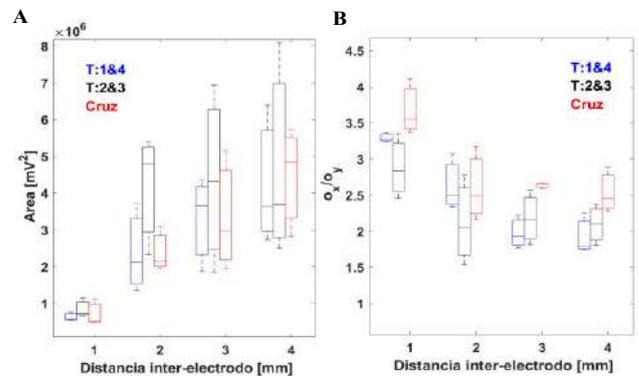
Para todas las distancias, la relación de componentes omnipolares obtenidas a través de la configuración en cruz de clique presenta un rango de valores superior a la obtenida mediante configuraciones triangulares.



**Figura 3.** Ejemplo de lazos generados por bipolos (en negro) y direcciones de propagación de onda en un mismo clique (rojo) con reconstrucción triangular y en cruz (distancias 1-4mm).



**Figura 4.** Ejemplo de omnipoles en un mismo clique reconstruidos a partir de configuraciones triangulares 1, 2 y en cruz mostrando componente horizontal (negro) y vertical (rojo) para 1-4 mm de distancia inter-electrodo.



**Figura 5.** Comparativa de distintos parámetros para distintas distancias inter-electrodo en tres configuraciones: **A.** Dispersión de áreas generadas por lazos de bipolos. **B.** Dispersión de valores normalizados de relación de componentes del omnipolo.

## 5. Discusión

Los métodos de reconstrucción de los EGM en el centro de un clique de registros unipolares se basan en configuraciones triangulares carentes de un centro alineado con el del clique. De ahí que consideremos que la forma actual de obtener los omnipolos es mejorable. La metodología propuesta presenta unas características que permiten reconstruir los EGM de forma más robusta y más inmune a las distancias interelectrodos.

En efecto, para distancias cortas donde se cumple la suposición de un frente de ondas con comportamiento plano y homogéneo en el *clique*, los resultados muestran una estimación de la dirección de propagación precisa para cualquiera de los métodos de configuración (triangulares y cruzado). Esto es coherente con los resultados obtenidos en la dispersión de áreas de los lazos (figura 4) que para distancias de hasta 2 mm que se orientan en la dirección de propagación y se alinean con esta; lo cual permite una detección más precisa del ángulo de propagación y por tanto, una reconstrucción más óptima del bipolo virtual.

El omnipolo obtenido se considera más preciso cuando su componente perpendicular a la dirección de propagación queda prácticamente anulada, dada la corrección del ángulo de las componentes bipolares. Por consiguiente, cuanto mayor es el valor de relación de componentes omnipolares, más precisa es su estimación.

En consecuencia podemos afirmar que para distancias cortas (<2mm) las configuraciones de los cliques estiman los oEGM con mayor precisión, especialmente el método cruzado. Siendo comparativamente mejor en detección de ángulo de propagación (figura 5A) y corrección de ejes bipolares. (figura 5B).

## 6. Limitaciones y Líneas Futuras

El estudio se ha realizado sobre señales epicárdicas obtenidas en corazón aislado perfundido de conejo. Se puede suponer un comportamiento y resultados similares en señales de especies de corazón más grande, asumiendo las mismas dimensiones del catéter y en base a la semejanza electrofisiológica de los corazones de conejo-humano. No obstante, sería deseable disponer de dichas señales a ser posible intracavitarias, que permitieran realizar un estudio similar en función de la distancia y demostrar las suposiciones.

## 7. Conclusiones

El método de estimación del omnipolo mediante una configuración de los ejes bipolares cruzada presenta resultados más precisos que cualquiera de las otras configuraciones triangulares de clique, utilizadas convencionalmente. En base a las distancias interelectrodo empleadas en los catéteres clínicos HD grid, la configuración en cruz tiene por consiguiente un gran potencial de uso, consistente con la adicional mejora en la caracterización del tejido arritmogénico.

## Agradecimientos

Este proyecto ha sido financiado por el Ministerio de Ciencia e Innovación del Gobierno de España (PID2019-109547RB), el Instituto de Salud Carlos III (CIBERCV CB16/11/00486) y la Generalitat Valenciana (PROMETEOII/2014/037).

## Referencias

- [1] N. Akoum *et al.*, “Assessment and Impact of Cardiac Fibrosis on Atrial Fibrillation,” *Current Cardiology Reports*, vol. 16, no. 8. Current Medicine Group LLC 1, Aug. 01, 2014.
- [2] H. Morita *et al.*, “Isolation of canine coronary sinus musculature from the atria by radiofrequency catheter ablation prevents induction of atrial fibrillation,” *Circ Arrhythm Electrophysiol*, vol. 7, no. 6, pp. 1181–1188, Dec. 2014.
- [3] R. Proietti *et al.*, “Enhanced ventricular tachycardia substrate resolution with a novel omnipolar high-density mapping catheter: the omnimapping study.” *Journal of Interventional Cardiac Electrophysiology*, May 2019
- [4] K. Magtibay *et al.*, “Reinserting Physiology into Cardiac Mapping Using Omnipolar Electrograms,” *Cardiac Electrophysiology Clinics*, vol. 11, no. 3. W.B. Saunders, pp. 525–536, Sep. 01, 2019.
- [5] S. Ruipérez-Campillo *et al.*, “Non-invasive characterisation of macroreentrant atrial tachycardia types from a vectorcardiographic approach with the slow conduction region as a cornerstone,” *Comput Methods Programs Biomed*, vol. 200, Mar. 2021.
- [6] S. Ruipérez-Campillo *et al.*, “Characterisation of Atrial Flutter Variants based on the analysis of spatial vectorcardiographic trajectory from standard ECG” 2016.
- [7] D. C. Deno *et al.*, “Orientation-Independent Catheter-Based Characterization of Myocardial Activation,” *IEEE Trans Biomed Eng*, vol. 64, no. 5, pp. 1067–1177, May 2017.
- [8] M. Hwang *et al.*, “Multiple factors influence the morphology of the bipolar electrogram: An in silico modeling study,” *PLoS Comput Biol*, vol. 15, no. 4, 2019.
- [9] P. Podziemski, *et al.*, “Far-field effect in unipolar electrograms revisited: High-density mapping of atrial fibrillation in humans,” in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, Nov. 2015, vol. 2015-November, pp. 5680–5683.
- [10] A. Guill *et al.*, “QT Interval Heterogeneities Induced Through Local Epicardial Warming/Cooling. An Experimental Study,” *Revista Española de Cardiología (English Edition)*, vol. 67, no. 12, pp. 993–998, Dec. 2014.
- [11] J. Riccio *et al.*, “Characterization of Atrial Propagation Patterns and Fibrotic Substrate With a Modified Omnipolar Electrogram Strategy in Multi-Electrode Arrays,” *Front Physiol*, vol. 12, Sep. 2021.

# Development of techniques to minimize the interference of muscle activity on the electroencephalographic signal during swallowing

J.M. Mira Tomás<sup>1a</sup>, J. Imaz Higera<sup>1b</sup>, Y. Ye Lin<sup>2a</sup>, F.J. García Casado<sup>2b</sup>, J.L. Martínez de Juan<sup>2c</sup>, M. Gutiérrez Delgado<sup>3</sup>, G. Más Sesé<sup>4a</sup>, A. Belda Calabuig<sup>4b</sup>, G. Prats Boluda<sup>2d</sup>

<sup>1</sup> Centro de Investigación e Innovación en Bioingeniería (Ci2b), UPV, Valencia, Spain, {juamito, jimahig}@etsii.upv.es

<sup>2</sup> Centro de Investigación e Innovación en Bioingeniería (Ci2b), UPV, Valencia, Spain, {yiye, jgarciac, jlmartin, gprats}@ci2b.upv.es

<sup>3</sup> Medical Rehabilitation Department. Medium- and Long-term care of hospital Pare Jofré, Valencia, Spain, martagutierrezdelgado@gmail.com

<sup>4</sup> Medical Rehabilitation Department. Hospital La Pedrera, Alicante, Spain, {mas\_gemses, belda\_marcala}@gva.es

## Abstract

*Neurofeedback based on electroencephalography (EEG) could be used as an alternative treatment to modify the progression of dysphagia in patients with a neurological origin. During the swallowing process, the physiological interferences that are always present in the EEG signal are muscle and eye activity. In this work we aim to develop a combined method of Multivariate Variational Mode Decomposition (MVMD) and Independent Component Analysis (ICA) techniques to minimize the physiological interferences of the EEG signal. Preliminary results obtained in 11 subjects suggest that the proposed method is effective in minimizing interference while preserving the signal of interest. As a result, it has been possible to identify the brain areas involved during swallowing in healthy subjects: prefrontal, premotor, motor, somatosensory, posterior parietal, occipital and temporal; as well as their temporal sequence of activation. These findings provide the basis for the development of neurorehabilitation therapies to reverse or slow down the dysphagia progression, as well as for its better monitoring.*

not appropriate for monitoring disease progression [4]. The treatments used are mainly palliative, although in recent decades clinical action has focused more on recovering the tone of the swallowing muscles with neuromuscular electrical stimulation or, at the research level, the application of techniques such as repetitive transcranial magnetic stimulation or transcranial direct current stimulation in dysphagia [5].

Another alternative technique in dysphagic patients of neurological origin would be to use neurofeedback to retrain neural circuits through brain plasticity. Neurofeedback is a type of biofeedback usually based on electroencephalography (EEG) that displays real-time visualizations of brain activity for the patient to learn self-regulation of brain function. This technique has been widely used as a treatment for attention deficit disorder, depression, epilepsy and hyperactivity [6]. To date, there is no evidence of its application as a treatment for dysphagia.

The first step towards neurofeedback is to record the electroencephalogram (EEG), which is the recording of the electrical potential variations of the pyramidal neurons of the cerebral cortex on the surface of the scalp. Different waves are distinguished in the EEG according to their spectral bandwidth: delta (1 - 4 Hz), theta (4 - 8 Hz), alpha (8 - 12 Hz), beta (12 - 30 Hz) and gamma (> 30 Hz). The amplitude of the EEG signal has a small order of magnitude (2 - 200  $\mu$ V) and is also embedded in several physiological interferences: eye movements, cardiac activity, muscle activity and movement artefacts.

Therefore, the aim of this work is to develop techniques to minimize physiological interference embedded in EEG recordings obtained from healthy subjects during swallowing. It also aims to establish common patterns of brain activation during the swallowing process, as well as their temporal sequence.

## 1. Introduction

Dysphagia is the difficulty in the progression of the alimentary bolus from the oral cavity to the stomach. Normal swallowing consists of 3 phases: oral (voluntary), pharyngeal and esophageal (both involuntary). Any impairment in one of these phases would lead to dysphagia. Although the etiology of this pathology is very varied, it is true that the most predominant causes are those due to lesions in the cerebral cortex caused by stroke, those derived from neurological diseases and those associated with ageing itself or presdysphagia [1].

Dysphagia prevalence in the world population is around 6 %, but in the elderly population, this dysfunction increases to 22 % [2]. Given the ageing of the population and the increasing prevalence of neurological diseases, these figures are expected to increase in the not-too-distant future. Moreover, this would place an additional burden on the National Health System (approximately € 3677/year per patient) [3].

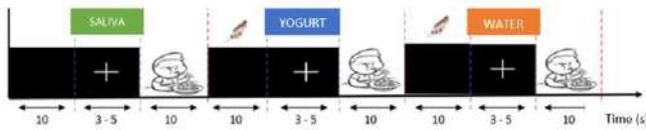
Currently, dysphagia diagnosis is made using the clinical exploratory volume-viscosity based method (V-VECM), although there are also image-based techniques such as videofluoroscopy that have better specificity and sensitivity. However, this has a higher associated cost and uses ionizing radiation that is harmful to the body, so it is

## 2. Materials and methods

### 2.1. Database

An own database of 32-channel EEG recordings was generated from 11 healthy volunteers (4 females and 7 males) aged 19-24 years during a swallowing protocol. For each recording, the 10-20 EEG recording configuration electrode headset (TMSi) was first attached, and the scalp was carefully prepared to reduce the electrode-scalp

impedance to below 10 k $\Omega$ . Additionally, an Ag/AgCl electrode was placed on the left mastoid as a reference electrode. The acquired EEG signals were conditioned with the TMSi SAGA 32+ amplifier and sampled at 2000 Hz.



**Figure 1.** Registration protocol.

For each subject, EEG signals were acquired in 20 trials following the protocol in Figure 1. Each trial consists of 3 boluses of saliva, 10 ml of yoghurt and 10 ml of water. In order to minimally disturb the subject, the experimenter delivered the bolus of yoghurt and water to the subject with a syringe at the pre-set time. The blank cross alerted the subject to be prepared for swallowing, the command for which was the appearance of the child's image on the screen. Okazolab's EventIDE software was used to design the sequential appearance of the different commands in the protocol and to generate the event markers at the same time as the signals were acquired.

## 2.2. Preprocessing

First, the signals were resampled at 200 Hz and a high pass filtered with a 1 Hz cut-off was applied to remove the baseline oscillations.

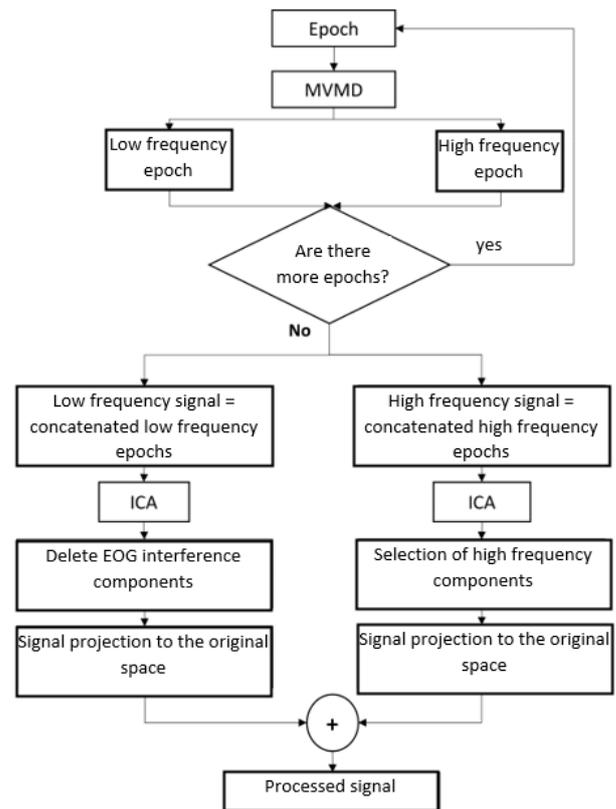
Next, the artefact channels were identified by the FASTER algorithm, and these signals were reconstructed by interpolation with information from adjacent channels.

Subsequently, epochs [-1 3] s around the bolus (saliva, yoghurt or water) swallowing order were extracted, giving rise to a total of 60 epochs/subject, and artefactual epochs were identified and removed by FASTER.

## 2.3. Minimization of Interference

Since the signal-to-interference ratio in the gamma band is very low, the decomposition of the EEG signal by ICA in the full frequency range would only allow a partial separation between brain activity and muscle interference. In this work, a combined MVMD/ICA method was proposed to minimize interference embedded in the EEG signal, the flowchart of which is shown in Figure 2.

First, the Multivariate Variational Mode Decomposition (MVMD) algorithm [7] was applied to decompose the EEG signal into multiple components in a physiological manner. Unlike other methods such as the Wavelet transform, the MVMD does not make any assumptions about the input data [7]. Furthermore, compared to other iterative algorithms such as empirical or variational mode decomposition, MVMD aims to obtain common components across all channels [7]. Specifically, MVMD was applied to each epoch and decomposed into 10 finite amplitude and frequency modulated functions. Subsequently, in order to emphasize the muscle interference to facilitate its subsequent separation with ICA of the gamma activity while preserving the signal of interest, two signals were obtained from each epoch: the "high frequency" signal by summing the components with a center frequency above 30 Hz (typically 4); and the "low frequency" signal by summing the remaining components.



**Figure 2.** Signal processing chart

After the MVMD signal decomposition of all the epochs, the interference embedded in both signals was removed using the Independent Component Analysis (ICA) technique [8]. In the low frequency range, the main aim was to remove the ocular interference that is always present in EEG signals. For this, several criteria had to be fulfilled. The topographic map of the independent component had to be concentrated on the frontopolar electrodes. The power spectral density (PSD) should be mainly concentrated below 10 Hz and have no characteristic frequency peaks in the alpha and beta range. Furthermore, in the Event Related Potential (ERP) image the activations were observed randomly. The low-frequency signal was then recomposed to the original space after removing the ocular interference.

The high-frequency signal, which was dominated by muscle interference, was similarly reconstructed in a very similar way. After applying ICA to all epochs, the components that were of brain origin were visually identified by experts considering that a decreasing trend with increasing frequency had to be observed in the PSD without showing abrupt peaks at high frequencies (> 60 Hz). Furthermore, the presence of activity in the first 500 ms after the swallowing command in the 30 - 60 Hz range in the spectrogram was associated with brain activity, while the increase in ERP amplitude after 500 ms was mainly associated with muscle interference. Then, similarly as low frequency components, the high-frequency signal was recomposed by projecting backwards only the components associated with brain activity.

Finally, the recomposed low-frequency and high-frequency signals were added over time to obtain the processed signals. All the steps were implemented with Matlab v19b.

## 2.4. Identification of the common pattern of brain activations

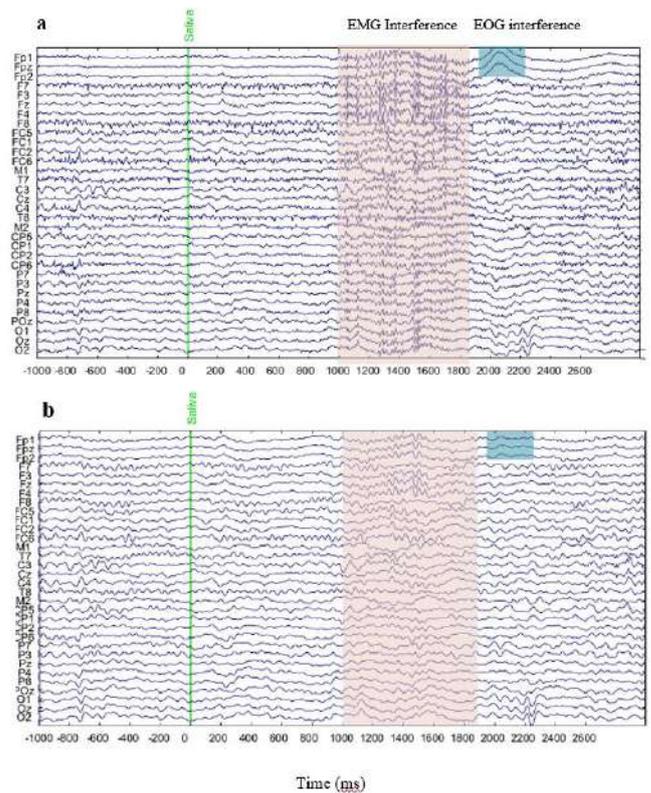
In order to identify the activation of brain areas during swallowing, ICA was first applied to the processed signal of each subject. The independent components of all subjects were then grouped into 15 clusters according to the distribution of their topographic maps. Also, for each subject, a new signal was recomposed using only the brain-derived independent components common to all subjects. The Grand Average ERP was then obtained as the average ERP of the 11 subjects, on which the temporal sequence of activation of brain areas during swallowing in physiological healthy subjects was analyzed.

## 3. Results and discussion

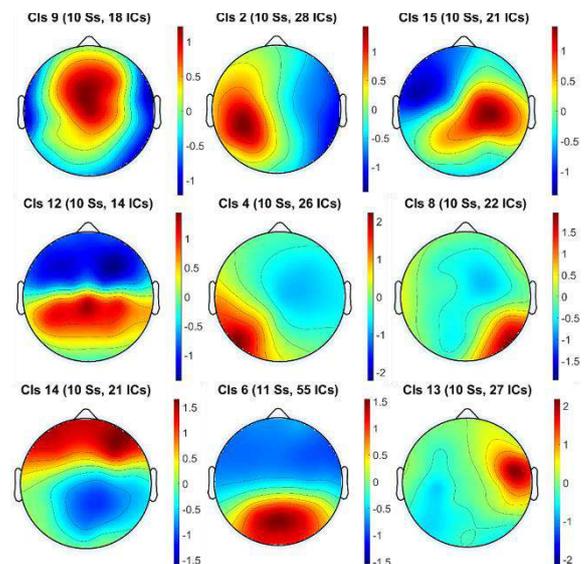
Figure 3 shows an epoch of the "original" EEG signal and the signal processed using the combined MVMD/ICA method. It can be seen in the original signal that the ocular interference has a large amplitude over the prefrontal channels (Fp1, Fp2 and Fpz) as reported in the literature [9]. Muscle interference is present in almost all channels as swallowing involves the voluntary activation of numerous facial and neck muscles. In contrast, neither muscle nor eye interference is visible in the processed signal (b). These results suggest the effectiveness of the MVMD/ICA method for the removal of interference embedded in the EEG signal while preserving the signal of interest. While it is true that this is not the first time that MVMD or ICA [10] have been used to remove interference from the EEG signal to our knowledge this is the first time they have been used together to remove ocular and muscular interference from the EEG signal during the swallowing process.

Figure 4 represents the topographical maps of the common components of the healthy subjects during swallowing. The areas that were clearly identified during swallowing in virtually all subjects (at least 10 of the 11 subjects) were prefrontal, premotor, motor, somatosensory, posterior parietal, temporal and occipital. These areas coincide with what has been shown in the literature on the areas of activation during swallowing in EEG [11, 12].

Figure 5 shows the temporal sequence of brain activations during swallowing obtained from the Grand Average ERP of the 11 study subjects. First, at 125 ms after the start of swallowing, an activation in the prefrontal area is observed, which is associated with the function of selective visual attention [13]. Then, at 175 ms, activation is observed in the occipital area, which may be involved in implicit memory [14]. At 250 ms, in the premotor and motor cortex activations are detected, responsible for planning and sending the swallowing command to the involved muscles respectively [12]. In the interval 300 - 550 ms, activation is identified in the posterior parietal area and in the precuneus (located in the occipital lobe), areas that modulate the motor response according to the sensory information received [15].

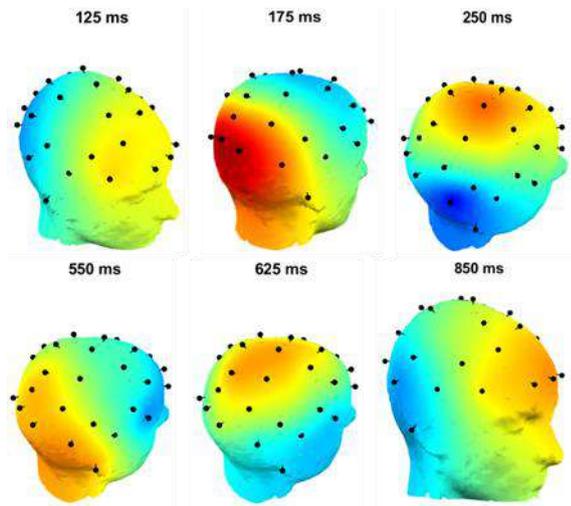


**Figure 3.** An EEG signal epoch during the swallowing of the saliva bolus of subject 1. (a) Epoch "original" signal. (b) Processed epoch signal. The green line in both images represents the time marker of the swallowing command.



**Figure 4.** Topographic maps of the independent components common to the vast majority of subjects. (a) Premotor area, (b-c) left and right motor, (d) somatosensory, (e-f) left and right posterior parietal, (g) prefrontal, (h) occipital and (i) temporal.

Again, at 625 ms, activation of the premotor and motor cortex is observed, which may correspond to the onset of the peristaltic movement of the esophagus [12]. Finally, between 700 and 850 ms there is a complete activation of the frontal lobe which is associated with short-term memory to make it easier to remember the actions performed during the recording protocol [16].



**Figure 5.** Temporal sequence of brain area activations during swallowing from the Grand Average of the 11 subjects.

Finally, it should be noted that this work is not without certain limitations. First, the sample size of the study is relatively small. The results need to be corroborated with an expanded database. Also, the subjects involved in the study are young people aged 20-25 years. However, most patients with dysphagia are older than 65 years. Future work on the characterization of the EEG signal in healthy subjects of similar age to the patients should be pursued. In addition, possible differences in the EEG signal of different boluses have not been analyzed in this work.

#### 4. Conclusions

A new method based on the combined use of MVMD and ICA has been developed for the minimization of physiological interference of the EEG signal during the swallowing process. In addition, a temporal sequence of the activated brain areas during swallowing in healthy subject has been identified. This study may help to better understand how the brain functions during the swallowing process, laying the groundwork for the development of monitoring or neurorehabilitation-aid systems and therapies to reverse or slow down dysphagia progression. The future of this project will aim to evaluate the differences between this process and the main processing methodologies described in the literature.

#### Acknowledgment

Supported by grants from the Generalitat Valenciana, the Polytechnic University of Valencia and the promoters POLISABIO and INVESTIGO. Thanks to all the recording volunteers. Also, thanks to the medical and clinical staff whose involvement permitted to develop this project.

#### References

- [1] Lata Caneda, M. C., & Mourelle, M. Disfagia en las enfermedades crónicas y degenerativas. *Disfagia orofaríngea*, 253.
- [2] Galindo, C. J. A., & Aguilar, D. M. R. (2020). Disfagia en el adulto mayor. *Universitas Medica*, 61(4).
- [3] Westmark, S., Melgaard, D., Rethmeier, L. O., & Ehlers, L. H. (2018). The cost of dysphagia in geriatric patients. *ClinicoEconomics and outcomes research: CEOR*, 10, 321.
- [4] Méndez Sánchez, I. M. (2017). Enfoque clínico-diagnóstico: técnicas diagnósticas. Evaluación y diagnóstico de la disfagia orofaríngea: presentación de casos clínicos.
- [5] Yang, S. N., Pyun, S. B., Kim, H. J., Ahn, H. S., & Rhyu, B. J. (2015). Effectiveness of non-invasive brain stimulation in dysphagia subsequent to stroke: a systemic review and meta-analysis. *Dysphagia*, 30(4), 383-391.
- [6] Fajardo, A., & Guzmán, A. L. (2016). Neurofeedback, aplicaciones y eficacia. *Interdisciplinaria*, 33(1), 81-93.
- [7] ur Rehman, N., & Aftab, H. (2019). Multivariate variational mode decomposition. *IEEE Transactions on Signal Processing*, 67(23), 6039-6052.
- [8] Méndez, F. J. M., & Seoane, M. S. (2008). Análisis de Componentes Independientes en separación de fuentes de ruido de tráfico en vías interurbanas. ETSI Telecomunicación, Universidad de Vigo. C./maxwell s/n, 36310.
- [9] Molla, M. K. I., Tanaka, T., & Rutkowski, T. M. (2012, March). Multivariate EMD based approach to EOG artifacts separation from EEG. In 2012 IEEE international conference on acoustics, speech and signal processing (ICASSP) (pp. 653-656). IEEE.
- [10] Gavas, R., Jaiswal, D., Chatterjee, D., Viraraghavan, V., & Ramakrishnan, R. K. (2020, March). Multivariate variational mode decomposition based approach for blink removal from EEG signal. In 2020 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops) (pp. 1-6). IEEE.
- [11] Jestrović, I., Coyle, J. L., & Sejdić, E. (2015). Decoding human swallowing via electroencephalography: a state-of-the-art review. *Journal of neural engineering*, 12(5), 051001.
- [12] Hamdy, S., Mikulis, D. J., Crawley, A., Xue, S., Lau, H., Henry, S., & Diamant, N. E. (1999). Cortical activation during human volitional swallowing: an event-related fMRI study. *American Journal of Physiology-Gastrointestinal and Liver Physiology*, 277(1), G219-G225.
- [13] Hillyard, S. A., & Anllo-Vento, L. (1998). Event-related brain potentials in the study of visual selective attention. *Proceedings of the National Academy of Sciences*, 95(3), 781-787.
- [14] Pernet, C., Basan, S., Doyon, B., Cardebat, D., Démonet, J. F., & Celsis, P. (2003). Neural timing of visual implicit categorization. *Cognitive Brain Research*, 17(2), 327-338.
- [15] Kern, M. K., Jaradeh, S., Arndorfer, R. C., & Shaker, R. (2001). Cerebral cortical representation of reflexive and volitional swallowing in humans. *American Journal of Physiology-Gastrointestinal and Liver Physiology*, 280(3), G354-G360.
- [16] Lee, J., Yamate, C., Taira, M., Shinoda, M., Urata, K., Maruno, M., ... & Iwata, K. (2018). Prefrontal cortex activity during swallowing in dysphagia patients. *Journal of oral science*, 17-0238.

# Altered pelvic floor muscle activity associated with the use of intravaginal probes in severe vulvodynia

M. Albaladejo-Belmonte<sup>1</sup>, J.M. Mira-Tomas<sup>1</sup>, P. Villa-Muñoz<sup>2</sup>, F.J. Nohales-Alfonso<sup>2</sup>, E. Guijarro<sup>1</sup>, J. Alberola-Rubio<sup>3</sup>, J. Garcia-Casado<sup>1</sup>

<sup>1</sup> Centro de Investigación e Innovación en Bioingeniería (CI2B), Universitat Politècnica de València, Valencia, Spain {moalbel, juamito, eguijarro, jgarcia}@ci2b.upv.es

<sup>2</sup> Servicio de Ginecología y Obstetricia, Hospital Universitari i Politècnic La Fe, Valencia, Spain, {vimupau, fnohales}@gmail.com

<sup>3</sup> Sonda Devices S.L., Valencia, Spain, palberola@sondadevices.com

## Abstract

*Vulvodynia is usually associated with a hypertonic dysfunction of the pelvic floor muscles (PFM), so surface electromyography (sEMG) can be used to assist patient's assessment. Although recordings are typically performed with intravaginal probes, they provoke pain and may alter PFM activity, so the aim of this study was to assess differences in sEMG signals recorded with self-adhesive electrodes in the presence vs. absence of an intravaginal probe according to pain intensity. Twenty-four patients with vulvodynia were treated with incobotulinumtoxinA and monitored after treatment. Their PFM activity and pain were assessed by monopolar and bipolar sEMG recordings with self-adhesive electrodes and by Visual Analogue Scale (VAS), respectively. Recordings performed before treatment and in a follow-up visit were classified into two groups: severe pain and moderate/mild/no pain. Root mean square (RMS), median frequency (MDF) and sample entropy (SampEn) of signals were compared when the probe was into vs. out of the vagina in both groups. In severe pain group, PFM contractile activity had a significantly lower power when the probe was into the vagina than out of it, as shown by lower RMS values in monopolar recordings, especially in M3 (median[interquartile range] in vs. out: 3.8[2.7]mV vs. 4.4[2.5]mV; p-value<0.01). This implies that deep PFM activation decreases during contractions when a probe is inserted into the vagina in severe vulvodynia, probably because of the patients' reluctance to perform maximum contractions, given the pain elicited by the pressure of the probe. Patient's evaluation should thus avoid using intravaginal probes and increase that of self-adhesive electrodes.*

## 1. Introduction

Vulvodynia has been estimated to affect up to 16% of women in their lifetime and has a broad impact on patients' work, couple and social life [1]. It relates to a widespread or localized pain at the clitoris, labia or vestibule for at least 3 months without any identifiable cause [2], for which there are different therapeutic options, such as hormone replacement therapy, amitriptyline or botulinum neurotoxin type A (BoNT/A) injections [3].

Vulvodynia may be associated with different pathophysiological factors, among which are hypertonic dysfunctions of the pelvic floor muscles (PFM) [2]. They consist of different muscle bundles that are typically grouped as superficial PFM and deep PFM, the latter

showing a more intense total electrical activity during contractions than the former [4]. PFM play a key role in vital functions such as pelvic organ support, micturition, defecation and sexual function, which may not be correctly fulfilled in PFM dysfunctions [5].

Given the relationship between vulvodynia and PFM dysfunctions, the patient's evaluation usually comprises an assessment of her PFM tonicity and voluntary activity, for which surface electromyography (sEMG) has proved to be a suitable tool [4]. sEMG recordings can be performed with intravaginal/rectal probes or self-adhesive electrodes attached to the perineum. While probes are usually preferred by researchers, they are uncomfortable and their insertion can even provoke pain [6]. Therefore, although previous studies have shown that they do not alter PFM activity in healthy subjects [7], this may not be true in patients with pelvic pain since this may prevent them from contracting their PFM at maximum, thus deriving in a wrong perception of their PFM function. The aim of the present study was to assess whether PFM myoelectrical activity monitored by sEMG with self-adhesive electrodes is different depending on the presence or absence of an intravaginal probe and pain intensity in vulvodynia.

## 2. Materials and Methods

### 2.1. Overview of the clinical study

This study was performed in the framework of a prospective and longitudinal study carried out in the Pelvic Floor Unit of Hospital Politècnic i Universitari La Fe (Valencia, Spain), which met the Declaration of Helsinki and was approved by the ethics committee of the hospital. Twenty-four patients with vulvodynia that reported an overall >3 pain score according to the Visual Analogue Scale (VAS) provided their informed consent to participate in the study. Additional exclusion criteria were ages under 18 or over 65, vulvar dermatological lesions, pudendal nerve entrapment and active pelvic/vulvovaginal infections. Patients' mean age and number of pregnancies were 42.7±12.5 years and 1.0±1.1, respectively, and 25% of them had menopause.

The study consisted of a first visit (Week 0), in which patients were treated with injections of BoNT/A into the

vulvar vestibule, and three additional follow-up visits scheduled 8, 12 and 24 weeks after. Patients' clinical status and PFM myoelectrical activity were monitored in all visits according to clinical questionnaires (including VAS) and sEMG, respectively.

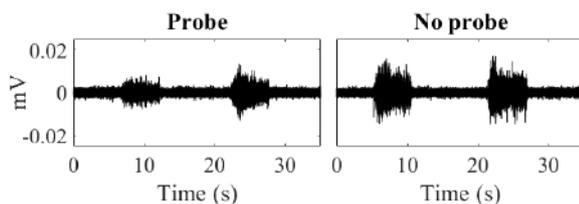
## 2.2. BoNT/A treatment

Vulvar vestibule tenderness in {1, 3, 5, 6, 7, 9, 11} clock positions was tested with a cotton swab by applying pressures of 0.2 – 0.4 kgf, which were measured with a digital algometer (Wagner FPIX™, WAGNER INSTRUMENTS, Greenwich, CT, USA). Points where the patient reported a moderate or severe pain (VAS>3) with pressure were infiltrated with 25-33 I.U. of incobotulinumtoxinA (Xeomin®, Merz Pharmaceuticals GmbH, Frankfurt, Germany).

## 2.3. sEMG signal acquisition

Patients' PFM electrical activity was monitored by sEMG performed with 4 self-adhesive Ag/AgCl electrodes (Red Dot 2660-5, 3M, St. Paul, MN, USA) attached to both labia majora and 2 to ischiatic spines (reference and ground), and an intravaginal probe (Periform®+, Neen Healthcare Mobilis Healthcare Group, Oldham, Lancashire, UK). Perineum skin was previously exfoliated with an abrasive gel (Nuprep 114g, Weaver and Company, Aurora, CO, USA) to reduce skin-electrode impedance and the probe was coated with lubricating gel (Kefus SL, Beniarbeig, Alicante, Spain). Self-adhesive electrodes were arranged as in [8], so that four monopolar (M1, M2, M3, M4) and two vertical bipolar (B1, B2) sEMG signals were recorded. Signals were acquired with a multipurpose amplifier (Grass 15LT+4 Grass 15A94, Grass Instruments, West Warwick, RI, USA) with a gain of 20.000 and a 3-1000 Hz band-pass bandwidth, and sampled at 10 kHz.

Patients were instructed to stay still for 1 minute and then to perform 5 PFM maximum voluntary contractions of 5s, with resting periods of 10s between them. This protocol was carried out twice after a minute of rest: 1) with the probe into the vagina and 2) with the probe out of it. Figure 1 shows two of the five contractions recorded by M1 channel when the probe was into and out of the vagina.



**Figure 1.** sEMG signal of M1 channel recorded of a patient when the probe into (left) and out of (right) the vagina.

## 2.4. sEMG signal preprocessing

Signals were filtered with a zero-lag 4<sup>th</sup> order Butterworth band-pass filter (cut-off frequencies: [30, 450] Hz) to remove undesired signal components such as motion artifacts and with a comb filter (notch frequency: 50Hz) to attenuate the power line interference in the signals. Then, the five contractions and a 10s-segment of baseline activity

before the first contraction were manually annotated in the recorded sEMG signals.

## 2.5. sEMG signal parametrization

sEMG signals power, spectral content and level of complexity/irregularity were assessed by three parameters that have proved to yield relevant information on PFM myoelectrical activity in previous studies [8], [9]:

*Root mean square (RMS)* is a temporal metric typically used to characterize signal amplitude, computed as the square root of the signal power. It is related to the number of recruited motor units during muscle activation [10].

*Median frequency (MDF)* is a spectral measure whose value depends on the conduction velocity of muscle fibers, the volume conduction effect of the tissue, among others [10]. MDF is the frequency of the signal bandwidth that divides its power spectral density (PSD) into two regions with the same total energy.

*Sample entropy (SampEn)* is a non-linear measure used to assess the complexity/irregularity of signals. It is computed as the negative natural logarithm of the conditional probability that two segments within a signal that are similar at  $m$  samples are also similar at  $m+1$  samples, according to a tolerance  $r$  and ignoring self-matches [11]. SampEn was computed as in [8], [9], i.e. on signals with zero mean and unit variance and with  $m=2$  and  $r=0.15$ .

*RMS*, *MDF* and *SampEn* were computed of contractile and relaxation segments annotated in signals. The median of the parameters in the five contractions was obtained.

## 2.6. Data analysis

Two of the four recording sessions performed per patient were selected and divided into two groups: 1) severe pain and 2) moderate/mild/no pain. The first group included recordings at Week 0 if the patient reported a severe pain (VAS>6), or alternatively the recording of one of the other visits (Week 8, 12 or 24) at which the patient reported the highest VAS, as long as it was >6. The second group included the recording of the follow-up visit at which the patient reported the lowest VAS, as long as it was ≤6. Groups comprised 24 recordings each and their mean VAS scores were  $8.5 \pm 1.1$  and  $3.7 \pm 1.8$ , respectively.

Statistically significant differences between parameter distributions when the probe was into vs. out of the vagina were assessed by a paired-sample T-test or Wilcoxon signed-rank test depending on normality of data, which was assessed by Kolmogorov-Smirnov tests (confidence level: 5%). Comparisons were independently carried out for the different signal channels.

## 3. Results

Table 1 shows the summarized values of the parameters computed from the sEMG signals acquired when the probe was into and out of the vagina, and highlights those cases when both distributions were significantly different. It can be seen that all statistically significant differences between both recording scenarios (probe into vs. out of the vagina) were obtained in monopolar rather than bipolar signals. Moreover, RMS during PFM contractions in severe pain

		CONTRACTION						RELAXATION					
		Severe pain			Moderate/mild/no pain			Severe pain			Moderate/mild/no pain		
		RMS ( $\mu$ V)	MDF (Hz)	SampEn	RMS ( $\mu$ V)	MDF (Hz)	SampEn	RMS ( $\mu$ V)	MDF (Hz)	SampEn	RMS ( $\mu$ V)	MDF (Hz)	SampEn
M1	IN	3.8 [1.8]	156 [28]	0.52[0.08]	4.0 [1.3]	152 [29]	0.51[0.08]	1.2 [0.5]	160 [50]	0.59[0.17]	1.0 [0.7]	161 [60]	0.62[0.17]
	OUT	4.1 [1.9]	146 [36]	0.50[0.09]	4.2 [1.9]	151 [33]	0.51[0.08]	1.3 [0.8]	158 [51]	0.60[0.14]	1.0 [0.8]	172 [50]	0.64[0.17]
M2	IN	5.8 [2.9]	147 [32]	0.53[0.07]	6.5 [1.7]	145 [34]	0.51[0.07]	1.7 [0.5]	136 [38]	0.52[0.09]	1.5 [1.0]	148 [39]	0.54[0.12]
	OUT	6.3 [2.2]	149[32]	0.52[0.08]	5.9 [3.1]	152 [33]	0.52[0.06]	1.7 [0.1]	139 [23]	0.54[0.08]	1.3 [0.9]	138 [49]	0.57[0.10]
M3	IN	3.8 [2.7]	123[27]	0.45[0.09]	4.3 [2.5]	128 [26]	0.45[0.05]	1.3 [0.6]	137 [38]	0.55[0.13]	1.1 [0.4]	146 [47]	0.60[0.14]
	OUT	4.4 [2.5]	125[22]	0.44[0.12]	4.4 [2.9]	127 [20]	0.45[0.04]	1.4 [0.8]	142 [56]	0.57[0.16]	1.0 [0.6]	156 [58]	0.63[0.15]
M4	IN	6.6 [2.5]	131[30]	0.50[0.09]	6.8 [3.0]	134 [27]	0.47[0.07]	1.9 [0.8]	122 [28]	0.49[0.10]	1.6 [0.8]	134 [29]	0.53[0.10]
	OUT	6.9 [2.0]	130[31]	0.47[0.07]	6.3 [4.0]	138 [23]	0.48[0.06]	1.8 [0.1]	126 [29]	0.50[0.10]	1.2 [0.9]	129 [52]	0.55[0.11]
B1	IN	3.2 [1.8]	126[24]	0.47[0.09]	3.5 [1.6]	126 [25]	0.47[0.08]	1.2 [0.5]	140 [41]	0.58[0.12]	1.1 [0.5]	154 [59]	0.63[0.15]
	OUT	3.6 [2.0]	125[22]	0.45[0.10]	3.4 [1.8]	127 [22]	0.47[0.09]	1.4 [0.7]	152 [55]	0.59[0.17]	1.0 [0.5]	161 [66]	0.66[0.17]
B2	IN	3.4 [2.2]	118[27]	0.44[0.12]	3.5 [1.7]	117 [27]	0.44[0.07]	1.2 [0.6]	141 [54]	0.60[0.17]	1.1 [0.7]	153 [60]	0.63[0.17]
	OUT	3.9 [1.9]	117[32]	0.42[0.10]	3.6 [2.0]	119 [24]	0.44[0.07]	1.3 [0.7]	148 [59]	0.59[0.18]	1.0 [0.7]	167 [78]	0.66[0.20]

**Table 1.** Median [interquartile range] of parameters computed with the probe into and out of the vagina (IN, OUT) during PFM contraction and relaxation according to pain intensity. Shaded cells: p-value (IN vs. OUT) <0.05

group was the only parameter that showed significantly different values in all monopolar channels when the probe was inserted. MDF and SampEn also showed some significant differences when the probe was into vs. out of the vagina, although they were only found in one of the four monopolar channels.

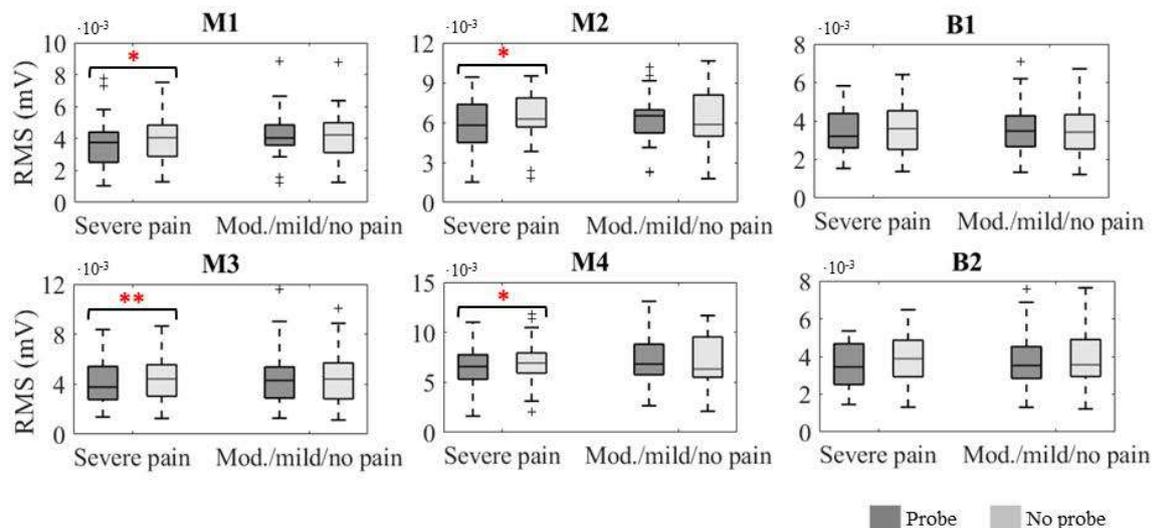
The distribution of RMS values in monopolar recordings during PFM contractions in both pain groups has been represented in Figure 2, and statistically significant differences have been highlighted with asterisks. It can be seen that statistical differences were associated with a lower p-value (<0.01) in M3 than in the other monopolar channels.

#### 4. Discussion

According to our results, PFM activity changes in the presence vs. absence of an intravaginal probe in patients with a severe pain. In these patients, sEMG monopolar signals had a lower power during PFM contractions when the probe was inserted, as shown by a lower RMS. The

value of this parameter is associated with the number of recruited motor units [10], so that it increases when a higher number of motor units are excited and develop action potentials. Therefore, the aforementioned differences suggest a lower PFM activation in patients with severe pelvic pain when a probe is inserted into their vagina.

Significant differences between both recording scenarios (probe into vs. out of the vagina) were obtained in monopolar signals, while bipolar signals were not influenced by the presence of the device. This could be related to the type of activity that is mostly detected with each configuration: unlike monopolar recordings, the contribution of far electrical fields to bipolar signals is minimum since they are almost equally detected by both electrodes and thus cancelled when their differential potential is obtained, implying that the signal is mostly originated by local myoelectrical activity. Considering that deep PFM activity has a higher total power than that of superficial PFM activity [4], external monopolar and bipolar recordings would mostly contain deep and



**Figure 2.** RMS of monopolar (M1, M2, M3, M4) and bipolar (B, B2) sEMG signals of both pain groups during PFM contraction when the probe was into (dark boxes) and out of (light boxes) the vagina. (\*): p-value<0.05. (\*\*): p-value<0.01.

superficial PFM activity, respectively, leading to the conclusion that the deep muscle bundles are the ones that show a different myoelectrical activity when a probe is inserted. This seems a reasonable outcome, since the pressure that it exerts on superficial PFM is minimum once it is in place.

Unlike PFM contractile activity, the power of resting activity was not influenced by the presence of the probe. Aunchincloss and McLean [7] assessed the PFM contractile activity of healthy women with fine wire electrodes when a probe was placed into and out of their vagina, and proved that the magnitude of any possible change induced by the device on the PFM sensory feedback, muscle length or tissue position was not high enough to translate into a significant change in the signal amplitude. This, together with the absence of significantly different RMS values in the moderate/mild/no pain group in the present study, implies that the differences observed in patients with severe pain would be associated with their reluctance to perform maximum voluntary contractions when the probe was inserted to avoid pain. Therefore, considering that PFM resting activity cannot be voluntarily controlled, the absence of significant differences during muscle relaxations would be an expected outcome. On the other hand, the intensity of the PFM activity is much lower during relaxations than during contractions, so that any possible differences between both recording scenarios during PFM relaxation could have been masked by common mode interferences and other noise sources [6].

Differences in signal power during PFM contractions were associated with a much lower p-value in M3 channel than in the other monopolar channels. M3 electrode was closer to the reference electrode than the others [8], [9], implying that crosstalk from neighboring muscles and other common mode signals were more similar between them than between the reference electrode and M1, M2 or M4, and thus more efficiently rejected in M3 signal. Therefore, differences shown by PFM activity when inserting an intravaginal probe would be greater in this signal. For this reason, other locations for the reference electrode should be considered in further studies to avoid its effect in the results obtained from the analysis. Further studies should also verify that the non-significant influence of the probe in the moderate/mild/no pain was associated with pain relief rather than other effects of BoNT/A treatment. Besides, they should make further efforts to increase the sample size to ensure that results can be reliably extrapolated to the global population.

## Conclusion

Intravaginal probes may alter deep PFM contractile activity in patients that suffer from a severe pelvic pain associated with vulvodynia. Given that self-adhesive electrodes can detect the activity of both deep and superficial PFM and they are not uncomfortable for patients, their use should be extended in clinical practice. Nonetheless, technical issues such as the reference electrode placement or rejection of common mode interferences in monopolar signals should be improved.

## Acknowledgements

This study was funded by Generalitat Valenciana in Programa para la promoción de I+D+i ACIF/2021/012, AICO/2021/126; and by private contracts from Merz Pharmaceuticals GmbH S.L.

## References

- [1] R. H. N. Nguyen, "The Prevalence and Relevance of Vulvodynia," in *Female Sexual Pain Disorders: Evaluation and Management*, A. T. Goldstein, C. F. Pukall, I. Goldstein, J. M. Krapf, S. W. Goldstein, and G. Goldstein, Eds. Oxford: John Wiley & Sons, Ltd, 2020, pp. 9–13. doi: 10.1002/9781119482598.CH2.
- [2] J. Bornstein *et al.*, "2015 ISSVD, ISSWSH, and IPPS Consensus Terminology and Classification of Persistent Vulvar Pain and Vulvodynia," *Journal of Sexual Medicine*, vol. 13, no. 4, pp. 607–612, 2016, doi: 10.1016/J.JSXM.2016.02.167.
- [3] N. O. Rosen, S. J. Dawson, M. Brooks, and S. Kellogg-Spadt, "Treatment of Vulvodynia: Pharmacological and Non-Pharmacological Approaches," *Drugs* 2019 79:5, vol. 79, no. 5, pp. 483–493, 2019, doi: 10.1007/S40265-019-01085-1.
- [4] E. Gentilcore-Saulnier, L. McLean, C. Goldfinger, C. F. Pukall, and S. Chamberlain, "Pelvic floor muscle assessment outcomes in women with and without provoked vestibulodynia and the impact of a physical therapy program," *J Sex Med*, vol. 7, no. 2 Pt 2, pp. 1003–1022, 2010, doi: 10.1111/J.1743-6109.2009.01642.X.
- [5] J. Hastings and M. Machek, "Pelvic Floor Dysfunction in Women," *Curr Phys Med Rehabil Rep*, vol. 8, no. 2, pp. 64–75, 2020, doi: 10.1007/S40141-020-00259-3.
- [6] N. Keshwani and L. McLean, "State of the art review: Intravaginal probes for recording electromyography from the pelvic floor muscles," *NeuroUrol Urodyn*, vol. 34, no. 2, pp. 104–112, 2015, doi: 10.1002/NAU.22529.
- [7] C. Aunchincloss and L. McLean, "Does the presence of a vaginal probe alter pelvic floor muscle activation in young, continent women? - ScienceDirect," *Journal of Electromyography and Kinesiology*, vol. 22, no. 6, pp. 1003–1009, 2012, doi: 10.1016/j.jelekin.2012.06.006.
- [8] M. Albaladejo-Belmonte, M. Tarazona-Motes, F. J. Nohales-Alfonso, M. De-Arriba, J. Alberola-Rubio, and J. Garcia-Casado, "Characterization of Pelvic Floor Activity in Healthy Subjects and with Chronic Pelvic Pain: Diagnostic Potential of Surface Electromyography," *Sensors* 2021, Vol. 21, Page 2225, vol. 21, no. 6, p. 2225, Mar. 2021, doi: 10.3390/S21062225.
- [9] M. Albaladejo-Belmonte, F. J. Nohales-Alfonso, M. Tarazona-Motes, M. De-Arriba, J. Alberola-Rubio, and J. Garcia-Casado, "Effect of BoNT/A in the Surface Electromyographic Characteristics of the Pelvic Floor Muscles for the Treatment of Chronic Pelvic Pain," *Sensors* 2021, Vol. 21, Page 4668, vol. 21, no. 14, p. 4668, Jul. 2021, doi: 10.3390/S21144668.
- [10] D. Farina, R. Merletti, and R. M. Enoka, "The extraction of neural strategies from the surface EMG: An update," *J Appl Physiol*, vol. 117, no. 11, pp. 1215–1230, Dec. 2014, doi: 10.1152/JAPPLPHYSIOL.00162.2014/ASSET/IMAGES/LARGE/ZDG0221412300009.JPEG.
- [11] J. S. Riehman and J. Randall Moorman, "Physiological time-series analysis using approximate and sample entropy," *Am J Physiol Heart Circ Physiol*, vol. 278, pp. H2039–H2049, 2000, doi: 10.1152/AJPHEART.2000.278.6.H2039.

# Cuantificación de la Heterogeneidad del Sustrato Electrofisiológico Cardíaco en Registros Obtenidos mediante Multielectrodos de Alta Densidad

L. Pancorbo<sup>1,2</sup>, S. Ruipérez-Campillo<sup>1,3,4</sup>, A. Guill<sup>1</sup>, A. Tormos, J. Chorro<sup>5,6</sup>, F. Castells<sup>1</sup>, J. Millet<sup>1</sup>

<sup>1</sup> Instituto ITACA, Universitat Politècnica de València, Valencia, España {saruicam, fcastells, jmillet}@upv.es

<sup>2</sup> Escuela Politécnica Superior, Universidad Carlos III de Madrid, Madrid, España

<sup>3</sup> Escuela de Medicina, Universidad de Stanford, Palo Alto, California, Estados Unidos

<sup>4</sup> Departamento de Ing. de Teleco. e Ing. Electrónica (D-ITET), Instituto Federal Suizo en Zúrich (ETH), Zúrich, Suiza

<sup>5</sup> D. de Medicina, Servicio de Cardiología, Hospital Clínico Universitario, Universitat de València, Valencia, España

<sup>6</sup> Centro de Investigación Biomédica en Red de Enfermedades Cardiovasculares (CIBERCV), Madrid, España

## Resumen

*En este estudio se propone una métrica para medir la heterogeneidad del sustrato cardíaco a partir de mapas vectoriales derivados de electrogramas omnipolares. Dicho parámetro cuantifica el nivel de desorganización de la propagación eléctrica con el fin de clasificar de forma instantánea el tejido cardíaco en contacto con el catéter. Con este fin, se calculan mapas vectoriales de registros obtenidos mediante modelo experimental, con y sin estimulación eléctrica, bajo la hipótesis de que los primeros son más heterogéneos. Los resultados muestran la capacidad de la métrica de discriminar entre ambos casos, dando valores significativamente mayores cuando no se aplica estimulación ( $p < 0.001$ ). Por tanto, el parámetro propuesto permite medir la heterogeneidad en la propagación de onda, disponiendo de un gran potencial para identificar tejido fibrótico.*

## 1. Introducción

Caracterizar el sustrato cardíaco es importante para el diagnóstico y tratamiento de las arritmias cardíacas, dada su vinculación con la presencia de tejido fibrótico [1]. Los electrogramas (EGMs), adquiridos mediante catéteres, permiten estudiar propiedades como los tiempos de activación, direcciones y velocidades de conducción de la señal electrofisiológica.

Existe una diversidad importante de catéteres de uso clínico que varían fundamentalmente en la distribución y el número de electrodos. Recientemente se ha introducido el uso clínico del catéter *Advisor HD Grid* que consiste en 16 electrodos dispuestos en una formación similar a una matriz equidistante 4x4 [2]. Este tipo de catéter de alta densidad de electrodos no solo ofrece mayor resolución, sino que permite registrar EGMs omnipolares.

Los EGM omnipolares se han propuesto en la última década por su capacidad de generar señales independientes a la orientación del catéter [3]. Además, la tecnología omnipolar es más robusta que los EGMs unipolares, dado que estos últimos se ven afectados por el ruido de campos lejanos. Igualmente, es más robusta frente a los EGMs bipolares, que son dependientes de la orientación del

catéter [4] – dada la geometría omnipolar que considera la contribución de dos direcciones perpendiculares de propagación. Una configuración de electrodos en *grid* permite implementar esta nueva técnica estimando los omnipolos a partir de los denominados *cliques*. Un *clique* se define como cuatro electrodos cercanos, a partir de los que se deriva una señal o EGM omnipolar [5].

Hasta el momento dos tipos de *cliques* han sido propuestos en la literatura: el triangular y el cuadrado. El primero utiliza tres electrodos en disposición triangular para obtener un par de bipolos ortogonales de los que se deriva el omnipolo, mientras que el segundo usa los cuatro electrodos del *clique* y calcula los dos bipolos haciendo una media entre las componentes paralelas.

Los EGMs unipolares y bipolares requieren analizar los tiempos locales de activación para estudiar la propagación. Sin embargo, la tecnología omnipolar es capaz de determinar las direcciones de propagación de forma instantánea sin necesidad de hacer primero un barrido de toda la cavidad cardíaca [6]. Además, con esta tecnología es posible visualizar patrones de propagación complejos, comunes en tejido arritmogénico [7].

Por tanto, dada la barrera estructural que presenta el tejido fibrótico en la propagación de onda [8], planteamos la hipótesis de que el sustrato cardíaco se puede caracterizar de acuerdo al nivel de heterogeneidad de los mapas vectoriales derivados de EGMs omnipolares.

Con este objetivo, diseñamos una métrica capaz de medir la heterogeneidad y por tanto de clasificar el tejido en contacto con el catéter de forma inmediata. Esta métrica se basa en el cálculo y comparación de los ángulos de propagación en el mapa vectorial, y se detalla en la sección de *métodos*. En este estudio, aplicamos la métrica a registros basales y estimulados, bajo la hipótesis de que la heterogeneidad es menor en los últimos ya que la estimulación reconduce la propagación del impulso eléctrico haciendo que los vectores estén más alineados.

## 2. Materiales

Las señales utilizadas provienen de un estudio de investigación previamente publicado [9] de una serie experimental de corazones de conejo aislados y perfundidos mediante sistema de Langendorff. Para adquirir las señales, se colocó un catéter *ad hoc* de mapeo multielectrodo (128 canales,  $d:1\text{mm}$  interelectrodo) en contacto con el epicardio del ventrículo izquierdo, obteniendo varios registros con el sistema MapTech® atendiendo al protocolo establecido. Para el presente estudio se han seleccionado registros basales y estimulados de 5 corazones a 37°C. La estimulación se llevó a cabo mediante el sistema GRASS S88 y un electrodo bipolar. En concreto, se seleccionaron 3 registros basales, 3 estimulados a 4 Hz y otros 3 a 6 Hz en distintos instantes por cada corazón.

## 3. Métodos

### 3.1. Creación de los Mapas Vectoriales

Con el objetivo de crear mapas vectoriales, se utilizan los registros adquiridos por la matriz central 4x4 de electrodos para simular la adquisición de un *Catéter Advisor HD Grid*.

Los bipolos ortogonales ( $\mathbf{b}_x, \mathbf{b}_y$ ) de los que se deriva la señal omnipolar provienen de una configuración en cruz de los *cliques*. La dirección de propagación estimada se fija en el centro del *click*, dando como resultado un mapa de 3x3 vectores.

Tal y como describe Deno *et al.* [5], el campo eléctrico local derivado de cada par de bipolos  $\mathbf{b}_x, \mathbf{b}_y$  traza un bucle durante la despolarización. La máxima amplitud del bucle corresponde con la dirección de propagación, formando un ángulo  $\theta$  con respecto al eje horizontal compuesto por  $\mathbf{b}_x$  (véase figura 1). La representación de señales en forma de bucle (VCG) y su análisis posterior es común en la literatura y permite diseñar distintos parámetros para su caracterización [9, 10].

En este caso, generalizando para un instante fijo en el tiempo en una matriz de electrodos  $G(t) \in \mathbb{R}^{m \times n}$  ( $\forall t = c$ ), se define una matriz  $A$  con los ángulos de propagación de cada *click*:

$$A \in \mathbb{R}^{p \times q}: p = m - 1, q = n - 1 \quad [\text{Ec. 1}]$$

$$\theta_{i,j}: 1 \leq i \leq p, 1 \leq j \leq q \quad [\text{Ec. 2}]$$

$$A := \begin{bmatrix} \theta_{1,1} & \cdots & \theta_{1,q} \\ \vdots & \ddots & \vdots \\ \theta_{p,1} & \cdots & \theta_{p,q} \end{bmatrix} \quad [\text{Ec. 3}]$$

Donde  $m$  y  $n$  son el número de filas y columnas de la matriz  $G$ ,  $p$  y  $q$  son el número de filas y columnas de la matriz  $A$  e  $i$  y  $j$  son las posiciones de un elemento en la matriz de ángulos (ilustrado en la figura 2).

La propagación se representa en un mapa vectorial con vectores unitarios  $\vec{u}_{i,j} = (\cos \theta_{i,j}, \sin \theta_{i,j})$ .

### 3.2. Derivación del Parámetro de Heterogeneidad

Para cada vector del mapa  $\vec{u}_{i,j}$ , se calcula la variación angular con cada uno de sus vectores contiguos ( $\alpha_{n_{i,j}}$ )

usando la definición geométrica del producto escalar. Para simplificar, a la variación angular entre dos vectores  $\Delta\theta$  se le designará como  $\alpha$  (véase una ilustración de las variaciones angulares calculadas en la figura 2).

$$\alpha_{n_{i,j}} = \cos^{-1}(\vec{u}_{i,j} \cdot \vec{u}_{i+c_1, j+c_2}): 1 \leq i + c_1 \leq p, 1 \leq j + c_2 \leq q$$

$$n = \{1, \dots, 9\}, c_1 = \{-1, 0, 1\}, c_2 = \{-1, 0, 1\} \quad [\text{Ec. 4}]$$

Donde  $c_1$  y  $c_2$  son constantes que se suman a  $i$  y  $j$  respectivamente para designar todos los elementos contiguos al elemento  $\theta_{i,j}$ , incluyéndose a sí mismo.

Dado que cada variación angular  $\alpha_{n_{i,j}}$  está comprendida en el intervalo cerrado  $[0, \pi]$ , el sumatorio de todas las variaciones para un elemento tendrá un límite superior  $\xi\pi$ . Sin embargo, teniendo en cuenta que una de las variaciones ( $\alpha_{5_{i,j}}$  de acuerdo a la figura 2) es del elemento consigo mismo, el límite superior será  $(\xi - 1)\pi$ . Esto es:

$$0 \leq \alpha_{n_{i,j}} \leq \pi \Rightarrow 0 \leq \sum_{n=1}^9 \alpha_{n_{i,j}} \leq \xi\pi \quad [\text{Ec. 5}]$$

$$\alpha_{5_{i,j}} = 0, \quad \forall i, j \Rightarrow 0 \leq \sum_{n=1}^9 \alpha_{n_{i,j}} \leq (\xi - 1)\pi \quad [\text{Ec. 6}]$$

Por tanto, el coeficiente de heterogeneidad para ese elemento  $\theta_{i,j}$  es:

$$v_{i,j} = \frac{\sum_{n=1}^9 \alpha_{n_{i,j}}}{(\xi - 1)\pi} : 0 \leq v_{i,j} \leq 1 \quad [\text{Ec. 7}]$$

El parámetro  $\xi$  se define según se trate de un elemento central, de una esquina, o de un lado (superior o inferior) de la matriz de electrodos, de modo que se puede representar como:

$$\xi = \begin{cases} 9, & \text{si } (1 < i < p) \text{ y } (1 < j < q) \\ 4, & \text{si } i = \{1, p\} \text{ y } j = \{1, q\} \\ 6, & \text{si } [i = \{1, p\} \text{ y } (1 < j < q)] \text{ o } [(1 < i < p) \text{ y } j = \{1, q\}] \end{cases} \quad [\text{Ec. 8}]$$

Donde la primera fila de la parte derecha de la ecuación se corresponde con los elementos centrales y la segunda y tercera fila se corresponde con los lados (o esquinas).

Por último, se obtiene el valor global de heterogeneidad  $H$ :

$$H = \frac{\sum_{i=1}^p \sum_{j=1}^q v_{i,j}}{p \times q} : 0 \leq H < 1 \quad [\text{Ec. 9}]$$

### 3.3. Análisis Estadístico

Con el objetivo de comprobar la capacidad del parámetro para distinguir entre poblaciones según la estimulación, se ha realizado un test estadístico de la forma *t-test pareado de cola izquierda* con un nivel de significación del 1% sobre las distribuciones de valores de heterogeneidad  $H$ , (véase la figura 4). Nótese que esta prueba asume que las poblaciones siguen una distribución normal y que las muestras son independientes.

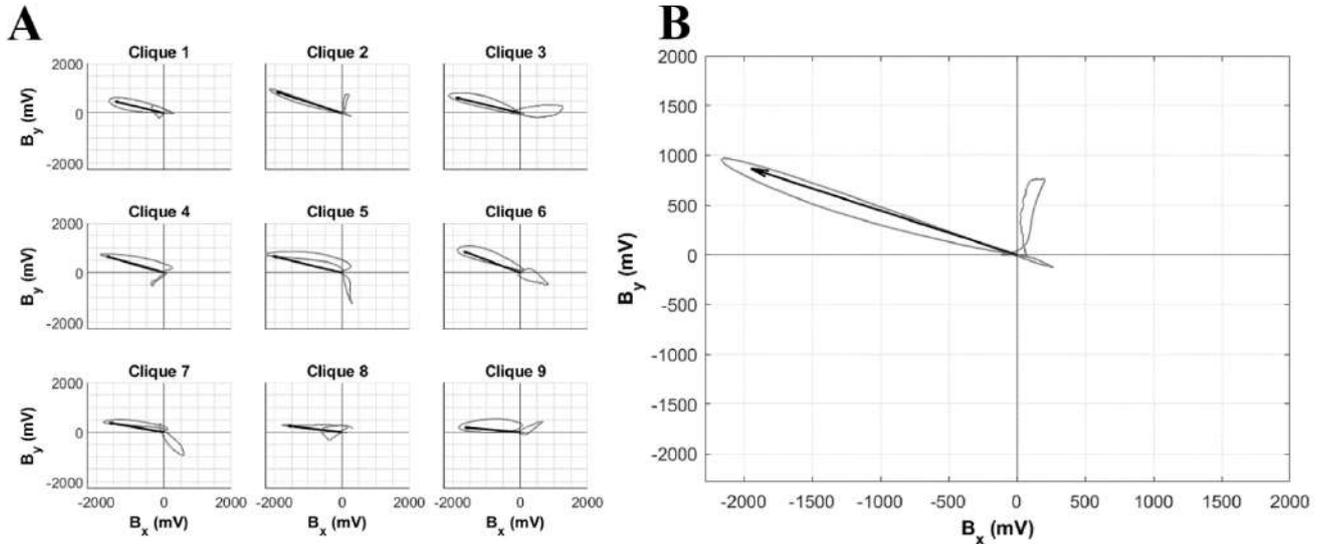


Figura 1. A. Bucle trazado en cada clique por los bipolos correspondientes  $b_x$ - $b_y$ . B. Ampliación al bucle trazado por los bipolos ortogonales del Clique 2. El vector indica la dirección de propagación estimada para ese clique.

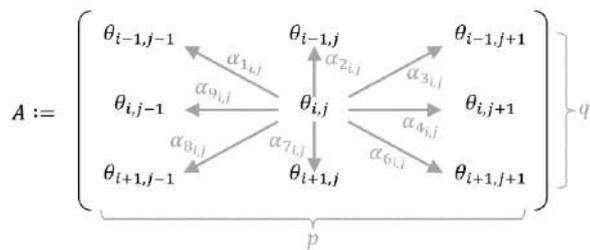


Figura 2. Variaciones angulares para un elemento central  $\theta_{i,j}$ .

#### 4. Resultados

En la figura 3 se muestran ejemplos de los mapas de heterogeneidad resultantes a 37°C para casos en los que se aplica estimulación (Figura 3.a) y casos en los que no hay estimulación (Figura 3.b).

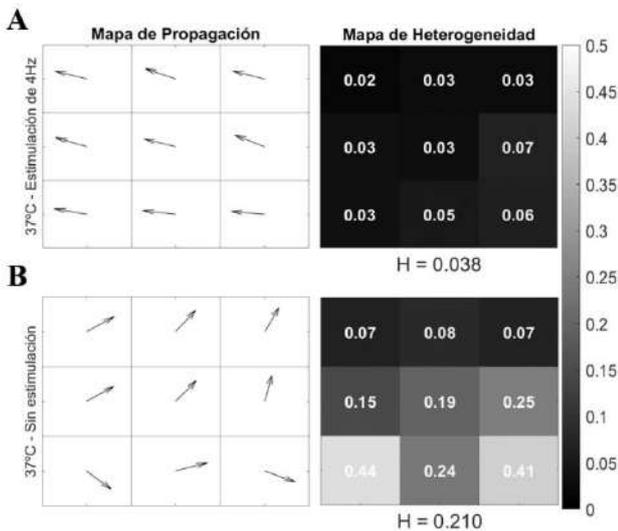


Figura 3. Ejemplos de mapas de heterogeneidad mostrando los coeficientes  $v_{i,j}$  de cada clique y el valor global de heterogeneidad  $H$ . A. Resultados para un registro con estimulación a 4 Hz tomado a 37°C. A la izquierda, mapa vectorial de propagación y a la derecha, mapa de heterogeneidad. B. Ídem para un registro sin estimulación.

Los valores de heterogeneidad se agrupan de acuerdo al tipo de estimulación (Sin, 4 y 6 Hz) obtenidos para cada experimento (tabla 1).

Las distribuciones de todos los valores obtenidos se muestran en la figura 4. Se da un valor de  $H$  cercano a 0.20 cuando no hay estimulación, frente a valores en torno a 0.06 con estimulaciones a 4 y 6 Hz respectivamente.

En general, se observan resultados significativos al comparar sin estimulación vs estimulados ( $p < 0.001$ ). Por otra parte, no se observan diferencias significativas entre los grupos estimulados a 4 y 6 Hz ( $p = 0.985$ ).

Experimento	Sin estimulación	Estimulación 4 Hz	Estimulación 6 Hz
Experimento 1	0.176±0.080	0.064±0.006	0.053±0.003
Experimento 2	0.283±0.175	0.044±0.011	0.049±0.012
Experimento 3	0.272±0.126	0.056±0.020	0.067±0.022
Experimento 4	0.151±0.051	0.075±0.012	0.093±0.014
Experimento 5	0.125±0.021	0.045±0.033	0.058±0.022

Tabla 1. Valor de heterogeneidad ( $H$ ) para cada experimento según el tipo de estimulación

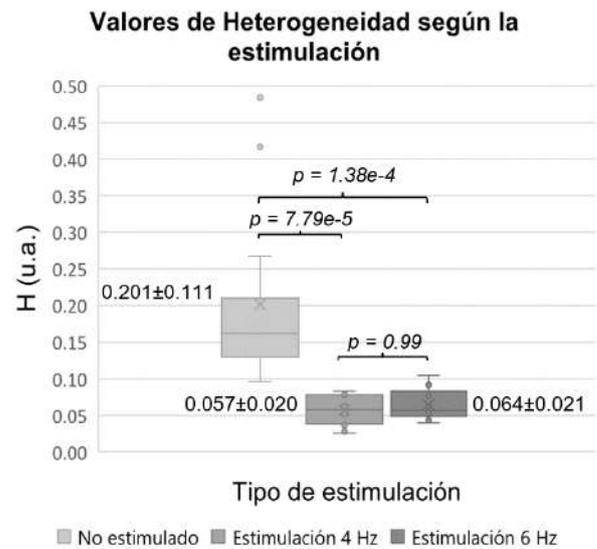
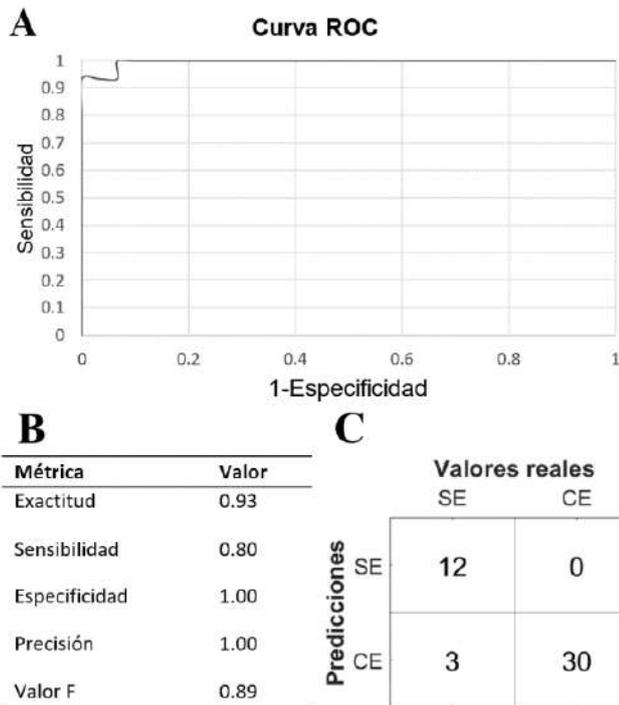


Figura 4. Representación de Cajas con bigotes de los valores de heterogeneidad según el tipo de estimulación.

Con estos valores se calcula una curva ROC (Figura 5A) con un área bajo la curva (AUC ROC) de 0.996 y un valor

de umbral óptimo de 0.113. Al aplicar este umbral se genera la matriz de confusión ilustrada en la figura 5C, cuyas métricas se recogen en la tabla 5B.



**Figura 5.** A. Curva ROC de H. B. Tabla con las métricas derivadas de la matriz de confusión para un umbral de 0.113. C. Matriz de confusión. CE: Con estimulación y SE: Sin estimulación.

## 5. Discusión

Los mapas vectoriales de propagación obtenidos de tejidos arritmogénicos presentan un alto nivel de desorganización. Este estudio propone una métrica para cuantificar dicha heterogeneidad y poder discriminar el tejido bajo el catéter como sano o fibrótico.

El parámetro de heterogeneidad se comporta de acuerdo con la hipótesis inicial, siendo capaz de diferenciar entre los grupos sin vs con estimulación. Ya que los mapas vectoriales con mayor desorganización (en los que no se aplicó estimulación) presentan valores de heterogeneidad significativamente más altos frente a los mapas con vectores más alineados (cuando sí se aplicó estimulación). Los p-valores del análisis estadístico así lo constatan.

Esta capacidad discriminadora del parámetro se refleja en el valor de AUC ROC y en las métricas de la matriz de confusión obtenidas para el umbral óptimo. Destacando que todos los casos con estimulación tanto a 4 como a 6 Hz son identificados correctamente.

Por otra parte, se constata que, al aumentar la frecuencia de estimulación, no hay diferencia significativa en la heterogeneidad (en ambos casos están alineados). Esto no contradice la hipótesis, sino que es un aspecto más que resaltar de los resultados.

## 6. Limitaciones y Trabajo Futuro

Se debe tener en cuenta que el mapa vectorial que utilizamos no es muy denso, permitiéndonos analizar nueve electrodos por cada activación, al emular la

distribución del catéter *Advisor HD Grid*. Aun así, qué duda cabe que los resultados obtenidos son prometedores.

Por ello, se espera ampliar el estudio usando mapas de mayor densidad. También sería deseable disponer de registros de especies más grandes y en la medida de lo posible señales intracavitarias.

## 7. Conclusión

En este estudio se define matemáticamente un parámetro capaz de medir la heterogeneidad en la propagación de onda. Los resultados muestran la capacidad discriminadora del parámetro, asignando valores más elevados a mapas desorganizados y valores más bajos a mapas homogéneos, postulándose como un biomarcador prometedor para identificar tejido fibrótico.

## Agradecimientos

Este proyecto ha sido financiado parcialmente por el Ministerio de Ciencia e Innovación del Gobierno de España (PID2019-109547RB), el Instituto de Salud Carlos III (CIBERCV CB16/11/00486) y la Generalitat Valenciana (PROMETEOII/2014/037).

## Referencias

- [1] M.-N. Nguyen, H. Kiriazis, X.-M. Gao, y X.-J. Du, «Cardiac Fibrosis and Arrhythmogenesis», *Compr. Physiol.*, vol. 7, n.º 3, pp. 1009-1049, 2017, doi: <https://doi.org/10.1002/cphy.c160046>.
- [2] K. Magtibay, A. Porta-Sánchez, S. K. Haldar, D. C. Deno, S. Massé, y K. Nanthakumar, «Reinserting Physiology into Cardiac Mapping Using Omnipolar Electrograms», *Card. Electrophysiol. Clin.*, vol. 11, n.º 3, pp. 525-536, 2019, doi: [10.1016/j.ccep.2019.05.003](https://doi.org/10.1016/j.ccep.2019.05.003).
- [3] J. Riccio *et al.*, «Characterization of Atrial Propagation Patterns and Fibrotic Substrate With a Modified Omnipolar Electrogram Strategy in Multi-Electrode Arrays», *Front. Physiol.*, vol. 12, p. 674223, sep. 2021, doi: [10.3389/fphys.2021.674223](https://doi.org/10.3389/fphys.2021.674223).
- [4] N. M. S. de Groot *et al.*, «Critical appraisal of technologies to assess electrical activity during atrial fibrillation: a position paper from the European Heart Rhythm Association and European Society of Cardiology Working Group on eCardiology in collaboration with the Heart Rhythm Society, Asia Pacific Heart Rhythm Society, Latin American Heart Rhythm Society and Computing in Cardiology», *EP Eur.*, vol. 24, n.º 2, pp. 313-330, 2022, doi: [10.1093/europace/euab254](https://doi.org/10.1093/europace/euab254).
- [5] D. C. Deno, R. Balachandran, D. Morgan, F. Ahmad, S. Massé, y K. Nanthakumar, «Orientation-Independent Catheter-Based Characterization of Myocardial Activation», *IEEE Trans. Biomed. Eng.*, vol. 64, n.º 5, pp. 1067-1077, 2017, doi: [10.1109/TBME.2016.2589158](https://doi.org/10.1109/TBME.2016.2589158).
- [6] D. C. Deno *et al.*, «High-resolution, live, directional mapping», *Heart Rhythm*, vol. 17, n.º 9, pp. 1621-1628, 2020, doi: [10.1016/j.hrthm.2020.04.039](https://doi.org/10.1016/j.hrthm.2020.04.039).
- [7] S. K. Haldar *et al.*, «Resolving Bipolar Electrogram Voltages During Atrial Fibrillation Using Omnipolar Mapping», *Circ. Arrhythm. Electrophysiol.*, vol. 10, n.º 9, p. e005018, 2017, doi: [10.1161/CIRCEP.117.005018](https://doi.org/10.1161/CIRCEP.117.005018).
- [8] L. J. van der Does y N. M. de Groot, «Inhomogeneity and complexity in defining fractionated electrograms», *Heart Rhythm*, vol. 14, n.º 4, pp. 616-624, 2017.
- [9] A. Guill *et al.*, «QT Interval Heterogeneities Induced Through Local Epicardial Warming/Cooling. An Experimental Study», *Rev. Esp. Cardiol. Engl. Ed.*, vol. 67, n.º 12, pp. 993-998, 2014, doi: [10.1016/j.rec.2014.02.026](https://doi.org/10.1016/j.rec.2014.02.026).
- [10] S. Ruipérez-Campillo *et al.*, «Non-invasive characterisation of macroreentrant atrial tachycardia types from a vectorcardiographic approach with the slow conduction region as a cornerstone», *Comput. Methods Programs Biomed.*, vol. 200, p. 105932, mar. 2021, doi: [10.1016/j.cmpb.2021.105932](https://doi.org/10.1016/j.cmpb.2021.105932).
- [11] S. Ruipérez-Campillo *et al.*, «Parametrización del Vectorcardiograma Auricular Para la Caracterización de Distintos Tipos De Flutter», 2020, pp. 393-396.

# Informática Biomédica

---

Miércoles, 23 de noviembre de 2022

# Segmentación de la capnografía durante la resucitación cardiopulmonar mediante el uso de redes neuronales

A. Elola<sup>1</sup>, E. Aramendi<sup>2</sup>, X. Jaureguibeitia<sup>2</sup>

<sup>1</sup> Departamento de Tecnología Electrónica, Universidad del País Vasco (UPV/EHU), Eibar, España, andoni.elola@ehu.eus

<sup>2</sup> Departamento de Ingeniería de Comunicaciones, Universidad del País Vasco (UPV/EHU), Bilbao, España

## Resumen

*La parada cardiorrespiratoria extrahospitalaria sigue siendo una de las causas principales de mortalidad en los países industrializados. Con el fin de ayudar en la terapia, las guías de resucitación actuales recomiendan el uso de la señal de capnografía, pero las compresiones torácicas aplicadas al paciente inducen un artefacto que dificulta su correcta segmentación automática. En este estudio se propone un algoritmo para segmentar la capnografía. Se han utilizado 251 episodios de paradas cardiorrespiratorias extrahospitalarias y se entrena una red neuronal profunda basada en la arquitectura encoder-decoder. El algoritmo obtuvo una media de F<sub>1</sub> de 92.7% para segmentar las fases de inspiración (clase minoritaria), lo cual está más de 5 puntos por encima del estado del arte.*

## 1. Introducción

La parada cardiorrespiratoria se define como el cese repentino de la respiración pulmonar y circulación sanguínea, y es una de las causas principales de mortalidad en los países industrializados. La resucitación cardiopulmonar (RCP) junto a la desfibrilación eléctrica constituyen la terapia básica de resucitación. La resucitación cardiopulmonar incluye la aplicación de compresiones torácicas y ventilaciones, las cuales deben ser aplicadas con ciertas características para que sean efectivas [1]. Monitorizar dichas características permite que el rescatador se adhiera mejor a lo que indican las guías de resucitación, aumentando así la probabilidad de supervivencia del paciente.

Actualmente una de las señales más utilizadas para monitorizar la frecuencia de ventilaciones es la capnografía, el cual representa la presión parcial del dióxido de carbono exhalado por el paciente. Además, las guías de resucitación actuales recomiendan el uso de la capnografía para confirmar la intubación correcta, monitorizar la calidad de las compresiones o para confirmar que el paciente ha recuperado la circulación espontánea. Sin embargo, las compresiones torácicas inducen un artefacto en la señal de capnografía que dificulta su correcta interpretación [2].

En el estado del arte se han propuesto algunos algoritmos para la detección de ventilaciones utilizando la capnografía en presencia de artefactos generados por la RCP [3], [4]. Sin embargo, dichos algoritmos no se evaluaron para segmentar la capnografía completamente, teniendo así un uso limitado.

El objetivo de este estudio es desarrollar un algoritmo para segmentar la capnografía utilizando redes neuronales.

## 2. Materiales

Los datos utilizados para este estudio provienen de una amplia base de datos de paradas cardíacas extrahospitalarias recogidas en Dallas, EEUU. Todos los episodios fueron recogidos durante 2016 con el monitor-desfibrilador Philips HeartStart MRx, el cual adquiere las señales de capnografía e impedancia torácica con una frecuencia de muestreo de 40/125 Hz y 200 Hz, respectivamente.

Un experto en ingeniería biomédica anotó las fases de inspiración en las capnografías de 251 pacientes basándose en la capnografía y en la impedancia torácica. Algunos intervalos se anotaron como ruidosos debido principalmente a la discordancia entre ambas señales, a la ausencia de alguna de ellas o al nivel de ruido presente.

En la Figura 1 se muestran algunos segmentos de la señal de capnografía. En el panel (a) se observa un ejemplo sin artefacto de RCP, mientras que los demás paneles muestran ejemplos con cierto nivel de artefacto. Se han marcado los intervalos de la fase de insuflado que se utilizarán como referencia.

## 3. Métodos

### 3.1. Preprocesado

La capnografía se remuestrea a 20 Hz para así reducir el coste computacional, y se segmenta en ventanas de 60 s descartando aquellos intervalos ruidosos. Estas ventanas de 60 s serán las entradas de la red neuronal.

### 3.2. Segmentación

Para realizar la segmentación, se utiliza una red neuronal profunda basada en la arquitectura de *encoder-decoder* que se muestra en la Figura 2. La entrada de la red es un segmento 1200 muestras, y la salida es un segmento de la misma longitud con valor de “0” durante las fases de inspiración anotados y valor de “1” para el resto de las muestras. Se puede apreciar que la red se basa principalmente en dos tipos de bloques, bloques del encoder (BloqueE) y bloques de decoder (BloqueD). Las arquitecturas de cada una de ellas se muestran en los paneles (b) y (c) de la figura, respectivamente. El número

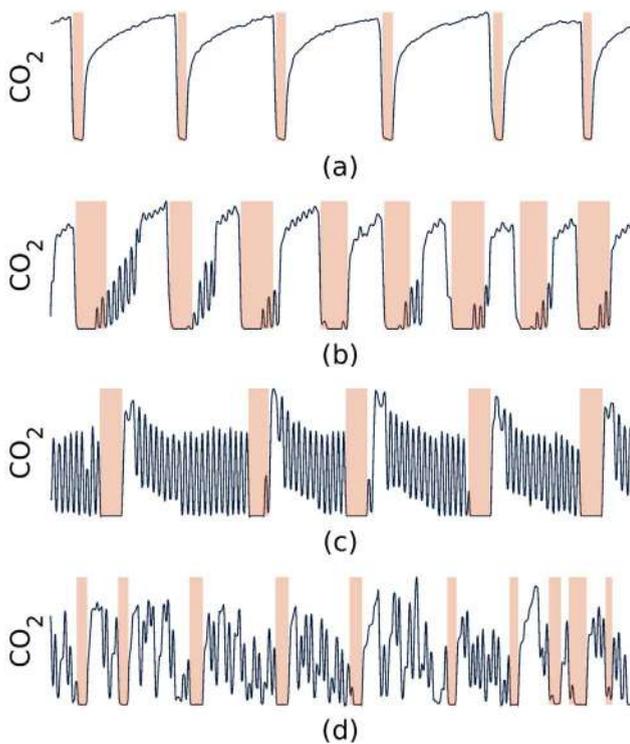


Figura 1. Ejemplos de segmentos de capnografía. Las fases de inspiración anotadas por el experto se muestran sombreadas

que se indica entre paréntesis para BloqueE y BloqueD muestra la cantidad de filtros que se utilizan en cada paso, mientras que entre bloques se muestran las dimensiones del tensor (cantidad de muestras x cantidad de canales).

A continuación se listan algunos detalles de implementación:

- En todos los casos se han utilizado convoluciones separables (SepConv) para reducir el número de parámetros en la fase de entrenamiento, y el orden escogido ha sido de 5. Para la primera capa convolucional (el de la entrada), se ha utilizado un stride de 2 para poder reducir la dimensión de la entrada a la mitad.
- En la figura BN significa "Batch Normalization", la capa "ReLU" aplica la función de rectificador lineal y "Sigmoid" aplica la función de sigmoide.
- Para todos los Dropout, la probabilidad establecida ha sido de 0.1.
- Para los bloques de reducción de dimensión (MaxPooling), se ha escogido el máximo entre 3 muestras con un stride de 2.
- Los tensores de entrada doblan la dimensión para todos los casos que se quiera sobremuestrear la señal (capas UpSampling).

Gracias a la función de sigmoide, se obtiene la secuencia  $p(n) \in (0,1)$  para cada segmento, que se redondea para obtener  $p(n) \in \{0,1\}$ . La red se ha entrenado utilizando el algoritmo Adam con los parámetros estándar y se ha adaptado el coeficiente de Dice como función de coste.

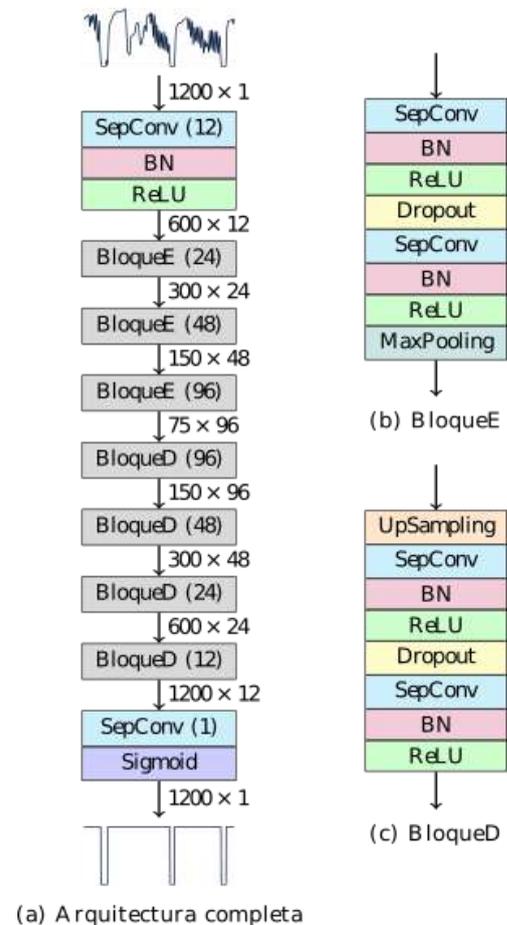


Figura 2. Arquitectura de la red neuronal profunda utilizada para segmentar la señal de capnografía

### 3.3. Evaluación

Para entrenar y evaluar el algoritmo, se ha utilizado una validación cruzada de 10 particiones. Todos los segmentos pertenecientes al mismo paciente se han asignado a la misma partición.

Las métricas de desempeño se han dividido en dos grupos. Por un lado, se ha definido una métrica para evaluar el desempeño de la red a la hora de segmentar la señal. Dicha métrica es el  $F_1$  para la detección de las fases de inspiración para cada segmento,  $F_{1,i}$ . Por otro lado, como caso de uso, se han medido la sensibilidad (Se), valor predictivo positivo (VPP) y el  $F_1$  para la detección de ventilaciones,  $F_{1,v}$ . El instante de ventilación se ha considerado al final de la fase de inspiración.

## 4. Resultados

En total se analizaron 3851 ventanas de 60 s provenientes de 251 pacientes, con  $15.3 \pm 10.4$  ventanas por paciente.

En la Tabla 1 se recogen las medias de todas las métricas de desempeño. El algoritmo actual se compara con un algoritmo propuesto en el estado del arte, y tal y como se puede apreciar, el algoritmo propuesto en este estudio muestra unos resultados significativamente mejores. Cabe mencionar que el propósito original del algoritmo propuesto en [3] era la detección de ventilaciones durante la RCP.

	$F_{1,i}$	Se	VPP	$F_{1,v}$
Aramendi et al. [3]	87.2%	88.3%	92.9%	92.3%
Este estudio	92.7%	96.8%	93.6%	94.7%

Tabla 1. Resumen de las métricas de desempeño obtenidas para toda la base de datos

En la Figura 3 se muestran diferentes rangos de las métricas de desempeño y la proporción de ventanas dentro de dicho rango. Se puede apreciar que para la mayoría de las ventanas todas las métricas de desempeño superan el 90%.

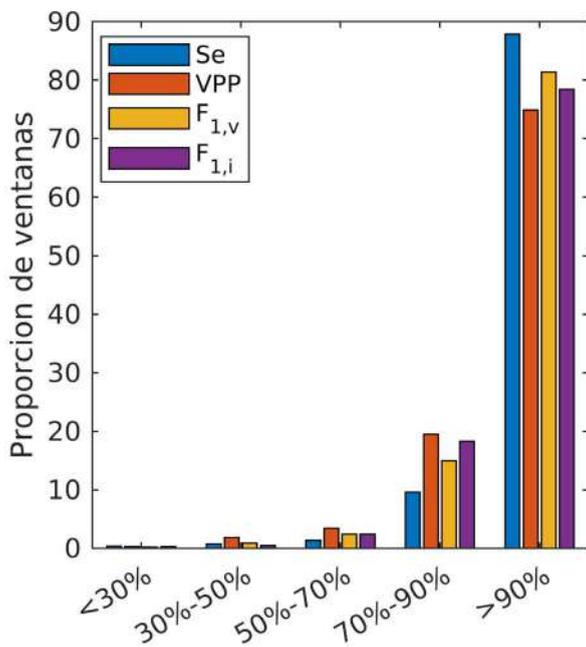


Figura 3. Proporción de ventanas para diferentes rangos de todas las métricas de desempeño

## 5. Discusión

En este estudio se propone el uso de las redes neuronales profundas para segmentar la capnografía durante la RCP. El algoritmo propuesto es capaz de segmentar las fases de inspiración con un  $F_1$  del 92.7%.

La segmentación de la capnografía puede tener varias aplicaciones en un escenario de parada cardiorrespiratoria extrahospitalaria. Uno de los usos más extendidos es la monitorización de la frecuencia de ventilaciones, para evitar así la hiperventilación o la hipoventilación del paciente. El algoritmo propuesto en este estudio es capaz de detectar las ventilaciones con un  $F_1$  de 94.7%. La Figura 4 muestra el error generado en el cómputo de la frecuencia de ventilaciones. Mediante las líneas horizontales se muestra el intervalo de confianza del 95%.

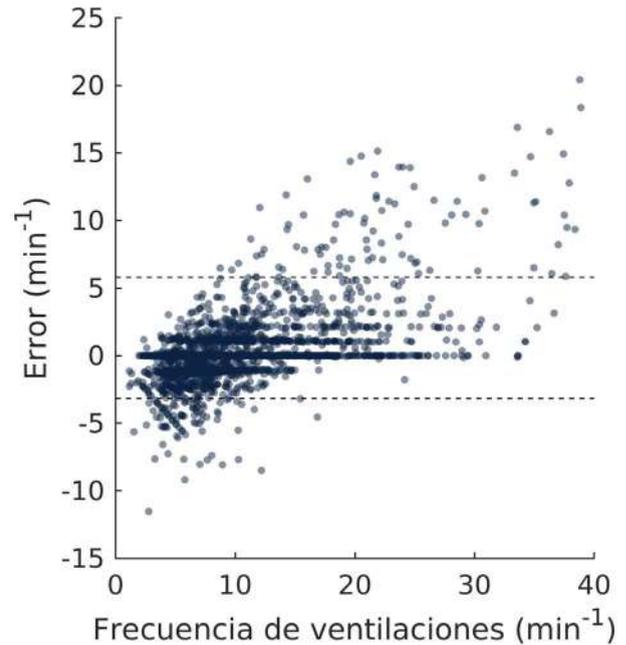


Figura 4. Error obtenido en el cómputo de la frecuencia de ventilaciones. Las líneas horizontales muestran el intervalo de confianza del 95%

Los algoritmos del estado del arte para la detección de ventilaciones se han evaluado utilizando episodios completos de resucitación [3], [4]. Sin embargo, la duración escogida para segmentar la capnografía en este estudio ha sido de 60 s, alineado con el intervalo utilizado habitualmente para proporcionar feedback sobre la frecuencia de ventilaciones durante la RCP. El segmentador propuesto puede además utilizarse para extraer características de las ventilaciones, que pueden utilizarse en estudio retrospectivos de análisis de la terapia de ventilación, con diferentes técnicas, y su relación con la supervivencia [5], [6].

Para realizar este estudio, se han utilizado datos provenientes de un solo centro y un solo monitor-desfibrilador. En futuros trabajos, interesa validar los resultados con bases de diferente origen.

## 6. Conclusiones

El presente estudio propone una red neuronal profunda para segmentar las fases de inspiración en la capnografía. Este segmentador puede ser útil para extraer características de la capnografía o monitorizar la frecuencia de ventilaciones. De este modo, se asiste al rescatador ayudando a adherirse a las guías de resucitación, aumentando así la probabilidad de supervivencia del paciente.

## Agradecimientos

Este trabajo ha sido parcialmente financiado por el Ministerio de Ciencia, Innovación y Universidades a través del proyecto RTI2018-101475-BI00, en conjunto con el Fondo Europeo de Desarrollo Regional (FEDER), por el Gobierno Vasco por medio del proyecto IT1717-22 y por la Universidad del País Vasco (UPV/EHU) a través del proyecto COLAB20/01.

## Referencias

- [1] G. D. Perkins *et al.*, «European Resuscitation Council Guidelines for Resuscitation 2015: Section 2. Adult basic life support and automated external defibrillation», *Resuscitation*, vol. 95, pp. 81-99, 2015.
- [2] «Abstract 83: High Incidence of Chest Compression Oscillations Associated With Capnography During Out-of-Hospital Cardiopulmonary Resuscitation | Circulation».  
[https://www.ahajournals.org/doi/abs/10.1161/circ.122.suppl\\_21.a83](https://www.ahajournals.org/doi/abs/10.1161/circ.122.suppl_21.a83) (accedido 9 de septiembre de 2022).
- [3] E. Aramendi *et al.*, «Feasibility of the capnogram to monitor ventilation rate during cardiopulmonary resuscitation», *Resuscitation*, vol. 110, pp. 162-168, 2017.
- [4] D. P. Edelson, J. Eilevstjønn, E. K. Weidman, E. Retzer, T. L. V. Hoek, y B. S. Abella, «Capnography and chest-wall impedance algorithms for ventilation detection during cardiopulmonary resuscitation», *Resuscitation*, vol. 81, n.º 3, pp. 317-322, 2010.
- [5] H. E. Wang *et al.*, «Airway strategy and ventilation rates in the pragmatic airway resuscitation trial», *Resuscitation*, vol. 176, pp. 80-87, jul. 2022, doi: 10.1016/j.resuscitation.2022.05.008.
- [6] A. Elola *et al.*, «Capnography: A support tool for the detection of return of spontaneous circulation in out-of-hospital cardiac arrest», *Resuscitation*, vol. 142, pp. 153-161, 2019.

# Simulación de la actividad electrofisiológica auricular mediante autómatas celulares

G. S. Romitti, A. Liberos, P. Romero, D. Serra, I. García Fernández, M. Lozano, R. Sebastian, M. Rodrigo

CoMMLab, Universitat de València, València, España {giada.romitti, alejandro.liberos, miguel.rodrigo}@uv.es

## Resumen

*Los modelos biofísicos para la simulación de comportamientos electrofisiológicos auriculares pueden aumentar la personalización y eficacia de las terapias (farmacológicas o quirúrgicas). Sin embargo, su uso está limitado por su alto costo computacional que, a su vez, las aleja de los tiempos requeridos en la práctica clínica.*

*Los autómatas celulares disminuyen el tiempo de cálculo ayudando a generalizar el uso de modelos personalizados. Esto es gracias a describir las características eléctricas cardíacas con estados discretos. En particular, estos modelos se pueden basar en el cálculo de la duración del tiempo en el que se encontrarán en estado activo ( $APD^{+1}$ ) a partir del intervalo en el que se han encontrado en estado inactivo hasta recibir un nuevo disparo (DI). Sin embargo, hay que preguntarse si este enfoque es suficiente en la simulación de arritmias.*

*El análisis de 992 protocolos de activación simulados mostró un aumento esperado de la  $APD^{+1}$  con el intervalo diastólico (DI) anterior. Sin embargo, se encontró además una dependencia de la  $APD^{+1}$  con la activación previa ( $APD^0$ ). Al emplear para la predicción del  $APD^{+1}$  ambas variables ( $APD^0$  y DI) el error en  $APD^{+1}$  respecto al modelo biofísico ( $14 \pm 10$  ms) se redujo con respecto a la predicción basada en el DI ( $29 \pm 19$  ms,  $p=0,029$ ).*

*Tener en cuenta la duración de las activaciones anteriores mejora la precisión de los autómatas celulares auriculares, especialmente en escenarios con altas frecuencias de activación.*

## 1. Introducción

Las arritmias cardíacas se encuentran entre las principales causas de enfermedad y mortalidad. La fibrilación auricular (FA) es la patología cardíaca más prevalente y afecta a >10% de la población anciana [1]. La falta de conocimiento a cerca de los mecanismos específicos que inician y perpetúan la FA, resultan en una eficacia mejorable de los distintos tratamientos de esta arritmia, resultando en un gran problema clínico y económico. Un mayor grado de personalización en el diagnóstico, teniendo en cuenta el sustrato específico del paciente, puede resultar en una mayor eficacia de los tratamientos. Los modelos *in silico* representan una gran herramienta para guiar la medicina personalizada, ya que estos brindan información interesante sobre la manifestación individual de la FA. En particular, se ha confirmado que son especialmente útiles en el diagnóstico de situaciones patológicas [2], y en la evaluación de resultados de estrategias ablativas [3] y farmacológicas [4]. Numerosos modelos matemáticos describen las funciones cardíacas en detalle tanto desde el punto de vista electrofisiológico [5] como hemodinámico [6].

En el contexto de las taquiarritmias auriculares, los modelos electrofisiológicos se pueden utilizar para predecir la respuesta del paciente frente a diferentes escenarios, como cambios en la localización o frecuencia de activación de rotores, concentraciones iónicas, temperatura, pH o efectos de fármacos [1]. En este sentido, los modelos biofísicos [7] se encuentran entre los más utilizados y existen diversos *solvers* como ELVIRA [8] o openCARP [9], que logran describir con precisión los potenciales transmembrana (TP), las concentraciones iónicas o variables asociadas a los canales iónicos. Sin embargo, estos modelos basados en un gran número de sistemas de ecuaciones diferenciales ordinarias a resolver traen consigo un alto costo computacional y, por tanto, largos tiempos de simulación que pueden comprometer los tiempos de diagnóstico clínico. Por esta razón, se demanda el uso de modelos electrofisiológicos más ligeros y compatibles con los tiempos de diagnóstico. Los autómatas celulares (CA) responden a esta demanda y describen las características eléctricas cardíacas en base a estados discretos [10].

El objetivo de este trabajo es extender el uso de un CA [11] a la electrofisiología auricular. Este CA auricular se calibrará frente a simulaciones biofísicas auriculares [5] bajo protocolos de estimulación que aseguren un comportamiento realista bajo condiciones de arritmia.

## 2. Materiales y métodos

### 2.1. Modelo biofísico

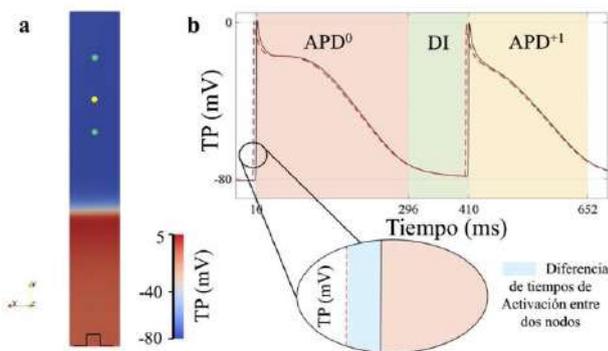
Las simulaciones biofísicas se ejecutaron sobre el *software* ELVIRA. Se empleó el modelo auricular de Courtemanche [12] sobre un tejido auricular rectangular (0,3x20x0,025 cm, 2106 células, 0,25 mm de distancia entre nodos, 0,01 ms de intervalo temporal) y estimulando desde el frente inferior, véase la Figura 1a. Se incluyó orientación de fibras, siendo el lado más largo el eje dominante con una difusión longitudinal de 0,0035 S/cm·pF y una relación transversal/longitudinal de 0,35. Para el análisis, y tal y como puede observarse en la Figura 1a, se consideraron tres puntos situados en el eje central del tejido: uno para el análisis de la curva de TP y dos para la medición de la Velocidad de Conducción (CV), como se muestra en la Figura 1a.

Se realizaron protocolos de estimulación S1-S2 de larga duración, correspondientes a trenes de 15 o 16 intervalos S1 (340 ms a 1000 ms, N=31 valores) seguidos de un estímulo S2 con valores de entre 100 ms y 1000 ms

( $N=32$  valores). Se simularon un total de 992 protocolos de entrenamiento S1-S2 empleando 31 simulaciones de larga duración, lo que resultó en un total de 3720 minutos simulados ( $120\pm 30$  minutos por simulación). El orden y la combinación de valores S1 y S2 se seleccionaron aleatoriamente.

Se utilizaron las curvas de TP para medir la Duración del Potencial de Acción (APD) y el Intervalo Diastólico (DI). El APD se midió como la diferencia entre el tiempo de activación (instante en el que se detecta la máxima pendiente en el TP) y el 90% de repolarización.  $APD^0$  y  $APD^{+1}$  denotan la última activación S1 y el potencial correspondiente al estímulo S2 respectivamente. El DI se midió como el intervalo desde el final de  $APD^0$  y el comienzo de  $APD^{+1}$  (ver Figura 1b).

La CV se midió como el espacio que separa los puntos de medida (6 mm) entre la diferencia en sus tiempos de activación (ver Figura 1b).



**Figura 1.** (a) Tejido auricular rectangular utilizado para simulaciones biofísicas. Se usaron los nodos amarillos y verdes para medir APD y CV respectivamente; (b) Potenciales transmembrana para dos nodos, se destaca el  $APD^0$  (rojo), DI (verde),  $APD^{+1}$  (amarillo) y la diferencia de tiempo de activación entre nodos (azul).

## 2.2. Autómata celular

Las simulaciones biofísicas que acabamos de describir demuestran precisión y facilidad en el análisis de los protocolos de estimulación. Sin embargo, el tiempo computacional que requieren es alto lo que dificulta su generalización en la práctica clínica. Serra [11] presenta un método alternativo para modelar la electrofisiología del tejido cardíaco, basado en un CA asíncrono, mediado por eventos, preparado para el análisis de tejidos y geometrías ventriculares.

Los CA, a diferencia de los modelos biofísicos, consideran sólo dos estados principales para cada elemento de tejido cardíaco: 0 (inactivo, es decir, repolarizado y excitable) y 1 (activo, es decir, capaz de activar a los vecinos). Cada evento simulado (activación, repolarización) se procesa en el momento exacto en que ocurre y, por lo tanto, no hay necesidad de un paso de tiempo granular.

La transición del estado 0 al 1 (activación) se desencadena por activaciones previas de nodos vecinos inmediatos (26 vecinos para mallas hexaédricas) o una

estimulación externa (manual). El instante exacto de activación se calcula utilizando el algoritmo *Fast-Marching* [13] considerando la velocidad de propagación de una onda plana. La CV de esta activación, así como el  $APD^{+1}$  del estado activo posterior, se calcula a partir de los estados anteriores y el DI. La transición del estado 1 al 0 (repolarización) se activa automáticamente una vez que finaliza el  $APD^{+1}$  calculado.

El cálculo de los valores de CV y  $APD^{+1}$  para cada nodo y activación es uno de los pasos más determinantes en los CA, y generalmente se ajustan resumiendo los datos experimentales en funciones numéricas. Este trabajo, a partir de los datos experimentales de simulaciones biofísicas, compara dos estrategias para el cálculo de CV y  $APD^{+1}$ : (1) considerando sólo el valor de DI anterior y obteniendo los valores de  $APD^{+1}$  de curvas de restitución o (2) considerando los valores de DI y  $APD^0$  y obteniendo los valores de  $APD^{+1}$  de superficies de restitución.

## 3. Resultados

### 3.1. Caracterización del modelo biofísico

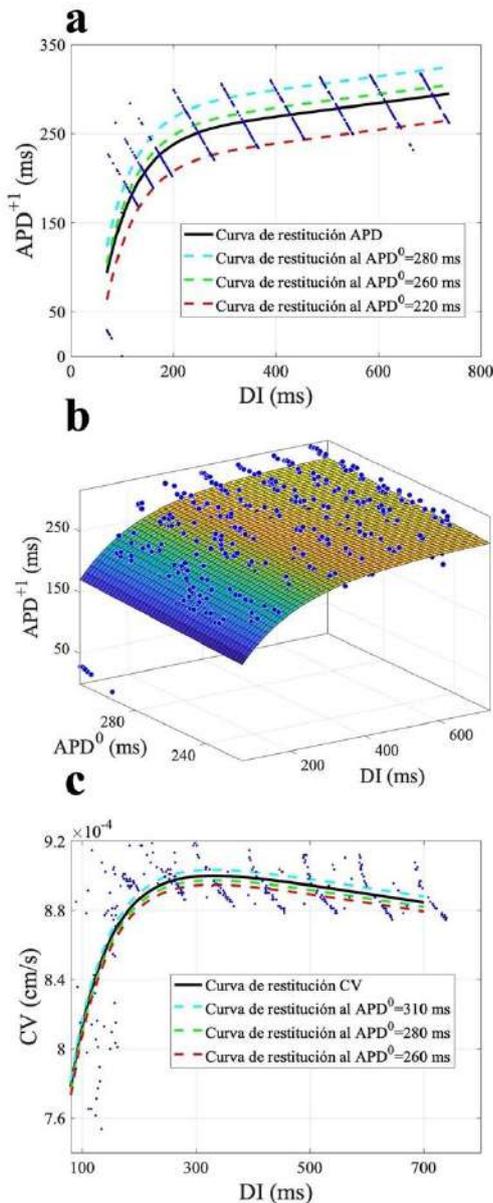
La Figura 2a describe las métricas de APD para las simulaciones biofísicas. Se puede observar la representación del  $APD^{+1}$  en función del DI. El ajuste a una sola curva produjo un error de  $23\pm 21$  ms. Esto se explica por el aumento del  $APD^{+1}$  con el  $APD^0$ , tal y como se puede observar en las diferentes curvas ajustadas para diferentes  $APD^0$ . Cuando el  $APD^{+1}$  se resume en una superficie y usando tanto DI como  $APD^0$  (Figura 2b), el error de ajuste fue considerablemente menor ( $10\pm 14$  ms,  $p<0.001$ ).

La Figura 2c muestra la caracterización de la simulación biofísica para la métrica CV. En este caso, el efecto del  $APD^0$  anterior fue casi inapreciable, y el ajuste de la curva más simple mostró un error similar ( $2,9\pm 4,6$  cm/s) al del ajuste por superficie usando DI y  $APD^0$  ( $3,0\pm 5,6$  cm/s,  $p=0,87$ ).

### 3.2. Simulaciones con el autómata celular

Para evaluar el comportamiento del CA, se reprodujeron las simulaciones del modelo biofísico empleando el CA. El tiempo de cómputo usando el CA fue en promedio 30 veces más rápido que el *solver* biofísico equivalente: 10,9 vs 327,3 segundos de tiempo de cómputo por segundo de tiempo simulado, respectivamente. El análisis de las simulaciones biofísicas tomó cinco veces más que en CA debido a la necesidad de calcular valores relacionados con el potencial de acción: 300 segundos en comparación con 60 segundos.

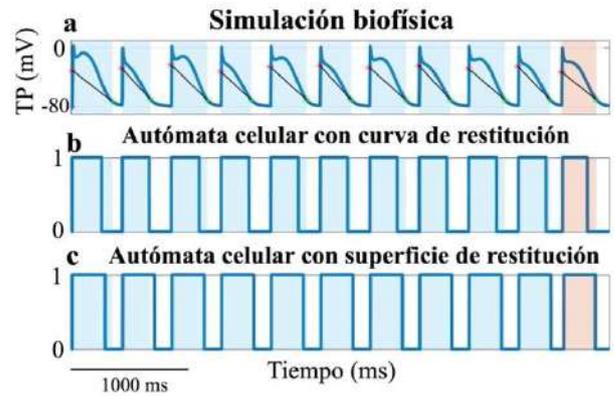
La Figura 3 muestra un ejemplo de una simulación biofísica (panel a) reproducida por el CA auricular utilizando ambas estrategias: (1, panel 3b) calcular el  $APD^{+1}$  a partir del DI y la curva de restitución y (2, panel 3c) calcular el  $APD^{+1}$  a partir del DI y el  $APD^0$ . Se observa como el CA basado en la superficie de restitución tiene una respuesta más cercana a lo observado en la simulación biofísica.



**Figura 2.** (a) Curva de restitución de APD, se pueden observar diferentes ajustes en función del APD<sup>0</sup> (b) Superficie de restitución de APD (c) Curva de restitución de CV.

Para confirmar este hallazgo individual en múltiples simulaciones, se realizaron dos análisis diferentes. La desviación teórica inicial se muestra en la Figura 4a, indicando la desviación producida por la caracterización del modelo biofísico. Este se calculó como la diferencia entre los valores de APD<sup>+1</sup> extraídos de las simulaciones biofísicas y los valores equivalentes obtenidos mediante la interpolación de la superficie o curva de restitución. Como era de esperar y de acuerdo con los resultados detallados en el punto 3.1, al emplear tanto DI como APD<sup>0</sup> el error en APD<sup>+1</sup> resultó menor (ver Figura 4a).

La segunda estrategia de validación consistió en la realización de nuevas simulaciones, tanto con los CA como con el modelo biofísico, que no fueron utilizadas previamente en el cálculo de las superficies/curvas de restitución. Las nuevas simulaciones realizadas (N=15) reprodujeron protocolos S1-S2 con diez S1 seguidos de



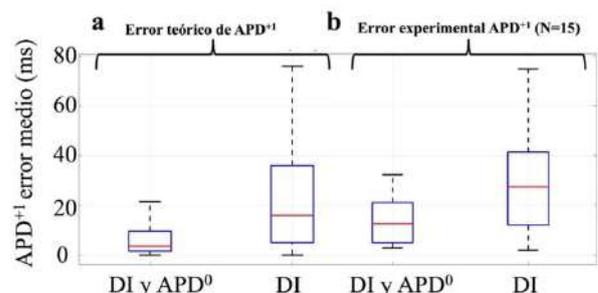
**Figura 3.** Ejemplo comparación entre CAs y simulación biofísica (a) Curva TP de la simulación biofísica (b) Respuesta del CA basado en curva de restitución o DI (c) Respuesta del CA basado superficie de restitución o DI + APD<sup>0</sup>.

un S2. Los valores de S1 se tomaron aleatoriamente entre 300 ms y 1000 ms y los valores de S2 entre 200 ms y 500 ms.

La verificación experimental confirmó la diferencia entre las dos estrategias de CA: considerando tanto APD<sup>0</sup> como DI, el error medido fue de 14±10 ms, mientras que, considerando solo DI, el error medio fue de 29±19 ms lo que resultó en una p=0.029 (ver Figura 4b). Esto implica que considerar el valor de APD anteriores mejora la estimación de APD posteriores.

#### 4. Discusión y conclusiones

Este trabajo detalla los resultados sobre el desarrollo de un autómata celular [11] capaz de reproducir la actividad electrofisiológica auricular. Los resultados de las simulaciones biofísicas en un tejido rectangular se caracterizaron utilizando valores DI, APD y CV. Estos marcadores sirvieron para evaluar la dependencia de APD<sup>+1</sup> y CV con respecto al APD<sup>0</sup>, lo que representa el efecto de memoria a corto plazo del modelo. APD<sup>+1</sup> depende del valor de APD<sup>0</sup> de una forma comparable a su dependencia con DI, en contraste, no se observaron dependencias relevantes con respecto al APD<sup>0</sup> en el caso de la CV. Como consecuencia, se decidió calcular el APD<sup>+1</sup> en función de DI y APD<sup>0</sup>, mientras que, por simplicidad, la CV dependerá únicamente de DI. Esto redujo el error en el valor estimado de APD<sup>+1</sup> en más



**Figura 4.** Error teórico (a) y experimental (b) en la estimación del APD.

de la mitad, acercándolo más a los resultados de las simulaciones biofísicas.

Los resultados obtenidos permiten disponer de un simulador con un tiempo de cálculo reducido (30 veces más rápido que el modelo biofísico) y con una alta precisión en los parámetros simulados, lo que lo hace interesante para uso clínico por la posibilidad de predecir resultados en tiempos de diagnóstico clínico.

Sin embargo, en el caso de la aurícula, el autómatas solo se ha aplicado a tejidos rectangulares. Deben extenderse los resultados a tejidos más realistas, incluidas geometrías auriculares tridimensionales. Además, el estudio se realizó con protocolos de estimulación de valores diferentes pero constantes, debiendo considerarse casos de ritmos irregulares los observados en episodios de fibrilación auricular.

Simular la electrofisiología de las aurículas es de gran interés en el diagnóstico de situaciones patológicas y en la evaluación preliminar de tratamientos de ablación o farmacológicos. Es necesario hacer esto en tiempos computacionales razonables compatibles con los tiempos clínicos. Este trabajo presenta un autómatas celular de las aurículas que permite predecir los APD con mayor precisión acercándose así a los resultados de modelos biofísicos. Esta mejora en la predicción no compromete la gran mejora en cuanto a tiempos de simulación que representan los CA frente a los modelos biofísicos.

## Agradecimientos

Este trabajo ha sido financiado por la Generalitat Valenciana AICO/2021/318 (Consolidables 2021), el Ministerio de Ciencia e Innovación y la Unión Europea a través de los fondos de desarrollo regional PID2020-114291RB-I00, PRE2021-100124.

## Referencias

- [1] Jalife J. et al. Basic cardiac electrophysiology for the clinician. John Wiley & Sons, 2011. (ISBN: 9781405183338).
- [2] Trayanova N.A. Mathematical approaches to understanding and imaging atrial fibrillation: significance for mechanisms and management. *Circulation research*, vol 114, sup 9, 2014, pp. 1516-31 (ISSN: 0009-7330).
- [3] Roney C. H., et al. Patient-specific simulations predict efficacy of ablation of interatrial connections for treatment of persistent atrial fibrillation. *EP Europace*, vol 20, 2018, pp. iii55-iii68 (ISSN: 1532-2092).
- [4] Grandi E., Maleckar M.M. Anti-arrhythmic strategies for atrial fibrillation: The role of computational modeling in discovery, development, and optimization. *Pharmacology & therapeutics*, vol 168, 2016, pp. 126-42, (ISSN: 0163-7258).
- [5] Liberos A. et al. Balance between sodium and calcium currents underlying chronic atrial fibrillation termination: An in silico intersubject variability study. *Heart Rhythm*. Vol 13, sup 12, 2016, doi: 10.1016/j.hrthm.2016.08.028 (ISSN: 2358-2365).
- [6] Zingaro A., et al. Hemodynamics of the heart's left atrium based on a variational multiscale-LES numerical method. *European Journal of Mechanics - B/Fluids*, vol. 89, 2021, pp. 380-400 (ISSN: 9977546).
- [7] Hodgkin A.L., Huxley A.F. A quantitative description of membrane current and its application to conduction and excitation in nerve. *J Physiol.*, vol. 117, 1952, pp. 500–44 (ISSN: 0022-3751).
- [8] Heidenreich E.A., Ferrero J.M., Doblaré M., Rodríguez J.F. Adaptive macro finite elements for the numerical solution of monodomain equations in cardiac electrophysiology. *Ann Biomed. Eng.*, vol. 38, 2010, pp. 2331-45, (ISSN: 0090-6964).
- [9] Plank G. et al. The openCARP simulation environment for cardiac electrophysiology. *Comput Methods Programs Biomed*, vol. 208, 2021 (ISSN: 0169-2607).
- [10] Von Neumann J., Burks A. W. Theory of self-reproducing automata. *IEEE Transactions on Neural Networks*, vol 5, sup 1, 1966, pp. 3-14 (ISSN: 2162-237X).
- [11] Serra D., et al. An Automata-Based Cardiac Electrophysiology Simulator to Assess Arrhythmia Inducibility. *Mathematics*, vol 10, sup 8, 2022, 1293 (ISSN: 2227-7390).
- [12] Courtemanche M., Ramirez R.J., Nattel S. Ionic mechanisms underlying human atrial action potential properties: insights from a mathematical model. *Am J Physiol*. vol 275, sup 1, 1998, pp. H301-21 (ISSN: 0002-9513).
- [13] Ethian J. A. Fast Marching Methods. *SIAM Review*. vol 41, 1999, pp. 199–235 (ISSN: 1095-7200)

# Subject-based association networks for clustering analysis and phenotyping of pediatric sleep apnea

J. Gómez Pilar<sup>1,2</sup>, D. Ferreira-Santos<sup>3,4</sup>, P. Pereira-Rodrigues<sup>3,4</sup>, D. Gozal<sup>5</sup>, R. Hornero Sánchez<sup>1,2</sup>, G.C. Gutiérrez Tobal<sup>1,2</sup>

<sup>1</sup> Grupo de Ingeniería Biomédica, Universidad de Valladolid, Valladolid, España

<sup>2</sup> Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), España

<sup>3</sup> MEDCIDS-FMUP – Community Medicine, Faculty of Medicine of the University of Porto, Porto, Portugal

<sup>4</sup> CINTESIS – Center for Health Technology and Services Research, Porto, Portugal

<sup>5</sup> Department of Child Health, Child Health Research Institute, The University of Missouri School of Medicine, Columbia, MO, United States

## Abstract

*Pediatric patients suffering from obstructive sleep apnea (OSA) show different responses to treatment, being the reasons for this heterogeneity not completely understood. The phenotypic characterization of this disease could shed new light to personalized treatments by identifying predictive factors of the patient's evolution. In this context, the objective of this study is to identify subgroups of patients, as well as the most relevant variables to differentiate them. This was conducted using a modularity analysis applied to a subject-based association network. The Childhood Adenotonsillectomy Trial (CHAT)-baseline dataset participated in the study. It comprises 464 pediatric patients ranging 5-10 years of age that includes a variety of sociodemographic and clinical data. A novel variation of the association network generation methodology allowed to obtain a network where each node represents a subject, and the links represent the associations between its characteristics. After applying a bootstrap-based approach and network modularity analysis following Blondel's algorithm, 3 subgroups of patients with similar intragroup phenotypic characteristics, but significantly different intergroup features, were identified. The main defining characteristics of each group were age, sex, presence of hypertension, and weight-related characteristics. Together, these findings highlight the need for acquiring clinical and sociodemographic data beyond those purely derived from polysomnography to provide a tailored diagnosis and treatment, while showing the ability of the new proposed methodology to find modules of subjects with well-differentiated phenotypical characteristics.*

## 1. Introduction

Pediatric obstructive sleep apnea (OSA) is a breathing disorder with high prevalence, affecting between 1 to 5% of the children worldwide [1]. It is characterized by the recurrent repetition of episodes of complete cessation (apneas) and decreases (hypopneas) of breathing during sleep that lead to hypoxemia, changes in intrathoracic pressure, surges in sympathetic activity, and changes in the heart rate regulation. Together, this get rise to the specific disruption of normal oxygenation and sleep patterns that depends on each patient [1].

OSA is diagnosed using standardized protocols based on overnight polysomnography (PSG) [2]. Then, adenotonsillectomy is the first line of treatment for the affected children. Nonetheless, the effectiveness of adenotonsillectomy is estimated to be no greater than 79%

of cases [3]. Those children who are not candidates for adenotonsillectomy typically require continuous positive airway pressure (CPAP) therapy. Adherence and tolerance to treatment is highly heterogeneous though, which suggests the need for tailored treatments. There is therefore a need to develop personalized clinical protocols based on precision medicine for an objective characterization of the disease leading to more accurate diagnosis and treatments.

In adult OSA, factors predicting treatment adherence have been scarcely studied, obtaining disparate results among the few works oriented to this problem. These findings varied among studies, with the most consistent ones being the severity of OSA based on the initial PSG parameters and the daytime sleepiness [4]. On the contrary, illness severity does not appear to be a predictive factor for adherence in children [5]. Although it has been suggested that OSA treatment maximize the reversibility of the adverse effects in the affected children [6], potential adherence predictors have remained even more unexplored to date. Therefore, the search of subgroups of pediatric OSA population with differentiating characteristics could help provide patient-oriented diagnosis and treatment.

In this context, the objective of this study is to identify subgroups of pediatric OSA patients, as well as the most relevant variables of each subgroup. Some previous studies in adults successfully used an automatic analysis based on a *k*-means algorithm extension to determine 3 adult clusters/phenotypes [7]. Alternatively in this work, we propose a novel technique via modularity analysis applied to associations networks. Particularly, our proposal is the use of a variation of this technique to identify subgroups of patients with characteristics that could have been hidden by other conventional clustering techniques. This family of techniques, widely known in fields such as genetics [8], [9], has not been applied in the context of pediatric OSA, except for a single previous study in which it showed enormous potential to differentiate electroencephalographic patterns [6].

Based on the previous evidence shown in studies involving adults [4], [7], our starting hypothesis states that it is possible to find groups of pediatric OSA patients with well-differentiated phenotypic characteristics.

## 2. Materials and methods

### 2.1. Participants

In this study, we used the multicenter Childhood Adenotonsillectomy Trial (CHAT)-baseline dataset [3], after obtaining proper approval ([www.sleepdata.org](http://www.sleepdata.org)). This dataset is comprised of 464 pediatric patients ranging 5-10 years of age. All these patients met the criteria for being considered for adenotonsillectomy. PSG was conducted and apnea/hypopnea events were scored according to the American Academy of Sleep Medicine 2007 guidelines [2]. The OAH1 was defined as the number of all obstructive apneas and hypopneas. In addition to PSG data, clinical, and sociodemographic variables were also acquired. Among all the information, a total of 26 variables were used to try to identify the phenotyping subgroups. These were selected to get the set of variables with the highest coincidence degree with previous studies in adults among all the variables included in the CHAT database [7].

### 2.2. Association networks

The association network analysis is a technique used in different fields to assess the relationship between categorical variables [6]. In this approach, the nodes in the network represent the variables under study, while the values of the connection between the nodes are an index of the association between each pair of variables.

We here propose a modification of this approach in which the nodes represent the subjects under study (unlike the usual technique where each node represents a variable). Each subject is then characterized by a vector of categorized variables (26 in this case). Therefore, the values of each link between nodes are computed as the  $X$ -square value (appropriate for categorical data) between the vector of a given pair of subjects. Thereby, subjects with higher degree of association between them will obtain a high value in their connection.

In order to remove the links with a residual association between the respective nodes, we only maintain those links with a statistical degree of association between them ( $p < 0.05$ ). This results in a semi-weighted network, i.e., the values of non-significant links are set to zero, while the other links are set to the  $X$ -square value of a given pair of nodes.

### 2.3. Bootstrap procedure

Aiming at increasing the robustness of the results, a previously validated bootstrap procedure was applied [6], [9]. We first randomly selected a subgroup of 26 variables with possible repetition. Then, the association network, i.e., association matrix, was estimated following the previously described computations between nodes (subjects). This procedure was repeated 1000 times, selecting the mean value matrix for subsequent analysis among all those obtained. This methodology provides more robust networks in the sense of higher reproducible results, while allowing the evaluation of the stability of the network. Thus, if the parameters obtained in the generated networks have a high variance, it implies that the network

is not very stable and, therefore, reaches less generalizable results [9], [10].

### 2.4. Network visualization

The resulting network were represented using *Gephy* (version 0.9.6) software. It allows to represent the network using ForceAtlas2 approach [11]. ForceAtlas2 is based on visualizing the associations between nodes (subjects) by mirroring physical attraction and repulsion forces. Considering both the distance and the node degree (sum of the  $X$ -square values reaching a node) of the connected nodes, this method turns structural proximities into visual proximities. In this way, subjects with a higher level of association between them are more likely to appear closer in the network.

### 2.5. Cluster analysis

Cluster analysis aimed at finding similar groups of subjects, where "similarity" means a global measure over the full set of characteristics. In this study, we used an unsupervised learning algorithm for clustering the data in the association network, meaning in this context that the algorithm neither has prior information about the number of clusters that exist before running the model nor does it make assumptions about relationships within the data. Particularly, we used Blondel's modularity [12]. This is a heuristic method based on modularity optimization. It is shown to outperform all other known community detection methods in terms of computation time, being important in large networks. The algorithm provides a modularity value that indexes the modular characteristic of the network and, more importantly, it gives a label for each node indicating the cluster to which said node belongs.

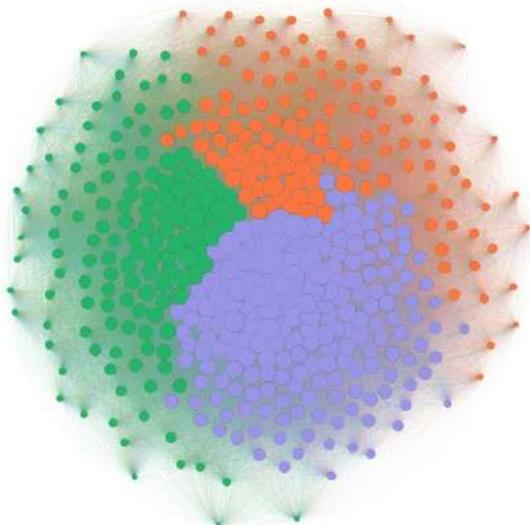
## 3. Results

The generated network model is shown in Figure 1. The color of the nodes (patients) indicates each of the 3 groups automatically identified by Blondel's algorithm. On the other hand, the size of each node represents the average level of association with the rest of the nodes in the network. In addition, after applying the ForceAtlas2 algorithm, the nodes are placed so that those with the highest level of association between their parameters will appear spatially closer. As both the size of the nodes and the position of them depends on the level of association, a relationship between them can be observed. Thus, the largest nodes tend to appear centered in the network, while the smallest nodes are located on the periphery of the network.

Among the 26 variables used, Table 1 shows the 8 with the highest differences among the 3 modules automatically defined. As observed, these modules showed well differentiated values among these variables. First, module A (green) consists of 136 subjects, all of them males. On the contrary, module B (orange) is composed of 123 female patients. Module C (purple) include both male and female patients with similar percentages (41.1% and 58.9%, respectively). Age also proved to be one of the most differentiating characteristics among modules. Interestingly, module C had a significantly higher mean

	Module A (green)	Module B (orange)	Module C (purple)	$\chi^2$ -square test ( $X p$ )
Subjects (#)	136	126	202	-
Sex (males%   females%)	100   0	0   100	41.1   58.9	267.8   <0.001
Age (% between 5-6 years   7-8 years   9-10 years)	77.2   18.4   4.4	68.3   26.2   5.6	34.2   43.6   22.2	76.8   <0.001
OAHI (% < 1 event/hour   > 1 event/hour)	1.5   98.5	11.9   88.1	3.0   97.0	12.1   0.002
BMI (underweight%   normal%   overweight%   obese%)	6.6   87.5   4.4   1.5	4.8   89.7   4.8   0.8	0   1.5   25.7   72.8	398.3   <0.001
Waist circumference (normal%   increased%)	73.0   27.0	60.3   39.7	2.5   97.5	202.8   <0.001
Neck circumference (normal%   increased%)	100   0	99.2   0.8	70.3   29.7	51.4   <0.001
Arterial hypertension (No%   Yes%)	94.8   5.1	95.2   4.8	81.7   18.3	21.2   <0.001
Epworth sleepiness scale (No%   Yes%)	83.8   16.2	80.9   19.0	74.75   25.25	27.1   <0.001

**Table 1.** Summary of the main differences between modules of the clinical, socio-demographic, and PSG data.



**Figure 1.** Subject-based association network of the pediatric OSA patients. The network is derived from clinical, socio-demographic, and PSG data after a bootstrap procedure. The size of each node (patient) is determined by the average level of association with the rest of the nodes in the network. The color represents the module to which it belongs according to Blondel's algorithm [12]. Its spatial distribution follows a heuristic of attractive and repulsive force to reach a stable position of all the nodes, following ForceAtlas2 approach [11].

age, with 65.8% of patients in this module older than 7 years of age.

Module C showed other noteworthy characteristics frequently related to weight. A noticeable example is that the 98.52% of patients in this module are overweight or

obese. Also, waist and neck circumference showed significant higher values compared to modules A and B. In fact, among all overweight/obese, increased waist, and increased neck subjects, 93.75%, 98.36%, and 69.37% of them belong to module C.

Finally, in line with the previously mentioned results, the subjects belonging to module C also showed differentiating clinical characteristics, as in the Epworth sleepiness scale (ESS). About a quarter of the subjects in this module showed excessive daytime sleepiness. Interestingly, OAHI was the characteristic with lower significant differences between the three modules. Together, these results highlight the noticeable association between different PSG, sociodemographic, and clinical parameters, as well as the ability of the proposed methodology to find modules of subjects with well-differentiated characteristics.

#### 4. Discussion

In this study, we applied a novel association network-based approach to identify phenotypic profiles in children with OSA. After the use of a robust bootstrap-based statistical assessment, the conducted unsupervised modularity analysis revealed 3 distinctive groups of pediatric OSA patients.

One of the identified subgroups (module C) showed a distinctive profile in obesity-related variables, such as BMI, and waist and neck circumferences. Interestingly, previous studies related obesity to the ability to adhere to treatment in children suffering from OSA [13]. Moreover, and again in line with previous studies [5], this group did not show a higher OSA-severity profile in terms of OAHI. Indeed, OAHI was the characteristic with the lowest differences among the 8 variables of Table 1.

In addition to the obesity-related variables, our subject-based association networks let identify new characteristics that, although potentially associated with weight, have not been related to phenotypic subgroups of pediatric patients. This is the case of the age, sex, sleepiness, and arterial hypertension. All of them showed great significant differences between the subgroups of patients found. Based on these findings, it seems feasible to establish the hypothesis that patients belonging to module C may have either lower response to adherence or treatment [5], thus benefiting the most of personalized interventions. This, however, remains unexplored, so future studies with a longitudinal design could be conducted to evaluate it.

Curiously, a recent work focused in adult OSA phenotyping, and using 40 variables, showed 3 different modules too [7]. Despite some of the variables defining the subgroups are not suitable for children, such as alcohol consumption, the authors also reported the high importance of obesity, neck circumference, sex, and age to define the groups [7].

Among the limitations of the study, it is worth highlighting the lack of cognitive variables for the identification of subgroups of subjects. In this regard, recent works show the importance of exhaustively characterize the phenotypic profile of patients and recommend decreasing the threshold for conducting adenotonsillectomy from 5 event/hour of OAH to 1 events/hour if the child presents associated neurocognitive symptoms [14]. Therefore, future studies should include these types of variables into account when establishing subgroups of patients with the aim of finding those who respond appropriately. In addition, a longitudinal study design could provide important clues regarding the adherence capability to treatment of each subgroup of pediatric patients. Finally, the use of variables tailored to children could derive in higher modularity degrees.

## 5. Conclusions

In this study, a new method grounded on subject-based association networks, and its modularity analysis, was proposed to identify subgroups of OSA patients with well-differentiated phenotypic characteristics. Our novel proposal reached results coherent with other studies conducted in both adults and children. One of the three subgroups automatically identified shows a profile of variables consistent with lower adhere/response to treatment. Additionally, the usefulness of previously unexplored variables relevant for OSA subgrouping were determined. This opens the door to the search for new predictive variables of adherence and response to pediatric OSA treatment, while highlighting the need for simultaneously acquiring clinical and sociodemographic variables beyond those purely derived from PSG to provide tailored diagnosis and interventions.

## Acknowledgements

This research has been developed under the grants PID2020-115468RB-I00 and PDC2021-120775-I00 funded by 'Ministerio de Ciencia e Innovación/Agencia Estatal de Investigación/10.13039/501100011033/' and

ERDF, A way of making Europe; and by 'CIBER en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN)' through 'Instituto de Salud Carlos III' co-funded with ERDF funds as well as under the project SleepyHeart from 2020 CIBER-BBN valorization call.

## References

- [1] C. L. Marcus *et al.*, "Diagnosis and Management of Childhood Obstructive Sleep Apnea Syndrome," *Pediatrics*, vol. 130, no. 3, pp. e714–e755, Sep. 2012, doi: 10.1542/peds.2012-1672.
- [2] C. Iber, S. Ancoli-Israel, A. Chesson, and S. Quan, "The AASM Manual for the Scoring of Sleep and Associated Events: Rules," *Terminology and Technical Specification*, 2007, Accessed: Oct. 13, 2021. [Online]. Available: <http://ci.nii.ac.jp/naid/10024500923/en/>
- [3] C. L. Marcus *et al.*, "A Randomized Trial of Adenotonsillectomy for Childhood Sleep Apnea," *New England Journal of Medicine*, vol. 368, no. 25, pp. 2366–2376, Jun. 2013, doi: 10.1056/NEJMoa1215881.
- [4] M. Kohler, D. Smith, V. Tippett, and J. R. Stradling, "Predictors of long-term compliance with continuous positive airway pressure," *Thorax*, vol. 65, no. 9, pp. 829–832, Sep. 2010, doi: 10.1136/thx.2010.135848.
- [5] N. DiFeo *et al.*, "Predictors of Positive Airway Pressure Therapy Adherence in Children: A Prospective Study," *Journal of Clinical Sleep Medicine*, vol. 08, no. 03, pp. 279–286, Jun. 2012, doi: 10.5664/jcsm.1914.
- [6] G. C. Gutiérrez-Tobal *et al.*, "Pediatric Sleep Apnea: The Overnight Electroencephalogram as a Phenotypic Biomarker," *Front Neurosci*, vol. 15, Nov. 2021, doi: 10.3389/fnins.2021.644697.
- [7] D. Ferreira-Santos and P. P. Rodrigues, "Obstructive sleep apnea: A categorical cluster analysis and visualization," *Pulmonology*, Nov. 2021, doi: 10.1016/j.pulmoe.2021.10.003.
- [8] P. J. Gutiérrez-Díez, J. Gomez-Pilar, R. Hornero, J. Martínez-Rodríguez, M. A. López-Marcos, and J. Russo, "The role of gene-to-gene interaction in the breast's genomic signature of pregnancy," *Sci Rep*, vol. 11, no. 1, 2021, doi: 10.1038/s41598-021-81704-8.
- [9] N. Jimeno *et al.*, "Main Symptomatic Treatment Targets in Suspected and Early Psychosis: New Insights From Network Analysis," *Schizophr Bull*, pp. 1–12, Jan. 2020, doi: 10.1093/schbul/sbz140.
- [10] S. Epskamp, D. Borsboom, and E. I. Fried, "Estimating psychological networks and their accuracy: A tutorial paper," *Behav Res Methods*, vol. 50, no. 1, pp. 195–212, Feb. 2018, doi: 10.3758/s13428-017-0862-1.
- [11] M. Jacomy, T. Venturini, S. Heymann, and M. Bastian, "ForceAtlas2, a continuous graph layout algorithm for handy network visualization designed for the Gephi software," *PLoS One*, vol. 9, no. 6, pp. 1–12, 2014, doi: 10.1371/journal.pone.0098679.
- [12] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, "Fast unfolding of communities in large networks," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2008, no. 10, p. P10008, Oct. 2008, doi: 10.1088/1742-5468/2008/10/P10008.
- [13] S. L. Katz *et al.*, "Factors related to positive airway pressure therapy adherence in children with obesity and sleep-disordered breathing," *Journal of Clinical Sleep Medicine*, vol. 16, no. 5, pp. 733–741, May 2020, doi: 10.5664/jcsm.8336.
- [14] H.-L. Tan, M. L. Alonso Alvarez, M. Tsaousoglou, S. Weber, and A. G. Kaditis, "When and why to treat the child who snores?," *Pediatr Pulmonol*, vol. 52, no. 3, pp. 399–412, Mar. 2017, doi: 10.1002/ppul.23658.

# A Cross-Platform Editor and Simulator of Magnetic Resonance Imaging Sequences: Design and Implementation

P. Villacorta-Aylagas<sup>1</sup>, I. Fernández-Arias<sup>1</sup>, C. Castillo-Passi<sup>2</sup>, P. Irarrazaval<sup>2</sup>,  
F. Simmross-Wattenberg<sup>1</sup>, M. Rodríguez-Cayetano<sup>1</sup>, C. Alberola-López<sup>1</sup>

<sup>1</sup> Laboratorio de Procesado de Imagen. ETSI Telecomunicación. Universidad de Valladolid.

<sup>2</sup> Instituto de Ingeniería Biológica y Médica, Pontificia Universidad Católica de Chile.

## Abstract

*Simulation plays an important role in MRI research and training. Several simulators exist. Nevertheless, they do not have an easy interface to design the MRI sequence to be simulated. In this work, we developed an application that provides a graphical interface for that purpose. We have developed software with a graphical interface in which the user is allowed to build any arbitrary MRI sequence by arranging basic blocks such as gradients and RF pulses as well as to carry out a physics-informed simulation. This editor, implemented in Qt, makes use of KomaMRI.jl, an existing GPU-accelerated MRI simulator. We chose this simulator because it is the fastest. The main feature of our application is that it supports web-based technologies so that it can run either as a desktop application or in a browser. It also includes other functionalities, such as saving/loading sequences or the possibility of visualizing the excited slice in three dimensions. The results of its development have been satisfactory and promising for future work.*

## 1 Introduction

Magnetic resonance imaging (MRI) is a non-invasive medical technique used for the diagnosis of pathologies related to soft tissues [1]. It is a versatile tool, so it has become the most appropriate choice in many clinical situations. However, it is not an easy task to operate with a MRI system and many factors contribute to these difficulties [2]. These obstacles, coupled with the high cost of real MRI systems as well as their high clinical demand, make the task of learning this technology an issue.

This motivation gave rise to MRI simulators, which are a breadth of tools that ease the understanding of all the elements involved in the process, and they also allow practitioners to carry out an unlimited number of uses. Since this idea emerged, several simulators have been described. One of these simulators is *KomaMRI.jl* [3]; this tool, developed in the Julia programming language [4], is meant to simulate general MRI scenarios that could arise in pulse sequence development. It has a GUI that allows for the visualization of the outcomes that arise from the simulation of a predefined sequence (reconstructed image, K-space, raw signal, pulses, etc.). However, the tool lacks an actual sequence editor as well as a graphical tool to observe the selected slice. This, coupled with the fact that, to the best of our knowledge, no MRI sequence editors are available on web-based technologies, is the main motivation for this paper.

What we propose here is an MRI sequence editor with the following features:

- **Simple to use**, versatile and both education- and research-oriented purposes.
- **Cross-platform**. Hosted on a server, the application can be accessed from any web browser.

This article aims to describe the sequence editor creation process from an engineering software perspective. Albeit in the paper we focus on the desktop-application, the technologies used enable us, with almost no code modification, to embed it in a client-server application; detailed steps on how to accomplish the latter are also provided.

## 2 Background

Nowadays, there is a wide variety of tools available for MRI simulation. Some of them (like the ones described in [5] and [6]) were created with the aim of teaching theoretical MR concepts such as spin dephasing (Bloch equation) or quantum mechanics, but had no direct medical purposes [2]. Other simulators, for example, Virtual MRI [7], can emulate some of the main elements of the acquisition process in an actual MRI scanner. Torheim [8] simulator has a very similar behaviour to the one we have just mentioned, and shares the didactic purposes [2].

As for the research-oriented simulators, we should talk about JEMRIS [9], an open-source simulator focused, like the one here presented, on sequence development. It considers many properties of interest in MRI, to be as realistic as possible; however, it presents the drawback of only using CPU multi-threading. Other alternatives could be BlochSolver [10], MRISIMUL [11] (both closed-source) or MRILab [12], one of the most recent MRI simulators.

*KomaMRI.jl* [3] is a simulator intended to be fast, easy-to-use, extensible, open-source, and cross-platform. This simulator has been developed in Julia, a nice language for fast prototyping which includes precompiling and GPU support; the former should make *KomaMRI.jl* as efficient as other simulators written in C/C++ languages and the latter makes fast execution possible.

## 3 Analysis of requirements

The main functional requirements are:

- A set of basic building blocks for a MRI sequence (essentially, RF pulses and gradients) should be available and ready to use out of the box.
- The application should allow for the creation of arbitrary MRI sequences by arranging these basic blocks through simple drag & drop operations. Every block should be individually configurable and sequences should be passed to the simulator.
- There must be the option to generate groups of blocks. These groups, from the user's perspective, will be seen as blocks in themselves, so all the operations that can be performed on simple blocks (create, move, delete) must also be available on groups.
- The functionality of saving and loading sequences must be

included.

- The user must be able to define some global parameters for the simulation (intensity of the main magnetic field  $B_0$ , sampling period  $\Delta t$ , maximum gradient  $G_{max}$ , etc.). Other acquisition parameters, such as TE (echo time), TR (repetition time) or ETL (echo train length) are implicitly defined by means of the sequence.
- The user interface should provide a view of the reconstructed image and its corresponding K-space, as well as a temporal diagram of the created sequence.
- A 3D-visualization of the excited slice should also be displayed.
- The simulation core may reside either on the local machine or on a remote one.

Regarding non-functional requirements, these include the development of a desktop application and its subsequent extension to a client-server web version, which shares the same code base, and thus is identical to the desktop version for the end user.

## 4 Design and implementation

### 4.1 Selected technologies

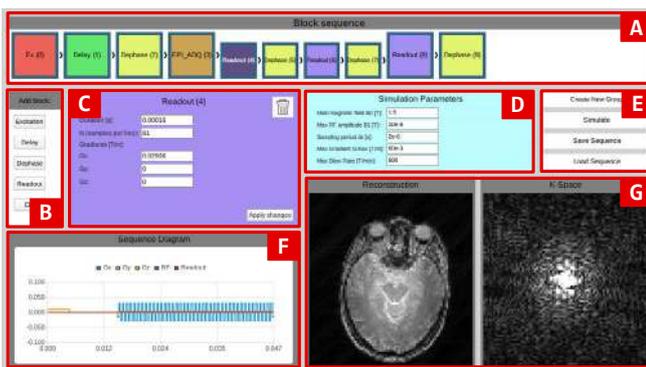
Once the requirements have been established, it is necessary to select the technologies to be used. As for the graphical user interface (this is where the user must define all the simulation parameters, including the sequence), Qt [13] is chosen. This C++ framework is oriented to cross-platform (desktop, web or even mobile) application development, so it seems the right option; specifically, Qt Quick, a module for interface development based on the QML language, has been used. The usefulness of QML mainly lies on the definition of the elements that make up the interface.

In addition, Qt permits compilation to WebAssembly [14], a type of low-level code that can be executed in modern browsers.

### 4.2 Implementation

#### 4.2.1 Interface overview

Here we present the design, both visual and functional, of the interface itself. Based on all requirements previously seen, the final application appearance is shown in figure 1 in which, by means of panels, all the functionalities are included. The interface has been designed to be simple and intuitive.



**Figure 1.** Graphical User Interface sketch. Each panel is labelled with a letter used in the main text.

The sequence panel (A) stands out by occupying the entire width of the application window; this is intended as a way to attract the user attention to the sequence, which is the main constituent of the whole process. In addition, it eases its complete visualization, without the need of trimming or drastically reducing the number of visible blocks; these blocks can be freely moved, modified and deleted. Panel (B) provides the library of available blocks,

which can be added to the sequence by simply clicking one of the panel buttons. Panels (C) and (D) contain, respectively, the configurable block and the global parameters. As for the former, clicking on a given block will display its configuration menu from which related parameters can be easily set; as for the latter, global parameters are directly tunable. Panel (E) is intended for sequence management; simple operations such as sequence loading and saving are available. In addition, the user is allowed to create a block group; groups are intended for providing repetition capabilities without the need of explicitly writing the repeated blocks. The button to launch the simulation is also allocated here for convenience. Finally, the lowermost panels, namely, panels (F) and (G), are intended for visualization; specifically, panel (F) will display the sequence pulse diagram while panel (G) will show the simulation result, both in image and k spaces.

#### 4.2.2 An overview of KomaMRI.jl

As stated above, KomaMRI.jl will be used for the simulation itself. This simulator presents three input arguments, which will be instances of the following structures:

- **Scanner:** this data structure contains the global parameters for the simulation, such as the main magnetic field strength  $B_0$ , the raster time  $\Delta t$  or the maximum slew-rate, among others.
- **Phantom:** it simulates what in a real scanner would be the patient's body (or part of the body), so it contains the data on which the sequence is applied. This data include the spatial location of the spins,  $T_1$ ,  $T_2$ ,  $T_2^*$  times and proton density  $\rho$  information, among others.
- **Sequence:** it contains the MRI sequence information itself. In the simulator, a `Sequence` object consists of:
  - A 2D array of `Grad` elements. One dimension (rows) is space (so that we can apply gradients in  $x$ ,  $y$  and  $z$ ) and the other (columns) is time. `Grad` structure mainly consists of two real scalars: `A`, corresponding to the gradient amplitude, and `T`, which refers to its time duration.
  - A 2D array of `RF` elements. As before, columns dimension refers to time; as for the other dimension, each row represents a different coil, so various RF pulses may be simultaneously applied. The `RF` structure is composed of these attributes: `A` is the amplitude of the RF pulse complex envelope, `T` corresponds to the pulse duration and  $\Delta f$ , included in latest versions, accounts for the difference in frequency with respect to the Larmor frequency of the RF pulse.
  - A 1D array of `ADC` elements, the dimension of which is time. `ADC` structure stands for data acquisition and contains information about when the readout is accomplished. It contains two values: `T`, which is the duration, and `N`, which accounts for the number of (evenly-spaced) samples to be acquired during time `T`.

#### 4.2.3 Interfacing with KomaMRI.jl

We need functions from KomaMRI.jl to be called from the GUI and their input arguments to be in the correct format. To carry out this task, file `QtSlots.jl` has been added to the source directory of KomaMRI.jl. This file contains the `sim` function, which acts as an “intermediary” between the GUI and the simulator; it is responsible for adapting the data related to the sequence and the global parameters to a suitable format. Our solution has been to provide a `Sequence` element (as described in the previous section) for the former and a `Scanner` object—which is, in essence, a vector—for the latter. The third input argument is a `Phantom` object, also described above. This phantom is obtained, not from the GUI, but from a `(.nii)` file. Figure 2 provides a simple outline of all this.

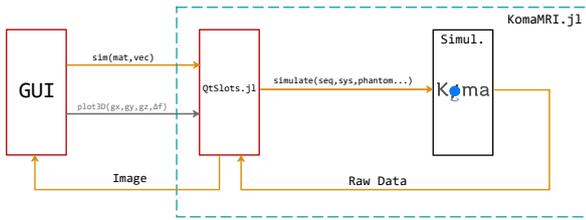


Figure 2. Desktop-based application operating diagram

In Qt/QML, sequence data is stored in a `ListModel` structure, which is nothing more than a container of definitions of the `ListElement` structure. Each `ListElement` instance contains information about a sequence block.

Once the `Simulate` button is pressed, the sequence information will be passed from QML to C++ by means of a Qt slot (see *Signals and Slots* in [13]). The next step is to call the Julia function, passing the sequence matrix and the global parameters vector (now, available in C++) as input arguments. The way to do this is by using a C/C++ API that allows embedding Julia (expressions, functions, variables...) in C++. Specifically, the following line of code is in charge of calling the –previously defined in C++– Julia function:

```
j1_array_t * output = (j1_array_t*)j1_call12(j1_sim, seq_arg, sys_arg)
```

Finally, a function called `plot3D` is defined in file `QtSlots.jl`. Its task is to show a 3D representation of the excited slice while displaying the entire 3D model by means of a volume rendering technique (see Figure 5). This functionality provides us with precise information of the slice selected for a choice of selection gradients and it is reactive in real time to the choice of these gradients from the GUI. The visualizer is implemented by using `PlotlyJS`, a Julia package oriented to the construction of graphs and diagrams, albeit with limited support for 3D rendering.

#### 4.2.4 Technical steps for a client-server application

As previously stated, despite the code is reusable from a browser-based perspective, the application is intended to be deployed in an application server so a client-server architecture is mandatory. Hence, additional elements are needed. Figure 3 provides a detailed diagram of what these elements are. Specifically:

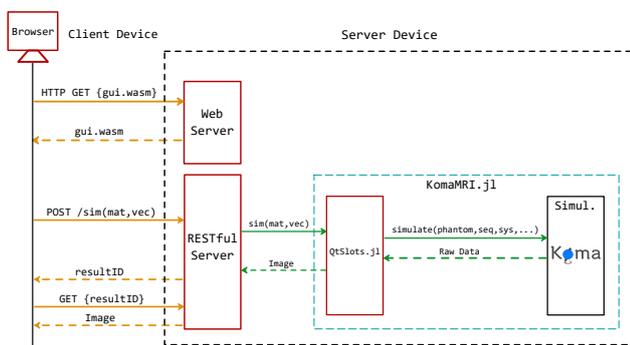


Figure 3. Web-based application operating diagram. In this case, orange arrows represent commands of the protocol established between the client and the server (HTTP), while the green arrows correspond to Julia function calls.

- A **web server**, capable of providing the client with the resulting files from the compilation to WebAssembly, which are, mainly, a `.wasm`, a `.html` and a `.js` file. These

three files will be loaded on the client’s browser, so the user will be able to see the GUI and interact with it.

- A **REST-based (RESTful) server**: this server will run remotely-requested commands and will return their results. A RESTful web service works similarly to a web page, so we only need to send requests and receive their corresponding responses. It is created using Flask, a Python framework for creating web applications in a simple way. The RESTful server supports the following functions:
  - `POST/commands`: `POST` executes a new command and saves its output in a temporary file whose identifier is by default returned in the Location header of the response. A JSON document with the fields `command` and `arguments` must be included in the request body.
  - `GET/commands/{resultID}`: `GET` returns the content of the file with local identifier `resultID` (the one obtained in the Location header of the `POST` response).
  - `DELETE/commands/{resultID}`: deletes the temporary file created as a result of the `POST` command.

## 5 Results and discussion

In this section we will show the results obtained from the execution of the program. We will also see the functional graphical interface, the 3D visualization tool and the current status of the web version.

### 5.1 Desktop version

As previously mentioned, Figure 1 showed the final appearance of the editor. In this particular case, an 81x81 pixel image was acquired by means of a GE-EPI sequence. The RF pulse was implemented by adding an `Ex` block; a delay block is then added to control TE. Then a `Dephase` block is added to shift the readout position to the lower-left corner of `k`space. Then the `EPI_ADQ` block turns out to be a group, the contents of which (when expanded) are shown as smaller rectangles than regular blocks. The group consists of a two pairs of readout and dephase blocks; the first pair is a rightward readout plus an upward shift in the phase encoding direction; the second pair is a leftward readout plus the upward shift. This group is repeated forty times. The final readout line (to complete the 81 `k` space line reading) is carried out by the last `Readout` block (index number 8) which is not part of the group. The final `Dephase` block is not mandatory, since it simply returns the readout point to the `k` space origin.

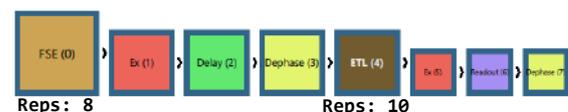
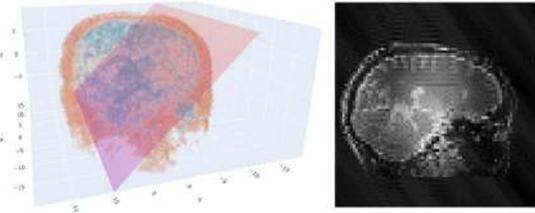


Figure 4. Block hierarchy for an FSE sequence

Other sequences are possible and they give us room to illustrate the block hierarchy that the GUI makes use of. Specifically, we will give rise to a Fast Spin-Echo (FSE) with 10 echoes per shot. The sequence has been built by two nested groups (see fig. 4). The outer group, referred to as `FSE`, starts with an excitation block generating the 90° RF pulse, followed by a `Delay` and a `Dephase` that allows the pointer to be placed at the initial position—the lower left corner of `K`-space, for example—. We set the number of repetitions of this group to 8, i.e., the sequence will consist of eight shots. The inner group referred to as `ETL` controls the readout positioning and the readout process itself; this group consists of an `Excitation` block that generates the 180° refocusing RF pulse followed by a –rightward– `Readout` block and a `Dephase` to move the pointer to the next line. Its number of repetitions—corresponding to the echo train length (ETL)—will be set to 10, so that a total of 80 lines will be read at the end of the simulation. The result is very similar to the one shown in figure 1.

## 5.2 3D Visualization

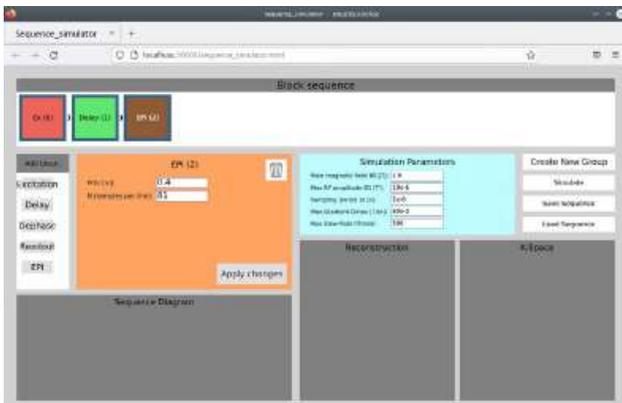
When clicking the `View 3D Model` button on the EX block configuration menu, a new window emerges. This window shows, simultaneously, the entire volume (on which the simulation is to be performed) and a plane through it, which represents the excited slice. Figure 5 shows the influence of the gradients activated during the excitation period (in our case, gradients in  $y$  and  $z$ ).



**Figure 5.** 3D visualization of the excited slice and its corresponding reconstruction

## 5.3 Web browser version

Figure 6 shows the browser version of the GUI. In this case, the test was performed in Mozilla Firefox; other popular browsers such as Chrome, Edge or Safari, have also been tested. As previously stated, graphical functionality so far has been implemented with Julia packages. This is a limitation for a web-based application since these packages are not currently supported by WebAssembly. However, well-known graphical tools, such as VTK [15] and, specifically, `vtk.js` (a rendering Javascript library made for scientific visualization on the web), are off-the-shelf material that can readily circumvent this problem.



**Figure 6.** Application running on a web browser

## 6 Conclusions and future work

This paper has presented a MRI sequence editor which, combined with `KomaMRI.jl` simulator, allows the user to create any arbitrary sequence and to visualize its results. The tool is intended to be cross-platform and simple to use. Its design and implementation (GUI-Julia interaction, client-server architecture operation, 3D visualization...) has also been described.

The GUI provides practitioners with easy access to `KomaMRI.jl` that would otherwise require programming skills and profound knowledge of the Julia language. Despite this language is similar to other popular prototyping languages, such as Matlab and R, it has some specificities (such as the a rigorous control of the variables types and its consequences in related operations, such as operator overloading, to give an example) that makes the learning curve an issue. Hence, our GUI hides this complexity to the user to that they can concentrate on their interests, whether edu-

cational or research-oriented. In addition, the aim is to offer this tool as free software for the scientific community.

Despite in its current state the client-server version is in working progress, we have provided clear technical steps on how this will be accomplished and facilitating technologies—both in terms of communications and visualization—have also been identified. In addition, interfacing with OpenCL is also a short-term goal, so that a parallelized multidevice version of `KomaMRI.jl` can also be provided.

## 7 Acknowledgements

This work has been funded in part by grant PID2020-115339RB-I00, Science and Innovation Ministry, Spain. The Social Council of the University of Valladolid should also be acknowledged for promoting student initiation in research activities.

## References

- [1] Z.-P. Liang y P. C. Lauterbur, *Principles of magnetic resonance imaging: a signal processing perspective*. “The” Institute of Electrical and Electronics Engineers Press, 2000.
- [2] D. Treceño-Fernández, J. Calabia-Del-Campo, M. L. Bote-Lorenzo, E. Gómez-Sánchez, R. de Luis-García, y C. Alberola-López, “A web-based educational magnetic resonance simulator: Design, implementation and testing,” *Journal of medical systems*, vol. 44, no. 1, 12 2019.
- [3] C. Castillo-Passi y P. Irarrazaval, “Komamri.jl,” sep 2021. Available: <https://doi.org/10.5281/zenodo.6627503>
- [4] J. Bezanson, A. Edelman, S. Karpinski, y V. B. Shah, “Julia: A fresh approach to numerical computing,” *SIAM review*, vol. 59, no. 1, pp. 65–98, 2017.
- [5] L. G. Hanson, “A graphical simulator for teaching basic and advanced mr imaging techniques,” *RadioGraphics*, vol. 27, no. 6, pp. e27–e27, 2007.
- [6] S. B. McKagan, K. K. Perkins, M. Dubson, C. Malley, S. Reid, R. LeMaster, y C. E. Wieman, “Developing and researching phet simulations for teaching quantum mechanics,” *American Journal of Physics*, vol. 76, no. 4, pp. 406–417, 2008.
- [7] T. Hackländer y H. Mertens, “Virtual mri: A pc-based simulation of a clinical mr scanner,” *Academic radiology*, vol. 12, no. 1, pp. 85–96, 2005.
- [8] G. Torheim, P. A. Rinck, R. A. Jones, y J. Kvaerness, “A simulator for teaching mr image contrast behavior,” *Magnetic Resonance Materials in Physics, Biology and Medicine*, vol. 2, no. 4, pp. 515–522, 1994.
- [9] T. Stöcker, K. Vahedipour, D. Pflugfelder, y N. J. Shah, “High-performance computing mri simulations,” *Magnetic resonance in medicine*, vol. 64, no. 1, pp. 186–193, 2010.
- [10] R. Kose y K. Kose, “Blochsolver: A gpu-optimized fast 3d mri simulator for experimentally compatible pulse sequences,” *Journal of Magnetic Resonance*, vol. 281, pp. 51–65, 2017.
- [11] C. G. Xanthis, I. E. Venetis, A. Chalkias, y A. H. Aletras, “Mrisimul: a gpu-based parallel approach to mri simulations,” *IEEE Transactions on Medical Imaging*, vol. 33, no. 3, pp. 607–617, 2013.
- [12] F. Liu, R. Kijowski, y W. F. Block, “MriLab: performing fast 3d parallel mri numerical simulation on a simple pc,” en *ISMRM Scientific Meeting & Exhibition*, vol. 2072, 2013.
- [13] T. Q. Company, “Qt documentation,” 2022. Available: <https://doc.qt.io>
- [14] A. Rossberg, *WebAssembly Specification*. WebAssembly Community Group, 2022. Available: <https://webassembly.github.io/spec/core>
- [15] W. Schroeder, K. Martin, L. Avila, y C. Law, *The VTK user’s guide*. Kitware, 2010.

# ARforIOERT: Una aplicación de Realidad Aumentada para el entrenamiento en la administración de Radioterapia Intraoperatoria

A. Sánchez López<sup>1</sup>, A. Pose Díez de la Lastra<sup>1</sup>, F. Calvo Manuel<sup>2</sup>, R.M. Meiriño Seoane<sup>2</sup>, J. Pascau González-Garzón<sup>1</sup>

<sup>1</sup> Departamento de Bioingeniería, Universidad Carlos III de Madrid, Madrid, España {[andrea.sanchezlopez@epfl.ch](mailto:andrea.sanchezlopez@epfl.ch), [apose@ing.uc3m.es](mailto:apose@ing.uc3m.es), [jpascau@ing.uc3m.es](mailto:jpascau@ing.uc3m.es)}

<sup>2</sup> Clínica Universidad de Navarra, Madrid, España {[fcavom@unav.es](mailto:fcavom@unav.es), [rmmeirino@gmail.com](mailto:rmmeirino@gmail.com)}

## Resumen

*La radioterapia intraoperatoria empieza a ser popular como tratamiento contra el cáncer en 1995. A diferencia de la radioterapia externa, este tratamiento dirige la radiación hacia el tejido tumoral sin ninguna atenuación por parte de otros tejidos. La radiación, invisible a los ojos, se dirige mediante el uso de aplicadores. De su correcto posicionamiento dependerá que tejidos sanos no reciban radiación. La distribución y profundidad de penetración de la radiación dependen tanto de la energía de irradiación como del diámetro y bisel del aplicador. En este estudio hemos desarrollado una aplicación para smartphones que utiliza realidad aumentada (AR) para mostrar dichos cambios en la radiación. Asimismo, proponemos la simulación de un entorno clínico con impresión 3D para que los radioterapeutas puedan practicar la irradiación de tumores en localizaciones complicadas.*

## 1. Introducción

La radioterapia se encuentra dentro de los tratamientos más comunes contra el cáncer [1]. La radioterapia intraoperatoria con electrones (IOERT) consiste en la irradiación del lecho quirúrgico durante la cirugía tras la resección del tumor, colimando el haz de electrones mediante un aplicador. De este modo, se maximiza la dosis sobre el lecho tumoral, limitando la irradiación de otros tejidos sanos [2]. Para conocer la distribución dosimétrica del haz de electrones, se realiza una medición en una cuba de agua desplazando un dosímetro por todo el volumen. Estas medidas se adquieren para aplicadores con distintos diámetros y bisels, así como para distintas energías de irradiación. Se obtiene de este modo una distribución volumétrica de la dosis, que se expresa como porcentaje de la dosis máxima depositada [3].

A pesar de aumentar la supervivencia, el tratamiento con IOERT también provoca efectos secundarios. En el cáncer de mama, un 71% de los pacientes presentaron necrosis grasa tras el tratamiento en el estudio [4]. Si se aumenta la precisión del tratamiento, maximizando la dosis en el tumor y disminuyendo la irradiación de tejido sano, se disminuye esa posible toxicidad. Este aumento de precisión se conseguiría con un mejor posicionamiento del aplicador [5].

La realidad aumentada (AR) y la impresión 3D son técnicas introducidas recientemente en el ámbito de la salud. La AR permite la superposición de modelos tridimensionales virtuales sobre el mundo real. El usuario

se siente completamente integrado en este entorno mixto, ya que esta combinación entre mundo real y virtual ofrece una sensación de profundidad y percepción que no existe en un entorno completamente virtual [8, 9]. La impresión 3D consiste en la fabricación de objetos de forma aditiva, capa a capa. Permite conseguir una customización de la producción de objetos a bajo coste [8].

La combinación de ambas tecnologías está teniendo resultados prometedores. Su uso en [9], propone un flujo de trabajo innovador para oncología ortopédica. Se proyecta mediante AR el modelo del tumor virtualmente, de forma que los cirujanos pueden verificar la resección completa, evitando dejar márgenes positivos y disminuyendo de este modo la posibilidad de recurrencia local en el paciente. Esta solución también permite a los cirujanos mostrar a los pacientes su patología antes de la cirugía de resección del tumor, permitiendo una mejor comprensión del procedimiento. El registro entre el mundo virtual y el real se realiza mediante una guía quirúrgica impresa en 3D que encaja únicamente en una posición específica sobre el hueso cercano al tumor. Esta incluye un marcador de referencia para registrar la escena de AR.

Inspirado en estos resultados, este trabajo presenta una aplicación para teléfonos inteligentes que permitiría mejorar la precisión en la IOERT. Esta mejora se conseguiría a través de un sistema de entrenamiento que guíe la correcta colocación del aplicador. Específicamente, la aplicación muestra al usuario cómo varía la distribución de la radiación al modificar parámetros como la energía del haz, el diámetro o el bisel del aplicador. Dado que los tumores son únicos para cada paciente, se utiliza la impresión 3D para conseguir la personalización necesaria. Por otro lado, la AR permite representar la radiación de forma virtual, de forma que el radioterapeuta puede aprender a evitar la irradiación innecesaria de tejido sano durante el entrenamiento, manteniendo la percepción de realidad. El sistema se ha evaluado en tres casos tratados con IOERT, mostrando cómo se mejora la planificación del tratamiento cuando se emplea la solución planteada.

## 2. Materiales y métodos

### 2.1. Selección de pacientes

En este estudio se han incluido 3 casos. Todos ellos corresponden a pacientes para los que la IOERT fue elegida como tratamiento más adecuado. Con el fin de

mantener el anonimato, a cada paciente se le asignó un código alfanumérico consistente en las siglas “ARIORT” seguidas de un número de cuatro dígitos. Para evaluar el beneficio de la aplicación (ARforIOERT), se seleccionaron pacientes que presentaban tumores en distintas localizaciones cercanas a estructuras vitales (como las venas hepáticas). En estos tejidos sanos la radiación debe ser siempre la mínima posible. Específicamente, los tumores incluidos se encontraban en cavidad pélvica (paciente ARIORT0001), páncreas (paciente ARIORT0002) y sobre el hueso maxilar (paciente ARIORT0003).

## 2.2. Procesamiento de los datos

Antes de cada intervención quirúrgica, se adquirieron estudios de tomografía axial computarizada (TAC) y tomografía de emisión de positrones (PET) para todos los pacientes. El TAC permite mostrar con buena resolución la anatomía de los órganos cercanos al tumor. Por otro lado, el PET, marcado con fluorodesoxiglucosa (FDG), ofrece información más detallada sobre la localización y extensión del tumor gracias a la elevada actividad metabólica del mismo. Para los pacientes ARIORT0002 y ARIORT0003 se adquirió además una resonancia magnética (RM) con contraste para mostrar la anatomía de los vasos sanguíneos cercanos al tumor. Se realizó un registro rígido automático con *Siemens Syngo.via* entre los distintos estudios de imagen utilizando como función de coste la diferencia de intensidad.

## 2.3. Segmentación

Se realizó una segmentación de los órganos sanos de interés y del tumor utilizando el software *3D Slicer* (versión 4.10.2) [10]. Para el paciente ARIORT0001 se segmentó la pelvis. En el caso del paciente ARIORT0002, los órganos segmentados fueron el páncreas, el duodeno, las venas hepáticas y la región de la columna vertebral más próxima al tumor. Para el paciente ARIORT0003 se segmentaron ambos ojos y sus correspondientes nervios ópticos, el cerebro, la mandíbula y las arterias carótidas. La segmentación de los tumores se realizó sobre el PET, la de los vasos sanguíneos se basó en la RM y la del resto de órganos se hizo a partir del TAC. La segmentación se realizó de forma semiautomática en los tres casos utilizando diferentes herramientas del módulo “Segment Editor” de *3D Slicer*. En una primera fase, se obtuvo un modelo inicial de cada órgano realizando una segmentación por umbral (herramienta “Thresholding”). Posteriormente, se eliminaron los grupos de píxeles que no eran adyacentes al órgano de interés (herramienta “Islands”). Las segmentaciones se repasaron siguiendo las indicaciones de los radiólogos para la identificación de las estructuras, y se perfilaron de forma manual (herramientas de “Paint” y “Erase”).

Para poder representar la radiación como modelos virtuales en AR, se emplearon medidas dosimétricas realizadas en cuba de agua para los aplicadores empleados en los tres casos: aplicador de diámetro 6 cm y bisel de 30° (pacientes ARIORT0001 y ARIORT0002) y aplicador de diámetro 5 cm y bisel de 15° (paciente ARIORT0003). A partir de estos datos, se utilizó *3D Slicer* para realizar una

segmentación por umbral y crear cinco modelos 3D. Estos correspondían al volumen entre los porcentajes de dosis máxima estándares (30%, 50%, 70%, 90% y 110%) [11].

## 2.4. Procesamiento de los modelos e impresión 3D

Para posicionar los modelos virtuales de los órganos y el tumor en AR, imprimimos en 3D el órgano más cercano al tumor de los previamente segmentados. Se incluyó un soporte y ocho marcadores en los modelos impresos para poder realizar la fase de validación descrita más adelante. La Figura 1 muestra los maniqués resultantes y sus dimensiones.

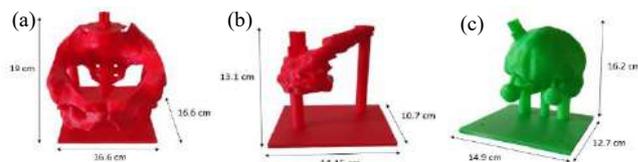


Figura 1. Modelos impresos para cada paciente: (a) ARIORT0001, (b) ARIORT0002 y (c) ARIORT0003.

Por otro lado, diseñamos dos modelos de aplicadores basados en los empleados en estos tres casos. Sobre los diseños añadimos un puntero y un soporte triangular para la validación del sistema mediante posicionamiento óptico. Además, se incluyó un pequeño soporte para colocar un marcador de AR. Este marcador se emplea para el registro entre el mundo real y el virtual, permitiendo posicionar los modelos virtuales de las curvas de radiación en la posición correcta. Este proceso se realizó con el software *Autodesk MeshMixer* (Autodesk Inc., San Rafael, CA, USA). Finalmente se imprimieron en 3D ambos modelos. La Figura 2 muestra uno de los aplicadores impresos.

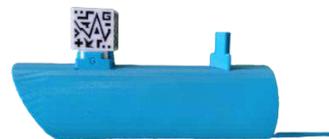


Figura 2. Aplicador de diámetro 6 cm y bisel de 30° usado para los pacientes ARIORT0001 y ARIORT0002. El marcador de AR aparece ya introducido en el soporte cuadrado.

Todos los procesos de impresión 3D se realizaron con la técnica de modelado por deposición fundida (FDM) usando como material ácido poliláctico (PLA) en impresoras *Ultimaker 3 Extended* (Ultimaker B.V., Utrecht, The Netherlands) y *Raise3D Pro 2* (Raise 3D Technologies Inc., Irvine, California, USA).

## 2.5. Desarrollo de la aplicación de AR

Desarrollamos ARforIOERT con *Unity* (versión 2020.3). Este software permite la creación de aplicaciones de AR para móviles compatibles con iOS y Android. El registro entre el mundo real y el virtual se realizó de forma automática mediante reconocimiento de patrones. Para ello, utilizamos *Vuforia* (Parametric Technology Corporation Inc., Boston, MA, USA) en dos modalidades distintas: reconocimiento de marcadores y reconocimiento de modelos. En el reconocimiento de marcadores, los patrones se extraen de un archivo de imagen 2D. En el reconocimiento de modelos, los patrones se extraen de un archivo de imagen 3D.

En ARforIOERT, usamos el reconocimiento de modelos con los archivos de imagen 3D de los modelos impresos para cada paciente. Esto permitió evitar el uso de un marcador para estos modelos y trabajar a unas distancias de reconocimiento mayores que con patrones 3D. Al evitar el uso de un marcador, el modelo impreso se puede reconocer desde distintos ángulos, permitiendo obtener una mayor información sobre la localización del tumor y los órganos virtuales. Esta solución permite también cubrir parcialmente el modelo impreso (lo que resulta necesario para simular la irradiación con el aplicador).

Para el reconocimiento del aplicador utilizamos un marcador cúbico de tamaño 3 x 3 x 3 cm. Este marcador se imprimió en 3D en PLA con la impresora *Ultimaker 3 Extended* (Figura 2). Cada una de las caras del marcador tiene un diseño único que *Vuforia* puede detectar, independientemente de su orientación o posición al mover el aplicador. Los patrones se extrajeron de imágenes 2D de cada una de las caras del marcador.

Las coordenadas relativas entre el modelo impreso y los órganos virtuales se conocían gracias al registro entre las distintas modalidades de imagen. En el caso de los aplicadores y las curvas de radiación, planteamos las coordenadas relativas entre el marcador y los modelos de radiación en función de sus dimensiones. Este procedimiento se llevó a cabo utilizando *3D Slicer* siguiendo el protocolo descrito en [12]. Una vez desarrollado este sistema, en cuanto ARforIOERT detecta los patrones correspondientes, proyecta automáticamente, y en tiempo real, todos los modelos virtuales.

## 2.6. Uso de la aplicación

ARforIOERT comienza con un menú inicial en el que el usuario puede seleccionar el paciente (Figura 3). Tras la selección inicial, se muestra el mundo a través de la cámara del dispositivo, y el contorno del órgano impreso en primer plano. Esto sirve para guiar al usuario en la colocación del móvil para el reconocimiento inicial del maniquí (Figura 4a). Tras el reconocimiento, se muestran los órganos y el tumor virtuales de manera automática. Se puede controlar la visualización de dichos modelos mediante botones con el nombre de cada uno de los órganos virtuales proyectados, tal y como se muestra en la Figura 4b.

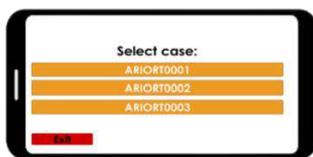


Figura 3. Menú inicial de la aplicación ARforIOERT.

Para el reconocimiento del aplicador, el usuario debe enfocar al marcador de AR. Una vez detectado, un marco de color verde rodea la pieza, sirviendo de guía visual para garantizar el seguimiento adecuado. Si en algún momento la detección se interrumpe, el marco pasa a ser de color rojo. Una vez detectado, en la esquina superior izquierda se muestran botones para elegir la energía de irradiación. La visualización de los modelos de irradiación puede cambiarse con una barra deslizadora situada en la parte inferior de la pantalla (Figura 5).

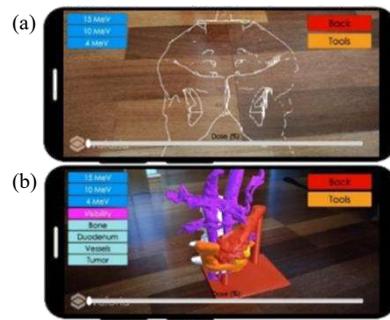


Figura 4. (a) ARforIOERT con la guía para el reconocimiento inicial en el paciente ARIORT0001. (b) ARforIOERT con los botones para controlar la visualización de los órganos virtuales en el paciente ARIORT0002.



Figura 5. Detección del marcador (con un marco verde) y proyección de los modelos de radiación en el aplicador.

Otras funcionalidades de ARforIOERT permiten cambiar la visualización de los modelos de radiación a 2D, visualizar los órganos virtuales tras los órganos impresos y ver la interacción entre la radiación, los órganos virtuales y el tumor. Esta interacción oculta la parte de los órganos virtuales y del tumor que recibe la radiación.

## 2.7. Validación

Para validar la solución de AR, comparamos el porcentaje de tumor y de tejido sano irradiado con la aplicación y sin ella, suponiendo que en cada caso el objetivo clínico es irradiar el máximo tejido tumoral protegiendo los órganos circundantes. Como referencia se empleó el sistema de posicionamiento óptico Polaris Spectra (Northern Digital Inc., Waterloo, ON, Canada). En el soporte triangular añadido a los modelos impresos y al aplicador, colocamos una estructura con esferas retroreflectantes. El sistema de posicionamiento recogía la posición y orientación de los modelos reales en base a la localización de esas esferas. Esta información se reprodujo en modelos virtuales análogos en *3D Slicer*. De este modo, es posible ver la porción de tejido irradiada en función de sus ubicaciones relativas. Para ello fue necesario un registro entre el mundo real (las posiciones recibidas por el sistema de posicionamiento óptico) y el mundo virtual de *3D Slicer*. Este registro se calcula seleccionando los marcadores de los modelos impresos con el puntero añadido al aplicador.

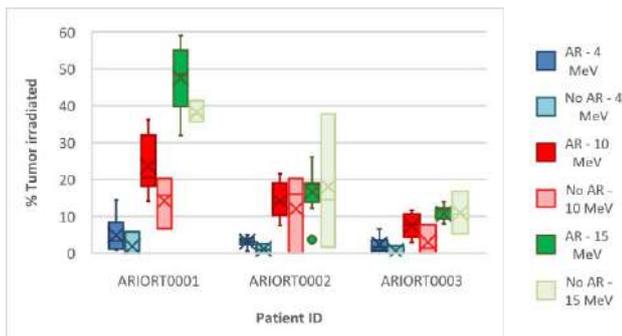
Tres usuarios no experimentados participaron en el experimento. Cada uno de ellos realizó cuatro repeticiones. En la repetición inicial, a los usuarios se les mostró brevemente la localización del tumor, pero no se les indicó la profundidad de la radiación (como ocurre en la situación quirúrgica). En las siguientes repeticiones, usaron ARforIOERT como guía.

## 3. Resultados

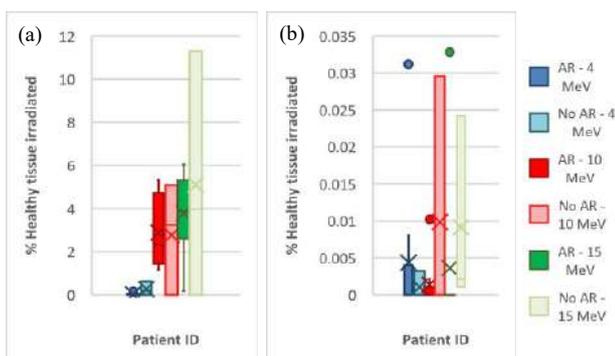
La aplicación registra el mundo real y virtual prácticamente de manera instantánea. El reconocimiento inicial del

órgano impreso y del marcador requiere una distancia de aproximadamente 30 cm hasta la cámara del móvil. Esta detección se mantiene hasta una distancia aproximada de 1,7 m en el caso del órgano impreso y 50 cm en el caso del marcador de AR.

Las Figura 6 y 7 muestran el porcentaje de tumor o tejido sano irradiado para cada paciente según el método de guiado empleado (con AR o sin ella) para diferentes energías de haz. Para el paciente ARIORT0001, se irradió siempre un mayor porcentaje de tumor utilizando ARforIOERT. Para los pacientes ARIORT0002 y ARIORT0003, solo en el caso de la energía de irradiación de 15 MeV se irradió más porcentaje de tumor sin ARforIOERT. No obstante, hay una mayor dispersión de los resultados cuando no se empleó guiado con AR. En el caso del paciente ARIORT0002, esto se debe a que el tumor se encontraba dentro del páncreas. Al emplear AR se consiguió mantener la irradiación en el tumor disminuyendo la dosis sobre el tejido sano. En el caso del paciente ARIORT0003, la diferencia en cuanto a porcentaje de tumor irradiado fue de un 0.2% de media. Sin embargo, se consiguió evitar la irradiación de tejido sano (bajó de un casi 0.1% de tejido irradiado a un 0%) con AR. En concreto, se evitó irradiar el ojo izquierdo, lo que evitaría una posible distrofia y necrosis corneal [13].



**Figura 6.** Porcentaje de tumor irradiado para cada paciente, energía de haz y método de guiado (sin/con AR).



**Figura 7.** Porcentaje de irradiación de tejido sano según el método de guiado para los pacientes (a) ARIORT0002 y (b) ARIORT0003. Los puntos representan los valores atípicos de la distribución.

#### 4. Conclusiones

En este estudio, valoramos el empleo de AR en una aplicación de smartphone para entrenar a los cirujanos en procedimientos de IOERT. Los resultados obtenidos en el

escenario experimental demuestran una mayor irradiación del tumor y una menor dosis en el tejido sano en la mayoría de los casos cuando se guía el procedimiento con AR. Además, la solución presentada permite a los cirujanos tomar decisiones para compensar la irradiación del tumor y del tejido sano. Como trabajo futuro, se ha propuesto la validación de la aplicación con usuarios experimentados, así como la inclusión de nuevos casos clínicos para ampliar las localizaciones anatómicas.

#### Agradecimientos

Este trabajo ha sido financiado mediante los proyectos PI18/01625 y AC20/00102-3, Era Permed PerPlanRT (Ministerio de Ciencia, Innovación y Universidades, Instituto de Salud Carlos III, Asociación Española Contra el Cáncer y Fondo Europeo de Desarrollo Regional “Una manera de hacer Europa”).

#### Referencias

- [1] Página web del Instituto Nacional contra el Cáncer (NCI). <https://www.cancer.gov/about-cancer/treatment> (Consultada: Junio 2022).
- [2] Pilar A, Gupta M, Ghosh Laskar S, Laskar S. Intraoperative radiotherapy: review of techniques and results. *Ecancermedicalscience*, vol. 11, p. 750 (ISSN: 1754-6605).
- [3] Grant D. G. Radiation Therapy Physics. *Johns Hopkins APL Technical Digest*, vol 4, no 1, 1983 (ISSN: 0270-5214).
- [4] Lemanski C et al. Electrons for intraoperative radiotherapy in selected breast-cancer patients: late results of the Montpellier phase II trial. *Radiation Oncology*, vol. 8, no 1, 2013, p. 191 (ISSN: 1748-717X).
- [5] Hensley F.W. Present state and issues in IORT Physics. *Radiation Oncology*, vol. 12, no 1, 2017, p. 37 (ISSN: 1748-717X).
- [6] Salamin P, Thalmann D, Vexo F. The benefits of third-person perspective in virtual and augmented reality? *Proceedings of the ACM symposium on Virtual reality software and technology - VRST '06*, Limassol, Cyprus, 2006, p. 27 (ISBN: 978-1-59593-321-8).
- [7] Berning M, Kleinert D, Riedel T, Beigl M. A study of depth perception in hand-held augmented reality using autostereoscopic displays. *2014 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, Munich, 2014, pp. 93–98 (ISBN: 978-1-4799-6184-9).
- [8] Schubert C, van Langeveld M. C, Donoso L. A. Innovations in 3D printing: a 3D overview from optics to organs. *Br J Ophthalmol*, vol. 98, no. 2, 2014, pp. 159–161 (ISSN: 1468-2079).
- [9] Moreta-Martinez R, Pose-Díez-de-la-Lastra A, Calvo-Haro J. A, Mediavilla-Santos L, Pérez-Mañanes R, Pascau J. Combining Augmented Reality and 3D Printing to Improve Surgical Workflows in Orthopedic Oncology: Smartphone Application and Clinical Evaluation. *Sensors (Basel)*, vol. 21, no. 4, 2021, p. 1370 (ISSN: 1424-8220).
- [10] Fedorov A et al. 3D Slicer as an Image Computing Platform for the Quantitative Imaging Network. *Magn Reson Imaging*, vol. 30, no. 9, 2012, pp. 1323–1341 (ISSN: 0730-725X).
- [11] Johansson S. A et al. Dosimeter Gel and MR Imaging for Verification of Calculated Dose Distributions in Clinical Radiation Therapy. *Acta oncologica*, vol. 36, no 3, 1997, pp. 283–90 (ISSN: 1651-226X).
- [12] Moreta-Martinez R et al. Combining Augmented Reality and 3D Printing to Display Patient Models on a Smartphone. *J. Vis. Exp.*, no 155, 2020 (ISSN: 1940-087X).
- [13] Blázquez Sánchez V. Efectos secundarios de la radioterapia a nivel ocular. *Gaceta óptica*, no 453, 2010, pp. 14-18 (ISSN: 0210-5284).

# Aprendizaje profundo para la segmentación de ventilaciones en impedancia durante la resucitación cardiopulmonar

X. Jaureguibeitia Lara<sup>1</sup>, E. Aramendi Ecenarro<sup>1</sup>, H.E. Wang<sup>2</sup>, A.H. Idris<sup>3</sup>

<sup>1</sup> BioRes, Dpto. de Ingeniería de Comunicaciones, Universidad del País Vasco (UPV/EHU), Bilbao, España, {xabier.jaureguibeitia, elisabete.aramendi}@ehu.eus

<sup>2</sup> Wexner Medical Center, Ohio State University, Columbus, Ohio, EEUU, Henry.Wang@osumc.edu

<sup>3</sup> University of Texas Southwestern Medical Center, Dallas, Texas, EEUU, ahamed.idris@utsouthwestern.edu

## Resumen

*Una resucitación cardiopulmonar (RCP) de calidad es esencial para mejorar la supervivencia a la parada cardiorrespiratoria extra-hospitalaria (PCREH), pero poco se sabe sobre la terapia óptima de ventilación. La impedancia torácica (IT), medida a través de los parches de desfibrilación, permite monitorizar las ventilaciones de cara al análisis retrospectivo de grandes registros de PCREH, pero presenta artefactos debidos a las compresiones torácicas y al movimiento de los electrodos. El objetivo de este trabajo fue evaluar si las arquitecturas de aprendizaje profundo podrían ofrecer ventajas respecto de los métodos disponibles para la detección de ventilaciones en IT con compresiones concurrentes. Se analizó un total de 367 episodios de PCREH, y se anotaron 20724 ventilaciones de referencia en capnografía. Se implementó una arquitectura de segmentación U-Net a la que se alimentó con segmentos no filtrados de IT de un minuto de duración. El desempeño de la red se evaluó usando una estrategia de validación cruzada. Las métricas de desempeño globales fueron de un 79,8% de sensibilidad, un 83,4% de valor predictivo positivo, y un 81,6% de valor F1, superando a soluciones anteriores propuestas en la literatura, aunque con una variabilidad importante entre casos/pacientes. Para las ventilaciones correctamente detectadas, los instantes de inspiración y expiración anotados en el capnograma se reprodujeron con un error medio (desviación típica) de 0,02s (0,51s) y -0,002s (0,43s), respectivamente.*

## 1. Motivación

La parada cardiorrespiratoria extra-hospitalaria (PCREH) constituye un problema de salud pública de primer orden, con en torno a 56 casos atendidos por 100.000 habitantes y año, y una tasa de supervivencia inferior al 10% [1]. Un paciente en PCREH pierde, generalmente de forma súbita, la circulación y respiración espontáneas, lo que desemboca en la muerte en ausencia de tratamiento. La resucitación cardiopulmonar (RCP), consistente en compresiones torácicas y ventilaciones, suministra un flujo mínimo de sangre y oxígeno al paciente, y es clave para mejorar las opciones de supervivencia a la PCREH [2]. En ese sentido, se han destinado importantes esfuerzos a mejorar la calidad de la RCP: los parámetros óptimos de compresión han sido estudiados extensamente [3], y se han desarrollado también nuevas tecnologías tales como dispositivos autónomos de compresión mecánica. La mayoría de avances han estado, sin embargo, centrados en las compresiones torácicas; la tecnología actual para monitorizar la ventilación durante la PCREH es limitada, y poco se sabe aún sobre su aplicación

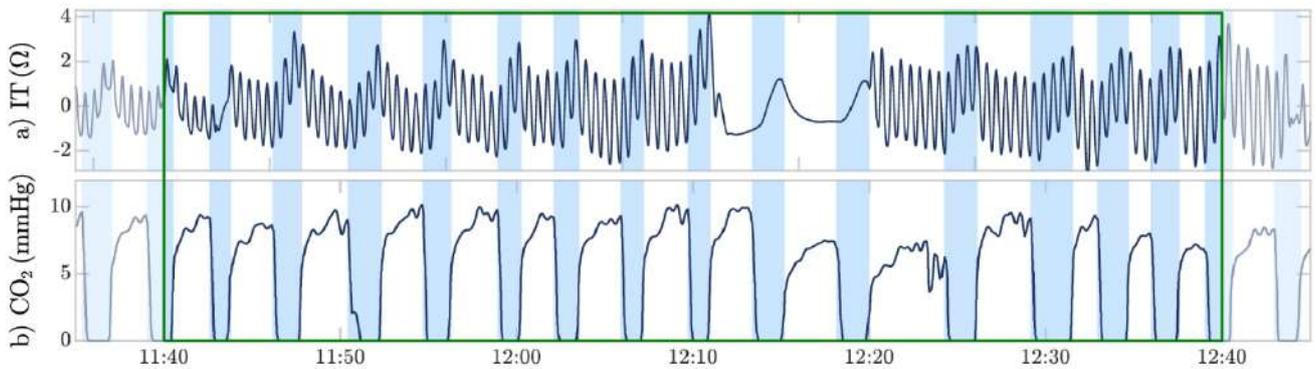
óptima [4]. Son necesarias nuevas técnicas y tecnologías que permitan un mejor análisis y una mayor comprensión del efecto de la ventilación durante la PCREH.

La capnografía, que monitoriza la concentración de CO<sub>2</sub> en los gases expirados, es la técnica habitual para evaluar la ventilación en la PCREH [5]. Sin embargo, ésta no suele estar disponible hasta fases avanzadas de la resucitación, tras la inserción de una vía aérea. La impedancia torácica (IT), que mide la resistencia del cuerpo al paso de corriente, es sensible a cambios de volumen de aire en los pulmones, y supone una alternativa para la detección de ventilaciones. La IT se adquiere junto al electrocardiograma a través de los parches de desfibrilación, siendo así una de las primeras señales en estar disponible durante la PCREH. Además, la amplitud de las formas de ventilación se relaciona con el volumen de aire insuflado [6], una información no disponible mediante las técnicas habituales de capnografía extra-hospitalaria. La detección de ventilaciones en IT no está, sin embargo, exenta de problemáticas: las formas de onda de ventilación son heterogéneas [7], y la propia señal es susceptible de ser contaminada por diversas fuentes de ruido, como el movimiento de electrodos [8]. Más aún, la IT es sensible a las compresiones administradas como parte de la RCP, lo que resulta en un artefacto que enmascara las ventilaciones. La supresión de este artefacto requiere el uso de filtros muy exigentes [9], con la consiguiente pérdida de información de ventilación, o de filtros adaptativos en base a señales adicionales [7,10], no siempre disponibles.

En los últimos años, las técnicas de aprendizaje profundo han obtenido resultados prometedores en diversos campos de la medicina, incluyendo la PCREH [11]. En esta línea, el objetivo de este trabajo fue evaluar si dichas técnicas podrían superar a los métodos actuales en la identificación de ventilaciones en IT durante compresiones concurrentes.

## 2. Materiales

Se analizaron los registros electrónicos de 367 pacientes de PCREH, tratados por los servicios de emergencia de Dallas - Fort Worth (Texas, EEUU) entre 2016 y 2017, y adscritos al ensayo clínico PART (Pragmatic Airway Resuscitation Trial) [12]. Todos los registros fueron adquiridos usando un monitor-desfibrilador HeartStart MRx (Philips Medical Systems, Andover, MA, EEUU), e incluían grabaciones de IT (200Hz, resolución de 2,5mΩ) y capnografía, así como de fuerza, aceleración y profundidad de compresión.



**Figura 1.** Segmento de un minuto de impedancia torácica (IT) y su capnograma asociado. Las ventilaciones de referencia, anotadas en la capnografía y correspondientes a la fase de inspiración, se muestran en azul. Ambas señales se alinearon temporalmente haciendo coincidir los flancos de subida del capnograma con los picos de las fluctuaciones de IT.

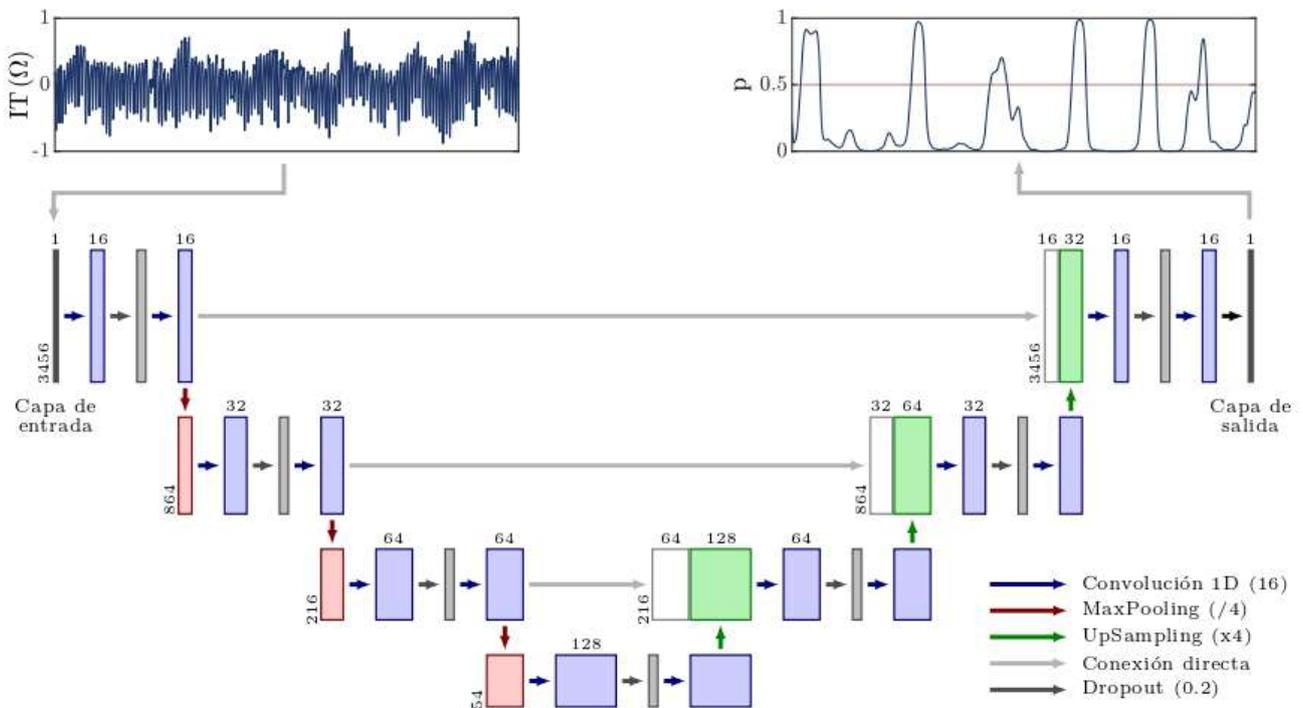
Los registros se procesaron y analizaron empleando Matlab (MathWorks Inc., Natick, MA, EEUU), y se seleccionaron intervalos de señal apropiados para el estudio. Se extrajeron intervalos durante compresiones torácicas manuales, y con concurrencia de todas las señales previamente citadas. Se descartaron pausas de compresiones superiores a 20s, así como excursiones abruptas en la IT y otros artefactos dominantes aislados. La capnografía se utilizó para anotar las ventilaciones de referencia para el entrenamiento y la evaluación de la red. Se corrigió también de forma manual el retardo temporal de la capnografía respecto de la IT (ver Figura 1). Las señales de fuerza, aceleración y profundidad de compresión se incluyeron exclusivamente para la implementación de soluciones similares en la literatura [7, 9, 10].

Los intervalos de señal seleccionados se subdividieron en segmentos no superpuestos de un minuto, sobre los que se ejecutaron los diferentes algoritmos. A fin de facilitar la detección de ventilaciones parcialmente incluidas en los segmentos, se incluyó también un relleno no evaluable de aproximadamente 4,5s en cada extremo de los mismos. La base de datos final incluyó un total de 2551 segmentos de un minuto y 20724 ventilaciones de referencia, el 97.1% de las mismas durante compresiones torácicas.

### 3. Métodos

#### 3.1. Preprocesado de señal

Se apostó por un preprocesado mínimo, evitando el uso de señales adicionales y delegando sobre la red de aprendizaje



**Figura 2.** Arquitectura de la red U-Net implementada. Los valores numéricos sobre cada capa indican el número de filtros o canales a la salida de la misma. Los valores en la parte inferior derecha de cada capa/nivel indican el número de muestras por canal a la salida de la misma. La red se alimenta con segmentos de un minuto de impedancia (IT), y devuelve por cada muestra de señal la probabilidad  $p$  de que esta corresponda a una ventilación.

profundo el tratamiento del artefacto de compresiones. Los segmentos de IT se remuestrearon a 50Hz, y se eliminó su componente continua mediante un filtrado digital paso alto de fase de cero (Butterworth, orden 4, frecuencia de corte de 0,06Hz).

### 3.2. Arquitectura de aprendizaje profundo

Se escogió una arquitectura de segmentación U-Net [13], adaptada para operar sobre señales unidimensionales. Una red U-Net sigue un esquema codificador-decodificador, en el que se incluyen conexiones directas entre niveles pares de codificación y decodificación. Esto permite reconstruir el nivel de detalle perdido durante la etapa de contracción o codificación, al tiempo que se mantiene su contribución en términos de contexto. La red implementada (ver Figura 2) constó de cuatro niveles, con dos capas de convolución por nivel. A fin de reducir el número de pesos entrenables, se emplearon convoluciones separables, todas con filtros de 16 muestras y función de activación rectificadora lineal (ReLU). Las capas de deconvolución del diseño original se sustituyeron también por capas de sobremuestreo o *upsampling*, sin parámetros entrenables. Dada la limitación de los datos y el considerable tamaño de la red, se incluyó en cada nivel una capa de olvido o *dropout* (20%) para combatir el sobreajuste. La red se alimentó con los segmentos de un minuto de IT, incluyendo 228 muestras ( $\sim 4,5s$ ) de relleno de señal, no evaluables, en cada extremo. La capa de salida consistió en una convolución puntual con activación sigmoide, produciendo una secuencia de valores de probabilidad  $p$  con tantas muestras como la entrada.

La optimización de la red se llevó a cabo en 50 épocas, con tandas de 64 segmentos de IT. Se utilizó como función de coste la entropía cruzada binaria, y un optimizador Adam con tasa de aprendizaje de 0,001. La implementación y la optimización de la red se realizaron en Tensorflow 2.8.

### 3.3. Evaluación como detector

El desempeño de la red de segmentación se evaluó usando una estrategia de validación cruzada de cinco grupos. Los segmentos de IT se asignaron a los distintos grupos en base a paciente, a fin de mantener separados segmentos con similitudes morfológicas. La asignación se realizó también de forma cuasi-estratificada, de modo que cada grupo recogiese un quinto (tolerancia máxima del 5%) de pacientes, segmentos y ventilaciones de referencia.

Dada la salida de la red, se consideró como una ventilación potencial cualquier intervalo continuo con valores  $p \geq 0,5$ . Estas detecciones se evaluaron como verdaderos positivos (VP) o falsos positivos (FP) en base a las anotaciones de

referencia realizadas en el capnograma. Se tomó como instante de detección el instante con mayor valor de  $p$  dentro de cada intervalo, y se consideró VP si quedaba comprendido dentro de una ventilación de referencia, con un margen de tolerancia de 0,5s. Este margen se redujo, de ser necesario, para evitar el solape de las ventilaciones de referencia. En caso de que varias detecciones calificasen para la misma ventilación de referencia, aquella más cercana a su punto medio se consideró VP, y el resto FP.

El desempeño se evaluó en términos de sensibilidad (Se), valor predictivo positivo (VPP), y valor  $F_1$ , dados por:

$$Se = \frac{VP}{N_{ref}} \quad VPP = \frac{VP}{VP + FP} \quad F_1 = 2 \frac{Se \cdot VPP}{Se + VPP}$$

siendo  $N_{ref}$  el número de ventilaciones de referencia. Las métricas se calcularon para toda la base de datos, así como para cada paciente individual. Las métricas por paciente se caracterizaron en términos de mediana (rango intercuartil). El desempeño se comparó con el de varias soluciones de la literatura [7, 9, 10], las cuales fueron implementadas en Matlab y evaluadas contra los datos de estudio.

### 3.4. Evaluación como segmentador

Para cada detección evaluada como correcta, se estableció el comienzo de la inspiración  $t_{insp}$  como el instante inmediatamente anterior con salida de la red  $p < p_{insp}$ , o como el instante con  $p$  mínima desde la ventilación precedente, lo que resultase más cercano. El valor  $p_{insp}$  se obtuvo usando como referencia las predicciones de la red sobre los datos de entrenamiento. Se escogió  $p_{insp}$  tal que se minimizara el error cuadrático medio entre los instantes  $t_{insp}$  y el inicio de la referencia en capnografía asociada (flanco de bajada). Se siguió un procedimiento análogo para determinar el valor frontera  $p_{exp}$  y establecer los inicios de expiración  $t_{exp}$ .

Se evaluaron, sobre las detecciones correctas en los datos de testeo, el error medio (desviación típica) de los instantes  $t_{insp}$  y  $t_{exp}$ , así como de la duración total de la inspiración, respecto de las referencias en capnografía asociadas.

## 4. Resultados

Tal como muestra la Tabla 1, la red implementada superó, en desempeño como detector, a las soluciones previas en la literatura, incluidas aquellas que requerían de señales adicionales en la etapa de filtrado [7, 10]. Las soluciones por Risdal et al. y Alonso et al. obtuvieron una mejor Se, pero a costa de reportar un elevado número de FP (VPP bajo), resultando en un valor  $F_1$  muy inferior. Esto se atribuyó principalmente a una falta de contexto para discriminar pequeñas fluctuaciones debidas al ruido de ventilaciones más

Algoritmo	Datos completos			Por paciente		
	Se (%)	VPP (%)	$F_1$ (%)	Se (%)	VPP (%)	$F_1$ (%)
U-Net	79.8	83.4	81.6	83.1 (65.3-94.2)	84.9 (71.4-95.0)	82.2 (66.7-93.3)
Risdal et al. [7]	80.8	62.0	70.2	87.3 (71.7-95.8)	65.2 (48.1-82.1)	70.0 (56.3-82.8)
Alonso et al. [9]	89.0	59.1	71.1	92.5 (84.6-97.1)	60.2 (43.4-79.5)	68.6 (56.1-85.5)
Jaureguibeitia et al. [10]	76.3	82.4	79.2	79.7 (58.6-91.3)	83.3 (69.1-93.8)	80.5 (63.9-91.0)

**Tabla 1.** Métricas de desempeño de la arquitectura U-Net y de soluciones previas en la literatura.

prominentes en el mismo segmento. La solución previa por Jaureguibeitia et al., que incluía un algoritmo de aprendizaje máquina y varias características de contexto, ofreció mejores resultados, pero inferiores a la U-Net propuesta pese a requerir de detecciones precisas de los instantes de compresión medidas en la profundidad de compresión.

Resulta importante observar que, con independencia del algoritmo empleado, existió entre pacientes una variabilidad considerable de las métricas de desempeño. Algunos segmentos de IT, e incluso registros completos, podrían no ser aptos para la identificación de ventilaciones. En este sentido, el enfoque alternativo ofrecido por el aprendizaje profundo podría servir para analizar relaciones de concordancia con los algoritmos clásicos y desarrollar índices de calidad de señal para la IT [8].

En lo referente a la segmentación, el procedimiento empleado obtuvo distribuciones de error muy centradas, con errores medios marginales, pero desviaciones importantes. Los errores medios para  $t_{\text{insp}}$  y  $t_{\text{exp}}$  fueron de  $-0,002\text{s}$  (0,43s) y  $0,02\text{s}$  (0,51s), respectivamente. El error medio en la duración de la inspiración fue de  $-0,02\text{s}$  (0,60s).

## 5. Conclusiones

La red U-Net propuesta mostró resultados prometedores de cara al uso de aprendizaje profundo para la segmentación de ventilaciones en la señal de IT. Se mejoró el desempeño de soluciones previas en la literatura, sin utilizar, además, señales adicionales para la cancelación del artefacto de compresiones. Esto representa una mejora significativa en cuanto a su aplicabilidad a un mayor rango de escenarios. Alternativamente, la red podría también alimentarse con dichas señales para una mejora potencial de los resultados. La red es también susceptible a otras mejoras. Por ejemplo, la frecuencia de muestro de la IT (50Hz) podría reducirse, reduciendo el uso de memoria y permitiendo acortar los filtros para una menor carga computacional. Si bien el uso de la red actual parece limitado al análisis retrospectivo de registros de PCREH, una red suficientemente ligera podría integrarse en el equipamiento de campo para proporcionar realimentación en tiempo real. Una red más ligera sería también menos susceptible a un potencial sobreajuste, algo que podría lograrse también mediante técnicas de aumento de datos, no empleadas en este estudio. Finalmente, cabría explorar diferentes arquitecturas, incluyendo algunas de las múltiples variantes de U-Net, pero también otros enfoques radicalmente diferentes que podrían adaptarse mejor a la segmentación de ventilaciones.

## Agradecimientos

Este trabajo ha recibido ayuda financiera del Ministerio de Ciencia, Innovación y Universidades, proyecto RTI2018-101475-BI00, junto con el Fondo Europeo de Desarrollo Regional (FEDER), así como del Gobierno Vasco a través de la subvención a grupos de investigación IT-1229-19 y la beca predoctoral PRE-2021-2-0126, y de la Universidad del País Vasco a través del proyecto COLAB20/01.

## Referencias

- [1] Gräsner JT, Wnent J, Herlitz J, et al. Survival after out-of-hospital cardiac arrest in Europe – Results of the EuReCa TWO study. *Resuscitation*, vol 148, 2020, pp 218-226.
- [2] Cummins RO, Eisenberg MS, Hallstrom AP, et al. Survival of out-of-hospital cardiac arrest with early initiation of cardiopulmonary resuscitation. *The American Journal of Emergency Medicine*, vol 3, 1985, pp 114-119.
- [3] Panchal AR, Bartos JA, Cabañas JG, et al. Part 3: Adult Basic and Advanced Life Support: 2020 American Heart Association Guidelines for Cardiopulmonary Resuscitation and Emergency Cardiovascular Care. *Circulation*, vol 142, 2020, pp S366-S468.
- [4] Neth MR, Idris A, McMullan J, et al. A review of ventilation in adult out-of-hospital cardiac arrest. *Journal of the American College of Emergency Physicians open*, vol 1, 2020, pp 190-201
- [5] Aramendi E, Elola A, Alonso E, et al. Feasibility of the capnogram to monitor ventilation rate during cardiopulmonary resuscitation. *Resuscitation*, vol 110, 2017, pp 162-168.
- [6] Chang MP, Lu Y, Leroux B, et al. Association of ventilation with outcomes from out-of-hospital cardiac arrest. *Resuscitation*, vol 141, 2019, 174-181.
- [7] Risdal M, Aase SO, Stavland M, et al. Impedance-based ventilation detection during cardiopulmonary resuscitation. *IEEE Transactions on Biomedical Engineering*, vol 54, 2007, pp 2237-2245.
- [8] Charlton PH, Bonnici T, Tarassenko L, et al. An impedance pneumography signal quality index: Design, assessment and application to respiratory rate monitoring. *Biomedical Signal Processing and Control*, vol 65, 2021.
- [9] Alonso E, Ruiz J, Aramendi E, et al. Reliability and accuracy of the thoracic impedance for measuring cardiopulmonary resuscitation quality metrics. *Resuscitation*, vol 88, 2015, pp 28-34.
- [10] Jaureguibeitia X, Irusta U, Aramendi E, et al. An impedance-based algorithm to detect ventilations during cardiopulmonary resuscitation. *Proc. Computing in Cardiology*, 2020, pp 1-4.
- [11] Jaureguibeitia X, Zubia G, Irusta U, et al. Shock decision algorithms for automated external defibrillators based on convolutional networks. *IEEE Access*, vol 8, 2020.
- [12] Wang HE, Schmicker RH, Daya MR, et al. Effect of a strategy of initial laryngeal tube insertion vs endotracheal intubation on 72-hour survival in adults with out-of-hospital cardiac arrest: a randomized clinical trial. *JAMA*, vol 320, 2018, pp 769-778.
- [13] Ronneberger O, Fischer P, y Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. *arXiv*, 2015.

# Neurotecnologías I

---

Miércoles, 23 de noviembre de 2022

# An Overview of M3Rob, a Robotic Platform for Neuromotor and Cognitive Rehabilitation Using Augmented Reality

Ana Ciscal<sup>1</sup>, Víctor Martínez-Cagigal<sup>2,3</sup>, Gonzalo Alonso-Linaje<sup>1</sup>, Selene Moreno-Calderón<sup>2</sup>, Javier Pérez Turiel<sup>1</sup>, Roberto Hornero<sup>2,3</sup>, Juan Carlos Fraile Marinero<sup>1</sup>.

<sup>1</sup> Instituto de las Tecnologías Avanzadas de la Producción, Universidad de Valladolid, Valladolid, España, {[ana.ciscal@uva.es](mailto:ana.ciscal@uva.es), [gonzalo.alonso.alonso-linaje@uva.es](mailto:gonzalo.alonso.alonso-linaje@uva.es), [jcfrail@uva.es](mailto:jcfrail@uva.es), [turiel@ei.uva.es](mailto:turiel@ei.uva.es)}

<sup>2</sup> Grupo de Ingeniería Biomédica, Universidad de Valladolid, Valladolid, España

<sup>3</sup> Centro de Investigación Biomédica en Red de Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), España, {[victor.martinez@gib.tel.uva.es](mailto:victor.martinez@gib.tel.uva.es), [selene.moreno@gib.tel.uva.es](mailto:selene.moreno@gib.tel.uva.es), [robhor@tel.uva.es](mailto:robhor@tel.uva.es)}

## Abstract

*Stroke is a major cause of disability worldwide and its prevalence has increased in recent decades. Typically, conventional physical therapies for stroke patients can be improved to promote neuroplasticity. Neurofeedback techniques have been proved to enhance not only brain plasticity, but also cognitive functions. However, most rehabilitation approaches do not integrate several types of feedback simultaneously. We hypothesize that the inclusion of neurofeedback in robotic physical therapy and the application of augmented reality should favor brain plasticity in the sensorimotor area, leading to a more successful rehabilitation of paretic limbs. For this reason, we aimed to develop a rehabilitation platform that combines different feedback techniques (haptic robotics, neurofeedback and augmented reality). M3Rob, a robotic platform for hand and wrist neuromotor rehabilitation and cognitive functions recovery, is presented.*

## 1. Introduction

Cerebrovascular accident (CVA) is a major public health problem and is considered the most common cause of permanent disability in adulthood [1]. The sequelae of stroke depend on the region of the brain affected and its extent, including motor and cognitive problems. Mobility problems are the most prevalent, occurring in more than 75% patients with moderate disability. Approximately, 30 % of stroke patients develop dementia within one year of stroke onset [2].

Classical motor rehabilitation methodologies for stroke patients do not target the origin of the problem (brain trauma), which limits the effectiveness of rehabilitation. They focus on repetitive exercises through an exogenous approach, in which the patient is not entirely involved in the lost function recovery. However, previous studies have demonstrated that helping users to learn to control their own brain rhythms through neurofeedback training generates neuroplasticity in the damaged areas, thus leading to an enhancement of the rehabilitation process [3]-[5].

Furthermore, neurofeedback training has been proved not only to enhance neuroplasticity [6], [7], but also enhances certain cognitive functions not directly related to the motor therapy, such as memory, language and visuospatial ability and recognition [8], [9].

Due to the limitations of traditional rehabilitation methods, we consider essential to investigate new strategies that directly address the source of impairment to improve the effectiveness of neuromotor and cognitive stroke rehabilitation, involving the ability of brain networks to restore lost functions and change (i.e., neuroplasticity).

For all the reasons mentioned, the development of a neuro-rehabilitation platform M3Rob (from the Spanish “Mente-Mano-Muñeca Robot”; in English Mind-Hand-Wrist Robot) is proposed. First, the combination of a robotic platform and a brain-computer interface (BCI) system allows stroke patients to perform new therapies that simultaneously address the rehabilitation of neuromotor and cognitive functions. Secondly, the platform is focused on the rehabilitation of the hand and/or wrist due to approximately 60% of patients experience upper limb dysfunction, which is especially prevalent in the hand [10].

Combining these two therapeutic approaches at the same time (neuromotor and cognitive rehabilitation) using augmented reality (AR), the patient would be fully involved in their recovery through constant feedback. Due to this level of feedback (motor, cognitive and AR), users will be aware of their improvement, generating new strategies to restore control of the lost limb and, therefore, favoring neuroplasticity in the affected cortical areas.

Currently, M3Rob is an on-going project. In this document, we detail the most important aspects of the platform, as follows: (i) system overview, (ii) motor, (iii) cognitive and (iv) AR modules

## 2. System overview

### 2.1. Background

M3Rob is a continuation of the RobHand project, which was aimed at developing a hand exoskeleton for neurorehabilitation of hand functions through active therapies for people who has suffered a stroke [11]–[13]. RobHand is a hand exoskeleton for performing virtual-based exercises using either passive or EMG-driven control modes.

The evolution of the M3Rob in relation to the RobHand rehabilitation platform integrates the following new aspects: (i) incorporation of wrist module and modification

of the hand exoskeleton to make it compatible with the wrist module; incorporation of new sensors (EEG and force), which allows (ii) to introduce neurofeedback providing cognitive rehabilitation and (iii) to develop new strategies based on assist-as-needed paradigms to model a haptic control; (iv) incorporation of AR to the current virtual reality (VR) environment to make the therapy more immersive.

## 2.2. Description

The M3Rob platform will use the measurement of electromyography (EMG) and electroencephalography (EEG) signals to detect the intentionality of the patient's hand and wrist movements. The EMG will be used to assist in motor rehabilitation, while the EEG through BCI technology will be used as a cognitive rehabilitation tool, aimed at achieving neuroplasticity [3]-[5]. in the affected areas and an indirect improvement in general cognitive functions [8], [9].

The monitoring and processing of EMG and EEG will allow the delivery of real-time commands, which will be used to control a modular mechatronic device consisting of a hand exoskeleton and a wrist rehabilitation module, allowing the patient to perform personalized assistive therapies for neuromotor and cognitive rehabilitation, through the use of AR environments oriented towards activities of daily living (ADL).

The mechatronic device will implement haptic control strategies to modulate the physical patient-device interaction, while dynamically adapting online the degree of assistance received by the patient during therapy, by applying the assist-as-needed paradigms. The platform is complemented by a software environment for the local and remote management of therapies, patients, historical data collected online during the execution of therapies, etc.

The M3Rob robotic platform is composed of three main functional modules:

- Motor rehabilitation module. The robotic system, comprising a hand exoskeleton and a wrist module, that provides dynamic assistance to the patient for performing hand-wrist movements during the therapeutic exercises.
- Cognitive rehabilitation module. Monitors and processes the EEG signals of the patient using a BCI system. The BCI system will detect the hand/wrist movement intention while the patient imagines the kinesthetic sensation of performing the movement.
- Augmented reality module. Provides an immersive environment in which a virtual scenario is combined with physical interaction with real elements in order to offer perceptually enriched experiences and enhance the effectiveness of rehabilitation therapies.

## 3. Motor Rehabilitation Module

The M3Rob rehabilitation platform is modular and can adopt three different configurations to perform hand and/or wrist rehabilitation exercises.

### 3.1. Configuration for Hand Rehabilitation

The hand exoskeleton assists in flexion/extension movements of the hand fingers using a 4-bar linkage underactuated mechanism for each finger. Each mechanism fixed to a platform is powered by a linear actuator (Actuonix) and has one active rotative degree of freedom (DoF) and one passive rotative DoF for the metacarpophalangeal (MCP) and the proximal interphalangeal (PIP) joints, respectively. The exoskeleton is attached to a forearm support to mitigate the negative effects generated by its weight. The connection is made by means of a spring-loaded interlocking positioning mechanism placed on the back the platform (Figure 1). The RoM of the MCP joint is 70°.



Figure 1. Hand exoskeleton and forearm support

### 3.2. Configuration for Wrist Rehabilitation

The wrist module moves the user's wrist using a mechanism (Figure 2) with three serial revolute active joints (3-DoF RRR). The pronation/supination (PS), flexion/extension (FE) and radial/ulnar deviation (RU) movements of the wrist are assisted by the first, second and third rotational joint, respectively. Each rotational joint is powered by one independent brushed DC motor using cable to transmit the motion. The mechanism has a removable cylindrical handle with a force sensor (Me-Systems). The range of motion of PS, FE and RU joints are 180°, 135° and 110°, respectively.

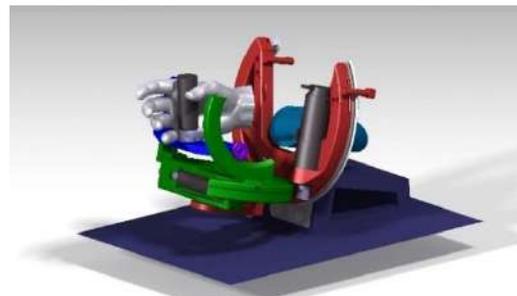
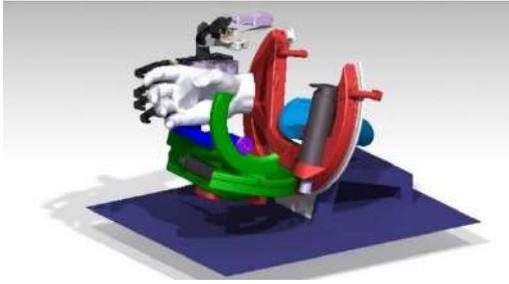


Figure 2. Wrist rehabilitation module

### 3.3. Configuration for Combined Hand and Wrist Rehabilitation

The original hand exoskeleton platform has undergone some modifications to make it compatible with the wrist module, so a linear rail and positioner blocking has been added to connect the hand exoskeleton to the wrist module. In this configuration, the handle is removed (Figure 3).



**Figure 3.** Hand exoskeleton and wrist rehabilitation module

#### 4. Cognitive Rehabilitation Module

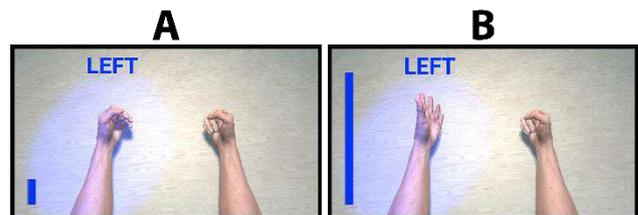
The most common method for monitoring brain activity is the EEG, given its portability, low cost, and reliability [14]. Since the goal is to improve hand and wrist motor functions, the BCI system should focus on analyzing the cortical area that controls the upper extremities: the sensorimotor region. The EEG signals collected over the sensorimotor cortex may reflect changes in certain frequencies when the user performs or imagines a movement, known as sensorimotor rhythms (SMRs). SMRs are defined as alpha and beta band oscillations that can be generated before, during and after a movement (or even imagined movement) over the sensorimotor cortex [14]. Through neurofeedback training, users can learn to regulate their EEG activity in this region to generate volitional SMRs. Thus, the BCI system can detect the intention to perform a movement even in people with motor impairment, promoting neurological activity in the affected area [4].

One of the most successful strategies to generate contralateral SMRs in both hemispheres is based on the Motor Imagery (MI) of right- and left-hand movements. The effectivity of MI-based BCIs to favor stroke neurorehabilitation has been demonstrated recently [15], not only by improving the movement of paretic hands [4], [16], but also by eliciting changes in functional brain organization due to Hebbian neuroplasticity [5]. Recent studies have also suggested that ipsilateral pre-frontal activity also plays an important role in stroke rehabilitation, although additional investigation is encouraged [16].

The reliability of the MI-based BCIs heavily depends on the user's ability to self-regulate their SMRs, which in turn is closely related to the quality of the training and the kinesthetic strategies used to elicit the desired brain activity [3]. One key aspect of MI training is the motivation of the user and the quality of the neurofeedback, which must reflect EEG changes in real-time to allow users to change their mental strategies to provoke them. The simpler feedback would be visual embodied feedback, in which the user is instructed to imagine the kinesthetic remembrance of a random hand movement, while the intensity of the SMRs is indicated in real-time. A deep neural network (i.e., EEGSym [17]) will be used to detect the probability that the user is imagining the movement of the right or left hand, moving the visual embodied feedback of the Figure 4 accordingly. The quality of this feedback can be improved by involving additional sensory outputs when the SMR is detected, e.g.,

using neuromuscular electrical stimulation, virtual and augmented reality, or haptic robotics [15].

In M3Rob, visual feedback is provided via a Unity app integrated in MEDUSA© ([www.medusabci.com](http://www.medusabci.com)), a novel Python-based software ecosystem to design BCIs and neuroscience experiments in real-time. Screenshots of the cognitive rehabilitation module are shown in Figure 4. The feedback of the first sessions will be always positive to help users learn to modulate their own brain activity on the sensorimotor region. As the sessions increase, the classifier will be updated to progressively show the real feedback.



**Figure 4.** Visual embodied neurofeedback is provided to reflect the intensity of the motor imagination.

#### 5. Augmented Reality Module

Gamification has been proved to positively influence different rehabilitation therapies, including those using exoskeletons [18]. This positive influence is directly related to an increase in patients' motivation and engagement, which results in patients performing their exercises in a more sustained fashion. RobHand used VR technologies to facilitate rehabilitation exercises. M3Rob takes this approach further, replacing VR with interoperable AR interfaces. While VR creates a whole virtual environment for the patient, AR extracts and leverages information from the real world, enhancing the properties of existing objects and allowing the creation of new interactive elements which are integrated in the real-world-environment.

Fusing gamification and AR, a set of interactive virtual objects are placed in the real world and visualized through an AR device. These objects and their interactions simulate activities that were previously carried out daily by the patient with additional game-related elements, increasing the motivation of the patient, not only due to the novelty of the therapy, but also due to the transfer from a rehabilitation environment to an enhanced but realistic activity of daily living environment.

In order to develop and easily customize augmented therapies, a set of customizable general-purpose interactive elements is needed. Furthermore, logical protocols and components defining how these elements should be interconnected and how they interact between them is also required. Additionally, it is beneficial to include support for additional hardware, such as controllers or haptic feedback devices. To properly design these components, a close collaboration between software developers and healthcare professionals is needed, ensuring that the virtual elements and the objectives planned within the augmented environments meet the therapeutic requirements for upper limb rehabilitation.

## 6. Discussion and conclusions

We have presented the M3Rob platform, currently under development, an evolution of the RobHand hand exoskeleton. It will allow patients to undergo therapies for rehabilitation of both neuromotor and cognitive functions damaged by stroke. To the best of our knowledge, this has only been addressed separately. With the proposed platform, the patient will be able to benefit from the combined effects of both therapeutic approaches, which may potentially improve the outcomes of the rehabilitation. Furthermore, the platform adopts three different configurations so that damaged neuromotor functions of the hand and wrist can be rehabilitated individually or simultaneously.

The platform integrates (i) a haptic control for reliable and natural HRI (human-robot interaction) by the modulation of involved forces (ii) assisted-as-needed strategies which not only take into account the exercises performance but also the patient's intention by processing their EMG and EEG signals (iii) the EEG processing will also allow to recover cognitive functions by using neurofeedback (iv) augmented reality so the patient can interact with virtual objects in the real environment.

The brain damage unit of the Benito Menni Hospital Center (Valladolid) is assisting with this project. The unit has defined the clinical protocol for testing M3Rob platform with stroke patients. Furthermore, the protocol has already been submitted to the ethics committee and we are awaiting approval for starting the clinical trial.

### Funding

This study was supported by the Ministry of Science and Innovation, through the research project RTC2019-007350-1, by the company TICCYL Digital S.L.U and by the support of the Regional Ministry of Education.

### References

- [1] G. A. Donnan, M. Fisher, M. Macleod, and S. M. Davis, "Stroke," *Lancet*, vol. 373, no. 9674, p. 1496, 2008, doi: 10.1016/S0140-6736(09)60833-3.
- [2] B. Cullen, B. O'Neill, J. J. Evans, R. F. Coen, and B. A. Lawlor, "A review of screening tests for cognitive impairment," *J. Neurol. Neurosurg. Psychiatry*, vol. 78, no. 8, pp. 790–799, 2007, doi: 10.1136/jnnp.2006.095414.
- [3] W. P. Teo and E. Chew, "Is motor-imagery brain-computer interface feasible in stroke rehabilitation?," *PMR*, vol. 6, no. 8, pp. 723–728, 2014, doi: 10.1016/j.pmrj.2014.01.006.
- [4] C. Guger et al., "Brain-computer interfaces for stroke rehabilitation: summary of the 2016 BCI Meeting in Asilomar," *Brain-Computer Interfaces*, vol. 5, no. 2–3, pp. 41–57, 2018, doi: 10.1080/2326263X.2018.1493073.
- [5] B. M. Young et al., "Changes in functional brain organization and behavioral correlations after rehabilitative therapy using a brain-computer interface," *Front. Neuroeng.*, vol. 7, no. JUL, pp. 1–15, 2014, doi: 10.3389/fneng.2014.00026.
- [6] D. C. Irimia, R. Ortner, M. S. Poboroniuc, B. E. Ignat, and C. Guger, "High classification accuracy of a motor imagery based brain-computer interface for stroke rehabilitation training," *Front. Robot. AI*, vol. 5, no. NOV, pp. 1–9, 2018, doi: 10.3389/frobt.2018.00130.
- [7] A. Caria et al., "Chronic stroke recovery after combined BCI training and physiotherapy: A case report," *Psychophysiology*, vol. 48, no. 4, pp. 578–582, 2011, doi: 10.1111/j.1469-8986.2010.01117.x.
- [8] E. Angelakis, S. Stathopoulou, J. L. Frymiare, D. L. Green, J. F. Lubar, and J. Kounios, "EEG Neurofeedback: A Brief Overview and an Example of Peak Alpha Frequency Training for Cognitive Enhancement in the Elderly," *Clin. Neuropsychol.*, vol. 21, no. 1, pp. 110–129, 2007, doi: 10.1080/13854040600744839.
- [9] J. Gomez-Pilar, R. Corralejo, L. F. Nicolas-Alonso, D. Álvarez, and R. Hornero, "Neurofeedback training with a motor imagery-based BCI: neurocognitive improvements and EEG changes in the elderly," *Med. Biol. Eng. Comput.*, vol. 54, no. 11, pp. 1655–1666, 2016, doi: 10.1007/s11517-016-1454-4.
- [10] H. C. Fischer, K. Stubblefield, T. Kline, X. Luo, R. V. Kenyon, and D. G. Kamper, "Hand Rehabilitation Following Stroke: A Pilot Study of Assisted Finger Extension Training in a Virtual Environment," *Top. Stroke Rehabil.*, vol. 14, no. 1, pp. 1–12, Jan. 2007, doi: 10.1310/tsr1401-1.
- [11] V. Moreno-San Juan, A. Ciscal, J. Fraile, J. Pérez-turiel, and E. de la Fuente, "Design and Characterization of a Lightweight Underactuated RACA Hand Exoskeleton for Neurorehabilitation," *Rob. Auton. Syst.*, vol. 143, pp. 1–32, 2021, doi: 10.1016/j.robot.2021.103828.
- [12] A. Ciscal, J. Perez-Turiel, J. C. Fraile, D. Sierra, and E. De La Fuente, "RobHand: A Hand Exoskeleton with Real-Time EMG-Driven Embedded Control. Quantifying Hand Gesture Recognition Delays for Bilateral Rehabilitation," *IEEE Access*, vol. 9, pp. 137809–137823, 2021, doi: 10.1109/ACCESS.2021.3118281.
- [13] A. Ciscal, V. Moreno-SanJuan, D. Sierra, J. P. Turiel, and J. C. Fraile, "An Embedded Implementation of EMG-Driven Control for Assisted Bilateral Therapy," in *Converging Clinical and Engineering Research on Neurorehabilitation IV. ICNR 2020. Biosystems & Biorobotics*, vol. 28, 2022, pp. 817–821, doi: 10.1007/978-3-030-70316-5\_130.
- [14] J. Wolpaw and E. Wolpaw, *Brain-Computer Interfaces: Principles and Practice*. 2012.
- [15] M. A. Cervera et al., "Brain-computer interfaces for post-stroke motor rehabilitation: a meta-analysis," *Ann. Clin. Transl. Neurol.*, vol. 5, no. 5, pp. 651–663, 2018, doi: 10.1002/acn3.544.
- [16] E. López-Larraz, A. Sarasola-Sanz, N. Irastorza-Landa, N. Birbaumer, and A. Ramos-Murguialday, "Brain-machine interfaces for rehabilitation in stroke: A review," *NeuroRehabilitation*, vol. 43, no. 1, pp. 77–97, 2018, doi: 10.3233/NRE-172394.
- [17] S. Pérez-Velasco, E. Santamaría-Vázquez, V. Martínez-Cagigal, D. Marcos-Martínez and R. Hornero, "EEGSym: Overcoming Inter-Subject Variability in Motor Imagery Based BCIs With Deep Learning," in *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 30, pp. 1766–1775, 2022, doi: 10.1109/TNSRE.2022.3186442.
- [18] O. Mubin, F. Alnajjar, N. Jishtu, B. Alsinglawi, and A. Al Mahmud, "Exoskeletons with virtual reality, augmented reality, and gamification for stroke patients' rehabilitation: Systematic review," *JMIR Rehabil. Assist. Technol.*, vol. 6, no. 2, pp. 1–11, 2019, doi: 10.2196/12010.

# Precisión de movimientos y gestos de la mano en terapias virtuales inmersivas

C. Sánchez del Valle<sup>1</sup>, M. Álvarez Rodríguez<sup>1</sup>, R. Perales Gómez<sup>2</sup>, V. Lozano Berrio<sup>1</sup>, A. Gil Agudo<sup>3</sup>,  
A. de los Reyes Guzmán<sup>1</sup>

<sup>1</sup> Unidad de Biomecánica y Ayudas Técnicas, Hospital Nacional de Paraplégicos, Toledo, España  
[vlozanob,adlos}@sescam.jccm.es](mailto:{vlozanob,adlos}@sescam.jccm.es)

<sup>2</sup> Unidad de Terapia Ocupacional, Hospital Nacional de Paraplégicos, Toledo, España [rpgomez@sescam.jccm.es](mailto:rpgomez@sescam.jccm.es)

<sup>3</sup> Servicio de Rehabilitación, Hospital Nacional de Paraplégicos, Toledo, España [amgila@sescam.jccm.es](mailto:amgila@sescam.jccm.es)

## Resumen

*Actualmente, hay un interés creciente en el uso de tecnología en el campo de la Neurorehabilitación con el objetivo de cuantificar y generar conocimiento acerca de los trastornos sensorimotores tras enfermedades neurológicas, entendiendo que la tecnología posee un elevado potencial para su uso como herramientas terapéuticas. Además, las enfermedades neurológicas producen en la mayor parte de los casos trastornos discapacitantes que perduran durante toda la vida del paciente afectando a su calidad de vida y participación social. En este contexto, se entiende que el proceso rehabilitador de los trastornos motores se debe extender más allá de la fase de ingreso hospitalario del paciente, siendo necesario involucrar tecnología de bajo coste en el proceso de rehabilitación con el objetivo de contar con soluciones tecnológicas que se puedan aproximar al domicilio del paciente. Así, se cuenta con el software RehabHand que, basado en aplicaciones virtuales manipuladas por medio de tecnología de bajo coste, como Leap Motion Controller, se ha diseñado y desarrollado para la rehabilitación de la destreza manipulativa de los miembros superiores en pacientes con afecciones neurológicas. Actualmente, existe una tendencia a incorporar metodologías inmersivas en los programas terapéuticos que reciben los pacientes, que ya han sido incorporadas por algunos centros clínicos. En este contexto resulta de interés analizar la precisión de los gestos y movimientos de la mano en ambas modalidades de terapia virtual, inmersiva y no inmersiva, inicialmente en sujetos sanos.*

## 1. Introducción

Actualmente, hay un interés creciente en el uso de tecnología en el campo de la Neurorehabilitación con el objetivo de cuantificar y generar conocimiento acerca de los trastornos sensorimotores tras enfermedades neurológicas, entendiendo que la tecnología posee un elevado potencial para su uso como herramientas terapéuticas [1]. Sin embargo, estas tecnologías, tanto la robótica como la basada en redes de sensores, no poseen actualmente una verdadera implantación en la práctica clínica diaria por varios motivos. Uno de ellos, de elevada importancia, es el elevado coste que poseen, que reduce su ámbito de aplicación únicamente al de los centros de rehabilitación especializados.

Por otro lado, las enfermedades neurológicas producen en la mayor parte de los casos trastornos discapacitantes que

perduran durante toda la vida del paciente afectando a su calidad de vida, nivel de independencia en ejecución de actividades de la vida diaria y participación social. Un ejemplo, es la lesión medular cervical, que carece de un proceso curativo completo y en la que la fuerza residual de músculos parcialmente paralizados es un importante determinante para poseer independencia y funcionalidad en los miembros superiores [2]. En este contexto, se entiende que el proceso rehabilitador se debe extender más allá de la fase de ingreso hospitalario del paciente, siendo necesario involucrar tecnología de bajo coste en el proceso de rehabilitación con el objetivo de contar con soluciones tecnológicas que se puedan aproximar al domicilio del paciente. La tecnología unida al uso de aplicaciones virtuales permitiría la ejecución de ejercicios terapéuticos de la intensidad y número de repeticiones necesarias para inducir la neuroplasticidad, proporcionando las fuentes de feedback necesarias como para motivar y asegurar la adherencia del paciente a las tareas [3]. Además, existe una tendencia al diseño y desarrollo de entornos virtuales inmersivos para rehabilitación motora. Sin embargo, antes de adoptar estos entornos con un propósito rehabilitador, es necesario determinar la equivalencia cinemática entre los movimientos realizados en el entorno virtual con aquéllos realizados en entornos físicos reales. Este aspecto es especialmente importante en los movimientos de alcance y agarre, que dependen de la información acerca del objeto a manipular con el objetivo de planificar las trayectorias de alcance y el tipo de agarre. Este estudio concreto lo realizaron en personas sanas utilizando head-mounted display junto con guantes como Cybergrasp y Cyberglove y un equipo de fotogrametría para analizar la cinemática [4].

Actualmente, esta modalidad inmersiva se está poniendo en práctica en algunos centros clínicos para el tratamiento de enfermedades neurológicas como Daño Cerebral Adquirido, con el objetivo de trabajar aspectos relacionados con la atención, percepción y control motor dentro de entornos realistas que simulan situaciones y actividades del día a día [5]. Sin embargo, pese a la realización de estudios en muestras de pacientes, es todavía mucha la investigación necesaria, en términos de efectividad, para poder utilizar estas terapias consistentemente con un propósito rehabilitador, y en

términos de viabilidad en cuanto a la duración óptima de estas terapias.

Entendemos que las terapias virtuales son el presente y el futuro, en complemento con las terapias convencionales, de forma que se ha desarrollado un software para rehabilitación del miembro superior basado en aplicaciones virtuales manipuladas por medio de Leap Motion Controller (LMC) [6] para la rehabilitación de los aspectos motores de los miembros superiores en pacientes con afecciones neurológicas [7]. Cada aplicación virtual posee un objetivo terapéutico en cuanto a los aspectos motores del miembro superior y, más concretamente la mano [8]. A partir de este software, el objetivo del trabajo que aquí se presenta es adaptar una de estas aplicaciones virtuales a un entorno inmersivo para su interacción por medio de HTC Vive junto con LMC y realizar un pequeño estudio en sujetos sanos para analizar la precisión de los gestos y movimientos de la mano en ambas modalidades de terapia virtual, no inmersiva por medio de LMC e inmersiva por medio de HTC Vive y LMC.

## 2. Métodos

### 2.1. Prototipo software para rehabilitación motora

Este prototipo se ha diseñado y desarrollado mediante la participación colaborativa y coordinada de un equipo multidisciplinar de trabajo formado por médicos, terapeutas ocupacionales e ingenieros. A partir de los requisitos clínicos y necesidades en el paciente definidas por el personal clínico se han definido y desarrollado distintos entornos virtuales. El contenido de las aplicaciones está orientado a lograr tareas funcionales completas que requieren la realización de gestos y movimientos del miembro superior y, en concreto de la mano, que son necesarios rehabilitar en el paciente para una mejora de la destreza manipulativa. Un ejemplo de estas aplicaciones se muestra en la Figura 1. El dispositivo de interacción con los entornos virtuales es el LMC. El desarrollo se ha realizado utilizando el motor de videojuegos Unity3D y la SDK de LMC (versión Orion 3.1.0).



**Figura 1.** Aplicación virtual RehabHand, manipulada mediante LMC para trabajar alcance y manipulación de objetos y pronación-supinación del antebrazo

Así, aspectos terapéuticos a trabajar dentro de las aplicaciones son distintos tipos de pinza, capacidad de agarre, prensión, coordinación bimanual, así como los movimientos de flexión-extensión de la muñeca y pronación-supinación del antebrazo.

Cada aplicación virtual está dotada de las fuentes de feedback visual y auditivo que necesita el paciente para apreciar su propia mejoría tras sesiones consecutivas (tiempo de ejecución, número de objetivos alcanzados, barra de progresión para mantener el gesto de pronación-supinación del antebrazo). Además, aquellas tareas que consisten en lograr el mayor número posible de objetivos en un tiempo determinado, disponen de una barra de progresión en la que se refleja el desempeño realizado relativo a un patrón de referencia formado por un grupo de personas sanas (Figura 1).

### 2.2. Versión de la aplicación basada en RV inmersiva

Para lograr el objetivo de este trabajo, se ha desarrollado una versión de la aplicación RehabHand, basada en RV inmersiva, a partir de la aplicación previa del prototipo desarrollado. Esta nueva versión inmersiva se manipula por medio de las gafas de RV HTC Vive [9].

Se utilizaron los assets de SteamVR para adecuar el juego al uso de las gafas y los controller de las gafas HTC Vive para controlar el juego. También se adaptaron los elementos de la interfaz de usuario para poder usarlos con el controlador.

Se ha desarrollado en un ordenador Lenovo, con un procesador Intel(R) Core(TM) i7-7500 CPU 3GHz, de 16 GB de RAM y tarjeta gráfica Nvidia GeForce GT 1050. Después se ha ejecutado en un ordenador HP, Intel(R) Core(TM) i5-8500 CPU 3GHz, de 8 GB de RAM y tarjeta gráfica Nvidia GeForce GT 1030. La SDK de LMC que se utiliza es Orion 3.1.0.

### 2.3. Intervención

En este estudio preliminar han participado 4 personas sanas. La intervención se realiza en un solo día. Todos los participantes habían realizado previamente alguna sesión experimental con las aplicaciones virtuales no inmersivas. El objetivo terapéutico de la aplicación del Laboratorio en la versión no inmersiva e inmersiva era exactamente el mismo. La tarea consistía en alcanzar, coger y manipular una serie de botes y probetas de forma que hay que verter el contenido dentro del recipiente central. Aparece una barra de progresión de manera que ésta se va completando si se realiza el movimiento de pronación-supinación del antebrazo que simula verter el contenido. Las cápsulas directamente se depositan dentro del recipiente. Un total de 10 objetos tienen que ser alcanzados y transportados al recipiente central para completar la tarea.

Todos los participantes tenían que cumplir los criterios de inclusión y firmar el correspondiente modelo de consentimiento informado. Los criterios de inclusión fueron: edad  $\geq 16$  años; no tener compromiso neurológico u ortopédico que afecte a la función de los miembros superiores.

Los participantes realizaron la prueba en sedestación. La duración de la prueba fue de aproximadamente 10 minutos para cada persona. La aplicación del Laboratorio se lanzó 3 veces en la versión no inmersiva. En la primera de ellas el participante realiza 3 veces el gesto de agarre

con la palma de la mano paralela al LMC; en la segunda se realiza 3 veces el gesto de agarre con la palma de la mano perpendicular al LMC; en la tercera se ejecuta el juego completo hasta completarlo o agotar el tiempo máximo fijado en 5 minutos. A continuación se repite esta metodología con la versión inmersiva por medio de HTC Vive y LMC.

### 2.4. Variables analizadas

El gesto del agarre se registra a través de LMC por medio de la variable cinemática *GraspStrength* que adopta un valor entre 0 (mano abierta) y 1 (mano cerrada). Asimismo se registra el nivel de precisión con el que se ofrece este dato en cada frame registrado. Por tanto, las variables a registrar consisten en la capacidad de ejecución del gesto de agarre correspondiente a la mano cerrada y el nivel de precisión en las dos primeras ejecuciones de la aplicación también medido entre 0 y 1. En la última, se ofrece el resultado obtenido en estas variables durante la ejecución de la tarea completa. Además, en este último caso se muestra los tiempos de ejecución de la tarea y el porcentaje de éxitos conseguidos. Los resultados se muestran como media y desviación estándar.

### 3. Resultados

En este apartado se describen los principales resultados de las pruebas realizadas.

En la Figura 2 aparece uno de los participantes durante la ejecución de la prueba de realidad virtual inmersiva.



**Figura 2.** Persona sana realizando la actividad terapéutica inmersiva mediante gafas RV y LeapMotion

En la tabla 1 se muestran los resultados cuantitativos de las pruebas realizadas. La ejecución 1 corresponde a la realizada con la mano paralela al sensor. La ejecución 2 corresponde a aquella en la que la mano está perpendicular y la ejecución 3 corresponde a la ejecución de la aplicación completa.

Los resultados obtenidos en ambas variables han sido muy similares en ambas modalidades terapéuticas sin significación estadística entre ellas. Los participantes ejecutaron un agarre completo en todas las pruebas realizadas. El agarre se ha detectado en su nivel máximo en el caso de la terapia inmersiva y en la condición en que la mano ha estado paralela al LMC. Sin embargo, en las pruebas realizadas con la mano perpendicular, el nivel de agarre detectado ha sido superior en la terapia no

inmersiva con LMC que en la inmersiva (0.97 (0.05) vs. 0.90 (0.19), respectivamente).

Agarre (0-1)	LMC		HTC Vive + LMC	
	Agarre	Precisión	Agarre	Precisión
Ejecución 1	0.98(0.03)	0.96(0.03)	1.00(0.00)	0.99(0.01)
Ejecución 2	0.97(0.05)	0.97(0.02)	0.90(0.19)	0.99(0.00)
Ejecución 3	0.56(0.07)	0.99(0.00)	0.58(0.08)	0.99(0.00)

**Tabla 1.** Nivel de agarre y precisión del gesto medidos a través de LMC

En la ejecución los resultados obtenidos en estas dos variables son muy similares en las dos modalidades terapéuticas. Todos los participantes completaron el 100% de la tarea alcanzando los 10 objetos en ambas terapias. Sin embargo, el tiempo de ejecución fue superior en la modalidad inmersiva (158.49 (49.96) s) que en la realizada solamente con el LMC (101.75 (22.20) s).

Ningún participante experimentó mareo o algún tipo de efecto adverso durante la realización de la prueba, ni en lo sucesivo durante el día de experimentación.

### 4. Discusión

En este trabajo se presenta la adaptación de una de las aplicaciones virtuales de escritorio del prototipo basado en LMC a una aplicación similar que, bajo el mismo contenido terapéutico, se pueda manipular mediante gafas de RV y LMC. Se ha medido el nivel de agarre y la precisión con la que se detecta este gesto por parte de LMC en ambas modalidades terapéuticas. En este trabajo se han utilizado las gafas HTC Vive. En otros trabajos encontrados en la literatura se ha utilizado head-mounted display [4]. A pesar de los estudios de viabilidad y efectividad todavía necesarios para su aplicación como soluciones terapéuticas en el entorno clínico, algunos centros ya han empezado a utilizar estas terapias con pacientes, sobre todo aplicadas al tratamiento de DCA [5].

Se ha detectado la necesidad de incorporar estas tecnologías al tratamiento de pacientes neurológicos, teniendo en cuenta el potencial que ofrecen las terapias virtuales. En el contexto de este estudio preliminar realizado en 4 personas sanas, el resultado obtenido ha sido satisfactorio en cuanto al nivel de detección de los movimientos y gestos de la mano realizados. Estos estudios son imprescindibles dada la información que comerciantes de este tipo de tecnologías ofrecen al consumidor en cuanto a la aparición de posibles efectos adversos. Un ejemplo es la información que muestra Amazon relativa a los dispositivos de Realidad Virtual, como es el caso de Oculus Rift, avisando del riesgo a que el usuario experimente una sensación de mareo conocida como “*Motion sickness*”. Este hecho no se debe a que el

producto esté defectuoso, sino más bien a que no todos los usuarios son aptos para su uso [10]. En este sentido, dentro del estudio que aquí se presenta, ningún participante ha experimentado sensación de mareo. Una ventaja es que este estudio se realiza en sedestación, lo que implica que no se utilizan las gafas al límite máximo de su funcionalidad, minimizando el riesgo de aparición de *motion sickness*.

Otro aspecto que es importante destacar, corresponde a la duración de la actividad terapéutica cuando estas metodologías se apliquen a pacientes neurológicos. Las gafas HTC Vive tienen un peso de 555g. Por ejemplo, pensando en pacientes con lesión medular cervical, esta característica de la tecnología, unida a que los pacientes presentan menos fuerza en la cabeza, obliga en principio a limitar la duración de estas terapias. En este estudio inicial la duración ha sido de aproximadamente 10 min en cada participante. Lo ideal para acciones y estudios futuros es usar esta nueva modalidad terapéutica en combinación con la terapia virtual de escritorio propuesta en el prototipo ya desarrollado, e ir incrementando la dosis de terapia inmersiva en función de la tolerancia del paciente.

No obstante, este estudio presenta limitaciones importantes. El entorno inmersivo se ha desarrollado a partir de la aplicación previa para LMC que utiliza la SDK de LMC versión 3.1.0. Se hizo una adaptación rápida del entorno inmersivo, de forma que se han observado deficiencias en el modelo de la mano y en los modelos de agarre y manipulación de objetos. Este hecho hace patente la necesidad de seguir invirtiendo esfuerzo en adaptar, mejorar y optimizar las aplicaciones virtuales desarrolladas para LMC, para que puedan ser adecuadamente manipuladas por medio de las gafas de RV y LMC.

## 5. Conclusiones

El estudio preliminar realizado en una muestra de personas sanas para comparar ambas modalidades terapéuticas ha sido satisfactorio.

Los siguientes pasos deben ser la reprogramación software de las aplicaciones virtuales que componen el prototipo previamente desarrollado y obtener mediciones cinemáticas para comparar los patrones de movimiento realizados en ambas modalidades y analizar la efectividad de las terapias inmersivas.

## Agradecimientos

Proyecto PID2020-117361RB-C22 financiado por MCIN/AEI/ 10.13039/501100011033.

## Referencias

- [1] Shirota C, Balasubramanian S, Melendez-Calderon A. Technology-aided assessments of sensorimotor function: current use, barriers and future directions in the view of different stakeholders. *Journal of neuroengineering and rehabilitation*, 16(1), 2019, 53.
- [2] Yozbatiran N, Francisco GE. Robot-assisted Therapy for the Upper Limb after Cervical Spinal Cord Injury. *Physical Medicine and Rehabilitation Clinics*, 2019, 367–384. <https://doi.org/10.1016/j.pmr.2018.12.008>
- [3] Dimbwadyo-Terrer I, Gil-Agudo A, Segura-Fragoso A, de los Reyes-Guzmán A, Trincado-Alonso F, Piazza S, Polonio-López B. Effectiveness of the virtual reality system Toyra on upper limb function in people with tetraplegia: a pilot randomized clinical trial. *BioMed research international*, 2016.
- [4] Magdalon EC, Michaelsen SM, Quevedo AA, Levin MF. Comparison of grasping movements made by healthy subjects in a 3-dimensional immersive virtual versus physical environment. *Acta psychologica*, 138(1), 2011, 126-134.
- [5] Fundación Lescer. <https://www.centrolescer.org/centro-lescer/la-realidad-virtual-al-servicio-de-la-neurorehabilitacion/> (Consultada: Junio 2019).
- [6] Leap Motion Controller. <https://www.leapmotion.com/> (Consultada: Mayo 2022).
- [7] de Los Reyes-Guzmán A, Fernández García L, Álvarez-Rodríguez M, Lozano-Berrio V, Domingo-García AM, Ceruelo-Abajo S. Realidad virtual de bajo coste. Una nueva aplicación para rehabilitación motora de los miembros superiores en patología neurológica: Estudio piloto [Low-cost virtual reality. A new application for upper extremity motor rehabilitation in neurological pathology: Pilot study]. *Rehabilitacion (Madr)*. 2022;56(3):173-181. doi:10.1016/j.rh.2021.07.001
- [8] de Los Reyes-Guzmán A, Lozano-Berrio V, Álvarez-Rodríguez M, et al. RehabHand: Oriented-tasks serious games for upper limb rehabilitation by using Leap Motion Controller and target population in spinal cord injury. *NeuroRehabilitation*. 2021;48(3):365-373. doi:10.3233/NRE-201598
- [9] VIVE™ | Discover Virtual Reality Beyond Imagination. <https://www.vive.com/us/> (Consultada: Julio 2022).
- [10] Oculus Rift en Amazon. [https://www.amazon.es/Oculus-Rift-Sistema-realidad-exposiciones/dp/B00ZFOGHRG/ref=pd\\_sbs\\_63\\_5/257-4835547-0527902?encoding=UTF8&pd\\_rd\\_i=B00ZFOGHRG&pd\\_rd\\_r=12a70af4-6f2e-47c3-a71b-f01dc7cb7f4e&pd\\_rd\\_w=9a55c&pd\\_rd\\_wg=AG0rK&pf\\_rd\\_p=f9384d3f-fa3d-4e25-8bc3-b0c7853cd8a6&pf\\_rd\\_r=TKPDYFDNS0YGA9JNNG2C&psc=1&refRID=TKPDYFDNS0YGA9JNNG2C](https://www.amazon.es/Oculus-Rift-Sistema-realidad-exposiciones/dp/B00ZFOGHRG/ref=pd_sbs_63_5/257-4835547-0527902?encoding=UTF8&pd_rd_i=B00ZFOGHRG&pd_rd_r=12a70af4-6f2e-47c3-a71b-f01dc7cb7f4e&pd_rd_w=9a55c&pd_rd_wg=AG0rK&pf_rd_p=f9384d3f-fa3d-4e25-8bc3-b0c7853cd8a6&pf_rd_r=TKPDYFDNS0YGA9JNNG2C&psc=1&refRID=TKPDYFDNS0YGA9JNNG2C) (Consultada: Julio 2022).

# Análisis de Imaginación Motora durante pedaleo a partir de señales EEG

J. V. Juan<sup>1,2</sup>, E. Iáñez<sup>1,3</sup>, M. Ortiz<sup>1,3</sup>, J. Tornero<sup>2</sup>, J. M. Azorín<sup>1,3,4</sup>

<sup>1</sup> Brain-Machine Interface Systems Lab, Universidad Miguel Hernández de Elche,

Elche, España, {javier.juanp,eianez,mortiz,jm.azorin}@umh.es, <http://bmi.edu.umh.es/>

<sup>2</sup> Center for Clinical Neuroscience Hospital Los Madroños, Hospital Los Madroños, Brunete, España, [jesus.tornero@lmh.es](mailto:jesus.tornero@lmh.es)

<sup>3</sup> Instituto de Investigación en Ingeniería de Elche-I3E, Universidad Miguel Hernández de Elche, Elche, España

<sup>4</sup> ValGRAI: Valencian Graduated School and Research Network of Artificial Intelligence, España

## Resumen

*El empleo de técnicas de neurorrehabilitación ha generado resultados muy prometedores a lo largo de los últimos años en el tratamiento de ciertos tipos de dolencias motoras. Algunos estudios dejan constancia de que la aplicación de estas terapias acorta los plazos de recuperación de los pacientes y en muchos casos contribuye a aumentar el grado de movilidad recuperado. Estos resultados son aún mejores cuando se combina estas técnicas con el uso de dispositivos clínicos como pueden ser exoesqueletos. Sin embargo, estos dispositivos conllevan un elevado coste económico; en este estudio se plantea la posibilidad de sustituir los exoesqueletos por dispositivos clínicos de pedaleo, más asequibles, en el empleo de estas terapias. Para ello, se ha desarrollado un algoritmo para la decodificación de la Imaginación Motora (MI, 'Motor Imagery') a partir de señales electroencefalográficas (EEG) y se ha testado con cuatro sujetos sanos, obteniendo alrededor de un 60% de acierto y sentando así la base para un futuro estudio clínico.*

## 1. Introducción

Las Interfaces Cerebro-Máquina (BMIs, *Brain-Machine Interfaces*) son dispositivos capaces de establecer una comunicación entre cerebro y computadora [1]. Asimismo, existen diversas maneras de establecer esta comunicación, ya sea mediante métodos invasivos o no invasivos, como por ejemplo la Imagen por Resonancia Magnética Funcional (fMRI, *functional Magnetic Resonance Imaging*) [2], o la Magnetoencefalografía (MEG) [3]. Una de las más extendidas, especialmente en neurorrehabilitación, es el Electroencefalograma (EEG) [4, 5]. Este consiste en cuantificar los impulsos eléctricos generados por las neuronas en el cerebro, obteniendo señales susceptibles de analizar en busca de información sobre el estado mental del usuario.

Estas técnicas tienen aplicaciones en muchos ámbitos, entre ellos la medicina. Una de las más prometedoras es la neurorrehabilitación, es decir, la aplicación de esta clase de lecturas fisiológicas a la rehabilitación de pacientes con disfunciones motoras debidas a daños nerviosos o neuronales, principalmente pacientes que han sufrido un accidente cerebrovascular o una lesión de médula espinal. Por lo general, estas terapias rehabilitadoras se fundamentan en la detección de las intenciones motoras del

paciente, en otras palabras, en analizar las señales obtenidas por la BMI para poder discernir en qué momentos el paciente muestra intención de movimiento (lo que se conoce como imaginación motora o MI). De este modo, durante esos instantes de tiempo pueden activarse dispositivos clínicos motores, generalmente exoesqueletos, que efectúen ese movimiento por el paciente. Con ello se logra una mayor implicación a nivel neurofisiológico del paciente en estas terapias, ya que no conseguirá ejecutar ningún movimiento si no se concentra en él.

De acuerdo con algunos estudios [6, 7], la implicación neuronal en terapias motoras ayuda a la recuperación de la neuroplasticidad, lo que permite una mayor y veloz restauración del movimiento perdido por el paciente. En consecuencia, se han desarrollado numerosos estudios enfocados a la decodificación de la MI para terapias motoras, muchos de ellos apoyados en exoesqueletos [8, 9], aunque su alto coste económico ha provocado el surgimiento de otras líneas de investigación basadas en dispositivos menos costosos, como por ejemplo dispositivos clínicos de pedaleo [10, 11]. Muchos de estos casos han resultado siendo, en cierta medida, exitosos, ya que se ha logrado alcanzar niveles relativamente altos de acierto en la decodificación de MI (75%, en ocasiones incluso 80% [12, 13]). No obstante, todavía no ha llegado a establecerse una terapia estandarizada y extendida cimentada en estos sistemas, lo que sugiere que aún es necesario avanzar un poco más con las investigaciones y afinar en la búsqueda de la manera óptima y más segura de emplear estos dispositivos basados en EEG.

Principalmente por este motivo, en este artículo se presenta un estudio preliminar dirigido a la detección de la MI de sujetos durante el pedaleo, tratando de sentar las bases para un posterior estudio clínico que permita crear una estrategia terapéutica robusta que ayude a mejorar la vida de los pacientes con disfunciones neuronales motoras.

## 2. Materiales y métodos

A continuación, se presenta en esta sección el protocolo seguido para la toma experimental de datos, así como todo el material y la sensorización empleados para tal fin.

Además, se detalla el procedimiento empleado para el procesamiento y análisis de los datos capturados:

### 2.1. Sujetos

Las pruebas seguidas han sido realizadas por cuatro sujetos, dos hombres (S2 y S4) y dos mujeres (S1 y S3), todos sin ningún tipo de disfunción motora ni neuronal diagnosticada, con edades comprendidas entre los 28 y los 47 años ( $33.25 \pm 9.18$ ). Por su parte, todos los sujetos recibieron información clara sobre el estudio, y firmaron un consentimiento, la declaración de Helsinki y el comité ético aprobado por el Hospital Los Madroños de Brunete, en Madrid.

### 2.2. Protocolo experimental

El control de dispositivos motores mediante EEG requiere de dos paradigmas: el modelo de control cuando el dispositivo está parado, y el modelo de control cuando el dispositivo está en movimiento. En consecuencia, el protocolo diseñado ha cubierto ambos paradigmas, ejecutando tanto repeticiones con el cicloergómetro parado, como repeticiones con el cicloergómetro activo (siempre en modo pasivo, es decir, asistiendo al sujeto en el pedaleo). El protocolo a seguir en todas estas repeticiones ha sido invariable: el sujeto comenzaba con 15 segundos de relax (tratar de evadir la mente), continuando con 30 segundos de MI (concentrarse en la imaginación motora del pedaleo) y finalizando con otros 15 segundos de relax para balancear las clases (30 segundos relax – 30 segundos MI). Con este protocolo, ejecutando ambos paradigmas, se dispone de datos tanto para arrancar la pedalina con MI cuando está parada, como para detenerla con relax cuando está en movimiento. Cada sujeto realizó 20 repeticiones de este protocolo, intercalando repeticiones en las que el cicloergómetro estaba parado con repeticiones en las que se haya en movimiento. Todo este proceso está sintetizado en la Figura 1.

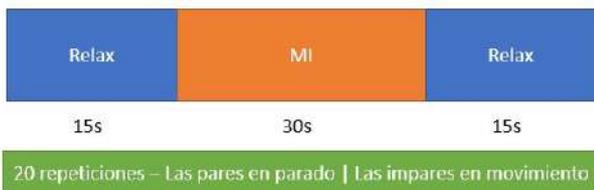


Figura 1. Protocolo experimental empleado en las pruebas.

### 2.3. Material empleado

Para los registros EEG, se ha usado el gorro de electroencefalogramas g.NAUTILUSPRO Flexible de 32 canales (g.tec medical engineering ©), el cual comunica por vía inalámbrica con el PC de registro mediante un transmisor (HEADSET) y un receptor (BASE STATION). La adquisición de estas señales EEG fue ejecutada a 500Hz.

Respecto al cicloergómetro, este ha sido el modelo CycleMotus™ A4 (Fourier Intelligence ©). Cuando la pedalina estuvo activa, siempre fue en modo pasivo, con una frecuencia de pedaleo de 20 revoluciones por minuto.

La Figura 2 muestra el montaje experimental seguido durante las sesiones de adquisición de datos.



Figura 2. Montaje experimental empleado durante las pruebas. El usuario se encuentra sentado cómodamente en una silla, pedaleando cuando el protocolo lo indica, mientras el equipo EEG registra sus señales cerebrales y las envía inalámbricamente al PC de registro.

### 2.4. Análisis realizados

Para extraer la imaginación motora de los registros tomados, se realizó un procesamiento de los datos basado en inteligencia artificial hasta lograr modelos predictivos.

En primera instancia, se efectuó un preprocesamiento de las señales, con el fin de eliminar toda la información no relevante para la decodificación de la imaginación motora que pudiera disminuir el rendimiento de los modelos predictivos y aumentar la carga computacional necesaria para su generación. Este preprocesamiento consistió en un filtro hardware Notch de 50Hz (48Hz-52Hz) de orden 4 aplicado por el equipo de EEG a todos los canales para eliminar la frecuencia de red. Tras ello, se efectuó una preselección de 19 canales empleados para el estudio situados en la zona del córtex motor, desechando el resto. Estos son, de acuerdo con el estándar 10/10: FZ, F4, F3, FC5, FC1, FC2, FC6, T7, C3, CZ, C4, T8, CP5, CP1, CP2, CP6, P3, PZ y P4 (ver Figura 3). Para finalizar el preprocesamiento, se aplicó mediante software un filtro paso-banda Butterworth de orden 2 entre 5Hz y 70Hz.

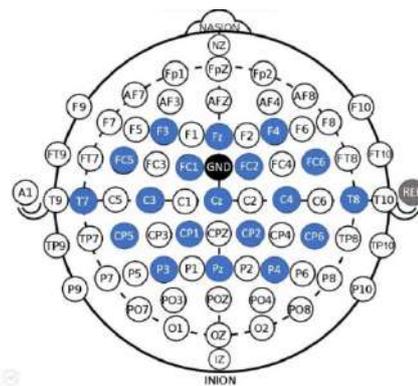
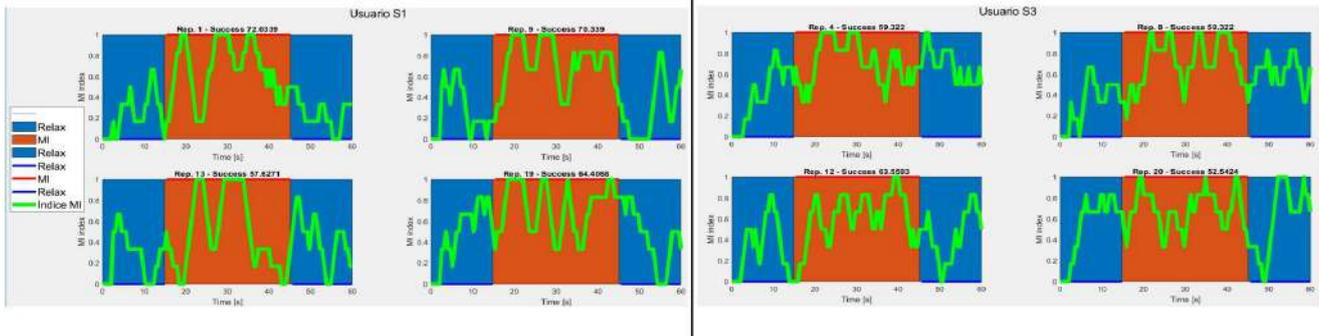


Figura 3. Distribución de electrodos empleada para los análisis (en azul). Como se puede observar, el electrodo de tierra (GND, 'Ground') estaba situado en la posición de FCz, mientras que el de referencia (REF) se colocó mediante pinza en el lóbulo de la oreja derecha del usuario.

Tras el acondicionamiento de las señales, se procedió con su análisis para la generación de los modelos predictivos. Para tal fin, se efectuó en primer lugar una extracción de características a las señales filtradas. Esta se ha basado en



**Figura 4.** Resultados de decodificación del índice de MI en varias repeticiones en parado del sujeto S1 (izquierda) y en movimiento del sujeto S3 (derecha). Idealmente, el índice debería valer 0 durante los dos tramos completos de relax (rectángulos azules) y 1 durante el tramo de MI (rectángulo naranja). ‘Success’ hace referencia al porcentaje de éxito del modelo predictivo en cada caso.

el filtro CSP o *Common Spatial Patterns* (Patrones Comunes Espaciales) [14]. CSP estima la transformación espacial que maximiza la discriminación entre dos patrones cerebrales. Para aplicar esta transformación y obtener las características, se tomaron ventanas de procesamiento de 1 segundo, que se fueron desplazando medio segundo por toda la señal (se obtienen características cada medio segundo, basadas en el segundo inmediatamente anterior de señal).

Con esto, fue posible aplicar un clasificador que determinara si estas ventanas de procesamiento correspondían con instantes de relax, o con instantes de MI. El clasificador seleccionado es el LDA o Linear Discriminant Analysis (Análisis Discriminante Lineal) [15]. Este recibe datos EEG de entrenamiento y aplica un análisis basado en el Teorema de Bayes para generar un modelo capaz de clasificar conjuntos de datos externos a los de entrenamiento, devolviendo valor 1 en las ventanas que detecte MI, y valor 0 en las que detecte relax. De esta manera se puede saber cuál es el porcentaje de ventanas clasificadas de manera correcta.

Para cada usuario se generaron 20 modelos, de los cuales 10 eran para el paradigma del dispositivo parado, y los 10 restantes para el aparato en movimiento. Se debe discriminar entre usuarios ya que cada persona tiene una forma de imaginar y relajarse distinta, de manera que los modelos predictivos de MI son personales. Para cada modelo se efectuó una validación cruzada de 10 repeticiones, (las 10 en parado o las 10 en movimiento que efectuó cada usuario), de las cuales 9 se empleaban para entrenamiento y la décima para test. Así, se tomaban nueve repeticiones a las que se les aplicaba el filtro CSP para calcular la transformación espacial discriminante óptima de ese conjunto de datos, aplicándola y mediante LDA se clasificaban las tareas mentales de la prueba de test. De este modo se puede comparar los resultados obtenidos con el clasificador con las tareas reales, ya que se conoce cuál era la tarea mental ejecutada por el usuario en cada momento. Todo este protocolo se repetía, empleando cada vez una repetición diferente para test.

### 3. Resultados y discusión

La Figura 4 (izquierda) muestra un resumen de los resultados obtenidos en la decodificación de la imaginación motora para el paradigma del dispositivo

cíclico de pedaleo en parado, mientras que la Figura 4 (derecha) muestra otro resumen del funcionamiento de los modelos generados para el paradigma en movimiento. En la Figura 4, los rectángulos azules representan tramos de relax, mientras que los naranjas hacen referencia a los tramos de MI. La línea verde muestra el índice de MI decodificado frente a relax (la media de los tres segundos inmediatamente anteriores, es decir, las 6 ventanas de procesamiento previas). De este modo, con una clasificación ideal, este índice debería ser mínimo en los tramos de relax (0), y máximo en los de imaginación motora (1). A su vez, la Tabla 1 muestra un resumen de los porcentajes de acierto conseguidos para ambos paradigmas en el caso de los cuatro sujetos del estudio.

**Tabla 1.** Resumen de los resultados obtenidos en el estudio. La tabla muestra los diferentes porcentajes de acierto medios obtenidos por los modelos predictivos, junto con sus respectivas desviaciones estándar.

Sujeto	Paradigma	Relax	MI
S1	Parado	60.0±11.5	67.5±14.0
	Movimiento	56.0±7.7	64.1±8.4
	Valor medio	58.0±9.7	65.8±11.2
S2	Parado	35.4±11.0	71.5±11.0
	Movimiento	35.6±12.9	63.4±10.4
	Valor medio	35.5±12.0	67.5±10.7
S3	Parado	30.0±11.8	73.1±16.2
	Movimiento	48.8±12.0	63.2±9.3
	Valor medio	39.4±11.9	68.1±12.7
S4	Parado	41.7±11.3	63.7±16.2
	Movimiento	53.7±16.0	53.6±22.6
	Valor medio	47.7±13.7	58.6±19.4
<b>Valores medios</b>	<b>Parado</b>	<b>41.8±11.4</b>	<b>68.9±14.4</b>
	<b>Movimiento</b>	<b>48.5±12.2</b>	<b>61.1±12.7</b>
	<b>Valor medio</b>	<b>45.2±11.8</b>	<b>65.0±13.5</b>

Como puede observarse, especialmente en la Tabla 1, los resultados no son especialmente buenos, principalmente en relax, sobre todo si se comparan con los logrados en otros estudios previos similares como [9] ó [13] donde los porcentajes de acierto rara vez quedan por debajo de 70%. No obstante, esta investigación no deja de ser un estudio preliminar al que se le pueden aplicar numerosas mejoras,

como estudios de preselección de electrodos, discriminación por bandas de frecuencia, añadir el concepto de atención a la imaginación motora [9], o probar otras estrategias de acondicionamiento de señal (filtrado y estandarizado). Estos progresos mejorarán la calidad de los modelos clasificadores y podrían permitir el control en tiempo real de estos dispositivos de pedaleo en estudios futuros, abriendo la puerta a desarrollar terapias de rehabilitación seguras y efectivas para pacientes reales.

#### 4. Conclusiones

A pesar de ser un estudio preliminar, es posible extraer algunas conclusiones de los datos obtenidos. Una de ellas es evidente, especialmente clara en la Tabla 1: este método de decodificación genera resultados mucho mejores en la MI (65% de acierto de media) que en el relax (45%). Esto sugiere que, para mejorar la clasificación del relax, sean necesarios más datos de relax para el entrenamiento, es decir, que aumentar el tiempo de relax en el protocolo experimental podría incrementar este valor. Además, parece que el relax funciona mejor en movimiento, mientras que la MI en parado, lo cual es una buena señal ya que se busca reconocer bien el relax en el paradigma de movimiento porque supone el comando de parada del cicloergómetro, mientras que la MI en parado corresponde con la orden de arranque del dispositivo; son las casuísticas que realmente interesa detectar. Respecto a los valores de dispersión, no son muy elevados, pero igualmente conviene buscar técnicas para disminuirlos ya que picos altos de dispersión podrían provocar activaciones involuntarias del dispositivo de pedaleo, que pueden resultar peligrosas, especialmente cuando lo manipule un paciente con disfunciones motoras.

En cualquier caso, para mejorar este estudio preliminar será necesario realizar nuevas pruebas experimentales con más sujetos para obtener una mayor cantidad de datos que permita obtener conclusiones más sólidas así como perfeccionar los procesamientos y clasificación para optimizar los resultados.

#### Agradecimientos

Esta investigación ha sido financiada por una ayuda para la formación predoctoral en colaboración con empresas - modalidad A: incentivos para contratos predoctorales suscritos por la Universidad Miguel Hernández de Elche y financiados por la Generalitat Valenciana para la elaboración de una tesis con mención industrial. Agradecer también a Javier Castilla su apoyo en el proyecto.

#### Referencias

- [1] M. Tangermann *et al.*, “Review of the BCI competition IV,” *Frontiers in Neuroscience*, no. JULY. 2012, doi: 10.3389/fnins.2012.00055.
- [2] M. Misaki *et al.*, “Beyond synchrony: The capacity of fMRI hyperscanning for the study of human social interaction,” *Soc. Cogn. Affect. Neurosci.*, vol. 16, no. 1–2, 2021, doi: 10.1093/scan/nsaa143.
- [3] E. Sareen, S. Zahar, D. Van De Ville, A. Gupta, A. Griffa, and E. Amico, “Exploring MEG brain fingerprints: Evaluation, pitfalls, and interpretations,” *Neuroimage*, vol. 240, 2021, doi: 10.1016/j.neuroimage.2021.118331.
- [4] G. Garipelli, R. Chavarriaga, and J. Del R Millán, “Single trial analysis of slow cortical potentials: a study on anticipation related potentials,” *J. Neural Eng.*, vol. 10, no. 3, Jun. 2013, doi: 10.1088/1741-2560/10/3/036014.
- [5] M. Ortiz, L. Ferrero, E. Iáñez, J. M. Azorín, and J. L. Contreras-Vidal, “Sensory Integration in Human Movement: A New Brain-Machine Interface Based on Gamma Band and Attention Level for Controlling a Lower-Limb Exoskeleton,” *Front. Bioeng. Biotechnol.*, vol. 0, p. 735, Sep. 2020, doi: 10.3389/FBIOE.2020.00735.
- [6] C. SC, “Repairing the human brain after stroke. II. Restorative therapies,” *Ann. Neurol.*, vol. 63, no. 5, pp. 549–560, May 2008, doi: 10.1002/ANA.21412.
- [7] A. Gharabaghi, “What Turns Assistive into Restorative Brain-Machine Interfaces?,” *Front. Neurosci.*, vol. 0, no. OCT, p. 456, Oct. 2016, doi: 10.3389/FNINS.2016.00456.
- [8] C.-V. JL *et al.*, “Powered exoskeletons for bipedal locomotion after spinal cord injury,” *J. Neural Eng.*, vol. 13, no. 3, Apr. 2016, doi: 10.1088/1741-2560/13/3/031001.
- [9] L. Ferrero, V. Quiles, M. Ortiz, E. Iáñez, and J. M. Azorín, “A BMI Based on Motor Imagery and Attention for Commanding a Lower-Limb Robotic Exoskeleton: A Case Study,” *Appl. Sci. 2021, Vol. 11, Page 4106*, vol. 11, no. 9, p. 4106, Apr. 2021, doi: 10.3390/APP11094106.
- [10] E. Ambrosini *et al.*, “Neuro-Mechanics of Recumbent Leg Cycling in Post-Acute Stroke Patients,” *Ann. Biomed. Eng.*, vol. 44, no. 11, pp. 3238–3251, 2016, doi: 10.1007/s10439-016-1660-0.
- [11] M. Seeber, R. Scherer, J. Wagner, T. Solis-Escalante, and G. R. Müller-Putz, “High and low gamma EEG oscillations in central sensorimotor areas are conversely modulated during the human gait cycle,” *Neuroimage*, vol. 112, pp. 318–326, May 2015, doi: 10.1016/J.NEUROIMAGE.2015.03.045.
- [12] V. Quiles, E. Ianez, M. Ortiz, N. Medina, A. Serrano, and J. M. Azorin, “Lessons Learned from Clinical Trials of a Neurorehabilitation Therapy Based on tDCS, BMI, and Pedaling Systems,” *IEEE Syst. J.*, vol. 15, no. 2, pp. 1873–1880, Jun. 2021, doi: 10.1109/JSYST.2020.3026242.
- [13] M. Ortiz, M. Rodríguez-Ugarte, E. Iáñez, and J. M. Azorín, “Analysis of algorithms for detection of pedaling intention in brain-machine interfaces,” *RIAI - Rev. Iberoam. Autom. e Inform. Ind.*, vol. 16, no. 2, pp. 222–231, 2019, doi: 10.4995/RIAI.2018.9861.
- [14] H. Ramoser, J. Müller-Gerking, and G. Pfurtscheller, “Optimal spatial filtering of single trial EEG during imagined hand movement,” *IEEE Trans. Rehabil. Eng.*, vol. 8, no. 4, pp. 441–446, 2000, doi: 10.1109/86.895946.
- [15] A. J. Izenman, “Modern Multivariate Statistical Techniques,” 2008, doi: 10.1007/978-0-387-78189-1.

# Propuesta de un sistema domótico de fácil configuración controlado por la actividad cerebral y aplicado a personas afectadas con ELA.

F. Velasco-Álvarez<sup>1</sup>, A. Fernández-Rodríguez<sup>1</sup>, JM. André<sup>2</sup>, R. Ron-Angevin<sup>1</sup>

<sup>1</sup>Dpto de Tecnología Electrónica, Universidad de Málaga, Málaga, España, {fvelasco, afernandezrguez, rron}@uma.es

<sup>2</sup>Laboratoire IMS, Bordeaux INP-ENSC, Talence, Francia, jean-marc.andre@ensc.fr

## Resumen

Las interfaces cerebro-ordenador (BCI, de *brain-computer interface*) permiten utilizar la actividad cerebral de un usuario como canal de comunicación para interactuar con el entorno. La herramienta UMA-BCI Speller permite una fácil configuración de un sistema BCI, permitiendo una fácil manipulación sin demasiados conocimientos técnicos. El objetivo del presente trabajo es presentar un sistema domótico de fácil configuración que permita controlar, a través de un sistema BCI, aplicaciones de gran interés para pacientes afectados de Esclerosis Lateral Amiotrófica (ELA). Para ello, el sistema domótica hará uso de comandos de voz, los cuales serán interpretados por un asistente virtual. Las aplicaciones y dispositivos que pueden ser controlados en este prototipo de sistema domótico son: WhatsApp, Spotify, Google Nest, una bombilla inteligente, un enchufe inteligente (para encender y apagar una radio) y un mando de infrarrojos (para controlar una televisión y un aire acondicionado). En este estudio concreto se evaluó el sistema sobre 3 pacientes. Los resultados obtenidos mostraron la usabilidad del sistema, demostrando con éxito el uso de un sistema BCI para el control de un sistema domótico flexible y que puede ser adaptado a las necesidades de los usuarios.

## 1. Motivación

Las interfaces cerebro-ordenador (BCI, de *Brain-Computer Interface*) son un tipo de tecnología que utiliza la señal cerebral de los usuarios para establecer un canal de comunicación entre ellos y un dispositivo externo [1]. Por tanto, esta tecnología puede ser una opción adecuada para aquellas personas que presentan severamente deterioradas sus capacidades motoras. La técnica de neuroimagen más utilizada por las BCI es la electroencefalografía (EEG), posiblemente debido a su portabilidad, relativo bajo coste y alta resolución temporal [2]. Debido a ello, para el presente trabajo se decidió usar esta señal para crear la BCI.

Las señales EEG más utilizadas son los potenciales visuales relacionados con eventos (ERP, de *Event-Related Potential*), cambios en la señal de EEG evocados como respuesta a un determinado evento visual. Estas BCI basadas en ERP (ERP-BCI) visuales han ofrecido en general un buen rendimiento y un elevado número de comandos disponibles [3]. Además, estos sistemas ya han mostrado resultados prometedores en usuarios con graves problemas de control muscular (p. ej., [4], [5]). Normalmente, los ERP se evocan a través del paradigma

*oddball*, en el que un conjunto de estímulos visuales se presenta de forma pseudoaleatoria mientras el usuario presta atención sólo a uno de ellos, lo que resulta en un ERP después de la estimulación de este elemento deseado (es decir, el estímulo objetivo). Por lo tanto, el objetivo de la BCI es detectar el ERP asociado a la presentación del estímulo objetivo, que tiene una orden de control asociada. Las aplicaciones BCI previas basadas en EEG han sido extensas y preparadas para el control de diversos dispositivos [6]. El presente trabajo se centrará en el control de un sistema domótico a través de una ERP-BCI.

Gracias a una ERP-BCI, el usuario puede seleccionar diferentes estímulos asociados a, por ejemplo, determinadas letras o palabras para construir una frase. A continuación, esta frase puede ser convertida a voz por el ordenador e interpretados por un asistente virtual. Esta idea de enlazar la BCI con un sistema domótico a través de comandos de voz fue empleada previamente por [7] utilizando la señal de respuesta auditiva de estado estable (ASSR, de *Auditory Steady State Response*). En concreto, dicho trabajo utilizó Alexa como asistente de voz, OpenBCI como software BCI, y dos placas Arduino (una utilizada para reproducir archivos de sonido predeterminados y otra equipada con un sensor de proximidad, ver [7] para más detalles). El inconveniente de este sistema era que, para adaptar su uso a nuevos dispositivos, se necesitaría personal técnico especializado o conocimientos de programación elevados. Por tanto, sería interesante utilizar una BCI que pueda ser fácilmente configurada por los cuidadores que, sin necesidad de conocimientos técnicos, les permita adaptar diversas aplicaciones y dispositivos del entorno útiles para el paciente. Esta adaptación podría ser proporcionada por el UMA-BCI Speller, un *software* fácilmente configurable que utiliza los ERP visuales y permite una gran flexibilidad en la creación de la interfaz gráfica de usuario [8].

En definitiva, el objetivo del presente trabajo es presentar un sistema de control domótico que sea flexible y, por tanto, pueda adaptarse a los dispositivos y aplicaciones que el usuario desee controlar. Hasta donde sabemos, este es el primer trabajo relativo a una BCI fácilmente configurable que propone el control domótico de varios dispositivos y aplicaciones.

## 2 Método

### 2.1 Participantes

La finalidad de este estudio era probar el sistema domótico desarrollado sobre pacientes afectados de Esclerosis Lateral Amiotrófica (ELA), siendo los potenciales usuarios de las interfaces cerebro-computadora. En el presente estudio participaron 3 pacientes (edad  $68 \pm 9$  años, 2 hombres y 1 mujer) denominados P01-P03. El estudio fue aprobado por el Comité de Ética de la Universidad de Málaga y cumplió las normas éticas de la Declaración de Helsinki.

### 2.2 Adquisición de datos y procesamiento de la señal

El EEG se registró a una frecuencia de muestreo de 250 Hz utilizando las posiciones de los electrodos: Fz, Cz, Pz, Oz, P3, P4, PO7 y PO8, según el sistema internacional 10/20. Todos los canales se referenciaron al mastoide izquierdo y se conectaron a tierra en la posición AFz. Las señales se amplificaron con un amplificador acti-CHamp (Brain Products GmbH, Gilching, Alemania). Todos los aspectos relativos a la recogida y el procesamiento de los datos del EEG fueron controlados por el software BCI2000 [9]. Dicho software empleó el análisis lineal discriminante de pasos hacia delante (SWLDA, de *Stepwise Linear Discriminant Analysis*) de los datos obtenidos a través del EEG para crear el clasificador que permitiría realizar la selección de comandos en la fase online.

### 2.3 Implementación del sistema

El objetivo de la BCI era generar comandos de voz que pudieran ser interpretados por el Asistente de Google, el cual controlaría las aplicaciones y dispositivos externos. Por un lado, en un ordenador portátil, se ejecutaba el UMA-BCI Speller que presentaba los estímulos y registraba y analizaba la señal de EEG. Por otro lado, en un *smartphone*, se ejecutaba el Asistente de Google que recibía, interpretaba y ejecutaba las órdenes de voz enviadas por el UMA-BCI Speller.

En el sistema domótico desarrollado se implementaron las siguientes aplicaciones y dispositivos: WhatsApp, Spotify, Google Nest, una bombilla inteligente, un enchufe inteligente (para encender/apagar una radio) y un controlador de infrarrojos (para controlar una televisión y un aire acondicionado). El Asistente de Google podía controlar directamente WhatsApp para enviar mensajes, y Spotify para reproducir música. También era posible comunicarse directamente con el asistente para hacer peticiones específicas (p. ej., poner una alarma a las 8:00 o preguntar qué temperatura hará mañana). Por otro lado, el Asistente de Google se comunicaba a través de una conexión Wi-Fi con diferentes dispositivos: una bombilla inteligente, un enchufe inteligente que permitía encender y apagar una radio, el dispositivo Google Nest y un controlador que emitía infrarrojos y que gestionaba otros dos dispositivos (un televisor y un aire acondicionado). El controlador de infrarrojos recibía las órdenes del Asistente de Google a través del Wi-Fi y emulaba un mando a distancia convencional. Además, el *software* UMA-BCI Speller incluía una función de predicción de texto – basada

en los caracteres ya escritos en esa palabra y la probabilidad de aparición – que puede ayudar a los usuarios a la hora de escribir las palabras que constituirán el comando a enviar al Asistente de Google.

Para evitar la influencia del ruido ambiental en la comprensión de la orden por parte del asistente virtual, las órdenes de voz provenientes del UMA-BCI Speller se enviaban al *smartphone* a través de una conexión por cable (minijack) que conectaba la salida de audio del portátil con la entrada del micrófono del *smartphone*. El volumen de salida del portátil se mantuvo fijo durante todo el experimento, por lo que el asistente siempre recibió el mismo nivel de audio. Se utilizó el Narrador de Windows 10 (una función de texto a voz), en particular la voz denominada “Microsoft Helena” del catálogo de voces en español. Como el asistente virtual utilizado en el experimento era el Asistente de Google, cada comando comenzaba con las palabras “Ok Google...”, que es una de las palabras clave de activación del asistente. El usuario recibía información auditiva por dos conjuntos de altavoces: en el primer conjunto, se reproducía el comando de voz enviado por el UMA-BCI Speller; mientras que, en el segundo conjunto, se reproducía la respuesta del Asistente de Google o la música de Spotify.

### 2.4 Paradigma de control

Para enviar una orden al asistente virtual, los usuarios tenían que seleccionar los elementos apropiados en diferentes menús. La selección de un elemento seguía el procedimiento habitual en un paradigma de fila-columna (RCP, de *Row-Column Paradigm*): los usuarios tenían que prestar atención al elemento deseado (p. ej., el elemento “Radio ON” para encender la radio) y contar mentalmente el número de veces que aparece una imagen superpuesta a ese elemento (p. ej., el icono de una radio en el caso del elemento “Radio ON”). En el RCP, las imágenes son presentadas sobre los elementos siguiendo un patrón de filas y columnas en orden aleatorio. La duración de cada presentación de imágenes era de 192 ms, y el intervalo entre la desaparición de un conjunto de imágenes (ya sea en fila o columna) y la aparición del siguiente conjunto era de 32 ms. El tiempo de cada selección para todos los menús era el mismo, ya que todas las interfaces consistían en una matriz de  $7 \times 7$ , aunque en tres de ellas había elementos ficticios (es decir, elementos que no tenían ningún comando asociado y, por tanto, no tenían ningún efecto al ser seleccionados). Para controlar el sistema domótico, se configuraron seis menús usando la herramienta UMA-BCI Speller (figura 1). Navegando por estos menús, los participantes seleccionaban diferentes elementos para formar una frase que finalmente se convertiría en audio para proporcionar una orden específica.

### 2.5 Procedimiento

Los experimentos se llevaron a cabo en el domicilio de cada uno de los pacientes. Para cada uno de ellos se dedicó el tiempo necesario para explicar el funcionamiento del sistema. Una vez explicada la tarea y preparada la instrumentación necesaria para el registro de la actividad EEG, el experimento constó de dos partes: i) una fase de

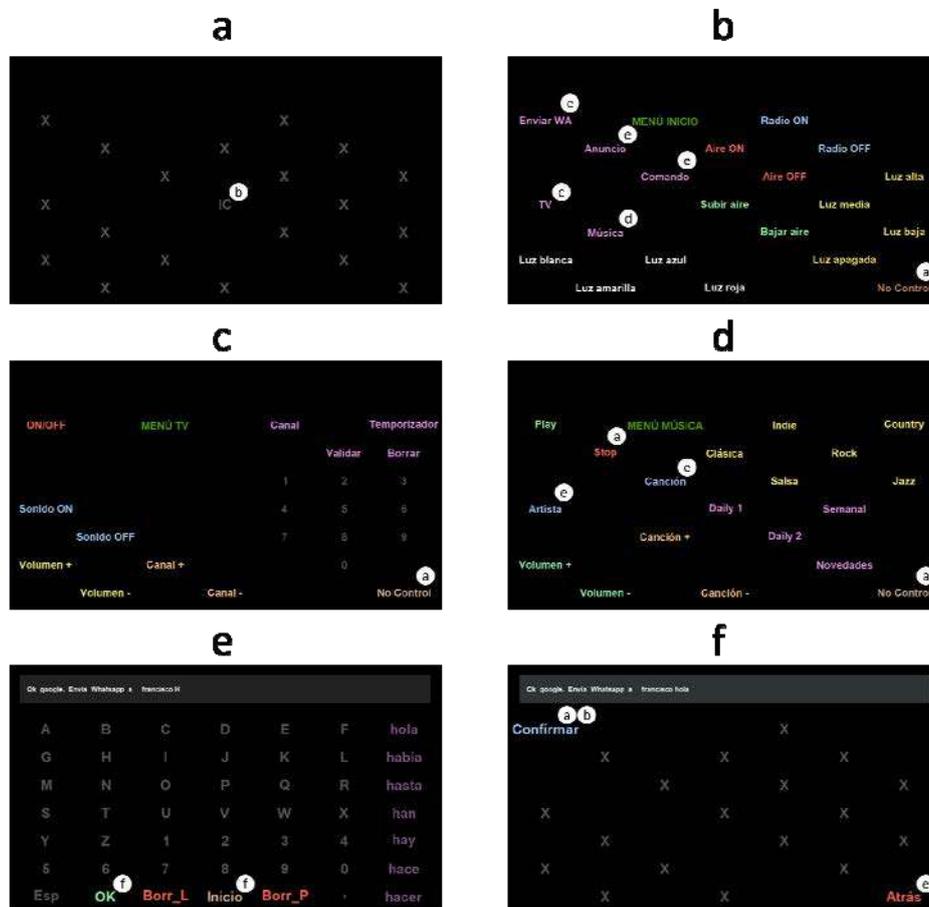
calibración utilizada para obtener los parámetros del clasificador del EEG dependientes del sujeto, y ii) una fase experimental *online* en la cual se ponía a prueba el sistema. El experimento tuvo una duración de 100-130 minutos. En el caso de los pacientes P01 y P03 fueron necesarias dos visitas, al no poder acabar la experimentación en una única intervención por las circunstancias de dichos pacientes

La fase de calibración consistió en prestar atención a 12 elementos predeterminados (“DOMOTICA2021”) en el menú de escritura (figura 1e). El número de secuencias (es decir, el número de veces que se resaltaba cada fila y columna) se fijó en diez, por lo que cada estímulo aparecía 20 veces. El tiempo de calibración duró, en promedio, 8:31 minutos. Tras la calibración, se realizó un SWLDA para obtener el clasificador P300 específico del usuario. Según los resultados de *accuracy* obtenidos con dicho clasificador, se ajustó también el número de secuencias que sería empleado en la fase *online* para cada sujeto. El criterio para elegir este número de secuencias fue maximizar la tasa de símbolos escritos (WSR, de *Written Symbol Rate*). Sin embargo, este criterio se adaptó para que el número de secuencias resultara cómodo a cada paciente. Para ello, se iba variando el número de secuencias y probando mediante la escritura de la palabra “hola” hasta que el paciente diera su aprobación definitiva.

La fase *online* consistió en realizar varias tareas con la finalidad de probar el control con las diferentes

aplicaciones y dispositivos. Todos los participantes realizaron las mismas tareas y en el mismo orden. En caso de que el participante se equivocara (aunque ese error completara una tarea posterior), debía corregir la selección y continuar por donde iba. Al no disponer de aparato de aire acondicionado en el domicilio de los pacientes, las tareas asociadas a este equipo fueron omitidas. De este modo, las tareas a realizar, detalladas por orden, fueron las siguientes (los comandos específicos requeridos para completar cada tarea están entre paréntesis):

- El participante comienza la tarea en el menú de no control.
- Ir al menú de inicio (“IC”) y encender la lámpara a alta intensidad (“Luz alta”).
- Encender la radio (“Radio ON”) y escucharla durante un minuto en el menú de no control (“NC”). Apagarla después de ese minuto (“IC”, “Radio OFF”).
- Acceder al menú Comando para introducir “tiempo hoy” (“Comando”, “T”, “tiempo”, “H”, “hoy”, “OK”, “Confirmar”).
- Acceder al menú Inicio (“IC”), luego al menú Música (“Música”), elegir un género libre (por ejemplo, “Jazz”), subir el volumen (“Volumen +”), seleccionar la siguiente canción (“Canción +”) y pasar al menú de no control (“No control”).
- Ir al menú Inicio (“IC”) y, a continuación, ir al menú de escritura relativo a WhatsApp para enviar un mensaje a un contacto llamado “Francisco” con el texto “hola”. Tanto el



**Figura 1:** Navegación entre los seis menús disponibles: a) menú de no control, en el cual el único comando disponible estaba destinado a comenzar el control intencional en el siguiente menú; b) menú de inicio, en el cual se selecciona el dispositivo o aplicación a controlar; c) el menú TV, relativo al control de la televisión; d) el menú música, relativo al control de Spotify; e) el menú de escritura, requerido para la construcción de texto personalizado (p. ej., para un mensaje de Whatsapp); y f) menú de confirmación, empleado para confirmar la ejecución del comando. La letra en el interior de los círculos blancos indica a qué menú redirigía dicho comando; este círculo no era presentado en la aplicación. 194

nombre del receptor como la palabra “hola” fueron propuestos como predicciones por el sistema al seleccionar la primera letra. Se necesitaron un mínimo de 8 acciones para completar esta tarea (“IC”, “Enviar WA”, “F”, “Francisco”, “H”, “hola”, “OK”, “Confirmar”).

– Acceder al menú de Inicio (“IC”), y luego al de Música (“Música”) para desactivar Spotify y volver a NC (“Stop”).

– Encender la luz a una intensidad baja (“Luz baja”) y ponerla en azul (“Luz azul”).

– Acceder al menú TV (“TV”) para encender la tv (“ON/OFF”), poner el canal 7 (“Canal”, “7”, “Validar”), subir el volumen (“Vol+”), poner un temporizador de 2 minutos (“Temporizador”, “2”, “Validar”) y permanecer en NC hasta que se apague la tv (“NC”).

– Ir al menú Inicio (“IC”), y apagar la lámpara (“Apagar luz”).

– Acceder al menú Anuncio (“Anuncio”) para escribir “fin” (“F”, “fin”, “OK”, “Confirmar”) y comunicar a través del dispositivo Google Nest que el experimento ha finalizado.

### 3 Resultados

En la fase de calibración los pacientes alcanzaron el 100% de *accuracy* con 4, 5 y 7 secuencias para P01, P02 y P03 respectivamente. Sin embargo, tras la adaptación, las secuencias elegidas fueron: P01, 7; P02, 6 y P03, 7.

En la fase on-line, la tasa de éxito obtenida para cada participante fue: P01, 78.87%; P02, 73.85% y P03, 60.56%. En cuanto al tiempo (min) requerido para realizar todas las tareas: P01, 40.66; P02, 34.3 y P03, 39.56

Tanto P02 como P03 tuvieron ciertas complicaciones en ejecutar algunas tareas. A P02 se le olvidó encender la bombilla al inicio del experimento. Aún así los experimentadores optaron por que el experimento continuara. En cuanto al paciente P03, tuvo varios fallos consecutivos a la hora de seleccionar el estilo de música. Debido a ello, y para no alargar demasiado la sesión, se decidió simplificar las tareas asociadas al TV, limitándose a encender, incrementa el volumen y establecer un temporizador de 2 minutos.

### 4 Discusión y conclusiones

En este trabajo se ha presentado un procedimiento que permite controlar un sistema domótico a través de la actividad cerebral de un usuario. Si bien ya existen algunos trabajos que reportan el control de algunos dispositivos a través de un ERP-BCI, por ejemplo [9], estos requieren una configuración compleja y específica para cada dispositivo. En este trabajo, se ha demostrado que es posible implementar una BCI para el control de diversos dispositivos y aplicaciones sin necesidad de elevados conocimientos técnicos, de forma flexible y que, por tanto, puede adaptarse a las necesidades de un usuario. En general, resulta complejo adaptar un sistema BCI para el control de dispositivos y aplicaciones externas. Sin embargo, el control por voz es cada vez más común en los dispositivos comerciales. Por lo tanto, la idea de esta propuesta facilita el trabajo a la hora de controlar estos dispositivos previamente adaptados a los comandos de voz. Es importante recordar que los usuarios potenciales

de estos sistemas (es decir, los pacientes o sus cuidadores) pueden no tener los conocimientos técnicos necesarios para manipular sistemas complejos. Además, el software UMA-BCI Speller ha demostrado ser una herramienta útil y adecuada para controlar numerosos dispositivos del entorno mediante la construcción de comandos de voz. El sistema ha sido probado sobre 3 pacientes afectados de ELA, obteniendo resultados satisfactorios teniendo en cuenta el elevado número de tareas implicadas en la experimentación. En ese sentido, si bien en trabajos futuros deberían hacerse test de usabilidad estandarizados, por ejemplo, el SUS, o el NASA-TLX para medir la carga de trabajo, los resultados obtenidos sí permiten adelantar que el sistema puede ser usable y, sin lugar a duda, ofrecer una mejor calidad de vida a estos pacientes, dotándoles de una herramienta que les permita incrementar sus opciones de comunicación, ocio y control.

### Agradecimientos

Este trabajo ha sido financiado conjuntamente por el Ministerio de Ciencia, Innovación y Universidades y el Fondo Europeo de Desarrollo Regional (FEDER) a través de los proyectos SICCAU, referencia: RTI2018-100912-B-I00, y SICODIS, referencia: PID2021-1272610B-I00.

### Referencias

- [1] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, “Brain-computer interfaces for communication and control,” *Clin. Neurophysiol.*, vol. 113, no. 6, pp. 767–791, 2002.
- [2] L. F. Nicolas-Alonso and J. Gomez-Gil, “Brain computer interfaces, a review,” *Sensors*, vol. 12, no. 2, pp. 1211–1279, 2012.
- [3] B. Z. Allison, A. Kübler, and J. Jin, “30+ years of P300 brain-computer interfaces,” *Psychophysiology*, vol. 57, no. 7, pp. 1–18, 2020.
- [4] K. Utsumi, K. Takano, Y. Okahara, T. Komori, O. Onodera, and K. Kansaku, “Operation of a P300-based braincomputer interface in patients with Duchenne muscular dystrophy,” *Sci. Rep.*, vol. 8, no. 1, pp. 4–11, 2018.
- [5] M. T. Medina-Juliá, A. Fernández-Rodríguez, F. Velasco-Álvarez, and R. Ron-Angevin, “P300-Based Brain-Computer Interface Speller: Usability Evaluation of Three Speller Sizes by Severely Motor-Disabled Patients,” *Front. Hum. Neurosci.*, vol. 14, no. October, pp. 1–10, 2020.
- [6] S. Saha *et al.*, “Progress in Brain Computer Interface: Challenges and Opportunities,” *Front. Syst. Neurosci.*, vol. 15, no. February, pp. 1–20, 2021.
- [7] V. K. K. Shivappa, B. Luu, M. Solis, and K. George, “Home automation system using brain computer interface paradigm based on auditory selection attention,” *I2MTC 2018 - 2018 IEEE Int. Instrum. Meas. Technol. Conf. Discov. New Horizons Instrum. Meas. Proc.*, pp. 1–6, 2018.
- [8] F. Velasco-Álvarez, S. Sancha-Ros, E. García-Garaluz, Á. Fernández-Rodríguez, M. T. T. Medina-Juliá, and R. Ron-Angevin, “UMA-BCI Speller: an Easily Configurable P300 Speller Tool for End Users,” *Comput. Methods Programs Biomed.*, vol. 172, pp. 127–138, 2019.
- [9] R. Corralejo, L.F. Nicolás-Alonso, D. Álvarez, R. Hornero, “A P300-based brain-computer interface aimed at operating electronic devices at home for severely disabled people”, *Med. Biol. Eng. Comput.*, 52 (2014), pp. 861-872.

# Un nuevo método de parada temprana no paramétrico para sistemas Brain–Computer Interface basados en c-VEP

Víctor Martínez-Cagigal<sup>1,2</sup>, Eduardo Santamaría-Vázquez<sup>1,2</sup>, Sergio Pérez-Velasco<sup>1</sup>,  
Diego Marcos-Martínez<sup>1</sup>, Selene Moreno-Calderón<sup>1</sup> y Roberto Hornero<sup>1,2</sup>

<sup>1</sup> Grupo de Ingeniería Biomédica, Universidad de Valladolid, Valladolid, España

<sup>2</sup> Centro de Investigación Biomédica en Red de Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), España  
{[victor.martinez@gib.tel.uva.es](mailto:victor.martinez@gib.tel.uva.es), [robhor@tel.uva.es](mailto:robhor@tel.uva.es)}

## Resumen

Los sistemas brain–computer interface (BCI) basados en potenciales evocados visuales modulados por código (c-VEP) destacan por alcanzar excelentes precisiones de selección de comandos con tiempos de calibración muy reducidos. Uno de los pasos naturales para democratizar su uso en entornos *plug-and-play* es desarrollar algoritmos de detección temprana que permitan detectar en tiempo real el número mínimo de repeticiones del código necesarias para ofrecer una selección fiable. Sin embargo, este tipo de técnicas son escasas en el estado de la técnica actual para sistemas c-VEP-BCI basados en el paradigma clásico de desplazamiento circular. En esta comunicación se propone un novedoso método de parada temprana no paramétrico, que aproxima la distribución de los comandos no atendidos a una distribución normal y emite una selección cuando la correlación del comando a seleccionar supera el percentil 99.87% de dicha distribución. La propuesta se ha evaluado de forma *offline* con 6 sujetos de control, alcanzando una precisión media del 97.92% y una velocidad de 1.24 s/comando. Asimismo, el algoritmo también se ha evaluado con un sujeto adicional de forma *online*, como prueba de concepto para validar su viabilidad técnica, alcanzando una precisión media de 96.88% con una velocidad de 1.67 s/comando. Estos resultados sugieren que la aplicación del algoritmo propuesto es factible de forma *online*, reduciendo significativamente el tiempo de selección necesario sin afectar negativamente a la precisión del sistema.

## 1. Introducción

Los sistemas brain–computer interface (BCI) no invasivos habitualmente monitorizan la señal de electroencefalografía (EEG) para detectar las intenciones del usuario y transformarlas en comandos de una aplicación o dispositivo externo [1]. Dado que dichas intenciones no se reflejan directamente en el EEG del usuario, su decodificación no es trivial, siendo necesario emplear distintas señales de control; i.e., estrategias que generan cambios medibles en el EEG mediante la realización de tareas cognitivas (i.e., endógenas) o la recepción de estímulos externos (i.e., exógenas) [1]. Dentro las estrategias exógenas, recientemente se han propuesto los potenciales evocados visuales modulados por código (c-VEP) como una nueva señal de control que permite alcanzar un alto rendimiento con tiempos de calibración muy reducidos [2].

Los sistemas c-VEP-BCI tradicionales emplean un código pseudoaleatorio cuya función de autocorrelación es prácticamente nula, lo que permite codificar cada uno de los comandos a seleccionar con versiones desplazadas

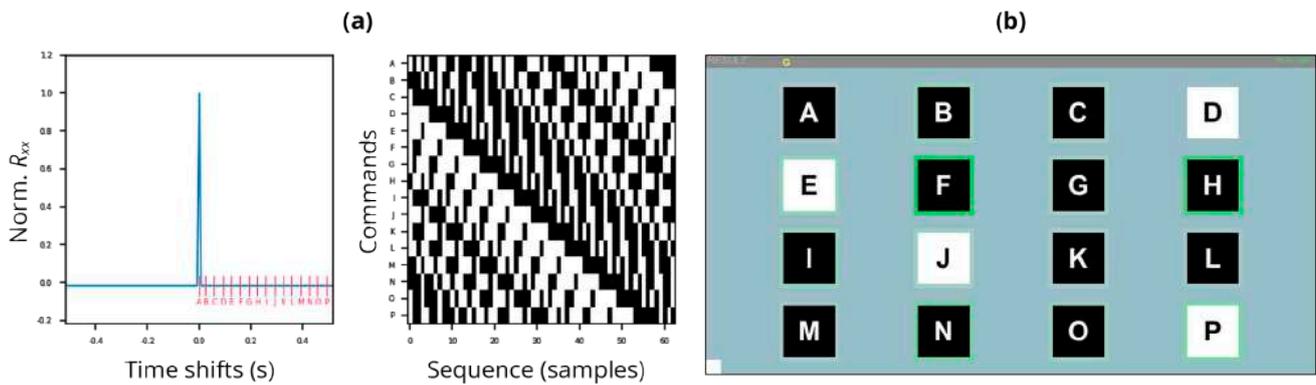
temporalmente de la misma secuencia [2]. En este paradigma, conocido como *circular shifting* o desplazamiento circular, el usuario presta atención al comando que desea seleccionar en cada momento, generando una respuesta en su EEG más correlada con el comando objetivo que con el resto, permitiendo su decodificación en tiempo real [2]. Este método usualmente alcanza precisiones superiores al 90% y tasas de información por minuto (ITR) de hasta 100 bpm (bits por minuto) con una calibración de 10–30 s [3].

El rendimiento de los sistemas c-VEP-BCI sugiere que tiene potencial para evolucionar como una tecnología *plug-and-play*. En este sentido, el desarrollo de técnicas de parada temprana (i.e., *early stopping*) permitiría optimizar el tiempo de selección de cada comando de forma adaptativa. Es decir, permitirían determinar automáticamente el número de repeticiones de la secuencia necesario para efectuar una selección fiable en tiempo real. A pesar de la popularidad de estos sistemas, las técnicas de parada temprana en paradigmas c-VEP son escasas [2]. Además, muchas de las técnicas existentes no están diseñadas para el paradigma de *circular shifting* [4], [5], dependen del clasificador o son paramétricas (i.e., requieren la optimización previa de un parámetro) [6], [7]. Aun así, la aplicación de la parada temprana se ha demostrado beneficiosa para el rendimiento del sistema, alcanzando precisiones superiores al 90% con tiempos de selección de comandos de 3.26 s [6], 4.2 s [5], o 6.17 s [8].

El objetivo de este estudio piloto es proponer una nueva técnica de parada temprana para sistemas c-VEP-BCI basados en *circular shifting*. Asimismo, la técnica propuesta es *filter* (i.e., no depende del clasificador), no es paramétrica (i.e., no necesita entrenar ningún parámetro), no requiere registrar señales EEG adicionales y puede aplicarse en tiempo real.

## 2. Sujetos

La caracterización y evaluación del algoritmo propuesto se ha llevado a cabo de forma *offline* con 6 sujetos de control (edad media:  $26.17 \pm 3.44$  años, 3 varones), y de forma *online* como prueba de concepto con 1 sujeto de control adicional (29 años, varón). Todos ellos dieron su consentimiento para participar en el estudio. La señal se registró empleando 16 canales EEG en las posiciones F3, Fz, F4, C3, Cz, C4, CPz, P3, Pz, P4, PO7, POz, PO8, Oz, I1 e I2, con tierra en AFz y referencia en el lóbulo de la



**Figura 1.** (a) Codificación de los comandos con la  $m$ -secuencia empleada: función de autocorrelación y desfase asociados (izquierda), y codificación visual de los comandos de acuerdo con sus desfases. (b) Captura de la aplicación desarrollada en MEDUSA© durante un ciclo de estimulación.

oreja. Se utilizó un equipo *g.USBamp* (*g.Tec, Guger Technologies*, Austria) con una frecuencia de muestreo de 256 Hz. Tanto la adquisición como el procesamiento de la señal en tiempo real, así como el paradigma c-VEP, se han programado en MEDUSA© ([www.medusabci.com](http://www.medusabci.com)), un sistema de propósito general para el desarrollo de sistemas BCI basado en Python [9].

### 3. Métodos

#### 3.1. Paradigma y procesamiento de señal

La secuencia empleada para codificar los comandos es una  $m$ -secuencia (i.e., de máxima longitud) de 63 bits binaria, generada mediante un registro de desplazamiento con retroalimentación lineal (LFSR, *linear feedback shift register*) utilizando el polinomio  $x^6 + x^5 + 1$  con estado inicial 110000 [2]. Se ha empleado una frecuencia de refresco del monitor de 120 Hz, con lo cual un ciclo completo de la secuencia dura 525 ms. El paradigma c-VEP empleado consta de 16 comandos, codificados con versiones desplazadas de la  $m$ -secuencia con retardos de  $\theta_i = i \cdot \tau$  muestras, donde  $i$  indica el comando y el paso es de  $\tau = 4$  [2]. En la Figura 1(a) se puede observar la codificación de los comandos utilizada.

En la etapa de calibración, se pide al usuario que preste atención al comando codificado con la  $m$ -secuencia original (i.e.,  $\theta = 0$ ), que se repite durante  $k$  ciclos. Tras aplicar el pre-procesado, compuesto por un banco de filtros paso-banda (1–60 Hz, 12–60 Hz y 30–60 Hz) y un *notch* a 50 Hz, se obtienen dos versiones de la respuesta EEG: (1) las épocas concatenadas  $\mathbf{A} \in \mathbb{R}^{[kN_s \times N_c]}$ , donde  $N_s$  es el número de muestras de un ciclo completo y  $N_c$  es el número de canales; y (2) las épocas promediadas a lo largo de los  $k$  ciclos  $\mathbf{B} \in \mathbb{R}^{[N_s \times N_c]}$ . Posteriormente, se aplica un análisis de correlación canónica (CCA, *canonical correlation analysis*) para encontrar los filtros espaciales que maximizan la correlación entre las proyecciones de  $\mathbf{A}$  y  $\mathbf{B}$ . Tras hallar el filtro espacial  $\boldsymbol{\omega}_b$  que maximiza la correlación entre las épocas concatenadas y la respuesta promediada, se proyecta la señal original para obtener la plantilla principal  $\mathbf{x}_0 = \mathbf{B}\boldsymbol{\omega}_b$ , mientras que las plantillas para el resto de comandos se obtienen mediante el desplazamiento circular de la original de  $\theta_i, i = 1 \dots 16$  muestras (para más información sobre el procesamiento de

referencia, consulte [2]). Adicionalmente, con el objetivo de eliminar épocas ruidosas, se descartaron las épocas de calibración que mostraban una desviación estándar 3 veces superior a la desviación estándar media de todas ellas.

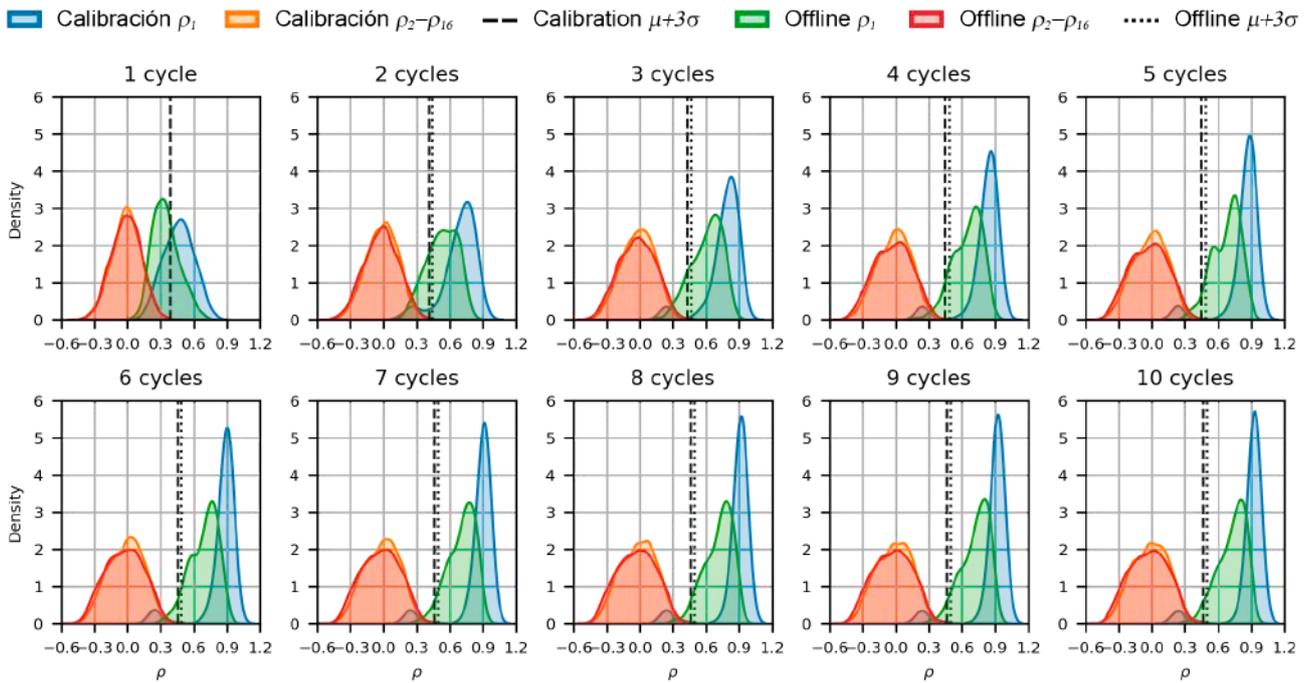
En la etapa de selección de comandos (i.e., test): (1) se extraen las épocas de cada ciclo, (2) se proyectan espacialmente con  $\boldsymbol{\omega}_b$ , y (3) se calcula la correlación  $\boldsymbol{\rho}$  de la proyección resultante con todas las plantillas. El comando seleccionado será el que pertenezca al índice  $\arg. \max_i(\boldsymbol{\rho})$ .

#### 3.2. Algoritmo de parada temprana propuesto

Un algoritmo de parada temprana debe tomar una decisión dicotómica cada vez que se muestra un ciclo en test: (1) efectuar la selección del comando más probable, o (2) repetir la estimulación un ciclo más. Cada vez que se efectúa la estimulación de un ciclo, la comparación entre la respuesta EEG desde el inicio de la estimulación visual hasta el final del ciclo actual y las plantillas de los comandos retorna un vector de correlaciones  $\boldsymbol{\rho} \in \mathbb{R}^{[1 \times 16]}$ . Tras ordenarlo de forma descendente,  $\rho_1$  indica la correlación del comando más probable. Si asumimos que, efectivamente, el usuario está prestando atención al comando con correlación  $\rho_1$ , entonces se puede considerar que  $\rho_2, \rho_3, \dots, \rho_{16}$  son correlaciones espúreas de comandos no deseados. Por tanto, podemos determinar que es fiable considerar que  $\rho_1$  corresponde al comando seleccionado si su valor es un *outlier* de la distribución a la que pertenecen los valores  $\rho_2, \rho_3, \dots, \rho_{16}$ . Existen diversas métricas para detectar *outliers* de una distribución, como aquellas basadas en el rango intercuartílico o en test de hipótesis. Sin embargo, nosotros proponemos el uso de *z-scores*. Asumiendo que la distribución de los valores espúreas es normal, i.e.,  $\rho_2, \rho_3, \dots, \rho_{16} \sim \mathcal{N}(\mu, \sigma)$ , podemos detectar si  $\rho_1$  es un *outlier* si  $\rho_1 - \mu > h\sigma$ , donde  $h = 3$ , lo que equivaldría a considerar que  $\rho_1$  supera el percentil 99.87% de la distribución. Si se cumple dicha inecuación, se efectuaría la selección del comando, mientras que en caso contrario se proseguiría con el siguiente ciclo.

#### 3.3. Protocolo de evaluación

Para evaluar la viabilidad y robustez del método, se realizará un análisis offline y un análisis online. En el



**Figura 2.** Distribuciones de las correlaciones del comando más probable  $\rho_1$  y espúreas  $\rho_2, \rho_3, \dots, \rho_{16}$  en el análisis offline para los datos de calibración (azul y naranja) y test offline (verde, rojo), así como el percentil 99.87% para cada una.

análisis offline se cuenta con los datos de los 6 sujetos de control sin aplicar parada temprana: 300 ciclos de calibración por sujeto ( $6 \text{ runs} \times 5 \text{ trials} \times 10$  ciclos); y 320 ciclos de test por sujeto ( $2 \text{ runs} \times 16 \text{ trials} \times 10$  ciclos), donde en cada *run* los usuarios seleccionaban todos los comandos en orden lexicográfico. En este análisis se evaluará si la distribución de las correlaciones espúreas sigue una distribución normal, así como el número de ciclos medio que hubiesen alcanzado si se hubiera aplicado el método propuesto. Por otro lado, en el análisis online se implementó el algoritmo de parada temprana, y se evaluó con un sujeto adicional como prueba de concepto: 300 ciclos de calibración ( $6 \text{ runs} \times 5 \text{ trials} \times 10$  ciclos), y 32 *trials* de selección *online* (donde el número de ciclos varía).

## 4. Resultados y discusión

### 4.1. Análisis offline

En la Figura 2 se muestran las distribuciones de las correlaciones objetivo y espúreas tanto para los datos de calibración como de test offline (sin parada temprana). Los resultados del test de Kolmogorov-Smirnov indican que la distribución de las correlaciones espúreas son normales ( $p$ -valor  $< 0.01$ ) para ambos *subsets* de datos. Por lo tanto, la estimación del percentil 99.87% cuando el  $z$ -score es mayor que  $h = 3$  veces la desviación estándar de los datos se puede considerar precisa. Asimismo, dado que la distribución espúrea en la calibración es significativamente similar a la del test offline ( $p$ -valor  $< 0.01$ ), podría optimizarse el valor de  $h$  por usuario sin necesidad de adquirir datos adicionales. Esta optimización podría realizarse mediante el análisis de una curva ROC (*receiver operating characteristic*) variando el valor de  $h$  y estudiando la separación de las distribuciones objetivo y espúreas de la calibración.

La Tabla 1 muestra las precisiones obtenidas por usuario durante ambos análisis. Además, se ha incluido el máximo teórico como el número de mínimo de ciclos posibles para alcanzar la precisión obtenida durante las pruebas. Como se puede observar, el uso del algoritmo de parada temprana ha permitido alcanzar una precisión media del 97.92% con una media de 2.37 ciclos (equivalente a 1.24 s/comando, ITR de 191.93 bpm) en el análisis offline. La diferencia con respecto a la precisión original de 98.96% (con 10 ciclos) no es significativa (*Wilcoxon signed-rank test*,  $p$ -valor = 1.04). Por tanto, podemos asegurar que el algoritmo propuesto permite reducir significativamente el número de ciclos necesarios para efectuar una selección sin perjudicar la precisión del sistema.

Cabe destacar que el máximo teórico indica que podría haberse reducido a una media de 1.39 ciclos, lo que implica una diferencia significativa con respecto a los 3.27 ciclos empleados (*Wilcoxon signed-rank test*,  $p$ -valor = 0.03). La razón por la cual se requiere aproximadamente un ciclo adicional con nuestro método radica precisamente en el primer ciclo de estimulación, pues teóricamente muchos *trials* podrían haberse decodificado tras la primera estimulación. Como se puede observar en la Figura 1, las distribuciones objetivo y espúreas no están claramente diferenciadas en el primer ciclo de estimulación, lo que provoca que el método propuesto no sea capaz de emitir una selección fiable de forma tan temprana. Aunque consideramos que los resultados obtenidos (97.92% con 1.24 s/comando) demuestran la utilidad del método, este hecho indica que aún hay rango de mejora.

### 4.2. Análisis online

En la Figura 1(b) se muestra una captura de pantalla de la app desarrollada para MEDUSA©. La app hace uso del motor gráfico Unity para poder mantener una frecuencia de

**Tabla 1.** Rendimiento obtenido en los análisis offline y online.

	Parada Temprana		Máximo Teórico*		
	Pre.(%)	#Ciclos	Pre.(%)	#Ciclos	
Offline	U1	96.88	2.03	96.88	1.66
	U2	100.00	3.03	100.00	1.34
	U3	100.00	2.25	100.00	1.50
	U4	100.00	2.28	100.00	1.09
	U5	93.75	3.06	96.88	1.59
	U6	96.88	1.56	100.00	1.16
	<i>Media</i>	<i>97.92</i>	<i>2.37</i>	<i>98.96</i>	<i>1.39</i>
<i>DE</i>	<i>2.33</i>	<i>0.53</i>	<i>1.47</i>	<i>0.21</i>	
Online	U7	96.88	3.19	n.a.	n.a.

\* El máximo teórico se ha calculado como el mínimo número de ciclos posibles para mantener la precisión de test (en la que se usaron 10 ciclos).

refresco a 120 Hz de forma constante, y se comunica con MEDUSA© mediante protocolo TCP/IP [9]. Tal y como se puede observar, la aplicación ofrece *feedback* relativo a la correlación acumulada del método de parada temprana a lo largo de los ciclos en forma de un marco verde que rodea los comandos. Cuanta mayor sea la correlación asociada a un comando, el marco mostrará un verde más intenso.

En la Tabla 1 se muestran los resultados del usuario adicional que evaluó el algoritmo *online* como prueba de concepto, alcanzando una precisión de 96.88% con una media de 3.19 ciclos (equivalente a 1.67 s/comando, ITR de 131.76 bpm). Los resultados demuestran que la aplicación del algoritmo propuesto en tiempo real es viable y beneficiosa para el sistema c-VEP-BCI, permitiendo disminuir drásticamente el tiempo necesario para efectuar una selección de 5.25s (10 ciclos) a 1.67 s (3.19 ciclos) sin perjuicio en la precisión obtenida. Asimismo, el rendimiento obtenido para este usuario es similar o supera los resultados de otros estudios relacionados, e.g., 3.26 s/comando [6], 4.2 s/comando [5], o 6.17 s/comando [8].

#### 4.3. Limitaciones y líneas futuras

A pesar de los excelentes resultados obtenidos con el algoritmo de parada temprana propuesto, se sugieren distintas líneas futuras de investigación para extender su utilidad y/o fiabilidad. En primer lugar, es necesario evaluar de forma *online* el método propuesto con una población mayor de sujetos de control y/o personas con grave discapacidad, con el fin de incrementar la potencia estadística de los resultados. Sería también interesante extender el algoritmo propuesto para desarrollar un algoritmo de asincronía (i.e., detectar la atención que presta el usuario a la estimulación). Actualmente, dado que las correlaciones se calculan de forma acumulativa, la aplicación del algoritmo para asincronía no es directa, pues los ciclos anteriores afectan sustancialmente a la decisión en tiempo real; mientras que lo deseable en un algoritmo de asincronía sería que la detección se produjese en *single-trial*. Otra futura línea de investigación podría ser la evaluación del algoritmo con un sistema c-VEP-BCI que emplee *m*-secuencias no binarias; o incluso conjuntos de códigos con baja correlación cruzada (e.g., códigos Gold, Kasami, etc.). Por último, consideramos que sería interesante analizar el rendimiento del algoritmo para desfases con saltos menores (e.g.,  $\tau = 2$ ), una estrategia útil para aumentar el número de comandos seleccionables.

## 5. Conclusiones

En este estudio se ha propuesto un nuevo algoritmo de parada temprana para sistemas c-VEP-BCI. Se ha demostrado que el algoritmo propuesto (1) no depende del clasificador, (2) no requiere optimizar ningún parámetro, (3) no requiere registrar señales EEG adicionales y (4) puede aplicarse en tiempo real. El análisis offline concluye que la aplicación del método permite disminuir el tiempo de selección de 5.25 s/comando a 1.24 s/comando sin afectar significativamente a la precisión (de 98.96% a 97.92%). Estos resultados se han reafirmado en la prueba de concepto *online*, alcanzando una precisión del 96.88% con 1.67 s/comando y una ITR de 131.76 bpm. Se concluye que el algoritmo propuesto es robusto y viable, permitiendo detectar en tiempo real el número de ciclos necesarios para efectuar una selección sin afectar negativamente al rendimiento del sistema BCI.

## Agradecimientos

Este trabajo ha sido financiado por los proyectos RTC2019-007350-1, PID2020-115468RB-I00 y TED2021-129915B-I00 del Ministerio de Ciencia, Innovación/Agencia Estatal de Investigación y European Regional Development Fund (ERDF), así como por CIBER-BBN a través del Instituto de Salud Carlos III.

## Referencias

- [1] J. Wolpaw and E. W. Wolpaw, *Brain-computer interfaces: principles and practice*. OUP USA, 2012.
- [2] V. Martínez-Cagigal, J. Thielen, E. Santamaría-Vázquez, S. Pérez-Velasco, P. Desain, and R. Hornero, "Brain-computer interfaces based on code-modulated visual evoked potentials (c-VEP): a literature review," *J. Neural Eng.*, vol. 18, no. 6, p. 061002, 2021, doi: 10.1088/1741-2552/ac38cf.
- [3] G. Bin, X. Gao, Y. Wang, B. Hong, and S. Gao, "VEP-based brain-computer interfaces: Time, frequency, and code modulations," *IEEE Comput. Intell. Mag.*, vol. 4, no. 4, pp. 22–26, 2009, doi: 10.1109/MCI.2009.934562.
- [4] S. Nagel and M. Spüler, "World's fastest brain-computer interface: Combining EEG2Code with deep learning," *PLoS One*, vol. 14, no. 9, pp. 1–15, 2019, doi: 10.1371/journal.pone.0221909.
- [5] J. Thielen, P. Marsman, J. Farquhar, and P. Desain, "From full calibration to zero training for a code-modulated visual evoked potentials brain computer interface," *J. Neural Eng.*, 2021, doi: 10.1088/1741-2552/abecef.
- [6] F. Gembler and I. Volosyak, "A novel dictionary-driven mental spelling application based on code-modulated visual evoked potentials," *Computers*, vol. 8, no. 2, 2019, doi: 10.3390/computers8020033.
- [7] F. Gembler *et al.*, "A Dictionary Driven Mental Typewriter Based on Code-Modulated Visual Evoked Potentials (cVEP)," in *Proceedings - 2018 IEEE International Conference on Systems, Man, and Cybernetics, SMC 2018*, 2018, pp. 619–624, doi: 10.1109/SMC.2018.00114.
- [8] F. Gembler, P. Stawicki, A. Rezeika, A. Saboor, M. Benda, and I. Volosyak, *Effects of monitor refresh rates on c-VEP BCIs*, vol. 10727 LNCS. Springer International Publishing, 2018.
- [9] E. Santamaría-Vázquez *et al.*, "MEDUSA©: A novel Python-based software ecosystem to accelerate brain-computer interface and cognitive neuroscience research," *Comput. Methods Programs Biomed.*, vol. Under rev, 2022.

# Competición FENIN-SEIB de Estudiantes de Grado en Ingeniería Biomédica

---

Miércoles, 23 de noviembre de 2022

# Tumor-on-a-chip: new strategies of in-vitro cancer cell culture

A.I. Cànaves Llabrés<sup>1</sup>, J. de Ciurana Gay<sup>2</sup>, T. Puig Miquel<sup>3</sup>

<sup>1</sup> University of Girona, Girona, España, toniignasi@hotmail.es

<sup>2</sup> Mechanical Engineering and Industrial Construction, University of Girona, Girona, España, quim.ciurana@udg.edu

<sup>3</sup> Medical Sciences, University of Girona, Girona, España, teresa.puig@udg.edu

## Abstract

*According to studies, generating organ-like cell cultures requires simulating the conditions and dynamics of the human body. Organs-on-a-Chip simulates the environment of the human body in order to develop human tissue models for disease modelling and pharmacological testing. Few manufacturers produce dynamic culture equipment, and it is expensive. The goal was to create a system with the same objectives and features as existing dynamic culture equipment, but at a lower cost and with production that could be easily and cost-effectively managed. This work creates microfluidic devices for cell culturing, influenced by organ-on-a-chip and tumor-on-a-chip. The device was required to meet biological requirements, such as simulating extracellular structures, as well as manufacturing and material constraints. Miniaturisation, material strength, size, modularity, thickness, and microscopic component complexity of the models were balanced to be manufactured using fused filament fabrication. It was crucial that the polymer was simple to print because it was meant to be manufactured by non-experts. This innovative method promises to reduce production costs, facilitate the adaptation of culture chips to existing applications, be wire-free, and be watertight. The use of stress testing enabled models to function despite miniaturisation. Chloroform was utilised to seal the components, guarantee that the system was impermeable, enhance layer adhesion (and strength), and smooth the texture of the components in order to prevent cell attachment. The majority of models are composed of PLA, PETG, and PCTG, whereas TPU is used for the filtering mechanism.*

## 1. Introduction

Developing organ-like cell cultures requires imitating the human body's environment and dynamics. Flow and pressure of liquids in contact with cells, as well as oxygen and carbon dioxide levels in the culture vessel and the human extracellular matrix must be reproduced in order to model the human interior environment. Dynamic culture methods were explored, finding "organ-on-a-chip," a micro-scale technology that simulates the human body to develop human tissue models for disease modelling and pharmaceutical testing. Few businesses produce dynamic culture equipment, which begins at \$20,000 per kit, despite its many advantages. After the research, the doubt arose of how universities or research institutions with fewer resources could implement these new concepts and technologies. The idea was to create a system with the same objectives and features as existing dynamic culture equipment, but at a cheaper cost and whose output could be easily and cost-effectively managed by the same organisations. Here is a low-cost system for creating 3D

cell cultures. AutoCAD is used for design, whereas FFF is used for manufacturing.

The contribution focuses on the design of a microfluidic device for tumor cell culture, especially paying attention to the graphic design of the many components and parts, as well as the choice of materials. Considered to present preliminary results on the development of a first prototype, but missing results demonstrating the proposal's feasibility.

## 2. Hypothesis and Objectives

### 2.1. Research question

Is it possible to fabricate a microfluidic system for the in vitro cell culture of tumor-like structures inspired by the technological concept of organ-on-a-chip and specifically tumor-on-a-chip using fused filament fabrication as the primary manufacturing method and inexpensive electronic components?

### 2.2. Hypothesis

Manufacture of an in-vitro cell culture device incorporating a microfluidics system using cost-effective components and fused filament fabrication as the primary technology in the manufacturing process is possible by adapting the conceptualization of the models to the limitations of machinery and material and applying the appropriate post-processing.

### 2.3. Objectives

Open interest in 3D design, the use of biocompatible polymers, and fused filament fabrication in the conceptualization and manufacture of devices for cell culture using microfluidics, providing a system that overcomes the major restrictions previously connected with this technique of production in terms of the manufacture of organs-on-a-chip and tumors-on-a-chip.

## 3. Materials and Methods

### 3.1. Design

The design of the gadget had to adhere to biomedical criteria, such as mimicking extracellular structures, as well as manufacturing and material limitations. The use of polymers in FFF printers necessitated striking a compromise between miniaturisation, material strength,

and tiny component complexity. Variables included the size, modularity, and thickness of translucent filament-printed surfaces. Multiple elements comprised the microfluidics system. For culture chips, the state-of-the-art models were reinterpreted for FFF with the highest miniaturisation factor. The culture medium tank was designed to accommodate the pump while preventing medium or blood loss. The remaining components fit these two.

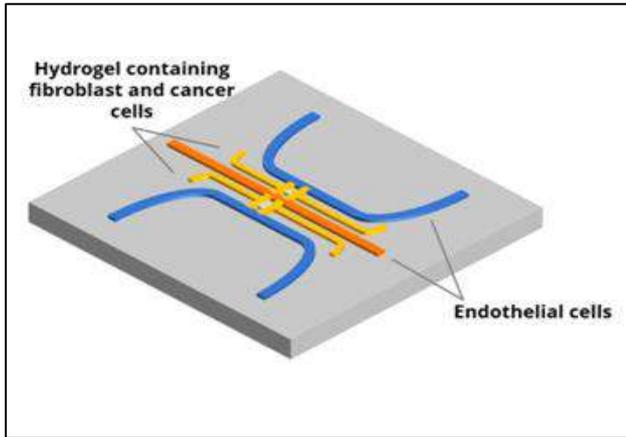


Figure 1. Example of state-of-the-art model for the general-purpose chip (breast tumor-on-a-chip) [1].

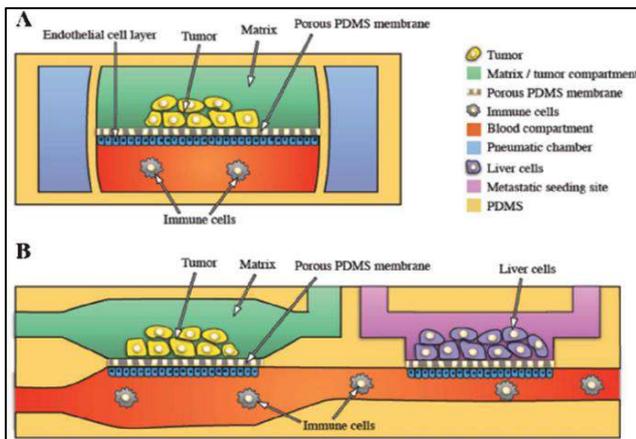


Figure 2. State-of-the-art model for metastasis tumor-on-a-chip [2].

### 3.1.1 Culture Chips

On the one side, we have a chip built to endure photometry, microscope, and other straightforward testing (general purpose tumor-on-a-chip). A metastatic tumor-on-a-chip chip, on the other hand, attempts to imitate its microstructure and cell distribution. There were three main challenges in the design of the chips: connecting the pump, printing the internal channels with sufficient detail, and making the bottom layer of the culture area transparent so the crop could be analysed. Adapting the size of the chip to current limitations resulted in a culture device that was smaller than a conventional culture plate but larger than organs-on-a-chip. Needles would not establish a secure connection with the polymer device (there would be leaks), and because to the size of the chip, perfusion would be

inefficient. Therefore, the connectors were printed on the same piece, allowing a medium-sized transport tube to be joined without the need for additional components. The external and interior dimensions of these connectors were modified to fit within a compact tube while maintaining their strength.

Using the "Difference" method in AutoCAD, chip covers were created to seal the crop, maintain device pressure, and avoid leaks.

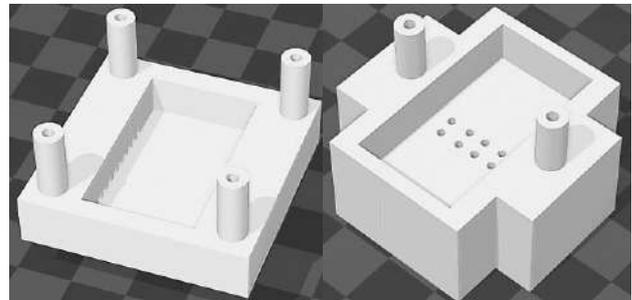


Figure 3. (right) Metastasis chip model overview. The two cylinders are the inlet and outlet for the medium, while the central portion is the culture area. The medium enters the chamber via the eight apertures at the chamber's base that are connected to the lower channel. (left) General purpose chip model overview. The four cylinders are medium inputs and outputs, while the region in the middle is the cultural area. The medium enters the camera through each side's nine apertures.

### 3.1.2 Tank

Originally, there was only one pump chamber in the culture medium tank. During the design process, obstacles arose, such as adjusting the tank size to accommodate alternative configurations or experiments, sealing the tank to manage oxygen and carbon dioxide concentrations, and establishing a filtration system prior to the microfluidic circuit cycle's conclusion. The final design was divided into two variants with the same fundamental design, two chambers, but distinct volumes, allowing this peripheral to be easily modified to accommodate the increasing number of chips.

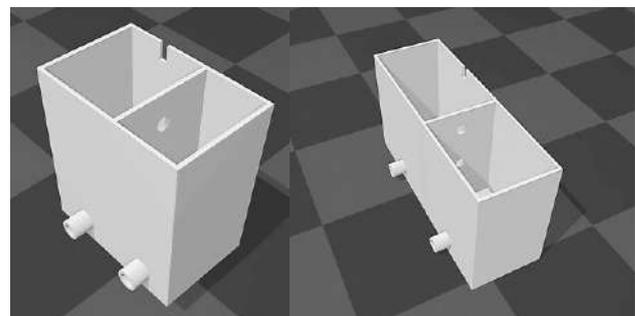
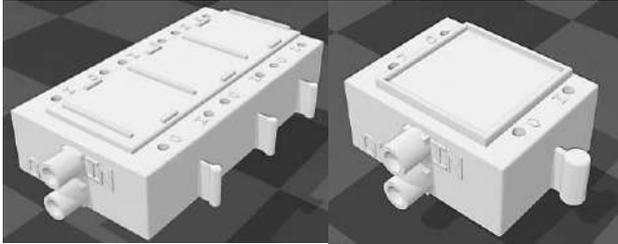


Figure 4. (left) Tank. Individual culture version. (right) Tank Extended culture version. The room above has the pump, while the room below contains the filters. One of the filtrate-connecting holes in the partition wall is visible

### 3.1.3 Channels Module

The purpose of this device was to stabilise the chip and its connections while retaining injection tubing via the

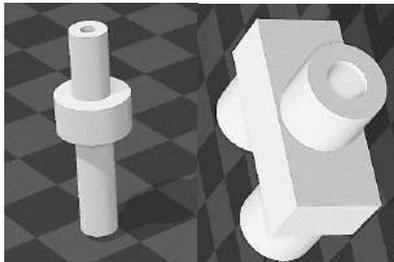
device's perforations. Later, it was realised that the two ideas might be combined, allowing for the uniform distribution and collecting of the pumped material. Internally, two 8 mm channels branch into 5 mm channels. It includes techniques for developing a single chip or up to three chips to simulate a body-on-a-chip, complex combinations, or a bioreactor. Without multiple tanks, the modular device's connections can be joined with a later-described adaptor. Each configuration includes two chip-based models.



**Figure 5.** Channels module overview. Configuration for the general-purpose chip. Multiple chip application on the left, individual chip application on the right. In the side openings, adaptors to connect with the chips are inserted. Within the implemented boxes, the chips are arranged.

#### 3.1.3.1 Adaptors

Adaptors connect infusion tubes and expand the capabilities of our design. Two types: the adapter that connects the 5 mm channels of the channel module to the tubes that lead to the chips (its use is required for optimal device operation) and the adapter that can be inserted between the tank and the channel module so that more than one of these can be used (its use is optional).



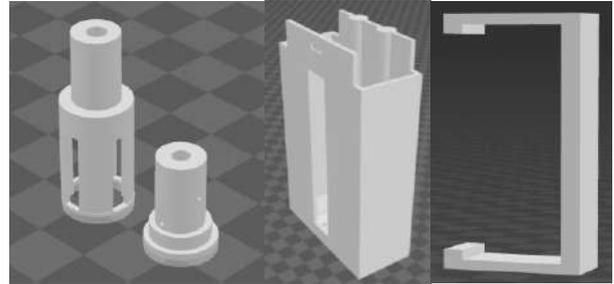
**Figure 6.** (left) Adapter between the 5 mm channels of the channels module and the chips. The 9 cm section is connected to the channel module's apertures, while the 1.5 cm section is connected to the medium perfusion tubes leading to the chip. (right) Adapter between the tank and the channels module. The tank tube is connected to the single input, while the inputs of the utilised channel modules are connected to the two outputs.

#### 3.1.4 Filters

A two-piece filter manages the connection between the tank's two portions, retaining any tumor spheroids for further study. It houses filtering fibre scaffolds. It fits into the slots in the secondary chamber's partition wall for optimal installation.

#### 3.1.5 Battery

Because the device will be in an incubator, a battery enclosure was made. This shell allows airflow between the culture basin and battery and encloses the battery to prevent humidity. A hole has been made at the back so that moisture can be easily detected.



**Figure 7.** (left) Filter for tumoroids / spheroids. One fibrous scaffold is positioned within the opening of the little component, while another is positioned surrounding it. Finally, it fits the larger section. Battery case (centre) and lock.(right).

### 3.2. Materials

Because they will interact with organic compounds, biocompatible materials must be used. According to the current state of the art, PDMS, glass, and thermoplastic polymers are the most viable options for creating organ-on-a-chip systems. Due to manufacture, only the third material fits the criterion. PLA is biocompatible, biodegradable, and the most user-friendly material for FFF printing. TPU is a flexible, biocompatible, and easy-to-print material; thus, it was utilised in several sections. PMMA is transparent, biocompatible, and simple to manufacture. Silicone is translucent, biocompatible, and flexible. Chloroform is a solvent used in biological labs.

#### 3.2.1 PLA / PETG / PCTG

PLA, PETG, and PCTG are typically used to print models. Several variables to consider:

- A. To print transparent chips, transparent filament is required. PETG has proven to be more durable and ideal for making this component, as we need to apply pressure to the connecting pipes when installing the medium distribution tubes; however, PLA has been used to produce a functional assembly.
- B. When printing the channel module, consider the 5 mm channel adaptor. If we need to remove the adaptors because they are broken or to seal the entry for a specific application, we will need to utilise heat to do so without causing damage to the channel module. PETG, which has a higher melting point than PLA, is therefore preferred for the channels module. The channel module will not melt if the adaptors are printed using this other polymer.

### 3.2.2 TPU

TPU was chosen for filter printing because PLA/PETG presented mounting and installation difficulties. This polymer absorbs more water than others, hence printed objects made from it decay faster than those made from PLA/PETG.

### 3.2.3 PMMA

Made of PMMA, chip covers are transparent. In the absence of laser cutters, testing polymer plates were cut with a cutter. The resulting components did not fit in the crop area hole, and the configuration was not as resistant to stress as the printed lid. PMMA will be utilised to produce optimal cell cultures.

### 3.2.4 Sealing

General chloroform treatment enhances layer adhesion, resulting in models that are more resilient and less fragile. Additionally, the texture of the infusion channel walls can be modified to prevent undesired cell adhesion.

### 3.2.5 Silicone

Silicone tubes are used to transport cellular media between the tank, channels module, and chips. Aquaponics circuit and respirator replacement part vendors provided the food- and medical-grade silicone.

## 4. Conclusions

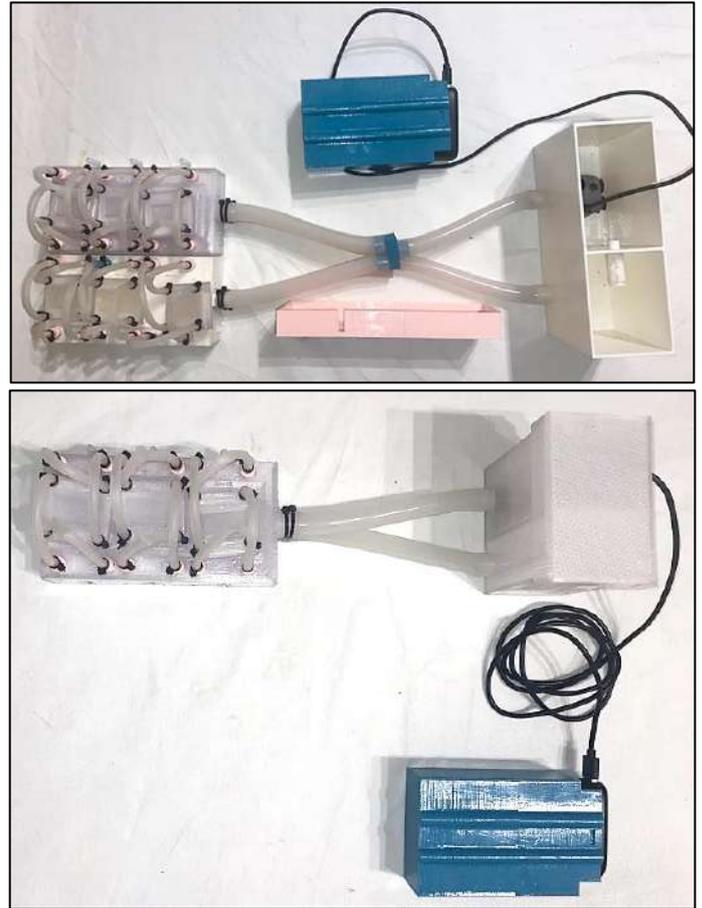
This project presents a framework for developing microfluidic devices for cell culture, inspired by organ-on-a-chip and tumor-on-a-chip. This novel technique promises to reduce production costs, make it easier to adapt culture chips to existing applications, be watertight, have an adjustable flow, and be wire-free. It will also serve as a platform for future bio-inspired designs. To make models work despite miniaturisation, component and section sizes were changed using stress testing. Chloroform was employed to seal the parts and guarantee the system was impermeable, improve layer adhesion (and hence strength), and smooth the pieces' texture to prevent cell adherence. PLA, PETG, and PCTG are utilised for most models, TPU for the filtering system. Before cell culture testing, PMMA caps for the chips must be created. Laser-cut and chloroform-joined. Cell culture and viability testing will determine the device's efficacy.

### 4.1.1 Improvement proposals

The current equipment needs three upgrades. First, design a control circuit with an Arduino board, pH, oxygen, carbon dioxide, temperature, pressure, humidity, and thermistor sensors. We could monitor the culture without a microscope or lab visit. The thermistor allows temperature control without an incubator. Another pump could provide more complex  $O_2$  gradients. Second, incorporate sensor data into a machine learning-based prediction model to automate cell culture management and forecast viability. Third, utilise PMMA as the chip bottom to increase

transparency and simplify microscopy. As a last expansion, a chloroform treatment chamber might be developed.

## 5. Results



**Figure 8.** (2) Assembly of the circuit in bioreactor mode; (1) assembly of the circuit in bioreactor mode utilising two channels modules and the extended tank.

## Acknowledgments

Special thanks to the Organizing Committees of the fortieth CASEIB for allowing this project to be presented and for its efforts to promote and encourage innovation in biomedical engineering.

## References

- [1] “Tumor-on-a-chip to study breast cancer with microfluidics: MTOAC.” <https://www.elveflow.com/microfluidics-research-horizon-europe/european-projects/organ-chip-multi-compartmental-tumor-chip-study-breast-cancer-mtoac-project/> (accessed May 14, 2022).
- [2] A. van de Stolpe and J. den Toonder, “Workshop meeting report Organs-on-Chips: Human disease models,” *Lab on a Chip*, vol. 13, no. 18, pp. 3449–3470, Sep. 2013, doi: 10.1039/C3LC50248A.

# Sistema de asistencia y monitorización para personas con Deterioro Cognitivo Leve

E.M. Solar Risueño<sup>1</sup>, D. Marcos Jorquera<sup>2</sup>

<sup>1</sup> Universidad de Alicante, Alicante, España esr38@alu.ua.es

<sup>2</sup> Depto. Tecnología Informática y Computación, Universidad de Alicante, Alicante, España, dmarcos@dtic.ua.es

## Resumen

*El envejecimiento de la población, así como la cada vez mayor tendencia de las personas mayores de 64 años a vivir en hogares unipersonales, hacen que se haga evidente la necesidad de disponer de herramientas que les permitan disfrutar de una mayor autonomía y una consecuente mejora en su calidad de vida. Conociendo que, una de las patologías comunes en este rango etario es el Deterioro Cognitivo Leve (DCL), se plantea, diseña y desarrolla una solución tecnológica en base a las necesidades más comunes detectadas para contribuir en su día a día desde el punto de vista de la Ingeniería Biomédica. Como propuesta de solución se propone un sistema basado en telemedicina que está compuesto por dos aplicaciones: una para los cuidadores que permite monitorizar en todo momento la actividad del paciente y gestionar sus actividades básicas como medicación, horas de despertar o citas; y otra que llevará el paciente, que se encargará, de forma totalmente destendida y usable, mediante una interfaz auditiva, de instruir sobre las acciones que tiene pendientes de realizar, así recabar datos de su actividad diaria mediante GPS o acelerómetro. Para validar la propuesta se ha realizado un prototipo y se han realizado pruebas que determinan la estabilidad y viabilidad del aplicativo.*

## 1. Introducción

De acuerdo con las estadísticas del Instituto Nacional de Estadística (INE) [1], existe una tendencia creciente en el aumento de personas mayores que viven solas en España. Por comunidades autónomas, Cataluña es la región con más mayores viviendo en soledad, seguida de Andalucía y Madrid. Las cifras ascienden a más de dos millones (2,1), en el caso de los mayores de 65 años y en un gran porcentaje se trata de viudos y viudas, con un claro predominio de éstas (1,5 millones) frente a 620.400 hombres en esta situación. El problema de la soledad en los mayores es la incidencia en problemas de salud: caídas no atendidas, depresión, fallos en la toma de medicamentos, etc.

La pandemia arroja datos que sugieren una incidencia significativa en el aumento de esta tendencia. En 2020 aumentaron en un 47% el número de hogares unipersonales de personas mayores en Ceuta, un 21% en Canarias, un 17% en Asturias y un 13% en Aragón. Al mismo tiempo se registraron ligeros descensos en otras comunidades como Navarra (9%), Baleares (8%) y Madrid (2%). En los casos de Cataluña y Madrid ha habido un aumento interanual del 9 y el 8% respectivamente [2].

Según algunos análisis demográficos, el proceso de envejecimiento español va a acelerarse en los próximos 30

años, cambiando significativamente la distribución de edad de la población, así como la relación entre los requerimientos de gasto frente a las capacidades productivas. Actualmente hay aproximadamente 3,3 habitantes de 15-64 años por cada uno mayor de 65, y la tendencia es que esta ratio disminuya a 1,7 para 2050. Este incremento de personas mayores implica nuevos retos para una sociedad envejecida en la que tendremos más personas dependientes, tanto social como económicamente, por cada persona joven. Dado que esta tendencia en el envejecimiento no se va a modificar a medio plazo, sus consecuencias estarán asociadas a diversos factores como: políticas públicas, avances tecnológicos, la educación y la sanidad, que ayuden a mitigar en gran parte sus consecuencias negativas. El envejecimiento condiciona sustancialmente la estructura de gasto público, ya que el creciente porcentaje de personas mayores provoca un incremento en la demanda recursos y servicios públicos.

Este incremento de personas de más de 65 años con respecto al resto de rangos etarios pone de manifiesto una relación proporcional con el aumento de personas diagnosticadas de Deterioro Cognitivo Leve (DCL) [3]. Se entiende por DCL a la alteración de las capacidades cognitivas, que, sin llegar a ser demencia, presenta síntomas tales como pérdida de memoria, afeción del lenguaje, o incluso declive de la capacidad visual o espacial. No obstante, las personas con esta condición, se considera que son independientes en la mayor parte de las acciones de su día a día. El presente proyecto plantea una herramienta de soporte diseñada para mejorar la calidad de vida de los pacientes y familiares, generando un sistema de actuación, que permite realizar un seguimiento semiautomatizado del día a día de los pacientes, con una implicación mínima de éstos en el funcionamiento y mantenimiento del servicio.

Se pretende poner a disposición de las personas a cargo de los pacientes, funcionalidades personalizadas que puedan emplear según necesidades, añadiendo alarmas y avisos programables, que ayudarán a tener un control telemático sobre las principales incidencias propias de la patología. Para ello, se desarrolla una aplicación distribuida basada en interfaces naturales que se comunican de forma verbal con las personas asistidas, lo cual permite salvar las barreras tecnológicas que encontramos en estos pacientes, consiguiendo una supervisión y asistencia en tiempo real.

## 2. Estado del arte

Una vez conocido el ámbito de estudio, se realiza un análisis de las aplicaciones existentes que ya se ocupan de dar solución a las carencias anteriormente descritas, estén o no diseñadas para su uso en personas mayores.

Se toman como referencia aquellas que están destinadas al seguimiento de la salud, permitiendo mayor control y autonomía sobre ella.

Las funcionalidades que ofrece Medisafe [4], se centran en la administración de medicamentos. Permite agregar medicaciones, recibir recordatorios y alarmas para las distintas tomas y darle seguimiento al estado de salud, así como obtener un informe de su progreso, que pueden posteriormente descargar y entregar a su médico o farmacéutico.

Life360 [5] se presenta como una App para localización a tiempo real de familiares. La tecnología GPS permite conocer situaciones simultáneas de varios familiares, así como avisos cuando salen o entran en una zona previamente delimitada.

Enfocado en el cuidado de personas mayores, Senda [6] ofrece un dispositivo geolocalizable con un botón de ayuda SOS que conecta mediante llamada con hasta tres números registrados. Destaca la comunicación bidireccional y autónoma, puesto que acepta la llamada sin necesidad de intervención del usuario y activa las manos libres para facilitar la comunicación. Consta de dos partes, el dispositivo ya mencionado y una aplicación móvil diseñada para que los familiares o cuidadores puedan hacer seguimiento de la localización, las caídas, la batería del dispositivo del usuario y alertas programables.

Durcal [7] está centrada en la teleasistencia de las personas mayores. Es la aplicación que más se acerca al modelo que se quiere conseguir con el presente proyecto. Dispone de un reloj GPS para la persona a cuidar y una app móvil para cada integrante de la familia o cuidador que se registre.

Tras estudiar estas aplicaciones, se puede determinar que, a pesar de que gran parte de ellas muestran soluciones al ámbito de la monitorización y asistencia, no existe ninguna que integre todas las funcionalidades en un único sistema.

## 3. Propuesta

El objetivo principal de este trabajo es el diseño y creación de un sistema de monitorización y teleasistencia continua a personas con DCL, dando así solución a uno de los grandes retos que plantea la estructura social actual (ver Figura 1).

El sistema está compuesto por dos elementos fundamentales:

- Una aplicación para los cuidadores que permite monitorizar en todo momento la actividad del paciente y gestionar actividades básicas como medicación, horas de despertar o citas.
- Una aplicación que llevará el paciente, que se encargará, de forma totalmente desatendida y 100% usable utilizando una interfaz auditiva, de

instruir sobre las acciones que tiene pendientes de realizar, o recabar los datos de su actividad diaria mediante GPS o acelerómetro.

Realizado el estudio sobre el campo de actuación, se establecen los siguientes objetivos específicos:

- Realizar un seguimiento de la localización a tiempo real del paciente.
- Gestionar los medicamentos asociados al paciente, de forma que se puedan realizar recordatorios referentes a su toma, así como a la compra de nuevas unidades
- Disponer de registro y recordatorios referentes a citas del paciente, que permita asistirle en los desplazamientos al lugar y hora indicado.
- Avisar al cuidador o familiar si se detecta una posible caída.
- Romper la barrera tecnológica presentando una interfaz sin interacción para el paciente.
- Asistir al paciente para despertarse en horas adecuadas.

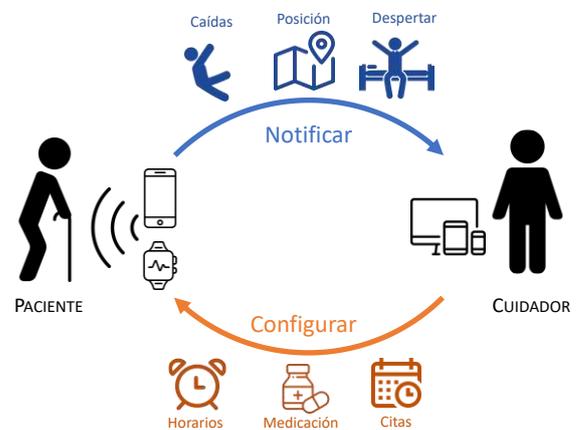


Figura 1. Modelo conceptual

Así mismo, es fundamental que la propuesta cumpla con una serie de características que la hagan viable en un entorno realista, destacando las siguientes:

- Que sea un sistema de bajo coste, utilizando dispositivos cotidianos para los ciudadanos de hoy en día como pueden ser *smartphones* y *wearables*.
- Que la propuesta sea flexible, y pueda ser adaptada y configurada a distintos entornos y circunstancias.
- Que la solución sea ubicua, permitiendo su uso, tanto por parte de los pacientes como de sus cuidadores, desde cualquier sitio, siempre que se cuente con una conexión a Internet.
- Que la solución sea totalmente usable y accesible, algo fundamental para poder ser implantado en un entorno donde los principales usuarios no tienen un alto nivel de conocimientos tecnológicos.

### 3.1. Arquitectura

Una vez establecidos los objetivos del proyecto e identificados los requerimientos que se van a abordar, se

ha decidido utilizar una Arquitectura Orientada a Servicios (SOA) mediante servicios de tipo REST (Figura 2).

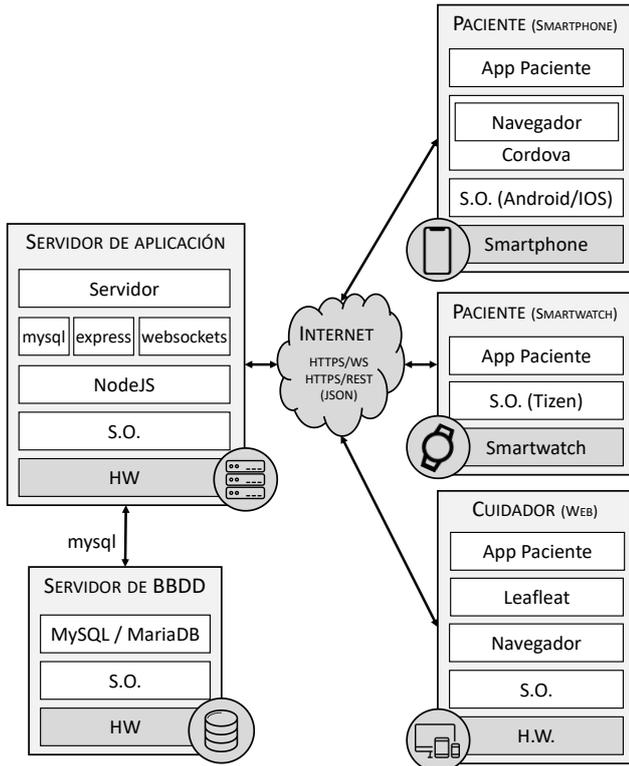


Figura 2. Arquitectura del sistema

Esta arquitectura permite desacoplar perfectamente los distintos componentes del sistema, utilizando un mecanismo de comunicación altamente interoperable basado en HTTP y en la codificación JSON. Esto permite integrar elementos heterogéneos en el sistema (*smartphones, tablets, PCs, wearables*) mediante una conexión a Internet.

En general, el API definido es una translación directa con el acceso a base de datos que se realiza en el servidor mediante una filosofía de CRUD (*Create, Read, Update y Delete*), donde las operaciones de lectura (HTTP GET) se implementan con sentencias SELECTS de SQL, las operaciones de creación (HTTP POST) con sentencias INSERT, las operaciones de actualización (HTTP PUT) con sentencias UPDATE y las operaciones de eliminación (HTTP DELETE) con sentencias DELETE.

El resto de las funcionalidades, que no tiene que ver tanto con la parte persistente del sistema, sino con la parte más reactiva (caídas, alarmas, etc.) si bien estarán integradas con los servicios que van a ser implementados, serán gestionadas mediante *WebSockets* (WS), lo cual permitirá a las aplicaciones un alto nivel de reactividad en tiempo real sin hacer constantes consultas (*pulling*).

En todo el desarrollo del software se ha optado por usar tecnologías web que solventan los problemas de heterogeneidad que se pudieran dar.

### 3.2. Aplicación de los pacientes

Para las personas asistidas se ha propuesto la implementación de dos tipos de clientes, uno basado en

*smartphones* y otro basado en dispositivos *wearables*, aportando flexibilidad a la hora de que un usuario escoja el dispositivo que más se ajuste a sus necesidades.

Para el primero, se ha utilizado la plataforma *Cordova*, que permite la creación de aplicaciones híbridas multiplataforma, utilizando un desarrollo web que posteriormente es empaquetado en una aplicación de Android o IOS.

Para el segundo, se ha utilizado un reloj modelo *Samsung Gear S3 frontier* (Figura 3), un smartwatch basado en el sistema operativo *Tizen*, cuya programación también se realiza mediante estándares web. Este modelo de reloj ha sido elegido ya que cuenta con conexión a internet independiente del teléfono y altavoz, que permite reproducir las instrucciones vocales de la aplicación.



Figura 3. Aplicación del paciente en un reloj

Con estas decisiones de diseño se ha conseguido realizar un desarrollo totalmente multiplataforma, creando una única aplicación web que posteriormente ha sido compilada a las distintas plataformas

### 3.3. Aplicaciones de los cuidadores

Para el uso por parte de los familiares de la persona asistida, se ha realizado una aplicación web que permita realizar la gestión y seguimiento de las acciones diarias de los pacientes. Se trata de una aplicación *responsive*, que permite su uso desde un navegador desde cualquier tipo de dispositivo (Figura 4).



Figura 4. Capturas de la aplicación del cuidador

### 3.4. Servidor

Para la parte del *backend*, se ha seleccionado Node.js, un motor de ejecución de *Javascript* para servidores, ya que, al tratarse de un sistema multiplataforma, permite la fácil migración del sistema a otras plataformas.

En cuanto al almacenamiento se ha optado por una base de datos *MariaDB*, una variante *OpenSource* de la base de datos relacional MySQL altamente extendida.

Para el acceso a base de datos se ha utilizado el módulo *mysql*, para la creación de los servicios REST el módulo *express*, y para WS el módulo *websockets*.

### 3.5. Procesos de monitorización

Para realizar, por parte de la aplicación del paciente, los procesos de monitorización y seguimiento se han utilizado un conjunto de estándares web que han permitido implementar los objetivos propuestos.

Para el caso del seguimiento del paciente, función necesaria tanto para la monitorización de la ubicación por parte del cuidador, como para instruir al paciente a la hora de dirigirse a la ubicación de una cita, se ha utilizado el API *Geolocation*, parte del estándar de HTML5, que permite recibir la ubicación actual del dispositivo en tiempo real. Este API utiliza varias tecnologías, principalmente *Global Position System* (GPS) para obtener la localización del dispositivo.

Junto con este API, y mediante la librería de *Javascript Leaflet*, se ha incorporado en la aplicación del cuidador un módulo de seguimiento visual basado en mapas. La comunicación en este caso se realiza mediante tecnología *WebSockets*, que nos permite una actualización en tiempo *real* de la posición del paciente según este se va desplazando.

Para las funcionalidades de detección de caídas y de identificación de que el paciente se ha despertado, se han utilizado dos APIs de HTML5, *Accelerometer*, que nos permite acceder a la aceleración en los tres ejes del dispositivo, y *Screen Orientation*, que permite acceder a los parámetros que determinan la orientación del dispositivo. Mediante ellas se puede identificar cuando el paciente está despierto, siempre que haya movimientos continuos, o cuando se produce una posible caída, si la aceleración en la vertical supera un determinado umbral. Esta información es transmitida mediante *WebSockets* para avisar al cuidador en tiempo real.

### 3.6. Interacción verbal

Uno de los principales objetivos de la propuesta es obtener un sistema altamente usable y accesible. Dado que las personas de edad avanzada suelen tener problemas para interactuar con dispositivos electrónicos, se ha decidido que toda la información que se le da al usuario sea mediante voz.

La generación automática de voz artificial a partir de texto es una técnica que lleva siendo utilizada durante décadas en multitud de aplicaciones, si bien en los últimos años, gracias a los asistentes de voz y los altavoces inteligentes, ha sufrido una significativa mejora en su calidad.

Para la implementación se ha utilizado el nuevo API estándar de síntesis de voz (*SpeechSynthesis*), que permite la conversión directa de texto a voz en multitud de idiomas.

## 4. Pruebas y validación

Una vez desarrollado el prototipo, se han diseñado y realizado un conjunto de pruebas tanto funcionales como no funcionales, para validar los objetivos del trabajo.

Se han realizado pruebas funcionales con diversos usuarios para validar todas las secciones de las aplicaciones (Alarmas, caídas, medicación, localización...).

También se ha probado la viabilidad del sistema con diversos dispositivos, sistemas operativos y navegadores.

Tras las pruebas se ha podido constatar la viabilidad del proyecto, pudiendo constatar que ha sido funcional y altamente usable en la mayor parte de los casos.

## 5. Conclusiones

En este trabajo se aborda un problema actual, la asistencia de personas con DCL, una patología con cada vez mayor prevalencia en la sociedad y cuyos índices se espera que sigan aumentando en los próximos años.

Se ha propuesto una solución basada en estándares, que ha demostrado ser viable, de bajo costo, ya que sólo necesita dispositivos de uso cotidiano en la actualidad, flexible, pudiendo ser adaptado a diferentes entornos, y altamente usable, con una interfaz basada en lenguaje natural, lo cual es fundamental en este tipo de propuestas.

Como trabajo futuro, se pretende incorporar reconocimiento de voz asociada a distintos comandos en la aplicación del paciente, para que éste pueda ponerse en contacto a sus familiares, o simplemente notificar de la toma de un medicamento de forma intuitiva y mediante lenguaje natural.

## Referencias

- [1] Instituto Nacional de Estadística. Una población envejecida [https://www.ine.es/prodyser/demografia\\_UE/bloc-1c.html?lang=es](https://www.ine.es/prodyser/demografia_UE/bloc-1c.html?lang=es) (Consultada: Marzo 2022)
- [2] Vértice salud, fundación juan cruzado entidad de 2000. <https://verticesalud.org> (Consultada: Mayo 2022)
- [3] Escarbajal De Haro, A., Martínez de Miguel López, S., y Romero Sánchez, J. (2015). El deterioro cognitivo leve en personas mayores usuarias de centros municipales de la región de Murcia (España). *Anales de Psicología / Annals of Psychology*, 32(1), 234–240. <https://doi.org/10.6018/analesps.32.1.196061>
- [4] Medisafe. <https://www.medisafe.com/>. (Consultada: Enero 2022)
- [5] Life360. <https://www.life360.com/intl/>. (Consultada: Enero 2022)
- [6] Senda. <https://byhs.eu/sendat/>. (Consultada: Enero 2022)
- [7] Durcal. <https://www.durcal.com/es/>. (Consultada: Enero 2022)

# EEG-Based Functional Connectivity during Progression from Infantile Spasms to Lennox Gastaut Syndrome

B. Romero Mila<sup>1,2</sup>, V. Liu<sup>3</sup>, D. Shrey<sup>3,4</sup>, B. Lopour<sup>1\*</sup>

<sup>1</sup> University of California, Irvine, Department of Biomedical Engineering, Irvine, CA USA, {beth.lopour@uci.edu}

<sup>2</sup> Universitat de Barcelona, Department of Electronics and Biomedical Engineering, Barcelona, Spain

<sup>3</sup> Children's Hospital Orange County, Division of Neurology, Orange, CA, USA

<sup>4</sup> University of California, Irvine, Department of Pediatrics, Irvine, CA, USA

## Abstract

*Infantile Spasms (IS) is a severe infantile epilepsy that can progress into Lennox-Gastaut Syndrome (LGS), which is associated with intellectual disability and lifelong epilepsy. Early diagnosis and treatment of LGS may improve prognosis [1], but there are currently no biomarkers to monitor for or predict the onset of this disease. Based on prior work [2], [3] we hypothesize that EEG functional connectivity strength is a robust biomarker of the severity of these epilepsies. Five children diagnosed with IS who progressed to LGS were included in this study. Each child had EEG studies performed at multiple timepoints, including when IS and LGS were diagnosed and following treatment. Functional connectivity networks were obtained by performing the statistical analysis of cross-correlation between EEG electrode pairs for each EEG study [2]. We found that the number of strong connections and the mean strength of the strongest 10% of all connections were correlated to disease state progression and response to treatment, but not the child's age at the time of the EEG recordings. The connectivity strength was high at timepoints when IS and LGS were diagnosed. Favorable treatment outcome was associated with a decrease in network strength, while an increase in strength reflected worsening of the disease or a lack of response to treatment. Connectivity strength was not correlated to age, suggesting that these differences are not due to age-related physiological changes. A second metric, the number of strong connections within the network, gave equivalent results. Overall, EEG functional connectivity analysis has potential as a clinical biomarker to improve the prognosis of LGS patients. However, it is critical to evaluate this biomarker in a larger cohort of subjects to increase statistical power and validate results.*

## 1. Introduction

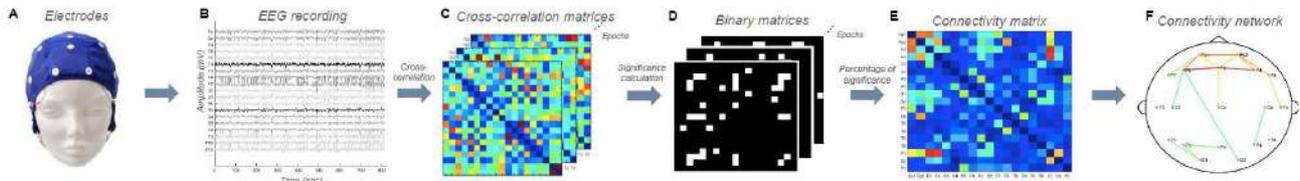
Infantile Spasms (IS) refers to a type of epileptic encephalopathy that occurs in children younger than 1 year old, with a prevalence of 2-5/10.000 live births. This syndrome is associated with epileptic spasms occurring in clusters, often with a characteristic interictal electroencephalographic (EEG) finding called hypsarrhythmia. Epileptic spasms are seizures characterized by brief, isolated muscle contractions, which involve the neck, the trunk, and the extremities in a symmetric bilateral manner. Individual spasms usually last between 1 and 5 seconds and occur in clusters of 3-20 spasms, often with several clusters occurring per day. The spasms are frequently associated with neurodevelopmental arrest or regression. Long term, patients with infantile spasms are at high risk for poor neurocognitive outcomes,

including chronic intractable epilepsy, cognitive impairment, and other neurodevelopmental disabilities [4].

Another potential sequela of IS is that 30-60% of children with IS will later develop Lennox-Gastaut Syndrome (LGS) [5], [6]. LGS is a severe epilepsy syndrome characterized by the triad of (1) encephalopathy, (2) multiple seizure types, and (3) a specific interictal EEG pattern with bursts of slow spike-wave (SSW) complexes and/or generalized paroxysmal fast activity (GPFA) [7]. This epileptic encephalopathy is usually diagnosed around 8 years of life and has a prevalence of 0.26/1000 children by 10 years of age [8]. For LGS patients, long-term neurocognitive outcomes are frequently poor, and complete seizure freedom is rarely achieved, often impacting social function, likelihood of independence, and quality of life. Timely diagnosis and appropriate treatment results in improved outcomes and less costly management of this disease [1]. Therefore, there is a need for biomarkers to monitor children for the progression from IS to LGS and rapidly assess treatment response.

Previous studies have analyzed the transition from IS to LGS in individuals, mainly focusing on treatment effects, risk factors [5], and the changes in the EEG [9]. Computational analysis has been applied to noninvasive scalp EEG recordings in patients with IS. These studies have demonstrated that computational EEG metrics are promising objective biomarkers for diagnosis, assessment of treatment response, and prediction of treatment response or relapse [10]. This includes the analysis of functional connectivity networks (FCNs), which are derived using a brain mapping technique based on the statistical interdependencies of spatially distinct time-varying neural signals [2]. FCNs have been shown to reveal key high-dimensional features of normal and abnormal nervous system physiology [2], as well as the potential to predict treatment outcome [11]. However, no studies have computationally analyzed the EEG signals during the evolution from IS to LGS. Therefore, the goal of this study is to evaluate the EEG-based functional connectivity of subjects during this progression. Based on prior work [2], [3], we hypothesize that the functional connectivity strength will be a robust biomarker to determine the presence of these epilepsies and assess treatment response.

## 2. Methods



**Figure 1** Summary of the functional connectivity analysis. (A) EEG scalp electrodes are used to (B) record EEG data. (C) Data is divided into one-second windows and the maximum cross-correlation between electrodes pairs is calculated for each time window. (D) The cross-correlation matrices are compared to a critical matrix, which contains thresholds determined using a baseline distribution. A binary matrix is obtained for each time window, which reflects whether a connection is significant (greater than the 95<sup>th</sup> percentile of the baseline distribution) or not. (E) The connectivity matrix shows the percent of one-second windows with significant cross-correlation values for each electrode pair. (F) Finally, the 10% strongest connections are plotted on a head map to visualize the connectivity network.

### 2.1. Data characteristics

The data used for this study consisted of serial EEG studies from five subjects diagnosed with IS at the Children’s Hospital of Orange County (CHOC), who later progressed into LGS. Approval to perform this study was obtained by the CHOC Institutional Review Board.

Data were recorded with nineteen EEG electrodes positioned on the scalp using the International 10-20 system and sampled at 200 Hz. Each subject had EEG recordings at the time of IS and LGS diagnoses, with varying numbers of EEGs in between, and at least one EEG following LGS treatment. For each recording, we clipped segments of awake, spontaneous EEG for analysis. Each EEG was clipped by a board-certified epileptologist at CHOC. The diagnosis of IS was clinically determined by the existence of epileptic spasms with abnormal EEG background, with or without hypsarrhythmia – a multifocal chaotic EEG pattern with asynchronous high amplitude slow waves. Further, LGS was assessed by the presence of slow-spike-and-wave (SSW) complexes, typically occurring at 1.5-2.5 Hz.

### 2.2. Data preprocessing

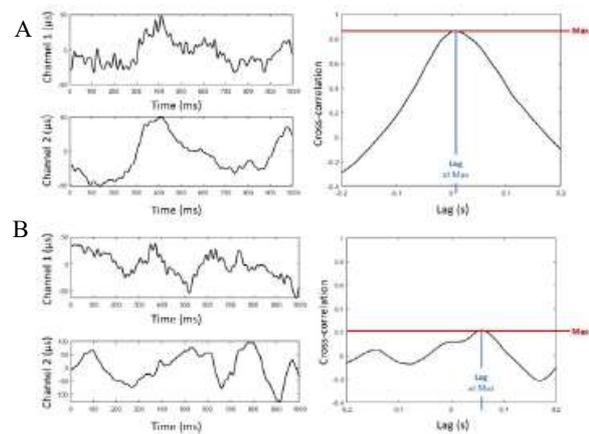
All EEG data were re-referenced to the common average and filtered with a broadband filter (0.5 – 55Hz). The data were windowed into one-second epochs, and epochs with artifacts were removed. Artifacts were identified using an amplitude-based automatic artifact detector described in more detail in Smith et al. [2]. Six hundred clean epochs during wakefulness (totaling 10 minutes) were randomly selected from each EEG recording and used for the connectivity analysis.

### 2.3. Connectivity calculation: Cross-Correlation

Here, we studied connectivity via cross-correlation using the method developed by Kramer et al. [12] and Chu et al. [13] and previously applied by our group [2] (Fig. 1).

First, each one-second epoch of EEG was normalized to have zero mean and unit variance. Then we calculated the absolute value of the cross-correlation for each electrode pair and each epoch, within a maximum lag of 200 ms, and we identified the maximum cross-correlation magnitude (Fig. 2). Then, significance was assessed by comparing the cross-correlation maximum values to a baseline distribution, generated under the null hypothesis that there was no temporal relationship between the two signals. A

connection was considered significant when its value exceeded the 95<sup>th</sup> percentile of the baseline distribution. Finally, the connectivity strength of an electrode pair equaled the percentage of one-second epochs with significant cross-correlation values. For visualization, network maps were generated showing the 10% strongest connections from the connectivity matrix.



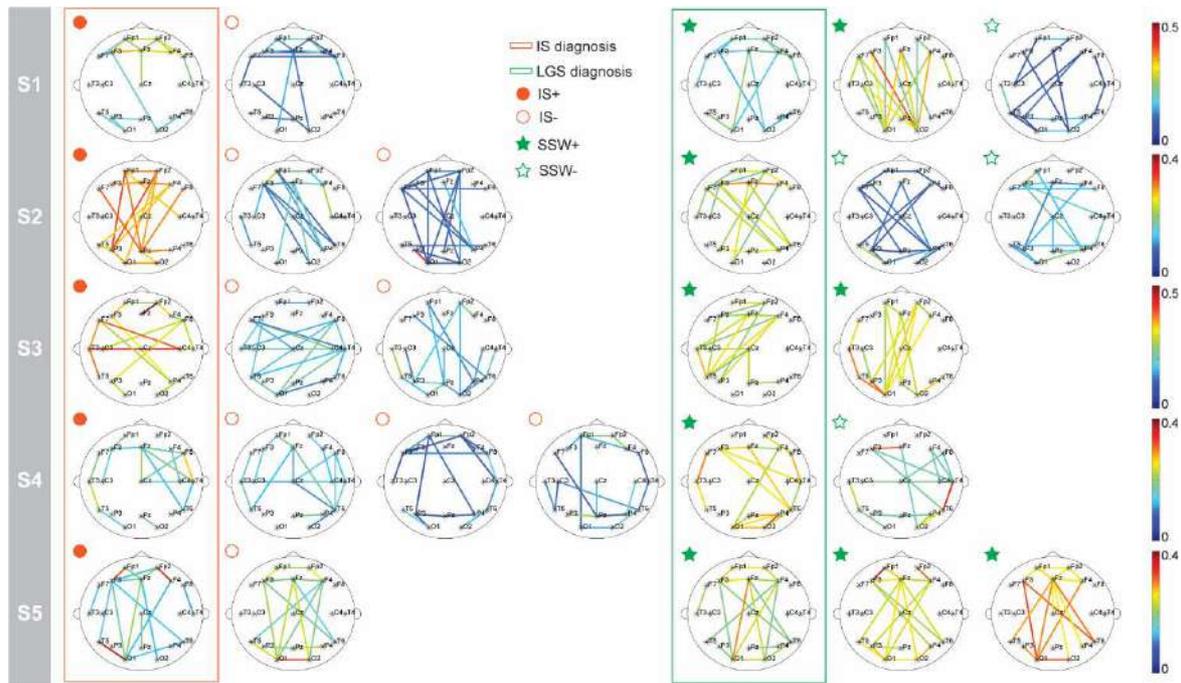
**Figure 2** Examples of (A) high and (B) low maximal cross-correlation values between pairs of EEG signals.

### 2.4. Network strength metrics

To quantify the network strength, the number of connections with a connectivity strength greater than 0.1 and the mean of the 10% highest connectivity values were computed for each EEG recording. Then these variables were compared to the disease state progression, including diagnosis of IS and LGS and positive/negative response to treatment. Finally, these two metrics were also compared to subject age at the time of the EEG recording.

## 3. Results

Figure 3 shows the EEG connectivity results obtained for the five patients in the study. The plots are organized with time increasing from left to right, with the EEGs associated with IS diagnosis and LGS diagnosis aligned vertically for comparison. Each row corresponds to one subject. Overall, we found that connectivity strength is high at the time of IS and LGS diagnoses. Disease resolution after treatment is associated with a decrease in connectivity, while the connectivity strength increases or is maintained with negative treatment outcome.

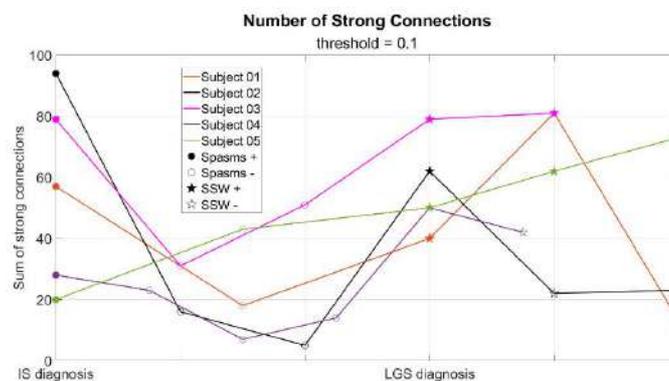


**Figure 3** EEG functional topographic maps for patients that were diagnosed with IS (red box) and were later diagnosed with LGS (green box). Each row is one subject, and time increases from left to right. For visualization purposes, each connectivity map shows the 10% strongest connections. Therefore, the same number of connections is shown in each connectivity map. The color of each connection represents the percentage of epochs with significant connectivity for the associated electrode pair. The symbol next to each map indicates whether spasms or the SSW pattern were present (filled circles and stars) or absent (open circles and stars), with an absence indicating a positive response to treatment.

For example, if we analyze the networks of subject 1 (Fig. 3, first row), we can see that the child was diagnosed with IS, followed by an IS resolution after treatment, which is characterized by a decrease in connectivity strength. Afterwards, the subject was diagnosed with LGS, and this was associated with an increase in connectivity strength. Initially, the treatment failed to control the slow spike-and-wave EEG activity, and the functional connections at this time point (fourth from the left) increase dramatically. Finally, the connectivity strength decreases with SSW resolution after LGS treatment.

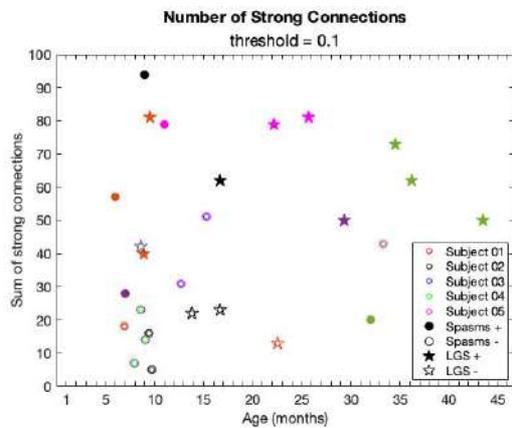
Subject 3 (Fig. 3, third row) follows a similar pattern. The main difference is that, after LGS treatment, there is never a resolution of the SSW activity, and this is associated

with continued high connectivity strength after treatment. We quantified these results using two metrics: the number of strong connections above a 0.1 threshold and the mean strength of connections above the 90th percentile. These metrics produced approximately equivalent results. Consistent with our hypothesis, IS diagnosis was associated with a high number of strong connections, and this number decreased with IS resolution (Fig. 4). LGS diagnosis was also associated with an increase in the number of strong connections, relative to each patient’s prior measurements. After LGS treatment, patients with SSW resolution exhibited a decrease in connectivity strength, while those with negative treatment outcome showed a maintenance or increase in connectivity strength (Fig. 4).



**Figure 4** The number of connections above a connectivity strength threshold of 0.1 is correlated to the disease state progression. The x-axis shows normalized time, where the time of IS and LGS diagnoses are aligned for all subjects. Each line represents one subject. The markers refer to the diagnosis, with a circle representing spasms and a star for LGS. Further the presence or absence of spasms and SSW are indicated by filled or open markers, respectively.

These two metrics were also analyzed as a function of subject age at the time of the recordings. We found that connectivity strength was not correlated to age, suggesting that the network differences in Fig. 4 are not due to normal physiological changes as the children grow (Fig. 5).



**Figure 5** Number of strong connections above a threshold of 0.1 as a function of subject age at the time of EEG recording. No correlation was found between the two variables. ( $P > 0.2$ ).

#### 4. Conclusions

Consistent with prior results, we found that functional connectivity strength reflected the presence of IS, as well as a positive response to treatment. IS diagnosis was associated with high strength, while a resolution of spasms after treatment was associated with a decrease in the connectivity strength. LGS diagnosis was similarly associated with an increase in strength. After LGS treatment, subjects could be divided into two groups. Those with a positive treatment outcome exhibited a decrease in the connectivity strength, while a negative outcome was associated with maintenance or increase of the connectivity strength.

Therefore, EEG functional connectivity has potential as a biomarker to support LGS diagnosis in patients with an IS history. It may also be useful to assess treatment response, which could improve the prognosis of LGS patients.

However, it is critical to repeat this analysis in a larger cohort of subjects to increase the statistical power of the study and validate these results. The subjects in this study needed to fulfill specific characteristics, the main one being a history of IS. Because this disease is rare, the number of subjects under analysis is a clear limitation of the study. However, larger multicenter studies may be able to overcome this limitation and enable analysis of subgroups based on etiology, which is likely a source of variance in the results.

#### Acknowledgments

This work was supported by the UC Irvine California-Catalonia Engineering Program, through the Balsells Mobility Fellowship to BRM. Funding was also provided

by the Children's Hospital of Orange County Physician Scientist Scholar's Program.

#### References

- [1] J. E. Piña-Garza et al., "Assessment of treatment patterns and healthcare costs associated with probable Lennox-Gastaut syndrome," *Epilepsy and Behavior*, vol. 73, pp. 46–50, 2017, doi: 10.1016/j.yebeh.2017.05.021.
- [2] R. J. Smith, D. K. Hu, D. W. Shrey, R. Rajaraman, S. A. Hussain, and B. A. Lopour, "Computational characteristics of interictal EEG as objective markers of epileptic spasms," *Epilepsy Research*, vol. 176, no. January, p. 106704, 2021, doi: 10.1016/j.eplepsyres.2021.106704.
- [3] W. Stacey et al., "Emerging roles of network analysis for epilepsy," *Epilepsy Research*, vol. 159, no. December 2019, p. 106255, 2020, doi: 10.1016/j.eplepsyres.2019.106255.
- [4] M. Wong and E. Trevathan, "Infantile spasms," *Pediatric Neurology*, vol. 24, no. 2, pp. 89–98, 2001, doi: 10.1016/S0887-8994(00)00238-1.
- [5] J. A. Nelson, S. Demarest, J. Thomas, E. Juarez-Colunga, and K. G. Knupp, "Evolution of Infantile Spasms to Lennox-Gastaut Syndrome: What Is There to Know?," *Journal of Child Neurology*, vol. 36, no. 9, pp. 752–759, Aug. 2021, doi: 10.1177/08830738211000514.
- [6] S. J. You, H. D. Kim, and H. C. Kang, "Factors Influencing the Evolution of West Syndrome to Lennox-Gastaut Syndrome," *Pediatric Neurology*, vol. 41, no. 2, pp. 111–113, Aug. 2009, doi: 10.1016/j.pediatrneurol.2009.03.006.
- [7] A. A. Asadi-pooya, "LGS a comprehensive review," pp. 403–414, 2018.
- [8] E. Trevathan, C. C. Murphy, and M. Yeargin-Allsopp, "Prevalence and descriptive epidemiology of Lennox-Gastaut syndrome among Atlanta children," *Epilepsia*, vol. 38, no. 12, pp. 1283–1288, 1997, doi: 10.1111/j.1528-1157.1997.tb00065.x.
- [9] A. Calvo, M. C. Buompadre, A. Gallo, R. Gutiérrez, G. R. Valenzuela, and R. Caraballo, "Electroclinical pattern in the transition from West to Lennox-Gastaut syndrome," *Epilepsy Research*, vol. 167, Nov. 2020, doi: 10.1016/j.eplepsyres.2020.106446.
- [10] B. Romero Milà, K. Remakanthakurup Sindhu, J. R. Myttinger, D. W. Shrey, and B. A. Lopour, "EEG biomarkers for the diagnosis and treatment of infantile spasms," *Front Neurol*, vol. 13, 2022, doi: 10.3389/fneur.2022.960454.
- [11] D. W. Shrey, O. Kim McManus, R. Rajaraman, H. Ombao, S. A. Hussain, and B. A. Lopour, "Strength and stability of EEG functional connectivity predict treatment response in infants with epileptic spasms," *Clinical Neurophysiology*, vol. 129, no. 10, pp. 2137–2148, 2018, doi: 10.1016/j.clinph.2018.07.017.
- [12] M. A. Kramer, U. T. Eden, S. S. Cash, and E. D. Kolaczyk, "Network inference with confidence from multivariate time series," *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, vol. 79, no. 6, pp. 1–13, 2009, doi: 10.1103/PhysRevE.79.061916.
- [13] Y. J. Chu, C. F. Chang, W. C. Weng, P. C. Fan, J. S. Shieh, and W. T. Lee, "Electroencephalography complexity in infantile spasms and its association with treatment response," *Clinical Neurophysiology*, vol. 132, no. 2, pp. 480–486, 2020, doi: 10.1016/j.clinph.2020.12.006

# Creación y optimización automática de arquitecturas de aprendizaje profundo aplicadas a imagen médica

A. Gómez<sup>1,2</sup>, M. Abella<sup>1,2,3</sup>, C. F. Del Cerro<sup>1,2</sup>, M. Desco<sup>1,2,3,4</sup>

<sup>1</sup> Departamento de Bioingeniería e Ingeniería Aeroespacial, Universidad Carlos III, Madrid, España,

<sup>2</sup> Instituto de Investigación Sanitaria Gregorio Marañón, Madrid, España

<sup>3</sup> Centro Nacional Investigaciones Cardiovasculares Carlos III (CNIC), Madrid, España

<sup>4</sup> Centro de Investigación Biomédica en Red de Salud Mental (CIBERSAM), Madrid, España

## Resumen

*El Aprendizaje Profundo es una herramienta que ha permeado todos los aspectos de la medicina y la investigación biomédica. Sin embargo, el diseño de redes neuronales convolucionales es una tarea complicada y bastante subjetiva. Aunque se ha intentado automatizar utilizando las llamadas “técnicas de neuroevolución”, los resultados no han sido muy positivos debido a la complejidad de los modelos usados en la actualidad. En este trabajo se propone un algoritmo de optimización que integra técnicas de inteligencia de colmena con métodos neuroevolutivos basados en la codificación genética de redes neuronales convolucionales. El optimizador resultante, Chimera Algorithm, es capaz tanto de crear modelos desde cero como de modificar una arquitectura dada para adaptarla al problema a tratar, de forma automática y requiriendo una interacción humana mínima.*

## 1. Introducción

La aplicación de técnicas de aprendizaje profundo está revolucionando el mundo de la medicina, desde la investigación hasta la clínica. Su potencial reside en la capacidad de extraer de forma automática patrones de los conjuntos de datos con que se entrenan. Sin embargo, la creación de modelos de aprendizaje profundo presenta sus propios desafíos. Por ejemplo, no existen unas pautas definidas para el diseño de arquitecturas de redes neuronales convolucionales, ni es posible predecir con exactitud el rendimiento de las mismas previo a su entrenamiento [1]. Esto hace de su desarrollo un proceso complejo y muy subjetivo.

Este trabajo propone un marco para el diseño o mejora automática de modelos de redes neuronales convolucionales basado en la integración de algoritmos metaheurísticos de inteligencia de colmena, específicamente el *Artificial Bee Colony Algorithm* [2], con técnicas de neuroevolución basadas en algoritmos genéticos [3]. Como resultado hemos diseñado un nuevo optimizador mixto —bautizado como *Chimera Algorithm* debido a su naturaleza híbrida— capaz de crear y modificar redes neuronales convolucionales de forma automática y con la mínima interacción humana. Como ejemplo, hemos aplicado este marco al problema de la calibración geométrica en Tomografía Computarizada.

### 1.1. Redes Neuronales Convolucionales

Las redes neuronales convolucionales consisten en una lista ordenada de capas —generalmente convolucionales y

de *pooling*— que alojan una serie de *kernels* que modifican secuencialmente la imagen de entrada. Los valores de cada kernel —llamados parámetros o pesos— evolucionan gracias a la exposición del modelo a un conjunto de datos de entrenamiento mediante un proceso denominado retropropagación. Sin embargo, el número, dimensiones, tipo y distribución de las capas —es decir, la arquitectura de la red— pertenecen al conjunto de hiperparámetros que la definen y que son seleccionados a priori por el diseñador de la misma. Para optimizar el entrenamiento se emplea comúnmente la técnica del aprendizaje transferido, que consiste en utilizar un modelo prediseñado y preentrenado para un problema similar a aquel a tratar, y reentrenarlo con el nuevo conjunto de datos. Algunas familias de modelos de aprendizaje transferido más comúnmente utilizados para el procesamiento de imágenes son Xception, Vgg y ResNet, todas ellas preentrenadas con la base de datos ImageNet, [4].

### 1.2. Algoritmos Metaheurísticos

Los algoritmos metaheurísticos son procedimientos de alto nivel diseñados para encontrar soluciones a problemas complejos de forma iterativa y experimental. La mayoría de estos algoritmos son independientes del problema a tratar, se inspiran en procesos naturales e incluyen un cierto grado de estocasticidad [5]. Una de las ramas de la metaheurística, denominada neuroevolución, se centra en aplicar dichos procesos para optimizar los parámetros o incluso los hiperparámetros de redes de aprendizaje profundo. Sin embargo, la complejidad y tamaño que presentan los modelos convolucionales actuales hacen imposible el uso de las técnicas de neuroevolución tradicionales, más centradas en el desarrollo de redes neuronales no convolucionales, mucho menos densas [6].

Los algoritmos metaheurísticos más utilizados para la optimización de modelos de aprendizaje profundo son los algoritmos genéticos [3] debido a su simplicidad y a su capacidad de trabajar con soluciones no numéricas mediante la mutación de sus parámetros —como la cantidad, tamaño o tipo de capas de una red neuronal convolucional—. Los algoritmos de inteligencia de colmena, pese a alcanzar altos niveles de complejidad y rendimiento mediante las interacciones simples de sus miembros [7], no se utilizan para optimizar estas arquitecturas ya que requieren de espacios numéricos para funcionar. El *Artificial Bee Colony Algorithm* [2] es un

optimizador basado en inteligencia de colmena que imita el comportamiento de las abejas melíferas y que puede explotar hiperespacios no continuos ni diferenciables. Tanto los algoritmos genéticos como el *Artificial Bee Colony Algorithm* están basados en una población de soluciones que evolucionan en paralelo y presentan un alto componente estocástico.

## 2. Materiales y Métodos

### 2.1. Desarrollo del *Chimera Algorithm*

El *Chimera Algorithm* se basa en el *Artificial Bee Colony Algorithm* [2], manteniendo su estructura pero sustituyendo la codificación numérica para las soluciones propuestas por una codificación genética capaz de trabajar con modelos convolucionales. Ambos algoritmos trabajan sobre una población compuesta por tres tipos distintos de agentes optimizadores —denominados abejas— que intercambian información para explorar eficientemente el hiperespacio de posibles arquitecturas. El *Chimera Algorithm* procede de la siguiente manera:

1. Se inicializa una población de ‘abejas empleadas’ y de ‘abejas espectadoras’. A cada abeja empleada se le asocia un modelo.
2. Cada abeja empleada optimiza su modelo asignado, añadiendo, destruyendo o modificando capas. Si el nuevo modelo es mejor que su predecesor, lo sustituye; de lo contrario, se añadirá un contador de extenuación al modelo anterior y se ignorará el nuevo. Las abejas empleadas vuelven a la colmena y comparten sus descubrimientos con el conjunto de abejas espectadoras que allí aguardan.
3. Cada abeja espectadora selecciona aleatoriamente uno de los modelos considerados, ponderados por su rendimiento. Los modelos que no sean seleccionados por ninguna abeja se mantienen inmutables. Cada abeja espectadora explora alrededor del modelo escogido, tratando de optimizarlo de la misma manera que las abejas empleadas en el paso 2.
4. Si un modelo excede el umbral indicado de contadores de extenuación, se almacena como presunto mínimo local y su abeja empleada asociada se reinicializa desde otro punto del hiperespacio de soluciones.
5. Los pasos 2-4 se repiten hasta que se identifique algún modelo que sobrepase un valor de rendimiento especificado por el usuario o hasta que se alcance un número máximo de iteraciones predefinido.

El *Chimera Algorithm* se puede utilizar como creador de arquitecturas desde cero, comenzando la evolución desde varios puntos aleatorios del hiperespacio, o como optimizador de un modelo dado, comenzando la evolución desde una arquitectura inicial introducida por el usuario. Para comprobar el estado del optimizador en tiempo real se ha desarrollado un visualizador que muestra el comportamiento y la evolución de los modelos propuestos, además de una representación gráfica de los mismos, tal y como se aprecia en la Figura 1.

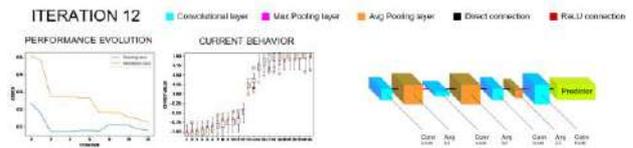


Figura 1. Visualizador desarrollado

El código del *Chimera Algorithm* ha sido desarrollado en Python. Las dependencias principales son las librerías de NumPy, PyTorch, pandas, SciPy, Matplotlib y GPUUtil.

### 2.2. Caso de uso

El problema propuesto para ejemplificar el comportamiento del *Chimera Algorithm* es la estimación mediante aprendizaje profundo del desalineamiento horizontal del detector en Tomografía Computarizada (TC) a partir del análisis de una imagen reconstruida con artefactos —dobles bordes en Figura 1—. Se han utilizado 7 volúmenes craneales de roedores obtenidos con el sistema de micro-TC de SEDECAL [8].

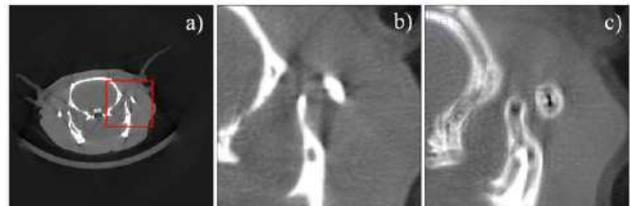


Figura 2. Corte axial de TAC de cabeza de roedor reconstruida con rango angular de 360 grados y sin desalineamientos (a) y detalle de la zona delimitada en rojo (b). Mismo detalle simulado y reconstruido con un desalineamiento horizontal en el detector de 1mm (c)

El esquema de trabajo se muestra en la Figura 2. Primero, se han simulado las proyecciones de cada uno de los volúmenes TC mediante la herramienta FUX-Sim [9] para 25 valores desalineamiento horizontal del detector tomados de una distribución aleatoria uniforme con un rango de  $\pm 1$ mm. Estas proyecciones se han reconstruido con FDK, generando volúmenes de  $256 \times 256 \times 200$  píxeles, que se han normalizado a los valores de ImageNet —media de 0.485 y desviación estándar de 0.225. La base de datos se ha dividido en 5 roedores —25000 imágenes— para el conjunto de entrenamiento, 1 roedor —5000 imágenes— para el de validación y 1 roedor —5000 imágenes— para el de test. Esta base de datos se introduce en el *Chimera Algorithm*, el cual trabaja a partir de un conjunto de 5 modelos del que selecciona como modelo final aquel que presente un mejor rendimiento con los datos de validación. Este proceso se ha realizado tanto creando los modelos desde cero —cuyos resultados se han denominado Chim— como a partir de evolucionar la arquitectura VGG11 —cuyos resultados se han denominado VGG<sub>EV</sub>—. Este modelo final es capaz de obtener una estimación de sus valores de desalineamiento a partir de los datos con artefactos y.

Para asegurar una convergencia rápida y fiable de los modelos durante el entrenamiento, se ha empleado una tasa de entrenamiento de un ciclo [10] de 32 épocas de longitud y delimitada entre  $5 \cdot 10^{-6}$  y  $5 \cdot 10^{-5}$ . Estos valores límite han sido obtenidos al pasar varios modelos generados

aleatoriamente por el test de tasa de aprendizaje propuesto en [10]. Los modelos abandonan la fase de entrenamiento en el momento en el que su rendimiento en el conjunto de validación comienza a converger, ahorrando tiempo y minimizando la posibilidad de sobreajuste. La función de pérdidas utilizada ha sido el error medio absoluto en la predicción del desalineamiento horizontal.

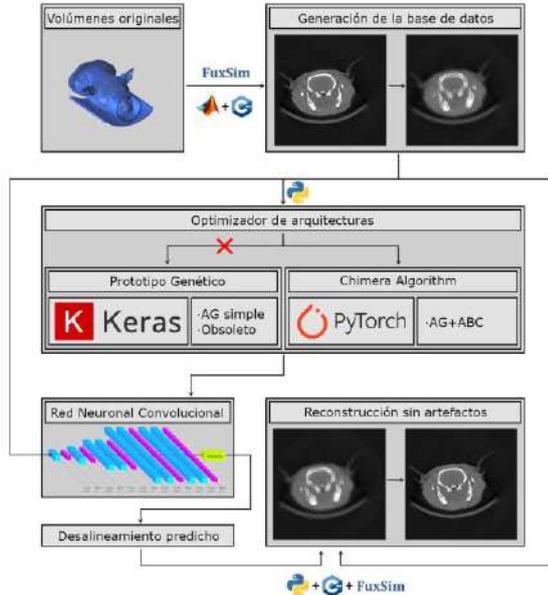


Figura 3. Esquema del flujo de trabajo propuesto

### 2.3. Evaluación

Para evaluar el rendimiento del *Chimera Algorithm* en el problema propuesto se ha comparado el error medio del valor de desalineamiento estimado en el conjunto de datos de validación de los modelos generados con los estimados por los modelos Xception, ResNet34, ResNet50, Vgg11 y Vgg19 descritos en [4]. Además, se ha evaluado el comportamiento de los distintos modelos generados a lo largo de las iteraciones.

Para comprobar su poder de generalización, todos estos modelos se han empleado para estimar los desalineamientos de dos volúmenes del conjunto de test. La mediana de los valores obtenidos para cada volumen se toma como el valor a utilizar para su reconstrucción, corrigiendo los artefactos si la predicción es correcta. Estos resultados se han analizado visualmente.

Durante el diseño y testado de los algoritmos considerados se ha utilizado una CPU Intel® Core™ i7-7700 y una GPU NVIDIA® GeForce® RTX 2060 Super™.

## 3. Resultados

### 3.1. Resultados en el conjunto de validación

El rendimiento de todos los modelos evaluados se muestra en la Tabla 1. Los modelos generados desde cero por el *Chimera Algorithm* presentan un rendimiento hasta un 25.05% mejor que el mejor modelo de aprendizaje transferido considerado, VGG19. Los modelos generados a partir de evolucionar VGG11 presentan un rendimiento hasta un 53.37% mejor que VGG19. Los cinco modelos generados a partir de VGG11 se han obtenido en 5

iteraciones del algoritmo, tomando entre 6 y 12 horas cada una, mientras que los cinco generados desde cero se han obtenido en un máximo de 18 iteraciones, tomando entre 30 minutos y 3 horas cada una.

Xception	ResNet34	ResNet50	VGG11	VGG19
0.27502	0.27256	0.35847	0.15268	0.13363
Chim I	Chim II	Chim III	Chim IV	Chim V
0.10015	0.11040	0.11934	0.12856	0.13809
VGG <sub>EV</sub> I	VGG <sub>EV</sub> II	VGG <sub>EV</sub> III	VGG <sub>EV</sub> IV	VGG <sub>EV</sub> V
0.06689	0.06231	0.09038	0.10022	0.10651

Tabla 1. Error medio del valor estimado (mm) por los distintos modelos en el conjunto de datos de validación

En la Figura 4 se puede apreciar cómo Chim I, a pesar de presentar un error alto en las primeras iteraciones, termina evolucionando hacia una arquitectura adecuada en menos de 10 iteraciones.

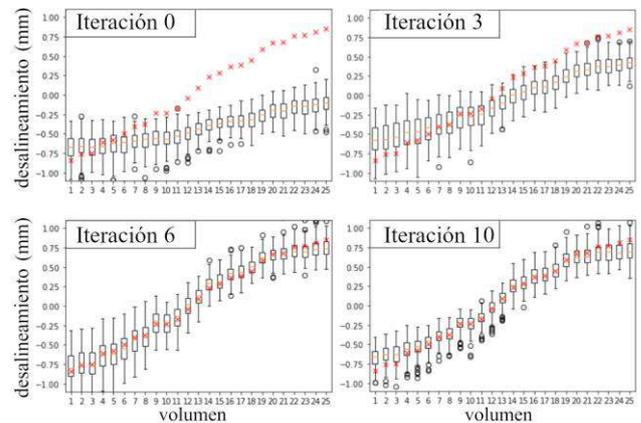


Figura 4. Predicción para las 200 rodajas de cada volumen del conjunto de validación por el modelo Chim I a lo largo de su evolución (cruces rojas: valor real del desalineamiento)

La Figura 5 muestra el comportamiento de VGG11 y VGG<sub>EV</sub> I en el conjunto de validación. Tal y como se puede apreciar, VGG11 parte de errores pequeños aunque presenta dificultades para caracterizar desalineamientos pequeños o desalineamientos grandes y positivos, que su evolución maneja sin ningún problema.

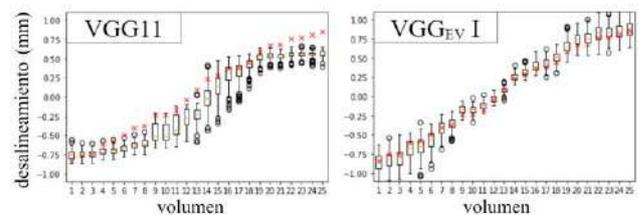
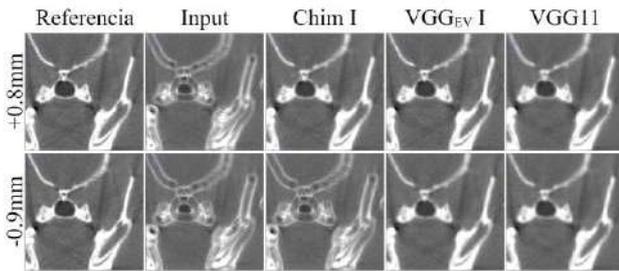


Figura 5. Predicción para las 200 rodajas de cada volumen del modelo VGG11 (izquierda) y de su mejor evolución final (derecha) en el conjunto de datos de validación

### 3.2. Resultados en reconstrucción

La figura 6 muestra un estudio del conjunto de test reconstruido teniendo en cuenta el valor predicho por las

distintas redes para un desalineamiento horizontal simulado de +0.8 y de -0.9 milímetros.



**Figura 6.** Detalle del corte axial para dos reconstrucciones del conjunto de test utilizando los valores predichos por los distintos modelos

Chim I es incapaz de predecir todos los desalineamientos del conjunto de test, reflejándose en los artefactos sin corregir en los volúmenes reconstruidos. Por otro lado, VGG11 produce en todos los casos resultados ligeramente peores que los obtenidos con su evolución, VGG<sub>EV</sub> I.

#### 4. Discusión y Conclusiones

En este trabajo se ha desarrollado el *Chimera* Algorithm, un creador y optimizador de redes neuronales convolucionales que combina las propiedades emergentes de la inteligencia de colmena con una codificación genética para las arquitecturas propuestas, obteniendo una exploración del hiperespacio de soluciones automática y con una intervención humana mínima.

Hemos comprobado en un caso real que las arquitecturas propuestas por el *Chimera* Algorithm presentan un mejor rendimiento que las técnicas de aprendizaje transferido tradicionales. Esto se debe a que el *Chimera* Algorithm permite alcanzar arquitecturas optimizadas que en un flujo de trabajo convencional no se suelen buscar de forma heurística debido a la cantidad de tiempo y trabajo que requeriría. Sin embargo, algunas de las arquitecturas generadas desde cero han presentado dificultades para generalizar al conjunto de datos de test. Esto puede deberse a que los modelos generados no son lo suficientemente complejos como para aprender los patrones reales del conjunto de datos. Por lo tanto, en lugar de forzar la evolución hacia arquitecturas más densas, se acaba estancando en modelos pequeños pero sobreajustados al set de validación, los cuales presentan óptimos locales no representativos del total del hiperespacio de arquitecturas. El diseño de una serie de rigurosas pruebas estadísticas, tanto durante como tras la evolución, ayudaría a garantizar la capacidad de generalización del modelo final seleccionado. Asimismo, la exploración del hiperespacio de arquitecturas podría mejorarse mediante la implementación de sistemas de puntuación por novedad [11] o la reinicialización de las arquitecturas mediante el cruce de modelos generados en lugar de utilizar reinicialización aleatoria.

Este trabajo muestra el potencial de los algoritmos basados en inteligencia de colmena para explotar y automatizar el proceso de creación de redes neuronales convolucionales, obteniendo arquitecturas que superan el rendimiento obtenido con aprendizaje transferido tradicional.

#### Agradecimientos

Este trabajo ha sido apoyado por el Ministerio de Ciencia e Innovación, Agencia Estatal de Investigación: PDC2021-121656-I00 (MULTIRAD), financiado por MCIN/AEI/10.13039/501100011033 y por la Unión Europea *NextGenerationEU*/PRTR; PID2019-110369RB-I00/AEI/10.13039/501100011033 (RADHOR), además de la Comunidad de Madrid: Acuerdo Multianual con la UC3M en la línea *Fostering Young Doctors Research* (DEEPC-TM-UC3M), y en el contexto del V PRICIT (Plan Regional de Investigación Científica e Innovación Tecnológica); S2017/BMD-3867 RENIM-CM, cofinanciado por el *European Structural and Investment Fund*. También parcialmente financiado por CRUE Universidades, CSIC y Banco Santander (Fondo Supera Covid19), proyecto RADC0V19. El CNIC es apoyado por el Instituto de Salud Carlos III, Ministerio de Ciencia e Innovación y la Fundación Pro CNIC.

#### Referencias

- [1] Jiang C et al. Neural Capacitance: A New Perspective of Neural Network Selection via Edge Dynamics. *arXiv*, 2022 (DOI: 10.48550/ARXIV.2201.04194)
- [2] Karaboga D, Basturk B. A powerful and Efficient Algorithm for Numerical Function Optimization: Artificial Bee Colony (ABC) Algorithm. *Journal of Global Optimization*, vol 39, 2007, pp 459-171 (ISSN:0925-5001)
- [3] Badan F, Sekanina L. Optimizing convolutional neural networks for embedded systems by means of Neuroevolution. *Theory and Practice of Natural Computing, Springer International Publishing*, 2019, pp. 109–121 (DOI: 10.1007/978-3-030-34500-6\_7)
- [4] Lendave V. A Comparison of 4 Popular Transfer Learning Models. *Analytics India Magazine*. <https://analyticsindiamag.com/a-comparison-of-4-popular-transfer-learning-models/> (Consultada: Julio 2022)
- [5] Blum C, Roli A. Metaheuristics in combinatorial optimization: Overview and conceptual comparison. *ACM Comput. Surv.*, vol. 35, 2001, pp. 268–308 (DOI: 10.1145/937503.937505)
- [6] Hutson M. Artificial intelligence can ‘evolve’ to solve problems. *Science*, 2018 (DOI: 10.1126/science.aas9715)
- [7] Luo Y et al. Swarm intelligence for complex energy networks. *Hybrid Systems and Multi-energy Networks for the Future Energy Internet, Academic Press*, chap. 9, sec. 4, 2021, pp. 217-234 (ISBN: 9780128191842)
- [8] Vaquero JJ et al. Assessment of a New High-Performance Small-Animal X-Ray Tomograph. *IEEE Transactions on Nuclear Science*, vol. 55, 2005, pp 898-905, 2008 (DOI: 10.1088/0031-9155/54/18/005)
- [9] Abella M et al. Fux-sim: Implementation of a fast universal simulation/reconstruction framework for x-ray systems. *PLOS ONE*, vol. 12, no. 7, 2017, pp. 1–22 (DOI: 10.1371/journal.pone.0180363)
- [10] Gugger S. The 1cycle policy. *Another data science student’s blog*. <https://sgugger.github.io/the-1cycle-policy.html> (Consultada: Julio 2022)
- [11] Sarti S, Adair J, Ochoa G. Recombination and Novelty in Neuroevolution: A Visual Analysis. *SN Computer Science*, vol. 3, art. 185, 2022 (DOI: 10.1007/s42979-022-01064-6)

# Quantification of the adhesion force of *E. coli* on Ti via single-cell force spectroscopy

Judith Zubia Aranburu<sup>1,2</sup>, Brunero Cappella<sup>1</sup>, Alaitz Zabala Eguren<sup>2</sup>, Lorea Buruaga Lamarain<sup>2</sup>, Andrea Aginagalde Lopez<sup>2</sup>, Jörn Bonse<sup>1</sup>, Karin Schwibbert<sup>1</sup>

<sup>1</sup> Bundesanstalt für Materialforschung und -prüfung (BAM), Berlin, Germany, {joern.bonse, karin.schwibbert}@bam.de

<sup>2</sup> Mondragon Unibertsitatea-Faculty of Engineering, Arrasate-Mondragón, Spain, judith.zubia@alumni.mondragon.edu, {azabalae, lburuaga, aaginagalde}@mondragon.edu

## Abstract

*Antibiotic resistance is a growing global problem which poses a massive threat to human health. Although human activity contributes to the acceleration of the process, bacteria have a self-driven stabilisation mechanism to protect themselves from such and other external threats: biofilm formation. Nonetheless, it is the adhesion of a single bacterial cell to a surface that triggers the formation of such network of biomolecules and microorganisms, as well as its hazardous consequences. The main objective of this work was to quantify the adhesion force of a single *E. coli* cell on a Ti substrate via the AFM-related single-cell force spectroscopy, with both the cell and the substrate material being of high clinical relevance. A set of 25 x 25 force-displacement curves was acquired with a maximum force of 3.2 nN without dwell time, yielding a topography map and an adhesion force map that showed to be correlated. A mean adhesion force of  $0.85 \pm 0.175$  nN was measured and the presence of cell appendages on the bacterial cell wall was verified through individual force-displacement curves. Bacterial viability was assessed after the measurements via live/dead staining.*

## 1. Introduction

According to the World Health Organization (WHO), antibiotic resistance poses a major global threat to both human health and development. As a matter of fact, more than 750,000 annual deaths are caused worldwide by resistant bacteria [1]; number estimated to increase 13-fold by 2050 [2]. The economic impact of this phenomenon does not go unnoticed either. Global healthcare costs are estimated to increase from EUR 265 billion to more than EUR 880 billion per year by 2050 [3].

According to the guidance for United Nations country teams with regard to antimicrobial resistance [4], this issue is present in 4 of the 17 Sustainable Development Goals (SDGs). First, and most importantly, antibiotic resistance is central to SDG 3 (Good Health and Well-Being) as 200,000 annual deaths in newborns and the struggle to achieve a universal health coverage due to high treatment costs are the result of antimicrobial resistance. Second, to SDG 1 (No poverty) as it is estimated that the chronification of infections and high treatment costs could be at the root of driving more than 28.3 million people to extreme poverty. Third, to SDG 2 (Zero hunger) since the overuse of antibiotics in animals calls into question food safety. Fourth, to SDG 8 (Decent work and economic growth) as worsening global health caused by antibiotic resistance could reduce economic activity up to 3 % [5].

Although human actions such as over-prescription and misuse of antibiotics, poor hygiene conditions, and mass population mobility led by globalisation increase the occurrence of antibiotic resistant bacteria, it is worth noting that antibiotic resistance is a naturally occurring phenomenon. Among others, biofilm formation is one of the mechanisms through which bacteria form stabilised colonies that allow to develop resistance towards antibiotics [6]. A biofilm is a complex sessile community of bacteria irreversibly embedded in a matrix of extracellular polymeric substances (EPS), mainly made of extracellular DNA, proteins, and polysaccharides [7].

Biofilm formation is seeded by the attachment of a single planktonic cell to a substrate. Indeed, it is this initial attachment of a single cell that gives rise to a series of physical and behavioural changes in bacteria, that lead to the formation of big communities. After the initial attachment, bacteria dispose of their cell appendages to give rise to an irreversible attachment followed by cell division and cellular aggregation. In parallel, bacteria produce the EPS for the acquisition of water and nutrients from the environment when either are scarce. Finally, dispersion of bacteria from the biofilm takes place, invading new surfaces and thus contributing to the diffusion of the biofilm.

Because of the importance of the initial adhesion of a single bacterial cell onto a surface for the biofilm formation process and its subsequent effects, major scientific attention has been drawn towards the analysis of the process and quantification of the adhesion forces of bacteria on a plethora of substrates.

The Atomic Force Microscope (AFM) has proved to be an effective and reliable tool for the quantification of such small forces. This high-resolution non-optical microscopic technique [8] offers the possibility to perform single-cell force spectroscopy (SCFS). SCFS relies on immobilizing a single living cell on an AFM cantilever and then measuring the interaction forces between the cell probe and a solid substrate by approaching and retracting the probe to and from the substrate, respectively [9].

In this context, the purpose of this work is to quantify the adhesion force of a single cell of the clinically relevant *E. coli* species on a pure titanium substrate via SCFS, being this material widely used in the medical field.

## 2. Materials and methods

### 2.1. Cultivation of bacteria

*Escherichia coli* TG1 DSM 6065 single colonies were grown in Luria-Bertani (LB) medium at 37 °C with shaking at 120 rpm on an incubating orbital shaker (3500I, VWR®). The overnight culture was diluted 1:100 in fresh LB medium and incubated under the same conditions for around 1.5 h, until cells had reached the exponential growth phase. 8 mL of the culture was centrifuged at 5,000 ref for 2.5 min and resuspended in 8 mL phosphate-buffered saline (PBS) (pH = 7.4). The optical density (OD<sub>600nm</sub>) was measured and adjusted to 0.1 in PBS (Ultrospec 10, Amersham Biosciences), followed by a 1:10 dilution for the acquisition of a bacterial suspension with an OD of 0.01. The bacterial suspension was stored in ice until the beginning of the single-cell attachment procedure.

### 2.2. Preparation of the bacterial probe

A rectangular tipless chromium-gold coated silicon cantilever of 0.03 N/m (HQ:CSC38/tipless/Cr-Au, MikroMasch®) was cleaned with argon plasma for 30 sec and 75 Watts, before it was functionalised with 4 mg/mL polydopamine (PDA) in Tris buffer (10 mM, pH = 8.5 at 25 °C) at 5.5 °C for 1 h, rinsed with ultrapure water (0.055 µS/cm at 25 °C) and vacuum-dried for 15 min. The probe was inserted in the Cypher S AFM (Asylum Research) and brought into contact with an isolated and stationary bacterium with 3.2 nN for 5 min.

### 2.3. Adhesion force measurements

A force volume of 25 x 25 force-displacement curves was measured in PBS on polished Ti (99.6 % purity, MaTecK GmbH) over an area of 20 x 20 µm<sup>2</sup>. Force-displacement curves were acquired with a maximum force of 3.2 nN without dwell time. Data was processed using the IGOR Pro software.

### 2.4. Viability of bacteria

Bacterial viability was assessed 6 hours after the measurements via life/dead staining (BacLight™ Bacterial Viability kit, Invitrogen™).

## 3. Results and discussion

### 3.1. Analysis of topography and adhesion force maps

As a result of the force volume measurement, a topography map (see Figure 1) and an adhesion force map (see Figure 2) were acquired. The topography map shows an overall flat surface with two protruding features with a maximum height of 650 nm. The adhesion force map visualises the adhesion force of the bacterium with the sample surface.

The adhesion force was homogeneous over the flat part of the surface. Also, as the contact area between the bacterium and the surface was significantly reduced at the topography peaks, the adhesion force was significantly reduced on the protrusions. The similarity between both maps proves the correlation between topography and adhesion force.

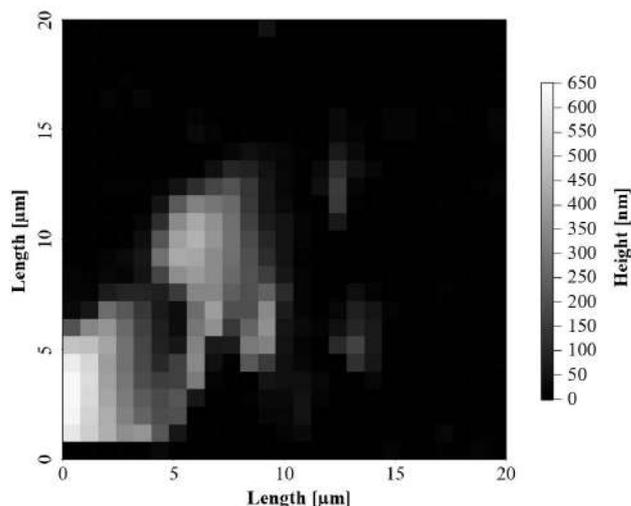


Figure 1. Topography map of polished Ti

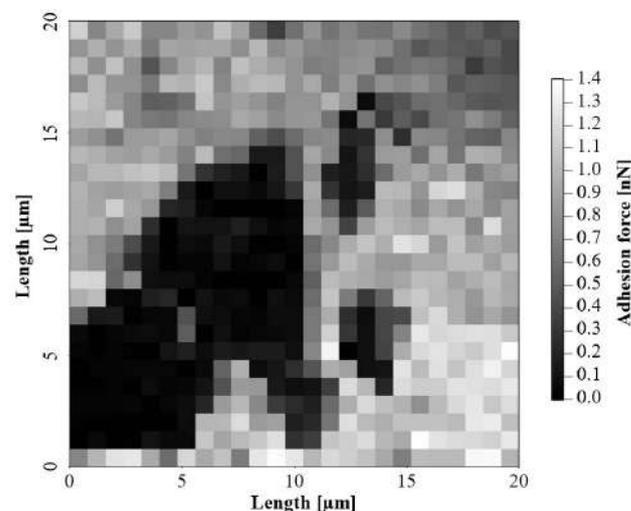


Figure 2. Adhesion force map of *E. coli* TG1 with polished Ti

Figure 3 shows the distribution of the adhesion force values and its corresponding Gaussian fit. Bright grey bars correspond to the adhesion force values that were not used for the Gaussian fit, because they correspond to the highest values of the topography, where a homogeneous contact between the bacterium and the surface was hampered.

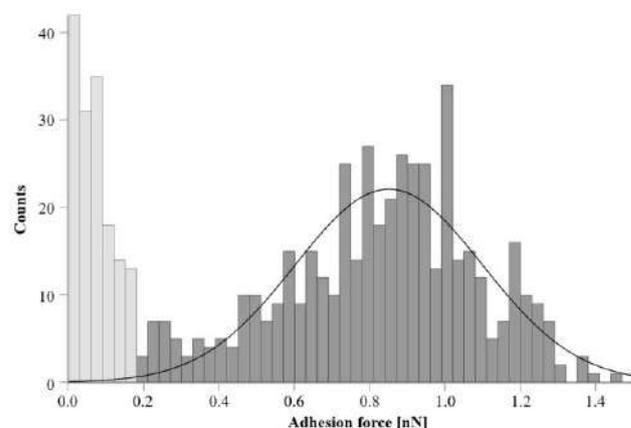


Figure 3. Histogram of adhesion force and its Gaussian fit

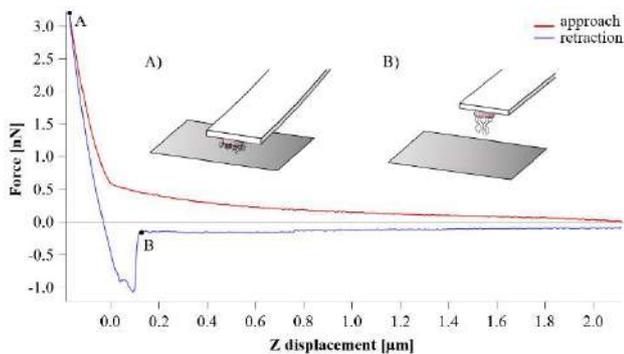
As it can be seen, the data is normally distributed, with a mean maximum adhesion force of  $0.85 \pm 0.175$  nN. This

value is somewhat different from the results obtained by Aguayo et al. on Ti with a force trigger of 500 pN without dwell time for *S. sanguinis*, whereby they measured a mean adhesion force of  $0.32 \pm 0.00$  nN [10]. The Ti substrates they used had an average roughness  $R_a$  value of  $0.61 \pm 0.01$   $\mu\text{m}$ , similar to the Ti sample used for this experiment.

However, it should be noted that the experimental conditions differ significantly from the ones used here: *S. sanguinis* is a gram-positive bacterium whereas *E. coli* is a less stiff gram-negative one; their bacterial probe was prepared using a microbead, whereas in this case it was not; their cantilever spring constant was of 0.3 N/m, whereas in this case it was of 0.03 N/m; and their force trigger was of 500 pN, whereas in this case it was of 3.2 nN.

### 3.2. Analysis of force-displacement curves

Figure 4 shows a representative force-displacement curve of the homogeneous and flat region of the polished Ti substrate. Several aspects should be discussed in the following.



**Figure 4.** Force-displacement curve on polished Ti.

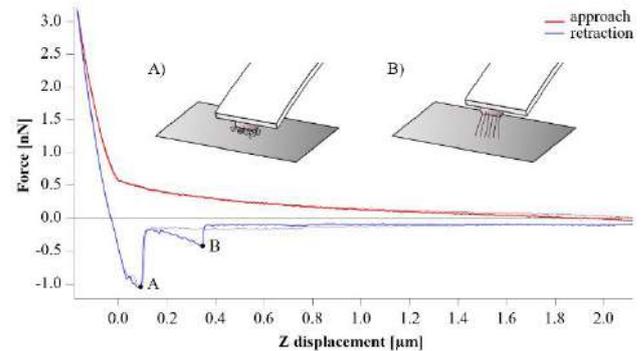
First of all, the so-called “zero line” of the approach curve was not strictly horizontal. This was due to the double-layer force. When a surface is immersed in a solvent, the charged molecules in the surface may get dissociated or the ones in the fluid may get adsorbed. When another surface with the same charge approaches, in this case the bacterial probe, the layers of charges get perturbed, resulting in a repulsive force [11][12][13]. This gave rise to the exponential shape of the approach curve.

Second, after reaching the force trigger of 3.2 nN (point A in Figure 4), the retraction procedure was initiated. However, the contact line of the retraction curve did not overlap with the one of the approach curve. This is the result of a plastic deformation on the “tip” of the cantilever, which in this case, is a bacterium. This detail confirms the presence of a bacterium on the apex of the cantilever.

Third, the retraction curve reached a minimum value, determined by the maximum adhesion force, in this case of around 1 nN. At that point, the bacterium detaches from the surface, leading to a sudden reduction in the adhesion force (point B in Figure 4).

Figure 5 shows another representative force-displacement curve on the homogeneous region of the Ti sample, but in this case, it indicates the presence of cell appendages (e.g., pili) at the cell wall of the bacterium. The thin approach and retraction curves are the force-displacement curve of

Figure 4, shown for comparison. As it can be seen, both force-displacement curves overlap almost completely, which demonstrates the high stability of the measurements. The two retraction curves only differ in the part of the curves where the bacterium is detached.

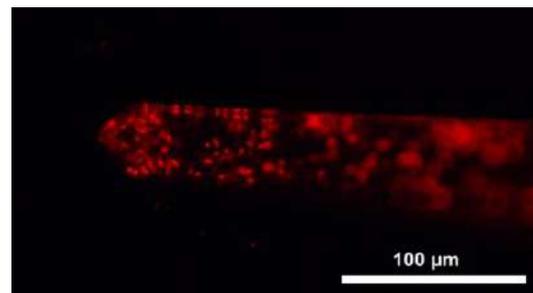


**Figure 5.** Overlapped force-displacement curves on polished Ti

Although both curves show an identical maximum adhesion force, the force-displacement curve in Figure 5 presents a secondary minimum at around 0.35  $\mu\text{m}$  above the surface. This means that, from the moment in which the cell body was totally detached from the surface (point A in Figure 5) till the second detachment event occurred (point B in Figure 5), an individual or group of cell appendages had been extending. After reaching the maximum extension, they could no longer be elongated and got therefore detached, giving rise to the secondary minimum with an adhesion force of approx. 0.4 nN. The distance between the maximum adhesion force and the second adhesion event was around 0.25  $\mu\text{m}$ , which corresponds to the length of pili cell appendages.

### 3.3. Assessment of bacterial viability

Figure 6 shows a fluorescence microscopy image of the bacterial probe life/dead stained 6 hours after the AFM measurements. Two main things can be seen: (i) that several bacteria are attached to the cantilever, and (ii) that all of them are dead.



**Figure 6.** Stained *E. coli* cantilever after the measurements

On the one hand, the attachment of several bacteria to the cantilever happened because the functionalised cantilever was immersed in a bacterial suspension, where attachment of unwanted bacteria was unavoidable. Although the attachment of several bacteria does not cause a deflection of the cantilever, at first sight, this observation may raise the question whether the measurements really correspond to a single bacterium. However, the scenario of multiple attached bacterial cells can be excluded here since due to the geometrical tilt of the cantilever only at its outmost tip,

a mechanical contact between an immobilised cell and the surface can be established during the AFM measurements.

Regarding bacterial death, to the best of the author's knowledge, three possible reasons explain why bacteria were dead after the live/dead staining. First, although bacteria were uninterruptedly under wet conditions, they did not have access to any nutrient source. Moreover, although this only affects to the bacterium getting in contact with the substrates, the measurement of 625 force-displacement curves is extremely invasive and prone to cause death. Additionally, as reported in literature: "bacteria in direct contact with the cantilever may be more subject to heating by the laser beam of the AFM optical detection system" [14]. The existence of this risk was unknown to the authors at the time of the experiment but contact between the bacterium and the cantilever could be avoided using a microbead attached to the cantilever.

However, it should also be mentioned that Figure 6 shows the state of the bacterial probe 6 hours after its preparation, which does not mean that during the measurements its state was the same. Anyhow, it is unclear to what extent it is relevant whether bacteria are alive when quantifying adhesion. While some studies limit the time of use of bacterial probes to ensure that cells are always alive [15], there is at least one study that states that viability does not affect the adhesion force of individual cells [16].

#### 4. Conclusions

A bacterial probe of *E. coli* TG1 was prepared and used to measure a force volume over a surface of polished pure Ti. With the obtained results, the correlation that exists between the topography and the adhesion force was demonstrated and a mean adhesion force of  $0.85 \pm 0.175$  nN was measured. However, due to a lack of similar studies, it was not possible to compare the results. The analysis of force-displacement curves proved the presence of a single bacterium attached to the cantilever, even going so far as to verify the presence of cell appendages. The rough estimation of the length of the cell appendages through a force-displacement curve implied that the cell appendages mainly involved in adhesive events were pili. The viability of bacteria was assessed via live/dead staining 6 hours after the preparation of the bacterial cantilever probe.

#### Acknowledgements

This work has been developed with the financial support from the Erasmus+ programme of the European Union. This study received funding through the European Union's Horizon 2020 research and innovation programme under FET Open grant agreement No. 862016 (BioCombs4Nanofibers).

#### References

- [1] F. Megraud, R. Bruyndonckx, S. Coenen, et al.. *Helicobacter pylori* resistance to antibiotics in Europe in 2018 and its relationship to antibiotic consumption in the community. *Gut*, vol. 70, no. 10, 2021, pp. 1815-1822 (ISSN: 0017-5749).
- [2] R. Wheatley, J. Diaz Caballero, N. Kapel, et al.. Rapid evolution and host immunity drive the rise and fall of carbapenem resistance during an acute *Pseudomonas aeruginosa* infection. *Nature Communications*, vol. 12, no. 2460, 2021 (ISSN: 2041-1723).
- [3] M. Ahmad, A. U. Khan. Global economic impact of antibiotic resistance: A review. *Journal of Global Antimicrobial Resistance*, vol. 19, 2019, pp. 313-316 (ISSN: 2213-7165).
- [4] World Health Organization et al.. Antimicrobial resistance and the United Nations sustainable development cooperation framework. *Guidance for United Nations country teams*, 2021, pp 2-6 (ISBN: 978-9-2400-3602-4).
- [5] Antimicrobial resistance threatens development, SDGs: Tripartite report (International Institute for Sustainable Development 2021). <https://sdg.iisd.org/news/antimicrobial-resistance-threatens-development-sdgstripartite-report/> (Accessed: April 2022).
- [6] S. Dincer, F. M. Uslu, A. Delik. *Bacterial Biofilms*. IntechOpen, 2020 (ISBN: 978-1-78985-900-3).
- [7] D. Sharma, L. Misba, A. U. Khan. Antibiotics versus biofilm: an emerging battleground in microbial communities. *Antimicrobial Resistance & Infection Control*, vol. 8, no. 76, 2019, pp. 1-10 (ISSN: 2047-2994).
- [8] G. Binnig, C. F. Quate, C. Gerber. Atomic force microscope. *Physical Review Letters*, vol. 56, no. 9, 1986, pp. 930-933 (ISSN: 0031-9007).
- [9] A. Beaussart, S. El-Kirat-Chatel, R. M. A. Sullan, et al.. Quantifying the forces guiding microbial cell adhesion using single-cell force spectroscopy. *Nature Protocols*, vol. 9, no. 5, 2014, pp. 1049-1055 (ISSN: 1754-2189).
- [10] S. Aguayo, N. Donos, D. Spratt, L. Bozec. Probing the nano-adhesion of *Streptococcus sanguinis* to titanium implant surfaces by atomic force microscopy. *International Journal of Nanomedicine*, vol. 11, 2016, pp. 1443-1450 (ISSN: 1178-2013).
- [11] B. Cappella, G. Dietler. Force-distance curves by atomic force microscopy. *Surface Science Reports*, vol. 34, no. 1-3, 1999, pp. 1-104 (ISSN: 0167-5729).
- [12] J. N. Israelachvili, D. Tabor. The measurement of van der Waals dispersion forces in the range 1.5 to 130 nm. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, vol. 331, no. 1584, 1972, pp. 19-38.
- [13] H. J. Butt, B. Cappella, M. Kappl. Force measurements with the atomic force microscope: Technique, interpretation and applications. *Surface Science Reports*, vol. 59, no. 1-6, 2005, pp. 1-152 (ISSN: 0167-5729).
- [14] A. Beaussart, S. El-Kirat-Chatel, P. Herman, et al.. Single-cell force spectroscopy of probiotic bacteria. *Biophysical Journal*, vol. 104, no. 9, 2013, pp. 1886-1892 (ISSN: 0006-3495).
- [15] D. Wilms, F. Schröer, T. J. Paul, S. Schmidt. Switchable adhesion of *E. coli* to thermosensitive carbohydrate-presenting microgel layers: A single-cell force spectroscopy study. *Langmuir*, vol. 36, no. 42, 2020, pp. 12555-12562 (ISSN: 0743-7463).
- [16] C. Spengler, F. Nolle, J. Mischo, et al.. Strength of bacterial adhesion on nanostructured surfaces quantified by substrate morphometry. *Nanoscale*, vol. 11, no. 42, 2019, pp. 19713-19722 (ISSN: 2040-3364).

# In vitro neuronal cultures and network analysis to build a proof-of-concept biological computing AI device

A. Fuster Palà<sup>1</sup>, J. Soriano<sup>2</sup>, D. Tornero<sup>3</sup>

<sup>1</sup> UPF, Barcelona, España, aubafp@gmail.com

<sup>2</sup> Neurophysics Group, UB, Barcelona, España, jordi.soriano@ub.edu

<sup>3</sup> Neural Stem Cells and Brain Damage, UB, Barcelona, España, daniel.tornero@ub.edu

## Summary

*The limitations of current electronic computation and artificial intelligence approaches have led to the emergence of biological computation, which has been studied in the last years as an alternative to traditional computation methods. The NeuChiP European project aims to develop a proof-of-concept for a biological computing AI device.*

*The first step towards biological computation is generating in vitro neuronal cultures with the appropriate characteristics for using them as new AI systems. This work aimed to study different culture setups using both primary cultures and human induced pluripotent stem cells (iPSCs)-derived cortical neurons. Also, the cultures' spontaneous activity was recorded using calcium imaging to allow the activity connectivity analysis of the networks.*

*The results suggest that it is possible to configure the culture functional connectivity by using physical constraints, chemical agents interfering in synaptic transmission, as well as mechanical agents like a random damage. Those have a direct effect on the activity and connectivity of the network and can be used to modulate it. However, after some time, cultures are able to adapt to the external changes thanks to the inherent plasticity of biological neuronal networks, which highlights the necessity of finding an effective and durable methods to retain the imprinted characteristics, possibly through stimulation. Reproducibility of results was verified in primary cultures but proved difficult in human iPSC-derived cortical neurons due to the novelty of the protocol. To conclude, this work is a starting point for future studies in biological-computing research fields.*

## 1. Introduction

Modern world is dependent upon artificial intelligence (AI) and machine learning. The estimated annual worldwide spend on AI is \$77.6B in 2022 with a business value of \$3.9T [1]. Despite the actual success of novel AI systems, important limitations are still present, being some of them potential obstacles for AI growth and long-term utility in some sectors. Indeed, AI architectures are defined heuristically based on engineering intuition, are not adaptable and do not fit naturally with conventional computer architectures; they are based on simple, typically information losing components; and are trained using ineffective learning rules. The steep increase in AI energy consumption and the computing power required for training, which has grown 300 thousand-fold between 2012-18, are both unsustainable [2, 3]. Worldwide, data centers consume about 200 terawatt hours of power per year—more than some countries. This enormous spend of energy is already being translated into environmental problems [4].

For this reason, biological computation is being studied in the last years as an alternative to traditional computation methods. It involves the design and development of computational techniques inspired by natural biota [5].

The present work has been developed as part of the *NeuChip* project, a European Union's Horizon 2020 research and innovation program granted in 2021. The main project objective is to use biological neurons to construct a “little brain”. More specifically, we aim to use human neural networks (hNNs) to produce a new paradigm for machine-learning devices that harnesses the mechanisms and rules of the human brain.

The first step towards biological computation is building *in vitro* neuronal cultures with the appropriate characteristics and studying them with data analysis techniques, with the final aim of using the produced cultures as AI systems. Many studies about *in vitro* neural cultures have been performed both using rat primary cultures derived from embryonic cortex and human iPSCs [6-8].

However, the principal issue that comes out when preparing *in vitro* neuronal cultures is that the rich repertoire of spontaneous spatiotemporal activity patterns present in the native tissue is absent. For this reason, there is an interest in inducing patterning in the connectivity of a neuronal culture. In this context, we hypothesize that it is possible to shape *in vitro* network dynamics to control spatiotemporal activity patterns and that we can mimic brain richness *in vitro* and some of its complex computations. In this work, both chemical and mechanical agents were used to shape neuronal connectivity and reproduce rich spatiotemporal activity patterns *in vitro*. Also, data analysis software was utilized for data extraction from the original calcium imaging recordings.

## 2. Methodology

For the development of neuronal cultures and the posterior analysis of the biological neuronal network, several techniques were used, including both wet lab protocols and dry lab computational tools. The neural cultures were build using both human iPSC-derived cortical neurons and rat primary cultures from embryonic cortices. The human cell line used for this work is termed long-term expandible neuroepithelial-like stem (lt-NES) cell [9]. It was modified using a lentivirus to stably express the green fluorescence calcium indicator GCaMP6s under the control of Synapsin I promoter. In experiments with primary cultures, neurons

from Sprague-Dawley rat embryonic cortices at day 14 of development were used. Rats were provided by Charles River Laboratories. In this case, GCaMP6s indicator was introduced into the cells using an Adeno-associated virus.

Cultures were built on the top of 2D topographical designs with mesoscale patterning of printed obstacles 100  $\mu\text{m}$  characteristic size on polydimethylsiloxane (PDMS). For the neuronal cultures spontaneous activity recording, a Zeiss Axiovert C25 inverted microscope equipped with a high-speed CMOS camera (Hamamatsu Orca Flash 4.1) was used. These recordings were the main data source for subsequent analysis.

Calcium imaging recordings were used as input to the first analysis tool: Netcal 8.6<sup>1</sup> software, through which activity events detection and data extraction were performed in selected regions of interest (ROIs) that were ascribed as neurons. A Schmitt trigger method was used to obtain time series of neuronal activity. Then, the outcome from Netcal was used as input in the next two tools: Visualize Fronts<sup>2</sup> used for activity analysis, and Generalized Transfer Entropy (GTE)<sup>3</sup> used for connectivity analysis. All software ran in Matlab 2018a<sup>4</sup>.

The main output from the activity analysis was the raster plot (e.g., Figure 1), which represents the activation time (x axis) for each neuron (y axis) [10]. A square moving window with predefined step and width was used to extract data from the raster plot, since static binning would give overlapping sections issues. Thus, the program provided a number of outputs. They included the global activity, which represents population activity over time, computed as the fraction of ROIs in the network that activated together; the network fraction distribution, which displays box plots with mean and distribution of global activity; and finally the inter-burst interval (IBI), which is the mean time elapsed between population activity events.

For connectivity analysis, the Generalized Transfer Entropy (GTE) was used, defined e.g. in Ref. [11]. With that, binarized vectors with a bin size of 20 ms (for 50 fps experiments) were built, with 1 indicating the presence of spike and 0 the absence of it. The main outcome of this analysis was the effective connectivity matrix, which displayed binary connections between neurons and their organization in functional modules. Modularity Q was also computed, which measures the tendency of neurons to group in assemblies that are more connected within their group than with other neurons outside the group.

Different experiments were designed and performed. Firstly, the temporal evolution of cultures on parallel lines patterned PDMS. Then, cultures under the effect of a chemical agent called bicuculline, a GABA<sub>A</sub> receptor competitive antagonist in inhibitory neurons which makes dynamics purely excitatory-driven [12]. Another chemical agent called CNQX was also used as a competitive AMPA/Kainite (non-NMDA) ionotropic glutamate receptor antagonist to reduce excitatory synaptic transmission [13]. Finally, an experiment with a physical damage through a controlled incision was performed to disrupt the system dynamics and study its robustness to external perturbations.

### 3. Results and discussion

The results of the temporal evolution experiments indicate a large increase in the number of spikes with culture maturation. This is confirmed by computing the average IBI, which decreased from 59.81 s at day *in vitro* (DIV) 7 to 12.14 s at DIV 13, which is expected since the network is supposed to be more interconnected as time evolves. Also, there is a rich repertoire of spatiotemporal activity patterns maintained over time. However, the connectivity analysis (specifically the probability distribution of connectivity angles) suggested that the directionality of connections imposed by the topographical PDMS pattern was lost with maturation. This highlights the need for more precise and durable physical constraint methods which can retain cultures connectivity in mature states. In other studies, the methods for constraining connectivity consisted of more accurate designs at a lower scale, as the one presented by Yurchenko *et al.*, who studied axonal growth on micro-patterned PDMS surfaces [14]. Finally, modularity Q decreases with maturation from 0.62 at DIV 7 to 0.42 at DIV 13, which indicates that there was an increase of functional modules size, meaning that the culture became less modular with time.

Regarding the bicuculline experiments, there is an increase in the number of neurons spiking at each timepoint from 300 s onwards in the drug raster plot displayed in Figure 1. This matches the expected effect of bicuculline, which silences the action of inhibitory synapses [15]. There is also a reduction of the repertoire of spatiotemporal activity patterns: less variety of neurons activations in bicuculline compared to control. This makes the bicuculline culture more rigid and indicates that inhibitory synapses have a key role in inducing variability and richness in neuronal cultures. Moreover, an increase in inter-burst interval from 7.40 s to 13.63 s was observed, which means that firing rate is lower in the presence of bicuculline. This might seem contradictory with the expected drug effect, but it has been demonstrated in some publications the heterogeneous effect of bicuculline on firing rate both with positive and negative changes [16]. Also, we hypothesize that this effect could be caused by phenomena like refractory period and short-term synaptic depression. The connectivity analysis (Figure 2.) shows an increase in modularity Q from 0.35 in the control case to 0.57 in the bicuculline experiment. This is reflected in the connectivity matrix with an increase in the functional modules' internal connection strengths and reduction of inter module connections (a larger presence of blank spaces outside the colored boxes). Finally, the connectivity matrix and map display changes in functional connections and synaptic strengths in the bicuculline experiments as compared to controls, thus demonstrating the feasibility of reproducing *in vitro* synaptic plasticity.

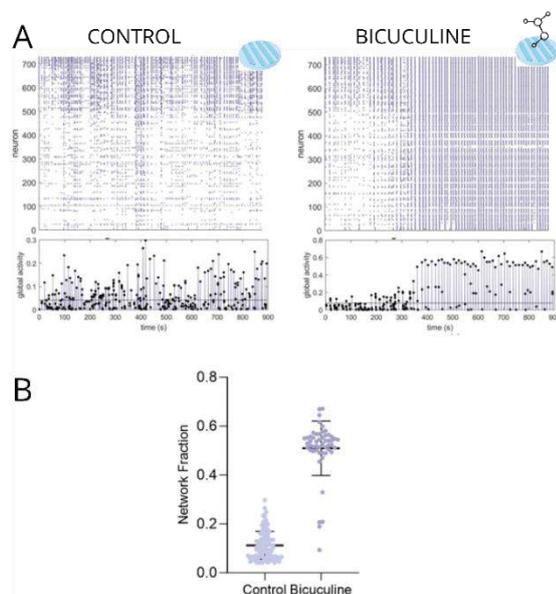
Regarding the CNQX experiment, our finding goes in line with previous literature studies. With a concentration of 1000 nM we show that the culture's activity totally ceases, because the concentration is above the previously described threshold of 600 nM [17]. For this reason, we decided to go a step further and keep the culture in the same

1: <http://www.itsnetcal.com/getting-started/>

2,3: Home-made programs from Jordi Soriano's lab

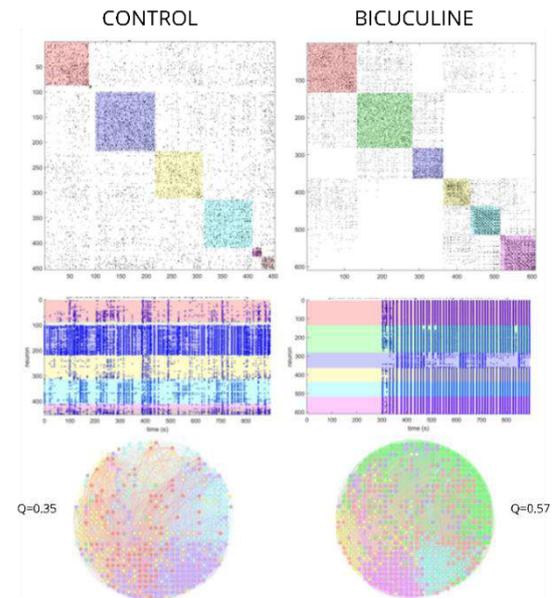
4: [https://es.mathworks.com/login?uri=%2Fdownloads%2Fweb\\_downloads](https://es.mathworks.com/login?uri=%2Fdownloads%2Fweb_downloads)

medium (with CNQX) for two days to see if it was able to partly recover the activity. With that, we demonstrated that the culture was able to fight against the chemical perturbation making efforts to go back to its initial dynamical state. It was able to activate again, though it was with a higher IBI (10.81 s compared to the 9.94 s in the control case) and a lower mean network fraction (0.1 compared to the 0.3 in control case) due to the high concentration of CNQX used. This highlights again the feasibility of reproducing *in vitro* homeostatic plasticity phenomena that are typically observed *in vivo*. We show how the network is able to rearrange and increase connections strength to cope with the chemical agent perturbation. Thus, we hypothesize that, if the culture media had been changed after the last recording to wash off CNQX and it had been kept during some more days, the firing rate would have substantially increased, since the system had already adapted by increasing the excitatory synaptic weights. This would have been the final experiment to reproduce *in vitro* a specific type of synaptic plasticity called synaptic scaling, which stands for the adjustment of excitatory synapses strengths to compensate for variations in activity [18].



**Figure 1.** (A) Raster plots (top) and population activity (bottom) in experiments with PDMS topographical lines. Events above the defined threshold of the monitored neurons (horizontal black lines) were considered significant (black dots) and shaped ‘network bursts’. (B) Distribution of bursts sizes, presented as box plots that contained all bursting events observed, comparing control and bicuculline conditions.

The mechanical perturbation experiment allowed us to demonstrate again that the culture is able to fight against a perturbation, in this case a mechanical one. The results suggest that the network fraction decreases from control (0.40) to cut (0.20) to post-cut (0.18). This is due to the physical alteration, since it breaks apart functional modules or makes them globally less connected, thus hindering the capacity of the culture to fire at unison. This is confirmed with the increase on variability and richness of activity, with a diverse repertoire of spike sizes (different number of ROIs involved).



**Figure 2.** Top, adjacency matrices of effective connectivity among neurons for control and bicuculline experiments of Figure 1. Color boxes along the diagonal of the matrices highlight functional modules, with their color intensity proportional to their strength. Center, corresponding raster plots. Bottom, corresponding network maps, where each dot is a neuron color coded according to the functional module it belongs to. The modularity  $Q$  values of the networks are indicated next to each map.

Moreover, as there are less ROIs spiking at each time point, this allows the IBI to decrease (meaning that the firing rate increases), because while some neurons are in the refractory period the others can fire. Additionally, connectivity parameters gave us more information about the network organization. Modularity  $Q$  increased in the cut experiment (0.54) with respect to the control (0.37) to then decrease again in the post cut (0.41). This illustrates a decrease in the functional modules size caused by the cut (e.g., a more modular culture), which then stabilizes with time. In the literature, S. Teller *et al.* hypothesized that the central role of the physical network after damage is to support the rerouting of dynamic interactions, reshaping information flow and functional reorganization [19]. They postulated this because their experiments were performed in few minutes time scale, and thus there was not time for structural remodeling of connectivity. However, since our experiments were performed on a longer time scale (over two days), reorganization of the culture was possible after damage, involving both structural and functional changes that compensate the lesion and remote effects of it, as demonstrated in different studies on brain function [20].

All those experiments were successfully performed using primary cultures, but when trying to replicate them with human iPSC, plating on PDMS did not work. Progenitors were successfully primed into cortical precursors but, when plating them on the top of the PDMS patterns, cell death occurred. We are not sure of the reasons why the cells were not able to survive in this last protocol step. Our main

hypothesis is that the initial number of cells for differentiation was very low (the cells plated in the final passaging) and thus the biological threshold was not enough to allow their proliferation to create a network which could give rise to neuronal activity. However, it could also have influenced that the coating was not well absorbed by the PDMS surface due to its hydrophobicity. Thus, we hypothesize that plasma-treatment of the PDMS to reduce de hydrophobicity and increasing one more day the incubation with the coating solution could solve these problems in future experiments.

#### 4. Conclusions

To conclude, this study is a first step in the biological computation working line. It provides basic findings for starting the development of a proof-of-concept *in vitro* AI computing device based on human iPSC-derived cortical neurons, which is the main objective of the *NeuChip* European project.

It has been demonstrated that it is possible to shape *in vitro* network dynamics to control spatiotemporal activity patterns. We also demonstrated that it is possible to mimic brain richness and some of its complex computations *in vitro*. This has been proved with specific findings in each experimental action.

Firstly, it was demonstrated that topographical patterns do have an effect on dictating young networks connectivity. Secondly, that inherent synaptic plasticity of the biological neuronal network is able to recover from intense external perturbations. This is one of the main differences between a conventional neural network used in current machine learning algorithms and a biological neuronal network: the conventional are not robust to huge perturbations while the biological ones are. Thirdly, that we do have the capacity to modulate synaptic strengths *in vitro*, which would be equivalent to changing the weights of a digital neural network using a learning rule. Finally, that it is totally feasible to use human iPSC-derived cortical neurons for the purpose of building *in vitro* biological HNNs.

It will be discovered in the next years whether biological computation will be the next chapter in the AI story. The advantages of this approach are clear and have been exposed during this work. Besides, it is important to highlight the interest of both academic and industry spheres on pushing up biological computation. In the academic side, the *NeuChip* European project is being funded by the European Union's Horizon 2020, which demonstrates the attractiveness of the topic. In the industry side, an example of a company working on biological computation is Cortical Labs [21], showing again the interest in the work main topic. In conclusion, biological computation is no longer seen as science fiction and experts are beginning to consider it as a key strategy for AI future.

#### References

- [1] Columbus, L. (2019). Roundup of Machine Learning Forecasts and Market Estimates for 2019. *Forbes*. <https://www.forbes.com/sites/louiscolombus/2019/03/27/roundup-of-machine-learning-forecasts-and-market-estimates-2019/>
- [2] E. Strubell, et al. (2019). Energy and Policy Considerations for Deep Learning in NLP. *Arxiv*. <https://arxiv.org/abs/1906.02243v1>
- [3] R. Schwartz, et al. (2019). Green AI. *Arxiv*. <https://arxiv.org/abs/1907.10597>
- [4] Jones, N. (2018). How to stop data centres from gobbling up the world's electricity. *Nature*, 561(7722), 163–166. <https://doi.org/10.1038/d41586-018-06610-y>
- [5] Chelly Dagdia, Z., Avdeyev, P., & Bayzid, M. S. (2021). Biological computation and computational biology: survey, challenges, and discussion. *Artificial Intelligence Review*, 54(6), 4169–4235. <https://doi.org/10.1007/s10462-020-09951-1>
- [6] Hernández-Navarro, L., Orlandi, J. G., Cerruti, B., Vives, E., & Soriano, J. (2017). Dominance of Metric Correlations in Two-Dimensional Neuronal Cultures Described through a Random Field Ising Model. *Physical Review Letters*, 118(20). <https://doi.org/10.1103/physrevlett.118.208101>
- [7] Orlandi, J. G., Soriano, J., Alvarez-Lacalle, E., Teller, S., & Casademunt, J. (2013). Noise focusing and the emergence of coherent activity in neuronal cultures. *Nature Physics*, 9(9), 582–590. <https://doi.org/10.1038/nphys2686>
- [8] Faci-Lázaro, S., Soriano, J., & Gómez-Gardeñes, J. (2019). Impact of targeted attack on the spontaneous activity in spatial and biologically-inspired neuronal networks. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 29(8), 083126. <https://doi.org/10.1063/1.5099038>
- [9] Grønning Hansen, M., Laterza, C., Palma-Tortosa, S., Kvist, G., Monni, E., Tsupykov, O., Tornero, D., Uoshima, N., Soriano, J., Bengzon, J., Martino, G., Skibo, G., Lindvall, O. & Kokaia, Z. (2020, 29 junio). Grafted human pluripotent stem cell-derived cortical neurons integrate into adult human cortical neural circuitry. *Stem Cells Translational Medicine*, 9(11), 1365-1377. <https://doi.org/10.1002/sctm.20-0134>
- [10] Chase, S. M., & Young, E. D. (2007). First-spike latency information in single neurons increases when referenced to population onset. *Proceedings of the National Academy of Sciences*, 104(12), 5175–5180. <https://doi.org/10.1073/pnas.0610368104>
- [11] Ludl, A. A., & Soriano, J. (2020). Impact of Physical Obstacles on the Structural and Effective Connectivity of *in silico* Neuronal Circuits. *Frontiers in Computational Neuroscience*, 14. <https://doi.org/10.3389/fncom.2020.00077>
- [12] Vathy, I., & ŠLamberová, R. (2009). PEPTIDES | Seizure Susceptibility is Altered by Exposure to Morphine during Mid to Late Gestation Period in the Rat. *Encyclopedia of Basic Epilepsy Research*, 1111–1116. <https://doi.org/10.1016/b978-012373961-2.00200-9>
- [13] Mitsikostas, D. D., Sanchez Del Rio, M., Waerber, C., Huang, Z., Cutrer, F. M., & Moskowitz, M. A. (1999). Non-NMDA glutamate receptors modulate capsaicin induced c-fos expression within trigeminal nucleus caudalis. *British Journal of Pharmacology*, 127(3), 623–630. <https://doi.org/10.1038/sj.bjp.0702584>
- [14] Yurchenko, I., Vensi Basso, J. M., Syrotenko, V. S., & Staii, C. (2019). Anomalous diffusion for neuronal growth on surfaces with controlled geometries. *PLOS ONE*, 14(5), e0216181. <https://doi.org/10.1371/journal.pone.0216181>
- [15] D R Curtis, A W Duggan, D Felix, G A Johnston. (1970). GABA, Bicuculline and Central Inhibition. *Nature*, 226(5252), 1222–1224. <https://doi.org/10.1038/2261222a0>
- [16] Sokal, D. M., Mason, R., & Parker, T. L. (2000). Multi-neuronal recordings reveal a differential effect of thapsigargin on bicuculline- or gabazine-induced epileptiform excitability in rat hippocampal neuronal networks. *Neuropharmacology*, 39(12), 2408–2417. [https://doi.org/10.1016/s0028-3908\(00\)00095-2](https://doi.org/10.1016/s0028-3908(00)00095-2)
- [17] Estévez-Priego, E., Teller, S., Granell, C., Arenas, A., & Soriano, J. (2020). Functional strengthening through synaptic scaling upon connectivity disruption in neuronal cultures. *Network Neuroscience*, 4(4), 1160–1180. [https://doi.org/10.1162/netn\\_a\\_00156](https://doi.org/10.1162/netn_a_00156)
- [18] Turrigiano, G. G., Leslie, K. R., Desai, N. S., Rutherford, L. C., & Nelson, S. B. (1998). Activity-dependent scaling of quantal amplitude in neocortical neurons. *Nature*, 391(6670), 892–896. <https://doi.org/10.1038/36103>
- [19] Teller, S., Estévez-Priego, E., Granell, C., Tornero, D., Andilla, J., Olarte, O. E., Loza-Alvarez, P., Arenas, A., & Soriano, J. (2019). Spontaneous Functional Recovery after Focal Damage in Neuronal Cultures. *eneuro*, 7(1), ENEURO.0254-19.2019. <https://doi.org/10.1523/eneuro.0254-19.2019>
- [20] Jiang, L., Xu, H., & Yu, C. (2013). Brain Connectivity Plasticity in the Motor Network after Ischemic Stroke. *Neural Plasticity*, 2013, 1–11. <https://doi.org/10.1155/2013/924192>
- [21] Cortical Labs. (2021). <https://corticallabs.com/>

# Design and implementation of a virtual reality training environment for health emergency technicians

B. Gutierrez-Bravo<sup>1</sup>, M. Rodríguez-Matesanz<sup>1</sup>, J. Rubio-Bolivar<sup>2</sup>, E.J. Gómez<sup>1,3</sup>,  
M. Quintana-Díaz<sup>2,4</sup>, P. Sánchez-González<sup>1,4</sup>

<sup>1</sup> Biomedical Engineering and Telemedicine Centre, ETSI Telecomunicación, Center for Biomedical Technology, Universidad Politécnica de Madrid, Madrid, Spain

<sup>2</sup> IdiPaz Instituto de Investigación Sanitaria del Hospital Universitario La Paz, Madrid, Spain

<sup>3</sup> Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina, Madrid, Spain

<sup>4</sup> Hospital Universitario La Paz, Madrid, Spain

## Abstract

The work of healthcare professionals is well known to be vital to our society. Not only do they need to master extensive knowledge and technical skills, but also to face difficult situations where they must decide how to apply their abilities under high levels of stress and pressure. However, the latter is often difficult to train before putting it into practice.

In the present work an immersive virtual reality training environment has been developed in which a health team collaborates to assist a cardiorespiratory arrest case. The team is constituted by a doctor, a nurse and two health emergency technicians and they must follow the Advanced Life Support protocol established by the European Resuscitation Council.

Additionally, the implementation of multi-user functionality allows that both a single user and several users can participate simultaneously. Finally, the integration of this Advanced Life Support simulation with a previously developed Basic Life Support scenario has been pursued, thus creating a new simulation platform where the user can choose the protocol to be performed.

## 1. Introduction

Cardio-respiratory arrest (CA) is defined as the sudden and unexpected cessation of the mechanical activity of the heart and spontaneous breathing, which stops the blood supply to vital organs. It is a situation of clinical death, susceptible to recovery if treated within the first few minutes of presentation. CA is the leading cause of premature death in Spain and surrounding countries [1]. Thus, due to its incidence, severity and mental and physical demand, the clinical case of out-of-hospital (OOH) Cardiorespiratory Arrest has been chosen for this project.

During an OOH CA, the first actions to be carried out are summarised in the so-called Chain of Survival, which the European Resuscitation Council (ERC) [2] defines as: “Early recognition and call for help”, “Early bystander cardiopulmonary resuscitation (CPR)”, “Early defibrillation in shockable rhythms”, “Early Advanced Life Support and standardised post-resuscitation care”. Once the emergency services arrive, Advanced Life Support (ALS) is performed. According to the European Resuscitation Council (ERC), the protocol to be followed is summarised in *Figure 1* [3].

## ADVANCED LIFE SUPPORT

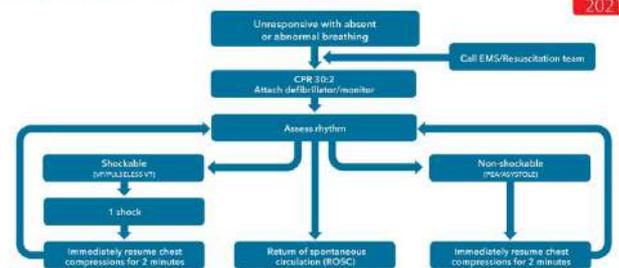


Figure 1. Advance Life Support protocol (ERC)[3]

The emergency units are made up of a medical graduate, a nursing graduate and two health emergency technicians (HET), as per Royal Decree 836/2.012, of 25 May [4] and the indications of the consulted health personnel. In this context, human factors such as teamwork, effective communication and decision making are strongly required, along with a great deal of theoretical and practical knowledge, all under an atmosphere of pressure and stress that is an additional obstacle for the healthcare worker. This is where virtual reality (VR) is key, as students can interact within a scenario that simulates a real-world scene and develop these skills before dealing with a real situation.

So much so that it is proposed by the ERC as a suitable method of training interpersonal factors [2]. Its learning recommendations are divided into three branches. First, theoretical knowledge can be acquired by reading manuals, attending seminars, or following interactive video-based films. Second, to develop practical skills, the use of high or low fidelity mannequins or attendance at practical workshops is proposed. Third, to build key interpersonal skills, such as teamwork or leadership, it is recommended to use simulations with standard patients (trained actors), interactive multimedia programmes, or other types of simulations such as virtual reality [2].

In particular, immersive virtual reality (IVR) allows the learner to have a sense of a real-life situation, which is as close as possible to the future cases they will have to deal with. IVR allows interaction with other colleagues, both simulated and real, which makes it a very strong resource for training communication skills and teamwork, as well as reinforcing theoretical knowledge. Also, if combined with

other tools, such as low-fidelity mannequins, practical skills can be trained at a low cost.

Some of the existing IVR simulators for CA are ALS-SimVR [5], or ACLS training simulator [6], but none of them allow the user to interact with real users nor with the medical equipment. Thus, a multiuser and interactive environment has been developed aiming to address some of these limitations and to provide a powerful tool for medical emergency training.

## 2. Materials and methods

This paper presents a multi-user and multi-platform IVR project for out-of-hospital ALS. For its development, Unity3D development engine has been used. Blender software has also been used to model some of the 3D objects used, such as, for example, the sanitary material.

Regarding the equipment needed to run the virtual reality application, the HCT Vive and Oculus Quest 2 solutions have been used, and since the application is cross-platform, there will be no incompatibility between the two devices.

Finally, regardless of the system used, a low fidelity mannequin is used for CPR training.

### 2.1. Characters

Different characters are involved in the simulation:

- The patient, who is suffering a CA.
- Spectators, who give realism to the simulation.
- Specialists, who represent the four members of the healthcare team. Depending on the number of users in the simulation, the healthcare workers will be either simulated characters or real users (e.g. when two users participate as the two technicians, the doctor and the nurse will be the only simulated specialists).

The tasks are distributed as follows: the doctor coordinates (through pre-recorded voice messages) all group actions and is also in charge of the airway; the nurse is responsible for managing the venous line and medication; one of the HETs assists the doctor and the other one assists the nurse (hereafter they will be called "blue" or "red" technicians).

### 2.2. Interactions

One of the main objectives of this training environment has been to offer a catalogue of interactions that favours the realism and immersive capacity of the simulation. However, a compromise must be found between the variety of controls and the learning difficulty.

The user can perform several interactions via the VR controllers:

- Picking up objects.
- Placing objects in specific areas (e.g., placing an electrode in the correct place on the victim's chest).
- Separating and joining parts of objects (e.g., separating the mask from the resuscitator).
- Stretching extendable items (e.g., stretching a cable that is in reality coiled up).

- Pressing buttons overprinted on the scene (e.g., selecting the role of a specialist)
- Interacting with a gripped object (e.g., loading or unloading defibrillation paddles).
- Teleporting (e.g., In addition to being able to walk, if there is not enough space the user can teleport)

Although there are a total of 7 interactions designed for the controller, only three buttons need to be used. For example, to pick up an object the users have to press one of the buttons, to place it they have to release that button and to separate two parts of an object they have to pick up one part with each hand and separate their hands.

One remarkable interaction is the voice recognition, which has led to a very realistic interaction with the environment. The system is able to recognise a number of predetermined phrases that are very frequently used according to the consulted professionals.

Finally, another action that the user performs is cardiopulmonary resuscitation on the physical mannequin, which adds realism and didactic value to the simulation.

### 2.3. Multi-User

One of the most ambitious functionalities of this IVR environment is the support of an online, real-time multi-user simulation. This allows two or more users to participate in the simulation simultaneously regardless of their actual location.

Photon Engine platform and its assets "PUN 2" and "Photon Voice 2", as well as the asset "VR Interaction Framework" were used to implement this functionality. Thanks to them, it has been possible to create a free cloud application that allows the connection of several users to a server. In order to have a multi-user experience it is necessary that the properties of all objects are synchronised in all instances of the application. Likewise, the events and functions need to be called in all instances.

### 2.4. Feedback

After interacting with the system, the user may receive two types of feedback, one related to the use of VR and the interactions themselves and the other one related to the quality of the executed protocol.

The first type of feedback aims to prevent the inexperience with the device from masking the user's real skills and knowledge. Both neutral sounds (e.g., indicate that a word or sentence has been recognised) and visual stimuli (e.g., the number of compressions counted by the system) have been incorporated.

Secondly, regarding the execution of the protocol, the user receives feedback during and after the simulation. During the simulation, the doctor makes corrections via pre-recorded audio if the user performs certain actions incorrectly, such as, indicating that before defibrillation it is necessary to separate from the patient if he or she has not done so. The user will also receive auditory stimuli depending on the quality of the compressions performed (e.g., if the depth and orientation are adequate).

Finally, at the end of the simulation, each user receives a results report, whose metrics will depend on the chosen specialist. The *Table 1* shows a summary of the metrics of both technicians grouped by the activity evaluated.

Metric group	#N of metrics
Cardiopulmonary resuscitation	4
Defibrillation	5
Assistance in securing the airway	4
Ventilator preparation	1
Correct placement of electrodes	1
Final post-arrest examinations	3

**Table 1.** Summary of performance report metrics

Besides the automatic feedback systems, there is also the possibility of having a supervisor as part of the simulation to check the students' performance or to suggest areas for improvement.

### 2.5. Simulation platform

Apart from developing the ALS simulation, a new simulation platform has been created that integrates the ALS scenario mentioned above and a simulation previously developed in the Biomedical Engineering and Telemicine Centre, "ECOE Station for CA on public roads", which is a Basic Life Support (BLS) [7]. Thus, the user can choose between simulating an ALS or a BLS.

## 3. Results and discussion

### 3.1. Final simulation

The result after the design and implementation of the simulation platform is as follows. First the user(s) enter the familiarisation scene (*Figure 2*), where they can interact with different objects and familiarise themselves with the interactions.



**Figure 2.** Familiarisation scene

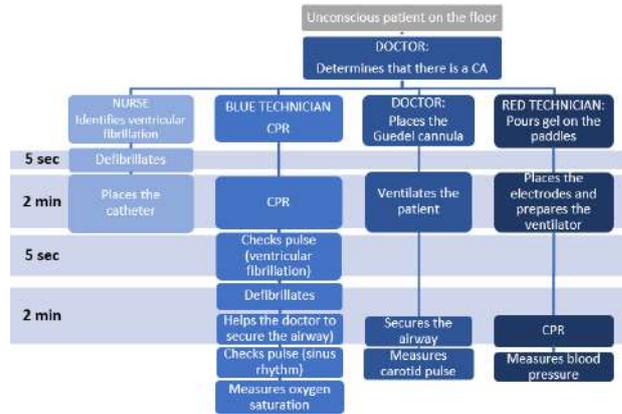
They will then be able to select from a selection panel whether one or several users are going to participate, as well as the specialist role they wish to undertake and the simulation to be carried out (ALS or BLS). Thus, the catalogue of possible scenarios is summarised in *Figure 3*.



**Figure 3.** Catalogue of possible scenarios in the environment

Once chosen, users will find themselves in the training scene where the corresponding scenario will be loaded.

The case of ALS, which is the one developed in this project, is complex from the clinical perspective. As a team of professionals is involved, the flow of the simulation needs to be carefully established. The aim is to ensure that the simulation adheres to the protocol, but at the same time also allows the user(s) to have some freedom, so that it is not a guided exercise. The established simulation flow is depicted in *Figure 4*, indicating the times within each step needs to be accomplished.



**Figure 4.** ALS simulation flow

When the scene loads, the team finds an unconscious patient on the floor and the doctor identifies the CA. He instructs the nurse to defibrillate the patient; the “red technician” (RT) to assist the nurse and then prepare the resuscitator; and the “blue technician” (BT) to start CPR. In the meantime, the doctor himself is going to place the Guedel cannula. The nurse, after placing the paddles on the patient, indicates ventricular fibrillation and defibrillates him (*Figure 5*).



**Figure 5.** Training Scene (1. One-user, 2. Multi-user)

The doctor now tells the nurse to place a catheter in the patient, while the RT connects the resuscitator to the oxygen tube. Once the resuscitator is attached, the doctor asks for the resuscitator and asks him to connect the electrodes to the patient.

Two minutes after defibrillation, the doctor indicates the RT to replace the BT with CPR and asks the BT to check the patient's pulse. Then the BT, after looking at the monitor, identifies that the patient is still in VF, so he or she picks up the paddles, tells the team to separate and defibrillates.

The doctor now informs that the airway is to be secured and that the BT should prepare the supraglottic cannula (apply lubricating gel). The doctor then asks for the cannula and tells him to prepare the ventilator to check if the cannula is correctly positioned (the capnography tube

should be attached to the ventilator and the mask needs to be separated from the rest of the ventilator).

After the doctor secures the airway, the BT must perform assisted ventilation to the patient. A capnography signal will then appear on the monitor and he or she should identify and report that it is well positioned.

Two minutes after the last defibrillation the doctor asks the BT to check the ECG signal, a sinus rhythm in this case. The doctor checks for a carotid pulse and the post-arrest care begins (the RT measures the blood pressure and the BT the oxygen saturation). After confirming that the values are appropriate, the doctor congratulates the team and the simulation ends.

### 3.2. Preliminary validation

Two surveys have been designed so that both students and health professionals can evaluate the simulation after having carried it out. The aim of these surveys is to assess whether the simulation is truly realistic and whether it meets the pedagogical objectives pursued. To this end, they contain a total of 34 questions each and are divided into different sections: (1) general information, (2) realism and immersion, (3) compression and intuitiveness, (4) learning capabilities, (5) comfort and (6) general opinion. As a first step to receive feedback on the environment, a preliminary validation has been carried out in collaboration with Puerta de Hierro hospital in Madrid. This trial was carried out with three doctors during the month of January 2022 and consisted in three users running the application in multi-user mode. After the simulation, the doctors were asked to fill in the aforementioned survey for professionals.

According to the results, the professionals reported to be very satisfied with the immersive capacity and realism of the case, the simulation and the voice instructions. They also indicated that they felt dizzy when using teleportation so it should be modified. As for the interactions, the professionals consider that there are too many, but that they are not complicated and that the familiarisation scene is adequate to learn about them. In general, they highlighted that it is a tool of adequate difficulty and that it can train skills that current methods do not.

After this positive feedback, it is necessary to collect input from more users on the positive aspects of the application and areas for improvement. Therefore, the next steps will be to carry out a technical validation, in which it will be corroborated that the metrics reflect student performance. Subsequently, it will be also necessary to carry out a clinical validation. For this, a cohort of between 25 and 30 students from the Emergency Health Technician course and between 5 and 10 instructors are being recruited.

Beyond the margin for improvement that the results of the validations might suggest, another of the future lines of action should aim to increase the variability of the simulation. Although the developed solution takes some variability into account, the system's responses to erroneous user actions could be increased, for example, if it takes too long to perform the tasks, the patient's survival could be compromised. Similarly, the ECG signal could be randomised and the simulation adapted to its values. In

addition, the number of available specialists could be increased, allowing the user to act as a doctor or nurse.

## 4. Conclusion

This work consisted in the development of an immersive virtual reality simulation of ALS for the training of health professionals, as well as the creation of a platform for simulating emergency health situations, which integrates the ALS project and a previously developed BLS project.

After a co-creation process an IVR training environment has been implemented. Realistic and simple interactions and multi-user functionality have been integrated so that another student and/or a supervisor can participate simultaneously in the activity.

Finally, the educational objective has been completed thanks to the incorporation of real-time feedback and the generation of a results report at the end of the activity, including objective performance metrics.

## References

- [1] CONSEJO ESPAÑOL DE RESUCITACIÓN CARDIO PULMONAR, “ESTATUTOS - CERCP.” <https://www.cercp.org/el-cercp/estatutos-2/> (accessed Dec. 16, 2021).
- [2] F. Semeraro *et al.*, “European Resuscitation Council Guidelines 2021: Systems saving lives Introduction and scope,” 2021, doi: 10.1016/j.resuscitation.2021.02.008.
- [3] J. Soar *et al.*, “European Resuscitation Council Guidelines 2021: Adult advanced life support,” *Resuscitation*, vol. 161, pp. 115–151, Apr. 2021, doi: 10.1016/J.RESUSCITATION.2021.02.010.
- [4] “Real Decreto 836/2012 de 25 de mayo,” *BOE núm. 137 disposición 7655*, 2012.
- [5] N. Moore, S. Yoo, N. Ahmadpour, R. Tommy, M. Brown, and P. Poronnik, “ALS-SimVR: Advanced life support virtual reality training application,” *Proceedings of the ACM Symposium on Virtual Reality Software and Technology, VRST*, Nov. 2019, doi: 10.1145/3359996.3365051.
- [6] P. Khanal *et al.*, “Collaborative virtual reality based advanced cardiac life support training simulator using virtual reality principles,” *J Biomed Inform.*, vol. 51, pp. 49–59, Oct. 2014, doi: 10.1016/J.JBI.2014.04.005.
- [7] M. Rodríguez-Matesanz, C. Guzmán-García, I. Oropesa, J. Rubio-Bolivar, M. Quintana-Díaz, and P. Sánchez-González, “A New Immersive Virtual Reality Station for Cardiopulmonary Resuscitation Objective Structured Clinical Exam Evaluation,” *Sensors*, vol. 22, no. 13, p. 4913, Jul. 2022, doi: 10.3390/S22134913/S1.

# Design, Development and Validation of an Orthosis for a Brachial Plexus Injury

M. Rollano Corroto

CEU San Pablo University, Madrid, Spain, m.rollano@usp.ceu.es

## Abstract

*Brachial plexus injury causes a break-off of the communication between the spinal cord and the arm, wrist, and hand. It affects 1-2 births out of 1000, and the current treatments do not ensure the total recovery of the full movement in the affected part. The use of active orthoses can help to carry out motor activities that the patient cannot perform without an external aid, however, most of them are expensive, heavy, or their access is very limited. In this context, a proof of concept of a simple, easy-to-use, lightweight, portable, and low cost (less than 70€) orthosis has been designed and developed for a particular case of brachial plexus injury. The validation of the orthosis was focused on three main aspects: the measurement of the range of movement for arm abduction, the registration of surface electromyography to study the response of the muscles involved in the movement, and finally, the assessment of the device through the Quebec User Evaluation of Satisfaction. The results of this proof of concept can be summarized as an increase in the abduction angle from 26° to 108°, a considerable reduction in muscular effort, and a total score of 4.5 out of 5 points in terms of usability. Nevertheless, a broader study will be necessary in order to determine if the proposed solution could be useful for more patients with different levels of brachial plexus injury.*

## 1. Introduction

The brachial plexus is a bundle of nerves which control the movement of wrists, hands, and arms. These nerves are rooted in the cervical (neck) and upper trunk (torso) sections of the spinal cord, extending through the shoulders and down the arms. An injury in the brachial plexus causes a break-off of the communication between the spinal cord and the arm, wrist, and hand. It can be partial or complete and the symptoms can vary depending on the part involved. If the injury is related to the upper part nerves of the spinal cord, the affected area is the shoulder. However, if it is rather related to the lower part nerves of the brachial plexus, the arm, wrist, and hand are the parts affected. Among others, the most common symptoms are loss of feeling, impossibility to control or move the part affected or pain in the shoulder or arm.

The causes of brachial plexus injury are diverse. On the one hand, in children, obstetric brachial plexus injury occurs in 1-2 births out of 1000. During birth, the brachial plexus is even more vulnerable. In the case that the baby is very large, when the head is stretched away from the shoulder, the brachial plexus nerves can get injured. As well, it is common in babies in breech position, where the bottom end comes first. There are two main types of brachial plexus injury: Erb's palsy and Klumke's palsy. The first one affects the upper brachial plexus nerves, the

second one, less common, affects the lower nerves. On the other hand, in adults, the main causes are trauma, athletic injuries, gunshot wounds, cancer or even radiation therapy [1].

In many cases, no surgery is needed, instead, a combination of physical and occupational therapies may be required [2]. However, for patients who have persistent neurological deficits and functional limitations, microsurgical repair and/or construction are recommended. Nonetheless, treatments do not ensure the recovery of full movement in the injured part. Because of this reason, and, with the aim of improving the daily life of these patients by supporting and enhancing their motor control in their functional activities, a series of orthoses (an orthosis is an external device attached to a part of the body that have several purposes, such as correct deformity, protect an injury, rehabilitation, reduce pain or increase mobility [3]) have been developed in recent years [4]. However, these solutions are commonly heavy, expensive, and difficult to access for the final patient.

In this context, the main purposes of this project are the design, development, and validation of a simple, easy-to-use, lightweight, portable, and low-cost mobilization orthosis for patients with brachial plexus injury. To achieve these objectives, a first approximation has been developed, as a proof of concept designed for a specific patient. This case study serves as a starting point for the solution to be subsequently generalized to other patients. Specifically, the patient (female, 31 years old) suffers an obstetric brachial plexus injury (Erb's palsy) and, as a consequence, she is not able to lift her arm. The proposed prototype has then been designed for helping in movements that imply shoulder abduction, such as touching the head or combing the hair.

## 2. Materials and methods

### 2.1. Design and development of the orthosis

As a basis for the later construction of the prototype, first, an exhaustive literature review has been carried out. From types of orthoses developed whose objectives are close to the one proposed [5-7] to materials and mechanisms that could be useful to solve the problem. Easy-to-purchase materials and basic but efficient control system were preferred so prior knowledge to start the orthosis was not required.

With the aim of achieving the objectives and requirements previously described, the orthosis was designed and developed as follows: the prototype consists

of two plastic bags reinforced by fabric bags sewn together, which will later be inflated to support the arm. The exoskeleton needed to be rigid enough to support the arm weight, but at the same time elastic so that when inflated it provided the desire shape without breaking. After several tries, the materials that satisfied both characteristics were polystyrene plastic and canvas fabric. The process followed in the construction is the one in Figure 1. Several iterations of the procedure described were needed until obtaining a feasible working prototype.

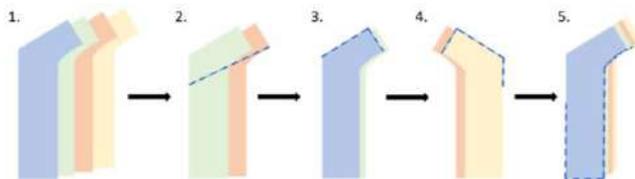


**Figure 1.** Iterative process in exoskeleton construction.

As the prototype was designed for a particular patient, all the dimensions of the device were adjusted to her. So that, the torso is 36 cm long, the arm 52 cm long and 24 cm wide.

The first five steps followed to produce the pieces and assemble the prototype were:

First of all (1), four pieces of canvas fabric were cut with the shape shown in Figure 2.1. Then (2), two of them (green and orange) were sewed together in the direction shown in Figure 2.2. This seam was made in such a way that it serves as support for the arm. Afterwards (3), a third piece of canvas was placed on top, the blue one, and sewn together with the green one as in Figure 2.3. Then (4), the three canvases were flipped (blue, green, and orange), and the last piece (yellow) was placed on top and sewn with the orange one as in Figure 2.4. In the last two steps, the different parts were sewn in order to allow the insertion of the plastic bags in subsequent steps. Finally (5), all of them were sewn together as shown in Figure 2.5. The general purpose of these canvases is to give shape to the orthosis as the plastics bags have cylindrical shape when inflated.



**Figure 2.** Steps in the exoskeleton construction.

For the following part oriented to the inflatable section, only two plastic bags and three plastic tubes were needed (see Figure 3.1).

First (1), two of the plastic tubes were inserted into the end of each of the plastic bags (Figure 3.2). Then (2), the plastic bag was heat sealed around all its sides (Figure 3.3). This step is really important in order that no air can escape through any hole. Afterwards (3), the plastic bags were inserted inside the fabric bags, one between the blue one and the green one, and the second one between the orange and the yellow one. Finally (4), the plastic tubes were connected to the compressor, however, as there are two of them, a T-junction was used to join them into a

single tube (see Figure 4). In order no air can escape through the junction they were glued to the T.



**Figure 3.** Steps in the exoskeleton construction.



**Figure 4.** T junction. The two plastic tubes are joined in a single tube using a T junction.

Finally, the completed exoskeleton was then sewn to a T-shirt that will hold to the torso and also to an elastic strap that will hold to the arm. In order to hold the air compressor, a simple belt was built that will be attached to the patient's waist.

## 2.2. Validation procedures

In order to validate that the designed prototype achieves shoulder abduction according to the initial requirements, three parameters were measured: the degree of elevation achieved (1), changes in muscle activity that take place when the prototype is used (2), and finally, the level of satisfaction of the patient (3).

(1) The range of shoulder abduction that the orthosis is able to achieve was measured. For this, a goniometer, that is an instrument that measures the available range of motion at a joint (the shoulder in this case) was used. The angle was measured three times in two different ways, one with respect to the vertical and another with respect to the horizontal.

(2) Electromyography (EMG), measures muscle response or electrical activity in reply to a nerve's stimulation of the muscle. The EMG of 5 muscles was captured: front deltoid, deltoid medial, rear deltoid, biceps, and triceps. The deltoid muscle was selected because it has a remarkable contribution in the shoulder abduction [8]. Also, biceps and triceps were used because, as stated before, the orthosis should help the patient in movements such as touching the head, where these muscles are involved. For that purpose, the mDurance system, consisting of a 4-channels, small and portable wireless sensor that integrates EMG was used. Two groups of muscles were established: shoulder muscles (front, medial and rear deltoid) and arm muscles (biceps and triceps). During validation, different exercises that require the activity of the groups of muscles previously mentioned were done, such as isometric holds in different positions, abduction movements or touches of head. Each exercise was repeated twice for each group of muscles, so that each exercise was performed four times in total. In addition, a comparison between the EMG registers of the patient and the ones from a healthy patient was also carried out.

(3) Whenever constructing an instrument, it is important to know the degree of satisfaction of the users, because of this, a survey-based validation was used to obtain

feedback so that the prototype can be improved in following versions. The selected survey was “The Quebec User Evaluation of Satisfaction with Assistive Technology (Quest 2.0)”, which is a 12-item outcome measure that assesses user satisfaction with two components, Device and Services (in this case only Device answers were considered).

### 3. Results

#### 3.1. Final prototype

As stated before, this prototype has been designed for a particular person, so that, its dimensions are adjusted to her. A first step of pattern making was necessary, where the pieces were designed according to her trunk, shoulder, and arm anatomy.

As previously mentioned, arm abduction will be achieved through the inflation of the system. This process was carried out with a portable air compressor, ADAZ, with lithium batteries of 2000 mAh in the inside. It weighs only 590 grams, is small, easy to transport and, in addition, only needs a USB cable for being charged.

It is important to highlight that, the requirement of developing a low-cost device was achieved, as its final cost is less than 70€ (canvas fabric x4: 2€/unit, plastic bag x2: 0.10€/unit, plastic tube x3: 0.10 €/unit, T-junction: 0.15 €/unit, T-shirt: 5€/unit, Elastic trap: 0.50€/unit, Air compressor: 54.15€/unit).

In the following figures (Figure 5, Figure 6, and Figure 7), the different parts of the final prototype are shown:



**Figure 5.** At the left, inflated exoskeleton sewn to T-shirt. At the right, T junction and connection to air compressor.

When inflated, the part that supports the arm looks as shown on Figure 5 left, Figure 6 and Figure 7.



**Figure 6.** View from the top of inflated exoskeleton.

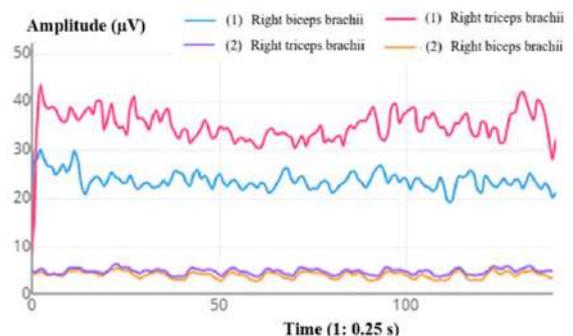


**Figure 7.** Patient wearing inflated exoskeleton

#### 3.2. Validation

(1) As stated before, one of the purposes of this prototype was to obtain a wider abduction angle. The difference in angle when lifting the arm of the patient to the maximal abduction position with and without the orthosis is remarkable; without the exoskeleton the patient is able to elevate the arm 26° respect the vertical and, with the help of the orthosis, this angle increases to 108°, what means that it provides approximately an extra angle of 82°.

(2) As mentioned previously, the activity of 5 muscles was captured in different exercises. The graphs selected to study the muscle activity were the ones which presented less movement artifacts. Two of the most relevant registers are shown below.

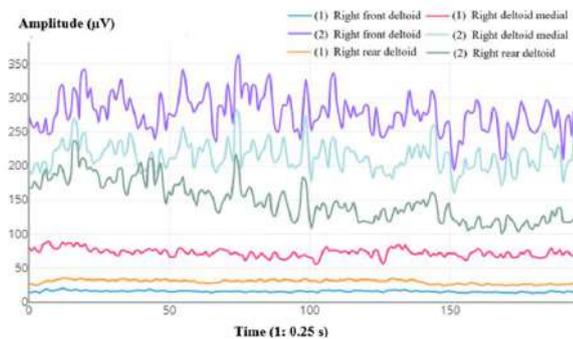


**Figure 8.** EMG in maximum abduction without orthosis (biceps and triceps) vs. with orthosis (biceps and triceps).

Figure 8 presents a comparison between the EMGs of the patient suffering the brachial plexus injury trying to reach her head with and without the orthosis. From the

observation of the graph, it can be stated that when the patient is wearing the device and, subsequently, the shoulder abduction is supported by the orthosis, a much lower muscular activity is registered.

In addition, the difference in the muscular activity between a patient with a brachial plexus injury wearing the orthosis and a healthy person without wearing it, is presented. As shown in Figure 9, it can be stated that when reaching the maximum abduction, the muscular activity of the patient (focusing on deltoid muscles) is much lower (thanks to the support of the orthosis). In addition, it is interesting to highlight that in this case, the muscle with a higher activity in the patient is the medial deltoid whereas in the healthy person, the muscle activity is greater in the front deltoid.



**Figure 9.** EMG in maximum abduction in patient with paralysis (front, medial and rear deltoid) vs. healthy patient (front, medial and rear deltoid).

(3) The results of the QUEST 2.0 questionnaire suggest that the most important items for the patient are weight, easy to use and comfort. In general, she is quite satisfied with the assistive device. The overall score is of 4.5 points. From the comments registered during the user's evaluation of satisfaction, it is important to highlight that adjustments in the fixation of the orthosis to the body are necessary, for example, in upper chest. Although the T-shirt adjusts the prototype to the torso, an additional point of union must be included.

#### 4. Conclusions and Future Work

Brachial plexus injury affects 1-2 births out of 1000, causing loss of feeling, and a lack in the control and movement of the part effected, as well as pain in the shoulder or arm. This project presents the design, development, and validation of a simple, easy-to-use, lightweight, and low-cost (less than 70€) proof of concept of an orthosis as an assistive technology for abduction movements. The device is affordable in comparison to the prototypes currently available in the market. Also, its high portability gives the patient the opportunity of using it anywhere in daily life activities. One of the main advantages of the proposed solution is that it leaves remaining degrees of freedom unhindered, allowing users to perform a wide range of tasks (i.e.: although it controls and assists the shoulder abduction movement, the elbow can move freely so that the arm can reach the head).

From the results obtained, it can be stated that the prototype increases considerably the range of shoulder angle with respect to the vertical. Also, when using the orthosis, the maximum abduction is reached with low levels of EMG activity in the involved muscles, which suggests that the designed prototype reduces muscular effort and efficiently supports the shoulder abduction. In addition, it can be concluded that muscle distribution in patients with brachial paralysis seems to be different to the one of a healthy person, although further research is needed to properly model this effect.

In addition, the results shows that the patient was quite satisfied with the tested solution. However, the suggestions of the patient regarding the fixation of the orthosis to the body should be considered in the following prototypes.

On the other hand, it is important to highlight that the presented prototype is only a proof of concept and, therefore, there is still plenty of room for improvement. Thus, for example, it would be interesting to analyze the behavior of all muscles involved in the abduction movement. Moreover, the identification of other disorders that could take advantage of this device would also be an interesting future research line. Moreover, as the proposed prototype has been designed not only for a particular case of brachial plexus injury but also for a particular person, further work should be carried out to study how it works with other patients with different body dimensions and characteristics to see which modifications could be done in order to make it adjustable to all individuals.

#### References

- [1] Web page of Johns Hopkins medicine. <https://www.hopkinsmedicine.org/health/conditions-and-diseases/brachial-plexus-injuries>. (Accessed: February 2022).
- [2] Web page of NYU Langone Health. <https://nyulangone.org/conditions/brachial-plexus-injuries/treatments/nonsurgical-treatment-for-brachial-plexus-injuries> (Accessed: October 2022).
- [3] B. J. Wilkinson Orthotics and Prosthetics in Rehabilitation (Fourth Edition), Orthoses in the Management of hand Dysfunction.
- [4] Condie DN. The modern era of orthotics. *Prosthetics and Orthotics International*, vol 32, issue 3, 2008, pp 313-323 (DOI: 10.1080/03093640802113006).
- [5] Sabin Copaci D, Flores Caballero A, Blanco Rojas MD, Moreno Lorente LE. Shoulder Exoskeleton for rehabilitation actuated with shape memory alloy. UC3M, 2016. *Conference on Future Trends in Robotics*, 2016 (ISBN 978-84-608-8452-1).
- [6] Li X, Li W, Li Q. Method, Design and Evaluation of an Exoskeleton for Lifting a Load In Situ. *Applied Bionics and Biomechanics*, 2021 (DOI: 10.1155/2021/5513013).
- [7] Simpson CS, Okamura A, Hawkes E. Exomuscle: An inflatable device for shoulder abduction support. *International Conference on Robotics and Automation (ICRA)*, Singapore, 2017, pp. 6651-6657 (DOI: 10.1109/ICRA.2017.7989785).
- [8] Lam J.H, Bordoni B. Anatomy, Shoulder and Upper Limb, Arm Abductor Muscles. StatPearls Publishing, 2021 (ISBN: NBK537148).

# Determinación de requisitos biomecánicos y funcionales en población sana con el software RehabHand

C. Sánchez del Valle<sup>1</sup>, M. Álvarez Rodríguez<sup>2</sup>, S. Ceruelo Abajo<sup>3</sup>, A. Gil Agudo<sup>4</sup>, C. Talayero<sup>5</sup>, A. de los Reyes Guzmán<sup>6</sup>

<sup>1</sup> Escuela de Arquitectura, Ingeniería y Diseño; Universidad Europea de Madrid; Madrid; España; clarasanchezdelvalle@gmail.com

<sup>2</sup> Unidad de Biomecánica y Ayudas Técnicas, Hospital Nacional de Paraplégicos (SESCAM), Toledo, España, maria.alvarez.rodriguez93@gmail.com

<sup>3</sup> Servicio de Rehabilitación, Hospital Nacional de Paraplégicos (SESCAM), Toledo, España, sceruelo@sescam.jccm.es

<sup>4</sup> Servicio de Rehabilitación, Hospital Nacional de Paraplégicos (SESCAM), Toledo, España, amgila@sescam.jccm.es

<sup>5</sup> Escuela de Arquitectura, Ingeniería y Diseño; Universidad Europea de Madrid; Madrid; España; carlosalberto.talayero@universidadeuropea.es

<sup>6</sup> Unidad de Biomecánica y Ayudas Técnicas, Hospital Nacional de Paraplégicos (SESCAM), Toledo, España, adlos@sescam.jccm.es

## Resumen

*RehabHand es una plataforma que integra diferentes serious games orientados a la rehabilitación del miembro superior en pacientes con lesión medular cervical. Entre ellos, se encuentra la versión virtualizada de la escala clínica Box and Block y el Explorador, que consiste en el seguimiento de una trayectoria en el aire con el dedo. RehabHand utiliza el Leap Motion Controller para captar y realizar un seguimiento del movimiento de las manos. En este trabajo se recogen y analizan datos de la ejecución de estos serious games (Block and Block virtual y Explorador) con población sana. En los resultados se ve una correlación entre la puntuación obtenida en la escala Box and Block original y la virtual, se elaboran unos patrones de normalidad o referencia y se estudian las diferencias en el rendimiento en varios equipos con distintas prestaciones.*

## 1. Introducción

En pacientes con tetraplejía, la rehabilitación del miembro superior (MS) es de vital importancia para lograr una mayor independencia. La incorporación de tecnologías utilizadas de manera complementaria a la terapia convencional logra intensificar la rehabilitación del paciente y motivarle a la realización de los ejercicios. RehabHand es una plataforma que integra diferentes *serious games* (SG) orientados a la rehabilitación del MS en pacientes con lesión medular cervical [1],[2].

RehabHand consiste en una aplicación de escritorio desde la que se puede acceder a los diferentes SG, cada uno orientado a unos patrones de movimiento concretos. El dispositivo utilizado para capturar el movimiento es el *Leap Motion Controller* (LMC). Con este sistema pueden obtenerse datos reales del desempeño de los pacientes. El LMC ya ha sido empleado anteriormente en proyectos de rehabilitación y SG con pacientes de Parkinson [3] o ictus [4], pero no en rehabilitación de lesión medular cervical.

Un patrón de referencia se utiliza para comparar los datos de un paciente con los de una persona sana y así saber cuánto dista el paciente de una función normal. Los datos biomecánicos y funcionales registrados con RehabHand pueden usarse para obtener valoraciones objetivas del MS, pero para ello es necesario conocer los registros obtenidos con personas sanas y así crear unos patrones de referencia.

## 2. Objetivos del proyecto

El objetivo general del proyecto es laborar un patrón de normalidad para RehabHand partir de los resultados biomecánicos y funcionales de población con una función del MS sana. Los objetivos específicos son:

- Elaborar una escala que permita evaluar de forma objetiva la capacidad funcional del MS en pacientes con lesión medular cervical.
- Analizar cómo influyen las prestaciones del ordenador en los resultados medidos. Para ello, se ejecutarán las actividades en tres ordenadores con distintas prestaciones.
- Analizar la correlación entre los resultados medidos y una escala clínica que se utilice actualmente en la clínica.

## 3. Metodología

### 3.1. Materiales y equipamiento

#### 3.1.1 Leap Motion Controller (LMC)

El LMC es un módulo óptico que captura el movimiento de las manos sin necesidad de contacto. Su funcionamiento se basa en tres LEDs infrarrojos, que emiten a una longitud de onda de 850 nm, y dos cámaras que detectan la luz emitida por los LEDs que es reflejada por las manos. Las cámaras son capaces de trabajar hasta a 200 fps, dependiendo del equipo que se esté usando [5].

#### 3.1.2 Características de los ordenadores utilizados

En las siguientes tablas se muestran las especificaciones técnicas de los equipos utilizados.

<b>Procesador</b>	Intel® Core™ i5-8500 CPU 3 GHz
<b>Sistema operativo</b>	Windows 10 Pro
<b>Tarjeta gráfica</b>	NVIDIA® GeForce® GT 1030
<b>Memoria RAM</b>	8 GB

*Tabla 1. Especificaciones técnicas del equipo 1.*

<b>Procesador</b>	Intel® Core™ i7-7500 CPU 2.7 GHz
<b>Sistema operativo</b>	Windows 10 Home
<b>Tarjeta gráfica</b>	NVIDIA® GeForce® GT 920MX
<b>Memoria RAM</b>	12 GB

Tabla 2. Especificaciones técnicas del equipo 2.

<b>Procesador</b>	Intel® Core™ i7-10750 CPU 2.6 GHz
<b>Sistema operativo</b>	Windows 10 Pro
<b>Tarjeta gráfica</b>	NVIDIA® Quadro T2000 con Max-Q Design
<b>Memoria RAM</b>	32 GB

Tabla 3. Especificaciones técnicas del equipo 3.

### 3.2. Diseño de los patrones de referencia

#### 3.2.1 Aplicaciones virtuales escogidas e índices medidos

##### Box and Block (BBT)

El test Box and Block es la herramienta de valoración de la destreza motora manual más estandarizada. Consiste en una caja dividida en dos compartimentos, uno de ellos con 100 bloques de forma cúbica (lado 2.5 cm). El paciente debe intentar pasar, de uno en uno, el máximo número de bloques de un compartimento a otro en sesenta segundos. Antes de comenzar la prueba, se otorgan quince segundos de práctica.

RehabHand cuenta con la versión virtualizada de esta escala clínica, desarrollada según las directrices del artículo general [6].



Figura 1. Test Box and Block

El índice medido fue el número de cubos que el sujeto es capaz de pasar de un lado al otro.

##### Explorador (EXP)

El juego consiste en seguir con el dedo o bolígrafo una trayectoria prefijada. La trayectoria se hace en el aire y con ello el objetivo es valorar la coordinación global del miembro superior y precisión y suavidad en el extremo distal.

Los índices medidos fueron:

- La longitud de la trayectoria recorrida por el dedo o bolígrafo, hallada como la suma de las distancias euclídeas de la posición de un frame al siguiente.
- El número de picos en el perfil de velocidad.

- El número optimizado de picos del perfil velocidad, que corresponde con aquellos picos cuya prominencia supere el 10% de la velocidad máxima.

### 3.3. Diseño del estudio

En este estudio participaron 16 sujetos entre 22 y 57 años ( $\bar{x} = 31,12$ ). Todas las pruebas se realizaron en las mismas condiciones de luz ambiental y en la misma posición (sedestación, con el ordenador y el LMC centrados frente al sujeto). El orden de ejecución de los equipos se determinó de forma aleatoria [7] y en todos los SG se otorgó un período de adaptación al entorno. Además, todos los sujetos realizaron el test real de Box and Block (Figura 1) a fin de comparar los resultados de la escala real y la virtual. Los participantes realizaron todas las pruebas con su mano dominante y antes de comenzar se les realizó la siguiente pregunta: “¿Estás habituado al uso de entornos virtuales, por ejemplo, videojuegos?”. De esta forma se agrupó a los participantes en dos grupos: 8 sujetos familiarizados con la interacción con diferentes entornos virtuales y 8 que no. En ambos grupos 7 sujetos eran diestros y 1 zurdo.

## 4. Resultados

### 4.1. Datos recogidos

#### 4.1.1 Box and Block

	Puntuación (uds.) Media(DE)	Coefficiente de Pearson con la escala clínica
<b>Equipo 1</b>	19,69(8,83)	0,76
<b>Equipo 2</b>	<b>19,00(7,10)<sup>a</sup></b>	0,53
<b>Equipo 3</b>	<b>22,06(8,68)<sup>a</sup></b>	0,66
<b>Escala clínica</b>	65,56(12,82)	

<sup>a</sup> p<0,05

Tabla 4. Resultados en BBT.

En las siguientes tablas se muestran los mismos datos en función del grupo al que pertenecen los participantes.

	Puntuación (uds.) Media(DE)	Coefficiente de Pearson con la escala clínica
<b>Equipo 1</b>	20,63(9,94)	0,92
<b>Equipo 2</b>	20,25(8,24)	0,63
<b>Equipo 3</b>	22,38(9,41)	0,92
<b>Escala clínica</b>	65,25(15,4)	

Tabla 5. Resultados en BBT del grupo familiarizado con los entornos virtuales.

	Puntuación (uds.) Media(DE)	Coefficiente de Pearson con la escala clínica
<b>Equipo 1</b>	18,75(8,14)	0,49

<b>Equipo 2</b>	17,75(6,04)	0,49
<b>Equipo 3</b>	21,75(8,51)	0,25
<b>Escala clínica</b>	65,88(10,71)	

Tabla 6. Resultados en BBT del grupo no familiarizado con los entornos virtuales.

4.1.2 Explorador

Nivel 1

	Trayectoria (cm) Media(DE)	Nº picos vel. (uds.) Media(DE)	Nº picos opt. (uds.) Media(DE)
<b>Equipo 1</b>	143,00(55,53)	161,00(84,22)	27,00(10,01)
<b>Equipo 2</b>	122,00(17,00) <sup>a</sup>	128,38(45,00) <sup>a</sup>	25,50(11,21)
<b>Equipo 3</b>	145,00(37,00) <sup>a</sup>	187,50(122,00) <sup>a</sup>	31,13(21,75)

Nivel 2

	Trayectoria (cm) Media(DE)	Nº picos vel. (uds.) Media(DE)	Nº picos opt. (uds.) Media(DE)
<b>Equipo 1</b>	113,00(14,00)	134,38(45,00)	24,19(9,84)
<b>Equipo 2</b>	106,00(15,58)	107,44(54,00) <sup>a</sup>	22,25(8,70)
<b>Equipo 3</b>	123,00(45,60)	146,56(73,00) <sup>a</sup>	28,13(13,91)

Nivel 3

	Trayectoria (cm) Media(DE)	Nº picos vel. (uds.) Media(DE)	Nº picos opt. (uds.) Media(DE)
<b>Equipo 1</b>	105,00(14,39)	136,00(41) <sup>a</sup>	22,56(8,85)
<b>Equipo 2</b>	97,00(12,20) <sup>a</sup>	103,31(40) <sup>a,b</sup>	20,88(10,62) <sup>a</sup>
<b>Equipo 3</b>	121,00(39) <sup>a</sup>	130,88(46) <sup>b</sup>	28,13(13,05) <sup>a</sup>

Nivel 4

	Trayectoria (cm) Media(DE)	Nº picos vel. (uds.) Media(DE)	Nº picos opt. (uds.) Media(DE)
<b>Equipo 1</b>	95,00(20,39)	124,94(35,00) <sup>a</sup>	20,00(8,67)
<b>Equipo 2</b>	86,00(10,17) <sup>a</sup>	95,81(36,00) <sup>a,b</sup>	20,94(8,32)
<b>Equipo 3</b>	95,00(11,44) <sup>a</sup>	122,44(57,00) <sup>b</sup>	25,63(22,63)

<sup>a, b</sup> p<0,05

Tabla 7. Resultados en EXP.

4.2. Correlación entre la puntuación en la aplicación virtual BBT y la escala clínica

En las Tablas 5 y 6 se observan las diferencias en la correlación con la escala clínica entre el grupo familiarizado y no familiarizado. A continuación, se muestran las gráficas de dispersión del equipo 1.

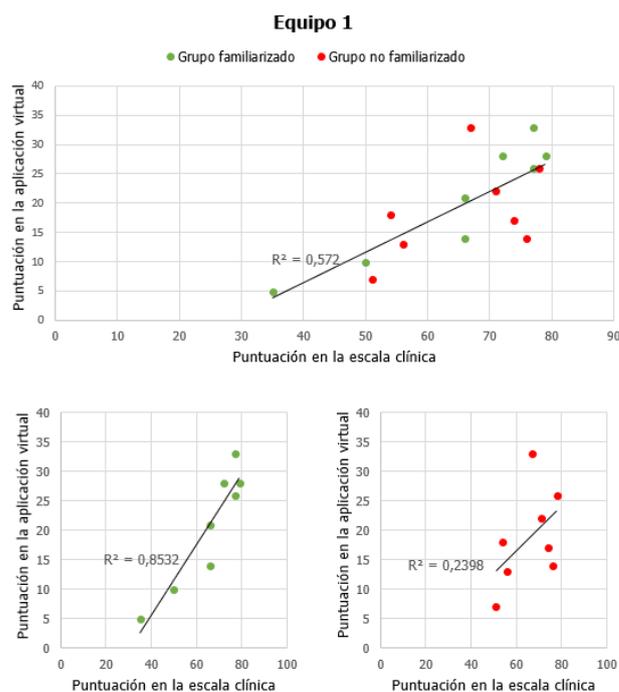


Figura 2. Gráficas de dispersión en BBT en el equipo 1.

4.3. Determinación de patrones de referencia en población sana

4.3.1 Patrones de referencia para Box and Block

Para la elaboración de los patrones de referencia se han diseñado dos modelos alternativos. Ambos modelos puntúan el desempeño del paciente entre 0 y 100, siendo 100 la puntuación correspondiente a la media obtenida en población sana. La variable independiente de las siguientes ecuaciones representa el número de cubos que consiguieron pasarse de un compartimento a otro.

En el primer caso, la función correspondiente al modelo estaría representada por la siguiente ecuación:

$$f(x) = \begin{cases} \frac{100}{\bar{x}}x & : 0 \leq x \leq \bar{x} \\ 100 & : x > \bar{x} \end{cases}$$

Siendo  $\bar{x}$  la media de cubos obtenida en población sana.

El segundo caso surge de un problema en el primero, ya que no tiene en cuenta la desviación estándar. Por ejemplo, véase el caso del equipo 1. La media de la población sana es 19,69 y la desviación estándar 8,83. Una persona sana podría conseguir pasar alrededor de 10,86 cubos ( $\bar{x}$ -DE). Sin embargo, según la función anterior esto supondría una puntuación de 55,15 sobre 100. Así pues, se propone una curva que pase por los puntos [0,0], [ $\bar{x}$ ,100] y [ $\bar{x}$ -DE,80].

A continuación, se muestra el ejemplo del equipo 1.

Equipo 1

- ♦ Función lineal

$$f(x) = \begin{cases} \frac{100}{19,69}x & : 0 \leq x \leq 100 \\ 100 & : x > 100 \end{cases}$$

♦ Curva

$$f(x) = \begin{cases} 10.1802x - 0.25909x^2 & : 0 \leq x \leq 100 \\ 100 & : x > 100 \end{cases}$$

**4.3.2 Patrones de referencia para Explorador**

Para las siguientes funciones  $f(x, y)$  las variables escogidas son la longitud de la trayectoria ( $x$ ) y el número de picos optimizado de la velocidad ( $y$ ). A diferencia de las escalas anteriores, estas puntúan entre 10 y 100. El motivo es que una puntuación de 0 resulta muy desmotivadora para el paciente. Esta puntuación es inevitable en BBT ya que cabe la posibilidad de que el paciente no consiga pasar ningún cubo. Sin embargo, sí que es evitable en este caso, dado que el paciente no es consciente de cómo es su rendimiento dentro de la aplicación.

Para elaborar estas escalas se tuvieron en cuenta ambas variables por separado. Debido a que son variables inversamente proporcionales a la puntuación que se desea asignar, se estableció la nota mínima (10) en cuatro veces la media de la población sana para cada una de las variables, y la máxima (100) en la media. De esta forma, se obtuvieron dos ecuaciones diferentes que puntuaban entre 10 y 100 y, para conseguir una ecuación única, se realizó la media de ambas. Así pues, la ecuación de esa función sería la siguiente:

$$h(f(x), g(y)) = \begin{cases} 100 & : x < \bar{x} \cup y < \bar{y} \\ \frac{100 + g(y)}{2} & : x < \bar{x} \cup \bar{y} \leq y \leq 4\bar{y} \\ \frac{f(x) + 100}{2} & : \bar{x} \leq x \leq 4\bar{x} \cup y < \bar{y} \\ \frac{f(x) + g(y)}{2} & : \bar{x} \leq x \leq 4\bar{x} \cup \bar{y} \leq y \leq 4\bar{y} \\ \frac{10 + g(y)}{2} & : x > 4\bar{x} \cup \bar{y} \leq y \leq 4\bar{y} \\ \frac{f(x) + 10}{2} & : \bar{x} \leq x \leq 4\bar{x} \cup y > 4\bar{y} \\ 10 & : x > 4\bar{x} \cup y > 4\bar{y} \end{cases}$$

Siendo  $\bar{x}$  la media de la trayectoria obtenida en población sana,  $\bar{y}$  la media del número de picos optimizado,  $f(x)$  la recta que pasa por los puntos  $[\bar{x}, 100]$  y  $[4\bar{x}, 10]$ , y  $g(y)$  la recta que pasa por los puntos  $[\bar{y}, 100]$  y  $[4\bar{y}, 10]$ .

**5. Discusión**

En este trabajo se observan las diferencias entre trabajar en el entorno real y hacerlo en el entorno virtual. En BBT, la puntuación obtenida por una misma persona en el entorno virtual es inferior a la escala clínica, aunque se esté realizando el mismo patrón de movimiento. Mientras que en el mundo físico una persona sana es capaz de realizar el movimiento prácticamente de forma automática, en el entorno virtual no se conoce la posición real del cubo, sino que a partir de lo que se muestra en la pantalla el sujeto debe modular el movimiento de su mano para alcanzar la posición del cubo. Mientras que en el mundo físico el sujeto sabe de primeras el lugar al que tiene que llevar su mano, en el entorno virtual únicamente puede saberlo por aproximaciones.

Si bien es cierto que se ha encontrado una correlación lineal entre los resultados en la aplicación virtual y el entorno

real, esta solo se da en el grupo familiarizado con los entornos virtuales. Teniendo esto en cuenta, sería recomendable disponer de un período de adecuación al entorno con los pacientes previamente a la realización de la prueba. Además, se considera que el tiempo proporcionado de adaptación no fue suficiente para la mayoría de sujetos.

Para la muestra estudiada, se han encontrado diferencias estadísticamente significativas entre los equipos utilizados. Por este motivo, el patrón de referencia diseñado para uno de ellos no podría utilizarse de forma universal.

**6. Conclusiones**

A partir de una muestra de 16 personas sanas, se han elaborado patrones de normalidad para los tres equipos utilizados y se ha observado la correlación entre trabajar en el entorno real y el virtual. El siguiente paso debe ser la validación de estos patrones en pacientes con lesión medular cervical, ya que aún no han sido probados con datos de pacientes y no se tiene información sobre su fiabilidad y repetibilidad.

Una posible línea de investigación consistiría en elaborar un patrón universal capaz de adaptarse a las prestaciones del ordenador que se vaya a utilizar.

**Agradecimientos**

Esta contribución se ha realizado en la Unidad de Biomecánica del Hospital de Paraplégicos de Toledo en el contexto del proyecto REHAB-IMMERSIVE, financiado por Ministerio de Ciencia e Innovación/Agencia Estatal de Investigación/10.13039/501100011033, proyecto REHAB-IMMERSIVE ref. PID2020-117361RB-C22.

**Referencias**

- [1] ÁLVAREZ RODRÍGUEZ, María. RehabHand: Juegos serios basados en el controlador Leap Motion para la rehabilitación del miembro superior en pacientes con lesión medular cervical. *Universidad de Castilla-La Mancha*. 2018.
- [2] DE LOS REYES-GUZMÁN, Ana et al. RehabHand: Oriented-tasks serious games for upper limb rehabilitation by using Leap Motion Controller and target population in spinal cord injury. *NeuroRehabilitation*, vol. 48 (abril de 2021), n.º 3, pp. 365–373. ISSN 1878-6448.
- [3] FERNÁNDEZ-GONZÁLEZ, Pilar et al. Leap motion controlled video game-based therapy for upper limb rehabilitation in patients with Parkinson’s disease: a feasibility study. En línea. *Journal of NeuroEngineering and Rehabilitation*, vol. 16 (noviembre de 2019), n.º 1. ISSN 1743-0003. [consultado el 18/02/2022].
- [4] IOSA, Marco et al. Leap motion controlled videogame-based therapy for rehabilitation of elderly patients with subacute stroke: a feasibility pilot study. En línea. *Topics in Stroke Rehabilitation*, vol. 22 (febrero de 2015), n.º 4, pp. 306–316. ISSN 1945-5119. [consultado el 18/02/2022].
- [5] MANTECÓN, Tomás et al. Hand Gesture Recognition Using Infrared Imagery Provided by Leap Motion Controller. En línea. *Springer*, vol. 10016 (2016), pp. 47–57. [consultado el 04/03/2022].
- [6] MATHIOWETZ, V. et al. Adult Norms for the Box and Block Test of Manual Dexterity. En línea. *American Journal of Occupational Therapy*, vol. 39 (junio de 1985), n.º 6, pp. 386–391. ISSN 0272-9490. [consultado el 10/03/2022].
- [7] Research Randomizer. En línea. *Randomizer.org*. [2007]. [consultado el 11/03/2022].

# Análisis multicanal de sonidos respiratorios mediante entropía muestral fija para la evaluación de la COPD

F. Aguilar Paredes<sup>1,2</sup>, M. Lozano-García<sup>1,2,3</sup>, R. Jané Campos<sup>1,2,3</sup>

<sup>1</sup> Universitat Politècnica de Catalunya (UPC)-Barcelona Tech. Escuela de Ingeniería de Barcelona Este (EEBE), Barcelona, España

<sup>2</sup> Institute for Bioengineering of Catalonia (IBEC), The Barcelona Institute of Science and Technology (BIST), Barcelona, España

<sup>3</sup> Centro de Investigación Biomédica en Red, Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), Instituto de Salud Carlos III, Madrid, España

## Resumen

*La enfermedad pulmonar obstructiva crónica (COPD por sus siglas en inglés) ha sido catalogada por la Organización Mundial de la Salud como la tercera causa de muerte a nivel mundial, siendo cada vez más las personas que padecen esta enfermedad. No obstante, la heterogeneidad clínica y la baja especificidad de algunos de los test diagnósticos utilizados en la actualidad para la COPD, dificultan a menudo el diagnóstico y seguimiento de esta enfermedad. El objetivo de este trabajo es explorar el uso potencial de las señales de sonidos respiratorios (RS) para derivar nuevos biomarcadores fisiológicos no invasivos para el diagnóstico y monitorización de la COPD.*

*Se han analizado las señales de flujo y RS de un total de 27 sujetos (15 controles y 12 COPD), registradas durante la realización de un protocolo de flujo inspiratorio variable. Se ha observado que los parámetros extraídos de los RS, mediante la entropía muestral fija, siguen un patrón similar al del flujo y volumen respiratorios y son sensibles al deterioro de la mecánica respiratoria en pacientes con COPD, por lo que podrían utilizarse para derivar nuevos biomarcadores fisiológicos no invasivos para la evaluación de estos pacientes en la práctica clínica.*

## 1. Introducción

La enfermedad pulmonar obstructiva crónica (COPD por sus siglas en inglés) es una enfermedad común, prevenible y tratable que se caracteriza por síntomas respiratorios persistentes y limitación del flujo respiratorio [1]. Dicha limitación proviene de la inflamación crónica de las vías respiratorias y/o la destrucción del parénquima pulmonar, lo que genera una obstrucción de las vías respiratorias de quien padezca la enfermedad.

Según la Organización Mundial de la Salud, la COPD es una de las principales causas de morbilidad y mortalidad a nivel mundial, induciendo un impacto económico y social que aumenta con el transcurso de los años. En el año 2019, esta enfermedad fue la causante de 3,23 millones de muertes a nivel mundial, colocándola como tercera causa de defunciones [2].

Son diversos los factores de riesgo que existen en la COPD, aunque el tabaquismo y la exposición a partículas contaminantes del aire se han considerado los factores de riesgo principales. A nivel sintomatológico, existe una gran variabilidad entre sujetos y periodos de tiempo, lo que hace

que tanto el diagnóstico como el seguimiento de la enfermedad sean complejos. El diagnóstico de la COPD está basado en la presencia de síntomas y en los resultados de pruebas espirométricas para determinar la limitación del flujo respiratorio, a partir de la ratio entre el volumen espiratorio forzado en el primer segundo (FEV<sub>1</sub>) y la capacidad vital forzada (FVC). Aunque la espirometría es la prueba más reproducible y objetiva para la determinación de la limitación del flujo de aire, no es una prueba que por sí sola sea suficiente para establecer un diagnóstico, dado que consta de baja especificidad. Este hecho suele llevar al infradiagnóstico de la COPD, lo que ha derivado en la investigación de nuevos biomarcadores que contribuyan a mejorar el diagnóstico y seguimiento de la enfermedad, como los biomarcadores fisiológicos derivados de la actividad muscular respiratoria [3]. Es por ello por lo que el objetivo de este estudio es explorar el uso potencial de las señales de sonidos respiratorios (RS), en comparación al flujo y volumen respiratorios, para derivar nuevos biomarcadores fisiológicos no invasivos que puedan ser integrados en el diagnóstico y monitorización de la COPD [4].

## 2. Métodos

### 2.1. Base de datos y protocolo

La base de datos utilizada para este estudio está formada por 27 sujetos: 15 controles sanos y 12 pacientes con COPD. El grupo de pacientes se encuentra subdividido en 7 pacientes leves o moderados (COPD1) y 5 pacientes severos o muy severos (COPD2).

Se registraron las señales de flujo y RS en cada uno de los participantes durante la realización de un protocolo de flujo inspiratorio variable y utilizando cuatro micrófonos piezoeléctricos de contacto (TSD 108, Biopac Systems, Inc.), tal y como se muestra en la Figura 1. En primer lugar, se registraron las señales durante 3 minutos de respiración normal (nivel de flujo basal). A continuación, se realizaron una serie de respiraciones máximas para registrar las señales durante la obtención del flujo inspiratorio máximo (MIF por siglas en inglés). A partir de estos dos valores se definieron tres niveles de flujo intermedio para cada participante y se le pidió realizar una serie de respiraciones manteniendo un nivel de flujo inspiratorio alrededor de los

tres niveles de flujo intermedio calculados, mediante *feedback* visual con la propia señal de flujo respiratorio del participante [5]. De esta manera se registraron las señales durante 5 niveles de flujo diferentes que se utilizaron posteriormente para el análisis de los datos.

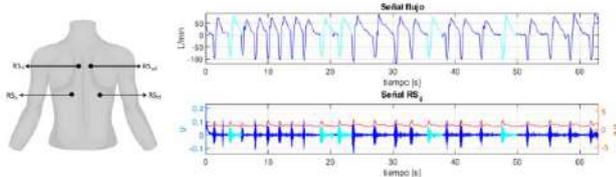
## 2.2. Preprocesamiento

El procesamiento de las señales se ha desarrollado en el entorno MATLAB.

Dado que las señales de RS son muy sensibles a múltiples fuentes de ruido, se establece un primer bloque de preprocesamiento de las señales de RS con el objetivo de reducir la información no relevante para el estudio. En primer lugar, se aplica un filtro pasa-banda tipo Butterworth de orden 4 con un ancho de banda de 70 a 1900 Hz. Seguidamente, se aplica un filtro tipo notch a 50 Hz y sus armónicos, con un ancho de banda de 1 Hz, para la eliminación de la interferencia de red.

## 2.3. Segmentación de ciclos respiratorios

Tras filtrar las señales de RS, se determinaron de manera automática los inicios y finales de los ciclos respiratorios (Figura 1), tanto para la fase de inspiración como de espiración, con un algoritmo basado en los pasos por el valor de offset, estimado previamente de forma semiautomática, de la señal de flujo respiratorio.



**Figura 1.** Distribución de los sensores (izquierda) y señales registradas de flujo y RS (derecha), con la selección automática de ciclos (cian) y la serie temporal de fSampEn (rojo).

## 2.4. Estimación de la intensidad de RS mediante entropía muestral fija

Los métodos convencionales de medida de amplitud de señales, como el *root mean square* (RMS), son muy sensibles a interferencias y ruidos impulsionales. La entropía muestral fija (fSampEn) ha sido propuesta en estudios previos como una técnica muy robusta frente a ruidos impulsionales para la evaluación de la actividad muscular respiratoria [3] o la intensidad de los RS [6]. La fSampEn es una medida de regularidad y complejidad de señales temporales que se calcula a partir de una ventana móvil y tiene como parámetros la longitud de ventana (N), la tolerancia (r) y la dimensión (m), los cuales deben fijarse previamente. De esta manera, la fSampEn no solamente mide la complejidad de una señal, sino que también es sensible a los cambios de amplitud de la señal, pudiendo utilizarse para estimar la intensidad de los RS. En este estudio, las series temporales de fSampEn de los RS se han obtenido utilizando unos valores de los parámetros N, r y m que fueron determinados en un estudio previo y que resultaron ser óptimos para la estimación de la intensidad

de los RS (N=0,35 s con un overlap de 0,3 s, r=0,2·0,0085 y m=2) [6].

## 2.5. Extracción de parámetros y selección de ciclos

Para la señal de flujo, se calcularon los valores pico del flujo inspiratorio, así como los volúmenes inspiratorios obtenidos a partir de la integración definida de la señal de flujo. En el caso de las señales de RS, se calcularon los picos de fSampEn de los tramos inspiratorios (por ser la fase activa de la respiración) para los 4 canales de RS.

Una vez obtenidos los parámetros, se llevó a cabo una selección automática de 5 ciclos respiratorios por cada nivel de flujo y participante, teniendo en cuenta dos criterios: para la señal de flujo, la desviación mínima del pico de flujo inspiratorio respecto al nivel de flujo fijado para cada paciente durante el protocolo; para las señales de RS, la desviación mínima del pico de fSampEn respecto a la mediana del pico de fSampEn de cada nivel de flujo. En la Figura 1 se muestran los ciclos seleccionados en azul cian para uno de los participantes y niveles de flujo.

## 3. Resultados

### 3.1. Base de datos

La Tabla 1 muestra un resumen de los datos antropométricos, espirométricos y de flujo respiratorio de los sujetos de la base de datos utilizada. Se observan diferencias significativas entre controles y pacientes COPD1 para el flujo base; entre controles y pacientes COPD2 para el MIF; y entre controles y ambos grupos de pacientes para la edad, el FEV<sub>1</sub> y la ratio FEV<sub>1</sub>/FVC.

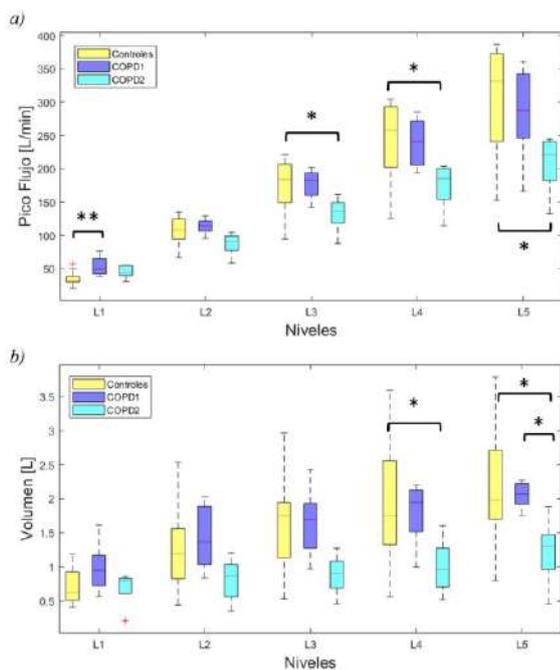
Parámetro	Controles	COPD1	COPD2
Edad	39.93 (30.25,43.50) *#	70.86 (66.75,75.75) *	66.00 (63.25,68.25) #
BMI	23.49 (20.90,25.50)	27.67 (25.53,31.30)	22.68 (18.71,26.05)
FEV <sub>1</sub> (% teórico)	100.91 (92.93,107.65) *#	60.80 (55.13,69.00) *	39.36 (33.18,43.60) #
FVC (% teórico)	103.29 (91.98, 114.23)	104.83 (98.15, 117.30)	82.74 (70.27,98.20)
FEV <sub>1</sub> /FVC (%)	79.93 (75.23,85.29) *#	44.89 (36.74,52.65) *	38.26 (28.76,46.81) #
Flujo base	35.27 (30.00, 39.50) *	54.00 (42.75,65.25) *	45.80 (40.25,55.00)
MIF	315.73 (262.50,373.75) *	305.71 (260,353.75)	216.80 (186.50,251) *

**Tabla 1.** Parámetros antropométricos, espirométricos y de flujo respiratorio en sujetos sanos y pacientes con COPD. Resultados significativos entre grupos mediante los símbolos \* y #. Índice de masa corporal (BMI), volumen espirado forzado en el primer segundo (FEV<sub>1</sub>), capacidad vital forzada (FVC), flujo inspiratorio máximo (MIF).

### 3.2. Flujos y volúmenes respiratorios

En la Figura 2a se observa una tendencia ascendente de los valores pico de flujo a lo largo del protocolo de flujo inspiratorio variable. El grupo control presenta los valores más altos de pico de flujo para los niveles L2 a L5. En esos mismos niveles, se observan unos valores de pico de flujo inferiores en el grupo COPD2 respecto a los del grupo COPD1 o el grupo control, siendo esa diferencia significativa entre controles y pacientes COPD2 en los niveles L3 a L5 ( $p=0,003$  en todos los casos). En contraposición, los pacientes COPD presentan valores pico de flujo más elevados que el grupo control en el nivel L1, siendo esa diferencia significativa entre controles y pacientes COPD1 ( $p=0,007$ ).

En cuanto al volumen inspiratorio (Figura 2b), éste aumenta linealmente a medida que se avanza en los niveles del protocolo, observándose valores mayores en los controles y los pacientes COPD1 que en los pacientes COPD2 en los niveles L2 a L5. Esa diferencia se va haciendo más notoria a lo largo del protocolo, siendo más significativa en el nivel L5 ( $p=0,018$  entre controles y COPD2,  $p=0,04$  entre COPD1 y COPD2). Como ocurre en el nivel L1 con el pico de flujo, se observa que los volúmenes de los pacientes con COPD son mayores que los de los controles en respiración normal, aun sin llegar a haber una diferencia significativa.



**Figura 2.** Gráfico de tendencias para el pico de flujo (a) y el volumen (b) durante el protocolo de flujo inspiratorio variable. Se muestran los resultados de Kruskal-Wallis con \* para  $0.01 < p < 0.05$  y \*\* para  $p < 0.01$ .

### 3.3. Intensidad de los RS

La Figura 3 muestra los valores pico de la fSampEn en los tramos inspiratorios para los 4 canales de RS. Dada la estrecha relación que existe entre los RS y el flujo, en todos los canales de RS se aprecia un comportamiento similar al visto en la Figura 2, con un incremento lineal del valor pico

de fSampEn a medida que aumenta el nivel de flujo demandado. Aunque no existen diferencias significativas entre grupos para ninguno de los niveles del protocolo y en ninguno de los canales de RS, sí se observa que, en general, el grupo control tiene valores pico de fSampEn superiores a los de los pacientes con COPD, especialmente en los niveles más altos (L3 a L5) y en los canales del lado izquierdo (RS<sub>si</sub> y RS<sub>ii</sub>).

De manera similar al pico de flujo y el volumen, se observa que los valores pico de fSampEn son mayores en los pacientes con COPD con respecto a los controles en el nivel L1 y en los 4 canales de RS.

## 4. Discusión

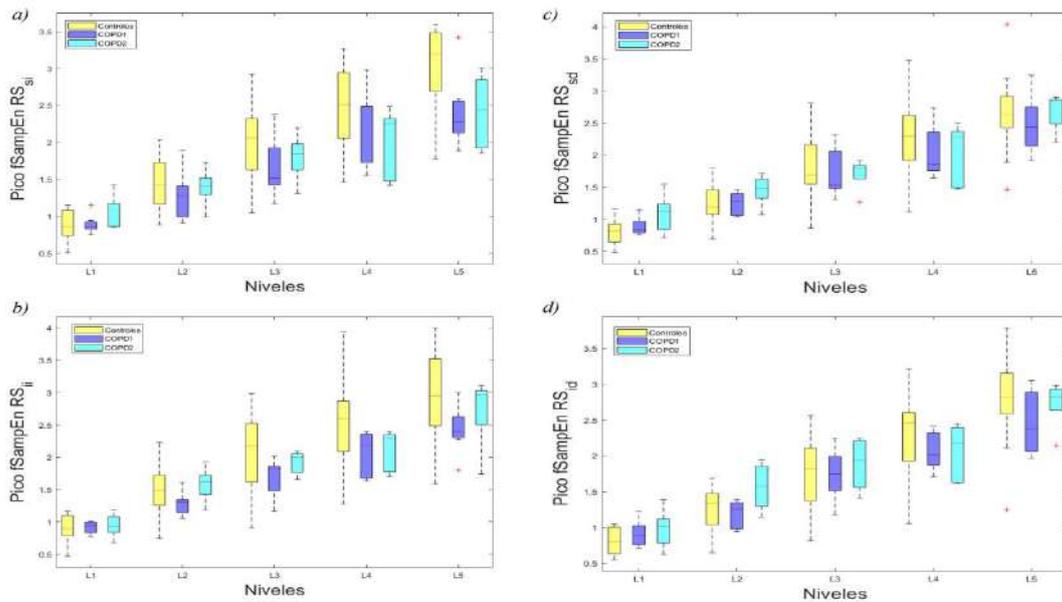
En primer lugar, en la Tabla 1 se ha observado que existen diferencias significativas entre los grupos de estudio tanto en parámetros espirométricos como en el nivel de flujo basal y los valores MIF. Estas diferencias permitieron suponer que *a posteriori* pudieran encontrarse diferencias significativas entre grupos en la intensidad de los RS durante el protocolo de flujo inspiratorio variable.

El análisis del flujo inspiratorio reveló la obstrucción de las vías respiratorias que presentan los sujetos con COPD, a diferencia de los controles sanos, en estado basal. Como se puede observar en la Figura 2, tanto el grupo COPD1 como el grupo COPD2 tuvieron valores más altos de flujo basal (L1), lo que permite deducir que a niveles más bajos de demanda respiratoria ya existe un sobreesfuerzo por parte de los pacientes para realizar una respiración normal. A niveles más altos de flujo respiratorio, esta misma causa es la que hace que no generen valores de flujo tan elevados como los individuos sanos, siendo los pacientes más severos (COPD2) los que menos flujo generan.

Los resultados del volumen son muy parecidos a los anteriores, observándose valores de volumen inferiores en los sujetos del grupo COPD2 con relación a los demás grupos. Una vez más, este comportamiento puede originarse por la hiperinflación pulmonar y el estrechamiento de las vías respiratorias, siendo esta afectación mayor en sujetos que padecen un grado severo de la enfermedad y cuanto mayor es la demanda de flujo.

En cuanto a los resultados obtenidos para la intensidad de los RS, se han observado patrones de comportamiento similares a los del flujo respiratorio, puesto que los sonidos respiratorios se generan a partir del flujo turbulento en las vías respiratorias. Los valores de intensidad de RS más elevados en el grupo control con respecto a los pacientes con COPD, demuestran que los RS son sensibles al deterioro de la mecánica respiratoria de estos pacientes y corrobora la relación estrecha que guardan RS y flujo respiratorio. El hecho de que las diferencias entre grupos sean más claras en el lado izquierdo que en el derecho puede deberse a la variación del tamaño de los pulmones y a la mayor intensidad de los RS en el lado izquierdo cuando se registran en la parte posterior torácica [5].

Por otro lado, cabe comentar que independientemente del lado de registro existe la posibilidad de que hubiera sonidos adventicios (principalmente sibilantes) en las señales de



**Figura 3.** Gráfico de tendencias para el pico de  $fSampEn$  de las señales de RS registradas en la parte superior izquierda ( $RS_{si}$ ) (a), inferior izquierda ( $RS_{li}$ ) (b), superior derecha ( $RS_{sd}$ ) (c) e inferior derecha ( $RS_{id}$ ) (d) del tórax posterior durante el protocolo de flujo inspiratorio variable.

RS, lo que puede provocar un aumento del valor de  $fSampEn$  en los ciclos con este tipo de sonidos. No obstante, este hecho no se ha comprobado, dado que el análisis de sonidos adventicios queda fuera del alcance de este estudio.

Finalmente, la dimensión de la base de datos ha supuesto una limitación en el análisis estadístico, lo que sugiere la posibilidad de llevar a cabo un posterior estudio con un número mayor de sujetos.

## 5. Conclusiones

Los resultados obtenidos en este estudio muestran que hay diferencias en los RS generados por controles sanos y pacientes con COPD durante un protocolo de flujo inspiratorio variable. Por lo tanto, los RS son sensibles a las alteraciones de la función respiratoria causadas por la COPD. Por ello, podrían derivarse nuevos biomarcadores fisiológicos no invasivos de las señales de RS para contribuir a mejorar el diagnóstico y seguimiento de los pacientes con COPD.

Se ha corroborado además la robustez de la técnica de  $fSampEn$  para el análisis de señales de RS. Este tipo de señales suelen presentar ruidos impulsionales a los cuáles son sensibles otros métodos más tradicionales, como puede ser el RMS.

Como trabajo futuro, sería interesante explorar la combinación de los parámetros propuestos con otros índices fisiológicos (por ejemplo, de la actividad muscular respiratoria) para derivar índices de eficiencia neuromecánica respiratoria [3] que puedan integrarse al diagnóstico y monitorización de la COPD.

## Agradecimientos

Este trabajo ha sido financiado por el Ministerio de Ciencia e Innovación y FEDER, en el proyecto PID2021-

126455OB-I00, por la European Respiratory Society (fellowship ERS LTRF 2015-5185) y por el programa CERCA de la Generalitat de Catalunya.

## Referencias

- [1] Global Initiative for Chronic Obstructive Lung Disease. 2022. “2022 GOLD Reports - Global Initiative for Chronic Obstructive Pulmonary Disease.” GOLD Reports, 1–130. <https://goldcopd.org/2022-goldreports-2/>
- [2] World Health Organization. 2022. “Chronic Obstructive Pulmonary Disease (COPD).” [https://www.who.int/news-room/fact-sheets/detail/chronic-obstructive-pulmonary-disease-\(copd\)](https://www.who.int/news-room/fact-sheets/detail/chronic-obstructive-pulmonary-disease-(copd)).
- [3] Lozano-García, Manuel, Luis Estrada-Petrocelli, Dolores Blanco-Almazan, Basak Tas, Peter SP Cho, John Moxham, Gerrard Francis Rafferty, Abel Torres, Raimon Jane, and C. J. Jolley. 2022. “Noninvasive Assessment of Neuromechanical and Neuroventilatory Coupling in COPD.” *IEEE JBHI* 1–1. <https://doi.org/10.1109/JBHI.2022.3166255>.
- [4] Aguilar Paredes, Francesca. 2022. “Enfoque multimodal para la evaluación de la enfermedad pulmonar obstructiva crónica basado en sonidos respiratorios y señales miográficas”. Trabajo Final de Grado en Ingeniería Biomédica. UPC. Directores: Manuel Lozano García y Raimon Jané Campos.
- [5] Lozano García, Manuel, Clare M. Davidson, Carlos Prieto-Ramón, John Moxham, Gerrard F. Rafferty, Caroline J. Jolley, and Raimon Jané. 2020a. “Spatial Distribution of Normal Lung Sounds in Healthy Individuals under Varied Inspiratory Load and Flow Conditions.” *IEEE EMBC* p. 2744-47 <https://doi.org/10.1109/EMBC44109.2020.9175992>
- [6] Lozano-García, Manuel, Jasna Nuhic, John Moxham, Gerrard F. Rafferty, Caroline J. Jolley, and Raimon Jané. 2020b. “Performance Evaluation of Fixed Sample Entropy for Lung Sound Intensity Estimation.” In *2020 EMBC* 2740–2743. IEEE. <https://doi.org/10.1109/EMBC44109.2020.9176215>

# Análisis de un nuevo método de posicionamiento de electrodos en el cuero cabelludo para las pruebas electrofisiológicas

A. Rigo-Vidal<sup>1</sup>, A. Fabregat-Sanjuan<sup>1</sup>, R. Pàmies-Vilà<sup>2</sup>, V. Pascual-Rubio<sup>3</sup>

<sup>1</sup> Departamento Ingeniería Mecánica, Universitat Rovira i Virgili, Tarragona, España, [agnes.rigo@alumni.urv.cat](mailto:agnes.rigo@alumni.urv.cat), [a.fabregat@urv.cat](mailto:a.fabregat@urv.cat)

<sup>2</sup> Departamento de Ingeniería Mecánica, Universitat Politècnica de Catalunya, España. Email: [rosa.pamies@upc.edu](mailto:rosa.pamies@upc.edu)

<sup>3</sup> Departamento de Neurofisiología Clínica, Hospital Sant Joan de Reus, España. Email: [vicenc.pascual@salutsantjoan.cat](mailto:vicenc.pascual@salutsantjoan.cat)

## Resumen

*En diferentes pruebas neurofisiológicas diagnósticas y terapéuticas, se requiere el posicionamiento de electrodos de registro o de estimulación en el cuero cabelludo según el sistema internacional 10/20. En la práctica clínica, para localizar las posiciones craneales, se utiliza una cinta métrica y un rotulador con el que se realiza el proceso de marcaje, previo a la colocación de los electrodos. En este estudio se ha comparado el marcaje del método convencional y un nuevo método desarrollado en la URV, el dispositivo EPlacement para distintos tests electrofisiológicos. La comparación se ha realizado teniendo en cuenta el tiempo utilizado, la precisión en la localización de las posiciones y una encuesta realizada al personal sanitario. En el estudio han participado 10 sanitarios, se han realizado 90 marcajes en maniqués y 10 marcajes en voluntarios. Los resultados muestran que con EPlacement, se incrementa la exactitud en el marcaje, se reduce el tiempo empleado para dicho marcaje, y el personal sanitario es receptivo a adoptar esta tecnología.*

## 1. Introducción

La Neurofisiología Clínica es una disciplina médica cuyo objetivo básico es el estudio fisiopatológico de los grandes síndromes y enfermedades del sistema nervioso central y periférico [1]. Utiliza equipos y técnicas de registro, estimulación y procesamiento de la señal bioeléctrica avanzados. Las diferentes técnicas pueden tener fines diagnósticos, pronósticos o terapéuticos: por ejemplo, la electroencefalografía, electromiografía, potenciales evocados, así como también la neuromodulación cerebral (eléctrica o magnética) [2].

Para realizar las técnicas neurofisiológicas diagnósticas y terapéuticas mencionadas anteriormente, se requiere del posicionamiento de unos dispositivos de registro o de estimulación en el cuero cabelludo próximo a la corteza cerebral subyacente. Para el posicionado de dichos dispositivos se utiliza el sistema 10/20 de la Federación Internacional en Neurofisiología Clínica (IFCN) [3]. El sistema se basa en dividir la superficie de la cabeza en segmentos obtenidos como porcentajes de medidas anatómicas craneales. Los puntos resultantes permiten localizar las áreas corticales de interés subyacentes.

En la práctica clínica habitual, para realizar el marcaje se utiliza una cinta métrica y un rotulador. Una vez determinados los distintos puntos craneales de cada prueba electrofisiológica, se marcan con un rotulador para ser detectados visualmente por el explorador y poder colocar

los electrodos. Por este motivo, al acto de determinar y marcar los puntos se le llama marcaje. Es muy importante que el marcaje se realice de forma precisa, ya que, de lo contrario, los resultados de la técnica neurofisiológica no serán óptimos o incluso podrían convertirse en erróneos o contraproducentes.

En ocasiones, el cálculo para el posicionado de los electrodos se realiza en situaciones de estrés del personal sanitario, como podría suceder durante una guardia nocturna en una unidad de cuidados intensivos con una carga asistencial elevada, lo que hace que la medida y los cálculos pertinentes no se realicen de forma adecuada. Hay también otras situaciones que requieren un marcado rápido y preciso, como la monitorización neurofisiológica intraoperatoria (MNIO) [5]. En la MNIO, el marcaje deber ser rápido para no demorar cirugías que de por sí, suelen ser complejas y de larga duración. Asimismo, hay que ser muy preciso en la colocación de los electrodos, para evitar la reubicación de estos electrodos ya que puede suponer detener la cirugía y poner en peligro las condiciones de esterilización del campo quirúrgico.

Cabe destacar que con la aplicación del método 10/20 pueden aparecer numerosas fuentes de error, ya que para cada prueba se requieren unos puntos diana determinados y se deben realizar unos cálculos de porcentajes individualizados, de forma manual, ya que se requieren en el momento en el que se tienen ambas manos ocupadas para efectuar la medición [3, 4].

Las medidas anatómicas craneales básicas son los segmentos Nasion-Inion (Ns-In), Tragus-Tragus (LPA-RPA) y el perímetro craneal. Cuando se posicionan los puntos que no están en estos segmentos (por ejemplo las posiciones F3, F4, CP3, CP4, P3, P4) el error en el posicionado es mayor por la acumulación de errores en las mediciones previas [6]. Es decir, la localización de F3, F4, CP3, CP4, P3 y P4 depende, entre otras, de la posición de electrodos realizada en un marcaje previo, con lo que eso significa de posible acumulación progresiva de errores.

Para dar solución a todos estos problemas, se requiere una herramienta de marcaje que guíe de forma rápida y precisa la colocación de los electrodos en la superficie craneal y que se adapte a los requerimientos de cada prueba neurofisiológica. Esta es la misión del dispositivo EPlacement, un nuevo dispositivo desarrollado por nuestro grupo de investigación. El nuevo dispositivo resuelve los problemas del marcaje convencional porque proporciona,

de forma rápida, sencilla, guiada y precisa, la localización de puntos craneales en el contexto de las diferentes pruebas o estudios neurofisiológicos.

En este estudio se ha evaluado la diferencia que hay entre el método convencional de marcaje y el que ofrece EPlacement.

## 2. Métodos, metodología

### 2.1. Sistema internacional 10/20

El sistema 10/20 es un método que determina la ubicación de puntos en el cuero cabelludo. Los números "10" y "20" se refieren a que las distancias entre electrodos adyacentes son el 10% o el 20% de la distancia total antero-posterior o derecha-izquierda del cráneo [7]. El método relaciona la ubicación de cada punto con la zona subyacente de la corteza cerebral. Cada posición tiene una letra para identificar la región cerebral subyacente y un número para identificar la lateralidad, correspondiendo los números pares al hemisferio derecho y los impares al izquierdo. Las áreas principales son frontopolar (Fp), frontal (F), central (C), temporal (T), parietal (P) y occipital (O). La "z" (cero) se refiere a un electrodo colocado en la línea media craneal.

Para el posicionamiento de los electrodos se utilizan cuatro puntos de referencia anatómicos: en primer lugar, el Nasion, que es el punto situado entre la frente y la nariz; en segundo lugar, el Inion, que es el punto más bajo del cráneo desde la parte posterior de la cabeza, y que normalmente se indica con una protuberancia prominente; y los Tragus, los puntos preauriculares anteriores a la oreja [8]. Una vez ubicados estos cuatro puntos se mide la distancia entre ellos y se realizan marcas cada 10 y/o 20% de la longitud medida. A partir de estas medidas, se sitúan el resto de los electrodos del sistema internacional.

### 2.2. EPlacement

El dispositivo EPlacement consta de una unidad electrónica (pantalla, microcontrolador y batería), un sensor de presión y un sistema de iluminación con alta densidad de leds. La unidad electrónica dispone de un menú de navegación donde el clínico escoge la prueba a realizar (potenciales evocados visuales (PEVP), somatosensoriales (PESS) y MNIO para este estudio) y el dispositivo informa con una guía paso a paso, de los distintos puntos requeridos en la prueba. El sensor de presión permite localizar las referencias craneales de forma directa y medir su longitud. Con esta medición, el microcontrolador realiza los cálculos e ilumina el led correspondiente para que el clínico marque la posición sobre el cuero cabelludo.

El sensor de presión y el sistema de iluminación están integrados en una tira que se adapta fácilmente a la forma de la cabeza. Existen dos versiones del dispositivo EPlacement, el primero está formado por una única tira (EP-1S), y el segundo está integrado por dos tiras con movimiento relativo entre ellas (EP-2S). El menú de navegación adapta el procedimiento a seguir según la versión del dispositivo.

### 2.3. Plan de investigación clínica

Antes de realizar esta investigación se recibió la autorización del "Comité de Ética del Institut d'Investigació Sanitària Pere Virgili" (IISPV). La investigación se ha realizado mediante un estudio experimental transversal donde han participado 10 sanitarios del Hospital Universitario Sant Joan de Reus (HUSJR).

El estudio se ha dividido en dos partes. La primera parte consiste en medir la precisión y el tiempo requerido de cada método de marcaje: EPlacement (de una y dos tiras) y cinta métrica. El personal sanitario realizó los marcajes con cada método sobre tres cabezas de maniqués adultos. El primer y segundo marcado se hicieron con dos rotuladores solo visibles con luz ultravioleta (azul y rosa) para no influir en el tercer marcado (cinta métrica) que se realizó con un rotulador negro. Cada proceso de marcaje fue cronometrado.

Los marcajes que se hicieron en la cabeza de maniqué corresponden a tres pruebas diagnósticas neurofisiológicas: un potencial evocado visual de patrón invertido PVEP (se requieren 3 electrodos), unos potenciales evocados somatosensoriales de extremidades superiores e inferiores PESS (se requieren 5 electrodos) y una monitorización neurofisiológica intraoperatoria MNIO (se requieren 9 electrodos).

En los maniqués, una vez realizado cada marcaje, se señalaron con agujas marcando los puntos del primer y segundo marcaje (iluminado con luz ultravioleta para su localización) para, posteriormente, calcular las diferencias en mm entre cada punto respecto a los puntos de referencia determinados mediante una plantilla impresa en 3D (explicada en detalle en la siguiente sección).

En la segunda parte del estudio se ha valorado, mediante una encuesta, la facilidad de uso y la comodidad de cada método. Antes de realizar la encuesta, también se realizó un marcaje (PESS) a voluntarios adultos para responder la encuesta teniendo en cuenta el uso en cabezas humanas. No se analizó el tiempo ni la exactitud en los voluntarios por la gran variabilidad que hay (densidad y longitud de pelo, además de forma cefálica).

### 2.4 Metodología para el análisis del error en el marcaje.

Para hacer el estudio se han utilizado 30 cabezas de maniqués idénticos. Previo al marcaje, se realizó un escaneado de precisión con luz estructurada (HP 3D Pro S3). La superficie resultante se ha importado al programa Fusion 360 de Autodesk, con el que se han determinado las ubicaciones reales de los electrodos del sistema 10/20. Estas se obtuvieron utilizando el método propuesto en el sistema 10/20 a partir de las medidas Ns-In y LPA-RPA obtenidas mediante la intersección del plano con la sección correspondiente. Con este mismo programa se ha creado un casco/plantilla de 2 mm de grosor con agujeros perpendiculares a la superficie en las posiciones del sistema 10/20 (Figura 1). El casco se ha imprimido en la impresora Original Prusa i3 MK3S+.

En cada maniquí, una vez realizado los marcajes manuales, se acopló el casco impreso en 3D y se marcaron las posiciones de referencia del sistema 10/20. A continuación, y mediante el uso de luz ultravioleta, se señalaron con agujas los puntos del primer y segundo marcado (obtenidos con las dos versiones del dispositivo EPlacement, Figura 1). Finalmente, para cada posición del sistema 10/20 se midieron las distancias entre los puntos experimentales (EPlacement y cinta métrica) y las posiciones de referencia (casco plantilla).



Figura 1. Maniqués con los marcajes y plantilla impresa

### 3. Resultados

#### 3.1. Exactitud en el marcaje

Para el análisis de la exactitud en el marcaje se han dividido las diferentes posiciones en dos grupos. Los electrodos que se encuentran en los planos Ns-In y LPA-RPA (Fpz, Fz, Cz, Oz, C1, C2, C3 y C4) y los que no (CP3, CP4 y CPz). Esta división se ha realizado porque el error ( $\epsilon$ ) en los electrodos que no pasan por los planos Ns-In y LPA-RPA en la práctica clínica, es mucho mayor debido a que para el marcaje se utilizan métodos aproximados o bien se parte de otras posiciones obtenidas previamente en el plano del perímetro craneal.

En la Figura 3 se muestra un gráfico radial agrupado del error por electrodo en distancia media sobre los tres tipos de marcaje utilizados, el EPlacement de una tira (verde), el de dos tiras (naranja) y finalmente la cinta métrica (gris). Se puede identificar claramente que el error en cinta métrica es mayor al obtenido con EPlacement y que el error con EPlacement es similar en ambas versiones. El error medio con cinta métrica es de 5,0 mm, con una tira de 3,1 mm, y con el sistema de dos tiras de 3,5 mm.

En los resultados de las posiciones CPz, CP3 y CP4 de la Figura 4 se puede observar que el error en cinta métrica es, de nuevo, mucho mayor al obtenido con EPlacement y que el error con EPlacement es similar en ambas versiones. En este caso, cabe destacar que los valores de error con cinta métrica son muy altos, con un error medio de 16,3 mm, en comparación con 3,8 mm de una tira y 3,6 mm de dos tiras.

Parte de esta gran diferencia se explica teniendo en cuenta que, en la práctica clínica usual, la metodología utilizada para la localización de estas tres posiciones no utiliza los pasos del sistema internacional 10/20 (por ser un procedimiento muy largo) sino que se colocan 2 cm detrás del electrodo Cz, C3, C4 respectivamente [9].

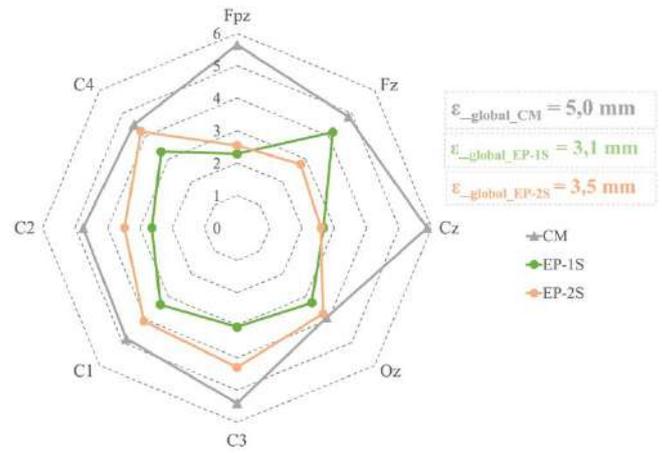


Figura 3. Errores (en mm) en el marcaje de las posiciones que están en el plano Ns-In y LPA-RPA.

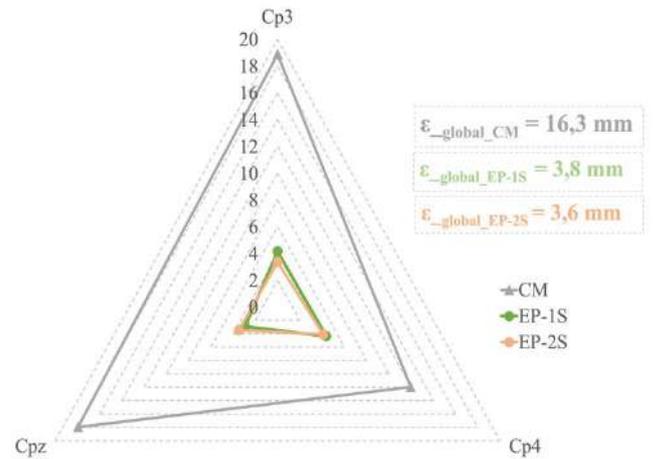


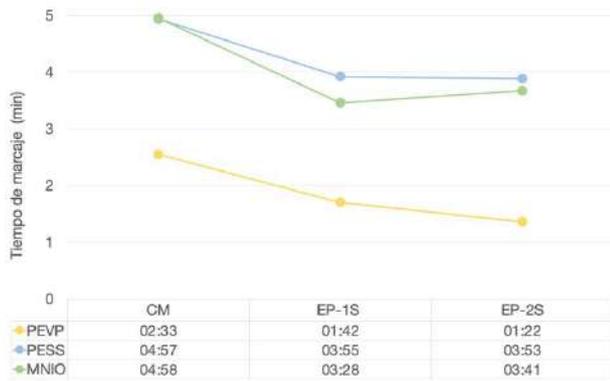
Figura 4. Errores (en mm) en el marcaje de las posiciones que no están en el plano Ns-In y LPA-RPA.

#### 3.2 Cronometraje

Para el análisis del tiempo utilizado para cada tipo de marcado, se realizaron dos mediciones de tiempo. En la primera se siguieron los pasos del sistema internacional 10/20 y en la segunda se utilizó la regla de los 2 cm para los electrodos que no se encuentran en los planos Ns-In o LPA-RPA.

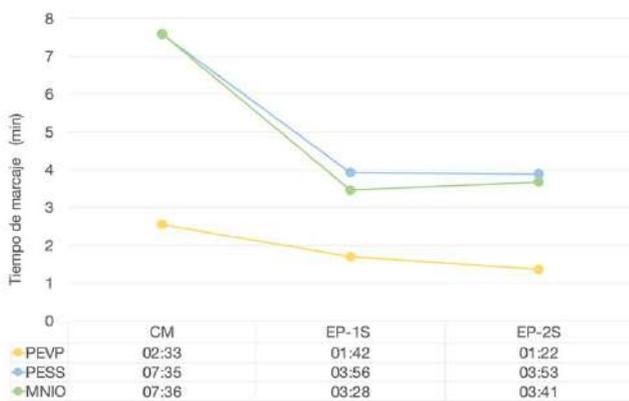
La figura 5 muestra el tiempo de marcado siguiendo el sistema internacional 10/20. Se detecta una diferencia significativa entre la cinta métrica y el EPlacement (media de EP-1S y EP-2S), con una reducción de 0'41'(27%) en el PVEP; 4'17''(52%) en el SEP; 4'02''(53%) en el MNIO. No se detectan diferencias de tiempo significativas entre las versiones de una y dos tiras de EPlacement.

El resultado de la segunda medición, utilizando la regla de los 2 cm en las posiciones que no se encuentran en los planos de Ns-In o LPA-RPA se han obtenido los siguientes resultados (ver Figura 6).



**Figura 5.** Tiempo de marcado de cada prueba (PEVP, PESS, MNIO) siguiendo el sistema internacional 10/20. Resultados de Cinta Métrica (CM) y EPlacement (EP-1s, EP-2s).

De nuevo, en este caso el nuevo dispositivo implica una reducción en los tiempos, aunque en un porcentaje menor. En cuanto a la PVEP, los resultados son los mismos que los anteriores, ya que los 3 electrodos necesarios están en el plano Ns-In. La diferencia detectada ahora entre la cinta métrica y el EPlacement tiene una reducción de 1'02" (21%) en el SEP y 1'23"(28%) en el MNIO.



**Figura 6.** Tiempo de marcado de cada prueba (PEVP, PESS, MNIO) siguiendo el sistema internacional 10/20 excepto en las posiciones fuera del plano Ns-In o LPA-RPA. Resultados de Cinta Métrica (CM) y EPlacement (EP-1s, EP-2s).

### 3.2. Encuesta de opinión

La encuesta de opinión demuestra que el personal sanitario está satisfecho con el nuevo método que proporciona el dispositivo EPlacement. En la pregunta cómo calificaría EPlacement en comparación con el sistema de marcado tradicional, los resultados fueron que un 90% contestaron que es la mejor opción que conocen. También se preguntó según el montaje que versión del EPlacement recomendaría. Las respuestas indican que, si el número de electrodos es inferior a 5 y especialmente si todos están en el mismo plano, se prefiere (80%) la versión de una tira, y, si el número de electrodos es mayor o igual a 5 y los electrodos están en diferentes planos, se prefiere (90%) la versión con 2 tiras.

## 4. Conclusiones

Los resultados demuestran que el nuevo método con el dispositivo EPlacement mejora el método convencional de marcaje. En posiciones EEG del sistema 10/20 de los planos Ns-In y LPA-RPA, se ha visto que el dispositivo disminuye el error de posicionado, de media, en un 34% (1,7 mm) respecto la cinta métrica. En posiciones que no se encuentran en los planos Ns-In y LPA-RPA, esta mejora es del 77% (12,5 mm). Este resultado es muy significativo y denota la deficiencia actual de marcaje en las posiciones que no se encuentran en el plano Ns-In o LPA- RPA.

Al mismo tiempo que mejora la exactitud, el dispositivo EPlacement reduce el tiempo de preparación por paciente en las tres pruebas diagnósticas neurofisiológicas estudiadas. La reducción es del 52% si se utiliza el método 10/20 para todas las posiciones, y un 24% comparado con el método aproximado de la regla de los 2 cm.

Finalmente, la encuesta del personal sanitario muestra una clara disposición a utilizar el nuevo dispositivo en la práctica clínica habitual.

## 5. Agradecimientos

Se agradece al personal sanitario colaborador del Servicio de Neurofisiología Clínica del HUSJR y la financiación obtenida en la convocatoria R2B2021-02, cofinanciada entre la URV y el Fondo Europeo de Desarrollo Regional.

## 6. Referencias

- [1] G. Morales and J. Artieda, "La neurofisiología clínica: pasado, presente y futuro.," *An Sist Sanit Navar*, vol. 32 Suppl 3, 2009, (ISSN: 1137-6627).
- [2] P. S. Bialik, "La neurofisiología clínica en la práctica médica.," *Medigraphic*, Vol. 62, Núm.2 Abr. -Jun. 2016 p.84-86, (ISSN:2683-2348).
- [3] M. Seeck et al., "The standardized EEG electrode array of the IFCN," *Clinical Neurophysiology*, vol. 128, no. 10, pp. 2070–2077, 2017, (ISSN: 1388-2457).
- [4] V. Jurcak, D. Tsuzuki, and I. Dan, "10/20, 10/10, and 10/5 systems revisited: Their validity as relative head-surface-based positioning systems," *Neuroimage*, vol. 34, no. 4, 2007, (ISSN: 1053-8119)
- [5] K. J. Nagle *et al.*, "Intraoperative monitoring of motor evoked potentials: A review of 116 cases," *Neurology*, vol. 47, no. 4, 1996, doi: 10.1212/WNL.47.4.999.
- [6] K. B. E. Böcker, J. A. G. van Avermaete, and M. M. C. van den Berg-Lenssen, "The international 10-20 system revisited: Cartesian and spherical co-ordinates," *Brain Topography*, vol. 6, no. 3, 1994, (ISSN: 15736792, 08960267).
- [7] A. Morley, L. Hill, and A. G. Kaditis, "10-20 System EEG Placement," *European Respiratory Society*, 2016.
- [8] Trans Cranial Technologies, "10/20 System Positioning MANUAL," 2012.
- [9] P.Shkurovich, A.Torres, O.Sánchez, G. Ángel, "Potenciales evocados somatosensoriales del nervio mediano como lateralizador en estado postictal temprano. Un estudio piloto", *Medigraphic*, Vol.61, Núm. 2 Abr.-Jun. 2016 p 112-116, (ISSN:2683-2348).

# Localización de Drivers en Fibrilación Auricular con Redes Convolucionales LSTM

S. Valiente Jaén<sup>1</sup>, M. Gutiérrez Fernández-Calvillo<sup>1</sup>, MA. Cámara-Vázquez<sup>1</sup>, I. Hernández-Romero<sup>2</sup>, M. S. Guillem<sup>2</sup>, A. M. Climent<sup>2</sup>, Ó. Barquero-Pérez<sup>1</sup>

<sup>1</sup>Escuela Técnica Superior de Ingeniería de Telecomunicación, Universidad Rey Juan Carlos, Madrid, España, s.valiente.2018@alumnos.urjc.es, miriam.gutierrez@urjc.es, miguelangel.camara@urjc.es, oscar.barquero@urjc.es

<sup>2</sup>Instituto ITACA, Universitat Politècnica de València, València, España, isherro@itaca.upv.es, mguisan@itaca.upv.es, acliment@itaca.upv.es

## Resumen

*La Fibrilación Auricular (FA) es una arritmia que se caracteriza por un ritmo auricular acelerado, causado por una conducción eléctrica anormal en el corazón, y puede desencadenar serias complicaciones médicas. La ablación es un procedimiento por el cual se aplican pulsos de radiofrecuencia a los mecanismos que mantienen la FA (drivers) con el objetivo de restaurar el ritmo sinusal. Sin embargo, la identificación del driver sigue siendo una cuestión científica sin resolver. Conscientes de esta limitación, proponemos en este trabajo un método no invasivo alternativo a la imagen electrocardiográfica (ECGI) para localizar los drivers de FA en el corazón a partir de registros eléctricos en el torso. El set de datos se compone de Potenciales de Superficie Corporal (BSPs), obtenidos a partir de modelos sintéticos del torso humano y de modelos computacionales de FA. Estos datos se organizaron en diferentes tipos de tensores para alimentar un modelo basado en redes Convolucionales LSTM (ConvLSTM). El modelo propuesto identificó correctamente la aurícula (derecha o izquierda) en la que se encontraba el driver el 99.1% de las veces, mientras que para un problema en el que se segmentaron las aurículas en 6 regiones el modelo identificó la región en la que se encontraba el driver el 99.7% de las veces. La metodología propuesta, por tanto, podría servir para identificar regiones diana para procedimientos de ablación utilizando BSPs antes de realizar el estudio electrofisiológico, permitiendo reducir el tiempo de mapeo eléctrico para la elección del tejido para la ablación.*

## 1. Introducción

La FA es una de las arritmias más prevalentes hoy en día, y dado que la edad es considerada uno de sus factores de riesgo, se prevé que la prevalencia de esta patología incremente en los próximos años [1]. La electrofisiología de la FA se caracteriza por un ritmo cardíaco rápido e irregular, con ausencia de onda P en el electrocardiograma (ECG) [2]. Aunque esta arritmia puede mostrarse asintomática, la contracción desorganizada e irregular del tejido cardíaco auricular incrementa el riesgo de formación de coágulos de sangre y, con ello, el riesgo de sufrir un fallo cardíaco, ictus, etc [3]. No existe consenso sobre las causas de FA, algunos investigadores apoyan la teoría de que la FA se origina por varios puntos focales que se activan de forma espontánea, mientras que otros mantienen que es causada por patrones de propagación eléctrica reentrantes (drivers), siendo los más importantes los denominados patrones cíclicos o rotores. Sin embargo, los diferentes mecanismos de perpetuación de FA descritos pueden coexistir [4].

El tratamiento más común para la FA es farmacoterapia, que suprime sus síntomas, pero no la cura. De forma alternativa,

existe un tratamiento quirúrgico llamado ablación cuya finalidad es devolver el corazón al ritmo sinusal, y que consiste en aplicar pulsos de radiofrecuencia para quemar ciertas regiones del corazón donde se cree que se perpetúan los drivers de FA [5]. Por ello, para que la ablación sea exitosa, es necesario conocer la ubicación anatómica de los puntos de inicio de dichos drivers, sin embargo, la identificación de los drivers en FA sigue siendo un problema sin resolver completamente.

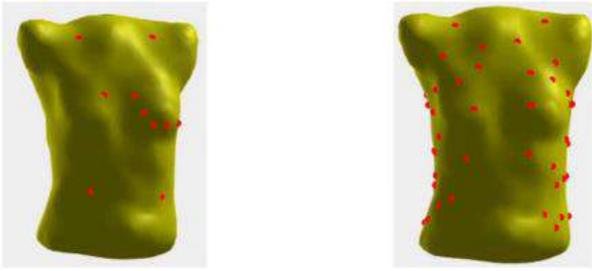
Tradicionalmente, la electrofisiología de la FA se ha estudiado por métodos de mapeo invasivo y no invasivo. El mapeo eléctrico invasivo consiste en insertar catéteres con electrodos intracavitarios en los vasos sanguíneos hasta el corazón [6] para medir la actividad eléctrica cardíaca. En los últimos años se ha trabajado en el mapeo eléctrico no invasivo que permite reconstruir la actividad eléctrica cardíaca resolviendo el problema inverso de la electrocardiografía (ECGI) a partir de medidas eléctricas del torso de los pacientes [7]. El ECGI es un problema matemáticamente mal condicionado, ya que no tiene una solución única y para resolverlo se deben aplicar técnicas de regularización, siendo la regularización Tikhonov la más popular [8].

En los últimos años, algunas tecnologías novedosas se han estado explorando con la intención de resolver las incógnitas presentes en el estudio de la FA. Métodos como el Machine Learning (ML) y Deep Learning (DL) se han empleado para diagnosticar FA [9,10], y también para decidir la mejor técnica de ablación [11]. En este trabajo proponemos mejorar la localización de drivers de FA utilizando medidas de BSPs planteado el problema como uno de clasificación supervisado y utilizando métodos de DL basados en redes convolucionales combinadas con redes recurrentes LSTM. De esta forma, el modelo permite explotar correlaciones espaciales (por la distribución de los BSPs en el torso del paciente) como correlaciones temporales de las señales eléctricas.

## 2. Materiales y Métodos

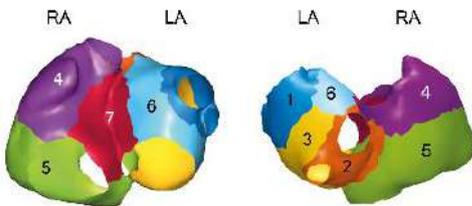
### 2.1. Modelos Computacionales

Para desarrollar este proyecto, se han usado 10 geometrías computacionales del torso. Sobre cada una se han seleccionado 10 y 64 nodos, los cuales representan la ubicación de los electrodos de BSPs.



**Figura 1.** Torso marcado con 10 (izquierda) y 64 (derecha) electrodos

Por otra parte, se han empleado 14 modelos de propagación eléctrica, con una frecuencia de muestreo inicial de 500 Hz, incluyendo un ritmo sinusal (SR) y varios modelos de FA, cada uno con sus señales EGM. Las aurículas se dividieron en 7 regiones, como se representa en la Figura 2, y las etiquetas correspondientes a la ubicación de los drivers se determinaron de forma manual. Finalmente, las 7 regiones iniciales se han reducido a 5 debido a la ausencia de drivers en la región 7 y a la decisión de unir las regiones 2 y 3 dado al desequilibrio entre etiquetas [13,14].



**Figura 2.** División de las aurículas en regiones

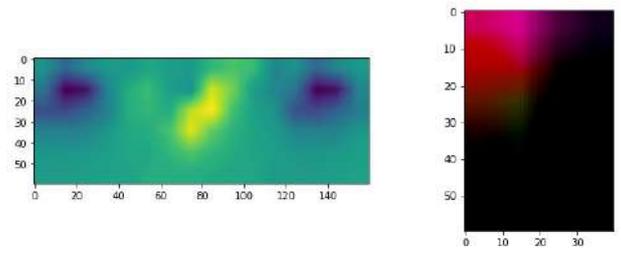
## 2.2. Obtención de los datos de entrada

En primer lugar, el procedimiento seguido para construir el conjunto de datos ha consistido en aplicar el problema directo de electrocardiografía sobre cada uno de los patrones de propagación y modelos de torso disponibles, resultando en 140 sets de BSPs distintos.

A continuación, se ha realizado un submuestreo de las señales BSP a 50Hz, ya que la FA se caracteriza por tener su banda de actividad en frecuencias inferiores a 25 Hz [15], con el objetivo de reducir la redundancia en los datos.

Finalmente, las señales remuestreadas se han organizado de tres formas diferentes, donde  $n\_times$  es el número de instantes temporales:

- Tensor de 3 canales (3-channel): Los BSPs obtenidos de 64 nodos del torso se organizan espacialmente en electrodos de torso frontal, laterales y espalda. De esta forma, se obtuvieron tensores de tamaño  $(n\_times, 6, 4, 3)$ .
- Tensor de 1 canal (1-channel): Los BSPs obtenidos de 64 nodos del torso se consideran como un cilindro desenrollado a una disposición de 16 columnas (12 asociadas a los BSPs, y 4 columnas espejo). De esta forma, se obtuvieron tensores de tamaño  $(n\_times, 6, 16, 1)$ .
- Tensor plano (flat): En este caso, los BSPs obtenidos de 10 nodos del torso no se organizan en forma de imagen, sino apilados. Al tener 10 electrodos disponibles, los tensores obtenidos tuvieron dimensiones  $(n\_times, 10, 1)$ .



**Figura 3.** Visualizaciones de los conjuntos 1-channel y 3-channel en el instante temporal  $t=25$

Finalmente, para aumentar la cantidad de datos en los tensores de 1 y 3 canales, se ha aplicado interpolación bilineal, aumentando las dimensiones de dichos tensores en un factor 10.

## 2.3. Arquitectura del Modelo ConvLSTM y Esquemas de Clasificación

Para localizar los drivers de los modelos se han usado redes ConvLSTM, que son una combinación de CNNs and LSTMs. Estas redes trabajan con información espacial y temporal dado que su estructura cíclica, propia de las redes LSTM, incorpora capas convolucionales.

El modelo ConvLSTM se ha construido utilizando Python, TensorFlow y Keras. La arquitectura del modelo desarrollado se basa en 2 capas ConvLSTM, de 1 o 2 dimensiones, dependiendo del tamaño del conjunto de datos de entrada. A continuación, se introduce una capa que reordena el tensor en un vector columna, 4 capas completamente conectadas con activación ReLU y por último una capa de salida con activación softmax que permite realizar la clasificación a partir de la estimación de probabilidad de pertenencia a las correspondientes clases. La capa softmax final depende del esquema de clasificación considerado:

- *Clasificación en ausencia de driver, driver en aurícula izquierda y driver en aurícula derecha.* En este caso, la capa softmax se compone de 3 unidades.
- *Clasificación en ausencia de driver y driver en diferentes 5 regiones anatómicas.* En este caso, la capa softmax se compone de 6 unidades: Una para modelar la ausencia de driver, y 5 restantes para cada una de las regiones anatómicas definidas.

## 2.4. Esquema de Validación

Para caracterizar estadísticamente el desempeño de los modelos, el modelo se ha entrenado con validación cruzada, usando 5 subconjuntos de validación. Esta técnica consiste en dividir los datos en  $k$  partes iguales. En cada iteración, una de estas partes es tomada como conjunto de test, mientras que las otras conforman el conjunto de entrenamiento. Para proteger la correlación temporal de los datos, la división en folds se ha realizado por sets de BSPs.

## 2.5. Métricas de Rendimiento

En primer lugar, se han representado las matrices de confusión, que contienen los valores reales y predichos para las categorías del clasificador. A partir de dichas matrices de

confusión se han calculado las siguientes métricas de rendimiento:

- Sensibilidad (recall) y especificidad: La primera es una medida del número de valores positivos reales, mientras que la segunda es una medida del número de valores negativos reales.

$$\text{Sensibilidad} = \frac{TP}{TP+FN}$$

$$\text{Especificidad} = \frac{TN}{TN+FP}$$

- Precisión: Define la fracción de valores positivos reales entre el total de positivos predichos.

$$\text{Precisión} = \frac{TP}{TP + FP}$$

- Exactitud (accuracy) y F1-score: Representan los drivers clasificados correctamente.

$$F1 = \frac{2 \cdot (\text{Precisión} \cdot \text{Sensibilidad})}{\text{Precisión} + \text{Sensibilidad}}$$

Por último, el coste (loss) mide el error introducido en las predicciones del modelo.

### 3. Resultados

En primer lugar, se muestran los resultados asociados al clasificador de 3 etiquetas y un único canal. Las curvas de entrenamiento de la Figura 4 muestran que la pérdida se reduce y la exactitud aumenta a medida que el entrenamiento avanza.

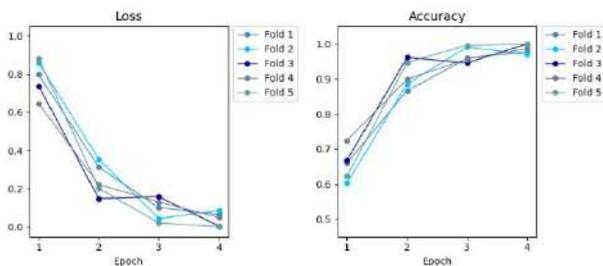


Figura 4. Curvas de coste (loss) y exactitud para el modelo 1-channel y el clasificador de 3 etiquetas

La matriz de confusión de la Figura 5 muestra el buen rendimiento del modelo 1-channel: Sus celdas representan la media y la desviación típica de la correlación entre predicciones y etiquetas reales, siendo 1 el valor más alto.

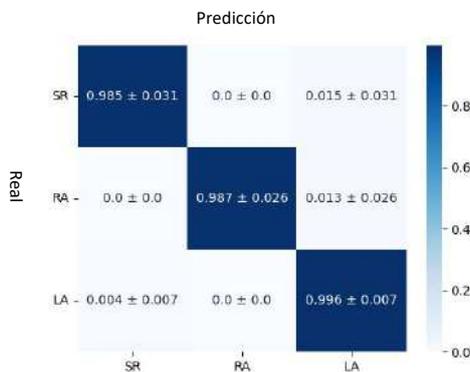


Figura 5. Matriz de confusión del modelo 1-channel y el clasificador de 3 etiquetas

La Tabla 1 contiene las métricas de sensibilidad, especificidad, precisión y F1 por región para el modelo 1-channel. Sus buenos resultados confirman su superioridad frente a los otros modelos.

	Sensibilidad	F1	Especificidad
SR	0.984 ± 0.032	0.984 ± 0.032	0.670 ± 0.006
RA	0.986 ± 0.028	0.994 ± 0.012	0.670 ± 0.006
LA	0.996 ± 0.008	0.992 ± 0.016	0.664 ± 0.005

Tabla 1. Métricas de rendimiento por región del modelo 1-channel y clasificación en 3 etiquetas.

En cuanto al clasificador de 6 etiquetas, el modelo 1-channel también obtuvo en general mejores resultados que los otros modelos. La Tabla 2 muestra sus métricas por región.

	Sensibilidad	F1	Especificidad
SR	1.000 ± 0.000	1.000 ± 0.000	0.833 ± 0.000
R1	0.994 ± 0.012	0.996 ± 0.008	0.834 ± 0.002
R2	1.000 ± 0.000	0.996 ± 0.008	0.833 ± 0.000
R3	1.000 ± 0.000	1.000 ± 0.000	0.833 ± 0.000
R4	0.988 ± 0.024	0.994 ± 0.012	0.835 ± 0.004
R5	1.000 ± 0.000	0.996 ± 0.008	0.833 ± 0.001

Tabla 2. Métricas de rendimiento por región del modelo 1-channel y clasificación en 6 etiquetas.

Los diagramas de cajas y bigotes representando la sensibilidad de ambos experimentos se presentan en las Figuras 6 y 7. Se puede observar que el modelo 3-channel tiene la variación más alta, y el modelo 1-channel, la más baja. Los modelos flat y 1-channel rinden mejor en el clasificador de 6 etiquetas, con valores muy altos de sensibilidad.

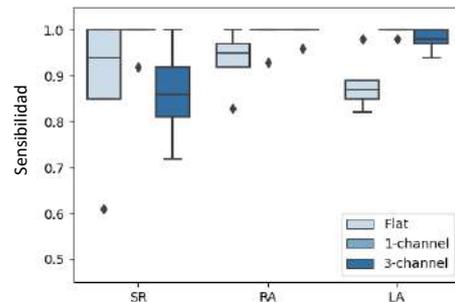


Figura 6. Diagrama de cajas de sensibilidad del clasificador de 3 etiquetas

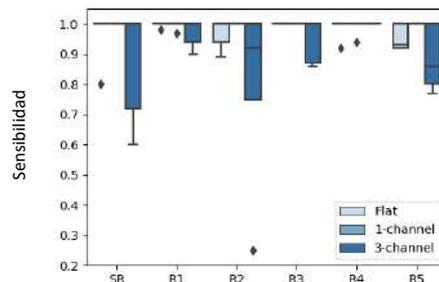


Figura 7. Diagrama de cajas de sensibilidad del clasificador de 6 etiquetas

Las métricas de rendimiento obtenidas para los 3 modelos y ambos experimentos se muestran en la Tabla 3.

	Modelo	Accuracy	F1	Loss
Clasificador 3 etiquetas	Flat	0.910 ± 0.051	0.895 ± 0.064	0.359 ± 0.249
	1-channel	0.991 ± 0.017	0.990 ± 0.021	0.033 ± 0.053
	3-channel	0.971 ± 0.020	0.901 ± 0.128	0.097 ± 0.078
Clasificador 6 etiquetas	Flat	0.984 ± 0.014	0.950 ± 0.060	0.085 ± 0.079
	1-channel	0.997 ± 0.007	0.997 ± 0.006	0.026 ± 0.023
	3-channel	0.936 ± 0.069	0.881 ± 0.085	0.197 ± 0.194

**Tabla 3.** Métricas de rendimiento medias para todos los modelos y ambas clasificaciones

#### 4. Discusión y conclusiones

El modelo 1-channel consigue los mejores resultados, aunque el modelo flat debería tenerse en consideración dado que su configuración es la más fácil y cercana al ECG clásico, y sus resultados sólo difieren ligeramente de los del modelo 1-channel. En cuanto a la clasificación de regiones, ambos experimentos tienen buen rendimiento, y por tanto sería interesante seguir trabajando en el clasificador de 6 etiquetas porque su clasificación es más diferenciadora.

La limitación más importante del proyecto han sido los requerimientos computacionales del mismo: El entrenamiento era muy costoso computacionalmente tanto en tiempo como en memoria. Los datos también han limitado la obtención de buenos resultados; contar con más modelos de propagación solventaría el desequilibrio entre clases e implicaría una mejor división de los datos en subconjunto de entrenamiento y validación, evitando el filtrado de información existente. Hemos de tener en cuenta que los excelentes resultados de este proyecto en parte se han dado porque los 140 modelos se originaban de sólo 14 modelos de propagación distintos.

Como trabajo futuro se deben explorar otras formas de manejar los desequilibrios entre categorías, así como mejorar la arquitectura de la red neuronal propuesta.

#### Agradecimientos

Este proyecto ha sido parcialmente financiado por: Ministerio de Ciencia e Innovación (PID2019-105032GB-I00), Instituto de Salud Carlos III, y Ministerio de Ciencia, Innovación y Universidades (con fondos FEDER-Fondo Europeo de Desarrollo Regional PI17/01106 y RYC2018-024346B-750), Consejería de Ciencia, Universidades e Innovación de la Comunidad de Madrid a través del programa RIS3 (S-2020/L2-622), EIT Health (Actividad código 19600, el EIT Health está apoyado por el EIT, un organismo de la Unión Europea) y el programa de investigación e innovación Horizonte 2020 de la Unión Europea en el marco de la beca Marie Skłodowska-Curie n° 860974.

#### Referencias

- [1] Colilla, S., Crow, A., Petkun, W., Singer, D. E., Simon, T., & Liu, X. (2013). Estimates of Current and Future Incidence and Prevalence of Atrial Fibrillation in the US Adult Population. *The American Journal of Cardiology*, 112(8), 1142-1147.
- [2] ECG Review. (2021). Atrial Fibrillation ECG Review. Learn the Heart. <https://www.healio.com/cardiology/learn-the-heart/ecg-review/ecg-topic-reviews-and-criteria/atrial-fibrillation-review>
- [3] Mayo Clinic. (2021, October 19). Atrial Fibrillation - Symptoms and Causes. <https://www.mayoclinic.org/diseases-conditions/atrial-fibrillation/symptoms-causes/syc-20350624>
- [4] Guillem, M. S., Climent, A. M., Rodrigo, M., Fernández-Avilés, F., Atienza, F., & Berenfeld, O. (2016). Presence and Stability of Rotors in Atrial Fibrillation: Evidence and Therapeutic Implications. *Cardiovascular Research*, 109(4), 480-492.
- [5] Schotten, U., Verheule, S., Kirchhof, P., & Goette, A. (2011). Pathophysiological Mechanisms of Atrial Fibrillation: A Translational Appraisal. *Physiological Reviews*, 91(1), 265-325.
- [6] Narayan, S. M., Krummen, D. E., Shivkumar, K., Clopton, P., Rappel, W. J., & Miller, J. M. (2012). Treatment of Atrial Fibrillation by the Ablation of Localized Sources: CONFIRM (Conventional Ablation for Atrial Fibrillation with or without Focal Impulse and Rotor Modulation) Trial. *Journal of the American College of Cardiology*, 60(7), 628-636.
- [7] Gulrajani, R. M. (1998). The Forward and Inverse Problems of Electrocardiography. *IEEE Engineering in Medicine and Biology Magazine*, 17(5), 84-101.
- [8] Oster, H. S., & Rudy, Y. (1992). The Use of Temporal Information in the Regularization of the Inverse Problem of Electrocardiography. *IEEE Transactions on Biomedical Engineering*, 39(1), 65-75.
- [9] Attia, Z. I., Noseworthy, P. A., Lopez-Jimenez, F., Asirvatham, S. J., Deshmukh, A. J., Gersh, B. J., ... & Friedman, P. A. (2019). An Artificial Intelligence-enabled ECG Algorithm for the Identification of Patients with Atrial Fibrillation during Sinus Rhythm: A Retrospective Analysis of Outcome Prediction. *The Lancet*, 394(10201), 861-867.
- [10] Xia, Y., Wulan, N., Wang, K., & Zhang, H. (2018). Detecting Atrial Fibrillation by Deep Convolutional Neural Networks. *Computers in Biology and Medicine*, 93, 84-92.
- [11] Muffoletto, M., Qureshi, A., Zeidan, A., Muizniece, L., Fu, X., Zhao, J., ... & Aslanidi, O. (2021). Toward Patient-specific Prediction of Ablation Strategies for Atrial Fibrillation using Deep Learning. *Frontiers in Physiology*, 12, 717.
- [12] Shi, X., Chen, Z., Wang, H., Yeung, D. Y., Wong, W. K., & Woo, W. C. (2015). Convolutional LSTM network: A Machine Learning Approach for Precipitation Nowcasting. *Advances in Neural Information Processing Systems*, 28.
- [13] Cámara, M. Á., Hernández, I., Morgado, E., Guillem, M. S., Climent, A. M., & Barquero, O. (2021). Non-invasive Estimation of Atrial Fibrillation Driver Position with Convolutional Neural Networks and Body Surface Potentials. *Frontiers in Physiology*, 12.
- [14] Gutiérrez, M., Cámara, M. Á., Hernández, I., Guillem, M. S., Climent, A. M. & Barquero, Ó. (2022). Non-Invasive Atrial Fibrillation Driver Localization Using Recurrent Neural Networks and Body Surface Potentials. *2022 Computing in Cardiology*.
- [15] Stridh, M., Bollmann, A., Olsson, S. B., & Leif, S. (2006). Detection and Feature Extraction of Atrial Tachyarrhythmias. *IEEE Engineering in Medicine and Biology Magazine*, 25(6), 31-39.

# Development of strategies for the generation and characterisation of a polymeric scaffold for in vitro bladder reconstruction.

M. Martínez Donázar, J. Paredes Puente, J. Aldazabal

Tissue Engineering Department, TECNUN (University of Navarra), San Sebastian, Spain,  
mmartinez.131@alumni.unav.es, jparedes@tecnun.es, jaldazabal@tecnun.es

## Abstract

*In the field of tissue engineering associated with regenerative medicine, numerous techniques for organ reconstruction are being explored. In this paper, we present an alternative method for the creation of a scaffold to replace the functions of a bladder, in which we have sought to imitate the elasticity that characterizes the walls of this organ, thanks to a specific pattern in the designed structure.*

## 1. Introduction

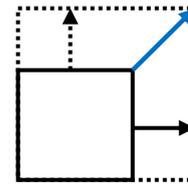
In 2021, 20,613 cases of urinary bladder cancer were diagnosed in Spain, making it the fifth most common cancer detected during that year [1]. Despite its relatively low mortality rate, cases of major normal tissue invasion by cancerous cells are commonly subjected to cystectomy treatment, where the organ (or part of it) is removed in order to eliminate the damaged tissue. In such cases, the patients' quality of life is markedly diminished, as the current treatment consists on placing an external bag to store the urine. Due to the discomfort derived, the possibility of fabricating a neo-bladder has been recently investigated, in order to try to bring normality back to these individuals. With this in mind, thanks to tissue engineering, several approaches that can be useful for regenerating this type of organ have arisen. In the case of this work, a new option was studied to create an elastic matrix that could simulate the walls of the bladder, where the appropriate cells are expected to be cultivated so that they could cover the scaffold, resulting in an implantable biomaterial that fulfils the functions of the organ.

## 2. State of the art

A review of the existing literature reveals several examples of various techniques in the field of tissue engineering used to develop an implantable bladder. One such example is the study carried out a few years ago by Atala, in which a 3D bladder was created for a child using bioprinting techniques [2]. This was actually the first 3D printed organ to ever be implanted. Years later, it has been confirmed that the transplant was successful, making this a promising field of research. Atala developed several studies around this organ, mainly creating it by using a 3D printer in order to do so [3].

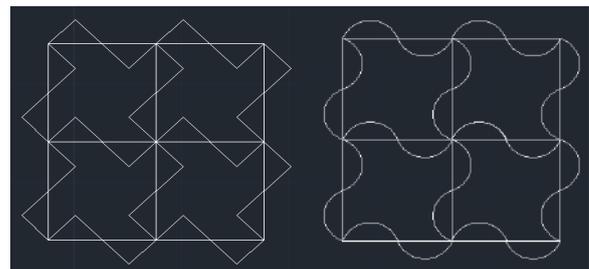
This work tries to take a different route in order to achieve a similar goal, as it will focus exclusively on reproducing the walls of this organ. The key to this is to

create a pattern that turns a structure made of a stiff material into a softer one. So, keeping with the literature, around a decade ago, certain materials and shapes were found to have a negative Poisson's ratio. This means that, when a compressive or tensile stress is exerted in one direction, the material reacts on the perpendicular direction with the same deformation, instead of the opposite, as is normally the case. This phenomenon occurs in materials known as auxetic (Figure 1).



**Figure 1.** Generalisation of the behaviour of an auxetic structure after applying a tensile stress on the horizontal axis, resulting in an elongation in both directions of the material, and also increasing its area.

With this in mind, structures with these characteristics were sought for the development of the work, among which those formed by zigzagging segments and/or sinusoidal segments seemed to be the most preferable, not only because of their simplicity, but also because of their usefulness.



**Figure 2.** Two of the possible auxetic structures to be analysed in this work are shown. On the left, four unit cells composed of zigzagging segments are sketched in AutoCAD. On the right, same structure, but with sinusoidal segments.

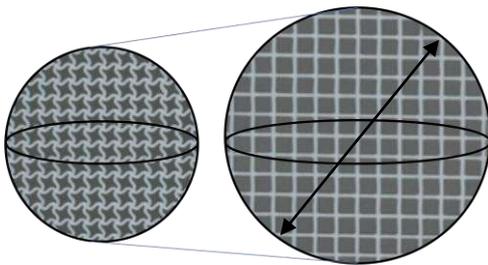
As can be seen in Figure 2, and as can be deduced from the literature, these two structures react with auxetic behaviour to tensile and compressive stresses and are therefore promising for this work.

Because of their auxetic behaviour, a printed mesh formed from a version of these structures could serve as

an innovative approach to give the needed elasticity to the walls of an in-vitro bladder.

### 3. Objectives

The main objective of this work was to imitate as closely as possible the elastic structure that characterizes the walls of a bladder. To do this, two issues had to be solved: on one hand, it was necessary to design and develop an auxetic pattern that would give the internal elasticity to the assembly. On the other hand, the external structure of the biomaterial should also imitate another important characteristic of the organ to be developed: its spherical shape. The combination of both aspects was expected to give rise to a sphere-shaped auxetic material (Figure 3).



**Figure 3.** Final design sketch. On the left, a representation of the sphere at rest, and on the right a representation of the sphere when stretched, simulating internal pressure, as the one a bladder would experience during filling.

Possible approaches to achieve both the auxetic internal pattern and the final spherical shape were studied. The structure was manufactured using a 3D printer, taking into account the facilities in the university's laboratory. Given this, a laminar structure made of polycaprolactone, biodegradable polyester better known as PCL, and compatible with the bioprinter, was devised, generating an "origami" sphere also possessing a specific auxetic internal design, thus expected to meet the objective of the work.

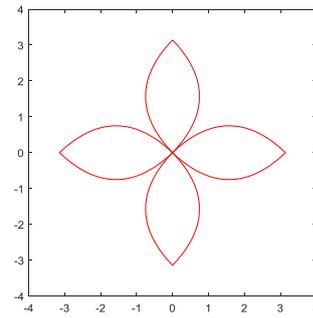
### 4. Materials and methods

The development of the work was mainly based on the fabrication of the desired product by means of 3D printing. This machine operates by CNC (Computerized Numerical Control), and this type of system is commonly controlled with a computer language named G code.

In view of the two points that had to be solved to achieve the final objective, which were the internal and external structure of the product, the Matlab computer program was used to generate codes that, converted to G-code, could give rise to the target design. This developed program was very useful because of its simplicity and ease to change parameters and perform several tests with the same code, and because it gave the possibility to preview the designs before converting them to G-code.

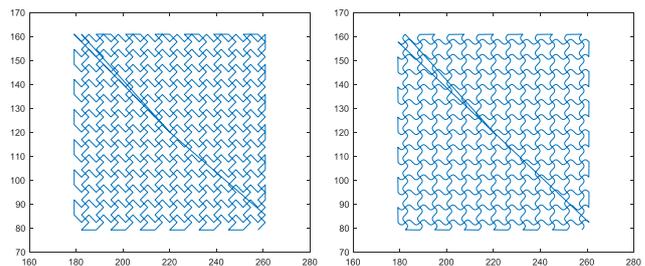
As for the external structure, in order for it to generate a sphere by means of folds and sutures (origami-like), the area of a sphere of radius  $r$  was divided into four equal

parts, giving rise to a flower-shaped structure with four petals on a 2D plane, as shown in Figure 4. Some extra material was added along the borders of the petals to avoid the formation of edges when suturing.



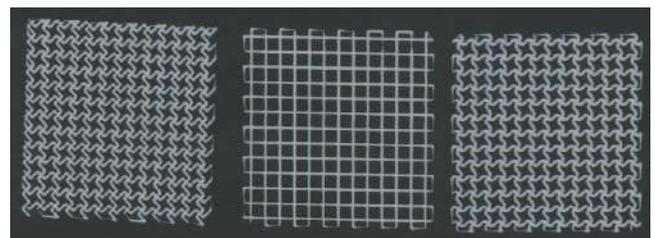
**Figure 4.** Design of the external structure in Matlab, without including the extra material in the petals, to appreciate the 2D shape that would result in a 3D sphere of radius  $r=1\text{cm}$ .

Additionally, in order to study the most suitable internal structure, several auxetic patterns were tested. With a linear pattern as a control sample, various meshes (both zigzag and sinusoidal, Figure 5) needed to be created in the same Matlab program to analyze their elasticity. Following the generic designs proposed in the state of the art section, an evaluation to see which of the two types of patterns was more suitable to simulate the elasticity of the bladder walls was carried out.



**Figure 5.** Design of two meshes in Matlab with the patterns to be studied, generated with the same conditions and parameters to make them comparable. The one on the left is the zigzag and the one on the right is the sinusoidal.

In order to decide which pattern was more suitable for the study, tensile tests were performed for both types of samples, plus the control sample (Figure 6). In addition, the effect of various parameters on the elasticity of the meshes was also studied in order to check the optimum printing conditions to achieve the objective. These parameters are the pore size, the amount of extruded material per millimetre and the angle at which the traction is made with respect to the vertical axis.



**Figure 6.** Printed mesh samples for the tensile test. From left to right: zigzag sample, control sample and sinusoidal sample.



**Figure 7.** Sample of a sinusoidal mesh located in the tensile testing machine.

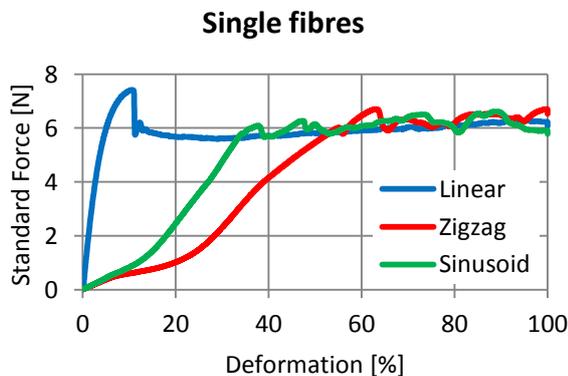
Different tensile tests were then performed to study the different stress-strain curves obtained, and to be able to draw useful conclusions for the project.

In addition, an electrospinning treatment was performed to the final product to create an outer layer.

## 5. Results and discussion

Considering that permanent deformation of the fibres was to be avoided, the most important part of the stress-strain curve was the initial slope. This linear part corresponds to the elastic behaviour of the material. Once the elastic limit was reached, the fibre was deformed in a permanent way, and would no longer return to its initial shape, disabling the main bladder function, desired to mimic.

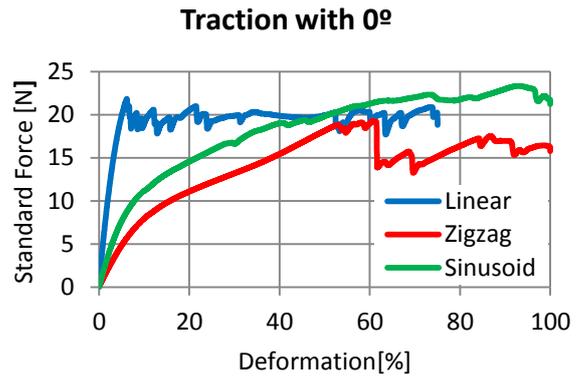
First, the elasticity of individual fibres was analyzed for each of the parameters studied (Figure 8).



**Figure 8.** Stress-strain curve obtained for the different single fibres.

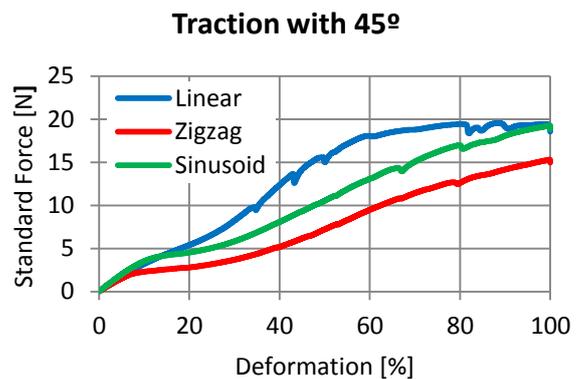
Figure 8 shows a higher elastic deformation for the zigzag and sinusoidal fibres. However, it should be noted that the zigzag fibre had an anomalous behaviour, showing a certain rotation in the fibre, perhaps explaining this marked curve prior to the rise due to plastic deformation, maybe due to its sharp angles.

In order to reproduce the reality closer, instead of a single fibre, the mechanical behaviour of these patterns was evaluated, this time in mesh format.



**Figure 9.** Stress-strain curve obtained for the different meshes, carried out with 0° traction.

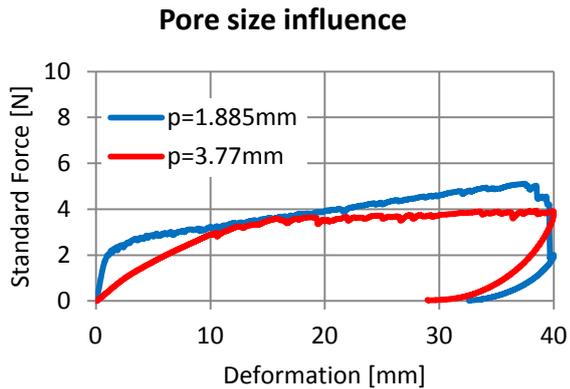
There is no difference between the percentage of deformation at the EL of the zigzag and sinusoid meshes (although there is a difference in the force required), but there is a difference with respect to the straight mesh, being this last one less elastic, as expected (Figure 9).



**Figure 10.** Stress-strain curve obtained for the different meshes, carried out with 45° traction.

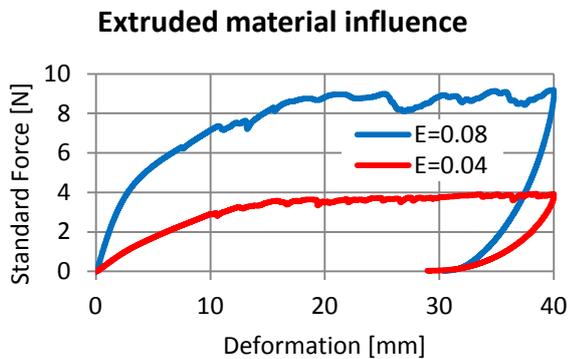
In general, it can be seen how the traction angle affects the stiffness of the assembly, increasing in this case (Figure 10) the percentage of deformation in the elastic region, compared to Figure 9. In addition, the behaviour of the linear mesh is surprising, perhaps due to the lack of sharp angles that affect this curve, as can be seen in Figure 8, although comparing the three in general it seems that the sinusoidal mesh achieves the highest percentage of deformation between the two under study.

Next, we went on to study the influence of the rest of the parameters on the elasticity of the printed samples, but only on the sinusoidal pattern, given that this was the one chosen to continue with the experiments. This was due to its good elastic behaviour, and also to it being advantageous when it comes to programming. That said, larger pore size and smaller amount of extruded material were expected to favour the elasticity of the assembly.



**Figure 11.** Stress-strain curve obtained for both small pore size (1.885mm width, represented by the blue line) and big pore size (3.77mm, width, represented by the red line).

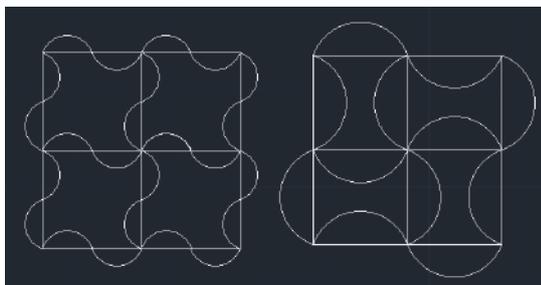
Looking at Figure 11, as expected, a larger pore size favours a higher percentage of reversible deformation.



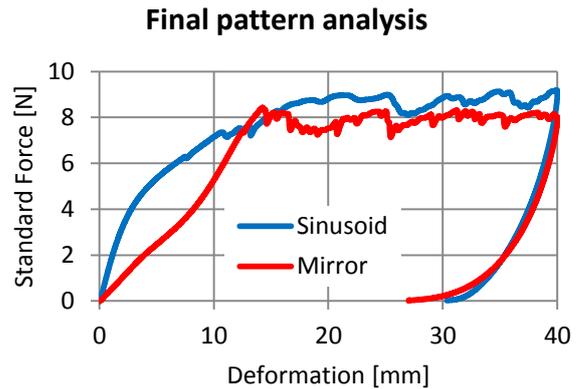
**Figure 12.** Stress-strain curve obtained for both extruded material quantities. Blue line represents the sample with more extruded material per mm, and the red one represents the other case.

In this case, the mesh made with more extruded material needs more force to deform, and appears less elastic than the other sample, as expected.

After this, within the selected pattern type, it was studied which of the two sinusoidal patterns shown in the image would be better (sinusoidal or mirror sinusoidal).

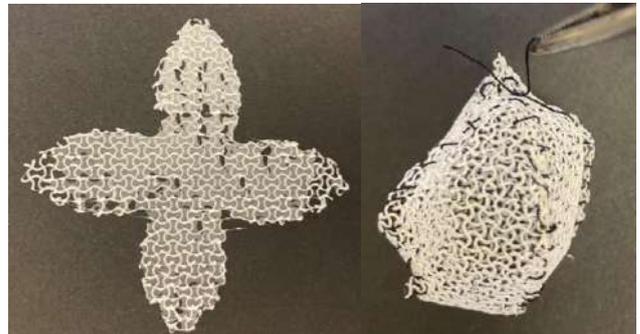


**Figure 13.** AutoCAD sketches of both patterns. On the left, the sinusoidal one and on the right the mirror sinusoidal. one



**Figure 14.** Stress-strain curve for both sinusoidal patterns.

As can be seen, the mirror sinusoid pattern has a higher elastic deformation, so it will be the one chosen for the final print. Once the electrospinning layer has been applied, the result will be as follows, bearing in mind that this process is still to be optimised.



**Figure 15.** Real images of the final design, with  $r=15\text{mm}$

## 6. Conclusions

In conclusion, an elastic structure made of the referred to as mirror sinusoidal pattern could work as an implantable mesh that could imitate the shape and functions of a human bladder wall, given that it was the best option among the ones considered for this work. Therefore, in the absence of cells and signals, a biocompatible scaffold with a certain range of elastic deformation has been created, obtaining some very promising results.

Also, these acknowledgements associated to auxetic materials working as printed scaffolds for tissue engineering are very encouraging if they were to be applied in other areas of investigation in the future.

## References

- [1] Web page: Sociedad Española de Oncología Médica ([www.seom.org](http://www.seom.org)). 'Las cifras del cáncer en España, 2021'
- [2] Pdraig Belton. 'A new bladder made from my cells gave me my life back'. BBC News, 11 September 2018
- [3] A.Atala. Tissue engineering of human bladder. *British Medical Bulletin* 2011

# Análisis de un Algoritmo de Deep Learning para Detectar Picos R en Electrocardiogramas

B. Lafuente Alcázar<sup>1</sup>

<sup>1</sup> Departamento de Ingeniería Electrónica, Universitat Politècnica de València (UPV), Valencia, España, bealaal@etsii.upv.es

## Resumen

Las señales biomédicas son resultado de la actividad eléctrica de la membrana celular, y de los cambios en su voltaje. Los electrocardiogramas (ECGs) son un tipo concreto de bioseñales, y constituyen el método principal para diagnosticar enfermedades cardíacas, ya que la mayoría de estas patologías se manifiestan como alteraciones tanto a corto como a largo plazo en la morfología del electrocardiograma. Sin embargo, el diagnóstico basado únicamente en la inspección visual de estas señales es potencialmente impreciso, pues contienen información y particularidades indefectibles para el ojo humano. Los algoritmos de inteligencia artificial son una herramienta de gran utilidad para extraer características relevantes y eliminar la subjetividad en el análisis de electrocardiogramas. Por ello, un algoritmo de deep learning ha sido entrenado para detectar los picos R en señales de ECG de manera automática, siguiendo una estrategia nunca utilizada hasta la fecha. Su correcto funcionamiento ha sido probado en diferentes bases de datos, poniendo especial énfasis en señales ruidosas o considerablemente alteradas. Se espera que ese código logre aumentar la efectividad y capacidad de los profesionales médicos a la hora de detectar alteraciones en la actividad cardíaca de los pacientes.

## 1. Introducción

### 1.1. Señales de electrocardiograma

El corazón contiene células eléctricas especializadas, que generan potenciales de acción durante la fase de contracción y pueden conducir los impulsos eléctricos a través del miocardio. Puesto que el cuerpo humano es equivalente a un medio puramente resistivo [1] las señales eléctricas originadas en el corazón se extienden hasta la superficie corporal y pueden ser captadas por electrodos en la piel del paciente. Las ondas, intervalos y segmentos característicos de un ECG pueden observarse en la figura 1.

La onda P representa la despolarización de las aurículas, el complejo QRS está relacionado con la despolarización ventricular y la onda T refleja la repolarización de los ventrículos.

Las duraciones fisiológicas de cada uno de los intervalos del ECG están indicadas en la tabla 1. La presencia de isquemia en el miocardio, la insuficiencia cardíaca o cualquier otra patología referente al corazón provocará una alteración en dichas duraciones o en la dirección de las ondas del ECG. Por ello, el diagnóstico basado en el

ECG consiste en analizar las posibles alteraciones en la forma normal de esta señal.

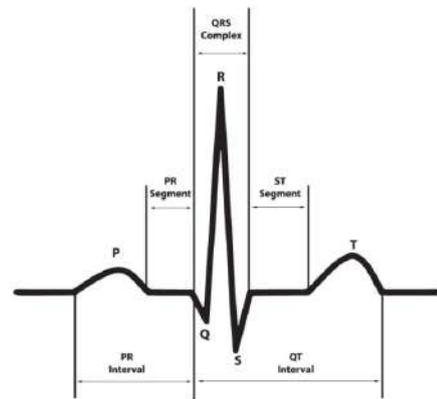


Figura 1. Morfología típica de una señal de ECG

Porción del ECG	Intervalo (s)
Intervalo PR	0.12 - 0.20
Intervalo QRS	0.06 - 0.10
Segmento ST	0.05 - 0.15
Intervalo RR	0.60 - 1.00

Tabla 1. Duraciones normales de las porciones del ECG

### 1.2. Aprendizaje automático

El término Aprendizaje Automático (AA) o Machine Learning (ML) refiere a una rama de la Inteligencia Artificial que engloba algoritmos capaces de clasificar datos y hacer predicciones sin estar específicamente programados para ello.

El proceso para desarrollar un algoritmo de AA es complejo y consta de una serie de pasos intermedios. En primer lugar, se compone una base de datos y se identifican las variables de interés. A continuación, esta base de datos se divide en tres subgrupos de datos: entrenamiento, test y validación (generalmente en proporción 60%, 20% y 20%, respectivamente) [2]. El conjunto de entrenamiento se usa para ajustar el modelo a los atributos de la base de datos. El conjunto de test permite comprobar el funcionamiento del modelo durante

el entrenamiento y, finalmente, el grupo de validación permite evaluar las clasificaciones que proporciona el modelo, usando datos diferentes a los usados durante el entrenamiento.

Así, cuando en un modelo de AA se introducen una serie de datos y sus referencias (o etiquetas), éste es capaz de adaptarse para alcanzar los resultados deseados. También es capaz de generalizar esta configuración *aprendida*, de manera que cuando se le presenta nueva información, sin etiquetar, puede proporcionar una vez más las predicciones requeridas.

### 1.3. Aprendizaje automático en cardiología

En los últimos años, el AA ha ganado popularidad en el ámbito de la salud, imponiéndose como una estrategia interesante para reducir la variabilidad en las decisiones médicas, minimizar el error diagnóstico y optimizar el cuidado del paciente. Asimismo, gracias a las historias clínicas electrónicas, gran cantidad de datos sanitarios están disponibles a nivel global. Dicha cantidad de información puede resultar compleja para ser analizada manualmente, pero es fácilmente procesable con algoritmos de AA.

El campo de la cardiología es una de las áreas que más puede beneficiarse del uso del AA. Las enfermedades cardiovasculares son la mayor causa de muerte en el mundo [3], y alrededor de 80% de las muertes súbitas cardíacas son debidas a arritmias o latidos irregulares [4]. La principal herramienta diagnóstica es el electrocardiograma (ECG). El ECG permite medir la actividad eléctrica del corazón y detectar ritmos anormales. Un cardiólogo es capaz de diagnosticar un amplio rango de patologías basándose en las alteraciones observables en la morfología del ECG, pero la interpretación humana es subjetiva y proclive a errores. Por otra parte, las señales ECG pueden tener una duración muy prolongada. Un ejemplo de ello son las señales recogidas con dispositivos ambulatorios Holter, que pueden captar y almacenar la actividad cardíaca de un paciente durante horas o incluso días. La extensa prolongación de estos ECGs complica el diagnóstico visual y retarda la detección de enfermedades.

Por todo ello, los algoritmos basados en AA están destinados transformar el ámbito cardíaco, puesto que son capaces de computar gran cantidad de datos en tiempos muy reducidos y pueden proporcionar diagnósticos tempranos y precisos.

## 2. Materiales y métodos

Para este trabajo, se ha utilizado un modelo de AA que combina una arquitectura de U-Net y bloques residuales Inception. Dicho modelo pertenece a un subgrupo particular del AA: el *Deep Learning* o aprendizaje profundo. Los algoritmos de esta clase permiten utilizar señales sin preprocesar, y son capaces de manejar eficientemente grandes bases de datos.

Como puede observarse en la Figura 2, el modelo escogido consta de un proceso inicial, conocido como *encoder* y una segunda etapa, llamada *decoder*.

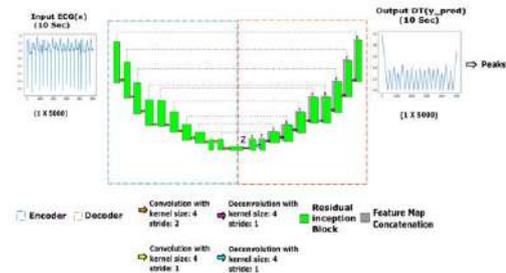


Figura 2. Arquitectura U-Net [5]

En primer lugar, se obtiene la transformada distancia (TD) de cada señal. La TD es una función que indica la distancia de cada muestra al pico R más próximo. Así, se obtiene:

$$X = (x_1, y_1), (x_2, y_2) \dots (x_m, y_m) \quad (1)$$

Donde  $x_i$  es un cierto fragmento de ECG compuesto por 5000 muestras e  $y_i$  es su TD.  $x_1, x_2, \dots, x_m$  constituyen los datos introducidos en la red.

Estos datos son muestreados a través de 8 capas, donde se aplican filtros convolucionales de 1 dimensión. El número de filtros aumenta de 64 hasta 1024, a partir de donde se mantiene constante. Además, el modelo contiene también capas de normalización, funciones de activación ReLU y bloques residuales. Estos últimos garantizan una convergencia rápida del modelo y permiten evitar el problema del desvanecimiento del gradiente. La estructura de estos bloques residuales es ilustrada en la Figura 3.

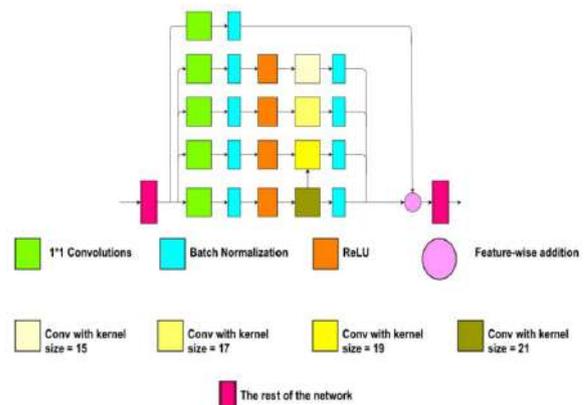


Figura 3. Estructura de los bloques residuales Inception usados en el modelo [5]

En resumen, el vector introducido  $x_i$  es muestreado y comprimido en el *encoder* en un vector de características  $z_i$ . Posteriormente, en la etapa de *decoder*, este vector es interpolado y se obtiene el resultado del modelo,  $y_{pred}$ .

Las operaciones que la red lleva a cabo son:

$$z_i = F_1(x_i; \theta 1) \quad (2)$$

$$y_{pred} = F_2(x_i; \theta 2) \quad (3)$$

Donde  $F_1$  y  $F_2$  representan el *encoder* y el *decoder*, con parámetros  $\theta 1$  y  $\theta 2$ , respectivamente. Los parámetros de la red son optimizados minimizando la siguiente función de pérdidas SmoothL1 entre  $y_{pred_i}$  e  $y_i$ :

$$L(x) = \sum_{i=1}^m \text{SmoothL}_1(y_{diff}) \quad (4)$$

$$y_{diff} = y_i - y_{pred_i} \quad (5)$$

$$\text{SmoothL}_1(y_{diff}) = \begin{cases} 0.5 (y_{diff}^2) & \text{if } \text{abs}(y_{diff}) < 1 \\ \text{abs}(y_{diff}) - 0.5 & \text{otherwise} \end{cases} \quad (6)$$

Los parámetros son inicializados aleatoriamente y el optimizador Adam es usado para minimizar las pérdidas entre la TD predicha ( $y_{pred}$ ) y el mapa de distancias original ( $y_i$ ). El modelo ha sido entrenado durante 100 épocas, con una tasa de aprendizaje inicial de 0.05.

Todo el proceso se ha realizado en el entorno Spyder, basado en Python 3.9.7.

### 2.1. Datos de entrenamiento

En este trabajo se ha entrenado al modelo siguiendo una técnica alternativa a la utilizada normalmente: se ha tomado como conjunto de entrenamiento una combinación de ECGs de diferentes bases de datos. A este método, se le ha acuñado como *entrenamiento mixto*. Se hipotetizó que de esta manera el algoritmo aprendería características más diversas, y sería capaz de reconocer mejor nuevos datos una vez entrenado.

Así pues, durante la etapa de entrenamiento se han utilizado señales procedentes del *2nd China Pysiological Signal Challenge (CPSC 2019)* [6] y de la base de datos de arritmias del MIT-BIH [7]. La primera contiene señales de 5000 puntos, muestreadas con una frecuencia de muestreo de 500 Hz. La segunda contiene señales de mucha mayor duración, por lo que han sido divididas en fragmentos de 5000 puntos, los cuales han sido considerados como ECGs independientes. La frecuencia de muestreo de las señales del MIT-BIH es 360 Hz.

Ambas bases de datos contienen también anotaciones sobre la posición de los picos R, que han sido utilizadas como referencia para entrenar al modelo.

### 2.2. Datos de validación

Para evaluar la capacidad de detección del modelo una vez entrenado, se han introducido datos procedentes de 3 fuentes distintas. Por una parte, se han utilizado señales del CPSC 2019 y del MIT-BIH que no habían sido usadas durante el entrenamiento. Asimismo, se ha introducido también una nueva base de datos sobre la que la red no había sido entrenada: la base de datos Europea ST-T [8]. La frecuencia de muestreo en este grupo de señales es de 500 Hz.

## 3. Resultados

Para examinar el funcionamiento del modelo, se ha elaborado un código adicional donde se utiliza las referencias de cada base de datos y las predicciones proporcionadas por el algoritmo para calcular la precisión, exhaustividad y puntuación F1 de acuerdo con las ecuaciones 7, 8 y 9.

$$\text{Precisión} = \frac{VP}{VP + FP} \quad (7)$$

$$\text{Exhaustividad} = \frac{VP}{VP + FN} \quad (8)$$

$$\text{Puntuación F1} = 2 * \frac{\text{Precisión} \times \text{Exhaustividad}}{\text{Precisión} + \text{Exhaustividad}} \quad (9)$$

Donde  $VP$ = verdaderos positivos,  $FP$  =falsos positivos, y  $FN$  =falsos negativos.

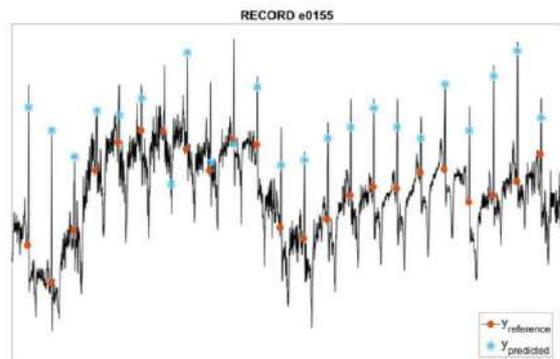
Se ha permitido una tolerancia de 40 muestras (80 ms, 110 ms y 160 ms en las bases de datos CPSC 2019, MIT-BIH y Europea, respectivamente) a la hora de considerar los picos como correctamente detectados.

Los resultados obtenidos quedan resumidos en la siguiente tabla:

	CPSC 2019	MIT-BIH	Europea
<b>Precisión</b>	92.6%	99.61%	100%
<b>Exhaustividad</b>	96.3%	97.87%	99.05%
<b>Puntuación F1</b>	94.4%	98.73%	99.52%

**Tabla 2.** Métricas de evaluación para las tres bases de datos utilizadas

Se ha puesto además especial atención en aquellas señales considerablemente afectadas por interferencias, como la que puede observarse en la Figura 4. Comparando los picos R de referencia (puntos rojos) con las predicciones del modelo (asteriscos azules), puede observarse que el correcto funcionamiento del algoritmo se mantiene incluso en situaciones desfavorables.



**Figura 4.** Detección del modelo en el archivo e0155 de la base de datos europea

Los resultados del modelo en cada una de las bases de datos son prometedores, y permiten afirmar el buen

funcionamiento de la red. Quizá el hecho más interesante es que las mejores métricas son obtenidas para la base de datos europea, en la cual el modelo no había sido entrenado. Esto demuestra la capacidad de generalización del código.

Por otro lado, se hace patente que la frecuencia de muestreo no afecta a las detecciones realizadas por el modelo. Además, se ha comprobado la capacidad de la red de trabajar con señales de baja calidad o afectadas por interferencias. Estas características del modelo hacen posible el uso de las señales ECG sin necesidad de aplicar un preprocesado previo.

#### 4. Discusión

En un mundo donde la principal causa de muerte son las patologías cardíacas, es imprescindible poseer herramientas de diagnóstico que permitan identificar de manera rápida y precisa cualquier irregularidad en el ritmo cardíaco de los pacientes. El análisis de la morfología del ECG es una manera rápida, precisa y no invasiva de analizar la función cardíaca. Sin embargo, el tradicional análisis visual es susceptible a errores, además de poco eficiente, y requiere de una considerable cantidad de tiempo. Una solución como la discutida en el presente estudio, basada en el aprendizaje automático, permite dar solución a todas estas cuestiones.

El modelo desarrollado puede manejar gran cantidad de información en un espacio de tiempo muy reducido, y consigue detectar propiedades de la señal incluso cuando su morfología aparece alterada debido a la presencia de ruido o patologías, algo de lo que el ojo humano no siempre es capaz. El entrenamiento mixto hace posible que el modelo aprenda a reconocer ECGs de diferente procedencia, y mejora sus predicciones cuando nuevos datos son introducidos.

Las líneas futuras de este proyecto consisten en tratar de incorporar el algoritmo descrito en un microchip, con tal de lograr detecciones en tiempo real.

En definitiva, se ha presentado una solución novedosa para la aplicación de la inteligencia artificial y el aprendizaje automático en el campo de la cardiología, que proporciona una manera rápida y exacta de identificar picos R, incluso en condiciones desfavorables. Esta información puede ser utilizada para calcular la variabilidad de la frecuencia cardíaca, un importante predictor de arritmias. Finalmente, una pronta identificación de estos ritmos anormales permitirá actuar de manera inmediata e identificar el tratamiento más adecuado para cada paciente, maximizando sus posibilidades de supervivencia y optimizando el pronóstico.

#### Agradecimientos

Por último, agradecer a Amit Nimunkar por su tutela y a al departamento de ingeniería biomédica de la universidad de Wisconsin-Madison por darme la oportunidad de terminar mis estudios con ellos y elaborar el presente trabajo.

#### Referencias

- [1] R. A. Malkin and B. K. Hoffmeister, "Mechanisms by which AC Leakage Currents Cause Complete Hemodynamic Collapse Without Inducing Fibrillation," *J Cardiovasc Electrophysiol*, vol. 12, no. 10, pp. 1154–1161, Oct. 2001, doi: [10.1046/j.1540-8167.2001.01154.x](https://doi.org/10.1046/j.1540-8167.2001.01154.x).
- [2] P. I. Dorado-Díaz, J. Sampredo-Gómez, V. Vicente-Palacios, and P. L. Sánchez, "Aplicaciones de la inteligencia artificial en cardiología: el futuro ya está aquí," *Revista Española de Cardiología*, vol. 72, no. 12, pp. 1065–1075, Dec. 2019, doi: [10.1016/j.recesp.2019.05.016](https://doi.org/10.1016/j.recesp.2019.05.016).
- [3] World Health Organization. (2021, June 11). Cardiovascular diseases (CVDs). Recuperado el 9 de julio de 2022, de <https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-cvds>
- [4] R. Mehra, "Global public health problem of sudden cardiac death," *Journal of Journal of Electrocardiography*, vol. 40, no. 6, pp. S118–S122, 2007.
- [5] S. Vijayarangan, V. R., B. Murugesan, P. S.P., J. Joseph, and M. Sivaprakasam, "RPnet: A Deep Learning approach for robust R Peak detection in noisy ECG," in *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, Montreal, QC, Canada, Jul. 2020, pp. 345–348. doi: [10.1109/EMBC44109.2020.9176084](https://doi.org/10.1109/EMBC44109.2020.9176084)
- [6] H. G. e. al, "An Open-Access ECG Database for Algorithm Evaluation of QRS Detection and Heart Rate Estimation," *j med imaging hlth inform*, vol. 9, no. 9, pp. 1853–1858, 2019.
- [7] G. B. Moody and R. G. Mark, "The MIT-BIH database," 1992. [Online]. Available: <https://physionet.org/content/mitdb/1.0.0/>.
- [8] A Taddei 1, G Distante, M Emdin, P Pisani, G B Moody, C Zeelenberg, C Marchesi, "The European ST-T database: standard for evaluating systems for the analysis of ST-T changes in ambulatory electrocardiography," *European Heart Journal*, vol. 13, no. 9, pp. 1164–72, 1992.

# Biomecánica

---

Jueves, 24 de noviembre de 2022

# Design and development of a robotic hand for finger spelling

C. Cabezaolías, R. Raya, V. Costa

Departamento de Tecnologías de la Información, Universidad San Pablo-CEU, CEU Universities, Madrid, España, cm.cabezaolias@usp.ceu.es, {rafael.rayalopez, vanina.costacortez}@ceu.es

## Abstract

*Nowadays, there are 70 million deaf people in the world who suffer from total or partial hearing loss. The communication method most used between them is sign language. The problem emerges when they have to deal with the hearing world. The number of hearing people that dominate sign language is reduced; therefore, communication with the deaf community can be complicated. The development of new communication tools and devices is of great relevance to contribute to the integration of people with disabilities. This project aims to design and develop a prototype of a robotic hand that can reproduce the letters of the sign alphabet based on voice command. The words said by the emitting person to a phone are sent via Bluetooth to a microcontroller that controls the hand where a sign will be performed. The word is executed letter by letter. The finger and wrist movements are carried out by servo motors controlled by an Arduino board. The final product must be low-cost and easily transportable. The design requirements for this project prioritized low cost and portability.*

## 1. Introduction

Oral communication is the method most used to get expressed. Not everything is about speaking, but also understanding and listening to others, for this reason the sense of hearing is essential in human communication. Deafness is defined as the disease that makes people unable to hear in a total (anacusis) or partial (presbycusis) way. They use sign language to communicate between them. Every time individuals with hearing disability need to communicate with the hearing world, they must look for alternatives, such as professional interpreters or family members who can act as interpreters. This scales down their autonomy and makes them feel more dependent. Despite a high percentage of the population in the world needs sign language to communicate low number of people is familiarized with it.

We live in a world that is changing about the inclusion of people with disabilities, favoring their integration and improving their quality of life through different solutions. Engineering must be at the service of society more than ever by providing solutions in this area. Robotics has given the possibility of creating new technology to increase the quality of life and bring new opportunities closer to patients in a more accessible way.

Focusing on hearing disability, some important solutions have been developed using robotics. Humanoids have been developed with the possibility of communicating with sign language: "Aiko Chihira" [1], "Asimo" [2], and "TEO Humanoid" [3]. There are also two robotic hands that

perform fingerspelling which is a technique that uses sign language alphabet to spell words. One was developed by the California Polytechnique State University [4], and the other is the Antwerp Sign Language Actuating Nod (ASLAN) Project" [5]. However, these solutions are too expensive and bulky. Therefore not so accessible for the disabled user.

In this project robotics will be used in the design and manufacture of a robotic hand that executes the letters of the American Sign Language (ASL) alphabet from voice commands. The fingers and the wrist of the hand will be moved with threads rolled up in servomotors. The control of the motors must be done through a programmable electronic board, in this case Arduino, with which it will communicate via Bluetooth for the voice commands. The working flow is shown in Figure 1. The aim is to get a low-cost, lightweight, easily portable, and reproducible solution, fulfilling three main requirements: functionality, accessibility, and portability.



Figure 1. Project working flow

## 2. Methodology

We have used a previously developed 3D design model [6], chosen over other existing ones based on the criteria of movement capacity. On this design and during the assembly stage, small modifications were made to adapt the model to the needs of the project.

### 2.1. Robotic hand movements

In order to reproduce a device inspired by a real hand and that fulfills the objective functionality of this project, prior to its development it is necessary to carry out a study of the physiology and biomechanics of the hand, reproducing a design similar in size and structure and function to that of a human hand.

Starting with the wrist, the robotic wrist has two different pair of movements. The first one is extension and flexion; this is wrist's ability to move the hand up and down. It was not developed at first and as it is essential for some sign, it was created afterward making the wrist able to flex and extend. The second one is the capability of rotating the wrist to the right and to the left, called is pronation and supination. The printed hand already had this possibility.

Alluding to the finger movements, this hand presents the joints for the flexion and extension finger phalanges motion. Moreover, this hand also has the presence of an extra joint that makes the thumb able to move itself to the palm's inner part.

### 2.2. 3D printing

3D printing achieves a low-cost and portable solution. This is due to the low-cost manufacturing process and inexpensive materials used in this technique. Moreover, with a low infill percentage and a low-density material the solution will be light and easily transportable. Polylactic acid (PLA) is the material that has been used because it fulfills these characteristics. The infill percentage used is 20%.

The hand design is divided into four parts: hand (fingers-palm assembly), wrist and forearm. Before proceeding with 3D printing, all the pieces were labeled with numbers to follow the traceability of the printing and subsequent assembly process.

### 2.3. Assembly

The hand, wrist and forearm follow distinct and independent assembly lines. They can be assembled in parallel and at the end join the three parts together.

#### Hand assembly

Once the pieces are labeled and printed, the steps that must be followed are shown in figure 2.

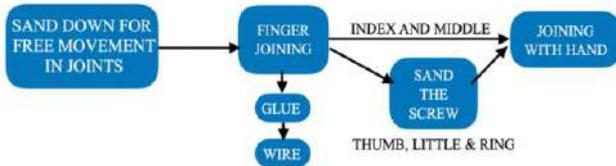


Figure 2. Finger assembly scheme

It is important to notice that the first two steps are the same for all the fingers. In these ones the pieces will be sanded down and joined together with glue or a wire depending on the type of piece. Afterwards, the assembly procedure will differ between index-middle finger group and the thumb-ring-little finger group in the last step because this second group have an extra piece before joining to the palm.

The up and down finger movements is performed thanks to the threads that goes through the hand joints. For each finger we are going to have two threads: one for flexion, other for extension. The flexion one goes through the holes of the joint in the anterior face. The extension one does the same but in the opposite face, the posterior. Both procedures are shown in Figure 3. Then, they will be led to the motors through the wrist, where holes that simulates the human carpal tunnel are place. Once in the forearm, the threads of the same finger are rolled up to the base of its motor finger with opposite directions. The thumb is the only finger that has the carpometacarpal movement, the presence of this joint will make the thumb able to move itself to the palm's inner part. Its threads are place inside the piece that is closest to the palm, in the wrist. They lead to its motor also throw the wrists.

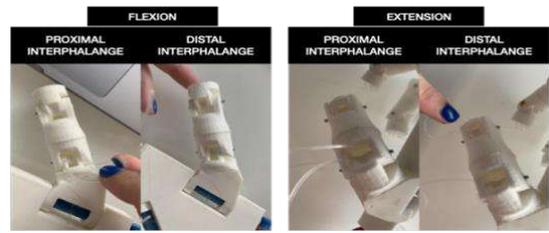


Figure 3. Placing inside the finger of threads that perform flexion and extension movement

#### Wrist assembly

The steps for the wrist assembly are shown in Figure 4.

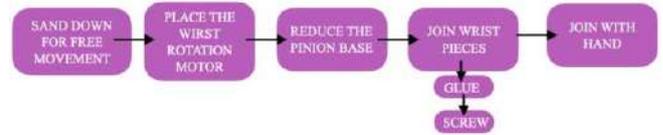


Figure 4. Wrist assembly scheme

The wrist is going to be able to perform two different movements. The first one is the extension and flexion of the wrist, rising the hand up and down, for this movement threads are also going to be used.

The second movement is the wrist rotation, performed with a toothed pinion- gear assembly. The mechanism is made up by a pinion located in the servomotor and the big gear in the wrist (see Figure 5).

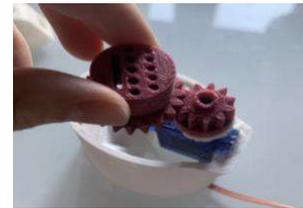


Figure 5. Wrist gear assembly

#### Forearm assembly

The steps for the forearm assembly are shown in Figure 6.

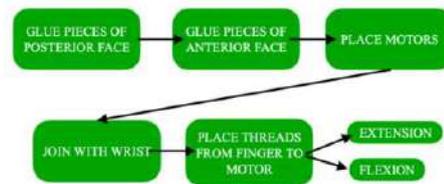


Figure 6. Forearm assembly scheme

All the motors for the flexion and extension of fingers and the wrist are placed in the posterior piece of the forearm face. They have to be placed in the same axis that the movement will be performed in because the thread has to be pulled in the same direction. For instance, the flexion and extension movements of the fingers are when the finger goes up and down in the palm, the Y axis of the palm, the motors are located in the same axis but in the forearm. However, the motor in charge of the carpometacarpal joint of the thumb has different movement, it brings the thumb closer or move it away from the palm. As this movement is done in the X axis, it is the only one that is located in this axis (see Figure 7).

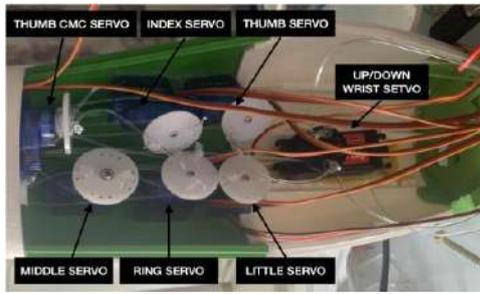


Figure 7. Motors distribution.

## 2.4. Control architecture

The hardware is made up by 3 main structures: eight electrical motors of continuous rotation, an Arduino Nano, and a Bluetooth module HC-05. In Figure 8 the scheme of the hardware is shown.

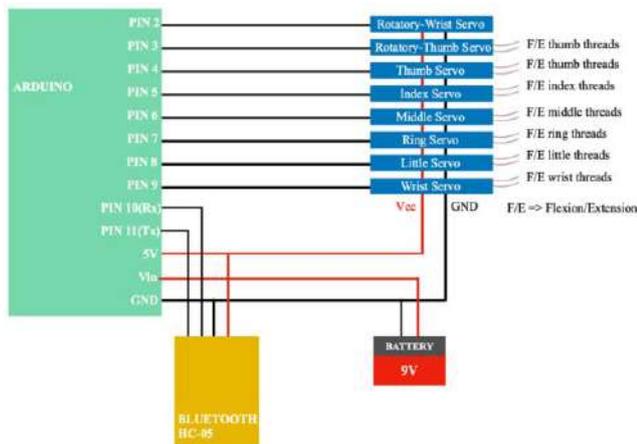


Figure 8. Hardware scheme

### Arduino Nano

A microcontroller will control all the motors, in this case Arduino. Arduino is a board that will be programmed with the instructions from the computer and will perform them when it obtains current. Arduino is an economical board that uses ATMEGA as microcontroller and can be easily programmable with its IDE.

To program the board a cable is needed, but once it is programmed it can work without it if and when it has a power source plugged. This is interesting for fulfilling one of our main goals: to make this hand portable. Different power supplies could be used, for us a battery will be the best solution to make it wireless as it has also low dimensionalities and can fit in the forearm. This battery must be around 7-12V to correctly supply the Arduino. It will be connected to the Vin and GND pins of the board [7]. In our case the battery used is of 9V.

### Motors

The motors used are servomotors, a type of electric motors that are precise and controllable [8]. There are different types of servomotors that rely on the degree of rotation. For the finger movement, it is essential that the servo can move 360° in both directions. Moreover, it is also critical the motor size as it has to be as small as possible to fit all of them in the forearm. Due to these factors and for fulfilling

the low-cost requirement, the motor chosen is MP-708-0002 [11].

Eight motors are needed for performing all the movements. Seven of them are going to work with a thread system. This system will be composed by two threads that are rolled up in the motor with opposite ways. Whenever the motor rotates one way, the thread rolled in the same sense will be pulled. The second thread will unroll, so it will stop staining and will give the pulled one more range of movement. This will make able the performance the movement. These seven motors are distributed as follows: 5 motors for fingers flexion and extension movement, 1 for the wrist extension and flexion, and the last one for moving the carpometacarpal thumb joint.

For each motor it is required 3 different connections: Vcc, GND, and a digital pin. These are going to be connected from Arduino. Vcc and GND are the connections that gives current to the servo. Vcc can go from 4.8V to 6V for the motors, in our case we will use the 5V pin that comes from Arduino. The digital pin is going to be the one in charge of sending the instructions.

In order to program the servomotors Arduino's library "Servo.h" is necessary. The first step is to initialize the servo as a "Servo" variable. Secondly, attach it to an Arduino digital pin. This is done with the function "attach(number)". To move the servo the velocity is fixed with "write(velocity)" function. Velocity equals to 0 means the fastest mode and in counterclockwise direction. To rotate in the opposite direction the velocity must be 180. Finally, the working time has to be setted. It is the time that the servo is performing the action. It is coded with the function "delay(1000)". For each letter, each of the motors will rotate in a different direction and speed.

### Bluetooth module

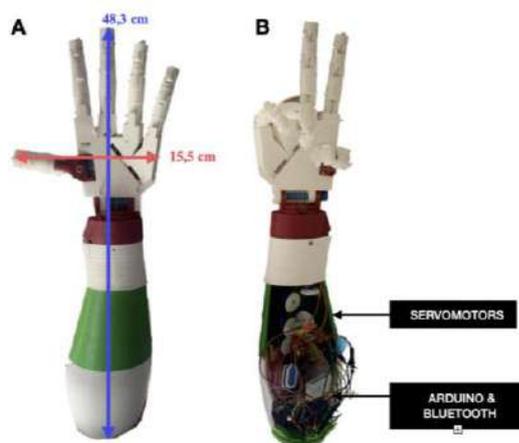
The Bluetooth module must be compatible with Arduino. The two modules that can connect Arduino and Bluetooth are HC-05 and HC-06.

The communication use to send and receive information from the Arduino is a bidirectional serial communication, that uses Rx and Tx pins. The Rx pin of Arduino is usually 0 and 1 for Tx, however this must be change because they are blocked during the programming process of the board. They are changed to 10 and 11 respectively with the library "SoftwareSerial.h". For the current of this module, the 5V pin of Arduino is also use.

It was chosen to work with a smartphone with microphone to send hand activation. The smartphone is in charge of grabbing the sentences said and send them to the Bluetooth module. To record the words said and automatically send them an app is necessary. There are multiple in the Play Store that has this functionality, we choose "Arduino Bluetooth Voice Controller" because it is free and easy to use [10]. However, before performing the communication with the Bluetooth module, it has to be configured. For this, AT commands are used in the terminal.

## 3. Results

Figure 9 shows the final product and its dimensions.



**Figure 9.** A. Hand general view with dimensions- B. Hand general view without the protective case

The 3D printing process, the manufacturing of all the elements took 48 hours and waisted 550g of PLA filament. The total cost of the product is around 100€ and its total weight is approximately 700g.

Concerning its functionality, 22 out of the 24 ASL letters are performed by the hand. The four letters that cannot be done are: "R", "M", "N", and "V". This is due to the restriction of the fingers of the hand to execute movements other than flexion and extension. For "R" and "V", their movements required that the middle and index fingers could be moved to the right and the left, respectively, in the same plane. This cannot be done because these fingers are joined straight to the hand and the carpometacarpal joint do not have all its possible movement. For "M" and "N" movements the thumb finger should be longer as it has to reach the middle phalange of the little finger. A new sign was created for each of them based on the movements it can perform.

Moreover, it is important to notice what happens when two consecutive letters are the same. In sign language, when this happens the person performing the signs makes the letter and moves the arm in a fast way. This has been replicated and coded for the hand, but instead of moving the whole arm, only the wrist will be moved.

The execution of a letter takes a maximum time of 4 seconds. This is because Arduino cannot work with all the servos at the same time; it can only attend to one or two at the simultaneously.

#### 4. Discussion and conclusions

This project aims to design and manufacture a prototype of a robotic hand that can execute the letters of the American sign language alphabet. The motivation was to use engineering and robotics to develop a new way of performing sign language that would allow deaf people to communicate with hearing ones.

For obtaining the complete hand functionality, the hand had to fulfill some movements capacities. This hand is able to move all the fingers with the degrees of mobility of a human hand during the flexion and extension movements. Its wrist is able to perform two movements: move up/down and rotate left/right. The movements are performed by

threads and servo motors with continuous rotation for the recreating the complete finger and wrist movement. These ones are controlled by an electronic board, Arduino, connected via Bluetooth. The motors are activated with human voice when the microcontroller is connected to the voice application in a phone.

Besides the fact that this prototype has the lowest cost between the other solutions, it is also lighter. Full robots are more cumbersome because of its weight and complexity of use. The other developed hands were less portable.

Some of the problems found where that some of the letters couldn't be performed and that the time of performance was too high. The first issues could be solved in a future prototype by developing extra pieces for the metacarpal joint of all the fingers. The second one would be solved with a servo shield that is a platform that makes able to work with multiple servos at the same time. In addition to these improvements, an extended technical and functional evaluation of the system with the deaf community would be interesting for future work.

#### References

- [1] Toshiba: *Toshiba Corporation Develops Lifelike Communication Android*. <https://www.global.toshiba/ww/news/corporate/2014/10/pr0601.html> (Accessed: April 2022).
- [2] S. Shigemi, ASIMO and Humanoid Robot Research at Honda. *Springer Nature*. Japan, 2019. (ISBN: 978-94- 007-6046-2) [https://doi.org/10.1007/978-94-007-6046-2\\_9](https://doi.org/10.1007/978-94-007-6046-2_9)
- [3] J.J. Gago et al., Sign Language Representation by TEO Humanoid Robot: End-User Interest, Comprehension and Satisfaction. *Electronics*, vol. 8, no. 1, 2019, pp. 57. <https://doi.org/10.3390/electronics8010057>
- [4] B. Fang, Robotic Fingerspelling Hand for Deaf-Blind Communication. *California Polytechnic State University*, June 2012
- [5] D Printing Industry: *UAntwerp 3D printed Project ASLAN crosses sign language barriers with 3D Hubs - 3D Printing Industry*. <https://3dprintingindustry.com/news/uantwerp-3d-printed-project-aslan-crosses-sign-language-barriers-3d-hubs-120296/> (Accessed: December 2021)
- [6] Thingiverse – Digital designs for Physical Objects: *Hand Robot Inmoov*. <https://www.thingiverse.com/thing:17773> (Accessed: December 2021)
- [7] F. Biendicho, Comunicaci n Bluetooth entre Arduino UNO y Android aplicado a un detector de mentiras. *Universidad Polit cnica de Valencia*, 2015.
- [8] Omron: *Integrated Servo Motors User's Manual*. (Cat. No. I103E-EN-02). [https://assets.omron.eu/downloads/manual/en/v2/i103e\\_integrated\\_servo\\_motors\\_users\\_manual\\_en.pdf](https://assets.omron.eu/downloads/manual/en/v2/i103e_integrated_servo_motors_users_manual_en.pdf).
- [9] MulticompPRO: *Analog Continuous Rotation Servo Motor, 6V, 360° Datasheet*. <https://www.farnell.com/datasheets/2914227.pdf> (Accessed: December 2021)
- [10] Circuit Digest: *Voice Control LEDs using Arduino and Bluetooth*. <https://circuitdigest.com/microcontroller-projects/arduino-based-voice-controlled-leds> (Accessed: March 2022)

# A virtual reality application for the assessment and training of neck sensorimotor control

J.A. Santos-Paz<sup>1</sup>, M.J. Patiño Mejido<sup>2</sup>, F. Moggioli<sup>1</sup>, E. Bocos-Corredor<sup>1</sup>, A. Martín-Pintado-Zugasti<sup>1</sup>, R. García-Carmona<sup>1</sup>, A. Otero<sup>1</sup>

<sup>1</sup> Escuela Politécnica Superior, Universidad San Pablo-CEU, CEU Universities, Urbanización Montepríncipe, 28660 Boadilla del Monte, España, {jose.santospaz, filippo.moggioli, elena.bocoscorredor, martinpintado, rodrigo.garciacarmona, aotero}@ceu.es

<sup>2</sup> Universidad Iberoamericana, A. C., Prolongación Paseo de Reforma 880, Lomas de Santa Fe, México, C.P. 01219, Ciudad de México, México, mariajosepatinom@gmail.com

## Abstract

Neck rehabilitation treatments include exercise routines that the subject must complete. These exercises can become monotonous and, thus, decrease the adherence to the treatment. To tackle this problem, we developed a virtual reality (VR) application for neck rehabilitation. VR was selected because it can be used as a reliable tool for measuring the cervical range of motion (CROM) and as an immersive application that can increase the patient's adherence to the treatment. The application is a game in which the user has to protect a space station from debris and asteroids that could damage it. The user must destroy these objects using only neck movements in the sagittal and transverse plane. Each object appears in a random position within the limits of the user's CROM. A crosshair that allows targeting the objects is in the center of the user's field of view. By successfully aiming this crosshair to an object for a certain period the object is destroyed. The design of the application was carried out alongside a team of physiotherapists from Universidad San Pablo-CEU. When it was finished, it was validated by the physiotherapists, who gave feedback verbally and through a questionnaire.

## 1. Introduction

There are more than 150 musculoskeletal conditions that affect the locomotor system of individuals. Neck pain ranks fourth among these conditions, behind low back pain, fractures, and osteoarthritis [1]. To assess the functional state of the neck, clinicians rely on the cervical range of motion (CROM). A reduced CROM is often associated with neck pain. An adequate measurement of CROM is necessary to design an appropriate treatment plan and thus to achieve a better prognosis of the subject. CROM is measured in three planes of movement: flexion-extension (sagittal plane), left and right rotation (transverse plane), and left and right lateral flexion (frontal plane) (see Figure 1). In the literature we can find a wide range of protocols to assess CROM; the CROM device, digital inclinometry, electromagnetic motion analysis, goniometry, gravity-plus-compass goniometry, inclinometry, optical motion, potentiometry, tape measure, and ultrasounds [2].

Recently, virtual reality (VR) has emerged as a novel setup for CROM assessment. The reliability of this tool has been demonstrated by a previous study [3]. VR has reported intraclass correlation coefficients (ICCs) that lie between 0.82 and 0.97 for intra-rater and inter-rater reliability. These ICC values are higher than those reported for the

goniometer, the most widely used instrument for evaluating CROM [3]. One of the main advantages that VR presents is that its cost has dropped significantly, facilitating access to this technology. In addition, VR can be used both to assess CROM and as an intuitive user interface to guide a subject through the assessment procedure. A VR system is typically composed of a head mounted display (HMD) and two controllers. There is a wide range of VR setups, from standalone devices like the Meta Quest 2™ to PC-connected devices like the HTC Vive Pro™. The difference between them is that in for the former the image rendering takes place inside the HMD, while in the latter it is performed by a computer.



**Figure 1.** Cervical movements. (A) flexion-extension (sagittal plane). (B) left and right rotation (transverse plane). (C) left and right lateral flexion (frontal plane)

### 1.1. Neck Rehabilitation

The most common treatment for neck pain is manual therapy, which includes manipulation and mobilization, combined with exercise and active movement. The exercise included in the treatment plan can be done at home as an alternative to the clinic. However, these exercises are monotonous, and this decreases adherence to treatment. Furthermore, there is a lack of supervision and feedback from a professional. Without this feedback, the patients cannot know if they are performing the exercises correctly. It is also important to consider the speed and precision of the movement, the subject's ability to maintain a fixed position, and the number and frequency of repetitions of each movement during the exercises.

A possible solution to the problems of performing exercises at the patient's home (monotony and lack of supervision) is telerehabilitation. However, current solutions are very limited and inefficient. They simply connect therapist and patient through phone calls, video conferencing systems or they require the patient to have hardware that is expensive or difficult to use [4,5]. VR applications or video games that guide the patient during their exercises can be used for rehabilitation, whether in a clinic or at home. These applications can be developed in such a way that their use is not only therapeutic but also entertaining by incorporating playful mechanisms to turn a routine task such as rehabilitation into a fun and entertaining task. Previous studies have shown benefits in the physical performance of users who play video games that encourage physical activity and that are motivating and promote adherence to the exercise [6].

VR makes it possible to create immersive and dynamic environments that allow the precise control of the interactions of the subjects within the virtual world [7]. Therefore, it should be possible to create a VR experience that increases adherence to the treatment, reduces the possibility of abandoning rehabilitation therapy, and simultaneously creates the sense of progress and fun. The physiotherapist and the patients could also have information about the state of the neck during each rehabilitation session by obtaining the CROM data during the exercises. This would enable a continuous monitoring of the patient, and thus a greater personalization of the therapy. VR is especially suitable for this goal, since the CROM information is captured by the HMD.

## 1.2. Virtual Reality and Neck Rehabilitation

Before designing the VR application, a bibliographic search of the exercises that should be commonly performed during neck rehabilitation was carried out. According to the results of the literature search, two studies caught our attention. The first by Humphreys and Irgens in which they used one eye-head-neck coordination exercise and one eye-head-neck-upper limb coordination exercises [7]. In the first exercise, the subject must look at a fixed point while performing movements of right and left rotation, flexion, and extension sequentially. The second exercise consists in extending both arms in front of them with the thumb extended up. They had to keep their eyes fixed on the thumb of the ipsilateral arm with which they performed the same sequential movements as in the first exercise, with the eyes and head following the movements in a slow and coordinated manner [7]. In a more recent study Alsultan et al. also used a procedure where cervical movements were performed within a limited range of motion (ROM) [8]. They used a limited ROM since performing functional tasks and activities of daily living do not usually require full active ROM. The subjects sat in a chair with their head in a neutral position and were instructed to avoid shoulder movement and to relax the arms. From this position, points were marked at 45° to the left, right, ceiling and floor of the subjects. These marks corresponded to left and right rotation and flexo-extension. Then they were asked to make movements between the 45° marks starting from the neutral position without stopping in the middle. They

performed 10 repetitions for flexo-extension, left and right rotation, and left and right lateral flexion movements [8].

The objective of this work is to design and develop a virtual reality experience for neck rehabilitation. Through gamification, the patient's adherence to the treatment can be increased, avoiding monotony. Additionally, continuous data on the patients' CROM can be measured, which can enable a more continuous adaptation of the therapy goals.

## 2. Materials and Methods

We decided to design and develop a neck rehabilitation VR experience where the subject should be sitting. The seated position promotes that the subject performs the movements only with the neck, avoiding shoulder movement and keeping the arms relaxed on the legs. The movements to be performed in the initial design were flexo-extension, left and right rotation, and left and right lateral flexion; the same movements that are performed during CROM assessment. We presented this design to a research group of physiotherapists from the Universidad San Pablo-CEU. The proposal was well accepted by the physiotherapists, but they recommended us to focus only on the flexo-extension and left and right rotation movements. Left and right lateral flexion is not a common movement in activities of daily living; thus, it can be excluded to focus on the rehabilitation of the other two movements.

### 2.1. Virtual Reality Setup

The first tests of the VR rehabilitation application will be with a setup made up of the HTC Vive Pro™ HMD, two wireless HTC Vive Controllers™ and two Lighthouse™ v2 base stations. The HMD has a resolution of 2880x1600 pixels, with a refresh rate of 90Hz, a field of view of 100° and built-in Hi-Res Certified headphones. The computer has SteamVR installed and an Intel® Core™ i7-7700K CPU, 16 GB of RAM, Windows 10 Home 64-bit, and a NVIDIA® GeForce® GTX 1080 video card with 8 GB of dedicated video RAM.

### 2.2. Development Environment

For the development of the visual assets of the application we have used a series of graphic design applications. First, Autodesk® Maya® was used to create the objects that appear in the main scene of the application. To perform UV mapping we used the standalone application UVLayout® for the UVs of the objects. UV mapping is the process that allows the mapping of textures on the surface of a 3D model from a 2D image. Adobe® Photoshop® was used to create the layouts and textures of the objects. Finally, Adobe® Substance 3D™ Painter was used to create the materials of the objects.

To develop the VR application itself, we have used the Unity real-time simulation and rendering development environment. Unity can easily compile VR applications for different platforms: iOS™, Android™, Microsoft® Windows®, MacOS®, Linux® and Khronos® WebGL™. We are using Windows to interact with the devices connected to the computer; the HTC™ devices.

### 2.3. VR Application for CROM Rehabilitation

In the VR application the user is in a space station that is gravitating around the earth (see Figure 2). The user's objective will be to defend the space station from collisions with approaching objects (asteroids and space debris). Using only neck movements, the user should be able to locate the objects and destroy them. To destroy such objects the user must keep the sight, which is associated with a crosshair, fixed on them during a certain time. After destroying an object, the next will appear in a random position, forcing the user to search for it within a cervical range of motion previously established by the therapist.

In total there are 4 configurable parameters that control the behavior of the objects to be destroyed. The main parameter is the maximum range at which objects that approach the space station can appear. The greater this range, the greater the amplitude of the neck movements that the user will have to perform to locate and destroy the objects. The range will be limited by each user's own range of motion, which will be previously assessed by the therapist. This is a safety mechanism to prevent the user from harming himself. To set the range within which the objects appear, the therapist can enter minimum and maximum values for flexo-extension and rotation measured in degrees. The next three parameters that the therapist can modify are the speed, the number of objects to destroy, and the maximum number of failures before the space station is destroyed.

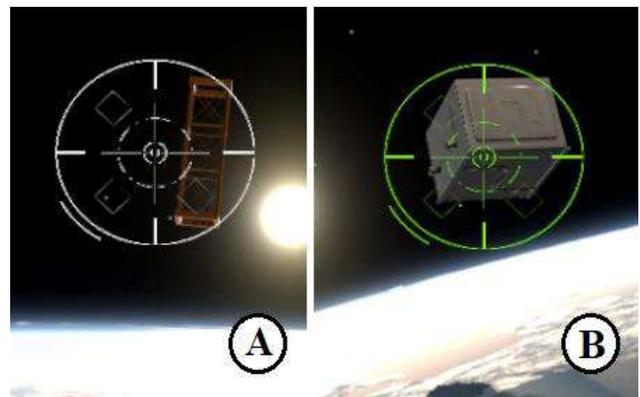


**Figure 2.** User point of view of the CROM rehabilitation VR experience. On the screen we can see the edge of the space station that is gravitating above the earth. An asteroid that approaches the space station and crosshair to aim at the objects can be appreciated.

The VR application for CROM rehabilitation will be comprised of two interfaces: the main VR application viewed by the user through the HMD, and an interface displayed on the computer monitor viewed only by the therapist. In the therapist interface, the parameters that directly influence the experience that the subject will have in rehabilitation can be modified. To use the application in standalone devices, the evaluator interface would have to be integrated into a home screen where the parameters can be modified by the user.

To carry out an adequate rehabilitation of the CROM, the movements should only be performed by moving the neck, avoiding movements other regions such as the trunk or the shoulders. The user should sit upright in a comfortable chair, leaning back against the back and resting the hands on the thighs so that the shoulders are relaxed. The user will remain in this position within the virtual environment as well as in the real world.

In the VR application the user will be seated in a space station that must be protected from incoming objects (see Figure 2). The main objective is that the user, through neck movements, directs shots to destroy meteorites, metal debris and other objects that can damage the station. To shoot the objects, the user must move the head to a point where the target is directly underneath the crosshair and keep them on the object for a configurable time. The crosshair have two colors, associated with two possible states: white if they are not aiming at the object to be destroyed, and green if they are aiming at the object (see Figure 3).



**Figure 3.** Crosshair changes color depending on whether we are correctly aiming at the object we want to destroy or not. (A) Crosshair color is white when the center is not aiming at the object. (B) Crosshair color is green when aiming at the object.

Objects will appear with random X and Y coordinates, within the angle range indicated by the therapist, at a fixed Z distance. Movements in the X axis correspond to flexion and extension. Movements on the Y axis correspond to left and right rotation (see Figure 1). Lateral flexion is not used on the game.

To control the difficulty of the game, the time during which the player has to keep the crosshair fixed on the center of the object to destroy it, and the frequency at which they manage to do so are considered. In the first (easiest) level the subject must maintain a fixed position for at least 4 seconds within the center of the object that is coming toward the space station at a low speed. As the difficulty level increases, the speed increases, but the time needed to destroy the object decreases proportionally, up to a point where the speed is high enough that it is only necessary to locate the center of the object to destroy it. In the first version of the application the difficulty is controlled directly by the therapist. Depending on the results of the pilot test, the possibility of automatically increasing and increasing the difficulty will be evaluated.

#### 2.4. VR Application Evaluation

Four physiotherapists from Universidad San Pablo-CEU were asked to test the VR application. Only the HTC Vive Pro™ HMD was used for this test. Controllers for this app are expendable as all interactions are guided by head and neck movements. The physiotherapists were instructed to always remain seated in the chair and to perform the movements only with the head, avoiding movements in other parts of the body.

To assess the application's ability to achieve neck rehabilitation, a human-computer interface acceptability qualitative response survey was designed [9]. The survey consists of five direct questions where physiotherapists must assess whether the application meets the previously established requirements. The requirements evaluated were speed of movement, accuracy of movement, user's ability to maintain a fixed position, number of repetitions of each movement, and frequency with which repetitions are performed. Each statement can only receive a yes or no response. In case of a negative answer, they were asked to indicate what modifications could be made to improve the application.

### 3. Discussion and Conclusions

The commitment and interest of the subjects undergoing rehabilitation can be increased by using games. In the case of physical therapy, the use of a game may help the subject to increase the adherence to the treatment. Through the gamification of rehabilitation exercises, we seek to achieve greater adherence to the treatment and to obtain better results for neck rehabilitation. Virtual reality permits creating scenarios within simulated environments that would not be possible to produce in the real world. Also, an online connected rehabilitation VR game allows the users to perform exercises in a supervised manner even if they are in their home.

We have proposed a VR game for the rehabilitation of patients with neck pain. The aim of the application is to allow subjects with neck pain to recover their CROM in an entertaining way. An application of these characteristics can be used from the patient's home, thus avoiding the problems of waiting lists or transport that performing therapy in the clinic entails. The VR experience proposed here allows rehabilitation through active movements on different axes of movement of the neck: flexo-extension and left and right rotation exercises. The HMD that we are using can reliably capture the user's movement and store it [3]. This permits a complete record of the evolution of the state of the patient to be kept, enabling the continuous adaptation of the therapy goals to each patient state and evolution.

Based on the test results and the feedback of the physiotherapists, the video game fulfills its objectives. In the survey that they filled after using the application they gave positive responses for all the statements. Therefore, we can conclude that the video game is ready to be validated with patients. A functional validation with healthy subjects will be carried out first. Then a clinical study with subjects with neck pain will be carried out.

During the validation study, the use of the VR device will always be supervised by a physiotherapist so that the user does not injure himself. To this end, the application will be migrated to a standalone VR device like the Meta Quest 2™, which is a standalone headset. Making the application portable will provide the option of remote rehabilitation using only one device and not requiring a high-end computer. Finally, a platform where several fun activities are integrated to help and guide the patient in the recovery process will be designed to complement the present video game.

#### Acknowledgments

This research was funded by the Ministry of Science, Innovation, and Universities of Spain and by the European Regional Development Fund of the European Commission; grant numbers RTI2018-095324-BI00 and PID2021-127096OB-I00.

#### References

- [1] Web page of World Health Organization. Musculoskeletal health (WHO 2022). <https://www.who.int/news-room/fact-sheets/detail/musculoskeletal-conditions> (Accessed: July 2022).
- [2] Williams, M. A., McCarthy, C. J., Chorti, A., Cooke, M. W., & Gates, S. (2010). A systematic review of reliability and validity studies of methods for measuring active and Passive cervical range of motion. *Journal of manipulative and physiological therapeutics*, 33(2), 138-155.
- [3] Santos-Paz1J.A., Sánchez-Picot A., Rojo A., Martín-Pintado-Zugasti A., Otero A., Garcia-Carmona R., (2022). A novel virtual reality application for autonomous assessment of cervical range of motion: Development and reliability study. *PeerJ* (In press).
- [4] Rocco P, Finkelstein J. Telerehabilitation for Patients with Cancer: A Scoping Review. *Stud Health Technol Inform.* 2022 Jun 6; 290:543-546. doi: 10.3233/SHTI220136. PMID: 35673075.
- [5] Oyama S, Saeki M, Kaneta S, Shimoda S, Yoneda H, Hirata H. Telerehabilitation Based on Markerless Motion Capture and IMT-2020 (5G) Networks. *Stud Health Technol Inform.* 2022 Jun 6; 290:1108-1109. doi: 10.3233/SHTI220291. PMID: 35673229.
- [6] Maillot, P., Perrot, A., & Hartley, A. (2012). Effects of interactive physical-activity video-game training on physical and cognitive function in older adults. *Psychology and aging*, 27(3), 589.
- [7] Humphreys, B. K., & Irgens, P. M. (2002). The effect of a rehabilitation exercise program on head repositioning accuracy and reported levels of pain in chronic neck pain subjects. *Journal of Whiplash & Related Disorders*, 1(1), 99-112.
- [8] Alsultan, F., Cescon, C., De Nunzio, A. M., Barbero, M., Heneghan, N. R., Rushton, A., & Falla, D. (2019). Variability of the helical axis during active cervical movements in people with chronic neck pain. *Clinical Biomechanics*, 62, 50-57.
- [9] Ozok, A.A. Survey design and implementation in HCI. In *The Human-Computer Interaction Handbook*; CRC Press: Boca Raton, FL, USA, 2012; pp. 1259–1277.

# Efecto del entrenamiento de la anastomosis microquirúrgica experimental en la calidad del procedimiento y la ergonomía del cirujano

Francisco M. González Nuño<sup>1</sup>, José Castillo Rabazo<sup>1</sup>, Juan A. Sánchez-Margallo<sup>1</sup>, Elena Abellán Rubio<sup>1</sup>, Laura C. Pires Louça<sup>1</sup>, María J. Jara Muriel<sup>1</sup>, Francisco M. Sánchez-Margallo<sup>1</sup>

<sup>1</sup>Centro de Cirugía de Mínima Invasión Jesús Usón, Cáceres, España, {fgonzalez, jasanchez, eabellan, msanchez}@ccmijesususon.com

## Resumen

*En la práctica microquirúrgica, la experiencia resulta un factor clave en el desempeño quirúrgico, no solo en calidad del procedimiento sino también en la actividad y fatiga muscular del cirujano, siendo el entrenamiento crucial en estas cuestiones. Para responder a esta hipótesis, se lleva a cabo un procedimiento experimental con una muestra de siete cirujanos con distintos niveles de experiencia en microcirugía, los cuales llevan a cabo una tarea de sutura en simulador. Durante su desarrollo se tomarán una serie de parámetros para evaluación de desempeño quirúrgico, así como la señal EMG. Tras completar la tarea, se les pidió a los participantes que rellenasen un cuestionario sobre carga física y mental percibida. Con estos datos, se estudió la carga real y percibida, la fatiga y el resultado quirúrgico de la actividad. Los participantes muestran una mejora en el desempeño quirúrgico durante el desarrollo del experimento, así como un aumento de la percepción de carga física. Asimismo, aparecen diferencias estadísticamente significativas en la actividad muscular entre los distintos niveles de experiencia.*

## 1. Introducción

La microcirugía es una disciplina quirúrgica que combina técnicas de magnificación de imágenes, herramientas de precisión especializadas y distintas técnicas quirúrgicas, las cuales son aplicadas principalmente en intervenciones en estructuras de pequeño tamaño tales como vasos sanguíneos, nervios o conductos linfáticos, entre otros.

Debido a las características de este tipo de intervenciones, ligadas a la necesidad de un microscopio quirúrgico, los microcirujanos son propensos a sufrir dolencias de índole musculoesquelética (ej. tendinosis). El análisis mediante electromiografía (EMG) es una herramienta ampliamente utilizada en diversos ámbitos de la práctica quirúrgica para medir la activación muscular y parámetros como la fatiga muscular localizada [1]. Sin embargo, son escasos los estudios que emplean este tipo de técnicas para el análisis ergonómico de los microcirujanos. Podemos destacar el estudio de Shimizu et al., quienes presentaron una nueva distribución de los interruptores de pie que controlan el microscopio quirúrgico [2]. Demostraron, mediante EMG, una reducción significativa de la fátiga muscular de los cirujanos en la mitad inferior del cuerpo durante procedimientos de microneurocirugía.

Normalmente, la experiencia favorece a reducir estos trastornos musculoesqueléticos de los cirujanos,

disminuyendo la fatiga muscular e incrementando la eficiencia del procedimiento. Aquí surgen las dos hipótesis objeto de estudio: (1) la actividad y fatiga muscular del cirujano en la anastomosis microquirúrgica experimental disminuirá tras la fase de entrenamiento y (2) la calidad del procedimiento se incrementará tras el entrenamiento de los cirujanos noveles. Por lo tanto, el objetivo principal de este estudio es evaluar el impacto del entrenamiento en la anastomosis microquirúrgica experimental sobre la carga del cirujano y la calidad del procedimiento.

## 2. Materiales y métodos

### 2.1. Protocolo de estudio

Un grupo de cirujanos con distintos niveles de experiencia en microcirugía vascular han participado en el estudio, todos ellos. Los participantes fueron clasificados con un nivel de experiencia novel, intermedio y experto, según el número de procedimientos microquirúrgicos vasculares y/o nerviosos realizados. Se les pidió realizar un total de 11 repeticiones de anastomosis (conexión quirúrgica entre dos estructuras) vascular en simulador con un total de 8 suturas y un máximo de 10 minutos por repetición; todas ellas debían realizarse el mismo día y bajo las mismas condiciones. Para analizar la evolución del estudio, se emplean las repeticiones 1, 6 y 11 para todos los sujetos. Asimismo, el procedimiento fue grabado tanto el microscopio como las manos del cirujano para su posterior evaluación.

De cada repetición, se analizó la calidad del desempeño quirúrgico, la actividad muscular, la calidad del resultado quirúrgico y la carga de trabajo tanto mental como física percibida por el participante. Para evaluar el desempeño quirúrgico se empleó la escala SMaRT [3], evaluada por un experto microcirujano. La actividad muscular fue analizada mediante EMG, la cual permite medir la respuesta muscular a un estímulo nervioso. Para evaluar la calidad del resultado quirúrgico, se analizó la permeabilidad de la anastomosis. La evaluación de la carga de trabajo fue realizada mediante el cuestionario SURG-TLX.

Para el registro de la actividad muscular (EMG) se empleó el sistema inalámbrico EMG TRIGNO™ Avanti de DELSYS. Los registros se llevaron a cabo con una frecuencia de muestreo de 1024 Hz y una resolución de

16-bits. Para este estudio, se analizaron de forma bilateral los siguientes grupos musculares: Músculos paravertebrales, trapecios superiores, lumbares y supinadores largos. El protocolo de colocación de sensores se ha llevado a cabo de acuerdo con el estándar SENIAM [4].

Se estudiará la actividad muscular respecto a la máxima contracción voluntaria (%MVC) y la fatiga muscular localizada. El registro de la MVC se realiza posterior a la tarea con el objetivo de no influir en el desempeño quirúrgico de los participantes. Las tareas microquirúrgicas requieren de una alta precisión en su desempeño, las cuales pueden verse afectadas por actividades previas a la intervención, como puede ser un elevado esfuerzo físico. Previo al registro del MVC, los participantes tuvieron un descanso de 15 min. Para el estudio de la fatiga, se analiza la evolución temporal de la frecuencia mediana ( $f_{MED}$ ) de las repeticiones 1, 6 y 11. Un desplazamiento de esta frecuencia hacia frecuencias menores es un indicativo de la aparición de fatiga muscular [5].

Por consiguiente, el flujo de trabajo en este estudio es el siguiente: (1) Puesta a punto del sistema; (2) Ejecución de las tareas y registro de datos; (3) Registro de la máxima contracción voluntaria; (4) Cuestionario SURG-TLX; (5) Análisis de datos.

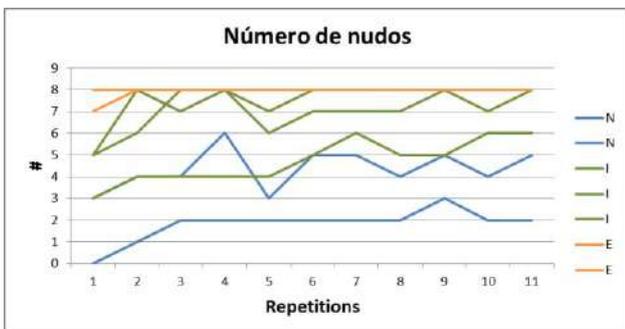
### 3. Resultados

La Tabla 1 muestra un resumen general de las tareas realizadas durante el estudio.

**Tabla 1.** Resumen de datos extraídos del estudio. NE: nivel de experiencia; TE: tiempo de ejecución medio (min.); NN: número de nudos medio; SMaRT: puntuación de la evaluación SMaRT.

Sujeto	NE	TE	NN	SMaRT
S01	NOVEL	10	4,36	24,67
S02	INTERMEDIO	10	4,73	29,67
S03	NOVEL	10	1,82	16,33
S04	EXPERTO	8,59	8	45
S05	INTERMEDIO	9,66	7,09	39,67
S06	INTERMEDIO	9,36	7,45	42
S07	EXPERTO	7,9	7,91	45

En general, se observa una reducción en el tiempo de ejecución (Figura 1), así como un incremento tanto del número de nudos como de la puntuación (Figura 2).

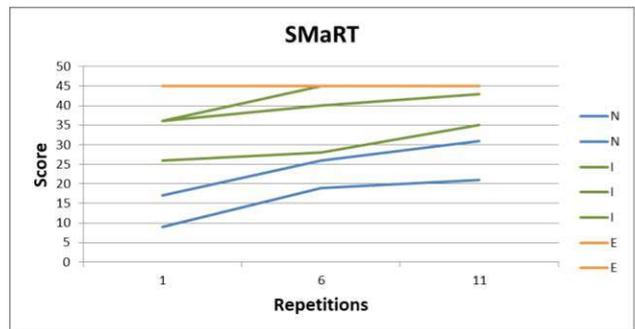


**Figura 1.** Número de nudos por categoría y repeticiones.



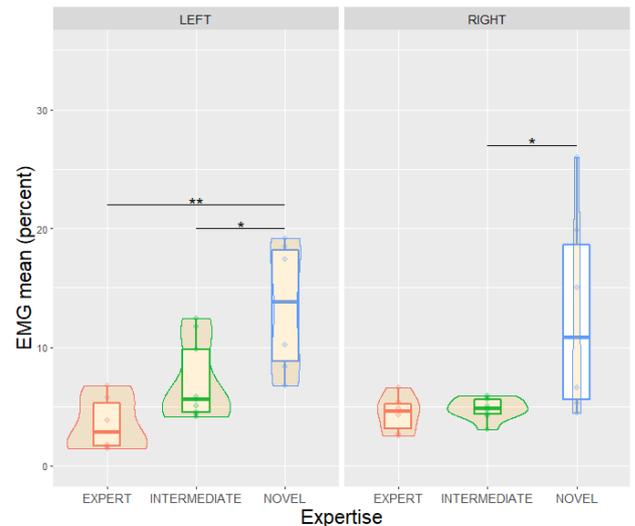
**Figura 2.** Tiempo de ejecución de las tareas por categoría y repeticiones.

Teniendo en cuenta los resultados del SURG-TLX, en general, existe una tendencia al alza para la demanda física para todos los grupos. Sin embargo, los cirujanos noveles muestran una tendencia a la baja en las demandas mental y temporal; los intermedios y expertos se mantienen más estables.



**Figura 3.** Evolución de la puntuación SMaRT por repetición.

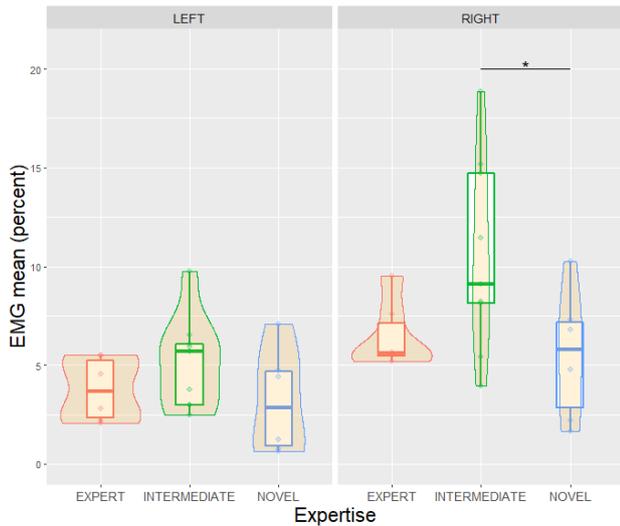
En las Figuras 4 a la Figura 7 se muestran los resultados de la actividad muscular (%MVC) para cada grupo muscular analizado y respecto a cada nivel de experiencia. Los grupos musculares paravertebral y supinador largo muestran una diferencia más notable en los niveles de actividad muscular entre las distintas categorías de experiencia.



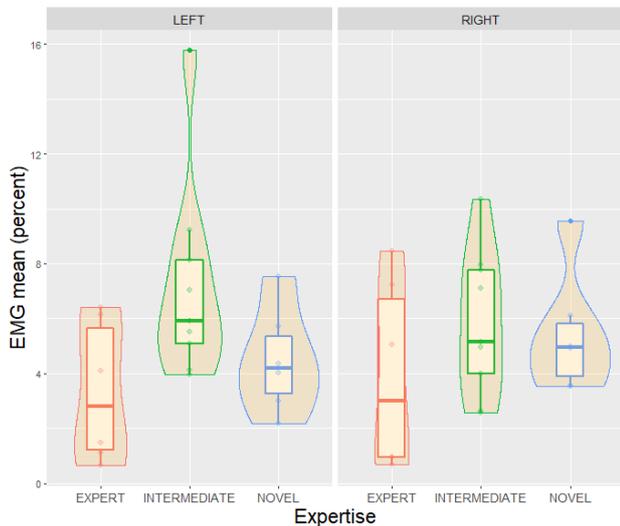
**Figura 4.** Actividad media por lado y experiencia del grupo muscular paravertebral.

En el grupo muscular paravertebral, existe una diferencia estadísticamente significativa entre cirujanos noveles e intermedios ( $p = 0.025$ ) y noveles y expertos ( $p = 0.004$ )

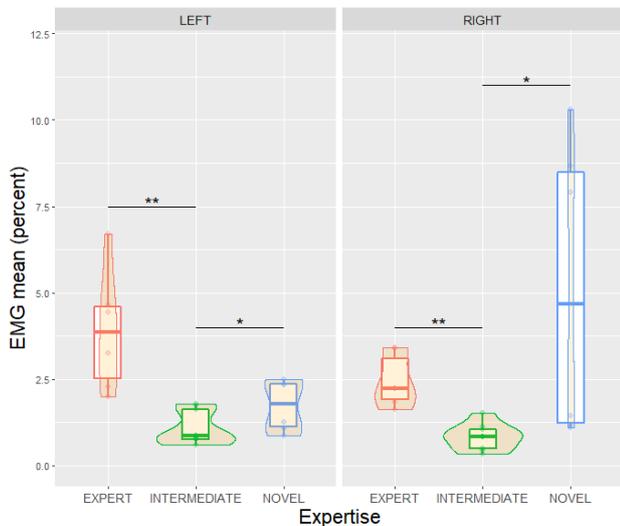
para el lado izquierdo y entre noveles e intermedios ( $p = 0.034$ ) para el lado derecho (Figura 4).



**Figura 5.** Actividad media por lado y experiencia del músculo trapecio superior.



**Figura 6.** Actividad media por lado y experiencia del grupo muscular lumbar.

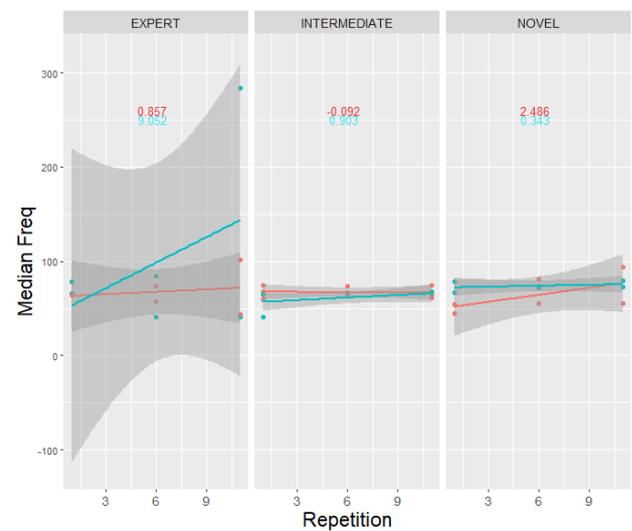


**Figura 7.** Actividad media por lado y experiencia del músculo supinador largo.

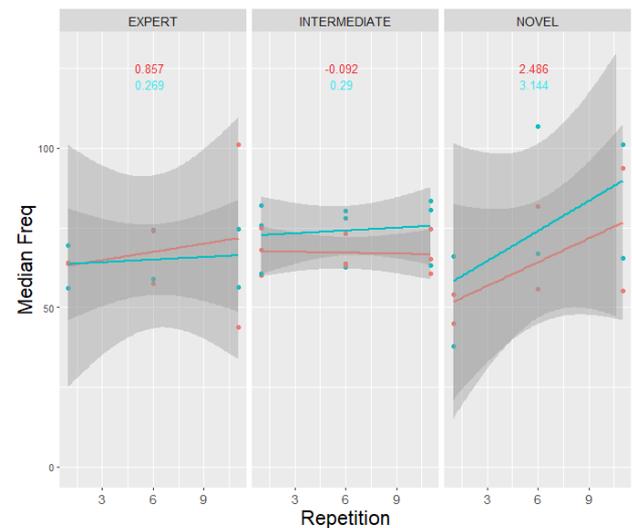
Para el caso del trapecio superior, se observan diferencias significativas entre noveles e intermedios ( $p = 0.045$ ) (Figura 5).

En el caso del supinador largo, en el lado izquierdo se observan diferencias significativas entre cirujanos noveles e intermedios ( $p = 0.034$ ) y entre el grupo de cirujanos intermedios y expertos ( $p = 0.001$ ). En lado derecho, se muestran diferencias entre cirujanos noveles e intermedios ( $p = 0.007$ ) e intermedios y expertos ( $p = 0.001$ ).

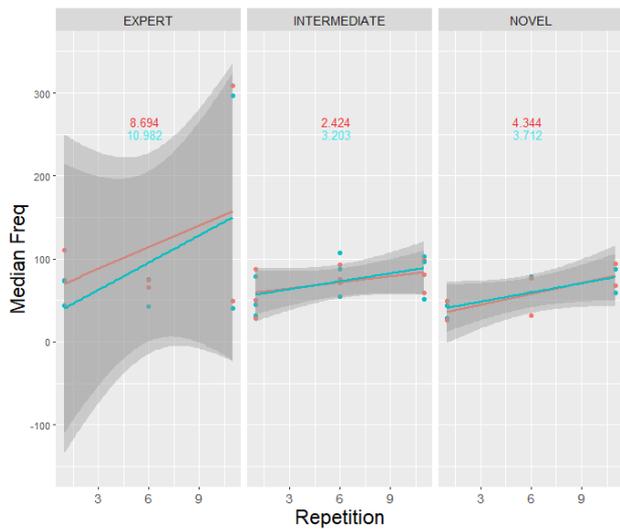
Respecto a la fatiga muscular, no aparecen indicios de fatiga para los grupos musculares paravertebral (Figura 8), trapecio superior (Figura 9) y lumbar (Figura 10). En el caso del supinador largo, se observa fatiga para los grupos de cirujanos expertos (bilateral) e intermedios (lado dominante).



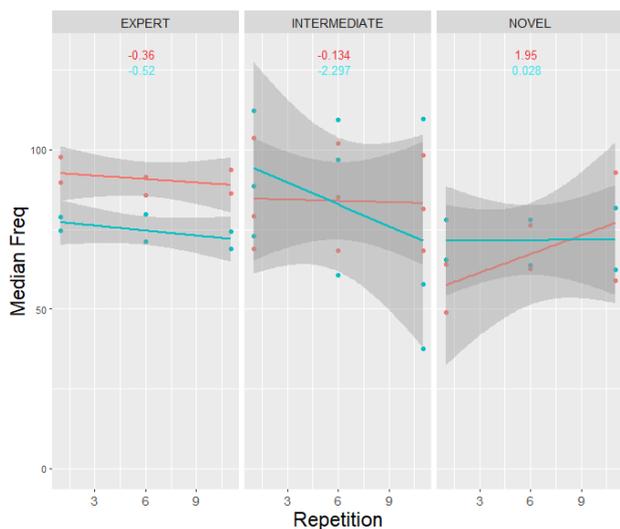
**Figura 8.** Evolución de la frecuencia mediana por lado y experiencia del grupo muscular paravertebral.



**Figura 9.** Evolución de la frecuencia mediana por lado y experiencia del trapecio superior.



**Figura 10.** Evolución de la frecuencia mediana por lado y experiencia del grupo muscular lumbar.



**Figura 11.** Evolución de la frecuencia mediana por lado y experiencia del músculo supinador largo.

#### 4. Conclusión y discusión

Tal y como hemos observado en el análisis de la carga de trabajo, la escasa experiencia de los cirujanos noveles en este tipo de técnicas quirúrgicas puede llevar a que sus primeras repeticiones perciban una elevada carga mental y temporal. Por otro lado, el aumento del número de nudos realizados, la calidad del desempeño quirúrgico (SMaRT), así como la disminución de la carga mental y temporal percibidas para estos cirujanos noveles, puede ser un indicativo de la mejora en el desarrollo de la tarea. Esto puede ser resultado del leve periodo de entrenamiento de los cirujanos participantes en el estudio. Este hecho también se observa en la disminución de tiempo de ejecución de los cirujanos expertos.

La diferencia de cargas percibidas entre los distintos grupos podría indicar que la experiencia, habilidades adquiridas y destreza son cruciales en el desempeño microquirúrgico.

La tendencia alcista general presente en la demanda física percibida puede ser un indicativo de la fatiga que va

apareciendo con el tiempo durante el desarrollo del estudio.

En cuanto a la actividad muscular, cabe destacar el valor elevado de los grupos paravertebrales, en especial en los cirujanos noveles, y su diferencia significativa con el resto de grupos (principalmente en el lado izquierdo). Esto podría ser debido a un inadecuado posicionamiento del microscopio durante la cirugía. En el caso del trapecio superior, aparecen valores superiores para el lado derecho, correspondiéndose con la mano dominante.

En el caso del músculo supinador largo, existe una diferencia notable entre el grupo intermedio con el resto de grupos. Este grupo muscular puede verse muy influenciado por el adecuado uso del instrumental microquirúrgico, así como la activación muscular empleada para llevar a cabo la tarea de sutura. De hecho, la fatiga muscular aparece principalmente en este músculo. Esta circunstancia aparece principalmente en el lado de la mano dominante, siendo la que habitualmente presenta mayor implicación tiene durante el desarrollo de la tarea.

Para poder confirmar las conclusiones obtenidas de este estudio preliminar se ha ampliado el tamaño de la muestra. Sin embargo, cabe destacar que se observan diferencias apreciables tanto en los datos extraídos directamente del procedimiento (Tabla 1), los extraídos de los cuestionarios y la actividad muscular (Figuras 4 a 7). Del mismo modo, para trabajos futuros, será de interés incluir distintos grupos musculares que pueden ser críticos para cirujanos en microcirugía (ej. abductor corto y largo del pulgar).

#### Agradecimientos

Este estudio ha sido parcialmente financiado por la Junta de Extremadura (España), el Fondo Social Europeo, el Fondo Europeo de Desarrollo Regional (FEDER) y el Ministerio de España de Ciencia e Innovación (IG122, TA18023, CPI-2019-2033-1-TRE -14).

#### Referencias

- [1]. C Lakhiani et al., Ergonomics in microsurgery. *Journal of Surgical Oncology*, 2018. DOI: 10.1002/jso.25197
- [2]. S Shimizu et al., Inclined Foot Switches for Surgical Microscopes: A Comfortable Design for Seated Surgeons. *Neurologia Medico Chirurgica*. DOI: 10.2176/nmc.51.260
- [3]. T Satterwhite et al., The Stanford Microsurgery and Resident Training (SMaRT) Scale, *Annals of plastic surgery*, 2014, DOI: 10.1097/SAP.000000000000139.
- [4]. SENIAM Project. Dirección web: <http://www.seniam.org/>
- [5]. LD Gilmore and CJ De Luca, Muscle fatigue monitor (MFM): second generation, *IEEE Transactions on Biomedical Engineering*, 1985, DOI: 10.1109/TBME.1985.325634.

# Automatización del método RULA para análisis ergonómico con sensores inerciales vestibles

Javier González-Alonso<sup>1</sup>, Cristina Simón-Martínez<sup>2</sup>, Francisco Javier Díaz-Pernas<sup>1</sup>, Miriam Antón-Rodríguez<sup>1</sup>, David González-Ortega<sup>1</sup> and Mario Martínez-Zarzuela<sup>1</sup>

<sup>1</sup> Grupo de Telemática e Imagen, Universidad de Valladolid, España, {javier.galonso, fjavier.diaz, mirant, david.gonzalez.ortega, mario.martinez}@uva.es

<sup>2</sup> University of Applied Sciences, Western Switzerland, Sierre (HES-SO), cristina.simon@hevs.ch

## Resumen

*Los sistemas de sensores inerciales se están convirtiendo en una de las principales soluciones para la captura del movimiento humano. Sobre estas soluciones se construyen diferentes capas de análisis de la cinemática humana, como aquellas que presentan una evaluación del riesgo postural. En este trabajo presentamos la implementación de una capa software de valoración ergonómica basada en el método RULA (rapid upper limb assessment) sobre un sistema de sensores inerciales de desarrollo propio. Las herramientas de software abierto y modelos humanos utilizados nos permiten obtener una puntuación instantánea para cada postura, según el grado de lesividad detectado. Además, el sistema presenta un análisis de adecuación en el que indicará cuánto tiempo se mantuvo el sujeto en una postura inadecuada con respecto al tiempo total de actividad. Para comprobar la viabilidad de este sistema se han realizado mediciones en un puesto de trabajo de la industria local, utilizando simultáneamente estos sensores y un sistema comercial. El análisis ergonómico muestra unos resultados comparables para ambos sistemas.*

## 1. Motivación

Actualmente, los métodos más utilizados para la evaluación postural están guiados por ergónomos. Se actúa en base a sus observaciones visuales para valorar la movilidad física de los pacientes y definir mecanismos de prevención. En el ámbito de la bioingeniería, el análisis del movimiento humano (human motion analysis) es especialmente relevante para la identificación de desórdenes biomecánicos y la reducción de posturas inadecuadas que pueden conllevar a futuras lesiones de origen musculoesquelético [1].

Con la popularización de sistemas de captura basados en IMUs (Inertial Measurement Units) que permiten registrar datos precisos a menor coste, surge un nuevo abanico de posibilidades para implementar herramientas metodológicas de evaluación a partir de los datos recogidos [2]. Algunas herramientas de evaluación ergonómica, como la evaluación rápida de las extremidades superiores (RULA, rapid upper limb assessment) [3], pueden ser construidas en base a estos datos, brindando al proceso de evaluación habitual objetividad, automatización y anonimización de datos. Tanto es así, que algunos sistemas comerciales como Xsens Awinda [4] ya han comenzado a incluir software adicional para la evaluación ergonómica de los datos recopilados.

Nuestro propósito es completar nuestro prototipo de sistema de captura de movimiento con sensores vestibles [6, 7], creado desde cero por nuestro equipo de investigación, con

una capa software de evaluación ergonómica que presente los datos recogidos de una forma fácilmente interpretable.

## 2. Sistema Propuesto

El principal objetivo del proyecto es la implementación de una solución para el análisis ergonómico con sensores vestibles, que permita evaluar el desempeño anatómico de actividades de distinta índole, mediante la presentación del grado de adecuación de las posturas producidas, facilitando su posterior corrección. Además, el software debe proporcionar una puntuación instantánea de las posturas producidas y un porcentaje de riesgo postural según el tiempo que pase en un rango articular del total de la actividad registrada. Esto facilitará la evaluación posterior de los ergónomos, agilizando el proceso de rediseño de puestos y fomentando la disminución de los accidentes laborales.

Aunque algunos estudios de investigación recientes presentan otras soluciones personalizadas no comerciales basadas en sensores inerciales, están limitados a un número reducido de sensores y sólo permiten mediciones de rango articular [5]. Sin embargo, nuestro sistema está avalado mediante un estudio comparativo con otros sistemas “low cost” comerciales, registra hasta 11 segmentos corporales, es inmune a interferencias electromagnéticas y a entornos saturados en la banda de 2,4 GHz [6, 7].

Para las grabaciones del presente trabajo, hemos utilizado una versión reducida del sistema conformada por 5 sensores vestibles, conteniendo cada uno, una unidad de medida inercial (BNO080), un núcleo de procesamiento (ARM Cortex M4) con transceptor de 2,4 GHz (7-8 m de alcance) y una batería de polímero de litio de 120mAh junto con un chip de carga de batería (MCP73831) sustituyendo la pila de botón recargable del modelo anterior para aumentar su autonomía. Todo en un mismo encapsulado de dimensiones totales de 45 × 28 × 10 mm.

Los datos generados se introducirán al sistema OpenSim [8], un sistema de acceso abierto desarrollado por la Universidad de Stanford que permite conectar los datos registrados a un modelo de estructuras musculoesqueléticas humano y generar simulaciones dinámicas del movimiento. Este sistema posibilita la obtención de resultados de cinemática inversa (Inverse Kinematics), validados con distintos sistemas de captura de movimientos incluyendo sensores inerciales [9]. La salida nos presentará los ángulos articulares

resultantes tras pasar por las restricciones del modelo humano y ajustándose lo mejor posible a los valores realmente alcanzables por cada sujeto según sus medidas corporales.

Por último, la capa de análisis ergonómico está compuesta de varios scripts de Python que asignarán valores según el método RULA [3] a los ángulos articulares instantáneos, generando gráficas e informes ergonómicos de manera semiautomática. Ésta podrá ser utilizada como capa final por cualquier archivo de salida del sistema OpenSim, facilitando la compatibilidad con otros sistemas como el de cámaras ópticas con marcadores o Vicon [10] entre otros.

La puntuación RULA se corresponde con los valores establecidos en las directrices del método [3] para el análisis local, teniendo en cuenta los siguientes puntos para los resultados obtenidos:

- No se ha considerado la elevación de hombros (+1).
- No se ha considerado si la persona está apoyando los brazos (-1).
- Se ha considerado la abducción de brazos (+1) desde un valor en abducción de +20 grados.
- No se ha considerado si los brazos cruzan la línea medial o si están hacia fuera del cuerpo.

La información resultante se le ofrece al profesional sanitario de una forma gráfica y simplificada, de manera que pueda apoyarse en estas mediciones para ofrecer una valoración ergonómica precisa y objetiva de la actividad registrada.

### 3. Procedimiento

Se escogió valorar la adecuación de un puesto de trabajo en el que se detectaron movimientos posturalmente inadecuados prolongados en el tiempo. El puesto consistía en la colocación de diferentes piezas de automóviles en una superficie de soldadura robotizada, y fue escogido entre 3 simulaciones de puestos registradas. La actividad analizada corresponde con la que concentraba los movimientos corporales de mayor carga lesiva, de las 4 subsecciones independientes en que estaba dividido el puesto.

Se escogieron las articulaciones de codo y hombro para realizar un análisis ergonómico local de acuerdo con el método RULA. Se utilizaron 5 sensores en total colocados en espalda, brazo y antebrazo izquierdos y brazo y antebrazo derechos. Los datos se recopilaban utilizando sensores inalámbricos en cada uno de los posibles ejes de movimiento de cada articulación, aunque siguiendo el método escogido se tuvieron en cuenta solo valores de rango articular en determinados ejes (flexión, abducción y pronosupinación). Se colocaron los sensores en la localización adecuada mediante unas cinchas de sujeción y un encapsulado especial para la espalda, como podemos apreciar en la **Figura 1**. Los datos fueron recopilados utilizando simultáneamente los sensores de desarrollo propio y los Xsens Awinda [4], colocados en la misma posición aproximada para buscar obtener resultados comparables, a sabiendas de que la colocación de los sensores es crucial en mediciones con sensores inerciales [11].



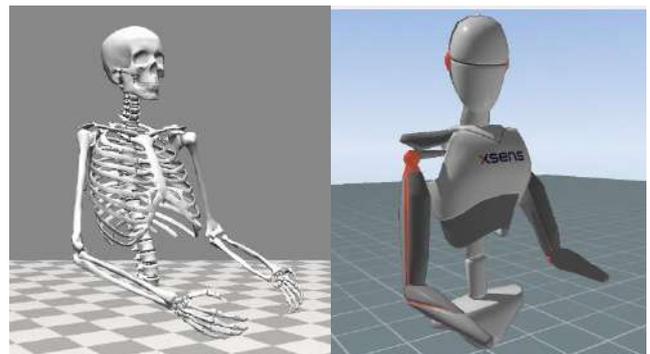
(a) Vista frontal (b) Vista trasera

**Figura 1.** Colocación conjunta del sistema de sensores personalizado y del sistema Xsens Awinda.

El sujeto elegido fue un trabajador que desempeñaba de forma habitual dicha actividad, caracterizado con las medidas descritas en la **Tabla 1**, inferidas por el software de Xsens MVN a partir de los datos de altura y pie del sujeto y según bases de datos normalizadas.

Medida corporal	Dimensiones (cm)
Altura del cuerpo	171
Longitud del pie o del zapato	26
Altura de los hombros	144
Anchura de los hombros	38
Longitud del codo	94
Envergadura de la muñeca	143
Envergadura del brazo	180
Altura de la cadera	87
Anchura de la cadera	24
Altura de la rodilla	49
Altura del tobillo	8

**Tabla 1.** Medidas del sujeto escogido para la adquisición.



**Figura 2.** Visualización de la actividad registrada: Modelo de OpenSim (izquierda) vs. Modelo Xsens MVN. (derecha)

Se utilizó el modelo musculoesquelético más adecuado para las mediciones de dichas articulaciones [12], modificándolo para permitir movimientos de brazos por encima de la línea de visión, que serían claves para nuestro estudio.

Podemos observar un instante concreto de las grabaciones de ambos sistemas en la **Figura 2**. Aquí se muestra la posición corporal del sujeto en el segundo 32 de la actividad, teniendo en cuenta que en el sistema de Xsens MVN se incluye además la articulación de cuello.

### 3.1 Rango articular

El objetivo de la medición de rango articular fue determinar si la representación de estas medidas por parte del modelo musculoesquelético escogido seguía fielmente los movimientos del sujeto registrado. Para ello se evaluaron los resultados de cinemática inversa de ambos sensores.

### 3.2 Puntuación RULA instantáneo

Con los datos anteriores, calcularemos los valores instantáneos de la puntuación RULA, teniendo en cuenta las limitaciones comentadas en la sección 2 Sistema Propuesto.

Obtendremos así de forma gráfica una primera valoración de la puntuación de riesgo local para cada articulación y mediante el modelo humano de OpenSim, podremos elegir situarnos en el instante deseado. Esta gráfica permitirá a los ergónomos colocarse, por ejemplo, en el instante de mayor riesgo postural y observar tanto el modelo como la puntuación RULA de un solo vistazo.

### 3.2 RULA global porcentual

Finalmente, las puntuaciones locales de las articulaciones se combinarán en un análisis global del total de la actividad. El sistema mostrará mediante una gráfica de secciones el porcentaje de tiempo (acumulado) en que el sujeto se ha encontrado dentro de un rango de puntuación postural. Estos rangos se han definido como sigue:

- 1-2 sin riesgo, no se requiere de cambios.
- 3-4 riesgo bajo, se pueden necesitar cambios.
- 5-6 riesgo medio, más investigación/cambio inmediato.
- 7 riesgo muy alto, eliminar la postura de la actividad.

Se proporcionará tanto de parte izquierda como derecha, aunque para calcular el riesgo postural total según el método RULA se toma habitualmente el peor de los casos.

## 4. Resultados

### 4.1 Rango articular

Cada una de las subsecciones del puesto fueron registradas con ambos sistemas. En este apartado mostraremos los resultados correspondientes al codo del lado derecho.

Rango articular codo derecho: Flex(+) / Ext(-) en grados

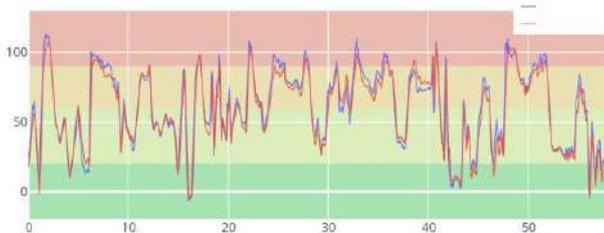


Figura 3. Comparación flexión-extensión codo derecho Sensores Custom vs Xsens Awinda

Como podemos apreciar en el análisis de cinemática inversa resultante, los movimientos registrados (Figura 3) son análogos con ambos sistemas de sensores para la articulación del codo y difieren escasamente en algún movimiento, lo que nos permite continuar con la siguiente capa de análisis.

Si bien es cierto que los movimientos de hombro son ligeramente menos precisos en las capturas de los sensores

de desarrollo propio, siguen entrando dentro de los mismos rangos definidos por el método RULA [3] en los distintos instantes de tiempo, luego podremos utilizarlos para el cómputo posterior del RULA global.

### 4.2 Puntuación RULA instantáneo

Podemos observar en la Figura 4 como en ambos sistemas se obtiene la misma puntuación a cada instante a excepción de algún movimiento puntual. Este análisis se corresponde con los datos de rango articular observables en la Figura 3. Una apreciación importante es que la puntuación toma valores esperados y no se sale del límite superior de 2 en el caso del codo y de 4 en el caso de los hombros.

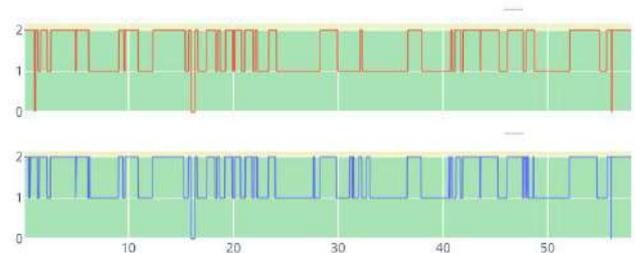


Figura 4. Comparación Codo Derecho actividad registrada con Xsens Awinda vs Sensores Custom

### 4.3 RULA global porcentual

Al combinar las puntuaciones de codo y hombro de ambos lados y mostrar el porcentaje del total según los rangos de puntuación establecidos, podemos observar a simple vista una representación fidedigna de la lesividad general.

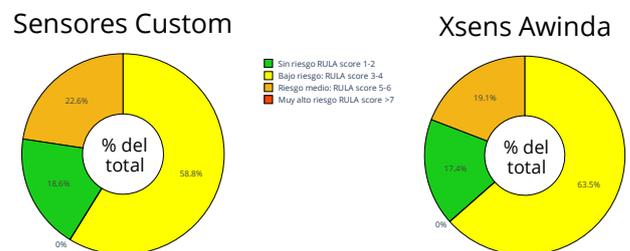


Figura 5. Comparación RULA total lado derecho: sensores personalizados (izquierda) vs Xsens Awinda (derecha)

En el caso de la actividad registrada observamos un comportamiento comparable en las grabaciones de ambos sistemas de sensores como vemos en la Figura 5.

## 5. Discusión

Las mediciones tomadas en el puesto muestran unos valores razonables de desempeño postural. La puntuación instantánea y los porcentajes de puntuación RULA están muy cercanos en las grabaciones de ambos sistemas.

La mayor diferencia observada pertenece a la articulación del hombro, ya que en los Xsens Awinda se incluye un sensor por cada lado en la escápula para obtener datos relativos a la articulación glenohumeral como vemos en la Figura 1. En posteriores iteraciones se añadirán estos sensores al sistema, lo que entra dentro de las posibilidades del sistema

personalizado. Para completar el análisis RULA de la parte superior del cuerpo, se pueden añadir sensores adicionales para medir tronco, cuello y muñecas entre otros. En futuros trabajos plantearemos la grabación de un mayor número de sujetos con diferentes medidas corporales y puestos de diferentes características, siendo esto clave para una generalización de la aplicabilidad del sistema.

Es preciso comentar que este estudio no se ha centrado en demostrar la precisión del sistema de desarrollo propio utilizado, ni el comportamiento de los sensores inerciales escogidos bajo determinadas condiciones, ya que han sido exploradas por otros estudios con anterioridad [13], [14].

## 6. Conclusiones

El enfoque de adquisición de movimientos corporales mediante sensores se beneficia de la fácil colocación y la monitorización continua de actividades para mejorar el proceso de evaluación y reeducación postural. Así mismo, la recopilación de datos de los sensores corporales sobre los movimientos humanos puede mejorar la prevención de lesiones incidiendo en el proceso de diseño de puestos de trabajo.

El propósito de este estudio fue realizar una primera validación de un sistema de valoración ergonómica basado en sensores vestibles de desarrollo propio previamente introducidos [7]. Para ello se realizó la adquisición de movimientos corporales de forma comparativa frente a un sistema comercial. Las medidas de rango articular obtenidas tras el análisis de cinemática inversa en ambos modelos resultaron muy cercanas entre sí, lo que permitió corroborar las posibilidades de una solución basada en los sistemas personalizados o *custom* complementados por software de acceso libre. El análisis RULA para las articulaciones elegidas obtuvo resultados comparables a los de una solución comercial.

Este tipo de sistemas de análisis basados en sensores puede ser muy relevante en el futuro, y acabar convirtiéndose en una herramienta fundamental para apoyar a los ergónomos en el proceso de valoración de puestos de trabajo.

## Agradecimientos

Este trabajo ha sido financiado parcialmente por la Consejería de Empleo e Industria de la Junta de Castilla y León, mediante la línea de subvenciones públicas dirigidas a la realización de acciones de proyectos de investigación en prevención de riesgos laborales por las Universidades Públicas de Castilla y León.

## Referencias

- [1] N. Özkaya, M. Nordin, D. Goldsheyder, y D. Leger, *Fundamentals of Biomechanics: Equilibrium, Motion, and Deformation*, 3.<sup>a</sup> ed. New York: Springer-Verlag, 2012. doi: 10.1007/978-1-4614-1150-5.
- [2] I. Poitras *et al.*, «Validity and Reliability of Wearable Sensors for Joint Angle Estimation: A Systematic Review», *Sensors*, vol. 19, n.º 7, Art. n.º 7, ene. 2019, doi: 10.3390/s19071555.
- [3] L. McAtamney y E. Nigel Corlett, «RULA: a survey method for the investigation of work-related upper limb disorders», *Appl. Ergon.*, vol. 24, n.º 2, pp. 91-99, abr. 1993, doi: 10.1016/0003-6870(93)90080-S.
- [4] M. Paulich, M. Schepers, N. Rudigkeit, y G. Bellusci, *Xsens MTw Awinda: Miniature Wireless Inertial-Magnetic Motion Tracker for Highly Accurate 3D Kinematic Applications*. 2018. doi: 10.13140/RG.2.2.23576.49929.
- [5] C. P. Walmsley *et al.*, «Validation of custom wearable sensors to measure angle kinematics: A technical report», *Health Technol.*, vol. 9, n.º 5, Art. n.º 5, nov. 2019, doi: 10.1007/s12553-019-00360-1.
- [6] J. González-Alonso, M. Martínez-Zarzuela, F. J. Díaz-Pernas, y H. J. Aguado, «Validación de prototipo con sensores vestibles para captura y análisis de movimientos del cuerpo humano con aplicación en medicina», en *XXXVIII CASEIB: Libro de actas, 2020, ISBN 978-84-09-25491-0, págs. 397-400*, 2020, pp. 397-400.
- [7] J. González-Alonso, D. Oviedo-Pastor, H. J. Aguado, F. J. Díaz-Pernas, D. González-Ortega, y M. Martínez-Zarzuela, «Custom IMU-Based Wearable System for Robust 2.4 GHz Wireless Human Body Parts Orientation Tracking and 3D Movement Visualization on an Avatar», *Sensors*, vol. 21, n.º 19, Art. n.º 19, ene. 2021, doi: 10.3390/s21196642.
- [8] S. L. Delp *et al.*, «OpenSim: Open-Source Software to Create and Analyze Dynamic Simulations of Movement», *IEEE Trans. Biomed. Eng.*, vol. 54, n.º 11, pp. 1940-1950, nov. 2007, doi: 10.1109/TBME.2007.901024.
- [9] M. Al Borno *et al.*, «OpenSense: An open-source toolbox for inertial-measurement-unit-based measurement of lower extremity kinematics over long durations», *J. NeuroEngineering Rehabil.*, vol. 19, n.º 1, p. 22, feb. 2022, doi: 10.1186/s12984-022-01001-x.
- [10] D. Panariello, S. Grazioso, T. Caporaso, A. Palomba, G. Di Gironimo, y A. Lanzotti, «Biomechanical analysis of the upper body during overhead industrial tasks using electromyography and motion capture integrated with digital human models», *Int. J. Interact. Des. Manuf. IJIDeM*, abr. 2022, doi: 10.1007/s12008-022-00862-9.
- [11] G. Wu *et al.*, «ISB recommendation on definitions of joint coordinate systems of various joints for the reporting of human joint motion—Part II: shoulder, elbow, wrist and hand», *J. Biomech.*, vol. 38, n.º 5, pp. 981-992, may 2005, doi: 10.1016/j.jbiomech.2004.05.042.
- [12] A. Rajagopal, C. L. Dembia, M. S. DeMers, D. D. Delp, J. L. Hicks, y S. L. Delp, «Full-Body Musculoskeletal Model for Muscle-Driven Simulation of Human Gait», *IEEE Trans. Biomed. Eng.*, vol. 63, n.º 10, pp. 2068-2079, oct. 2016, doi: 10.1109/TBME.2016.2586891.
- [13] R. Stanzani, P. Dondero, A. Mantero, y M. Testa, «Measurement Accuracy of an Upper Limb Tracking System Based on Two Hillcrest Labs BNO080 IMU Sensors: An Environmental Assessment», *IEEE Sens. J.*, vol. 20, n.º 17, Art. n.º 17, sep. 2020, doi: 10.1109/JSEN.2020.2992733.
- [14] L. Zhou *et al.*, «Validation of an IMU Gait Analysis Algorithm for Gait Monitoring in Daily Life Situations», en *2020 42nd Annual International Conference of the IEEE Engineering in Medicine Biology Society (EMBC)*, jul. 2020, pp. 4229-4232. doi: 10.1109/EMBC44109.2020.9176827.

# Biomechanical modelling of the behaviour of the human breast under the effect of gravity for tumor detection

Natalia Gorrin<sup>1\*</sup>, Felicia Alfano<sup>1,2\*</sup>, Pedro Navas<sup>3</sup>, Pablo Lamata<sup>4</sup>, Karla Ferreres García<sup>5</sup>, Juan E. Ortuño<sup>2,1</sup>, Oscar Bueno Zamora<sup>5</sup>, Santiago Lizarraga<sup>5</sup>, Andrés Santos<sup>1,2</sup>, Javier Pascau<sup>6</sup>, José María Goicolea<sup>3</sup>, María Jesús Ledesma Carbayo<sup>1,2</sup>

<sup>1</sup> Biomedical Image Technologies, Universidad Politécnica de Madrid. Madrid, España <sup>2</sup> CIBER-BBN, ISCIII, Madrid, España <sup>3</sup> Department of Continuum Mechanics and Theory of Structures, ETSI Caminos, Canales y Puertos, Universidad Politécnica de Madrid. Madrid, España <sup>4</sup> School of Biomedical Engineering and Imaging Sciences, King's College London, London, United Kingdom <sup>5</sup> Instituto de Investigación Sanitaria Gregorio Marañón. Madrid, España <sup>6</sup> Departamento de Bioingeniería e Ingeniería Aeroespacial. Universidad Carlos III de Madrid. Madrid, España

\* Contributed equally

## Abstract

*Magnetic Resonance Imaging (MRI) has been used as an effective tool for breast cancer evaluation due to its high sensitiveness to detect breast tumors. However, these MRIs are performed with the patient lying prone, while the tumor resection surgery is performed with the patient lying in supine position. The breast undergoes large deformations due to the change in the direction of gravity, and nonpalpable breast lesions must be localized before surgery for their removal. Modelling the breast deformation under gravity loading (from prone to supine position) could allow for the localization of the tumor and help the planning of breast cancer treatment. This study presents biomechanical simulations by the Finite Element (FE) method using ABAQUS software. Patient-specific breast geometries of two patients with different breast sizes were obtained from prone MR images and used in the simulations. Computed Tomography (CT) images in the supine position were also available for the same patients and used for evaluation purposes. The evaluation of the methodology has been carried out by considering the distance between the nipple and the tumor centroid identified in both images. Furthermore, the mean Euclidean distance between the estimated deformed breast surface and the supine surface extracted from CT image was calculated and finally, the percentage overlap between the deformed breast mask and the CT breast mask was obtained. The proposed method achieves an average localization error for the tumor of 4.27 mm.*

## 1. Introduction

Breast cancer is a disease with a major global impact. In fact, it is one of the most prevalent pathologies in women. Almost 2.3 million new cases are diagnosed every year worldwide [1] and produce about 500,000 deaths per year, making it one of the leading causes of cancer death among women [2]. The breast of an adult female is mainly composed of adipose tissue, fibro glandular tissue, and skin. The breast is posteriorly attached to the chest by the pectoralis fascia over the pectoral muscle, while the skin and supporting Cooper's ligaments define and maintain its shape. [3].

The main treatment for breast cancer is surgery. For years, mastectomy has been chosen as the treatment for breast cancer. Lately, improved mammography screening led to the diagnosis of nonpalpable tumors. In these cases, the preferred treatment is breast conserving therapy or

lumpectomy, whose goal is the complete resection of the tumor and the preservation of the shape of the breast. It is used for early-stage breast cancer and requires localization of the tumor prior to surgery. This necessary pre-localization is done since the diagnostic imaging, in this case MR imaging, is performed in the prone position, whereas surgery is performed in the supine position, so the breast undergoes a large deformation from one position to the other. One of the most popular techniques for preoperative localization is Wire Guided Localization: this procedure it is performed by a radiologist on the same day as the patient's planned surgery and consists in the insertion of a hooked wire to tag the location of the lesion under radiological guidance.

Biomechanical modelling and simulation of the breast is a promising area of research with potential uses in healthcare applications, specifically in the diagnosis and treatment planning of breast cancer. In recent years, several research groups have made important improvements in the field of biomechanics, modelling the breast with the FE method.

Previous works have modelled the breast as an isotropic and incompressible material [4,5,6,7,8,9,10]. All of them have considered a homogeneous tissue, except [5,6], which have considered a heterogeneous tissue differentiating between adipose tissue and fibroglandular tissue. Moreover, all these works have used non-linear elastic (hyperelastic) models, specifically, the Neo-Hookean form, and in [4] the 5-parameter Mooney-Rivlin form was also used.

Despite these advances, there are still several issues limiting the clinical application of biomechanical models, such as the automation of patient-specific model generation, robust image segmentation and mesh generation, and the accuracy of biomechanical models [11].

In this work, a biomechanical model of the breast has been designed by applying the gravity load that exists between the prone position (MR image) and the supine position (surgery). In addition, the viscoelasticity of the breast tissue has been explored and the breast has been modelled as a visco-hyperelastic material.

## 2. Methods

### 2.1. Data

Taking into account the burden of the simulation (long simulation set up time and running time) the proposed methodology has been applied in this work to two cases provided by the Hospital General Universitario Gregorio Marañón (HGUGM) in Madrid. Two cases with different breast volume were chosen: Case 1, a patient with a large breast and Case 2, a patient with a small breast.

Each case included a prone MR image, from which the 3D volume model of the breast was obtained and a supine CT image, that has been considered as our reference of the surgical position for evaluation purposes, an example of the two positions is shown in Figure 1.

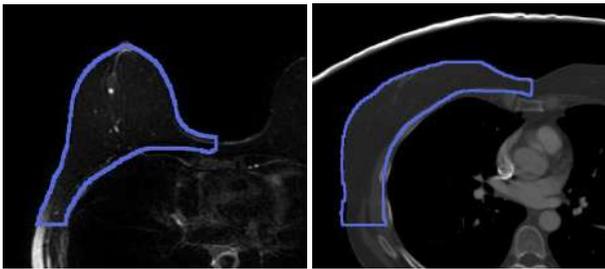


Figure 1. Left: prone MR image. Right: supine CT image.

### 2.2. Segmentation and Meshing

The patient-specific imaging data were segmented using Slicer 3D software [12]. In this procedure, the breast is delineated from the background and the border with the chest wall as a homogeneous tissue.

After the first phase, the segmentation was exported as a triangular mesh (STL format). The mesh obtained was postprocessed using the Iso2mesh library in MATLAB to generate a more suitable mesh for the solver.

The surface meshes were then imported into ABAQUS and converted into a tetrahedral mesh. Specifically, a linear 4-node tetrahedron element (C3D4 in ABAQUS nomenclature) was employed. The size of the for each case was:

- Case 1: 53,323 elements; 10,596 nodes.
- Case 2: 36,939 elements; 7,457 nodes.

### 2.3. Finite Element Model

#### 2.3.1. Formulation and Constitutive model

A literature review of previous work was carried out to find the best biomechanical model to be used in this work.

The breast was modelled as a homogeneous and isotropic tissue with a density of 1000 kg/m<sup>3</sup>. Like most biological tissues, breast tissue exhibits hyperelastic characteristics under deformation. Therefore, we used a Neo-Hookean hyperelastic constitutive model to describe the stress-strain relationship of breast tissues. The neo-Hookean first potential strain energy potential is defined as:

$$U = \frac{\mu_0}{2} (\bar{I}_1 - 3) + \frac{2}{K_0} (J^{el} - 1)^2$$

where  $\mu_0$  is the initial shear modulus and  $K_0$  is the initial bulk modulus:

$$\mu_0 = 2C_{10} \quad K_0 = \frac{2}{D_1}$$

$C_{10}$  is the neo-Hookean parameter;  $D_1$  is an incompressible parameter;  $\bar{I}_1$  is the first invariant of deviatoric deformation and  $J^{el}$  the determinant of the deformation gradient.

In this work, we assumed that the parameter  $C_{10}$  ranges from [40, 80] Pa [8,9].

In addition, the viscoelastic properties of the tissue have been explored in this work. To calculate the viscoelastic behaviour, the Prony series is used to describe the transient response of the material. The parameters of the Prony series are defined as:

$$g_i = \frac{G_i}{G_0}$$

where  $g_i^P$  is the  $i$ -th value of the ratio of the modulus of relaxation at one shear stress of the material;  $G_i$  is the value of the modulus of stiffness of the  $i$ -th element,  $G_0$  is the value of the instantaneous modulus of stiffness. Similarly,

$$k_i = \frac{K_i}{K_0}$$

$k_i^P$  is the  $i$ -th value of the ratio of the volumetric modulus of the material;  $K_i$  is the value of the volumetric modulus of the  $i$ -th element and  $K_0$  is the value of the instantaneous volumetric modulus.

Taylor et al. [13] suggested that visco-hyperelastic models could describe the behaviour of soft tissues subjected to large deformations. However, no reference was found for the adjustment of the viscoelastic parameters. Thus, an exhaustive calibration of the parameters was carried out to obtain a range of plausible values.

#### 2.3.2. Finite Element Simulation in ABAQUS

- *Parts*: Two parts were created: a rigid internal surface simulating the chest wall and the breast as a deformable part (Figure 2).

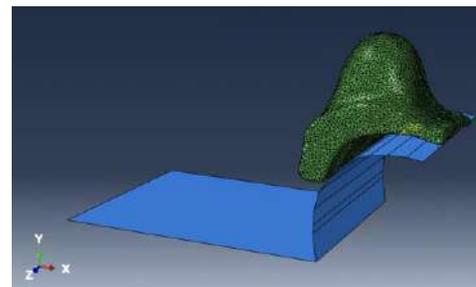


Figure 2. Rigid internal surface (blue) and breast volumetric mesh (green).

- *Material*: Visco-Hyperelastic.
- *Step*: Dynamic-explicit (period 0.35s,  $\Delta t=0.01$ )
- *Interaction*: The breast tissue was observed to slide on the chest wall under the effect of gravity. Hence, the interaction was modelled as a surface-to-surface contact between the internal rigid surface and the contact surface of the breast with a friction coefficient of 0.8.

- **Load:** a gravity load was applied.
- **Boundary Conditions:**
  - **Internal surface:** Fixed in all directions.
  - **Breast:** A medial symmetry boundary condition was applied to simulate the junction with the contralateral breast (Figure 3). Furthermore, the medial area of the breast was restricted in movement in the y-direction (Figure 4) a previously proposed in Vavourakis et al. [14].

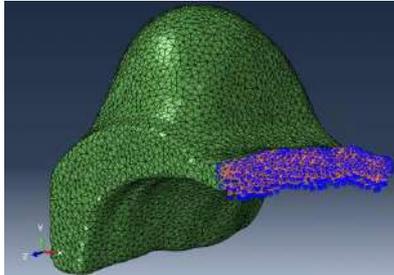


Figure 3. Medial Symmetry Boundary Condition.

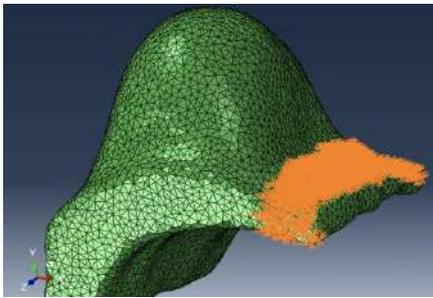


Figure 4. Medial Area Boundary Condition.

### 3. Results

#### 3.1. Viscoelastic Parameter Range

As mentioned above, the optimal range of the viscoelastic parameters was obtained by an exhaustive calibration using as reference the distance between the estimated position of the nipple in the surgical position and the reference supine position of the tumor in the CT used for evaluation. The optimal range identified for the viscoelastic parameter was between (0.65-0.85), neglecting those that provide values higher than 15 mm, as we can see in Figure 5.

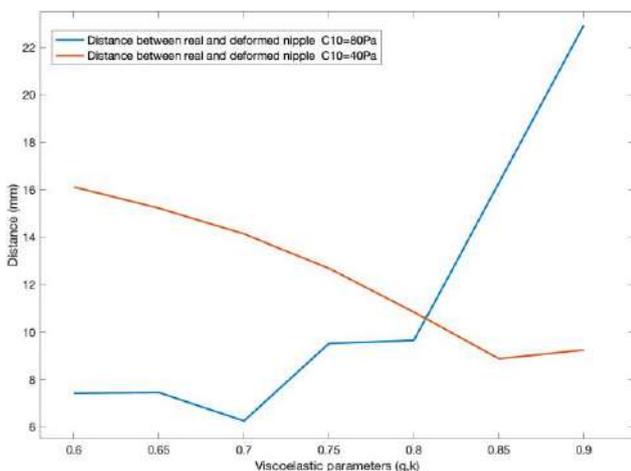


Figure 5. Viscoelastic parameters vs Nipple distance.

#### 3.2. Landmarks Distance

Two landmarks were identified in the prone and supine images: the position of the nipple and the centroid of the tumor extracted from the lesion segmentation. The Euclidean distance between the position of the supine landmarks in the CT and the position estimated by the simulation was calculated.

#### 3.3. Mean Distance Between Surfaces

The mean distance between surfaces has been calculated as the average value of the minimum distance between each node on the frontal surface portion of the breast in supine position and the frontal surface of the simulation mesh.

#### 3.4. Breast Overlapping

The percentage of overlapping was calculated between the supine breast mask from the CT and the deformed mask obtained from the simulation mesh.

Table 1 and Table 2 present the results obtained for the two cases. Visual results are shown in Figure 6 and Figure 7.

#### • Case 1

$C_{10}$ (Pa)	g;k	Nipple distance (mm)	Tumor distance (mm)	Surface distance (mm)	Overlapping (%)
80	0.7	6.26	8.83	11.47	75.82
40	0.85	8.89	10.11	10.48	75.61
40	0.8	10.86	3.13	10.85	77.65

Table 1. Results Case 1.

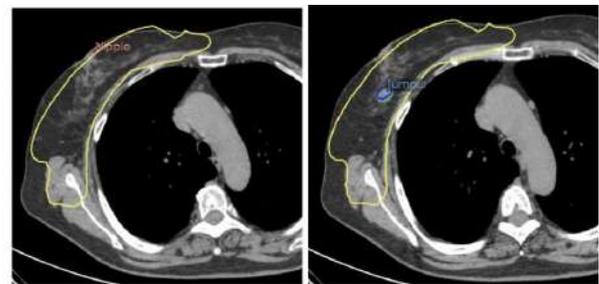
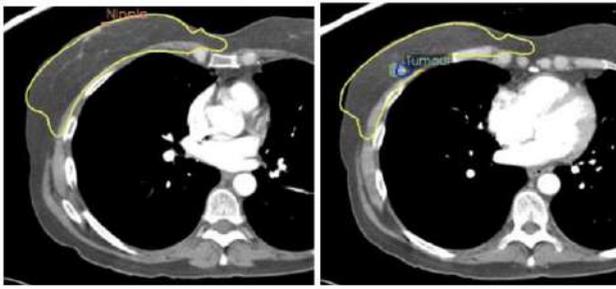


Figure 6. Results of simulation for Case 1 with  $C_{10}=40\text{Pa}$  and  $g=k=0.8$ . Left: supine CT image, estimated nipple position and deformed mesh (yellow). Right: Supine CT image, supine tumor (light blue), estimated tumor (blue) and deformed mesh (yellow).

#### • Case 2

$C_{10}$ (Pa)	g;k	Nipple distance (mm)	Tumor distance (mm)	Surface distance (mm)	Overlapping (%)
80	0.7	18.17	10.49	7.99	71.51
40	0.85	15.74	7.48	6.86	74.58
40	0.8	15.21	5.40	6.30	74.83

Table 2. Results Case 2.



**Figure 7.** Results of simulation for Case 2 with  $C_{10}=40\text{Pa}$  and  $g=k=0.8$ . Left: supine CT image, estimated nipple position, and deformed mesh (yellow). Right: Supine CT image, supine tumor (green), estimated tumor (blue) and deformed mesh (yellow).

#### 4. Discussion and Conclusions

This work proposes and validates a visco-hyperelastic biomechanical model to support preoperative localization of breast lesions. The patient-specific breast models were obtained from a preoperative MR image acquired in the prone position. A supine CT image has been used as reference in surgical position for the validation of the proposed method.

The analysis has been carried out in two cases: Case 1, a patient with large breast and Case 2, a patient with small breast. For both patients, better results in terms of distance between the predicted location of the tumor and the one segmented in the supine configuration have been obtained for a hyperelastic parameter  $c_{10}=40$  and viscoelastic parameter  $g=0.8$  with a mean value of 4.27 mm.

The results, obtained with a simplified model of the breast that does not consider its heterogeneous structure, are considered very promising. Similar results have been achieved with other proposals in the literature that also involve an image intensity-based registration step [10].

#### 5. Future work

The skin covering the breast provides some anatomical support to the breast. Therefore, future work could obtain the relationship between the skin and the breast and get the most appropriate model of the skin.

It could also be studied the breast tissue as a heterogeneous tissue consisting of adipose and fibroglandular tissue, being able to estimate the proportion of both tissues when segmenting them and thus, defining a specific biomechanical model that considers material properties for each patient.

It should also be noted that the proportion of breast tissue varies according to the woman, her age, size of the breast, whether she is premenopausal or postmenopausal, and, in general, the specific constitution of the patient, which is why the results vary from patient to patient. An important breakthrough would be to optimize the model parameters to create patient-specific models.

#### References

- [1] International Agency for Research of Cancer. (Jan. 2022). Globocan 2020: New global cancer data [Accessed: Jun. 17, 2022]. <https://www.uicc.org/news/globocan-2020-new-global-cancer-data>
- [2] Siegel, R. L., Miller, K. D., & Jemal, A. (2016). Cancer statistics, 2016. *CA: a cancer journal for clinicians*, 66(1), 7–30.
- [3] Babarenda Gamage, T. P., Rajagopal, V., Nielsen, P. M., & Nash, M. P. (2011). Patient-specific modeling of breast biomechanics with applications to breast cancer detection and treatment. *Patient-Specific Modeling in Tomorrow's Medicine*, 379–412.
- [4] Tanner, C., Schnabel, J. A., Hill, D. L., Hawkes, D. J., Leach, M. O., & Hose, D. R. (2006). Factors influencing the accuracy of biomechanical breast models. *Medical physics*, 33(6Part1), 1758–1769.
- [5] Carter, T. J., Tanner, C., Crum, W. R., Beechey-Newman, N., & Hawkes, D. J. (2006). A framework for image-guided breast surgery. *Proceedings of the Third International Conference on Medical Imaging and Augmented Reality*, 203–210. [https://doi.org/10.1007/11812715\\_26](https://doi.org/10.1007/11812715_26)
- [6] Del Palomar, A. P., Calvo, B., Herrero, J., López, J., & Doblaré, M. (2008). A finite element model to accurately predict real deformations of the breast. *Medical engineering & physics*, 30(9), 1089–1097.
- [7] Lee, A. W., Rajagopal, V., Gamage, T. P. B., Doyle, A. J., Nielsen, P. M., & Nash, M. P. (2013). Breast lesion colocalisation between x-ray and mr images using finite element modelling. *Medical image analysis*, 17(8), 1256–1264.
- [8] Mira, A., Carton, A. K., Muller, S., & Payan, Y. (2018). A biomechanical breast model evaluated with respect to MRI data collected in three different positions. *Clinical Biomechanics*, 60, 191–199.
- [9] Rajagopal, V., Lee, A., Chung, J.-H., Warren, R., Highnam, R. P., Nash, M. P., & Nielsen, P. M. (2008). Creating individual-specific biomechanical models of the breast for medical image analysis. *Academic radiology*, 15(11), 1425–1436.
- [10] Han, L., Hipwell, J., Mertzaniidou, T., Carter, T., Modat, M., Ourselin, S., & Hawkes, D. (2011). A hybrid fem-based method for aligning prone and supine images for image guided breast surgery. *2011 IEEE international symposium on biomedical imaging: from nano to macro*, 1239–1242
- [11] Hipwell, J. H., Vavourakis, V., Han, L., Mertzaniidou, T., Eiben, B., & Hawkes, D. J. (2016). A review of biomechanically informed breast image registration. *Physics in Medicine & Biology*, 61(2), R1.
- [12] Slicer, 3. (Apr. 2022). 3d slicer image computing platform [Accessed: Jun. 18, 2022]. <https://www.slicer.org>
- [13] Taylor, Z. A., Comas, O., Cheng, M., Passenger, J., Hawkes, D. J., Atkinson, D., & Ourselin, S. (2009). On modelling of anisotropic viscoelasticity for soft tissue simulation: Numerical solution and gpu execution. *Medical image analysis*, 13(2), 234–244.
- [14] Vavourakis, V., Eiben, B., Hipwell, J. H., Williams, N. R., Keshtgar, M., & Hawkes, D. J. (2016). Multiscale mechanobiological finite element modelling of oncoplastic breast surgery—numerical study towards surgical planning and cosmetic outcome prediction. *PloS one*, 11(7), e0159766.

# Estudio preliminar de la variación del rango de movimiento y la actividad muscular en pacientes de avanzada edad usando el andador SWalker

V. Costa Cortez<sup>1,2</sup>, L. Perea<sup>3</sup>, A. Velásquez<sup>3</sup>, S. López López<sup>4</sup>, H. Pareja Galeano<sup>5</sup>, C. Sánchez López de Pablo<sup>2</sup>, R. Raya López<sup>1,2</sup>

<sup>1</sup> Werium Assistive Solutions Ltd., Madrid, España, {vanina.costa, rafael.raya}@weriumsolutions.com

<sup>2</sup> Departamento de Tecnologías de la Información, Universidad San Pablo-CEU, CEU Universities, Madrid, España, {vanina.costacortez, cristina.sanchezlopezpablo, rafael.raya@ceu}@ceu.es

<sup>3</sup> Albertia Servicios Sociosanitarios S.A., Madrid, España., {lperea, avelasquez}@albertia.es

<sup>4</sup> Escuela de Doctorado e Investigación, Universidad Europea, Madrid, España, sergio.upgo@gmail.com

<sup>5</sup> Departamento de Educación Física, Deporte y Motricidad Humana, Universidad Autónoma de Madrid, España, helios.pareja@uam.es

## Resumen

*Este trabajo tiene como objetivo presentar las métricas rango de movimiento y actividad muscular durante la marcha de sujetos mayores de 65 años usando la plataforma de rehabilitación SWalker ante las distintas configuraciones de descarga de peso y velocidad que posee el dispositivo. El estudio se realizó con 21 sujetos del centro Albertia Majadahonda. En este artículo se presenta el análisis preliminar de los datos de los participantes que mejor capacidad funcional de deambulación poseían según los test Up and Go y SPPB, realizados previamente. Los datos obtenidos en este estudio pretenden contribuir con futuros protocolos clínicos que contemplen el uso del andador SWalker, así como con la literatura científica en el uso de dispositivos robóticos de rehabilitación para usuarios de edad avanzada.*

## 1. Introducción

La fractura de cadera es una de las lesiones más comunes en los adultos mayores de 65 años [1]. Es considerada la lesión más grave ocasionada por una caída, teniendo lugar en el 95% de los casos [2]. Este suceso, provoca una pérdida de movilidad y una reducción en la función de la marcha. La alta mortalidad y los costes derivados que presenta la fractura de cadera son indicadores alarmantes para el sistema sanitario [3]. Para los adultos mayores, el riesgo de fallecimiento es un 30% mayor que lo que cabría esperar tan solo un año después del ingreso.

La rehabilitación robótica puede proporcionar una terapia integradora implicando a diferentes grupos óseos y musculares de forma precisa mientras se realizan tareas funcionales. En esta línea han sido desarrollados varios exosqueletos y andadores robóticos como Indego [4], ReWalk [5] o Andado [6]. Varios estudios han demostrado la eficacia de estos dispositivos en neurorrehabilitación [7], pero no en lesiones traumáticas como la fractura de cadera. Así mismo, su diseño tampoco se dirige a la rehabilitación de esta patología específicamente en el adulto mayor. Estos fueron los motivos que llevaron a la creación del andador

SWalker, desarrollado y validado para la rehabilitación de fractura de cadera [8].

### 1.1. Descripción del dispositivo SWalker

Este dispositivo cuenta con un sistema de tracción que permite su movimiento en todas las direcciones y a distintas velocidades. Un sistema de soporte de peso eleva al paciente descargándole parcialmente del peso sobre sus piernas, favoreciendo así la movilidad temprana. El rango de movimiento (ROM) en flexoextensión de cadera, indicador principal de la evolución funcional del paciente, se mide gracias a dos potenciómetros [8]. Tras su validación clínica, se desarrolló una segunda versión del prototipo (ver Figura 1), con la que se llevó a cabo este estudio.

### 1.2. Efecto de la descarga de peso y la velocidad de marcha en la deambulación

Los sistemas de soporte parcial de peso corporal (PBWS) han demostrado ser prometedoras herramientas de rehabilitación para personas con discapacidades neurológicas. Han sido publicados numerosos artículos que estudian la influencia de la descarga de peso en las características de la marcha con distintos dispositivos, a diferentes velocidades e incluso comparando la marcha sobre suelo y sobre cinta o tapiz rodante.

La descarga de peso y la velocidad de marcha son dos parámetros clave para la terapia y configurables en el dispositivo SWalker. Por ello es de especial interés el estudio del efecto de estas variables en los parámetros de la marcha (ROM y actividad muscular) del adulto mayor para ofrecer un tratamiento más preciso con SWalker ante la fractura de cadera.

Los estudios clínicos llevados a cabo hasta la fecha coinciden en que, a mayor soporte de peso, menor es el ROM y menor es la longitud del paso. Así mismo, menor es la actividad muscular. Se ha demostrado que en

condición de soporte de peso bajo y velocidades medias y altas, tanto la estructura temporal del patrón de la marcha, como el patrón de activación muscular se mantienen normales [9]. A partir de un soporte de peso superior al 30% los patrones de marcha se ven alterados y no se contribuye a la recuperación de la marcha. En cuanto a los efectos de la variación de la velocidad de marcha, los estudios no reportan cambios significativos en la estructura temporal de la marcha, con soportes de peso inferiores al 20% y a la velocidad de confianza del sujeto pero sí a velocidades superiores [10].

Estos estudios se han realizado, en su mayoría, con sujetos sanos, con una media de edad de 33 años. En la mayor parte de las investigaciones que se han realizado, se utilizan dispositivos para rehabilitación de patologías neurológicas que funcionan sobre cinta rodante y no sobre suelo [11]. Apenas hay estudios que involucren a personas mayores de 65 años y no existen estudios en los que se prueben plataformas sobre suelo para rehabilitar fractura de cadera. Las personas mayores evitan un mayor balanceo y por lo tanto velocidades altas, así ganan estabilidad y una mayor confianza al andar. Esto produce rigidez en los miembros inferiores, que está directamente relacionada con el incremento de co-contracciones de los grupos musculares antagonista/agonista [12].

El presente trabajo es un estudio preliminar que tiene como objetivo la observación de las variaciones rango de movilidad de cadera y comportamiento muscular durante la marcha dentro de SWalker con pacientes de edad avanzada como efecto de la descarga de peso y las variaciones en la velocidad. Este estudio se llevará a cabo con pacientes sin fractura de cadera como fase previa a pruebas clínicas posteriores que sí incluirán pacientes con fractura de cadera y que requiere un protocolo preciso de uso de SWalker así como recomendaciones terapéuticas.

## 2. Diseño del estudio

### 2.1. Instrumentación

La plataforma de rehabilitación utilizada en este trabajo fue el andador SWalker. La recogida de datos de ROM de cadera se realizó por medio de los potenciómetros que integra el dispositivo. La configuración y visualización de los datos se realizó desde la propia interfaz gráfica del andador.

El comportamiento muscular se recogió con electromiografía (EMG) de superficie. Las señales se obtuvieron gracias a un dispositivo vestible ultraligero (mDurance™) que integra un sensor inercial tridimensional con un electromiógrafo de dos canales [13]. El dispositivo es controlado por una aplicación móvil Android™ vía Bluetooth. Cada registro se almacena en el servicio en la nube del dispositivo. Allí, los registros pueden ser descargados para su posterior análisis.

### 2.2. Experimentación

En este estudio participaron 21 sujetos sin fractura de cadera (13 mujeres y 8 hombres) de  $85,43 \pm 6,23$  años y  $IMC = 27,28 \pm 4,13$ ; residentes en el centro Albetria Majadahonda. Los criterios de inclusión impuestos

establecieron que los sujetos debían ser mayores de 65 años, de una estatura inferior a 180 cm y peso inferior a 90 kg y que no tuvieran lesiones o incapacidades motoras en miembro inferior.

Las pruebas se dividieron en dos sesiones. La primera tuvo el objetivo de conocer el estado funcional y de fragilidad del paciente en actividades que implicaran miembro inferior. También averiguar la velocidad de marcha media de la muestra para configurar las velocidades de tracción del andador. Se realizó la prueba Timed Up and Go [14], el test de Guraknik o SPPB [15] que evalúan con una serie de tareas funcionales el equilibrio y la fragilidad. Además, el test de velocidad de marcha en diez metros (10MWT) [16].



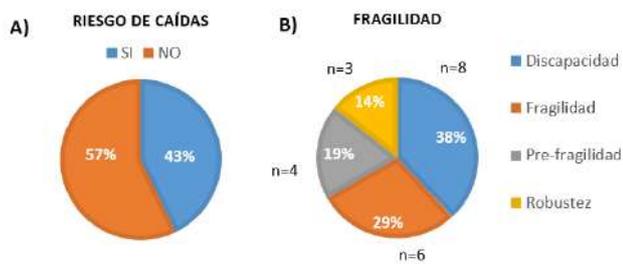
Figura 1. Participante utilizando SWalker durante el ensayo.

En la segunda sesión los sujetos utilizaron el andador SWalker completando un total de nueve recorridos cada uno, que se realizaron en un pasillo recto de diez metros de largo a las distintas velocidades (baja=0,08 m/s, media=0,2 m/s y alta=0,3 m/s) y distintas descargas de peso (0%, 25%, 50%). Todas las pruebas se llevaron a cabo en orden aleatorio. El fisioterapeuta fue el responsable de colocar los electrodos de EMG en el sujeto en el tibial anterior y gemelo.

El rango de velocidades configurado en el andador fue menor que la velocidad media alcanzada por los participantes en el estudio funcional (ver sección de resultados), siguiendo las recomendaciones del fisioterapeuta para preservar la confianza del sujeto dentro del andador.

## 3. Resultados

La Figura 2 muestra los resultados de los test de funcionalidad de la marcha de todos los sujetos (n=21). Se presentan las métricas *riesgo de caídas* y grado de *fragilidad*, evaluadas por los test Up and Go y SPPB respectivamente. La Tabla 1 presenta las velocidades media y máxima para los distintos grupos identificados según el grado de fragilidad.



**Figura 2.** Distribución de los pacientes según el resultado de los test *Up and Go* y *SPPB*, respectivamente.

Estado	Velocidad cómoda (m/s)	Velocidad máxima (m/s)
Discapacidad	0,44	0,6
Fragilidad	0,67	0,75
Pre-fragilidad	0,68	0,91
Robustez	0,87	1,03

**Tabla 1.** Velocidades promedio resultantes del test 10MWT.

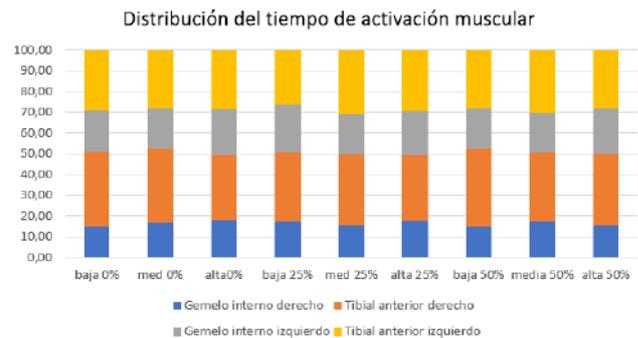
Como se observa en la Figura 2, son tres los sujetos que presentan “Robustez” en su deambulación. En este artículo se presentan los resultados de este grupo ( $n=3$ ;  $79,33 \pm 8,62$  años y  $IMC = 29,35 \pm 3,45$ ).

La Tabla 2 muestra un resumen de los rangos de flexión y extensión de ambas piernas para los pacientes sin fragilidad a velocidad media (0,2 m/s) y distintas descargas de peso.

Descarga de peso	Flexión promedio	Extensión promedio
0%	13°	6°
25%	8°	3°
50%	6°	2°

**Tabla 2.** Rangos promedios de flexión y extensión a velocidad media y distintas descargas de peso.

Los datos analizados de EMG fueron los tiempos de activación muscular. Durante las pruebas, se observó en tiempo real y con cada paciente que a mayor descarga de peso disminuye la activación de gemelos y tibial anterior. A una mayor velocidad de marcha, aumenta la amplitud de la señal muscular. En la Figura 3 puede verse que el tiempo de activación de los gemelos de ambas piernas es siempre inferior que el de los tibiales. También que el tibial derecho se activa más que el izquierdo y el gemelo izquierdo más que el derecho, aunque la distribución de tiempos es muy similar y puede decirse que hay simetría entre piernas.



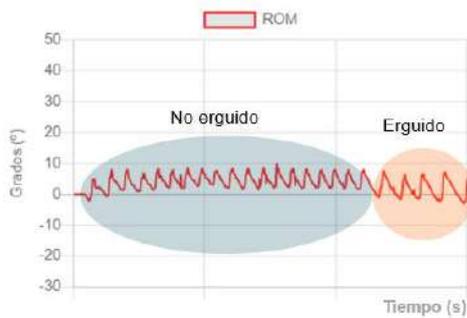
**Figura 3.** Gráfica de distribución del tiempo de activación de gemelo y tibial anterior de ambas piernas.

#### 4. Discusión y conclusión

Los resultados de ROM obtenidos en el presente estudio preliminar para los pacientes sin fragilidad tienen coherencia con los valores normativos en flexoextensión de cadera durante la marcha en el tiempo. Esto indica que el andador permite la obtención de un ROM de cadera fiable. Este hecho no se había comprobado hasta el momento con pacientes y proporciona una métrica muy valiosa para la rehabilitación de fractura de cadera, donde los pacientes poseen un ROM disminuido a causa de la lesión. No se observaron cambios significativos en el ROM con las variaciones de velocidad, pero sí con la descarga de peso. A la velocidad media y distintas descargas de peso, sí se observan variaciones, como muestra la Tabla 1, que además concuerdan con la bibliografía científica. A mayor soporte de peso, menor rango de movilidad. Hay una mayor variación entre 0% y 25% que entre 25% y 50%. De cara a aportar recomendaciones terapéuticas con SWalker, sería conveniente trabajar en un rango de soporte de peso inicial como mucho entre el 25% y 50% y no superior, ya que el ROM se ve muy limitado.

La adquisición del ROM también ha ayudado a detectar eventos sobre los que realizar recomendaciones para el equipo clínico para una rehabilitación eficaz. La más relevante de ellas afecta es relativa a la corrección postural. Al poseer el andador un arnés inguinal, los usuarios tienden a relajar la postura y apoyarse en el mismo, en posición semi sentado, con las piernas inclinadas hacia adelante. Este evento puede ser detectado y corregido por el clínico en la búsqueda de un entrenamiento eficaz para la recuperación temprana de la función de deambulación. La Figura 4 muestra cómo vería el clínico el efecto de una mala colocación o relajación de la postura del paciente (no erguido). Como puede apreciarse, la curva de rango articular queda por encima de la línea de calibración o línea neutra desde donde comenzó el paciente la deambulación.

En relación con los datos obtenidos de EMG, sería interesante ampliar el estudio de las señales hacia la cuantificación de simetrías y co-contracciones. Estas últimas se espera que estén presentes en el grupo de participantes que presenta mayor fragilidad.



**Figura 4.** Visualización del ROM de la cadera derecha de uno de los participantes que no mantiene una postura correcta.

Dado que este es un estudio preliminar, el número de pacientes fue reducido, pero sirvió para conocer los efectos del andador en la marcha y las métricas objetivo que se espera podrían alcanzar en las fases finales de la rehabilitación los pacientes. Convendría aumentar la muestra en futuros ensayos. Por otro lado, para algunos pacientes la velocidad de marcha fue inferior a su velocidad normal, por lo que habría que analizar si esto puede restringir el ROM y su actividad muscular.

Este trabajo contribuye a aportar recomendaciones en el uso del andador SWalker y presenta una nueva aportación a la literatura científica en el ámbito de dispositivos robóticos de rehabilitación para tercera edad.

### Agradecimientos

Agradecemos a los residentes y al equipo clínico de Albertia Majadahonda toda la colaboración recibida para este estudio. Este proyecto está financiado por el Centro para el Desarrollo Tecnológico Industrial (CDTI) bajo la referencia IDI-20220400.

### Referencias

[1] “Falls.” [Online]. Available: <https://www.who.int/news-room/fact-sheets/detail/falls>. [Accessed: 09-Sep-2022].

[2] R. Mitchell, L. Harvey, H. Brodaty, B. Draper, and J. Close, “Hip fracture and the influence of dementia on health outcomes and access to hospital-based rehabilitation for older individuals,” *Disabil. Rehabil.*, vol. 38, no. 23, pp. 2286–2295, 2016.

[3] A. Konnopka, N. Jerusel, and H. H. König, “The health and economic consequences of osteopenia- and osteoporosis-attributable hip fractures in Germany: Estimation for 2002 and projection until 2050,” *Osteoporos. Int.*, vol. 20, no. 7, pp. 1117–1129, Jul. 2009.

[4] R. J. Farris, H. A. Quintero, and M. Goldfarb, “Preliminary evaluation of a powered lower limb orthosis to aid walking in paraplegic individuals,” *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 19, no. 6, pp. 652–659, Dec. 2011.

[5] G. Zeilig, H. Weingarden, M. Zwecker, I. Dudkiewicz, A. Bloch, and A. Esquenazi, “Safety and tolerance of the ReWalk™ exoskeleton suit for ambulation by people with complete spinal cord injury: A pilot study,” *J.*

*Spinal Cord Med.*, vol. 35, no. 2, p. 96, Mar. 2012.

[6] H. J. A. van Hedel, I. Rosselli, and S. Baumgartner-Ricklin, “Clinical utility of the over-ground bodyweight-supporting walking system Andago in children and youths with gait impairments,” *J. NeuroEngineering Rehabil.* 2021 181, vol. 18, no. 1, pp. 1–20, Feb. 2021.

[7] D. Shi, W. Zhang, W. Zhang, and X. Ding, “A Review on Lower Limb Rehabilitation Exoskeleton Robots,” *Chinese J. Mech. Eng. (English Ed.)*, vol. 32, no. 1, Dec. 2019.

[8] V. Costa *et al.*, “Development and Clinical Validation of a Rehabilitation Platform for Hip Fracture in Elderly Population,” *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 30, pp. 1340–1349, 2022.

[9] K. Van Kammen, A. Boonstra, H. Reinders-Messelink, and R. Den Otter, “The combined effects of body weight support and gait speed on gait related muscle activity: A comparison between walking in the lokomat exoskeleton and regular treadmill walking,” *PLoS One*, vol. 9, no. 9, 2014.

[10] E. E. Thomas, G. De Vito, and A. Macaluso, “Physiological costs and tempo-spatial parameters of walking on a treadmill vary with body weight unloading and speed in both healthy young and older women,” *Eur. J. Appl. Physiol.*, vol. 100, no. 3, pp. 293–299, Jun. 2007.

[11] S. Apte, M. Plooij, and H. Vallery, “Influence of body weight unloading on human gait characteristics: A systematic review,” *J. Neuroeng. Rehabil.*, vol. 15, no. 1, pp. 1–18, 2018.

[12] H. J. Lee, W. H. Chang, B. O. Choi, G. H. Ryu, and Y. H. Kim, “Age-related differences in muscle co-activation during locomotion and their relationship with gait speed: a pilot study,” *BMC Geriatr.*, vol. 17, no. 1, pp. 1–8, Jan. 2017.

[13] A. Molina-Molina *et al.*, “Validation of mDurance, A Wearable Surface Electromyography System for Muscle Activity Assessment,” *Front. Physiol.*, vol. 11, Nov. 2020.

[14] D. Podsiadlo and S. Richardson, “The Timed ‘Up & Go’: A Test of Basic Functional Mobility for Frail Elderly Persons,” *J. Am. Geriatr. Soc.*, vol. 39, no. 2, pp. 142–148, Feb. 1991.

[15] L. Rodríguez-Mañas *et al.*, “An evaluation of the effectiveness of a multi-modal intervention in frail and pre-frail older people with type 2 diabetes - the MID-Frail study: Study protocol for a randomised controlled trial,” *Trials*, vol. 15, no. 1, pp. 1–9, Jan. 2014.

[16] G. Scivoletto, F. Tamburella, L. Laurenza, C. Foti, J. F. Ditunno, and M. Molinari, “Validity and reliability of the 10-m walk test and the 6-min walk test in spinal cord injury patients,” *Spinal Cord* 2011 496, vol. 49, no. 6, pp. 736–740, Jan. 2011.

# Señales Biomédicas II

---

Jueves, 24 de noviembre de 2022

# Improving the quality of combined TMS-EEG neural recordings: artefact removal and time analysis

G. Mijancos-Martínez<sup>1,3</sup>, A. Bachiller<sup>1,3</sup>, I. Fernández-Linsenbarth<sup>4</sup>, J.F. Alonso<sup>1,2,3</sup>, S. Romero<sup>1,2,3</sup>, V. Molina<sup>4,5</sup> and M.A. Mañanas<sup>1,2,3</sup>

<sup>1</sup> Biomedical Engineering Research Centre (CREB), Department of Automatic Control (ESAII), Universitat Politècnica de Catalunya (UPC), Barcelona, Spain, {gema.mijancos, alejandro.bachiller, joan.francesc.alonso, sergio.romero-lafuente, miguel.angel.mananas}@upc.edu

<sup>2</sup> CIBER of Bioengineering, Biomaterials and Nanomedicine (CIBER-BBN), Madrid Spain

<sup>3</sup> Institute of Research Sant Joan de Déu, Barcelona, Spain

<sup>4</sup> Department of Psychiatry, Medicine Faculty, University of Valladolid, Valladolid, Spain, {vicente.molina, ines.fernandez.linsenbarth}@uva.es

<sup>5</sup> Institute of Neuroscience of Castilla y León (INCYL), University of Salamanca, Salamanca, Spain

## Abstract

*Transcranial magnetic stimulation and electroencephalography (TMS-EEG) recordings are key to perform a direct evaluation of cortical function in a non-invasive and task-free manner. For the past few years, it has been used to study different brain disorders, such as schizophrenia (SCZ). Nonetheless, TMS-EEG signals present a huge artefact due to the magnetic pulse making its cleaning a challenge.*

*In this study, TMS-EEG is used to assess SCZ patients following different types of pulses. The objective of the study is twofold: to propose a different approach to obtain a clean signal; and to characterise it to compare the inhibitory activity of different groups of subjects regarding the TMS-evoked potentials (TEPs). The novelty independent component representation of TMS-EEG has led us to obtain a clean signal whose characterisation has allowed us to study the inhibitory activity. TEPs obtained match those from the literature: healthy control group shows more inhibition than the schizophrenic patients' group. Allowing the validation of the proposed TMS-EEG cleaning methodology.*

## 1. Introduction.

Schizophrenia (SCZ) is a chronic brain disorder, characterised by cognitive alterations, that affects 20 million people worldwide, according to the World Health Organisation [1]. To date, there is no cure to such disorder but different treatments, such as psychotherapy and drug treatments, are used to mitigate the symptoms [2-3].

Transcranial Magnetic Stimulation (TMS) is a novel, non-invasive method of altering the neuronal excitability via electromagnetic pulses that is increasingly being used in neuropsychiatric research and clinical psychiatry [4]. The combination of both TMS and electroencephalography (EEG), TMS-EEG, allows to further characterize the brain function. Nonetheless, when the magnetic pulse is triggered, a large artefact with amplitudes values 4 order of magnitude larger than normal EEG is produced, masking the signal of interest and making it impossible to recover some part of it. Nowadays, its removal is still a challenge [5-6].

TMS-EEG data has been used to understand SCZ providing a better insight on the brain interactions. The

TMS pulse activates the inhibitory interneurons, among others, allowing the assessment of the cortical inhibition processes [5]. Thus, it has been employed not only for the characterisation of the EEG regarding the cortical excitability and inhibition but also to study the oscillatory activity as well as the connectivity of inter- and intra-hemispheric regions [7].

Although there is no unique brain area to stimulate, the dorsolateral prefrontal cortex (DLPFC) is a great candidate since it seems to be closely related with SCZ [2,7]. In addition, apart from the motor cortex (which has also been widely studied), the TMS-evoked potentials (TEPs) are well characterized in the aforementioned region, compared to other non-motor areas [8].

There exist different procedures to administer the pulse, depending on: i) the number of consecutive pulses, either single pulse (SP) or two paired pulses (PP); ii) the intensity of the pulses (always referred as a percentage of the resting motor threshold (RMT)); and iii) the interstimulus interval in the case of PP (ISI), the time that passes between two consecutive pulses. In a paired-pulse signal, the ISI can be of a few milliseconds (4 to 6ms) or of a greater amount of time (more than 100ms). The short-interval intracortical inhibition (SICI) protocol has a short ISI. The conditioning pulse (CS, the first one) has a subthreshold intensity whereas the test pulse (TS, the second pulse) has a suprathreshold intensity [7]. On the other hand, the long-interval intracortical inhibition (LICI) protocol has a larger ISI and both pulses have a suprathreshold intensity. Following these protocols (SICI and LICI) and combining SP with PP it is possible to assess the level of inhibition of the subject through the TMS-EEG signal [9-10].

Therefore, in this work we propose a framework to discard the TMS artefact on EEG data, and then we hypothesise that it is possible to evaluate the differences between type of pulses and subjects, in terms of inhibitory activity, by characterising a processed TMS-EEG signal.

## 2. Materials and methods

### 2.1. Signal acquisition

The signals analysed in this study were acquired at Hospital Clínico Universitario of Valladolid (HCUV) with a 64-channel EEG amplifier (BrainVision actiCHamp 64) following the 10-10 International System. Sixty-one EEG channels were recorded (using channel Cz as a reference) along with the vertical and horizontal electrooculograms (VEOG and HEOG). The sampling frequency was set to 25 kHz.

Database was composed by 31 subjects: 16 SCZ patients (chronic and first-episode) and 15 healthy controls (HC). For the HC group, the 37 % were male and the mean age was  $25.6 \pm 10.5$  years whereas for the SCZ patients' group the 47 % were male and the mean age was  $31.8 \pm 10.5$  years.

A magnetic stimulator (MagVenture MagProX100) and a butterfly coil (MCF-B70) were used to stimulate the DLPFC.

Both types of pulses, SP and PP, were triggered. For the paired-pulse signals, the LICI protocol with an ISI of 100ms and an intensity of 120 % resting motor threshold (RMT) was employed. The SICI protocol was followed as well with an intensity of 80 % RMT and 120 % RMT for the CS and TS pulses, respectively. The ISI for SICI protocol had a value of 4ms. Nevertheless, in this work only the signals acquired following the SICI protocol will be assessed.

### 2.2. Pre-processing

EEG segments were selected with the TMS pulse(s) centred (-1 to 1s). The TMS pulse creates a huge artefact on the EEG [5-6], which was first identified by its great

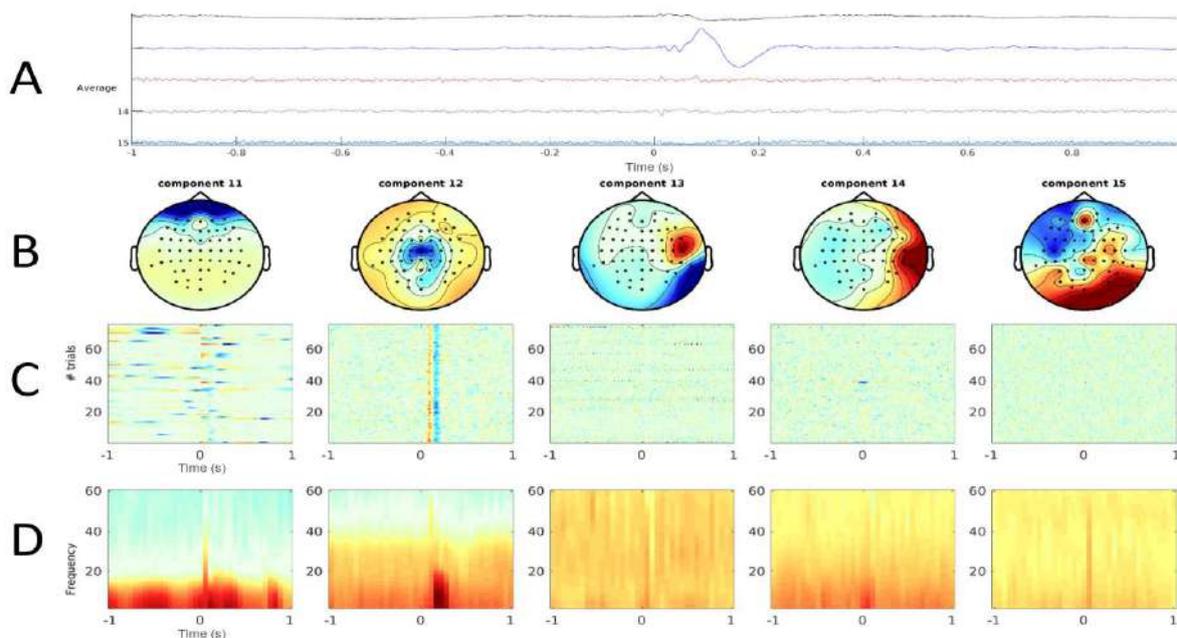
amplitude, deleted (from -1ms to 10ms related with TMS pulse onset) and cubic interpolated [6]

The second step was to reduce the contamination from the different artefacts (TMS-induced, muscle, ocular, auditory and noise related artefacts) on the data. Independent Component Analysis (ICA) was applied followed by a manual selection of the independent components (IC). Manual selection is usually performed based on either the trial-averaged amplitude, the spatial distribution, or the time-frequency maps of the ICs [6,11-12]. To facilitate manual selection, we developed a graphic interface that evaluates the IC components in all these properties. Figure 1 shows the trial-averaged amplitude, its spatial distribution, the trial activation map and the time-frequency map. In this way, the selected ICs are based on more than one property, making its rejection more objective. Moreover, three different researchers did the IC selection following the criteria based on Rogasch et al., 2014 [6] and those IC marked as artefact from at least two experts were discarded.

After the removal of the ICs selected, the signal was reconstructed. Subsequently, a kurtosis-base automatic thresholding was used to identify bad channels, which were rejected and interpolated by their neighbours. Besides, the algorithm considers whether or not the neighbouring channels of a bad channel were valid, in order not to generate interpolation artefacts. Next, trial rejection was performed based on amplitude. Finally, signals were downsampled to 5kHz and band-pass filtered between 0.5 and 70Hz to obtain a clean signal.

### 2.3. Processing

The processing of the signal was performed in a region of interest (ROI) composed by the channels covering the DLPFC, namely: FP1, AF7, AF3, F7, F5, F3, F1, FC5,



**Figure 1.** Generated figure representing five IC to analyse them. A) trial-averaged temporal signal, B) spatial distribution, C) trial activation map and D) time-frequency map.

FC3, FC1 [13]. Thus, all the computations in this study took into consideration only these channels.

Averaging is needed to obtain the TEPs for each subject and type of pulse (SP and PP). First, an average across all trials and channels of the ROI was performed. Afterwards, a grand average was obtained by averaging signals across both type of subjects. Lastly, the TEPs were assessed in a predefined time windows based on the SICI SP signal and their latencies were determined for every different TEP.

To evaluate the inhibitory activity, the mean amplitude of the signal in each time window, and hence for each TEP, was calculated. With this measurement, a modulation ratio (the mean amplitude of the TEP in the PP signal divided by the mean amplitude of the TEP in the SP signal) was computed. In addition, for visual comparison, a subtraction (SP minus PP signals) was computed for each group of subjects.

Lastly, a Wilcoxon statistical test ( $\alpha=0.05$ ) was applied to compare the HC and the SCZ patients. All processing and statistical analysis were performed with the software MATLAB (R2021b; The Mathworks Inc., Natick, MA).

### 3. Results and discussion

After the pre-processing procedure, an artefact-free signal was obtained where the TMS-pulse was no longer present.

Figure 2 depicts an example of the result after the pre-processing steps, showing the differences between a raw signal and the pre-processed signal for the same channel and subject.

Figure 3 depicts the grand-average for SICI SP and SICI PP data. From these waves it was possible to determine the main TEPs: P30, N40, P55, N100, P140 and P200, matching the peaks described in the literature [6-7,9,14]. Comparing both types of pulses (SP and PP), the TEPs in the PP signals appeared a few milliseconds earlier than the ones in the SP signals. This might be due to the fact that the cortex was already excited by the CS, so the TEPs

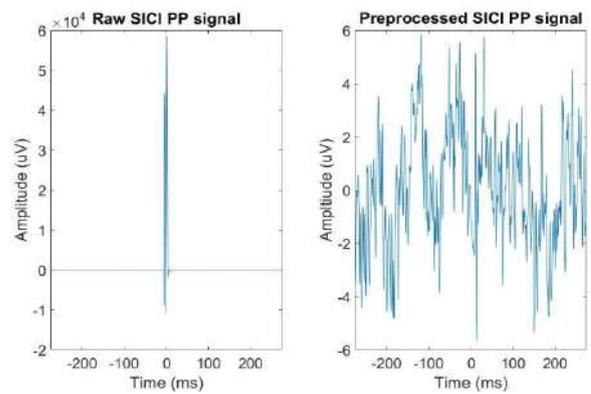


Figure 2. Left: raw signal (channel FC3) from one subject. Right: Same signal, pre-processed. Note the different scales.

emerged by the TS are influenced by the first pulse [9,12,15].

To assess the inhibition response on different subjects the difference between SP and PP was computed. If the resulting signal was closer to 0, it meant that there was no difference in the level of inhibition regarding the type of pulse (single or paired).

As Figure 3 shows, the values of the subtracted signal were close to 0 and the morphology for SCZ patients and HC was quite similar, which was expected. Nevertheless, as the values were alike, the level of inhibition seems not to change significantly neither within the type of pulse nor within the subject group.

Regarding the modulation ratio, values higher than 1 (the reference value) meant that there was cortical facilitation, otherwise there was cortical inhibition [13]. Our results showed that there was always inhibition (no ratio had a value greater than 1). As Figure 4 shows, the ratios in some TEP components (P30, P55, N100, and P200) are higher on HC group. It suggests more inhibition in controls than in patients, which was also expected [7,10,16-17].

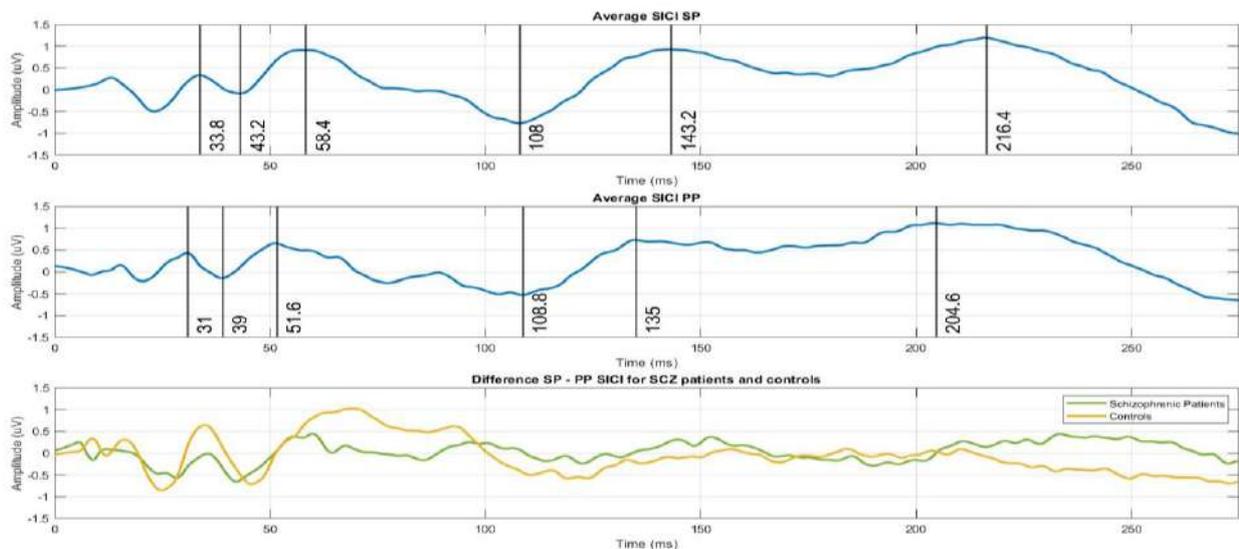


Figure 3. Top: Average signal across subjects of SICI SP paradigm with latencies of the TEPs marked. Centre: Average signal across subjects of SICI PP paradigm with latencies of the TEPs marked. Bottom: Subtracted signal from SCZ patients and HC.

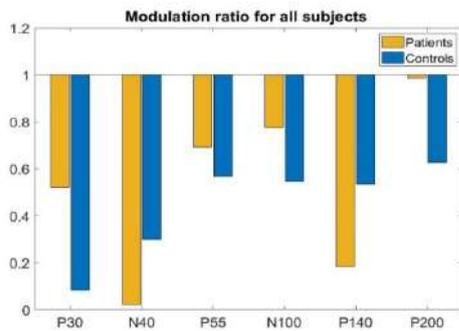


Figure 4. Modulation ratio for SCZ patients and HC.

Regarding the statistical analysis, the results did not reach significance but they showed a tendency, probably due to the sample size. Future studies will increase the number of SCZ patients and HC. Likewise, we will assess alternative methodologies to measure the effect of inhibition on paired-pulse protocols. Premoli et al. [18] proposed an alternative correction procedure on PP signal. It is expected that with more data and alternative corrections, the results will meet those from other studies and will reach significance.

#### 4. Conclusions

Firstly, a robust cleaning of the signal was possible thanks to this novel representation of ICs and the other steps of the proposed pre-processing (channel rejection and interpolation and trial rejection procedure).

Then, different TEPs were calculated from artefact-free EEG data. TEPs latencies and shape matched those from the literature. A comparison between subjects and type of pulses was performed, but no large differences were found between type of pulses. It appeared that the level of inhibition for the controls was slightly greater than the level of inhibition for the SCZ patients. Futures studies with an increased population will allow the assessment of significant differences with respect to cortical inhibition in SCZ patients compared to healthy controls.

#### Acknowledgments

This research was partially supported by: the ‘Ministerio de Ciencia e Innovación’ under project PID2020-117751RB; the ‘Junta de Castilla y León’ under project GRS 2121/AA/2020; and the ‘Instituto de salud Carlos III’ under project PI 18/00178. GMM is in receipt of a FI-2022 grant from ‘Agència de Gestió d'Ajuts Universitaris i de Recerca (AGAUR)’; IFL is in receipt of a predoctoral grant from the "Consejería de Educación, Junta de Castilla y León" (Spain), the European Social Fund (grant ID VA-183-18); finally, AB and JFA are Serra Húnter Fellows.

#### References

- [1] W. H. Organisation, “Web Page of World Health Organisation. Schizophrenia.” 2022. [Online]. Available: <https://www.who.int/news-room/fact-sheets/detail/schizophrenia>.
- [2] Edemann-Callesen H, Winter C, Hadar R. Using cortical non-invasive neuromodulation as a potential preventive treatment in schizophrenia - A review. *Brain Stimul.*, vol. 14, no. 3, 2021, pp. 643–651 (ISSN : 1876-4754).
- [3] Cole JC, Bernacki CG, Helmer A, Pinninti N, O’Reardon JP. Efficacy of transcranial magnetic stimulation (TMS) in the treatment of schizophrenia: A review of the literature to date. *Innov. Clin. Neurosci.*, vol. 12, no. 7–8, 2015 pp. 12–19 (ISSN: 2158-8341).
- [4] Haraldsson HM, Ferrarelli F, Kalin NH, Tononi G. Transcranial Magnetic Stimulation in the investigation and treatment of schizophrenia: A review. *Schizophr. Res.*, vol. 71, no. 1, 2004 pp. 1–16 (ISSN: 0920-9964).
- [5] Ilmoniemi RJ, Kičić D. Methodology for combined TMS and EEG. *Brain Topogr.*, vol. 22, no. 4, 2010, pp. 233–248 (ISSN: 0896-0267).
- [6] Rogasch NC, et al.. Removing artefacts from TMS-EEG recordings using independent component analysis: Importance for assessing prefrontal and motor cortex network properties. *Neuroimage*, vol. 101, 2014, pp. 425–439 (ISSN: 1095-9572).
- [7] Cao KX, et al.. TMS-EEG: An emerging tool to study the neurophysiologic biomarkers of psychiatric disorders. *Neuropharmacology*, vol. 197, no. April, 2021, p. 108574 (ISSN: 1873-7064).
- [8] Tremblay S, et al.. Clinical utility and prospective of TMS-EEG. *Clin. Neurophysiol.*, vol. 130, no. 5, 2019, pp. 802–844 (ISSN: 1872-8952).
- [9] Premoli I, et al.. Short-interval and long-interval intracortical inhibition of TMS-evoked EEG potentials. *Brain Stimul.*, vol. 11, no. 4, Jul. 2018, pp. 818–827 (ISSN: 1876-4754).
- [10] Li X, et al.. TMS-EEG research to elucidate the pathophysiological neural bases in patients with schizophrenia: A systematic review. *J. Pers. Med.*, vol. 11, no. 5, 2021 (ISSN: 2075-4426).
- [11] Cline CC, Lucas MV, Sun Y, Menezes M, Etkin A. Advanced artifact removal for automated TMS-EEG data processing. *Int. IEEE/EMBS Conf. Neural Eng. NER*, vol. 2021-May, pp. 1039–1042 (ISSN: 1948-3554).
- [12] Rogasch NC, Daskalakis ZJ, Fitzgerald PB. Cortical inhibition of distinct mechanisms in the dorsolateral prefrontal cortex is related to working memory performance: A TMS-EEG study. *Cortex*, vol. 64, Mar. 2015, pp. 68–77 (ISSN: 1973-8102).
- [13] Cash RFH, et al.. Characterization of Glutamatergic and GABA A-Mediated Neurotransmission in Motor and Dorsolateral Prefrontal Cortex Using Paired-Pulse TMS-EEG. *Neuropsychopharmacology*, vol. 42, no. 2, 2017, pp. 502–511 (ISSN: 1740-634X).
- [14] Hoy KE, Coyle H, Gainsford K, Hill AT, Bailey NW, Fitzgerald PB. Investigating neurophysiological markers of impaired cognition in schizophrenia. *Schizophr. Res.*, vol. 233, no. June, 2021, pp. 34–43 (ISSN: 1573-2509).
- [15] Opie GM, Sidhu SK, Rogasch NC, Ridding MC, Semmler JG. Cortical inhibition assessed using paired-pulse TMS-EEG is increased in older adults. *Brain Stimul.*, vol. 11, no. 3, May 2018, pp. 545–557 (ISSN: 1876-4754).
- [16] Kaskie RE, Ferrarelli F. Investigating the neurobiology of schizophrenia and other major psychiatric disorders with Transcranial Magnetic Stimulation. *Schizophr. Res.*, vol. 192, 2018, pp. 30–38 (ISSN: 1573-2509).
- [17] Mehta UM, Naik SS, Thanki MV, Thirthalli J. Investigational and Therapeutic Applications of Transcranial Magnetic Stimulation in Schizophrenia. *Curr. Psychiatry Rep.*, vol. 21, no. 9, 2019 (ISSN: 1535-1645).
- [18] Premoli I, et al.. Characterization of GABAB-receptor mediated neurotransmission in the human cortex by paired-pulse TMS-EEG. *Neuroimage*, vol. 103, Dec. 2014, pp. 152–162 (ISSN: 1053-8119).

# Análisis espectral de la actividad electroencefalográfica basal a nivel de fuente para la diferenciación de subtipos de migraña

J. Gómez Pilar<sup>1,2</sup>, V. Gutiérrez de Pablo<sup>1,2</sup>, V. Rodríguez González<sup>1,2</sup>, C. Gómez<sup>1,2</sup>, Á.L. Guerrero Peral<sup>3</sup>, M. Alves-Ferreira<sup>4,5</sup>, N. Pinto<sup>4,6</sup>, R. Hornero Sánchez<sup>1,2</sup>

<sup>1</sup> Grupo de Ingeniería Biomédica, Universidad de Valladolid, Valladolid, España

<sup>2</sup> Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), España

<sup>3</sup> Unidad de Cefaleas, Servicio de Neurología, Hospital Clínico Universitario de Valladolid, Valladolid, España

<sup>4</sup> Instituto de Investigação e Inovação em Saúde (i3S), Universidade do Porto, 4200-135 Porto, Portugal

<sup>5</sup> Instituto Ciências Biomédicas Abel Salazar (ICBAS), Universidade do Porto, 4050-313 Porto, Portugal

<sup>6</sup> Centro de Matemática da Universidade do Porto (CMUP), 4169-007 Porto, Portugal

## Resumen

*Las Clasificación Internacional de Cefaleas clasifica la migraña como crónica (MC) o episódica (ME) en función del número de días de dolor de cabeza al mes (más o menos de 15 días, respectivamente). A pesar de la gran prevalencia de la migraña y que el tratamiento puede variar en función de la situación de MC o ME, no disponemos de biomarcadores que nos permitan diferenciar entre estos subtipos más allá del convenio establecido por la CIC. El objetivo de este estudio es analizar la actividad electroencefalográfica (EEG) basal para identificar características que aporten nueva evidencia sobre el distinto sustrato neural de estos subtipos de migraña. Dada la literatura previa que relaciona la migraña con una fuerte especificidad topológica, se consideró el análisis a nivel de fuente (actividad cortical origen de la actividad eléctrica) en lugar de a nivel de sensor (actividad eléctrica en el cuero cabelludo). Para ello, se registró el EEG en 30 pacientes con MC y 30 con ME en estado de reposo. Tras el preprocesado y la resolución del problema inverso (localización de fuentes cerebrales), se realizó un análisis espectral en cada una de las 68 regiones de interés (ROIs) obtenidas. Los resultados mostraron una potencia significativamente mayor en los pacientes con MC, principalmente en la banda beta. Además, las ROIs identificadas con mayores diferencias entre grupos correspondieron a regiones relacionadas con el procesamiento del dolor, el estrés y la ansiedad. Estos resultados proporcionan nuevas evidencias para la diferenciación objetiva de estos tipos de migraña.*

## 1. Introducción

La migraña es una enfermedad con una alta prevalencia que afecta a alrededor del 15% de la población mundial [1]. Además, es la principal causa de discapacidad en personas menores de 50 años [2]. Su impacto socioeconómico se hace patente al comprobar que es el trastorno neurológico que genera la mayor cantidad de años vividos con discapacidad [3]. Todo ello implica un elevado sufrimiento y pérdida de oportunidades para los pacientes y sus familias. Por ello, un diagnóstico eficaz basado no solo en la sintomatología referida por los pacientes, sino también en sustratos neurológicos objetivos y reproducibles, ayudaría a proporcionar un tratamiento adecuado y personalizado a cada paciente.

La migraña se cataloga de acuerdo con la Clasificación Internacional de Cefaleas (CIC) como migraña crónica (MC) y migraña episódica (ME) [4]. Esta división se basa

únicamente en la frecuencia de aparición de la cefalea, definiendo la MC como una “cefalea que se presenta 15 o más días al mes durante más de tres meses y que, al menos ocho días al mes, tiene las características migrañosas” [4]. Por su parte, la ME se diagnostica cuando la frecuencia de la cefalea es menor a 15 días/mes. Puesto que no existen biomarcadores de ambos tipos de migraña, la comunidad científica se ha llegado a plantear si el constructo clínico propuesto por la CIC es procedente. A pesar de ello, como se refleja en las guías de práctica clínica [5], el tratamiento puede diferir entre los dos subtipos de migraña, lo que pone de manifiesto la necesidad de un diagnóstico personalizado basado en sustratos neurológicos objetivos.

Partiendo de la hipótesis razonable de que la MC y la ME deberían ser diferentes a nivel bioeléctrico, bioquímico y/o anatómico, un gran número de grupos de investigación han buscado en los últimos años biomarcadores específicos de estas entidades, como los niveles sanguíneos del péptido relacionado con el gen de la calcitonina (CGRP) [6], el depósito de hierro en la materia gris periacueductal en la resonancia magnética (RM) de difusión [7], o diferencias en la conectividad en la resonancia magnética de difusión (dMRI) [8]. Dado que el uso excesivo de medicación sintomática es uno de los factores de riesgo para la cronificación de la migraña [9], y que el uso de medicamentos es una causa bien conocida de cambios en los patrones de ondas eléctricas cerebrales [10], cabe preguntarse si la MC y la ME también podrían distinguirse a nivel funcional eléctrico, es decir, a través del análisis de la electroencefalografía (EEG).

Prácticamente no existen trabajos previos que analicen de manera directa o indirecta las diferencias electroencefalográficas o magnetoencefalográficas entre MC y ME. Sin embargo, estos estudios coinciden en que las diferencias aparecen en bandas frecuenciales rápidas, principalmente en la banda beta [11] [12], [13]. Así, un análisis espectral exhaustivo podría reforzar o contradecir las evidencias preliminares encontradas.

A diferencia de otras enfermedades neurológicas o psiquiátricas con alteraciones deslocalizadas, la migraña y en particular la diferenciación entre los subgrupos MC y ME parece residir en regiones cerebrales específicas [11]. Se ha especulado que estas diferencias podrían deberse al

proceso de cronificación del dolor [14], frecuentemente relacionado con el uso excesivo de medicamentos sintomáticos [9]. Por este motivo, la topología cerebral debe tenerse en cuenta en el estudio de la migraña.

Con estos precedentes, nuestra hipótesis de partida es que el contenido espectral del EEG en regiones específicas relacionadas con el procesamiento del dolor mostraría diferencias significativas entre la MC y la ME. Por tanto, el objetivo de este estudio es analizar la actividad EEG en el dominio frecuencial a nivel de fuente para identificar características objetivas que aporten nueva evidencia sobre el distinto sustrato neural de la MC y la ME.

## 2. Materiales y métodos

### 2.1. Participantes

En total, 60 pacientes fueron reclutados para este estudio. Los pacientes fueron reclutados en su primera visita ambulatoria (libres de tratamiento preventivo) en la Unidad de Cefalea del Hospital Clínico Universitario de Valladolid (España) y en el Instituto de *Investigação e Inovação em Saúde da Universidade do Porto* (Portugal). Dada la prevalencia desbalanceada de la migraña (aproximadamente, el 75% de los pacientes con migraña son mujeres [1]), la relación de factores hormonales en la cefalea y con el fin de evitar posibles sesgos debidos al sexo de los participantes, todos los pacientes fueron mujeres entre 18 y 40 años. Todas fueron diagnosticadas de acuerdo con los criterios de la CIC [4]. Se excluyeron del estudio las pacientes con otras enfermedades neurológicas o psiquiátricas, abuso de drogas o antecedentes de traumas que puedan afectar al EEG. A fin de evitar pacientes en situaciones limítrofes entre ambos tipos de migraña, las pacientes incluidas con ME presentaban menos de 10 días de cefalea al mes.

Tras el cribado, se reclutaron 30 pacientes con MC con edad mediana de 33.50 años y rango intercuartil (27.25, 38.75), así como 30 pacientes con ME con edad mediana 31.50 años y rango intercuartil (29.25, 35.75). La edad no mostró diferencias significativas entre ambos grupos. Estos pacientes fueron considerados por primera vez en el presente estudio, no solapándose total ni parcialmente con nuestro estudio previo [12].

### 2.2. Adquisición del EEG y preprocesado

Los registros de EEG se adquirieron usando un sistema de 32 canales (BrainVision), con electrodos activos colocados en un gorro elástico de acuerdo con el sistema internacional 10-10 y utilizando el electrodo Cz como referencia. La impedancia se mantuvo por debajo de 5k $\Omega$  durante el registro. Se registraron 10 minutos de actividad EEG basal mientras las participantes estaban sentadas con los ojos cerrados. Las señales se obtuvieron con una frecuencia de muestreo de 500 Hz.

Para la eliminación de ruido y artefactos que no correspondían con actividad neural, se llevó a cabo un preprocesado que constó de los siguientes pasos [12]. Primero, las señales se referenciaron a la actividad promedio de todos los sensores. Después, se restó la media de la señal con el fin de eliminar la componente continua.

A continuación, los datos se filtraron entre 0.4 y 98 Hz con un filtro de fase cero y de respuesta finita al impulso (FIR) de orden 2000 y ventana de Hamming. La frecuencia de la red eléctrica también fue eliminada mediante un filtro de ranura a 50 Hz (FIR, orden 2000, ventana Hamming). Las componentes relacionadas con el movimiento ocular, parpadeos y artefactos musculares fue eliminado tras un análisis de componentes independientes (ICA). Finalmente, la señal se segmentó en épocas de 5 segundos y se rechazaron aquellas que, tras su inspección visual, presentarían algún artefacto.

### 2.3. Localización de fuentes

Las señales a nivel de fuente se obtuvieron mediante el *toolbox Brainstorm*, documentado y extensamente validado [15]. Para ello, se estimó un modelo con 15000 fuentes mediante *Boundary Elements Method* (BEM) utilizando la plantilla ICBM152 (Montreal Neurological Institute) y el software OpenMEEG [16]. Las fuentes se restringieron a la corteza cerebral con una orientación normal a la misma [17]. Las 15000 fuentes se proyectaron en las 68 ROIs definidas por el atlas de Desikan-Killiany. Las posibles soluciones al problema inverso se restringieron mediante el método de norma mínima ponderada (wMNE), que minimiza la energía ponderando las fuentes profundas para facilitar su identificación [18].

### 2.4. Análisis espectral

En primer lugar, se calculó la densidad espectral de potencia (PSD) para cada época de 5 segundos como la transformada de Fourier de la función de autocorrelación. Así, cada período 2500 muestras de longitud se tradujo en una función de autocorrelación de 4999 muestras. Esto proporcionó una resolución espectral de 0.1 Hz.

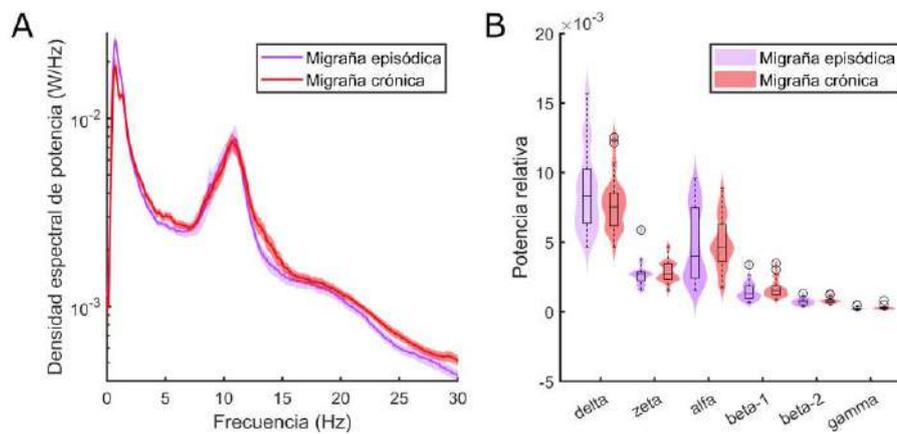
El contenido espectral se caracterizó mediante la potencia relativa (RP). La RP es una característica espectral clásica que representa el área bajo la curva de la PSD normalizada (ponderada para que la suma de sus coeficientes sea 1) en un rango específico de frecuencias. En este estudio, la RP fue calculada en cada una de las bandas convencionales del EEG: delta ( $\delta$ , 1-4 Hz), zeta ( $\theta$ , 4-8 Hz), alfa ( $\alpha$ , 8-13 Hz), beta-1 ( $\beta_1$ , 13-19 Hz), beta-2 ( $\beta_2$ , 19-30 Hz) y gamma ( $\gamma$ , 30-70 Hz). Este análisis se realizó de manera independientemente en cada una de las 68 ROIs obtenidas.

### 2.5. Análisis estadístico

Tras evaluar la normalidad (test de Shapiro–Wilk) y homocedasticidad (test de Levene) de los datos y comprobar que no se cumplían las hipótesis paramétricas, se optó por usar una prueba no paramétrica (test  $U$  de Mann-Whitney) para evaluar las diferencias entre grupos.

## 3. Resultados y discusión

En primer lugar, se optó por realizar un análisis global, es decir, evaluar el gran promedio de las 68 ROIs. La representación de la PSD de ambos grupos, así como la distribución de los valores de RP en cada banda, se muestran en la Figura 1. Salvo en la banda  $\alpha$ , los valores de RP muestran una típica disminución a medida que aumenta la frecuencia. Esto está en línea con el extendido supuesto de que el EEG tiene una distribución de potencia



**Figura 1.** Análisis espectral global (gran promedio de las 68 ROIs). (A) Densidad espectral de potencia para los grupos de ME y MC. La transparencia denota el intervalo de confianza al 95%. (B) Potencia relativa en cada banda del EEG. Los resultados no muestran diferencias significativas en ninguna banda.

similar al ruido rosa [19]. La banda  $\alpha$  es la excepción, mostrando un claro pico de potencia debido a que los registros se hicieron con los ojos cerrados [20]. En este análisis global, la RP no mostró diferencias significativas entre grupos en ninguna de las bandas del EEG.

A continuación, se procedió a un análisis local, evaluando cada una de las 68 ROIs por separado. En este caso, los resultados mostraron diferencias significativas en varias regiones de interés. El número de regiones con diferencias significativas entre grupos, así como su localización aproximada se muestran en la Tabla 1. En el caso de la banda  $\beta_1$ , por ser la banda con mayor número de ROIs con diferencias significativas (15 ROIs), se muestra su distribución espacial en la Figura 2. Cada centroide de las ROIs se representó con un nodo. Mientras que el color de cada nodo muestra el valor específico de RP, el tamaño de dicho nodo representa el grado de significancia en la comparativa entre ME y MC (a mayor significancia estadística, mayor tamaño de nodo). Solo se representaron aquellas ROIs con diferencias significativas entre grupos.

Para las bandas de frecuencia analizadas, se aprecia un solapamiento en las regiones con diferencias significativas entre grupos. Así, 7 de las 9 regiones que mostraron diferencias significativas en  $\beta_2$  también lo hicieron en  $\beta_1$ . Esto indica la alta especificidad que existe en los circuitos neurales involucrados en las diferencias entre grupos. Independientemente de la banda de frecuencia, las regiones con mayores diferencias significativas ( $p < 0.01$ ) son la ínsula, dos regiones del córtex cingulado anterior (ACC) y dos regiones cercanas al ACC (la región frontal media rostral, RMF). Mientras que la ínsula y el ACC son regiones asociadas con el procesamiento del dolor [21], el córtex RMF forma parte de la corteza prefrontal, involucrada en el estrés y la ansiedad [22].

Curiosamente, la banda beta (que fue la que mostró mayor número de regiones cerebrales con diferencias significativas) ha sido asociada tanto con el dolor [23], como con el estrés y la ansiedad [24]. Que los circuitos de dolor, ansiedad y estrés emerjan en la comparativa MC y ME no parece casual. Aunque de manera especulativa, se podría razonar que un mayor número de días de cefalea (como ocurre en MC) implica mayor estrés y ansiedad por

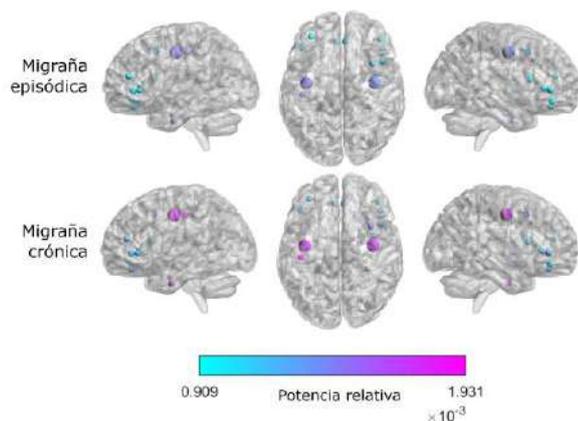
Banda	# ROIs	Topología
$\delta$	4	Regiones frontales interhemisféricas
$\theta$	9	Regiones frontales y temporales izquierdas
$\alpha$	0	-
$\beta_1$	15	Regiones frontales, centrales y temporales interhemisféricas
$\beta_2$	9	Regiones frontales y temporales izquierdas
$\gamma$	1	Región occipital derecha

**Tabla 1.** Distribución por bandas de frecuencia y localización espacial de las ROIs que mostraron diferencias significativas entre grupos.

parte del paciente. Esto produciría un aumento de la actividad, particularmente en la banda  $\beta$ . Los resultados mostraron precisamente estos resultados, es decir, una mayor RP, principalmente en banda  $\beta$ , en circuitos relacionados con el dolor, estrés y ansiedad.

En conjunto, los resultados sugieren que: (i) existen evidencias preliminares que confirman la diferencia entre los subtipos migrañosos de MC y ME; y (ii) estas diferencias parecen tener una alta especificidad, mostrando localización tanto frecuencial (banda  $\beta$ ) como espacial (circuitos relacionados con el dolor, el estrés y la ansiedad).

Entre las limitaciones del estudio se debe destacar, en primer lugar, que el tamaño muestral, aunque similar a estudios previos, es reducido. Esto ofrece una potencia estadística moderada y, por tanto, un tamaño del efecto medio. Importantes regiones cerebrales involucradas en las diferencias entre MC y ME podrían pasar desapercibidas por este hecho. Además, todos los pacientes fueron mujeres, por lo que se debe ser extremadamente cauto al extrapolar los resultados encontrados aquí a la población masculina. Finalmente, la población solo comprendía pacientes con menos de 10 días de cefalea al mes. La inclusión de pacientes con ME de alta frecuencia (entre 10 y 15 días de cefalea al mes) permitiría observar una evolución progresiva y no una puramente dicotómica de la enfermedad.



**Figura 2.** Distribución espacial de las ROIs que mostraron diferencias significativas en la banda  $\beta$ . El color indica la potencia relativa, mientras que el tamaño del nodo es proporcional al nivel de significancia estadística.

#### 4. Conclusiones

El análisis espectral de la actividad EEG a nivel de fuente identificó diferencias significativas entre MC y ME. Estas diferencias tienen un alto grado de especificidad, concentrándose en la banda  $\beta$  del EEG, así como en regiones cerebrales relacionadas con el procesamiento del dolor, el estrés y la ansiedad. Estos resultados proporcionan nueva evidencia para la diferenciación objetiva de estos tipos de migraña.

#### Agradecimientos

Esta investigación ha sido financiada por la Comisión Europea y el Fondo Europeo de Desarrollo Regional (FEDER) en el marco del proyecto “Análisis y correlación entre la epigenética y la actividad cerebral para evaluar el riesgo de migraña crónica y episódica en mujeres” (Programa de Cooperación Interreg V-A España-Portugal, POCTEP 2014–2020), por el Ministerio de Ciencia e Innovación / Agencia Estatal de Investigación /10.13039/501100011033/ y FEDER 'A way of making Europe' bajo los proyectos PDC2021-120775-I00, PID2020-115468RB-I00 y PGC2018-098214-A-I00, por 'CIBER en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN)' a través del 'Instituto de Salud Carlos III' y por fondos portugueses a través de FCT—Fundação para a Ciência e a Tecnologia/Ministério da Ciência, Tecnologia e Ensino Superior en el marco del proyecto POCI-01-0145-FEDER-029486 (PTDC/MEC-NEU/29486/2017).

#### Referencias

[1] L. J. Stovner, K. Hagen, M. Linde, and T. J. Steiner, “The global prevalence of headache: an update, with analysis of the influences of methodological factors on prevalence estimates,” *The Journal of Headache and Pain*, vol. 23, no. 1, p. 34, Dec. 2022.

[2] M. Ashina et al., “Migraine: epidemiology and systems of care,” *The Lancet*, vol. 397, no. 10283, pp. 1485–1495, Apr. 2021.

[3] T. Vos et al., “Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010,” *The Lancet*, vol. 380, no. 9859, pp. 2163–2196, Dec. 2012.

[4] “Headache Classification Committee of the International Headache Society (IHS) The International Classification of Headache Disorders, 3rd edition,” *Cephalalgia*, vol. 38, no. 1, pp. 1–211, Jan. 2018.

[5] S. Santos Lasaosa and P. Pozo Rosich, “Manual de Práctica Clínica en Cefaleas. Recomendaciones diagnósticoterpéuticas de la Sociedad Española de Neurología en 2020,” Madrid, 2020.

[6] C. Ramón, E. Cernuda-Morollón, and J. Pascual, “Calcitonin gene-related peptide in peripheral blood as a biomarker for migraine,” *Current Opinion in Neurology*, vol. 30, no. 3, pp. 281–286, Jun. 2017.

[7] K. Ito et al., “Detection of changes in the periaqueductal gray matter of patients with episodic migraine using quantitative diffusion kurtosis imaging: preliminary findings,” *Neuroradiology*, vol. 58, no. 2, pp. 115–120, Feb. 2016.

[8] Á. Planchuelo-Gómez, D. García-Azorín, Á. L. Guerrero, S. Aja-Fernández, M. Rodríguez, and R. de Luis-García, “Structural connectivity alterations in chronic and episodic migraine: A diffusion magnetic resonance imaging connectomics study,” *Cephalalgia*, vol. 40, no. 4, pp. 367–383, Apr. 2020.

[9] S. Cevali et al., “Family History for Chronic Headache and Drug Overuse as a Risk Factor for Headache Chronification,” *Headache: The Journal of Head and Face Pain*, vol. 49, no. 3, pp. 412–418, Mar. 2009.

[10] W. T. Blume, “Drug Effects on EEG,” *Journal of Clinical Neurophysiology*, vol. 23, no. 4, pp. 306–311, Aug. 2006.

[11] J. Gomez-Pilar, V. Martínez-Cagigal, D. García-Azorín, C. Gomez, Á. L. Guerrero, and R. Hornero, “Headache-related circuits and high frequencies evaluated by EEG, MRI, PET as potential biomarkers to differentiate chronic and episodic migraine: Evidence from a systematic review,” *The Journal of Headache and Pain (in press)*.

[12] J. Gomez-Pilar, D. García-Azorín, C. Gomez-Lopez-de-San-Roman, Á. L. Guerrero, and R. Hornero, “Exploring EEG Spectral Patterns in Episodic and Chronic Migraine During the Interictal State: Determining Frequencies of Interest in the Resting State,” *Pain Medicine*, vol. 21, no. 12, pp. 3530–3538, Dec. 2020.

[13] F. J. Hsiao et al., “Migraine chronification is associated with beta-band connectivity within the pain-related cortical regions: a magnetoencephalographic study,” *Pain*, vol. 162, no. 10, pp. 2590–2598, Oct. 2021.

[14] L. H. Schulte, A. Allers, and A. May, “Hypothalamus as a mediator of chronic migraine,” *Neurology*, vol. 88, no. 21, pp. 2011–2016, May 2017.

[15] F. Tadel, S. Baillet, J. C. Mosher, D. Pantazis, and R. M. Leahy, “Brainstorm: A User-Friendly Application for MEG/EEG Analysis,” *Computational Intelligence and Neuroscience*, vol. 2011, pp. 1–13, 2011.

[16] V. Rodríguez-González et al., “Consistency of local activation parameters at sensor- and source-level in neural signals,” *Journal of Neural Engineering*, vol. 17, no. 5, p. 056020, Oct. 2020.

[17] L. Douw, D. Nieboer, C. J. Stam, P. Tewarie, and A. Hillebrand, “Consistency of magnetoencephalographic functional connectivity and network reconstruction using a template versus native MRI for co-registration,” *Human Brain Mapping*, vol. 39, no. 1, pp. 104–119, Jan. 2018.

[18] F.-H. Lin, T. Witzel, M. S. Hämäläinen, A. M. Dale, J. W. Belliveau, and S. M. Stufflebeam, “Spectral spatiotemporal imaging of cortical oscillations and interactions in the human brain,” *Neuroimage*, vol. 23, no. 2, pp. 582–595, Oct. 2004.

[19] S. Wainio-Theberge, A. Wolff, J. Gomez-Pilar, J. Zhang, and G. Northoff, “Variability and task-responsiveness of electrophysiological dynamics: Scale-free stability and oscillatory flexibility,” *Neuroimage*, vol. 256, p. 119245, Aug. 2022.

[20] M. X. Cohen, *Analyzing neural time series data: Theory and practice*. 2014.

[21] H. Nakata, K. Sakamoto, and R. Kakigi, “Meditation reduces pain-related neural activity in the anterior cingulate cortex, insula, secondary somatosensory cortex, and thalamus,” *Frontiers in Psychology*, vol. 5, Dec. 2014.

[22] D. S. Jacobs and B. Moghaddam, “Medial prefrontal cortex encoding of stress and anxiety,” 2021, pp. 29–55.

[23] J. Bismuth, F. Vialatte, and J.-P. Lefaucheur, “Relieving peripheral neuropathic pain by increasing the power-ratio of low- $\beta$  over high- $\beta$  activities in the central cortical region with EEG-based neurofeedback: Study protocol for a controlled pilot trial (SMRPain study),” *Neurophysiologie Clinique*, vol. 50, no. 1, pp. 5–20, Feb. 2020.

[24] H. Díaz M., F. M. Cid, J. Otárola, R. Rojas, O. Alarcón, and L. Cañete, “EEG Beta band frequency domain evaluation for assessing stress and anxiety in resting, eyes closed, basal conditions,” *Procedia Computer Science*, vol. 162, pp. 974–981, 2019.

# Protocolo de registro de EEG para identificación de patrones cerebrales asociados a anosmia por COVID-19

D.I. Gracia Laso<sup>1</sup>, T. Candela Muries<sup>1</sup>, E. Iáñez Martínez<sup>1,2</sup>, M. Ortiz García<sup>1,2</sup>, J.M. Azorín Poveda<sup>1,2,3</sup>

<sup>1</sup> Brain-Machine Interface Systems Lab, Universidad Miguel Hernández de Elche, Elche, España  
{dgracia, tcandela, eianez, mortiz, jm.azorin}@umh.es

<sup>2</sup> Instituto de Investigación en Ingeniería de Elche – I3E, Universidad Miguel Hernández de Elche, Elche, España

<sup>3</sup> Valencian Graduated School and Research Network of Artificial Intelligence - ValGRAI, España

## Resumen

*La pérdida del olfato es un síntoma común de la infección por COVID-19. A pesar de que en la mayoría de los casos el paciente se recupera tras haber pasado menos de un mes después de la infección, también hay pacientes cuyos problemas olfativos persisten después de un año. Las causas de la disfunción olfativa siguen siendo desconocidas. Una mayor comprensión de la fisiología involucrada permitirá desarrollar nuevos procedimientos con los que ayudar a los pacientes a recuperar su sentido del olfato normal. Por ello se ha diseñado un novedoso protocolo para estudiar las señales de EEG mientras se presenta un estímulo olfativo (los Sniffin' Sticks) apoyado en la adquisición de la señal respiratoria para una correcta segmentación temporal. Con dicho protocolo se pretende caracterizar, en trabajos futuros, los patrones cerebrales que identifiquen las áreas de la corteza afectadas como resultado de la infección por COVID-19.*

## 1. Introducción y objetivo

A finales del año 2019 se declaró una nueva pandemia mundial, causada por una nueva variante de coronavirus (SARS-CoV-2) que se denominó COVID-19 (coronavirus disease 2019). Inicialmente fue considerada como una patología respiratoria asociada con síntomas neumológicos. Sin embargo, posteriores estudios [1], [2] detectaron condiciones neurológicas causadas por la enfermedad. Se observaron diferencias en los patrones cerebrales, detectados mediante electroencefalografía (EEG), en pacientes infectados por el COVID-19 y con síntomas neurológicos causados por dicha enfermedad [1]. Algunos de estos patrones fueron encontrados en las áreas frontales del cerebro, lo cual refuerza la teoría que hipotetiza la entrada del virus SARS-CoV-2 al sistema nervioso central mediante las estructuras olfativas y se esparce a partir de los lóbulos frontales [2]. Esta hipótesis coincide con el significativo porcentaje de pacientes que han sufrido una pérdida del olfato, anosmia, en los estados iniciales de la infección [3].

La causa de la disfunción olfativa sigue siendo desconocida. Aunque ha sido probado que no existe una relación directa entre la severidad de la infección o la carga vírica [4], [5]; algunas teorías proponen que la afección se debe de forma directa a lesiones celulares o indirectamente a la inflamación causada por la infección viral de las rutas olfativas [6]–[9].

Si bien en la mayoría de los casos los síntomas revierten de forma espontánea en el margen de semanas [4], [10], [11],

este no siempre es el caso. Un estudio multicéntrico [4], en el que participaron 18 hospitales europeos, estableció que el 24.1 % de los pacientes seguían padeciendo anosmia un mes después de haber sido infectados, un 15.3 % tras 2 meses y un 4.7 % después de medio año. Un estudio del hospital universitario de Estrasburgo [12] evaluó 97 pacientes con pérdida aguda del olfato, tras un año solo dos de ellos no habían recuperado la capacidad olfativa (uno presentaba un umbral olfativo anormal y el otro no identificaba correctamente los olores, es decir, padecía parosmia). Las teorías que defienden la aparición de la disfunción olfativa a causa de los daños que la infección patológica provoca en la mucosa nasal fallan en explicar la anosmia patológica, ya que el daño celular en este tejido es generalmente reparable [4], [10]–[12].

La falta de evidencia en los tratamientos actuales para la anosmia o hiposmia, pérdida parcial de la capacidad olfativa, [13] generan la necesidad de desarrollar nuevos procedimientos que permitan a los pacientes recuperar su capacidad olfativa normal. Sin embargo, para ello es necesario aumentar el conocimiento sobre las anomalías en la fisiología afectada en primer lugar.

Así el objetivo de este estudio es diseñar un protocolo para obtener señales EEG al presentar un estímulo olfativo. Este protocolo será usado en un futuro en pacientes anósmicos con el objetivo de encontrar patrones en su análisis que permitan determinar las áreas del córtex que podrían estar dañadas a causa de la infección del COVID-19.

## 2. Materiales y métodos

### 2.1. Participantes

2 sujetos, un hombre y una mujer con una edad media de 22 años y desviación de 4.24 años, sin enfermedades reportadas, participaron voluntariamente en los experimentos, habiendo firmado previamente un consentimiento informado. Todos los procedimientos fueron aprobados por la Oficina de Investigación Responsable de la Universidad Miguel Hernández de Elche (DIS.JAP.03.21).

### 2.2. Equipo

Los datos de EEG se registraron utilizando el casco g.Nautilus PRO FLEXIBLE (g.tec medical engineering GmbH, Austria) con un sistema de 34 electrodos húmedo no invasivos. Los 32 electrodos seleccionados para la adquisición se colocaron siguiendo la distribución 10-10

del sistema internacional, siendo estas las posiciones: FP1, FP2, AF3, AF4, F7, F3, FZ, F4, F8, FC5, FC1, FC2, FC6, T7, C3, CZ, C4, T8, CP5, CP1, CP2, CP6, P7, P3, PZ, P4, P8, PO7, PO3, PO4, PO8 y OZ. Los otros dos electrodos se emplearon como electrodos de tierra y referencia, colocados en el lóbulo de la oreja derecha y en la posición del sistema internacional CPZ, respectivamente. Los datos se registraron con una frecuencia de muestreo de 500 Hz.

Se emplea la plataforma Matlab para el almacenamiento, procesamiento y análisis de datos EEG, siendo los datos obtenidos, amplificados, digitalizados y transmitidos inalámbricamente mediante el propio equipo EEG.

La señal respiratoria se registró utilizando la banda BioHarness 3.0 (Zephyr Technology Corporation, EE.UU.) con una frecuencia de muestreo de 18 Hz. Los datos fueron adquiridos mediante la plataforma propia Bluetooth BioHarness Test Application. Se empleó además esta interfaz para representar las señales. La señal representada cuenta con un retraso de 3 segundos y los datos son actualizados cada segundo.

Como estímulos olfativos se empleó la prueba Sniffin' Sticks (Burghardt, Alemania). La versión empleada fue la prueba de olor extendida de 2-feniletanol, que constaba de tres pruebas:

1) Test de umbral, con 48 Sniffin' Sticks (16 con diluciones de 2-feniletanol y 32 neutros a diferentes concentraciones, teniendo la fragancia menor concentración en el stick número 16 y mayor en el número 1).

2) Test de discriminación, con 48 Sniffin' Sticks (16 tripletes en los que cada uno de ellos se conforman por dos sticks con el mismo olor y otro con un aroma diferente -target-).

3) Test de identificación, con 16 Sniffin' Sticks (con diferentes olores cotidianos).

### 2.3. Diseño experimental

El experimento se dividió en ocho fases consecutivas.

En las fases 1, 4, 6 y 8, el sujeto permaneció dos minutos relajado con los ojos cerrados. Estas fases proporcionaron información sobre el estado basal de los datos del EEG. También fueron relevantes como tiempo de descanso de las diferentes pruebas para evitar la saturación olfativa.

En la segunda fase se realizó una prueba inicial del test de umbral en la que se le presentaban al sujeto los olores presentes en los diferentes tipos de sticks, neutros y con aroma, empleando el triplete de sticks con mayor concentración. El objetivo de esta etapa es presentar al sujeto los diferentes aromas que deberá identificar en la etapa siguiente. Inicialmente se presentó el stick con aroma durante 5 segundos. Tras pasar 30 segundos, permitiendo la adaptación olfativa, se presentaron los otros dos sticks, neutros, durante 5 segundos cada uno con un intervalo de tiempo de 5 segundos entre ellos, necesarios para asegurar que no hubiera saturación olfativa.

El test de umbral se realizó en la tercera fase. Esta prueba permite determinar la concentración mínima que el sujeto es capaz de oler. Las pruebas se realizaron siguiendo un

procedimiento de escalera, es decir, la concentración de la barra se aumentó o disminuyó después de cada prueba. Las concentraciones se aumentaron o disminuyeron un stick a la vez, excepto en el primer set donde el aumento fue de dos sticks. La prueba se dividió en siete series de trials, cambiando la dirección de la escalera al finalizar cada serie, lo cual ocurría cuando se alcanzaba un punto de inflexión. En las series con sentido creciente (números impares), la concentración aumentaba cuando el sujeto no era capaz de identificar correctamente qué stick tenía aroma y se producía un punto de inflexión si el stick se identificaba correctamente dos veces seguidas. En los conjuntos con una dirección decreciente (números pares), la concentración se reducía después de cada correcta identificación (de nuevo, era necesario identificar correctamente dos veces) y el punto de inflexión se definió al producirse una identificación errónea. La primera serie comenzaba con el stick de menor concentración y las siguientes series empezaban con el stick un número mayor si la dirección de la serie anterior era decreciente o un número menor si la dirección anterior era creciente. El número de trials no se encontraba prefijado ya que la duración de cada una de las siete series se definió en función de las respuestas correctas del participante. La secuencia de tiempos fue similar a la utilizada en la segunda fase, si bien esta vez todos los intervalos entre sticks fueron de aproximadamente 5 segundos y se agregó un descanso en el cual el sujeto daba una respuesta sobre qué stick pensaba que tenía el aroma. El orden de presentación de los tripletes de sticks fue aleatorizado de forma fija (se aplicará la misma aleatorización para todos los sujetos futuros). La puntuación de esta prueba se calculó como el promedio del número de sticks para los últimos 4 puntos de inflexión.

En la quinta fase se realizó el test de discriminación, donde el sujeto debía identificar qué stick presentaba un olor diferente. La secuencia utilizada fue la misma que en la tercera fase: Los tres sticks se presentaron durante 5 segundos cada uno, con un intervalo de 5 segundos entre ellos. Posteriormente, el usuario respondió qué stick pensaba que tenía un olor diferente. Un intervalo mínimo de 30 segundos debía transcurrir antes de empezar el siguiente trial. La puntuación de esta prueba se calculó como el número de respuestas correctas.

El test de identificación se realizó en la séptima fase. En esta prueba, el sujeto debía reconocer 16 olores cotidianos, presentándose un único stick en cada trial durante 5 segundos. Posteriormente debía identificar el olor percibido entre 4 opciones mostradas durante el intervalo entre trials de 30 segundos. La puntuación de esta prueba se calculó como el número de respuestas correctas.

El participante necesitaba una señal para saber cuándo debía comenzar la inspiración. Convencionalmente esta indicación se realiza mediante un comando vocal, pero como podría producir un potencial evocado adicional al olfativo, se decidió advertir tocando al sujeto en una zona con menor representación en la corteza somatosensorial como es el hombro.

La ventana de tiempo de interés para el análisis de los datos del EEG se definió a partir de la señal respiratoria,

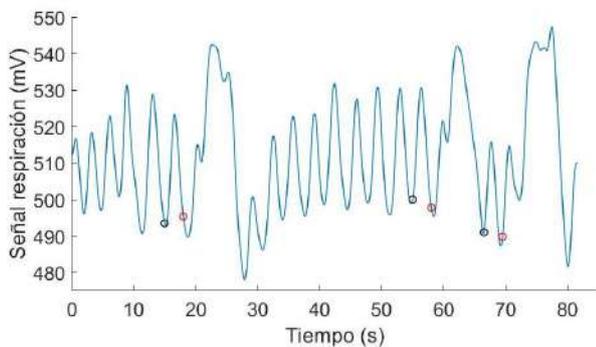
buscándose evaluar la respuesta del sujeto mientras se encontraba inspirando. Para identificar la inspiración a evaluar de forma más sencilla se guardó además un trigger adicional. Su señal es correspondiente a una galga extensiométrica posicionada en el hombro del sujeto y activada al producirse la señal de inicio al participante.

Dado que algunos de los Sniffin' Sticks se distinguen visualmente, las dos primeras pruebas requieren que el sujeto tenga los ojos cerrados. Sin embargo, se exigió a la persona que cerrara los ojos en todas las pruebas para adquirir las señales del EEG en las mismas condiciones.

Si bien esta prueba permite evaluar el desempeño de cada fosa nasal de forma individual y conjunta, solo se realizó la última opción para evitar elongar demasiado la duración de la prueba. Al presentar el stick, este se colocó a una distancia de 2 cm de las fosas nasales, evitando siempre tocar la nariz o las manos del operador ni del sujeto.

Después de cada prueba, incluso si el sujeto no estaba seguro, debían obligatoriamente dar una respuesta. La cooperación de la persona fue muy importante ya que se trata de una prueba semiobjetiva y las respuestas aleatorias influirán en la fiabilidad de los resultados [14].

Debido al retraso presente en la señal respiratoria representada, se estudia el ritmo de respiración del participante con el objetivo de buscar el momento idóneo para dar la señal de inicio de la inspiración. Es decir, se identifica el patrón morfológico del instante de la señal que se encuentra a 3 segundos del inicio de la inspiración, este suele ser el inicio de la inspiración de la respiración anterior. Para clarificación, véase la Ilustración 1.



**Ilustración 1.** Señal de respiración del sujeto C01 durante la fase 2. Los puntos rojos representan el instante en el que se activó la galga extensiométrica, mientras que los puntos negros representan la zona de la señal que se observaba en el instante en el que se activó la galga.

### 3. Análisis

Todos los análisis se realizaron de forma offline.

Inicialmente se resampleó la señal de respiración para que tuviese la misma frecuencia de muestreo que la señal EEG.

Seguidamente se definió una ventana temporal de 4 segundos. Se consideró como instante inicial la primera muestra después de haberse activado la galga en la que el participante se encontrase inspirando por al menos medio segundo.

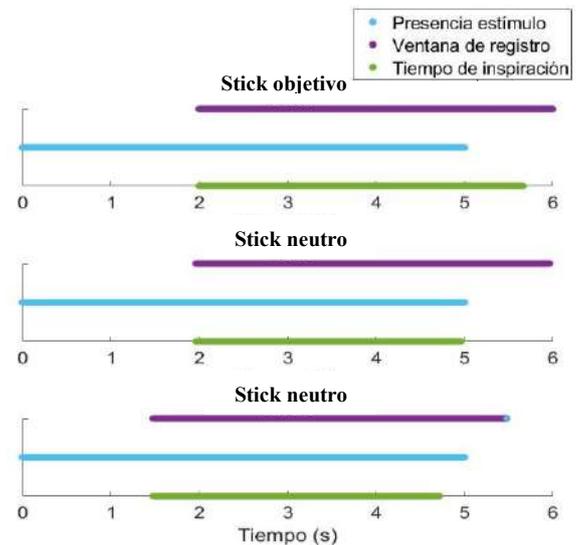
Para medir la eficiencia del proceso experimental se obtuvieron tres variables: 1) diferencia temporal entre el inicio del estímulo (momento en el que se activa la galga) y el inicio de la ventana temporal obtenida, 2) porcentaje de la ventana temporal en el que se estaba aplicando el estímulo olfativo y 3) porcentaje de la ventana temporal en el que el sujeto se encontraba inspirando.

### 4. Resultados

Se obtienen las 3 variables para cada una de las pruebas de los participantes, se representa un ejemplo en la Ilustración 2. Puesto que la cantidad de muestras registradas varía en función del desempeño del sujeto, se calculan los valores de cada inspiración, pero se presenta la media para poder realizar una comparación entre sujetos (véase la Tabla 1).

**Tabla 1.** Datos obtenidos por cada sujeto para cada una de las tres variables.

	C01	C02
Variable 1	0.783 ± 0.802 s	0.436 ± 0.692 s
Variable 2	93.58 ± 10.93 %	96.56 ± 7.88 %
Variable 3	78.55 ± 13.16 %	94.22 ± 9.99 %



**Ilustración 2.** Ilustración de las medidas analizadas en la fase 2 del sujeto C02. Este caso presenta unos valores de 2, 1.97 y 1.48 segundos para la variable 1; 75.05, 75.95 y 88.05 % para la variable 2; y 89.55, 72 y 78.5 % para la variable 3, empleando el stick rojo, verde y azul como estímulo, respectivamente.

### 5. Discusión

A partir de los resultados obtenidos se observa que el protocolo diseñado se adapta a los objetivos presentes, pues permite el registro de señales encefalográficas mientras se emplean estímulos olfativos, definiéndose las ventanas de interés mientras el sujeto se encuentra inspirando con el estímulo presente.

La diferencia temporal entre la señal dada al participante y el momento en el que inicia la inspiración es de 1-1.5 segundos en el 75 % de las muestras. No se ha observado un efecto temporal en los valores, puesto que la actuación del sujeto no se ve afectada por el número de trials realizados previamente. Tampoco se observa una notable

diferencia interpersonal, pues ambos participante presentan una distribución similar. Por lo tanto, la diferencia temporal se debe principalmente a la fase del proceso de respiración en la que se encuentra el participante cuando se le da la señal; siendo los tiempos mínimos cuando el sujeto se encontraba inspirando y máximos cuando acababa de iniciar la expiración. Por ello es crucial una monitorización de la respiración simultánea a la realización de la prueba experimental.

Si bien en el manual de empleo de Sniffin' Sticks [15] se indica un tiempo de presentación de los sticks de unos 4 segundos, el hecho de haberlo aumentado a 5 segundos posibilita que el tiempo de estímulo durante la ventana de registro se vea afectado en menor medida. Esto se observa en el hecho de que en el 75 % de los casos el porcentaje de la ventana temporal en el que se estaba aplicando el estímulo olfativo es superior al 90 %. Así pues, para mejorar los valores de esta variable es de mayor interés mejorar la variable 1 que aumentar el tiempo de estimulación, factor que además adelantaría la saturación olfativa.

Finalmente, el porcentaje de la ventana temporal en el que el sujeto se encontraba inspirando se encuentra especialmente influenciado por el factor interpersonal, pues cada persona presenta un ritmo de respiración diferente. El sujeto C01 presenta valores menores para la variable 3 pues realiza respiraciones de menor duración. A lo largo de la prueba experimental la respiración media de C01 era de  $2.7 \pm 0.99$  s, mientras que la de C02 era de  $3.39 \pm 1.39$  s. Dichos valores son mucho menores a las inspiraciones registradas, ya que estas se encuentran forzadas por el sujeto. Sin embargo los valores obtenidos se consideran suficientes, pues permiten al sujeto reconocer el estímulo.

Tras el diseño y evaluación del protocolo experimental, se analizará en trabajos futuros a sujetos con pérdida olfativa como causa de infección por COVID en búsqueda de patrones cerebrales identificativos.

## Agradecimientos

Esta investigación se ha llevado a cabo en el marco del proyecto "Estudio exploratorio de los efectos del uso de la neuroestimulación no invasiva en pacientes con anosmia persistente post COVID" (GVA-COVID19/2021/062), financiado por la Consellería de Innovación, Universidades, Ciencia y Sociedad Digital (Generalitat Valenciana, España).

## Referencias

- [1] G. Vellieux, R. Sonnevill, S. Vledouts, P. Jaquet, A. Rouvel-Taltec, y M. P. d'Ortho, «COVID-19-Associated Neurological Manifestations: An Emerging Electroencephalographic Literature», *Frontiers in Physiology*, vol. 11, 2021, doi: 10.3389/fphys.2020.622466.
- [2] A. S. Galanopoulou *et al.*, «EEG findings in acutely ill patients investigated for SARS-CoV-2/COVID-19: A small case series preliminary report», *Epilepsia Open*, vol. 5, n.º 2, 2020, doi: 10.1002/epi4.12399.
- [3] A. A. Agyeman, K. L. Chin, C. B. Landersdorfer, D. Liew, y R. Ofori-Asenso, «Smell and Taste Dysfunction in Patients With COVID-19: A Systematic Review and Meta-analysis», *Mayo Clin. Proc.*, vol. 95, n.º 8, 2020, doi: 10.1016/j.mayocp.2020.05.030.
- [4] J. R. Lechien *et al.*, «Prevalence and 6-month recovery of olfactory dysfunction: a multicentre study of 1363 COVID-19 patients», *J. Intern. Med.*, vol. 290, n.º 2, 2021, doi: 10.1111/joim.13209.
- [5] R. H. W. Cho *et al.*, «COVID-19 Viral Load in the Severity of and Recovery From Olfactory and Gustatory Dysfunction», *Laryngoscope*, vol. 130, n.º 11, 2020, doi: 10.1002/lary.29056.
- [6] M. F. V. V. Aragão, M. C. Leal, O. Q. Cartaxo Filho, T. M. Fonseca, y M. M. Valença, «Anosmia in COVID-19 associated with injury to the olfactory bulbs evident on MRI», *Am. J. Neuroradiol.*, vol. 41, n.º 9, 2020, doi: 10.3174/ajnr.A6675.
- [7] M. Eliezer *et al.*, «Loss of smell in patients with COVID-19: MRI data reveal a transient edema of the olfactory clefts», *Neurology*, vol. 95, n.º 23, 2020, doi: 10.1212/WNL.0000000000010806.
- [8] T. Laurendon *et al.*, «Bilateral transient olfactory bulb edema during COVID-19-related anosmia», *Neurology*, vol. 95, n.º 5, 2020, doi: 10.1212/WNL.0000000000009850.
- [9] J. Meinhardt *et al.*, «Olfactory transmucosal SARS-CoV-2 invasion as a port of central nervous system entry in individuals with COVID-19», *Nat. Neurosci.*, vol. 24, n.º 2, 2021, doi: 10.1038/s41593-020-00758-5.
- [10] S. T. Moein, S. M. R. Hashemian, P. Tabarsi, y R. L. Doty, «Prevalence and reversibility of smell dysfunction measured psychophysically in a cohort of COVID-19 patients», *Int. Forum Allergy Rhinol.*, vol. 10, n.º 10, 2020, doi: 10.1002/alr.22680.
- [11] C. Hopkins, P. Surda, E. Whitehead, y B. N. Kumar, «Early recovery following new onset anosmia during the COVID-19 pandemic - An observational cohort study», *J. Otolaryngol. - Head Neck Surg.*, vol. 49, n.º 1, 2020, doi: 10.1186/s40463-020-00423-8.
- [12] M. Renaud *et al.*, «Clinical Outcomes for Patients with Anosmia 1 Year after COVID-19 Diagnosis», *JAMA Netw. Open*, vol. 4, n.º 6, 2021, doi: 10.1001/jamanetworkopen.2021.15352.
- [13] S. Boesveldt *et al.*, «Anosmia-A clinical review», *Chemical Senses*, vol. 42, n.º 7, 2017, doi: 10.1093/chemse/bjx025.
- [14] B. Iravani, «Novel Measure of Olfactory Bulb Function in Health and Disease», 2021.
- [15] «How to use the Sniffin' Sticks - MediSense | Smelltest.eu». <https://www.smelltest.eu/en/smell-and-taste/how-to-use-the-sniffin-sticks/> (accedido ago. 31, 2022).

# Sincronía, metaestabilidad y dimensión fractal del EEG durante diferentes estados conscientes

A. J. Ibáñez-Molina<sup>1</sup>, S. Iglesias-Parro<sup>1</sup>

<sup>1</sup> Departamento de Psicología, Universidad de Jaén, Jaén, España, aibanez@ujaen.es, siglesia@ujaen.es

## Resumen

*En este estudio hemos investigado la electrofisiología global del cerebro durante diferentes estados conscientes no patológicos. El objetivo fue describir la sincronía global del EEG, su metaestabilidad y su complejidad medida con la dimensión fractal de Higuchi durante una tarea de procesamiento audiovisual. En esta tarea, los participantes sanos veían fragmentos cortos de películas y otros fragmentos con ruido blanco audiovisual. Durante el visionado, se preguntaba a los participantes si estaban atendiendo al vídeo (Focus) o pensando en otras cosas (mind wandering, MW). El objetivo general fue comprobar si el funcionamiento global del cerebro es diferente para distintos estados subjetivos y si está modulado por el tipo de estimulación. Los resultados obtenidos indicaron que los estados subjetivos Focus vs MW sólo fueron fisiológicamente distintos durante la presentación de películas, pero no ante el ruido blanco. En concreto, la sincronía y la metaestabilidad aumentaron cuando los participantes informaron estar atendiendo a los fragmentos de películas. En el resto de condiciones se observó un nivel más bajo en estas medidas. Por otro lado, en esta misma condición (presentación de películas) se observó que la dimensión fractal disminuyó, mostrando un patrón similar de datos, pero en la dirección contraria. En resumen, nuestro estudio muestra que los estados conscientes cuando están orientados hacia una estimulación externa con sentido están diferenciados globalmente del resto de estados. Esta diferenciación consiste en más sincronía global acompañada de una menor complejidad, y una mayor fluctuación en los estados globales tal como indica el aumento en la metaestabilidad.*

## 1. Introducción

El interés por el funcionamiento del cerebro como un sistema complejo ha aumentado en los últimos años [1]. El estudio de la activación neural local se ha sustituido por un enfoque basado en redes neuronales de carácter macroscópico [2]. En este contexto, se realizan estudios empíricos y simulaciones del cerebro completo que exploran la dinámica global del sistema como resultado de la interacción compleja de sus distintas regiones o áreas locales [3].

El enfoque holístico del funcionamiento cerebral se está aplicando con éxito en el estudio de la consciencia humana [4]. Las teorías más aceptadas consideran que la consciencia humana no está localizada en un área específica del cerebro, sino que requiere de una interacción compleja entre distintas zonas entre las que se encuentran el área temporal, parietal y occipital [5]. Es por eso que consideramos importante que las aproximaciones empíricas para entender mejor la consciencia y los distintos estados conscientes deben incluir indicadores que caractericen en la mayor medida posible el sistema al completo.

Entre las aproximaciones metodológicas más importantes para caracterizar el cerebro como un sistema completo se encuentran los conceptos de conectividad y de metaestabilidad [6]. La conectividad a nivel macroscópico se utiliza normalmente para calcular si dos señales distantes en la corteza están estadísticamente relacionadas [7]. En el estudio de la consciencia se ha propuesto la sincronía como mecanismo básico de conectividad ya que posibilita la integración y coordinación entre la actividad neuronal de áreas distantes del cerebro [8]. La sincronía se mide calculando si dos ritmos que provienen de dos conjuntos neuronales débilmente ligados, producen oscilaciones que se encuentran en fase [9].

Además, la metaestabilidad es un concepto que en neurociencia se utiliza más recientemente para referirse a el nivel dispersión estadística que presenta la sincronía global del sistema [6].

Además de la sincronía y la metaestabilidad el cerebro como sistema completo se suele caracterizar por medidas de complejidad como por ejemplo la dimensión fractal o la entropía [10]. Estas medidas, en general, cuantifican la irregularidad o el grado de desorden presente en la señal. La dimensión fractal es una de las medidas de complejidad más utilizadas y es sensible al número de componentes presentes en la señal. En este contexto, un componente se refiere a un ritmo o actividad oscilante que se produce a una escala temporal determinada. Cuantas más oscilaciones a diferentes escalas tenga la señal, más dimensión fractal presentará.

En este estudio tratamos de investigar la actividad global del cerebro a través del EEG durante distintos estados conscientes. En lugar de tomar un enfoque de carácter clínico, abordamos el problema desde un punto de vista cognitivo, y exploramos estados conscientes relacionados con la atención o distracción de distintos estímulos audiovisuales. La distracción durante cualquier tarea, en concreto, se suele denominar *mind wandering* (MW) y hace referencia a aquella pérdida atencional que hace divagar a una persona cuando está realizando una tarea o prestando atención a un evento [11]. Nuestra hipótesis de partida es que distintos estados conscientes asociados a la estimulación audiovisual tendrán una caracterización específica en el funcionamiento global de la corteza cerebral. En general, esperamos que las medidas de sincronía y dimensión fractal den lugar a patrones opuestos, ya que en otros estudios que hemos realizado hemos observado que la sincronía suele estar negativamente relacionada con la complejidad de la señal [12]. Además, esperamos que los estados conscientes que

requieren de una mayor especificidad como los de atención a vídeos de películas presentarán mayor sincronía que los estados más difusos como pueden ser aquellos en los que los participantes están distraídos en *MW*.

## 2. Método

### 2.1. Participantes

Un total de 23 personas (20 mujeres) participaron en nuestro estudio. Todas eran estudiantes del Grado de Psicología en la Universidad de Jaén, con una edad media de 20,2 años ( $SD=2,6$  años). Todas eran diestras, con visión/audición normal o corregida.

### 2.2. Registro de EEG y procedimiento experimental

El experimento fue realizado en un laboratorio adaptado para el registro de EEG. Los participantes, después de firmar los consentimientos informados, se sentaban en una silla cómoda y se les colocaba un gorro con 32 electrodos activos usando el sistema Brain Amp. Las posiciones de los electrodos se disponían según el sistema 10-20 y estaban referenciados a Cz: FP1, FP2, F7, F3, Fz, F4, F8, FT9, FC5, FC1, FC2, FC6, FC10, T7, C3, C4, T8, TP9, CP5, CP1, CP2, CP6, TP10, P7, P3, Pz, P4, P8, O1, Oz, y O2. Las impedancias de los electrodos se mantuvieron por debajo de 5 KOhm. La tasa de muestreo de los registros fue de 500 Hz. Con posterioridad, por medio de una computadora portátil se les presentaba una serie de videoclips audiovisuales pertenecientes a películas o ruido blanco (similar a tv fuera de señal). A través del visionado de los vídeos se paraba el vídeo cada  $60 \pm 5$  segundos y se les preguntaba a los participantes sobre en qué habían centrado su atención. Para este experimento clasificamos las posibles respuestas en *Focus* (atendiendo al vídeo) o bien *MW* (distraído pensando en otras cosas). Estas respuestas junto al tipo de vídeo fueron tenidas en cuenta para etiquetar los segmentos de EEG a analizar con posterioridad.

### 2.3. Procesamiento del EEG y cálculo de medidas

Para cada participante seleccionamos 25 segmentos de EEG, con una duración de 50 segundos. Limpiamos los artefactos de cada segmento usando funciones implementadas en EEGLAB y MATLAB. La señal fue filtrada usando un filtro de paso de banda de 0,5-60 Hz. Los artefactos restantes fueron extraídos con ICA *infomax* con ayuda de la inspección visual de los componentes obtenidos.

En el cálculo de la sincronía global, la metaestabilidad y la dimensión fractal usamos un procedimiento de ventana deslizante de 4 segundos con un solapamiento del 90%. Aunque no observamos una tendencia en las series de EEG, el procedimiento de ventanas móviles nos ayuda a controlar cualquier desviación de la estacionariedad. Cada medida era calculada en cada ventana y finalmente obteníamos la mediana como medida representativa del segmento completo.

La sincronía global la obtuvimos mediante el uso del parámetro de orden de Kuramoto (R) [13]. En lugar de

usarlo con la solución en forma de fases de un sistema de ecuaciones diferenciales acopladas, obtuvimos las fases de la señal EEG en 5 pasos de banda: 0.5-1.5 (delta); 4.5-5.5 (theta); 8-13 (alpha); 15-25 (beta); 35-45 (gamma). En concreto, para obtener las fases usamos la transformada de Hilbert con los diferentes pasos de banda. Las señales EEG con diferentes pasos de banda fue obtenida mediante la función *eegfilt* de EEGLAB, aplicando en primer lugar un paso alto y después un paso bajo. Para cada una de estas bandas calculamos R teniendo en cuenta todos los electrodos, y obtuvimos un valor de sincronía y metaestabilidad. El valor de la sincronía nos indica el grado en que las fases extraídas de todos los electrodos están en sincronía. Un valor unitario indicaría una sincronía perfecta entre todos los electrodos mientras que un valor cercano a cero indicaría un gran desfase entre las señales. La metaestabilidad es la desviación típica de la sincronía, y su valor nos indica en qué medida la sincronía fluctúa en el tiempo.

La dimensión fractal fue calculada usando el algoritmo de Higuchi (HFD) [14] para cada electrodo. Esta medida es un indicador de la rugosidad o densidad de la señal representada en el tiempo. Una señal muy plana dará valores bajos mientras que una señal con muchos componentes oscilatorios con distintas características dará valores más altos. El algoritmo de Higuchi estima la dimensión fractal de la señal explorando un número máximo de escalas ( $k_{max}$ ). En este estudio escogimos  $k_{max} = 16$ , el cual, fue el nivel en el que la medida de Higuchi comenzó a mostrarse estable en todos los electrodos.

Tanto las medidas de sincronía (R) como de fractalidad (HFD) fueron escogidas siguiendo la línea de investigación de estudios previos que muestra tanto su utilidad como las posibles relaciones que se pueden encontrar entre ambas [12].

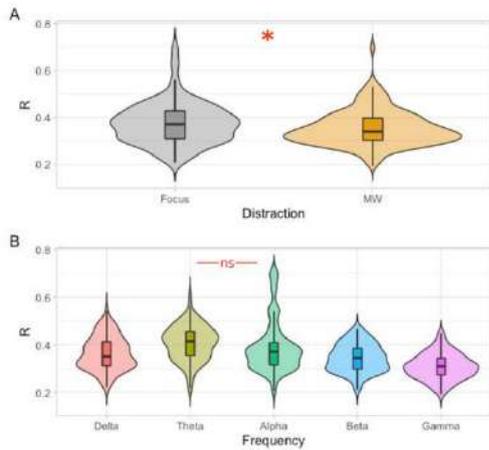
## 3. Análisis estadísticos y resultados

Para cada una de las variables dependientes, sincronía global -R- y metaestabilidad, ajustamos un modelo mixto con *Type* (Noise vs Video), *Distraction* (Focus vs MW) y *Frequency* (Delta, Theta, Alpha, Beta y Gamma) como factores fijos y distinto intercepto para cada participante como efecto aleatorio. Para la dimensión fractal -HFD-, incorporamos como factor fijo la *Regions* (Frontal, Central, Temporo-parietal derecha e izquierda y Posterior). Los análisis se realizaron en R. Para la estimación de los parámetros utilizamos el procedimiento de máxima verosimilitud restringida. Para la estimación de los grados de libertad de las *F* en los análisis de varianza, empleamos el método de Satterthwaite. Para controlar el incremento de la probabilidad de error tipo I debido al número de comparaciones utilizamos el procedimiento de false discovery rate (FDR).

### 3.1. Sincronía (R)

Para la variable Sincronía (R) obtuvimos un efecto significativo para *Distraction* ( $F(1,502.6)=15.33$ ;  $p<0.01$ ) y para *Frequency* ( $F(4,478.52)=41.52$ ;  $p<0.01$ ). El resto de

los efectos principales e interacciones no resultaron significativas, y por ello las interacciones entre distintas bandas de frecuencia y los estados de distracción no se consideraron. En la Figura 1 se presentan las medianas obtenidas en cada una de las condiciones experimentales.

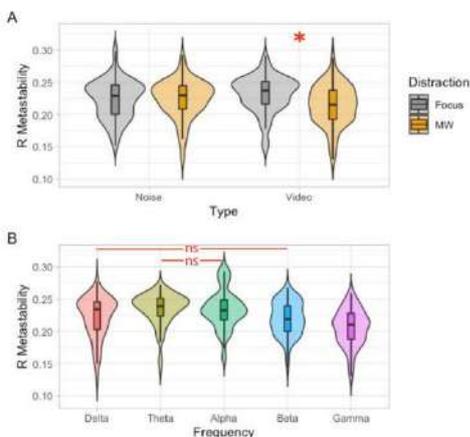


**Figura 1.** Gráfico de violín y box-plot para cada uno de los efectos significativos para Sincronía.

Como se puede ver en la Figura 1 (panel A), la sincronía resultó significativamente superior en Focus que en MW. Además (ver panel B), se encontraron diferencias significativas en la sincronía en todas las frecuencias, excepto Theta y Alpha, que resultaron iguales. Las frecuencias con mayor sincronía fueron Theta y Alpha, seguidas de Delta, Beta y Gamma, en ese orden.

### 3.2. Metaestabilidad

Para esta variable, obtuvimos un efecto significativo de la interacción *Distraction x Type* ( $F(1,486.75)=11.08$ ;  $p<0.01$ ), así como un efecto principal significativo para *Frequency* ( $F(4,479.57)=38.90$ ;  $p<0.01$ ) y para *Distraction* ( $F(1,494.90)=23.34$ ;  $p<0.01$ ). En la Figura 2 se presenta una representación de las medias para cada condición experimental de la interacción y, del efecto principal de los factores no implicados en interacciones.

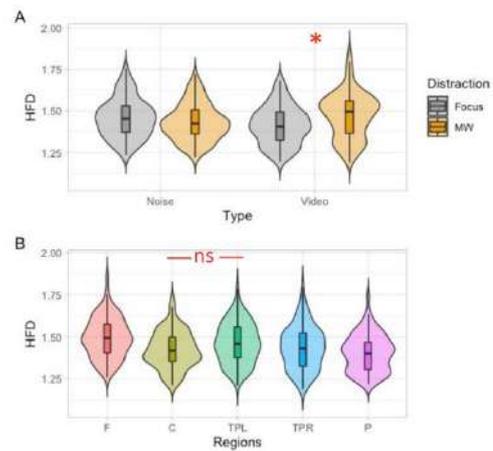


**Figura 2.** Gráfico de violín y box-plot para cada uno de los efectos significativos para Metaestabilidad.

Como se puede observar en la Figura 2, panel A, no se observaron diferencias significativas en la metaestabilidad en la condición Noise. Sin embargo, cuando los participantes estaban viendo Video, la metaestabilidad resultó significativamente superior en Focus respecto a MW. Respecto al efecto principal de *Frequency* (panel B), la metaestabilidad se redujo significativamente en las frecuencias más rápidas (Beta y Gamma) respecto a las más lentas (Delta, Theta y Alpha).

### 3.3. Dimensión fractal (HFD)

Para esta variable, encontramos una interacción significativa entre *Distraction x Type* ( $F(1,760.21)=24.13$ ;  $p<0.01$ ), así como efectos principales significativos para *Distraction* ( $F(1,502.60)=15.33$ ;  $p<0.01$ ), y *Frequency* ( $F(4,478.52)=41.52$ ;  $p<0.01$ ). En la Figura 3 se presenta una representación de las medianas para cada condición experimental de la interacción y del efecto principal del factor no implicado en interacciones.



**Figura 3.** Gráfico de violín y box-plot para cada uno de los efectos significativos para HFD. : frontal, C: central, TPL: temporoparietal left, TPR: temporoparietal right, P: posterior.

Como se puede observar en el panel A de la Figura 3, no se encontraron diferencias en HFD en la condición de ruido (Noise), sin embargo, en la condición de Video, la complejidad fue mayor en MW que en Focus. Por otro lado (ver panel B), la complejidad resultó más alta en la región frontal (F) y más baja en la región posterior (P).

## 4. Discusión

En nuestro estudio hemos hallado que los estados atencionales centrados en videoclips de películas generan más sincronía global, metaestabilidad y menos dimensión fractal que los estados de MW. Este resultado replica y extiende estudios anteriores realizados en nuestro laboratorio [15]. Es posible que los estados conscientes centrados en estímulos externos con sentido requieran una alta integración y especificidad, y que, por tanto, sea necesaria una mayor sincronía global con capacidad de cambio para ir adaptando dichos estados a estímulos que no cesan de transformarse sobre la pantalla. El incremento de HFD en MW encontrado en este y otros estudios previos

refuerza la idea de que un estado de distracción se corresponde con un patrón de activación cortical heterogéneo poco sincronizado. De hecho, en este estudio comprobamos que las medidas de HFD presentaban valores de correlación negativos con el resto de medidas ( $p < 0.05$  para R en Delta, Beta y Theta, y para todos los valores de Metaestabilidad).

Los resultados asociados a estados conscientes relacionados con la presentación de ruido blanco (tanto Focus como MW) presentaron una sincronía y metaestabilidad similar a los estados de MW durante la presentación de videoclips de películas en un estado Focus. Este resultado refuerza nuestra explicación anterior y nos lleva a pensar que es la especificidad en un estado consciente, su distintividad psicológica, lo que marca los valores globales que hemos utilizado en nuestro estudio.

Nuestros resultados van en la línea de las teorías que consideran la consciencia como un fenómeno que requiere información coordinada entre áreas distantes del cerebro y que además cambia de un modo fluido [16]. Aquellos estados conscientes más definidos requerirán de mayor coordinación y cambio, y los asociados a estados más difusos se mantendrán a un nivel de coordinación menor y por tanto una necesidad de cambio también menor.

A pesar de que los resultados observados en este estudio son prometedores, para incrementar su validez, sería necesario replicarlos utilizando nuevas muestras de participantes y distintas medidas de sincronía y complejidad. Además, consideramos que otra ampliación empírica relevante podría ser utilizar o diseñar medidas para detectar la integración y especificidad funcional en los patrones de activación del EEG. Debido a su relevancia teórica, nos daría información muy valiosa sobre los correlatos fisiológicos de distintas experiencias conscientes

## Agradecimientos

Este trabajo ha sido financiado por el Ministerio de Ciencia e Innovación de España (Proyecto PID2019-105145RB-I00/AEI/1013039/501100011033).

## Referencias

- [1] Turkheimer FE., Rosas FE, Dipasquale O, Martins D, Fagerholm ED, Expert P, ..., Leech, R. A complex systems perspective on neuroimaging studies of behavior and its disorders. *The Neuroscientist*, Vol 28, sup 4, 2021, pp 382-99 (ISSN:10738584).
- [2] Stam CV, Van Straaten, ECW. The organization of physiological brain networks. *Clinical neurophysiology*, Vol 123, sup 6, 2012, pp 1067-87 (ISSN: 1872-8952).
- [3] Deco G, Kringelbach ML. Great expectations: using whole-brain computational connectomics for understanding neuropsychiatric disorders. *Neuron*, Vol 84 sup 5, 2014, pp 892-905 (ISSN: 1097-4199).
- [4] Deco G, Tononi G, Boly M, Kringelbach M L. Rethinking segregation and integration: contributions of whole-brain modelling. *Nature Reviews Neuroscience*, Vol 16, sup 7, 2015, pp 430-439 (ISSN: 1471-0048).
- [5] Storm JF, Boly M, Casali AG, Massimini M, Olcese U, Pennartz CM, Wilke M. Consciousness regained: disentangling mechanisms, brain systems, and behavioral responses. *Journal of Neuroscience*, Vol 37, sup 45, 2017, pp10882-93 (ISSN: 0270-6474).
- [6] Deco G, Kringelbach ML. Metastability and coherence: extending the communication through coherence hypothesis using a whole-brain computational perspective. *Trends in neurosciences*, Vol 39, sup 3, 2016, pp 125-35 (ISSN: 1878-108X).
- [7] Bakhshayesh H, Fitzgibbon SP, Janani AS, Grummett TS, Pope KJ. (2019). Detecting synchrony in EEG: A comparative study of functional connectivity measures. *Computers in biology and medicine*, 105, 1-15 (ISSN: 1879-0534).
- [8] Steinmann S, Leicht G, Ertl M, Andreou C, Polomac N, Westerhausen R, ... Mulert C. Conscious auditory perception related to long-range synchrony of gamma oscillations. *Neuroimage*, Vol 100, 2014, pp 435-43 (ISSN: 1095-9572).
- [9] Lachaux JP, Rodriguez E, Martinerie J, Varela FJ. Measuring phase synchrony in brain signals. *Human brain mapping*, Vol 8, sup 4, 1999, pp 194-208 (ISSN: 1065-9471).
- [10] Ibáñez-Molina AJ, Iglesias-Parro S, Escudero J. Differential effects of simulated cortical network lesions on synchrony and EEG complexity. *International Journal of Neural Systems*, Vol 29, sup 4, 2019, pp 1850024 (ISSN: 1793-6462).
- [11] Smallwood J, Schooler JW. The science of mind wandering: empirically navigating the stream of consciousness. *Annual review of psychology*, 2015, pp 487-518 (ISSN: 0066-4308).
- [12] Ibáñez-Molina AJ, Iglesias-Parro S. Neurocomputational model of EEG complexity during mind wandering. *Frontiers in computational neuroscience*, Vol 10, sup 20, 2016, (ISSN: 1662-5188).
- [13] Rodrigues FA, Peron TKD, Ji P, Kurths J. The Kuramoto model in complex networks. *Physics Reports*, Vol 610, 2016, pp 1-98 (ISSN: 1873-6270).
- [14] Higuchi T. Relationship between the fractal dimension and the power law index for a time series: a numerical investigation. *Physica D: Nonlinear Phenomena*, Vol 46, sup 2, 1990, pp 254-64 (ISSN: 0167-2789).
- [15] Ibáñez-Molina AJ, Iglesias-Parro S. Fractal characterization of internally and externally generated conscious experiences. *Brain and cognition*, Vol 87, 2014, pp 69-75 (ISSN: 1090-2147).
- [16] Northoff G, Lamme V. Neural signs and mechanisms of consciousness: is there a potential convergence of theories of consciousness in sight?. *Neuroscience & Biobehavioral Reviews*, Vol 118, 2020, pp 568-587 (ISSN: 1873-7528).

# Cognitive stimulation leads to a long-term brain activation with a spectral shift to higher frequencies in Rett syndrome patients

Ana Tost<sup>1,\*</sup>, Alejandro Bachiller<sup>1,2,3</sup>, Ángeles García-Cazorla<sup>3,4</sup>, Inés Medina-Rivera<sup>3</sup>, Sergio Romero<sup>1,2,3</sup> and Miguel A. Mañanas<sup>1,2,3</sup>

<sup>1</sup> Biomedical Engineering Research Center (CREB), Department of Automatic Control (ESAI), Universitat Politècnica de Catalunya (UPC), 08028, Barcelona, Spain; alejandro.bachiller@upc.edu (A.B.); sergio.romero-lafuente@upc.edu (S.R.); miguel.angel.mananas@upc.edu (M.A.M)

<sup>2</sup> CIBER de Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), 28029, Madrid, Spain

<sup>3</sup> Institut de Recerca Sant Joan de Déu, 08950 Bsrcelona, Spain; inesfernanda.medina@sjd.es (I.M.-R); agarcia@sjdhospitalbarcelona.org (Á.G.-C.)

<sup>4</sup> Neurometabolic Unit and Synaptic Metabolism Lab, Neurology Department, Hospital Sant Joan de Déu, metabERN and CIBERER-ISCI

\*Correspondence: ana.tost@upc.edu

## Abstract

*Rett syndrome (RTT) is the second leading cause of mental disability in women and causes a serious neurodevelopmental disorder that leads to a global disability of the individual. Cognitive training has been shown to improve neuropsychological and neurophysiological parameters in similar pathologies. However, knowledge about neurophysiological training in RTT is scarce. The objective of the study was to evaluate the result and duration of the effect of cognitive stimulation over four consecutive days of training, in order to determine the viability of a long-term effect. Nineteen patients performed a cognitive task using an eye tracking device during the first three days of training. Prior to each training, a resting-state EEG with eyes open was recorded to assess changes in spectral variables in the brain. An EEG recording at rest with eyes open was also acquired on the fourth day, in order to evaluate the differences with respect to the first day. Results showed a spectral shift to higher frequencies in a long-term period of four days. Significant positive alterations of brain rhythms were observed: decrease in low-frequency (delta) oscillations and an increase in high-frequency (beta) rhythms, resulting in more complex signals closer to non-pathological brain behavior. In conclusion, this study showed the potential long-term benefits of cognitive stimulation in RTT patients.*

## 1. Introduction

Rett syndrome (RTT), considered a rare disease, causes a severe neurodevelopmental disorder that leads to a global disability of the individual. It affects almost all aspects of a person's life: the ability to speak, walk, eat or breath, among others. The most common form of the disease is known as classic RTT. It is caused by a MECP2 gene mutation and mainly affects females, specifically 1 in 10,000/15,000 newborns. The MECP2 gene is involved in the maturation, neural growth and dendritic ramification. It is essential for the proper function of nerve cells and neural synapsis, being fundamental for brain development and for the control of the functions of other genes [1]. The abnormal production of this protein is clinically manifested after apparently normal psychomotor development that lasts until the first 6 to 18 months of age. From this period, a regression process begins where all the skills acquired up to that moment are lost, highlighting hand skills and expressive language [2].

Speech difficulties and loss of purposeful hand movements make it difficult to explore cognitive processes and social aspects of RTT patients. Therefore, assessing their cognitive skills is extremely challenging [3]. One way to provide effective communication and evaluate cognitive and social aspects in RTT is by using eye-tracking technology. The eye-gaze response to visual tasks has allowed the possibility of evaluating the cognitive function of patients in a measurable way. Previous studies have demonstrated its utility to verify the level of understanding during cognitive tasks and to evaluate non-verbal cognitive social abilities and attention aspects, showing a preference for socially weighted stimuli and difficulty in endogenous attention [3].

Electroencephalography (EEG) can be used as a low-cost neuroimaging tool to evaluate heterogeneity of cognitive spectrum and cognitive treatment responses. It provides an excellent temporal resolution and during last years, it has been widely used in ADHD and autism disorders [4]. For what concerns this study, the aim of cognitive stimulation is to enhance the neuropsychological abilities in RTT patients. Therefore, it is necessary to find objective methods of evaluating cognitive tasks, as well as their effectiveness.

Previous studies in RTT syndrome have shown topographical changes of different EEG markers before and after the cognitive task showing significant reductions and improvements of delta and beta powers, respectively [5], [6]. It was concluded that the effect of cognitive stimulation remained after the end of the task, linking these changes to a general increase in cognitive competences [5]. This is due to the power distribution, since the predominance of low-frequency oscillations is associated with pathological states, whereas higher frequency oscillations are associated with cognitive functions such as attention, arousal, object recognition and top-down modulation of sensory processes. Similar results were obtained for a long-term training, indicating a positive effect on behavioral parameters and an improvement of neurophysiological measures [6].

The objective of this study was to assess whether the activation was maintained in the long term by means of topographical analysis, which had not been done

previously. The spectral changes of the EEG activity were studied in different test sessions (four consecutive days) to evaluate the effect of cognitive stimulation on the EEG topographical patterns, as well as its differences between the first and last day of cognitive training.

## 2. Methods

This is a cohort study, a prospective observational study without a control group. The study was conducted by the Department of Neurology of Sant Joan de Déu Children's Hospital in Barcelona. It was approved by the local ethics committee following the Declaration of Helsinki and written and informed consent was obtained from all parents or legal caregivers.

### 2.1. Participants

Nineteen young girls clinically diagnosed with RTT were included in the study. The stimulation protocol consisted of three different activities: learning, playing and Peabody Picture Vocabulary Test. The acceptance condition for this work was to have a minimum of four consecutive training days performing the same task. All participants underwent a full medical and neurological evaluation in order to determine, according to the signs and symptoms, the stage of the disease (I to IV). The molecular analysis confirmed the presence of a pathogenic mutation in MECP2 in all patients.

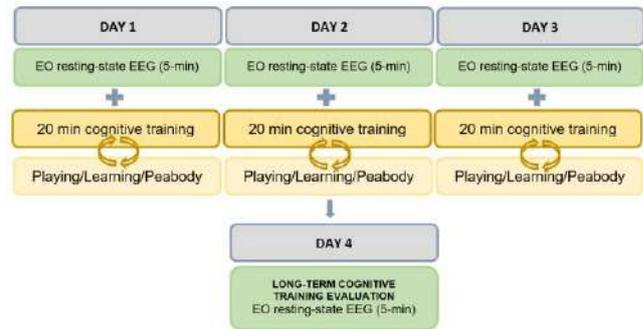
### 2.2. Experimental protocol

The experiment was conducted in a quiet room with reduced ambient light to diminish distractions. Participants were seated approximately 45 cm from a 29-in computer monitor. A portable eye-tracker (Tobii Dynavox PCEye Explore) was used as a computer mouse, allowing the navigation and control of the computer by tracking the gaze. The cognitive processes targeted by the tasks were related to attention, memory and language comprehension, which are severely impaired in RTT patients. The experimental session was divided into two stages, including in each session, an initial eyes-open (EO) resting-stage period (passive) followed by training activities (active). Three different activities were assessed: playing, learning and Peabody Picture Vocabulary Test. Figure 1 shows an outline of an example of the experimental procedure.

First, a resting-state EEG of approximately 5 min ( $5.8 \pm 1.7$  min) with eyes open was recorded. The active intervention, with a mean time of 20-min, consisted on the presentation of three different activities using the software Look to Learn from Smartbox Assistive Technology to encourage the user to engage the whole screen. In this phase of the study, the work focused on analyzing only playful activities that teach "cause and effect" (duration:  $8.6 \pm 1.4$  min). Patients were guided by a neuropsychologist who regularly evaluated whether the patient was following the task. The cognitive training was performed during three consecutive days and a resting-state EEG for evaluating its long-term effect, was recorded the fourth day.

Electroencephalography signals were continuously recorded using a 20-channel EEG system (Starstim 20

wireless device, Neuroelectronics, Barcelona, Spain). Dry electrodes were placed in accordance with the International 10-10 system (Fp1, Fp2, F7, F3, Fz, F4, F8, T7, C3, Cz, C4, T8, CP5, CP6, P7, P3, Pz, P4, P8, Oz). EEG signals were acquired during 25 min referenced to the right ear lobe with a sample frequency of 500 Hz.



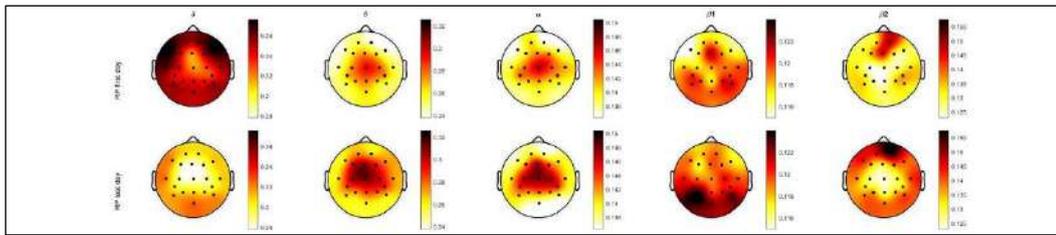
**Figure 1:** "Schematic diagram of the time course of the experimental procedure. At the beginning of each session day, a 5-min resting-state EEG recording with eyes open was acquired."

### 2.3. Electroencephalography analysis

As RTT is characterized by a large series of abnormal movements, it was necessary to design an artifact rejection procedure. It was designed a robust energy-based function

to detect outliers with entropy as its main component. The result was a method that was more robust to outliers but more stable to signal-to-noise ratio. The k-value (threshold) that worked best for the data was the 99.5% of the signal baseline calculated from the entropy value distribution. Together, sudden and strong movements (jerks) can cause impedance faults. A FieldTrip triangulation algorithm was used to define neighboring nodes, and channels were then repaired by replacing affected sections with the distance-weighted average of their neighbors. The last step of EEG preprocessing, Independent Component Analysis (ICA), was performed by Matlab EEGLab toolbox. Runica algorithm was employed for decomposition and a visual inspection of the components was realized by two experts to decide which components were considered ocular and muscle artifacts and, therefore, needed to be removed. Finally, the data was reconstructed without the artifact components, obtaining an artifact-free signal ready for the analysis.

A novel method to evaluate the effect of cognitive stimulation is through topographic maps. It allows us to observe the distribution of changes in spectral variables in the brain. The spectral analysis consisted of evaluating different parameters to observe the evolution of the cognitive response to the training protocol. The short-time Fourier transform was used to estimate power spectral density (PSD) of the EEG data. From the PSD data, the absolute and relative powers (AP, RP) were evaluated in the conventional EEG frequency bands: delta ( $\delta$ , 1-4 Hz), theta ( $\theta$ , 4-7 Hz), alpha ( $\alpha$ , 7-13 Hz) and beta ( $\beta$ , 13-30 Hz). Beta was divided into beta1 (13-20 Hz) and beta2 (20-30 Hz) to analyze if there were greater changes in a certain frequency range within the band.

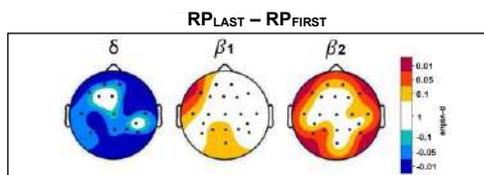


**Figure 2:** Topographic analysis of RP (from left to right): Mean RP of all patients ( $N=19$ ) for the five frequency bands, for the first day (top-row) and last day (bottom-row).

One way to summarize those changes is by using the median frequency (MF) and the spectral entropy (SE). These measurements have been applied in different brain disorders such as schizophrenia, showing a deficit in MF and SE modulation as a response to a cognitive task [7]. MF was computed as the frequency at half of the accumulative PSD. An increase in the MF would be associated with an increase in the signal power, expected to be caused by a simultaneous decrease in slow states (delta) and increase in fast states (alpha and beta). In order to measure signal complexity, SE was used to evaluate irregularity of signals due to the cognitive tasks, demonstrating its usefulness in EEG studies for several pathologies, such as epilepsy, schizophrenia and autism [7]. Decreases in entropy indicate decreases in signal complexity, which are consequently associated with pathological states of the brain.

### 3. Results

In order to analyze the effect of the training, the resting-state EEG data of the first day was compared against the resting-state EEG data of the last day. To ensure enough signal information representative of each stage condition, it was established a minimum of 24 free-artefact epochs (2 min). The average number of 5-s-free artefact epochs of the nineteen patients was  $43.01 \pm 16.98$  for the resting-state EEG of the first day of training and  $48.95 \pm 13.65$  for the resting-state EEG of the last day.



**Figure 3:** P-value representation showing the statistical differences between the first and the last day of EO resting-state EEG recording for delta and beta (Negative p-value: decrease. Positive p-value: increase).

Figure 2 shows the mean RP of the nineteen patients for the main frequency bands of interest. In order to evaluate the changes between average power topographies, paired-sample Wilcoxon ranked tests were performed. Since previous research has highlighted the relevance of modulation in delta and beta as an effect of cognitive training, the following results and discussion will focus solely on these two bands [5], [6]. Figure 3 shows the significant differences found in delta and beta bands. In delta, fifteen channels showed significant differences ( $P < 0.0486$ ,  $Z = -2.543 \pm 0.307$ ) indicating a general power decrease. In beta1, frontal F7 channel showed significant

differences ( $P = 0.0269$ ,  $Z = 2.213$ ), indicating an increase in power. Likewise, in beta2, significant differences were observed in a more generalized area. Five channels had p-values lower than 0.05 with a positive z-statistics for all channels ( $P < 0.0442$ ,  $Z = 2.205 \pm 0.2$ ). Therefore, cognitive training seems to induce a spectral shift to higher frequencies in a long-term period of four days.

These changes can be deeply evaluated in Figure 4. The violin plot, a hybrid of a box plot and kernel density plot, was used to visualize the distribution of numerical data. It was obtained only for the channels that showed statistically significant differences in Figure 3. In delta band, the median power decreased from 35.43% to 31.13%, whereas in beta1 and beta2, increased from 11.94% to 12.53% and from 12.91% to 14.34%, respectively. Likewise, it is observed that there is no overlap between the values corresponding to the first and last day of cognitive training. Which indicates that the effect of reduction of delta and increase of beta is fulfilled for all the participants.

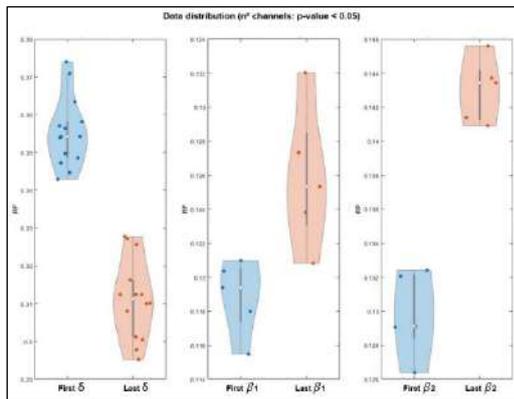
Finally, Figure 5 shows the topographic maps of the MF and SE, both for the mean difference between the first and last day and for its statistical differences. In MF, an average increase of almost 2 Hz was observed in areas with significant differences of  $P < 0.0442$  and  $Z = 2.079 \pm 0.08$ . Likewise, significant differences were also obtained in entropy with  $P < 0.0442$  and  $Z = 2.132 \pm 0.17$ . An increase in MF indicated an energy shift to higher frequencies, caused by a decrease in the slow states (delta) and an increase in the fast states (beta). On the other hand, an increase in SE implied that the spectral content spread out, resulting in an increase in complexity, thus moving away from a pathological state of the brain characterized by low complexity.

### 4. Discussion and conclusion

Rett syndrome, despite being considered a rare disease, is the second most common cause of mental disability in women. However, the literature is scarce when evaluating the disease at the neurocognitive level. In a previous phase of the study, with a single cognitive training session, we already observed statistically significant reductions and improvements of delta and beta powers, respectively, concluding that the effect of cognitive stimulation remained after the end of the task [5]. This opened the possibility of linking those changes to an overall increase in cognitive skills.

Pathological states are characterized by the predominance of low-frequency oscillations, spanning larger neuronal populations and general brain functions. Instead, higher-frequency oscillations originate from a smaller neuronal

population and have been associated with cognitive functions such as attention, arousal, object recognition, and top-down modulation of sensory processes.



**Figure 4:** violin plot of the mean RP in delta and beta for the first and last session. The mean of all patients ( $n=19$ ) has been performed across the electrodes with significant differences for each band. In delta: Fp1, Fp2, F7, F4, F8, C3, C4, T8, CP5, P7, P3, Pz, P4, P8, Oz. In beta1 and beta2: F7, F8, P7, P3, P8. For beta1, in order to obtain a general area representation, it has been performed with F7 (which had significant differences) and the rest of significant electrodes from beta2.

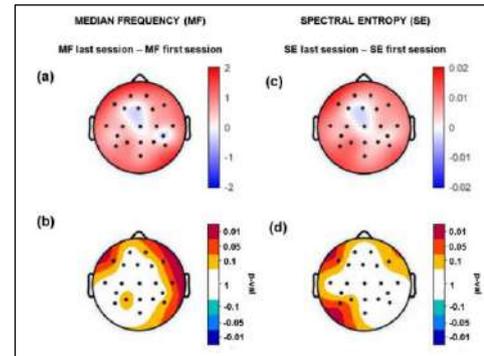
For this reason, the improvement in the beta/delta ratio after cognitive stimulation proved that, carrying out activities using eye-tracking, could lead to cognitive activation that involve reaching more active brain states that are farer from the pathological pattern.

In this phase of the study, the objective was to see if these changes were maintained over time and, therefore, had a prolonged effect on the patients. A positive effect of long-term cognitive training on brain and behavioral parameters in RTT patients was observed in a previous study, showing increased beta activity and decreased theta activity [6]. However, there are no studies that have evaluated topographic changes before and after long-term cognitive stimulation treatment in RTT, which allow us to observe the distribution of changes in spectral variables on the brain.

The evaluation of RP, MF and SE concluded a decrease of delta oscillations and increase of beta rhythms after four days of cognitive training. The possibility that patients reach a more active brain state for a prolonged time is demonstrated by the increase in SE. Since, more complex signals were also associated with less pathological states. Moreover, these findings may lead to the hypothesis that cognitive stimulation triggers a generalized functional interconnectivity in the brain of RTT patients, that should be tested in future studies. On the other hand, the increase in MF confirmed the increase in signal power, leading to states of intense mental activity and moving away from pathological states.

In conclusion, this study showed the potential long-term benefits of cognitive stimulation in RTT patients. After three consecutive training days, the differences between the EO resting states EEG of the first and fourth day, showed significant positive alterations of brain rhythms: EEG analysis showed a decrease in low-frequency (delta) oscillations and an increase in high-frequency (beta)

rhythms, resulting in more complex signals closer to normal brain behavior. Moreover, the current evaluation of the effectiveness of RTT therapies is made by clinical states. However, this study showed that EEG could add a new biomarker to analyze patient performance.



**Figure 5:** (a) Difference between the mean MF power of the first and last day resting-state EEG. (b) Statistical differences between the MF of the first and last day resting-state EEG. (c) Difference between the mean SE of the first and last day resting-stage EEG. (d) Statistical differences between the SE of the first and last day resting-state EEG.

It offers the possibility of having an objective measurement of the cognitive evolution. The spectral analysis of the signal before, during and after the completion of the therapy can be of great help to evaluate the preferences of the patients and find those therapies with the greatest impact. Finally, it should be noted that this research has been carried out based on previous works limited to the delta and beta bands. Future work will address the analysis of all frequency bands.

## 5. Acknowledgement

This research is supported by the ‘Ministerio de Ciencia e Innovación’ (PID2020-117751RB). AT in receipt of a FI-2021 grant from ‘Agència de Gestió d’Ajuts Universitaris i de Recerca’. A.B. is Serra Hünter Fellows.

## 6. References

- [1] M. Vu *et al.*, “Neurobiology of Disease MeCP2 Is Critical for Maintaining Mature Neuronal Networks and Global Brain Anatomy during Late Stages of Postnatal Brain Development and in the Mature Adult Brain,” 2012, doi: 10.1523/JNEUROSCI.1316-12.2012.
- [2] B. Hagberg, “Clinical manifestations and stages of Rett syndrome,” *Ment Retard Dev Disabil Res Rev*, vol. 8, no. 2, pp. 61–65, 2002, doi: 10.1002/MRDD.10020.
- [3] J. Ahoniska-Assa *et al.*, “Assessing cognitive functioning in females with Rett syndrome by eye-tracking methodology,” *European Journal of Paediatric Neurology*, vol. 22, no. 1, pp. 39–45, 2018, doi: 10.1016/j.ejpn.2017.09.010.
- [4] R. Aldemir *et al.*, “Evaluation of Two Types of Drug Treatment with QEEG in Children with ADHD,” *Transl Neurosci*, vol. 9, p. 106, 2018, doi: 10.1515/TNSCI-2018-0017.
- [5] C. Migliorelli *et al.*, “Cognitive stimulation has potential for brain activation in individuals with Rett syndrome,” *Journal of Intellectual Disability Research*, vol. 66, no. 3, pp. 213–224, Mar. 2022, doi: 10.1111/JIR.12902.
- [6] R. A. Fabio, L. Billeci, G. Crifaci, E. Troise, G. Tortorella, and G. Pioggia, “Cognitive training modifies frequency EEG bands and neuropsychological measures in Rett syndrome,” *Res Dev Disabil*, vol. 53–54, pp. 73–85, 2016, doi: 10.1016/j.ridd.2016.01.009.
- [7] A. Bachiller *et al.*, “Decreased spectral entropy modulation in patients with schizophrenia during a P300 task,” *Eur Arch Psychiatry Clin Neurosci*, 2014, doi: 10.1007/s00406-014-0488-6.

# Aplicación de *Deep Learning* para el procesado automático de componentes ICA de registros de magnetoencefalografía

L. Gutiérrez-de Pablo<sup>1</sup>, S. Pérez-Velasco<sup>1,2</sup>, V. Rodríguez-González<sup>1,2</sup>, V. Gutiérrez-de Pablo<sup>1,2</sup>, C. Gómez<sup>1,2</sup>, J. Poza<sup>1,2,3</sup>

<sup>1</sup> Grupo de Ingeniería Biomédica, Universidad de Valladolid, Valladolid, España, laura.gutierrez.pablo@alumnos.uva.es

<sup>2</sup> Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), España

<sup>3</sup> IMUVA, Instituto de Investigación en Matemáticas, Universidad de Valladolid, Valladolid, España

## Resumen

La magnetoencefalografía (MEG) permite registrar los campos magnéticos inducidos por la actividad neuronal. Sin embargo, los registros MEG están contaminados por señales procedentes de fuentes no neuronales. Para un posterior análisis de los datos es necesario mitigar la presencia de estos artefactos. El análisis de componentes independientes (ICA) es un método estadístico que permite separar las diferentes componentes que forman los datos MEG. Sin embargo, ICA no realiza la clasificación de dichas componentes; esta tarea la realiza normalmente un técnico mediante un análisis visual, por lo que el grado de subjetividad es elevado. Para solucionar este problema, en este trabajo se ha diseñado un sistema de *Deep Learning*, basado en la red EEG-Inception, para diferenciar entre componentes ICA de origen cerebral y componentes contaminadas por artefactos. Para ello se utilizó una base de datos de componentes ICA ya etiquetadas, procedentes de 473 registros de actividad MEG espontánea. Los resultados de la clasificación binaria de componentes muestran una precisión del 87% mientras que en la clasificación multiclase se obtuvo una precisión del 85%. Un posterior análisis visual de una selección aleatoria de las muestras mal clasificadas por la red mostró que algunas de ellas estaban mal etiquetadas en la base de datos. Estos resultados sugieren que el uso de la herramienta propuesta podría también mejorar el etiquetado manual de los técnicos.

## 1. Introducción

La magnetoencefalografía (MEG) y la electroencefalografía (EEG) son técnicas que permiten estudiar la dinámica electrofisiológica del cerebro de forma no invasiva. La primera hace referencia al estudio de los campos biomagnéticos emitidos por las corrientes eléctricas del cerebro, mientras que la segunda hace referencia al estudio de la actividad bioeléctrica cerebral [1], [2]. La MEG se emplea para estudiar diferentes enfermedades neurológicas como la epilepsia, para la planificación de la extracción de tumores cerebrales, para la valoración de pacientes con enfermedades cerebrovasculares, etc. [3]. Para poder analizar las señales de actividad MEG, es necesario un preprocesado previo debido a que normalmente están contaminadas por artefactos de diverso origen. Estos artefactos son interferencias indeseadas que pueden sesgar la interpretación de los resultados obtenidos a partir de las señales MEG y, por ello, es necesaria su eliminación. Los artefactos de MEG se pueden clasificar en dos categorías distintas: artefactos no biológicos, que pueden deberse a

problemas técnicos en la colocación de los sensores o interferencias externas; y artefactos relacionados con ruido biológico, que no depende de la actividad eléctrica cerebral y que englobarían a artefactos oculares o electrooculograma (EOG), artefactos cardíacos o electrocardiograma (ECG), movimientos de la cabeza del paciente, actividad muscular, etc.

Para un posterior procesado de las señales MEG, se requiere identificar los distintos tipos de ruido para posteriormente eliminarlos. Uno de los grandes problemas de la identificación de artefactos es que, en la mayor parte de los casos, se lleva a cabo de manera visual por un técnico especializado. Si el registro es adquirido con pocos canales y, además, es de corta duración, el análisis de la señal y la identificación de sus artefactos son relativamente sencillos. Sin embargo, si se utiliza un sistema de gran densidad en canales para el registro, además de ser este de larga duración, el preprocesado se convierte en una tarea compleja y que consume mucho tiempo. Por ello, son necesarias técnicas que faciliten esta tarea. Un método estadístico muy utilizado para solucionar este problema es el análisis de componentes independientes (ICA). Este método se basa en extracción y separación de fuentes estadísticamente independientes de señales biomédicas como EEG y MEG, pudiendo de esta forma diferenciar entre componentes con actividad cerebral y componentes contaminadas por artefactos.

Otro gran problema a la hora de preprocesar señales biomédicas es el posible sesgo a la hora de identificar artefactos, debido a la subjetividad inherente a su identificación manual. Es por ello que los métodos automáticos de clasificación han crecido en popularidad en este ámbito. Actualmente existen sistemas que son capaces de solucionar este problema, al menos con señales EEG, pero son sistemas basados en *Machine Learning* (ML) [5]. En campos como la visión artificial o el procesamiento natural del lenguaje, el ML ha sido sobrepasado por los sistemas basados en *Deep Learning* (DL), debido a que pueden llegar a conseguir resultados más precisos en tareas de reconocimiento de patrones, como lo es el etiquetado de componentes ICA. A día de hoy sólo encontramos trabajos, en los que se utilicen *higher order statistics* y medidas de entropía o *Deep Learning* para realizar clasificaciones de un simple tipo de artefacto, como EOG o ECG, y no una clasificación multiclase entre componentes MEG con varios artefactos [6]-[8].

El objetivo de este trabajo es implementar un sistema de *Deep Learning* utilizando la red EEG-Inception [9], adaptándola a señales MEG, de forma que consiga diferenciar entre componentes ICA correspondientes a actividad cerebral o asociadas a artefactos.

## 2. Materiales y métodos

### 2.1. Bases de datos

La base de datos que se ha empleado está compuesta por componentes ICA de 473 registros, de cinco minutos cada uno. Estos registros se realizaron en el hospital de Hokuto (Obihiro, Japón) mediante un sistema de 160 canales (MEG Vision PQ1160C, Yokogawa Electric). La frecuencia de muestreo original fue de 1000 Hz. Para cada sujeto se disponía de una matriz de datos que incluía las 160 componentes ICA obtenidas con el algoritmo Extended Infomax ICA. Cada componente ICA estaba etiquetada por un técnico especialista, con más de tres años de experiencia, como actividad cerebral o artefacto: “0” para componentes cerebrales, “1” para componentes cardíacas, “2” para componentes de red eléctrica, “3” para componentes oculares y, por último, “4” para otro tipo de componentes artefactuadas. La figura 1 ilustra algunas de estas componentes. Todos estos datos fueron analizados con MATLAB®.

### 2.2. EEG- Inception

EEG-Inception es una red neuronal convolucional (CNN) utilizada para el procesamiento de EEG y la detección de potenciales relacionados a eventos [9]. Aunque originalmente esta red fue diseñada para el procesado de señales EEG, en este trabajo se ha extendido su uso para clasificar componentes ICA extraídas de señales MEG. Para ello, se realizaron varios cambios en su estructura, que se resumen en la figura 2:

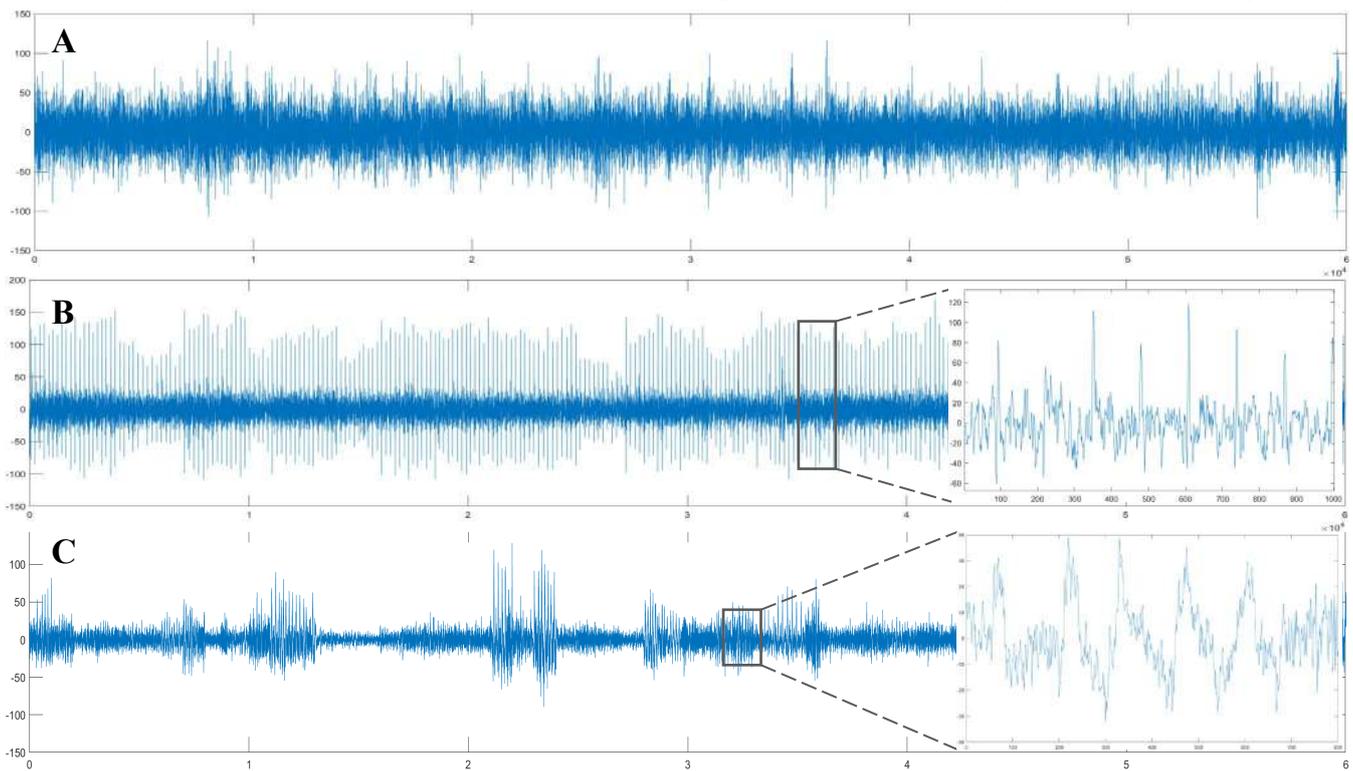
- Bloque I. En este bloque se aplican tres convoluciones 2D de  $100 \times 1$ ,  $50 \times 1$  y  $25 \times 1$ , respectivamente, utilizando las mismas ventanas temporales que en la configuración inicial de EEG-Inception. En este caso, se ha eliminado la convolución *depthwise* debido a que en este caso sólo tenemos un canal de entrada.
- Bloque II. La estructura de este bloque es igual a la del anterior. Su función es extraer características temporales adicionales en un mayor nivel de abstracción, debido a que se aplican los mismos filtros que en el bloque I, pero sobre la señal transformada que pasa al bloque II.
- Bloque III. Está compuesto por dos convoluciones y dos *average pooling*. Este bloque consigue extraer los patrones más complejos usados en la clasificación final.

## 3. Resultados y discusión

### 3.1. Clasificación binaria

La primera aproximación de la red tuvo como objetivo la clasificación binaria de las componentes ICA. Esta clasificación necesitaba que la red fuera capaz de diferenciar entre componentes correspondientes a actividad cerebral y componentes correspondientes a artefactos, sin que tuviera que diferenciar entre los diferentes tipos de estos.

Para obtener los valores de precisión se realizó *k-fold* con 5 *splits*, de manera que la precisión obtenida fuera para todas las muestras presentes en la base de datos. Para evitar el *overfitting* de la red, en cada conjunto de entrenamiento se usó *early-stopping*, al igual que se hacía en la implementación original de EEG-Inception. Este se aplicó cuando no había mejora en la función de pérdidas



**Figura 1.** Ejemplos de componentes ICA de MEG utilizadas como entradas para EEG-Inception: (A) componente de actividad cerebral, (B) componente de actividad cardíaca y (C) componente de actividad ocular.

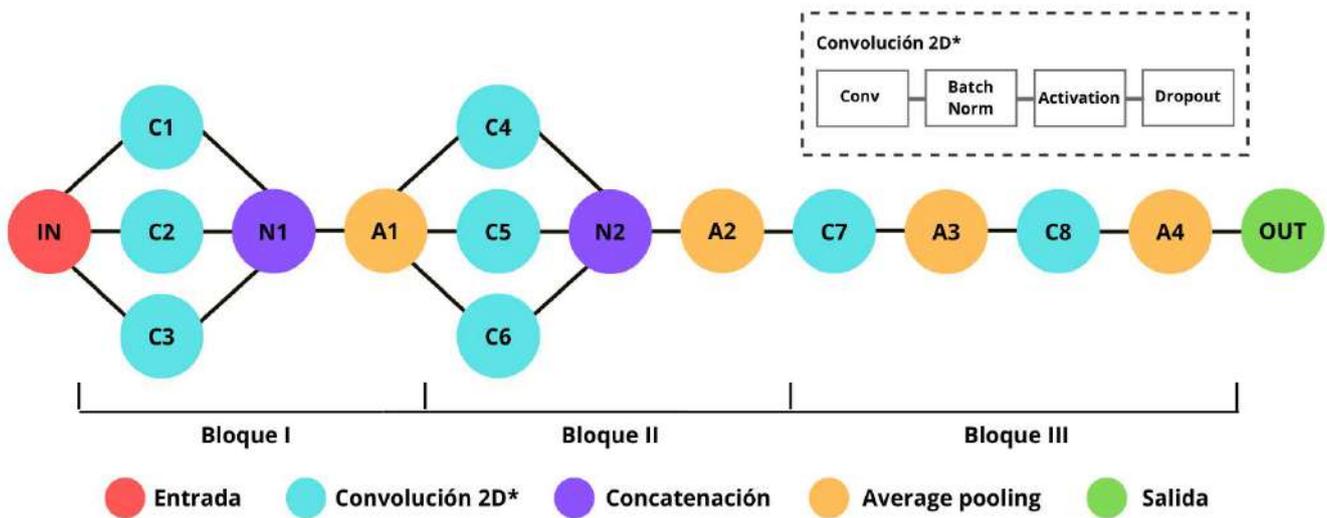


Figura 2. Arquitectura de la red EEG-Inception para la clasificación de componentes ICA de origen cerebral y artefactuadas a partir de datos MEG.

sobre un conjunto aleatorio de validación correspondiente a un 20% del conjunto de entrenamiento. Por otra parte, se hicieron múltiples pruebas para analizar el comportamiento de la red ante distintos valores de los hiperparámetros; sin embargo, los mejores resultados se obtuvieron con la configuración original de la red. Por último, antes de empezar el entrenamiento de la red, se normalizaron los datos como técnica de regularización adicional.

Durante el entrenamiento de la red se balancearon las clases, dándole más peso a las componentes contaminadas por artefactos, debido a que eran una minoría en comparación con las componentes de origen cerebral. Asignar un peso diferente en la función de pérdidas a las distintas clases permitió evitar que la red predijese todas las componentes como actividad cerebral. Una vez finalizado el entrenamiento de la red, se analizó el comportamiento de la misma sobre el conjunto de test. La matriz de confusión se muestra en la figura 3; puede observarse que la red tuvo una sensibilidad del 81% y una especificidad del 87%. La precisión final obtenida fue de un 87%. Para comprobar el rendimiento final de la red, se seleccionaron de forma aleatoria 40 componentes de los falsos negativos; un segundo análisis visual por un técnico reveló que el 72,5% de las componentes fueron clasificadas de forma correcta por la red, y pasadas por alto por el técnico que primeramente las clasificó. Lo mismo pasó con los falsos positivos, de los cuales se eligieron 20 componentes, y se observó que el 70% parecían haber sido predichas de forma correcta por la red. Esto nos indica que la precisión real de la red podría ser mucho mayor que ese 87% que se obtuvo en un primer momento.

### 3.2. Clasificación multiclase

Otra aproximación consistió en entrenar a la red con las cinco clases inicialmente detalladas, es decir, una clase para las componentes con actividad neuronal y las otras cuatro clases para los distintos artefactos. Se utilizaron los mismos parámetros que en la clasificación binaria. El objetivo de esta clasificación fue entrenar a la red con distintos tipos de artefactos y observar cambios en el rendimiento de la red. Los resultados finales se pueden

encontrar en la figura 4. Tras la clasificación multiclase se obtuvo una precisión del 85%. De nuevo, para comprobar el rendimiento final de la red, un técnico analizó aleatoriamente 35 componentes de cada una de las cinco clases clasificadas por la red como falsos positivos. Los resultados fueron los siguientes: un 67% de las componentes cardíacas, un 86% de las componentes con red eléctrica, un 80% de las componentes oculares y un 89% de los artefactos de otro tipo, fueron etiquetados de forma correcta por la red, aunque previamente pasados por alto por el primer técnico que las clasificó.

En estudios previos se propuso realizar una clasificación multiclase de componentes a partir del estudio de *higher-order statistics* y medidas de entropía [6]. Finalmente N. Kamal Al-Qazzaz *et al.* mostraron que la técnica conjunta de ICA con la Transformada Wavelet Discreta conseguía en gran medida conservar la actividad cerebral tras la eliminación de artefactos. Otros autores han estudiado el rechazo de artefactos a partir componentes ICA mediante el uso de redes neuronales, utilizando a su vez las señales de ECG o EOG tomadas al mismo tiempo que las de MEG para acelerar el preprocesado de las señales [7], de esta forma se obtuvo una correcta eliminación de los artefactos, preservando el contenido frecuencial de las señales con actividad cerebral. Garg *et al.* emplearon redes neuronales para clasificar un solo tipo de artefacto, en este caso el ocular sin el uso de EOG [8], obteniendo una precisión del 99%. En este estudio se han conseguido

		Predicción	
		Artefacto (1)	No artefacto (0)
Clasificación	Artefacto (1)	353	81
	No artefacto (0)	1966	12736

Figura 3. Matriz de confusión obtenida sobre el conjunto de test para la clasificación binaria de componentes ICA.

		Predicción				
		(1)	(2)	(3)	(4)	(0)
Clasificación	(1)	157	17	10	5	35
	(2)	1	32	11	5	9
	(3)	0	6	9	2	7
	(4)	25	21	28	24	30
	(0)	308	768	417	579	12630

**Figura 4.** Matriz de confusión obtenida sobre el conjunto de test para la clasificación multiclase de componentes ICA.

unos resultados muy satisfactorios sin el uso de esas dos señales, lo que disminuye el número de datos a utilizar. Además, sabe destacar que se han conseguido diferenciar entre cuatro tipos de artefactos, a diferencia de los estudios comentados anteriormente.

### 3.3. Limitaciones y líneas futuras

A pesar de la ayuda que la herramienta supone para mejorar el preprocesado de señales MEG, hay que tener en cuenta que las componentes ICA han sido etiquetadas por un solo investigador, lo cual puede implicar cierto sesgo. Por ello, sería interesante observar la precisión final que se obtiene con la red, tras un primer entrenamiento de la esta y una posterior inspección visual de las componentes mal etiquetadas. Un paso más allá podría ser un estudio comparativo entre el etiquetado de varios profesionales y el etiquetado de un solo profesional ayudado de esta herramienta. Por otra parte, este estudio podría extrapolarse al preprocesado de señales EEG, sobre todo porque la red neuronal utilizada fue originalmente creada para el uso de este tipo de señales. Este posible nuevo estudio podría ahorrar mucho tiempo de preprocesado a todos los investigadores que trabajan con esta señal. Por último, sería muy interesante aplicar métodos de inteligencia artificial explicable (XAI) para comprender el funcionamiento interno de la red, aumentando así la confianza en este tipo de sistemas [10].

## 4. Conclusiones

En este estudio se ha adaptado la red neuronal EEG-Inception para el análisis de componentes ICA obtenidas a partir de señales MEG, con el objetivo de diferenciar los diferentes tipos de componentes. Las precisiones obtenidas han sido de un 87% para la clasificación binaria y de un 85% para la multiclase. Además, cabe destacar que en el posterior análisis visual de los falsos positivos y negativos se pudo observar que la red había clasificado de forma correcta en torno a un 70% de las componentes que en el primer etiquetado no fueron correctamente clasificadas, lo que muestra que un uso conjunto del etiquetado del profesional junto con el uso de DL podría mejorar la detección de artefactos.

## Agradecimientos

Este estudio ha sido financiado por el ‘Ministerio de Ciencia e Innovación – Agencia Estatal de Investigación’

y el Fondo Europeo de Desarrollo Regional (FEDER) bajo el proyecto PGC2018-098214-A-I00, por la Comisión Europea y el FEDER bajo el proyecto ‘Análisis y correlación entre la epigenética y la actividad cerebral para evaluar el riesgo de migraña crónica y episódica en mujeres’ (‘Programa de Cooperación Interreg V-A España-Portugal POCTEP 2014-2020’), y por el CIBER en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN) a través del Instituto de Salud Carlos III. S. Pérez-Velasco, V. Gutiérrez-de Pablo y V. Rodríguez-González son beneficiarios de una ayuda PIF-UVa de la Universidad de Valladolid.

## Referencias

- [1] M. X. Cohen. Where Does EEG Come From and What Does It Mean?, *Trends in Neurosciences*, vol. 40, pp. 208–218, 2017. doi: 10.1016/j.tins.2017.02.004.
- [2] P. Mandal, A. Banerjee, M. Tripathi, A. Sharma, A comprehensive review of magnetoencephalography (MEG) studies for brain functionality in healthy aging and Alzheimer’s disease (AD), *Frontiers in Computational Neuroscience*, vol. 12, 2018. doi: 10.3389/fncom.2018.00060.
- [3] Fred, A.L.; Kumar, S.N.; Kumar Haridhas, A.; Ghosh, S.; Purushothaman Bhuvana, H.; Sim, W.K.J.; Vimalan, V.; Givo, F.A.S.; Jousmäki, V.; Padmanabhan, P.; et al. A Brief Introduction to Magnetoencephalography (MEG) and Its Clinical Applications. *Brain Sci.* 2022, 12, 788. <https://doi.org/10.3390/brainsci12060788>.
- [4] V. Rodríguez-González, C. Gómez, H. Hoshi, Y. Shigihara, R. Hornero, J. Poza, Exploring the Interactions Between Neurophysiology and Cognitive and Behavioral Changes Induced by a Non-pharmacological Treatment: A Network Approach, *Front Aging Neuroscience*, vol. 13, 2021, doi: 10.3389/fnagi.2021.696174.
- [5] L. Pion-Tonachini, K. Kreutz-Delgado, S. Makeig, ICLabel: An automated electroencephalographic independent component classifier, dataset, and website, *Neuroimage*, vol. 198, pp. 181–197, 2019, doi: 10.1016/j.neuroimage.2019.05.026.
- [6] N. Kamal Al-Qazzaz, S. Hamid Bin Mohd Ali, S. Anom Ahmad, M. Shabiul Islam, and J. Escudero, Automatic artifact removal in EEG of normal and demented individuals using ICA-WT during working memory tasks, *Sensors (Switzerland)*, vol. 17, no. 6, 2017, doi: 10.3390/s17061326.
- [7] L. Breuer, J. Dammers, T. P. L. Roberts, and N. J. Shah, Ocular and cardiac artifact rejection for real-time analysis in MEG, *J Neurosci Methods*, vol. 233, pp. 105–114, 2014, doi: 10.1016/j.jneumeth.2014.06.016.
- [8] P. Garg, E. Davenport, G. Murugesan, B. Wagner, C. Whitlow, J. Maldijan, A. Montillo, Using convolutional neural networks to automatically detect eye-blink artifacts in magnetoencephalography without resorting to electrooculography, in *Lecture Notes in Computer Science*, vol. 10435 LNCS, pp. 374–381, 2017, doi: 10.1007/978-3-319-66179-7\_43.
- [9] E. Santamaria-Vazquez, V. Martinez-Cagigal, F. Vaquerizo-Villar, R. Hornero, EEG-Inception: A Novel Deep Convolutional Neural Network for Assistive ERP-Based Brain-Computer Interfaces, *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 28, pp. 2773–2782, 2020, doi: 10.1109/TNSRE.2020.3048106.
- [10] P. Linardatos, V. Papastefanopoulos, S. Kotsiantis, Explainable AI: A Review of Machine Learning Interpretability Methods, *Entropy*, vol. 23, pp. 18, 2021, doi: 10.3390/e23010018.

# Neurotecnologías II

---

Jueves, 24 de noviembre de 2022

# Evaluación del nivel de atención durante el entrenamiento de una BMI con un exoesqueleto de miembro inferior

L. Ferrero <sup>1</sup>, V. Quiles <sup>1</sup>, M. Ortiz <sup>1,2</sup>, E. Iáñez <sup>1,2</sup>, J.M. Azorín <sup>1,2,3</sup>

<sup>1</sup> Brain-Machine Interface Systems Lab, Universidad Miguel Hernández de Elche, Elche, España {lferrero, vquiles, mortiz, eianez, jm.azorin}@umh.es

<sup>2</sup> Instituto de Investigación en Ingeniería de Elche - I3E, Universidad Miguel Hernández de Elche, Elche, España

<sup>3</sup> ValGRAI: Valencian Graduated School and Research Network of Artificial Intelligence, España

## Resumen

*Las interfaces cerebro-máquina (brain-machine interfaces, BMIs) decodifican patrones cerebrales y los convierten en comandos para controlar dispositivos externos, tales como exoesqueletos robóticos de asistencia. Por ello, se han utilizado recientemente como herramientas de neurorrehabilitación. Uno de los patrones cerebrales más utilizado consiste en la imaginación de un movimiento sin ejecutarlo de forma física, lo que se denomina imaginación motora (IM). La realización de esta tarea mental promueve mecanismos de neuroplasticidad que fomentan la recuperación de los pacientes. Sin embargo, para que la recuperación sea efectiva, los pacientes deben poner toda su atención. En este trabajo se ha estimado el nivel de atención durante IM de los usuarios durante una serie de sesiones de entrenamiento con una BMI. Sin embargo, no se han hallado diferencias significativas que pudiesen estar relacionadas con la fatiga o cansancio que experimentan los usuarios.*

## 1. Introducción

Las interfaces cerebro-máquina (brain-machine interfaces, BMIs) se han utilizado recientemente como herramientas de neurorrehabilitación para personas que han sufrido un accidente cerebrovascular o una lesión en la médula espinal. Estos sistemas decodifican patrones cerebrales y los convierten en comandos para controlar dispositivos externos, tales como exoesqueletos robóticos de asistencia o escenarios de realidad virtual. Existen dos tipos de BMI en base a los patrones cerebrales a detectar: exógenas y endógenas [1]. Las exógenas se basan en la respuesta cerebral a determinados estímulos externos y entre los más utilizados están los estímulos visuales como P300 [2] o SSVEP (steady state visual evoked potentials) [3]. Por otro lado, las interfaces endógenas utilizan patrones cerebrales que se generan de forma espontánea y voluntaria. El patrón más utilizado es la imaginación motora (IM) [4] que se define como el proceso de imaginar un movimiento sin ejecutarlo [5].

Para la rehabilitación de la movilidad se emplean generalmente interfaces basadas en la imaginación motora. Los pacientes deben imaginar el movimiento de su extremidad afectada con el objetivo de promover el reclutamiento de neuronas de áreas cerebrales involucradas en el control del movimiento y así, mejorar los mecanismos de neuroplasticidad. La neuroplasticidad es una propiedad del sistema nervioso que le permite reestructurarse y adaptarse a nuevas situaciones [6].

Las BMIs registran las señales cerebrales del sujeto mientras éste realiza la tarea de imaginación motora y esta información se emplea para controlar un dispositivo que proporcionará retroalimentación. La retroalimentación permite fortalecer el entrenamiento, ya que el sujeto sabe si lo está realizando correctamente. Además, en el caso de que el dispositivo a controlar sea un exoesqueleto robótico, éste permite generar una tarea de movimiento real en la extremidad afectada, lo que también promueve la rehabilitación [7].

Para que la recuperación sea exitosa, es crucial que los pacientes participen cognitivamente en el tratamiento. Por esta razón, el objetivo de esta investigación fue evaluar el nivel de atención de los usuarios mientras realizaban imaginación motora de la marcha [8].

Las interfaces BMI generalmente tienen dos fases, una primera fase de entrenamiento en la que el sistema se ajusta al usuario y una segunda fase de testeo en la que se produce el control del dispositivo externo. En este trabajo, se ha analizado la evolución del nivel de atención de los usuarios durante el entrenamiento de una BMI con un exoesqueleto de miembro inferior a lo largo de las distintas pruebas, con el objetivo de estudiar si la atención decae conforme avanzan las sesiones y los usuarios están más cansados.

## 2. Material y métodos

### 2.1. Participantes

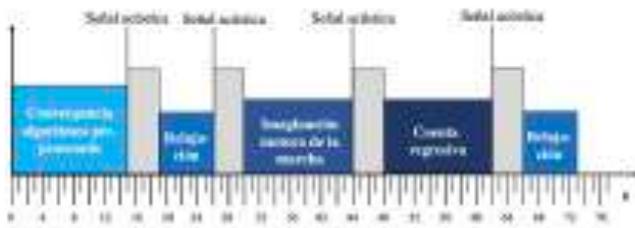
Tres usuarios sanos participaron en este estudio (edad media, 24±2). No tenían limitaciones motoras y no padecían ninguna enfermedad conocida. Además, no tenían experiencia previa con BMIs. Fueron informados sobre la experimentación que se iba a llevar a cabo y firmaron un consentimiento informado de acuerdo con la declaración de Helsinki. Los procedimientos fueron aprobados por la Oficina de Investigación Responsable de la Universidad Miguel Hernández de Elche (España).

### 2.2. Diseño experimental

Cada sujeto participó en 5 sesiones y durante cada sesión realizaron 22 pruebas con un exoesqueleto. Las pruebas consistieron en el entrenamiento de la BMI, por lo tanto, el exoesqueleto estaba controlado por los técnicos del laboratorio y no por el propio usuario. Se evaluaron dos condiciones: cuando los pacientes estaban en estático y cuando estaban en movimiento. En 11 pruebas los usuarios

permanecieron de pie inmóviles y durante las otras 11, anduvieron asistidos por el exoesqueleto.

En cada prueba se les pidió que realizaran 3 tareas mentales: imaginación motora de la marcha, cálculos de cuenta regresiva y un estado de relajación. Durante la IM, se les pidió que imaginaran el movimiento de caminar, pero centrándose en las piernas. En cuanto a la cuenta regresiva, se les pidió que realizaran una serie de restas. Por ejemplo, en el caso de 1000-7, debían calcular mentalmente la serie de 993, 986, 979, 972, etc. Finalmente durante el estado de relajación, se les pidió que estuvieran lo más relajados posible, intentando no centrar su atención en ningún pensamiento. La Figura 1 muestra la secuencia de cada prueba.



**Figura 1.** Secuencia de las tareas mentales a realizar. Durante los primeros 15 s los usuarios no debían realizar nada, se utilizaron estos datos para la convergencia de los algoritmos de pre-procesado. Una señal acústica indicaba el inicio de cada tarea.

### 2.3. Equipo

Las señales cerebrales se registraron con electroencefalografía (EEG). Se utilizó un equipo con 32 electrodos, Starstim R32 (Neuroelectrics, Barcelona, España). La configuración de los electrodos seguía el modelo internacional 10-10. De los 32 electrodos sólo se utilizaron 27: F3, FZ, F4, FC5, FC3, FC1, FCZ, FC2, FC4, FC6, C5, C3, C1, CZ, C2, C4, C6, CP5, CP3, CP1, CPZ, CP2, CP4, CP6, P3, PZ, P4. Los electrodos de referencia y tierra se colocaron en el lóbulo de la oreja derecha. Se utilizó el software de registro NIC (neuroelectrics instrument controller) (Neuroelectrics, Barcelona, España) y las señales fueron registradas a una frecuencia de muestreo de 500 Hz.

Los sujetos utilizaron el exoesqueleto H3 (Technaid, Madrid, España). Este dispositivo robótico proporciona asistencia a la movilidad emulando de forma completa la marcha humana. Un técnico del laboratorio sujetaba el exoesqueleto por la estructura de la espalda y los sujetos utilizaron muletas para garantizar un mayor apoyo. Se enviaron comandos de arranque y parada vía Bluetooth desde un ordenador para el control.

### 2.4. BMI

EEG se describe en términos de actividad rítmica en diferentes rangos de frecuencia. El pre-procesado, procesado y clasificación de la señal se hizo por ventanas de 1.5 s con un solapamiento de 0.5 s. Es decir, se tomó una ventana de 1.5 s que se iba desplazando de medio segundo en medio segundo y en cada desplazamiento, la señal resultante fue analizada y se asoció con un nivel alto o bajo de atención a la marcha. Los datos se muestrearon a

200 Hz para conseguir una mayor velocidad computacional y que todo el análisis de cada ventana se realizase en menos de 0.5 s.

Las señales fueron pre-procesadas para aumentar la ratio señal-ruido. Primero fueron filtradas mediante un filtro paso alto a 0.1 Hz para eliminar la componente continua de la señal. A continuación, se aplicó un filtro Notch a 50 Hz para mitigar los artefactos introducidos por la red eléctrica. Una vez los datos fueron filtrados en frecuencia, la señal de cada electrodo fue estandarizada siguiendo el proceso presentado en [8]. En cada canal se calculó el MVT (maximum visual threshold) como el valor promedio de los 6 valores con mayor amplitud de la señal. Este valor se calculó para cada desplazamiento de la ventana de tiempo, se almacenó en un vector y el MVT actualizado por ventana se calculó como el promedio de este vector. Las señales fueron estandarizadas como:

$$SV(t)_{ch} = \frac{V(t)_{ch}}{\frac{1}{Ch} \sum_{j=1}^{Ch} MVT_j}$$

La señal de cada canal  $V(t)_{ch}$  se normalizó considerando la información de todos los canales. Por último, se utilizó un filtro laplaciano de superficie para reducir el ruido espacial y aislar la actividad local de cada electrodo.

La actividad cerebral en la banda de frecuencia gamma se ha asociado con el nivel de atención a la marcha [3], [4], [9]. Por esta razón, la potencia de la señal se estimó con el método de máxima entropía (MEM) [5], y se calculó la potencia promedio de las frecuencias en la banda gamma. Este valor se calculó para cada electrodo para formar un vector de características.

Finalmente, un clasificador de análisis discriminante lineal (linear discriminant analysis, LDA) predijo el nivel de atención entre baja atención a la marcha y alta atención a la marcha, es decir, 0 para baja atención y 1 para alta atención. Los periodos de relajación y operaciones de cuenta regresiva se consideraron como baja atención a la marcha y los periodos de IM como alta atención a la marcha. Luego, este valor se promedió cada 4 segundos para obtener una línea continua durante toda la prueba. Para evaluar el nivel de atención, se calculó el área de esta línea durante el período IM.

La evaluación se realizó con validación cruzada. Para cada sujeto y sesión, se utilizaron los datos de todas las pruebas menos una para entrenar el modelo y éste se testeó con la restante. Este proceso se repitió de forma iterativa utilizando todas las pruebas una vez como testeó.

## 3. Resultados y discusión

El nivel de atención se midió para cada prueba y en promedio para cada sesión durante la clase IM. La Figura 2 muestra el nivel de atención estimado a lo largo de una prueba (la clase IM corresponde al periodo en rojo). Para calcular el nivel de atención correspondiente al periodo de IM, se calculó el área de la curva en el intervalo rojo. Este valor puede ir de 0, cuando el nivel fue 0, a 18 cuando el nivel fue 1 durante los 18 s del periodo.

Las Figura 3 y 4 muestran los niveles promedio de atención estimados para cada sujeto. Se muestran por separado los resultados de las pruebas en estático y las pruebas en movimiento. Se muestra el porcentaje del nivel de atención calculado como el área de la curva/18 \* 100.

A medida que los sujetos se fatigan, las BMI no proporcionan buenos resultados porque es más difícil para ellos mantener una alta atención a la marcha y se vuelven más propensos a las distracciones [10]. Sin embargo, en los resultados obtenidos no se aprecia una decaída del nivel de atención a lo largo de las sesiones. Éste se mantiene en el mismo rango. Y, en el caso del sujeto S2, se aprecia un incremento progresivo de este nivel en las pruebas en movimiento.

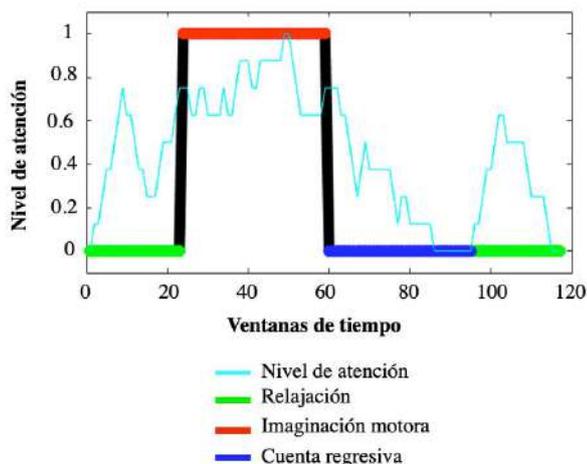


Figura 2. Ejemplo del nivel de atención a la marcha estimado para un usuario y una prueba.

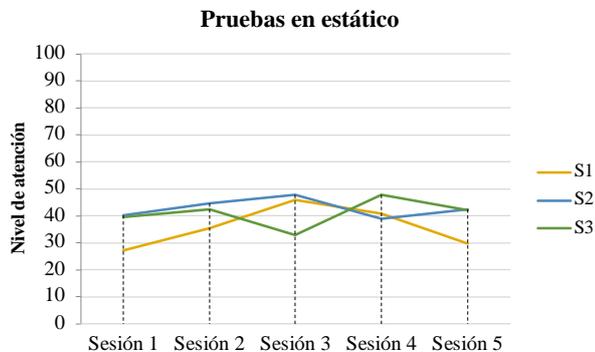


Figura 3. Nivel de atención estimado en las pruebas en estático para cada usuario y sesión.

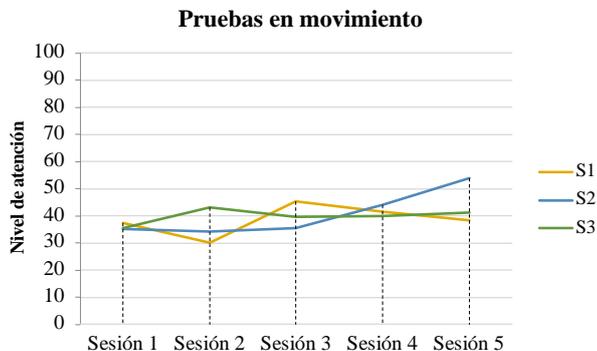


Figura 4. Nivel de atención estimado en las pruebas en movimiento para cada usuario y sesión.

A diferencia de las conclusiones obtenidas en trabajos previos con BMI [11], en los que se aprecian diferencias significativas entre usuarios; los resultados de este trabajo muestran que el nivel de atención a la marcha no difiere en gran medida entre ellos.

Las Figura 5, 6 y 7 muestran la evolución del nivel de atención a lo largo de cada sesión individual de entrenamiento. No hay una tendencia clara que indique que la atención fue variando con el tiempo y/o el nivel de fatiga de los usuarios. Algunas sesiones empezaron con un nivel de atención en 0, lo que indica que los usuarios no estaban concentrados en la tarea a realizar en la primera prueba. Sin embargo, este valor fue distinto de cero en las pruebas consecutivas, aunque no mostró una progresión lineal. Esto podría explicarse porque los usuarios necesitaban un tiempo para centrar su atención en las tareas mentales y no en su entorno u otro tipo de pensamiento. Este resultado fue más evidente en el sujeto S2.

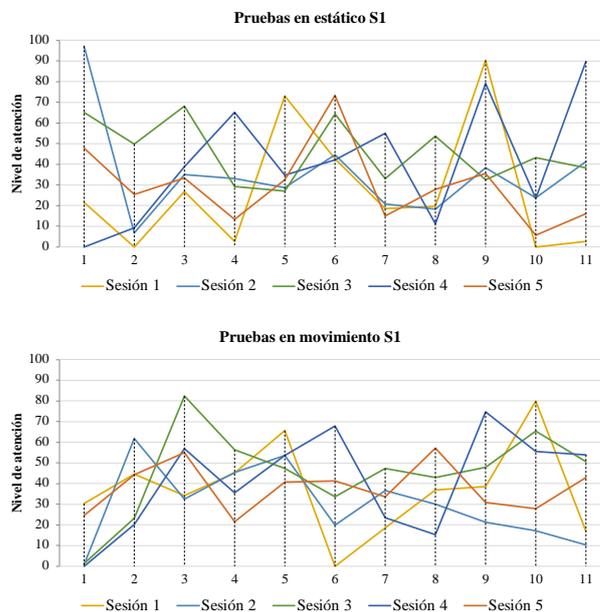


Figura 5. Evolución del nivel de atención a lo largo de cada sesión para el sujeto S1.

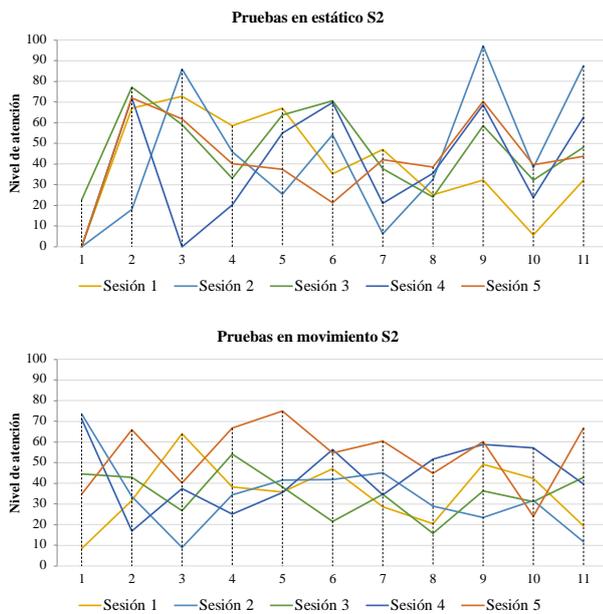


Figura 6. Evolución del nivel de atención a lo largo de cada sesión para el sujeto S2.

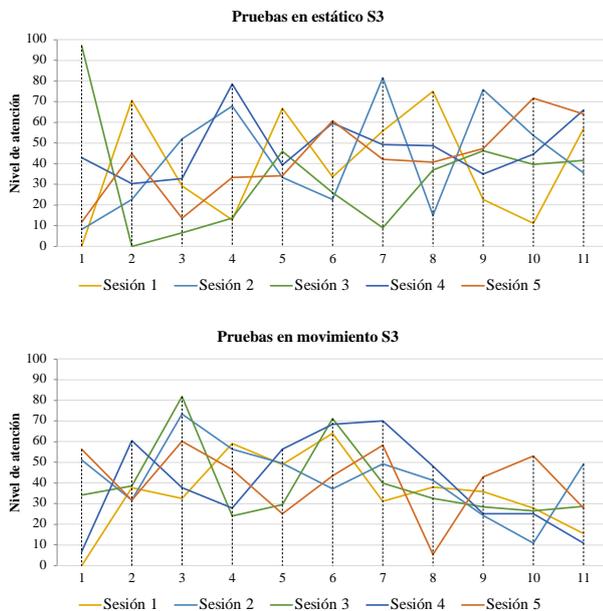


Figura 7. Evolución del nivel de atención a lo largo de cada sesión para el sujeto S3.

#### 4. Conclusión

El objetivo de este trabajo fue estudiar la evolución del nivel de atención de los usuarios durante el entrenamiento de una BMI con un exoesqueleto de miembro inferior. Se analizó si este nivel varía a lo largo de las sesiones de entrenamiento debido al esfuerzo que suponen y/o si existían diferencias entre las sesiones. Además, se estudiaron las diferencias entre sujetos.

No se apreciaron diferencias a lo largo de las sesiones ni entre ellas. Sin embargo, en algunos casos el nivel de atención inicial fue muy bajo debido a que los sujetos todavía no estaban concentrados en las tareas a realizar. Tampoco se hallaron diferencias relevantes entre los sujetos.

#### Agradecimientos

Esta publicación es parte del proyecto de I+D+i RTI2018-096677-B-I00, financiado por MCIN/AEI/10.13039/501100011033 y por “FEDER Una manera de hacer Europa”. Esta investigación ha sido financiada por el Ministerio de Ciencia, Innovación y Universidades a través de la Ayuda para la Formación de Profesorado Universitario FPU19/03165.

#### Referencias

- [1] Pfurtscheller G, Neuper C, Flotzinger D, Pregenzer M. EEG-based discrimination between imagination of right and left hand movement. *Electroencephalography and Clinical Neurophysiology*, vol 103, no 6, 1997, pp 642–651.
- [2] Kaplan A, Zhigulskaya D, Kiriyanov DA. Studying the ability to control human phantom fingers in P300 brain-computer interface, *Bulletin of Russian State Medical University*, 2016.
- [3] Kwak NS, Müller KR, Lee SW. A convolutional neural network for steady state visual evoked potential classification under ambulatory environment, *PLoS ONE*, vol 12, no 2, 2017, pp 1–20.
- [4] He Y, Eguren D, Azorín JM, Grossman RG, Luu TP, Contreras-Vidal JL. Brain-machine interfaces for controlling lower-limb powered robotic systems, *Journal of Neural Engineering*, vol 15, no 2, 2018.
- [5] Jeannerod M. Mental imagery in the motor context, *Neuropsychologia*, vol 33, no 11, 1995, pp 1419–1432.
- [6] Gharabaghi A. What Turns Assistive into Restorative Brain-Machine Interfaces? *Frontiers in Neuroscience*, vol 10, 2016, pp 456.
- [7] Van Dokkum LEH, Ward T, Laffont I. Brain computer interfaces for neurorehabilitation-its current status as a rehabilitation strategy post-stroke, *Annals of Physical and Rehabilitation Medicine*, vol 58, no 1, 2015.
- [8] Costa Á. *et al.* Decoding the Attentional Demands of Gait through EEG Gamma Band Features, *PLoS ONE*, 2016.
- [9] Ortiz M, Ferrero L, Iáñez E, Azorín JM, Contreras-Vidal JL. Sensory Integration in Human Movement: A New Brain-Machine Interface Based on Gamma Band and Attention Level for Controlling a Lower-Limb Exoskeleton. *Frontiers in Bioengineering and Biotechnology*, vol 8, 2020, pp 735.
- [10] Singh A, Hussain AA, Lal S, Guesgen HW. A Comprehensive Review on Critical Issues and Possible Solutions of Motor Imagery Based Electroencephalography Brain-Computer Interface. *Sensors*, vol 21, no 6, 2021.
- [11] Ferrero L, Quiles V, Ortiz M, Iáñez E, Azorín JM. A BMI Based on Motor Imagery and Attention for Commanding a Lower-Limb Robotic Exoskeleton: A Case Study. *Applied Sciences*, vol 11, no 9, 2021.

# Neuroprótesis para evaluar el efecto de estimulación vibratoria mecánica en el freezing de pacientes con Parkinson

Inés Martín Martínez <sup>1</sup>, J.S. Lora-Millán <sup>2</sup>, E. Rocon <sup>1</sup>

<sup>1</sup> Centro de Automática y Robótica, CSIC-UPM, Madrid, España, e.rocon@csic.es

<sup>2</sup> Departamento de Tecnología Electrónica, Universidad Rey Juan Carlos, Madrid, España, julio.lora@urjc.es

## Resumen

La enfermedad de Parkinson (EP) es el segundo trastorno neurodegenerativo más común en el mundo y afecta aproximadamente al 1% de la población mayor de 60 años. Dependiendo del estadio de la enfermedad, entre el 20 y el 60% de los individuos con enfermedad del Parkinson sufren de congelamiento episódico de la marcha (freezing). El freezing es un fenómeno motor caracterizado por periodos transitorios, que suelen durar varios segundos, en los que se detiene el intento de movimiento. Este artículo describe el diseño conceptual y desarrollo de una neuroprótesis capaz de analizar la marcha de un paciente con EP y estimular mecánicamente sus vías aferentes en sincronía con su estado de marcha con el objetivo de reducir los episodios de 'freezing'.

## 1. Motivación

La enfermedad de Parkinson (EP) es el segundo trastorno neurodegenerativo más frecuente en el mundo y afecta a alrededor del 1 % de la población mayor de 60 años [1]. Está caracterizada por una pérdida progresiva de neuronas dopaminérgicas de la sustancia negra pars compacta (SNpc) y por la presencia de cuerpos de Lewy. Según el estadio de la enfermedad, entre el 20 y el 60% de las personas con EP sufren episodios de bloqueo de la marcha (freezing). El freezing es un fenómeno motor caracterizado por periodos transitorios, que suelen durar varios segundos, en los que se detiene el intento de deambulación [2]. Esta perturbación en la marcha disminuye la calidad de vida de los pacientes y puede conllevar caídas del paciente provocando traumatismos [3].

Uno de los síntomas principales de la enfermedad de Parkinson (EP) es la incapacidad para modular la ganancia de esta información sensorial, lo que genera bradicinesia y lentitud del movimiento [4]. Los pacientes con EP muestran un déficit en la activación de los potenciales evocados somatosensoriales, que se sabe que se atenúan con el movimiento [5]. Hallazgos recientes han evidenciado que la modulación de la incertidumbre de la señal aferente propioceptiva mejora potencialmente el rendimiento motor al controlar la señal sensorial entrante y permitir predicciones propioceptivas de arriba hacia abajo. Keijser et al. [6] demostraron que los pacientes con EP tienen déficits no solo en el rendimiento motor, sino también en el procesamiento de la información propioceptiva. Se observan déficits propioceptivos incluso con síntomas muy leves de EP, y los déficits en el uso de la retroalimentación visual se desarrollan progresivamente en etapas posteriores de la enfermedad [6].

La aproximación propuesta en este artículo se basa en la hipótesis de que la estimulación mecánica aferente puede influir en el control motor. Durante la locomoción ficticia o real, el estiramiento o la vibración periódica de los músculos proximales y distales pueden arrastrar o restablecer el ritmo locomotor, modulando la amplitud y la fase de la actividad extensora y flexora [7].

Para evaluar esta hipótesis, este artículo propone el desarrollo de una neuroprótesis capaz de generar estímulos vibratorios (sobre músculos específicos del miembro inferior) sincronizados con las fases de la marcha. De esta manera, podremos estudiar este efecto en el freezing de pacientes con Parkinson.

Esta neuroprótesis tomará la forma de un textil que un paciente puede ponerse en su extremidad inferior (Figura 1). Esta plataforma comprenderá: 1) actuadores hápticos colocados en áreas específicas de los miembros inferiores, y 2) una prenda que integra sensores inerciales (para detectar los parámetros y eventos de la marcha), y la electrónica para el control de la estimulación vibratoria. Esta plataforma proporcionará un registro a largo plazo de los datos biomecánicos de las extremidades inferiores y controlará la estimulación durante la vida diaria del paciente mediante la estimulación aferente de las vías sensibles.

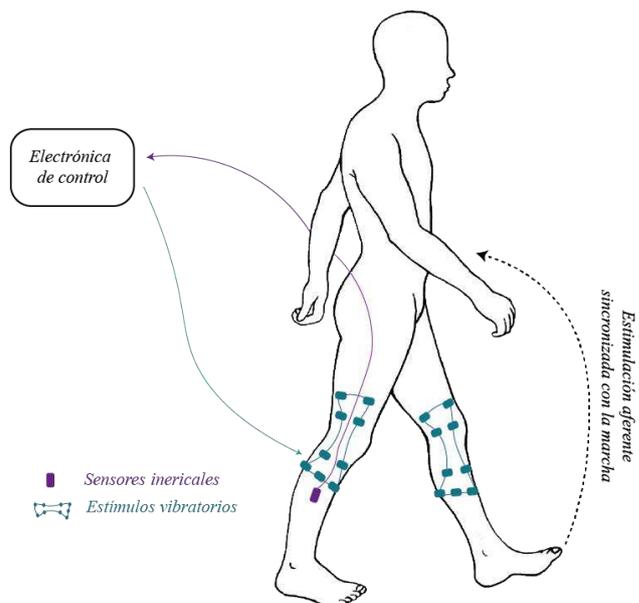
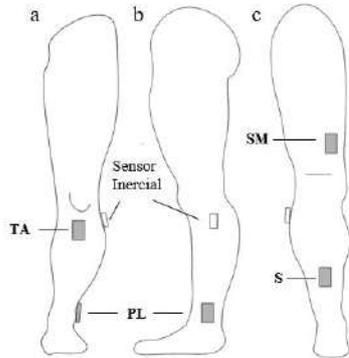


Figura 1. Diseño conceptual de la neuroprótesis.

## 2. Diseño del dispositivo

El sistema diseñado actualmente consta de un sensor inercial (modelo BNO055, Adafruit, E.E.U.U.), posicionado en cada pierna (Figura 1). Para la estimulación mecánica hemos seleccionado los motores hápticos modelo 304-116 (Precision Microdrives, Reino Unido), que se distribuyen en los miembros inferiores como se ilustra en la Figura 2:



**Figura 2.** Configuración del experimento en pierna izquierda. (a) Vista anterior. Actuadores hápticos en tibial anterior (TA) 3 cm por debajo de la rótula y en peroneo lateral (PL). (b) Vista lateral. Actuador háptico sobre el peroneo lateral (PL). (c) Vista posterior. Actuadores posicionados sobre el sóleo (S) y sobre el inicio del nervio tibial a la altura de la terminación del músculo semitendinoso (SM).

La arquitectura de control se basa en una *Single Board Computer* Raspberry Pi 4. Esta procesa los datos medidos por el sensor inercial, obteniendo la fase y la frecuencia de la marcha, así como el contacto inicial del talón con el suelo (CI). Esta electrónica también es responsable de la activación de los actuadores en la fase de la marcha precisa que estarán fijados al músculo por bandas elásticas para la comodidad de los sujetos.

## 3. Detección de la fase de la marcha

La sincronización de la estimulación con la marcha del paciente se basa en dos algoritmos. El primero es responsable por la estimación de la fase y frecuencia de la marcha. Se base en un oscilador de frecuencia adaptativo (AFO), capaz de adaptarse a los cambios del ritmo de la marcha [8]. El segundo algoritmo realiza la detección del CI al inicio de la marcha y permite la sincronización de la fase de marcha estimada con el ciclo de marcha. Este algoritmo se basa en detectar el ángulo mínimo de flexión y extensión [9].

### 3.1. Detección del contacto inicial

Siguiendo el enfoque propuesto por Chia et al. [9], el algoritmo presentado está basado en el análisis del ángulo que describe la pierna del sujeto con respecto a la vertical. Siguiendo esta aproximación, el CI se produce en el mínimo del ángulo analizado.

La detección de este mínimo se realiza en tiempo real cuando el mínimo de la señal es inferior a un umbral determinado. Este umbral se calcula como el valor medio más tres veces la desviación típica de los ángulos mínimos de los cinco últimos pasos registrados.

### 3.2. Oscilador de frecuencia Adaptativo

Para extraer la fase y la frecuencia de la marcha del paciente en tiempo real se utiliza un oscilador de frecuencia adaptativo (AFO) [8]. Para iniciar el funcionamiento del AFO, se considera que la marcha ha comenzado cuando el ángulo de la pierna sea mayor que diez grados y se haya detectado un contacto inicial en ese momento.

Para extraer la fase del ciclo de la marcha usamos se utiliza el ángulo obtenido por los sensores inerciales  $\theta_m(t)$ . La fase de la marcha  $\varphi(t)$  es una señal continua que aumenta con el tiempo y se resetea cada vez que se completa un ciclo de la marcha.

La entrada del oscilador es el error de la señal  $e(t) = \theta_m(t) - \theta_{rec}(t)$ , donde  $\theta_{rec}(t)$  es el ángulo estimado que se obtiene gracias a la descomposición en series de Fourier. Para esta descomposición, se utilizan cinco términos definidos por coeficientes  $\alpha_k$  y  $\beta_k$  que se actualizan y calculan siguiendo las siguientes ecuaciones:

$$\theta_{rec} = \sum_{k=0}^{N_f} \alpha_k \cos(k\varphi) + \beta_k \sin(k\varphi) \quad (1)$$

$$\dot{\alpha}_k = \eta \cos(k\varphi) e(t) \quad (k = 0, \dots, N_f) \quad (2)$$

$$\dot{\beta}_k = \eta \sin(k\varphi) e(t) \quad (3)$$

$$\dot{\varphi} = \omega - \epsilon e(t) \sin \varphi \quad (4)$$

$$\dot{\omega} = -\epsilon e(t) \sin \varphi \quad (5)$$

Después de hacer los cálculos anteriores se normaliza la fase y la frecuencia del oscilador de acuerdo a la ecuación (6). Este cambio de variables se realiza para que la fase sea una función creciente en el rango  $[0, 1]$  y la frecuencia se interprete en pasos (o ciclos) por segundos (Hz).

$$\sigma(t) = \frac{\varphi(t)}{2\pi}, \quad f(t) = \frac{\omega(t)}{2\pi} \quad (6)$$

Variables	Descripción
$\theta$	Ángulo de la pierna
$\varphi$	Fase del oscilador
$\omega$	Frecuencia oscilador
$\eta$	Constante de aprendizaje
$\epsilon$	Fuerza de acoplamiento
$\alpha, \beta$	Coefficientes de Fourier
$\sigma$	Fase del ciclo de la marcha
$f$	Frecuencia ciclo de la marcha (Hz)

**Tabla 1.** Variables de las ecuaciones.

### 3.3. Sincronización

La convergencia del AFO no asegura que una fase de  $\varphi=0$  se corresponda con el inicio del ciclo de marcha. Para solventarlo, se utiliza la aproximación propuesta por Van Dijk et al. [10]. Así, el sistema corrige la fase mediante un offset que es adquirido cuando se detecta el CI:

$$\sigma_c = \sigma_s - \rho \quad (7)$$

Siendo  $\rho$  el offset que se corresponde con la fase  $\sigma$  en el momento del CI,  $\sigma_c$  la fase corregida y  $\sigma_s$  la fase sin corregir.

Para evitar cambios bruscos en el valor de la fase corregida, se aplica un filtro IIR Butterworth paso bajo de orden 1 y frecuencia de corte 0.5 Hz. Esto permite que la corrección de la fase sea gradual gracias al filtrado del offset.

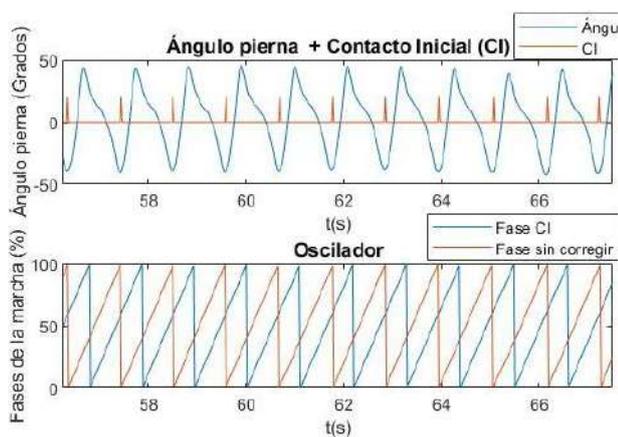
## 4. Resultados

Para evaluar esta plataforma se llevaron a cabo dos sesiones experimentales con objetivos diferentes. En primer lugar se evaluó el algoritmo para detección de la marcha y en segundo lugar se llevaron a cabo ensayos experimentales para analizar los efectos de la estimulación aferentes en los miembros inferiores.

### 4.1. Detección de la marcha

Para validar el procesamiento de los datos de la marcha se realizó un experimento con un sujeto sano. Se colocó sobre la pierna izquierda el sensor inercial en la posición descrita en la Figura 1. El sujeto partía de una posición en bipedestación con ambos pies a la misma altura. Comenzó a andar por una superficie lisa 20 metros mientras se registraban, en tiempo real, los datos de la marcha. Se especificó al sujeto que debía caminar a una velocidad en la que sintiera cómodo.

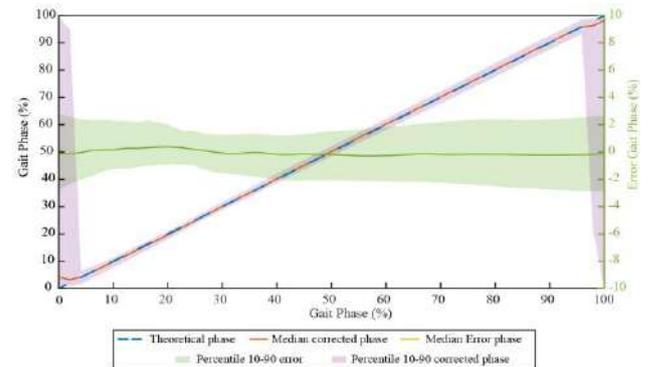
El resultado de la medición se ilustra en la Figura 3. En la gráfica superior se puede apreciar el ángulo descrito por la pierna y el contacto inicial determinado. En la gráfica inferior se observa la fase estimada por el AFO, así como esta señal corregida con la detección del CI. Se aprecia la alineación de la fase corregida con el contacto inicial de la gráfica superior.



**Figura 3.** Experimento de entrenamiento. Gráfica superior: El ángulo de la pierna medido con el sensor inercial y el contacto inicial del talón con el suelo (CI). Gráfica inferior: Fase de la marcha sin corregir y fase sincronizada con el CI.

La Figura 4 ilustra la fase corregida y el error en la estimación de la fase instantánea. Las señales fueron segmentadas en ciclos de marcha de acuerdo con el CI detectado, luego se interpolaron del 0 % al 100 % en 100

muestras equidistantes. En base a estos datos interpolados, se calculan la mediana de la fase y el error, así como sus percentiles 10 y 90. La mediana de la fase corregida estimada con el AFO (línea naranja) en un experimento se comparó con la fase teórica (línea azul discontinua). Los percentiles 10 – 90 de la mediana del error (sombra verde), de uno de los intentos del sujeto, oscila entre  $\pm 3\%$ . Además, los percentiles 10 y 90 de las fases corregidas (sombra púrpura) están entre el 5 % cuando el cambio de fase descende del 100 % al 0 % de la fase. En el resto de la señal la dispersión entre el percentil 10 y 90 de la fase corregida es despreciable. Los cambios bruscos al inicio y al final del percentil de la fase de la marcha se deben a la corrección de la fase en este instante del ciclo de marcha.



**Figura 4.** Mediana de la fase obtenida con el algoritmo AFO comparada con la mediana de la fase teórica. Mediana del error de fase. Percentiles 10-90 de la fase corregida y del error.

### 4.2. Estimulación de los miembros inferiores

Se realizaron una serie de pruebas para validar la utilización de dos actuadores vibratorios para estimular la propiocepción y su colocación, en las posiciones indicadas en la Figura 1. Cuatro sujetos sanos participaron en el experimento. Se colocó un actuador en cada pierna a cuatro sujetos sanos, para ver su efecto en la propiocepción estando en bipedestación. La sujeción de los actuadores a la pierna se realizó con una banda elástica. La frecuencia a la que el actuador estimula son 80Hz, que es la frecuencia a la que se maximiza la respuesta de los husos musculares [11]. En cada posición los actuadores vibraron durante dos tiempos distintos, 8s y 15s, para comprobar cuando empezaban a ver afectada su propiocepción.

Inicialmente se colocaron los dos sobre la misma pierna, pero los sujetos reportaron no ver afectada su propiocepción debido a que la otra pierna les servía de referente de equilibrio. Esto sucedía también cerrando los ojos, para evitar referencia visual.

Cuando se colocó un actuador en cada pierna, se probó en las posiciones del tibial anterior, sóleo, peroneo lateral (tobillo externo) y músculo semitendinoso. De los cuatro sujetos, todos reportaron mayor sensación de propiocepción alterada cuando la duración de la vibración era de 15 segundos. Aunque en el peroneo lateral y tibial anterior a los 8s ya notaban “una ligera inclinación del cuerpo hacia delante” según describieron. Tres de los 4 sujetos sentían una inclinación hacia delante en las cuatro

posiciones probadas, que incrementaba si cerraban los ojos. Esto ocurre por eliminar la referencia visual. Pero uno de los sujetos comentó que sentía una “ligera inclinación hacia atrás” al colocarle los vibradores en el sóleo.

## 5. Conclusión

El objetivo de esta contribución fue presentar el diseño conceptual de una neuroprótesis para estimular las vías aferentes mediante estimulación mecánica vibratoria para promover la marcha en pacientes con enfermedad de Parkinson. Para lograr este objetivo, se ha definido y desarrollado la electrónica y los algoritmos para implementar este concepto. Los algoritmos se desarrollaron para procesar en tiempo real la marcha de un sujeto y estimando su fase y detectando el evento del contacto inicial con el suelo. Esto permitió la sincronización de la estimulación mecánica con la fase de marcha del paciente.

En este experimento se han validado los algoritmos para la monitorización de la marcha y la extracción de la fase en tiempo real. Además, se comprobó que la localización de los estimuladores en el tibial anterior, sóleo, peroneo lateral y músculo semitendinoso parece la adecuada para estimular la propiocepción de los sujetos. Como pasos futuros se propone evaluar posiciones alternativas para los vibradores y medir empíricamente la influencia que tiene en la propiocepción y sincronizar la activación de los actuadores piezoeléctricos con el ciclo de la marcha de los pacientes. Además, se pretende evaluar los efectos de esta plataforma en pacientes con EP para analizar de manera fehaciente la idoneidad de la técnica presentada.

## Referencias

- [1] S. L. Kowal, T. M. Dall, R. Chakrabarti, M. v Storm, and A. Jain, “The current and projected economic burden of Parkinson’s disease in the United States.,” *Mov Disord*, vol. 28, no. 3, pp. 311–8, Mar. 2013.
- [2] F. Fröhlich, “Parkinson’s Disease,” in *Network Neuroscience*, Elsevier Inc, 2016, pp. 291–296.
- [3] B. R. Bloem, J. M. Hausdorff, J. E. Visser, and N. Giladi, “Falls and Freezing of Gait in Parkinson’s Disease: A Review of Two Interconnected, Episodic Phenomena,” *Movement Disorders*, vol. 19, no. 8, pp. 871–884, 2004.
- [4] P. Limousin *et al.*, “High- frequency peripheral vibration decreases completion time on a number of motor tasks,” *European Journal of Neuroscience*, vol. 48, no. 2, pp. 1789–1802, 2018.
- [5] A. Macerollo *et al.*, “High-frequency peripheral vibration decreases completion time on a number of motor tasks,” *European Journal of Neuroscience*, vol. 48, no. 2, pp. 1789–1802, 2018.
- [6] N. L. W. Keijsers, M. W. I. M. Horstink, and S. C. a M. Gielen, “Ambulatory motor assessment in Parkinson’s disease.,” *Mov Disord*, vol. 21, no. 1, pp. 34–44, Jan. 2006.
- [7] Y. P. Ivanenko, R. Grasso, and F. Lacquaniti, “Influence of Leg Muscle Vibration on Human Walking,” *J Neurophysiol*, vol. 84, no. 4, pp. 1737–1747, 2020.
- [8] G. Aguirre-Ollinger, A. Narayan, and H. Yu, “Phase-Synchronized Assistive Torque Control for the Correction of Kinematic Anomalies in the Gait Cycle,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 27, no. 11, pp. 2305–2314, Nov. 2019.
- [9] N. Chia Bejarano, E. Ambrosini, A. Pedrocchi, G. Ferrigno, M. Monticone, and S. Ferrante, “A novel adaptive, real-time algorithm to detect gait events from wearable sensors.,” *IEEE Trans Neural Syst Rehabil Eng*, vol. 23, no. 3, pp. 413–22, May 2015.
- [10] W. van Dijk, C. Meijneke, and H. van der Kooij, “Evaluation of the achilles ankle exoskeleton,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, no. 2, pp. 151–160, Feb. 2017.
- [11] E. Ribot-Ciscar, J. M. Aimonetti, and J. P. Azulay, “Sensory training with vibration-induced kinesthetic illusions improves proprioceptive integration in patients with Parkinson’s disease,” *J Neurol Sci*, vol. 383, no. June, pp. 161–165, 2017.

# Caracterización espacio-temporal de la clasificación de imaginación motora con herramientas de *explainable artificial intelligence* (XAI)

S. Pérez Velasco<sup>1</sup>, D. Marcos-Martínez<sup>1</sup>, E. Santamaría-Vázquez<sup>1,2</sup>, V. Martínez-Cagigal<sup>1,2</sup>, S. Moreno-Calderón<sup>1</sup>, R. Hornero<sup>1,2</sup>

<sup>1</sup> Grupo de Ingeniería Biomédica (GIB), Universidad de Valladolid, Valladolid, España, sergio.perez@gib.tel.uva.es

<sup>2</sup> Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), España

## Resumen

El uso de la imaginación motora (MI) para controlar interfaces cerebro ordenador (BCI) se ha realizado generalmente a través de la señal eléctrica del cerebro registrada mediante electrograma (EEG). En concreto, a partir de los ritmos sensoriomotres (SMR) presentes en la señal de los electrodos centrales C3 y C4. A partir de los SMR se ha fundamentado la clasificación de MI tradicional basada en métodos de machine learning. Sin embargo, recientemente las redes convolucionales (CNN), una clase de red de deep learning, se han mostrado como una alternativa superior para la decodificación del EEG. Esta superioridad puede radicar en la extracción de características consideradas anteriormente irrelevantes. En este trabajo presentamos una herramienta de explainable artificial intelligence basada en shapley additive explanations (SHAP). Con ella representamos la información que ha sido más relevante para la decodificación de MI a partir del EEG en una red CNN. Nos basamos en la implementación abierta de EEGSym para obtener los valores SHAP en una base de datos de 109 sujetos para los eventos de MI de mano derecha e izquierda. Gracias a esta herramienta, descubrimos que EEGSym presta más atención a la señal de los electrodos frontales F7 y F8. Asociamos esta relevancia a la decodificación de la planificación motora que se da en esa región del cerebro. Estos hallazgos pueden mejorar la selección de canales en la decodificación de MI y permitir el uso de sistemas más eficientes.

## 1. Introducción

La actividad eléctrica cerebral se puede registrar de manera no invasiva con la electroencefalografía (EEG) a través de electrodos colocados en el cuero cabelludo del usuario. Las interfaces cerebro ordenador (BCI) hacen uso del EEG para crear un camino de comunicación alternativo entre el cerebro y el entorno [1]. La facilidad de uso, alta resolución temporal y bajo coste relativo del EEG con respecto a sus alternativas, ha llevado a su amplio uso en la investigación del cerebro [1]. Sin embargo, el EEG también es muy ruidoso y tiene baja resolución espacial, por lo que los sistemas BCI tradicionalmente utilizan señales de control, que se activan a través de estrategias conocidas como paradigmas que aumenten la señal frente a este ruido siempre presente en los registros [1].

En este trabajo nos centramos en el paradigma de imaginación motora (MI) [1], en concreto en la imaginación del movimiento de mano izquierda o derecha, recientemente implementado en terapias de rehabilitación de ictus con prometedores resultados [2]. A pesar del uso

de un paradigma como MI para conseguir una señal de control discernible del ruido presente en el EEG, decodificar la intención del usuario sigue teniendo margen de mejora. Para solventar los problemas de variabilidad inter-sesión e inter-sujeto presentes en los sistemas de clasificación tradicionales basados en *machine learning* (ML), se ha investigado aprovechar la capacidad de generalización de los sistemas de *deep learning* (DL) [3]. En nuestro anterior trabajo presentamos EEGSym, una red convolucional (CNN) que permite clasificar con gran precisión la intención de nuevos usuarios [3].

A pesar de las ventajas que ofrecen las nuevas técnicas de DL frente a ML, se considera que son una “caja negra” de la cual no se puede extraer información sobre su proceso de decisión. El campo de investigación de *explainable artificial intelligence* (XAI) surge para dar respuesta a este problema, permitiendo descubrir en qué información se basan para realizar su clasificación. Para los registros de EEG, se han adaptado herramientas XAI basadas en métodos de propagación como *layerwise relevance propagation* [4] y Grad-CAM [5], o métodos de perturbación como easyPEASI [6] y *occlusion sensitivity analysis* [7]. Pero estas aplicaciones no cumplen ciertos principios matemáticos que aseguren la obtención de una única atribución de importancia de los datos usados para la clasificación. Otros investigadores [8] han usado ya la herramienta *shapely additive explanations* (SHAP) [9] para registros de EEG. Sin embargo, se usaba sobre un clasificador basado en árboles de decisión, que tiene una capacidad más limitada que las redes de DL. En este trabajo adaptamos la herramienta SHAP, para explicar una red de DL entrenada con registros de EEG durante MI. SHAP unifica previos métodos de XAI y asegura que la explicación de la clasificación realizada sea única y no dependa de parámetros de inicialización.

El objetivo de este trabajo es representar los valores SHAP asociados a la toma de decisiones de EEGSym en una base de datos de 109 sujetos [10]. Mediante esta representación, observaremos qué información es más relevante para EEGSym a la hora de clasificar eventos de MI.

## 2. Métodos

### 2.1. Dataset

Se ha analizado la base de datos pública Physionet [10]. Esta base de datos cuenta 109 usuarios y entre 42 y 46

eventos de MI por sujeto. Tras un aviso sonoro, una flecha indicaba el tipo de MI a realizar (i.e., imaginación de apertura y cierre de mano izquierda o derecha) durante los siguientes 3 segundos de manera continuada. Se realizaba una única sesión en la que no se les presentaba realimentación de la tarea mental que estaban realizando.

## 2.2. EEGSym

Se utiliza la implementación abierta de EEGSym [3] para 16 canales. EEGSym es una nueva red CNN de clasificación de MI inter-sujeto. Incluye nuevos avances en la aplicación de las técnicas de DL: conexiones residuales, técnicas de *data augmentation*, *transfer learning* entre sujetos y una disposición siamesa de la red que aprovecha la simetría del cerebro en el plano sagital medio.

La red ha demostrado ser significativamente superior a cuatro redes previas en la tarea de clasificación de un problema binario de MI inter-sujeto [3]. Junto con el conjunto de novedosas técnicas de DL con las que se aplica establece unos resultados que son un nuevo estado del arte en este tipo de clasificación binaria.

Para obtener las representaciones de importancia de las zonas del EEG en la clasificación de MI, utilizamos los pesos de EEGSym adaptados a la base de datos de Physionet [10]. Para ello, partimos de la inicialización disponible en la implementación abierta para 16 canales, y con ella realizamos una etapa de *fine tuning* a la base de datos de Physionet [10], como en el artículo original.

## 2.3. SHAP

La herramienta SHAP en Python [9] ha tenido gran aceptación debido a unificar y adaptar previas herramientas de XAI. Utiliza una adaptación de herramientas populares de XAI para producir una aproximación de valores Shapley, haciendo que estos cumplan tres propiedades: precisión local, pérdidas, y consistencia [9]. Cumpliendo estas propiedades, se obtiene una representación única de la importancia de las características de entrada.

Se utiliza la herramienta del paquete SHAP llamada “GradientExplainer”, una adaptación de *integrated gradients* y SmoothGrad para obtener una aproximación de los valores Shapley. En modelos de DL no se pueden aplicar los métodos exactos por a su alto coste computacional y es necesario el uso de un cálculo aproximado. Esta herramienta calcula el gradiente (i.e., cambio en la predicción) entre el *background* (en nuestro caso un EEG con señal 0 en todos los canales e instantes temporales), la señal de entrada y las diferentes representaciones intermedias resultantes de interpolar entre *background* y señal. El cambio entre la predicción realizada y la esperada se atribuye a este gradiente por cada punto de interpolación. Sumando todas las atribuciones intermedias se obtiene los valores SHAP para esa señal. Posteriormente, se hace la media para todas estas matrices de valores SHAP obtenidas para una misma clase de MI en toda la base de datos. Solo se incluyen los ejemplos correctamente clasificados, es decir, aquellos ejemplos en los que la red predice correctamente la clase a explicar.

Las matrices de valores SHAP se corresponden con el EEG usado para la clasificación. En el eje vertical se representan los distintos canales, y en el eje horizontal la magnitud de tiempo. Gracias a ello, esta representación tiene una correspondencia directa con los canales y periodos de tiempo. Los valores SHAP positivos incrementan la probabilidad de predecir la clase a explicar, mientras que los valores SHAP negativos reducen la probabilidad de predecir dicha clase. Con estos valores se puede calcular la importancia relativa que tiene cada canal o instante temporal en la predicción de cada clase.

## 3. Resultados

Podemos observar en la Figura 1 a y b, la representación de los valores SHAP obtenidos para los eventos de MI de mano izquierda y mano derecha, respectivamente. Como hemos explicado anteriormente, se representan las explicaciones de la predicción obtenida sólo en los eventos en los que la red neuronal ha acertado. En Physionet, EEGSym tenía una precisión inter-sujeto de  $88.6 \pm 9.0$ , por lo que se representa la media para casi todos los eventos disponibles en esta base de datos. Cada evento contiene los tres primeros segundos tras el *onset* de MI para los 16 canales usados en la predicción. De las figuras, se obtiene información tanto a nivel temporal como a nivel espacial.

En rojo se representan los valores positivos SHAP y en azul los valores negativos. Además, incluye también las medias por canales con un porcentaje que indica cuánto influye ese canal a la predicción en valores absolutos, ya sea por incrementar la probabilidad de predecir esa clase o por disminuirla. En el eje temporal, también se representa la media que todos estos valores tienen en todos los canales. Las medias temporal y espacial nos permiten observar la influencia que los distintos instantes de tiempo y los distintos canales tienen en la clasificación.

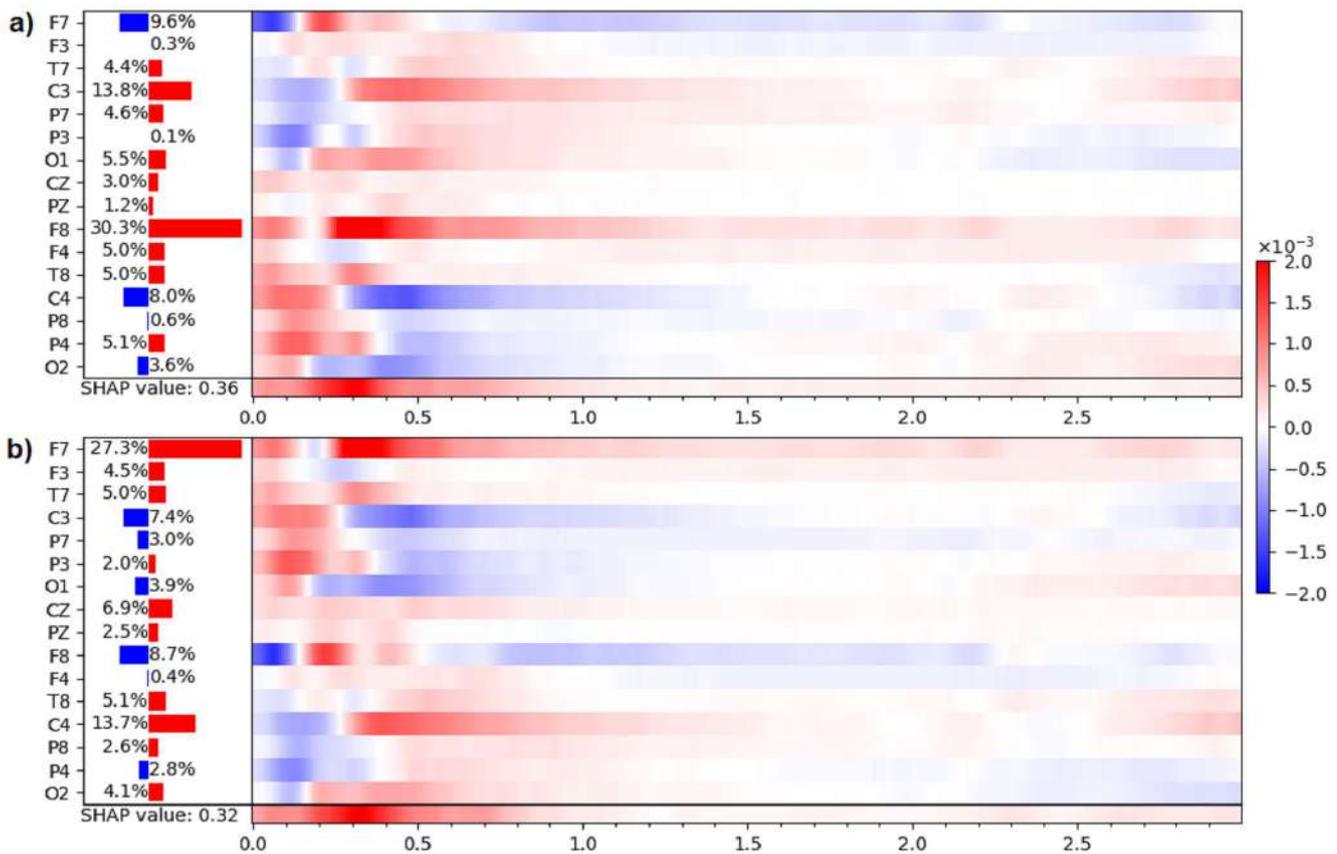
### 3.1. Análisis temporal

Se observa que, para la clasificación binaria, la región temporal más determinante para clasificar la señal con EEGSym se encuentra en el primer segundo de la señal. Más concretamente, se puede observar cómo la mayor contribución a la clasificación se encuentra entre 0.2 y 0.4 segundos, que coincide con el rango de tiempo de reacción humana. Los últimos dos segundos, aunque en algunos canales tienen una influencia similar a los primeros segundos, de manera global son menos relevantes al anularse las aportaciones positivas de unos canales con las aportaciones negativas de otros.

Analizando temporalmente cada canal observamos que ese primer segundo es el más relevante en la mayoría, sin embargo, en los canales C3 y C4 se extiende hasta segundo y medio tras el *onset*. Este comportamiento también se observa en F7 y F8, cuya aportación negativa a la predicción de clase izquierda y clase derecha, respectivamente, se encuentra entre los 0.7 y 2.2 segundos.

### 3.2. Análisis espacial

Los porcentajes presentados para cada canal representan la importancia absoluta relativa de cada canal en la predicción. A partir de esta información descubrimos que



**Figura 1.** Representación de las atribuciones de valores SHAP para eventos de imaginación motora (MI) en Physionet. En el eje vertical se representan los 16 canales del sistema 10-20. En el eje horizontal los tres segundos correspondientes al evento de MI tras el onset. En rojo se indican los valores SHAP positivos, que indican que esa región tiempo-espacial contribuye positivamente a la predicción de la clase. En azul se indican los valores SHAP negativos, que por el contrario indican que contribuye negativamente a la predicción de esa clase. a) MI de mano izquierda b) MI de mano derecha

EEGSym encuentra los canales frontales, F7 y F8, cómo los más determinantes para determinar la clase de MI realizada. Después se fija en los canales centrales, C3 y C4. El resto de canales realizan aportaciones netas más marginales a estos cuatro canales, que concentran el 60% de la influencia sobre la decisión de clasificación. Los siguientes canales en importancia serían F3, F4, T7 y T8.

Para ambos eventos de MI, utiliza una atribución prácticamente simétrica de los canales. Para clasificar mano izquierda, EEGSym encuentra a F8 y C3 como los canales más importantes. Mientras que la señal de los canales F7 y C4 de manera neta disminuye la probabilidad de predecir dicha clase. En cambio, para mano derecha, este comportamiento es exactamente el contrario.

#### 4. Discusión

A la hora de interpretar estos mapas de valores SHAP nos encontramos con que, a diferencia de la literatura previa, EEGSym encuentra como más relevantes los canales frontales F7 y F8 frente a los canales C3 y C4 típicamente escogidos. Tradicionalmente, cuando se aplican métodos de clasificación de ML se debe realizar una etapa previa de selección de características que no suele ser automática. Los eventos de MI se relacionan con la presencia de los ritmos sensoriomotores (SMR), una señal presente en la actividad eléctrica del surco central del cerebro, que para eventos de MI de mano izquierda y derecha se encontraba presente en la señal de los electrodos C4 y C3,

respectivamente. Los SMR se caracterizaban por un evento de desincronización en el electrodo contralateral, seguido de un evento de sincronización [1], que se puede observar de manera visual al promediar la señal de varios eventos de MI. De esta manera, los métodos de ML confiaban en los SMR y utilizaban la señal de los canales C3 y C4 (y en ocasiones también de los adyacentes) para realizar la clasificación de MI. Gracias a la combinación de XAI con la red EEGSym de DL, podemos observar cómo, al menos en esta base de datos de 109 usuarios, hay información incluso más relevante en los electrodos frontales F7 y F8. La decisión de predicción con respecto a estos canales se podría realizar incluso más rápido que el tradicional uso de C3 y C4, ya que la mayor parte de la aportación a la decisión se encuentra en el primer segundo, con respecto a la señal de C3 y C4 más distribuida durante segundo y medio. Estos descubrimientos se podrían aplicar para realizar clasificación de MI con épocas de menor duración.

Hemos visto como el mapa de valores SHAP obtenido para el registro de EEG utilizado en la clasificación es útil para obtener información de qué regiones espacio-temporales son importantes para clasificar eventos de MI con EEGSym. EEGSym recientemente ha conseguido un rendimiento superior a los modelos anteriores en la clasificación de MI inter-sujeto, lo que da mayor seguridad de la generalización de estas interpretaciones [3]. Estos valores SHAP nos permiten tener información de la relevancia espacio-temporal de los registros de EEG de MI.

Así encontramos como para EEGSym es más relevante el primer segundo de registro, y espacialmente la señal de los canales F7 y F8 seguidos de C3 y C4. En concreto, la señal que se obtiene de F7 y F8 está relacionada con la planificación del movimiento [11]. Es posible que EEGSym haya aprendido a identificar en la señal EEG de estos dos electrodos cómo se codifica la decisión del movimiento a imaginar. De la misma manera que los SMR de la zona central, esta decisión entre planificación de mano izquierda y mano derecha parece presentarse de manera simétrica en F7 y F8. Si nos fijamos temporalmente, la decisión en estos dos electrodos es más rápida que la de los electrodos C3 y C4, que contendrían la señal del SMR asociado a la imaginación del movimiento.

Una línea futura de investigación a partir de estos resultados, sería comprobar como un clasificador centrado en la señal de estos cuatro canales pueda alcanzar valores similares de clasificación a los obtenidos con estos 16 canales. En la publicación original de EEGSym con 8 canales se conseguía una precisión de 84.5% frente a la de 88.6% de 16 canales. Estos 8 canales no incluían los electrodos F7 y F8, que tan determinantes son en la clasificación de 16 electrodos, por lo que es muy probable que con una elección de 8 canales que los incluya, la precisión pueda llegar a ser más semejante con solo 8 canales. Además, también se podría intentar reducir el tiempo de clasificación a ese primer segundo que se ha mostrado como el más importante en la clasificación de MI sin realimentación. Por otro lado, también sería interesante estudiar si los valores SHAP son similares entre MI con realimentación y sin ella. Además, al haber encontrado en el paradigma de MI la correspondencia con el conocimiento previo, se podrían extraer información de las redes DL en otros paradigmas menos estudiados.

A pesar de estos interesantes resultados y su correlación con conocimientos previos sobre los eventos de MI, este trabajo tiene algunas limitaciones que es necesario apuntar. Para aprovechar el trabajo previo realizado, se han utilizado los 16 canales presentes en la publicación de EEGSym que buscaban homogeneizar los presentes en las 5 bases de datos sobre las que trabajaba. Sería interesante comprobar con los 62 electrodos presentes en la base de datos de Physionet si ciertamente F7, F8, C3 y C4 son los electrodos más óptimos para realizar la clasificación, o hay algún otro que pueda introducir más información relevante respecto a los 16 elegidos en la publicación original. Por otro lado, la red ha sido entrenada con señales que comparten el mismo *onset*. En este tipo de registros la información de la planificación motora presente en F7 y F8 puede estar sobrerrepresentada con respecto a eventos de MI, que no contengan un inicio compartido.

## 5. Conclusión

La herramienta XAI presentada para DL aplicado a EEG permite obtener los valores SHAP con correspondencia directa con la distribución espacial y temporal de un EEG. Gracias a estas representaciones, se observa como para EEGSym los instantes temporales más relevantes a la hora de realizar una clasificación de MI están concentrados en el primer segundo tras el *onset* de MI. Además, identifica

los canales frontales F7 y F8 como los más importantes para su decisión, seguidos de C3 y C4. Esta herramienta de XAI nos permite descubrir como las redes de DL se fijan en características que habían pasado desapercibidas en el uso tradicional de ML para la clasificación de MI.

## Agradecimientos

Este estudio ha sido financiado por los proyectos PID2020-115468RB-I00, RTC2019-007350-1 y PDC2021-120775-I00 del Ministerio de Ciencia e Innovación/Agencia Estatal de Investigación/10.13039/501100011033/ y FEDER Una forma de hacer Europa; por el proyecto I+D+i ‘Análisis y correlación entre la epigenética y la actividad cerebral para evaluar el riesgo de migraña crónica y episódica en mujeres’ (‘Programa de Cooperación Interreg V-A España-Portugal POCTEP 2014-2020’) y FEDER; y por ‘Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN)’ a través de ‘Instituto de Salud Carlos III’. Diego Marcos-Martínez, Eduardo Santamaría-Vázquez y Sergio Pérez-Velasco son beneficiarios de una ayuda PIF de la Consejería de Educación de la Junta de Castilla y León.

## Referencias

- [1] J. R. Wolpaw *et al.*, *Brain-Computer Interfaces: Principles and Practice*. Oxford University Press, 2012.
- [2] M. Sebastián-Romagosa *et al.*, “Brain Computer Interface Treatment for Motor Rehabilitation of Upper Extremity of Stroke Patients—A Feasibility Study,” *Front. Neurosci.*, vol. 14, no. October, pp. 1–12, 2020.
- [3] S. Perez-Velasco *et al.*, “EEGSym: Overcoming Inter-Subject Variability in Motor Imagery Based BCIs With Deep Learning,” *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 30, pp. 1766–1775, 2022.
- [4] I. Sturm *et al.*, “Interpretable deep neural networks for single-trial EEG classification,” *J. Neurosci. Methods*, vol. 274, pp. 141–145, 2016.
- [5] H. Choi *et al.*, “Non-human primate epidural ECoG analysis using explainable deep learning technology,” *J. Neural Eng.*, vol. 18, no. 6, p. 066022, Dec. 2021.
- [6] D. O. Nahmias *et al.*, “Easy Perturbation EEG Algorithm for Spectral Importance (easyPEASI),” in *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, Aug. 2020, pp. 2398–2406.
- [7] C. Ieracitano *et al.*, “A novel explainable machine learning approach for EEG-based brain-computer interface systems,” *Neural Comput. Appl.*, vol. 0123456789, no. DI, Mar. 2021.
- [8] H. Alsuradi *et al.*, “Explainable Classification of EEG Data for an Active Touch Task Using Shapley Values,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 12424 LNCS, 2020, pp. 406–416.
- [9] S. Lundberg *et al.*, “A Unified Approach to Interpreting Model Predictions,” *NIPS 2017*, vol. 32, no. 2, pp. 1208–1217, May 2017.
- [10] A. L. Goldberger *et al.*, “PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals,” *Circulation*, vol. 101, no. 23, Jun. 2000.
- [11] E. Beam *et al.*, “A data-driven framework for mapping domains of human neurobiology,” *Nat. Neurosci.*, vol. 24, no. 12, pp. 1733–1744, 2021.

# Diseño de una BCI mediante redes convolucionales profundas para la detección de la intención de parar ante un obstáculo inesperado

V. Quiles<sup>1</sup>, V. Ferrero<sup>1</sup>, E. Iáñez<sup>1,2</sup>, M. Ortiz<sup>1,2</sup>, J.M. Azorín<sup>1,2,3</sup>

<sup>1</sup> Brain-Machine Interface Systems Lab, Universidad Miguel Hernández de Elche, Elche, España {vquiles, lferrero, eianez, mortiz, jm.azorin}@umh.es

<sup>2</sup> Instituto de Investigación en Ingeniería de Elche - I3E, Universidad Miguel Hernández de Elche, Elche, España

<sup>3</sup> ValGRAI: Valencian Graduated School and Research Network of Artificial Intelligence, España

## Resumen

*Las tecnologías de asistencia aunadas con las tecnologías de registro de la actividad humana son una potente herramienta, tanto para la asistencia como para la rehabilitación de pacientes, los cuales tienen las funciones motoras dañadas. En este trabajo se diseñó una interfaz cerebro máquina (de las siglas en inglés, BCI), cuyo objetivo es detectar la intención del usuario de parar ante la aparición de un obstáculo inesperado. El objetivo de esta interfaz es que en un futuro pueda ser combinada con un dispositivo de asistencia, un exoesqueleto. Para conseguir crear una BCI sólida con pocas repeticiones se ha optado por clasificar la actividad cerebral mediante dos redes convoluciones que son parametrizadas con transfer learning de otros usuarios y optimizadas mediante técnicas de fine-tuning. Tras el registro de la actividad electroencefalográfica (EEG) de 5 usuarios sanos, la BCI consiguió acertar en una evaluación pseudo-online el 55,6% de las repeticiones sin ninguna falsa activación.*

## 1. Introducción

Las BCI permiten la comunicación directa entre humanos y dispositivos externos a través del análisis de señales neuronales [1]. Las BCI pueden ser utilizadas para tareas cognitivas y de comunicación de pacientes impedidos, pero también para establecer una conexión entre la intención y el control motor. De esta forma se puede mejorar la autonomía y calidad de vida de las personas con discapacidad motora severa. Discapacidades causadas por lesión de la médula espinal, accidente cerebrovascular, entre otros [2]. En la práctica, la mayoría de las BCI utilizan el EEG para registrar la actividad cerebral por su bajo costo relativo, facilidad de uso y por ser no-invasivos [2]. Esta técnica utiliza electrodos colocados sobre el cuero cabelludo para registrar la actividad eléctrica de las neuronas superficiales (principalmente las piramidales del cortex), reflejando los procesos cerebrales en curso con una resolución temporal excepcional [3]. No obstante, el EEG también tiene importantes inconvenientes. Por ejemplo, presenta baja resolución espacial [3]. Además, es muy sensible a una amplia gama de artefactos, como movimientos musculares, movimientos oculares o actividad cardíaca [3]. Por lo tanto, el EEG es un señal extremadamente compleja y ruidosa, lo que hace que la decodificación directa de los procesos cerebrales individuales sea realmente difícil de llevar a cabo. No obstante se puede utilizar una señal de control, en este caso dos potenciales serán mencionados: Los potenciales

visuales relacionados con eventos (ERP) son formas de onda que reflejan la actividad cerebral en curso que ocurre justo después de la percepción de un estímulo visual externo [4]. Otro potencial, es el potencial de error, el cual aparece frente a un evento inesperado o una respuesta errónea ante un estímulo [5]. Estos potenciales son bastante identificables en el promedio del EEG donde la variabilidad del single-trial desaparece. No obstante diseñar paradigmas de clasificación que de forma robusta que sean capaces de interpretar esta variabilidad sigue siendo un reto dentro del campo de las BCI [6].

Un gran número de estudios abordaron este problema mejorando algoritmos de reconocimiento de patrones para la detección de ERP [7]. Métodos basados en el análisis discriminante lineal (LDA) [9] y las máquinas de vectores de soporte son los enfoques más clásicos [10] logran un rendimiento razonable, especialmente cuando un gran cantidad de ejemplos de entrenamiento está disponible. Sin embargo, estos métodos son sensibles al ruido y muestran pobre rendimiento inter-sujeto y generalización entre sesiones [7].

El aprendizaje mediante redes profundas que ha sido aplicado en otros campos está ganando popularidad en el campo de las BCI. Algunos desafíos, como la falta de grandes conjuntos de datos han obstaculizado su desarrollo, no obstante, técnicas como el *transfer learning* y el *fine-tuning* [11] tratan de subsanar esta problemática.

El objetivo de este estudio es explorar como crear un modelo eficaz de BCI, basado en redes convoluciones y que tenga una tasa baja de falsas activaciones por minuto. Posteriormente analizamos la influencia del número de repeticiones y como crear protocolos y modelos de BCI que sean aplicables en pacientes (pocas repeticiones) y que su rendimiento sea suficientemente estable para su aplicación en tiempo real.

## 2. Material y Métodos

A continuación, se describe el equipo software y hardware utilizado en las pruebas, así como el protocolo experimental mantenido. Además, se detalla el método de preprocesamiento para el filtrado de la señal EEG, los diferentes métodos de extracción de características propuestos y el clasificador seleccionado. Por último, se

describe la estrategia seguida para el análisis pseudo-online.

### 2.1. Configuración experimental

En la Figura 1, se muestra el montaje experimental. Para el registro encefalográfico (EEG), se utilizó un gorro actiCAP (Brain Products GmbH, Alemania) con 27 electrodos. Los electrodos seguían la distribución 10-10 del sistema internacional Fz, FC5, FC1, FCz, FC2, FC6, C3, Cz, C4, CP5, CP1, CP2, CP6, P3, Pz, P4, PO7, PO3, PO4, PO8, FC3, FC4, C5, C1, C2, C6, CP3, CPZ, CP4, P1, P2 y POz. 4 electrodos fueron puestos para registrar EOG. La referencia se colocó en el lóbulo izquierdo y la tierra en el lóbulo derecho. La señal EEG fue registrada a 500Hz y amplificada mediante el equipo actiCHamp (Brain Products GmbH, Alemania).

Se empleó una cinta de andar Performance 750 de la marca Pro-form así como una línea láser proyectada en la parte delantera de la cinta de correr, con una longitud de onda de 635 nm (color rojo) y una potencia de salida de 3 mW.

Para detectar la parada del usuario, se han utilizado 3 unidades de medición inercial (IMUs) de la marca WIT Motion dispuestas de la siguiente manera: cabeza, pie izquierdo y pie derecho. Para cada sensor se han tenido en cuenta las aceleraciones en cada uno de los ejes: ‘Acc X’, ‘Acc Y’ y ‘Acc Z’.

La sincronización del EEG, IMUs y las activaciones del láser han sido controladas mediante una arquitectura customizada en Matlab.

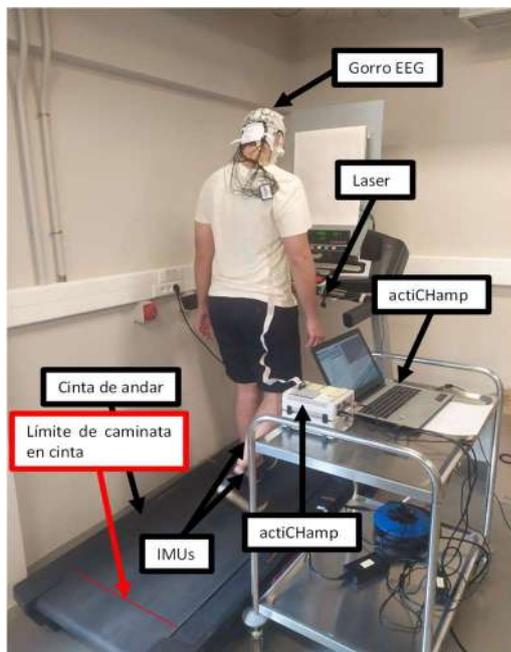


Figura 1. Montaje utilizado para llevar a cabo el protocolo de la experimentación.

### 2.2. Procedimiento experimental

Los registros se han llevado a cabo con 5 sujetos. Los registros se han llevado a cabo en 4 sujetos sanos, sin ningún desorden neurológico. Cada uno de ellos ha realizado entre 8 y 10 registros. Cada registro tiene una

duración de 120 segundos, durante los cuales se proyecta de manera variable la línea láser con una duración de encendido de 1 segundo y un tiempo entre dos estímulos sucesivos aleatorio comprendido entre 10 y 12 segundos. El número total de láseres por sesión es 10, pero el último fue eliminado por lo que el número total de láser es 9.

La prueba inicia con la cinta de correr parada, instante en el cual se procede al calibrado de las IMUs; los primeros 15 segundos son guardados para la convergencia del algoritmo de eliminación de artefactos oculares, a continuación, se activa la cinta y el sujeto comienza a andar a un ritmo prefijado constante de 2 km/h y una inclinación de 0 grados. Cuando la persona se encuentra andando de manera libre (sin cogerse a la máquina) y estable, se procede a iniciar el registro.

### 2.3. Etiqueta de la parada según las IMUs

Para etiquetar el momento en el que el usuario paró se utilizaron las dos IMUs posicionadas en cada uno de los pies. Se calculó el módulo de aceleración, esta señal se filtró mediante el algoritmo EMD y se seleccionaron las frecuencias que se asocian con la marcha desechando las altas y bajas frecuencias. Se seleccionó el instante del pie izquierdo o derecho dependiendo del último pico detectado antes del periodo de parada, ver Figura. 2.

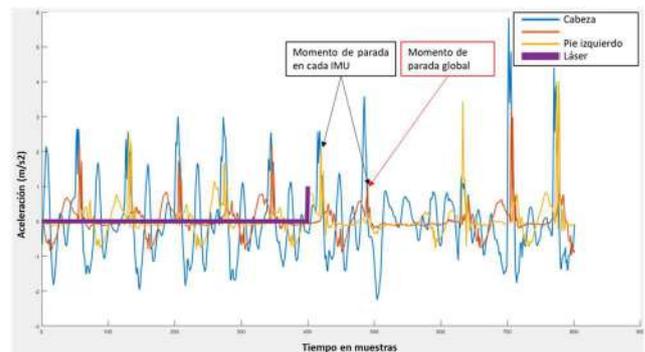


Figura 2. Señales pertenecientes al filtrado del módulo de aceleración de cada uno de las IMUs.

### 2.4. Preprocesamiento

A cada uno de los 27 canales de EEG se les aplicó el algoritmo de filtrado de artefactos oculares  $H_{\infty}$  [12].

Sobre la señal filtrada con  $H_{\infty}$  se aplican 4 filtros de orden 2, por variables de estados, las bandas escogidas fueron: 0.4-3 Hz, 2-4 Hz, 3-6 Hz, 5-8 Hz.

### 2.5. Segmentado de clases

Las dos clases son escogidas según al periodo que pertenezcan: antes de la aparición del láser, de la clase Andar o después de la aparición de él, etiquetadas como clase Láser. Por cada láser se escogieron 4 ventanas de señal de cada clase. Para la clase Andar la ventana temporal fue antes del láser, para la clase Láser después del Láser. La ventana es de 300 muestras, es decir, 0.6 segundos.

### 2.6. Clasificador

El primer clasificador es entrenado con ventanas de ambas clases mediante una red convolucional profunda. Los

parámetros de la red están especificados en la Tabla 1 y los parámetros de entrenamiento de la red están especificados en la Tabla 2.

	<i>Kernel</i>	Numero de filtros
Entrada		
conv	64x2	6
BN		
ReLu		
Max-Pool	1x2	6
conv	1x11	12
BN		
ReLU		
Max-Pool	1x2	12
conv	1x10	12
BN		
Relu		
Max-Pool	1x2	12
Dropout		
FC	60	60
ReLU		
FC	2	2
softmax		

**Tabla 1.** Parámetros de la red convolucional.

Parámetros	Valor
Optimización	<i>Stochastic Gradient Descent with momentum</i>
Impulso	0.9
Índice de aprendizaje	0.01
Regularización L2	0.0001
Factor de <i>Drop</i>	0.1
Periodo de <i>Drop</i>	10
Numero de épocas	500
Tamaño de lote	100

**Tabla 2.** Parámetros del entrenamiento de la red convolucional..

Se añade un segundo clasificador para disminuir las falsas activaciones, evaluando en que ventanas son las que más falla el primer clasificador cuando trata de predecir la clase Láser. El primer clasificador es evaluado con las mismas repeticiones con las que se creó el modelo, pero en modo pseudo-online, es decir no por ventanas, desde que empezó a andar el usuario hasta que paró por voluntad propia al ver el láser. La ventana tiene un tamaño de 300 muestras y se desplaza cada 0,1 segundos. Para este segundo modelo se

consideran dos clases análogas a Andar y Laser, se considera la primera clase como aquellas ventanas en las que el primer clasificador etiquetó la clase Andar como Laser y la segunda clase como aquellas ventanas en las que el primer clasificador acertó la clase Laser. Como el número de ventanas puede ser diferente para cada clase, se eliminan de forma aleatoria repeticiones de la clase que exceda en número de ventanas a la otra.

## 2.7. Validación cruzada Offline y Pseudo-Online

La validación cruzada se realiza siempre con 9 repeticiones de testeo y el primer modelo se crea con 4 ventanas de cada clase extraídas de cada una del resto de repeticiones, el segundo modelo se crea con la evaluación del primero en modo pseudo-online. Tanto el primero y el segundo se evalúan con las 4 ventanas de cada clase de las repeticiones de testeo.

Ambos modelos también se evalúan en modo pseudo-online (ventana móvil cada 0.1 segundos) con las repeticiones de testeo según la siguiente regla: si el primer clasificador etiqueta la ventana como clase Andar no se evalúa el segundo clasificador. Si el primer clasificador etiqueta la ventana como Laser, se evalúa en el segundo clasificador, este segundo clasificador puede etiquetar esta ventana como Andar o Laser. En el modo pseudo-online se evalúa la señal desde que empezó a andar (se estimaron 8 segundos antes del láser) hasta que la persona paró (etiquetado en las IMUs). Una moda de 5 ventanas se aplicó para disminuir los falsos positivos.

Se considera falso positivo (FP) si se etiquetó la ventana como Láser cuando en realidad esa ventana pertenecía a Andar y se considera acierto (de las siglas en inglés, TP) si se etiquetó como clase Laser y en realidad esa ventana pertenecía a esa clase. Las métricas para el modelo offline es el acierto total, es decir, el porcentaje de repeticiones acertadas, frente al total de repeticiones. En el modelo pseudo-online hay cuatro métricas: porcentaje de repeticiones en las que hubieron 0 FP y en la parte de Laser hubo al menos un TP (NOFP/TP), porcentaje de repeticiones en los que no hubo FP (NOFP) independientemente de si luego hubo TP, porcentaje de repeticiones que hubo (TP) independientemente de si hubo o no FP y por último el número de FP por minuto (FP/min).

## 2.8. Validación de modelos

En una prueba en tiempo real, el modelo siempre se testearía con las últimas repeticiones. Para evaluar como varía el rendimiento de la BCI dependiendo del número de repeticiones utilizadas para entrenar el modelo, se crearon BCIs con distinto número de repeticiones: “todas excepto las 9 últimas”, “40”, “20”, “10”.

Para subsanar este problema se aplica *transfer learning* y *fine-tuning*. La red se entrena con las repeticiones de todos los usuarios con “todas las repeticiones excepto las 9 últimas”, del usuario que se evalúa solo se introducen las “20 repeticiones anteriores a las 9 últimas” (TL+FT). El *fine-tuning* se realiza con estas 20 repeticiones y todos los pesos se vuelven a calcular.

### 3. Resultados y Discusión

De la validación cruzada el análisis offline tiene resultados muy altos, y se observa que el segundo clasificador tiene mayor variabilidad que el primero (Tabla 3). Los NOFP indican que un número alto de repeticiones no hubo FP, lo cual indica que los FP/min pueden estar acumulados en algunas repeticiones y no repartidos homogéneamente, este último es un indicador que puede falsear el rendimiento de la BCI, esto se evidencia más en la Tabla 4.

Además el porcentaje de NOFP/TP, indica que la interfaz puede tener un rendimiento aceptable detectando la parada sin generar FP en varias repeticiones.

	Offline		Pseudo-Online			
	C1	C2	NOFP/TP	NOFP	TP	FP/min
$\bar{x}$	90,4%	80,1%	67,9%	87,2%	77,4%	3,9
$\sigma$	4,3	10,8	15,9	10,6	14,4	3,1

**Tabla 3.** Promedio y desviación de las validaciones cruzadas de los 5 usuarios.

En la Tabla 4, se aprecia la tendencia en la disminución de los indicadores tanto NOFP/TP, NOFP y TP a medida que se disminuyen el número de repeticiones usadas para crear el modelo. Como solución a este problema el modelo de TL+FT creado con 20 repeticiones consigue mejorar al modelo de 20 repeticiones normal, pero además mejora el resto de modelos.

	Offline		Pseudo-Online			
	C1	C2	NOFP/TP	NOFP	TP	FP/min
Todos	87,8%	78,6%	44,4%	71,1%	64,4%	8,5
40,0	84,4%	66,4%	35,6%	55,6%	62,2%	11,7
20,0	76,1%	70,3%	28,9%	66,7%	40,0%	18,7
10,0	71,1%	61,7%	20,0%	53,3%	44,4%	29,8
TL+FT	<b>89,9%</b>	<b>79,0%</b>	<b>55,6%</b>	<b>77,8%</b>	<b>75,0%</b>	<b>6,0</b>

**Tabla 4.** Promedio de todos los usuarios para los distintos tipos de modelo. En negrita el mejor modelo.

### 4. Conclusión

En este trabajo se ha explorado un modelo de red convolucional para crear una BCI que mediante dos clasificadores consigue distinguir el patrón de marcha del potencial evocado generado en el EEG al decidir el usuario parar, debido a un obstáculo inesperado. Ha sido probado que el rendimiento de la BCI propuesta depende del número de repeticiones con las que se entrena. El modelo de *transfer learning* más *fine-tuning* ha sido capaz de subsanar este inconveniente. Mediante esta metodología la interfaz acierta con éxito y sin comandos erróneos en el 55,6% de las repeticiones. En este estudio se han mantenido resultados comparables con los reportados en [9] (3,9FP/min vs 2,6FP/min), pero las redes neuronales permiten generalizar la señal EEG entre sujetos (mientras que LDA no), lo cual facilita el entrenamiento del modelo (pocas repeticiones de ese usuario). En futuros trabajos se validará esta metodología en tiempo real, añadiendo un exoesqueleto como dispositivo de asistencia. La BCI deberá enviar el comando de parada al exoesqueleto cuando detecte la intención del usuario de parar.

### Agradecimientos

Esta publicación es parte del proyecto de I+D+i RTI2018-096677-B-I00, financiado/a por MCIN/AEI/10.13039/501100011033 y por “FEDER Una manera de hacer Europa”.

### Referencias

- [1] J. D. R. Millán *et al.*, “Combining brain-computer interfaces and assistive technologies: State-of-the-art and challenges,” *Frontiers in Neuroscience*, vol. 4, no. SEP. Frontiers Media SA, 2010.
- [2] K. Wada *et al.*, “Development of a Brain-machine Interface for Stroke Rehabilitation Using Event-related Desynchronization and Proprioceptive Feedback,” *Advanced Biomedical Engineering*, vol. 8, pp. 53–59, 2019.
- [3] M. Bamdad, H. Zarshenas, and M. A. Auais, “Application of BCI systems in neurorehabilitation: A scoping review,” *Disabil Rehabil Assist Technol*, vol. 10, no. 5, pp. 355–364, Sep. 2015.
- [4] R. Fazel-Rezai *et al.*, “P300 brain computer interface: Current challenges and emerging trends,” *Front Neuroeng*, vol. 0, no. JUNE, pp. 1–30, Jun. 2012.
- [5] G. Hughes and N. Yeung, “Dissociable correlates of response conflict and error awareness in error-related brain activity,” *Neuropsychologia*, vol. 49, no. 3, pp. 405–415, Feb. 2011.
- [6] B. Blankertz, C. Schäfer, G. Dornhege, and G. Curio, “Single trial detection of EEG error potentials: A tool for increasing BCI transmission rates,” *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 2415 LNCS, pp. 1137–1143, 2002.
- [7] F. Lotte, M. Congedo, A. Lécuyer, F. Lamarche, and B. Arnaldi, “A review of classification algorithms for EEG-based brain-computer interfaces,” *J Neural Eng*, vol. 4, no. 2, Jun. 2007.
- [9] M. Elvira, *et al.*, “Pseudo-Online BMI Based on EEG to Detect the Appearance of Sudden Obstacles during Walking,” *Sensors 2019, Vol. 19, Page 5444*, vol. 19, no. 24, p. 5444, Dec. 2019.
- [10] A. Rakotomamonjy and V. Guigue, “BCI competition III: Dataset II- ensemble of SVMs for BCI P300 speller,” *IEEE Trans Biomed Eng*, vol. 55, no. 3, pp. 1147–1154, Mar. 2008.
- [11] I. Goodfellow, Y. Beng, and A. Courville, *Deep Learning*. MA, USA: MIT Press, 2016.
- [12] A. Kilicarslan, R. G. Grossman, and J. L. Contreras-Vidal, “A robust adaptive denoising framework for real-time artifact removal in scalp EEG measurements,” *J Neural Eng*, vol. 13, no. 2, p. 026013, Apr. 2016.

# Evaluación de un videojuego multijugador basado en Brain Computer Interfaces utilizando c-VEPs

Selene Moreno-Calderón <sup>a</sup>, Víctor Martínez-Cagigal <sup>a,b</sup>, Eduardo Santamaría-Vázquez <sup>a,b</sup>, Sergio Pérez-Velasco <sup>a</sup>, Diego Marcos-Martínez <sup>a</sup> y Roberto Hornero <sup>a,b</sup>

<sup>a</sup>Grupo de Ingeniería Biomédica (GIB), E.T.S. Ingenieros de Telecomunicación, Universidad de Valladolid;

<sup>b</sup>Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN).

([selene.moreno@gib.tel.uva.es](mailto:selene.moreno@gib.tel.uva.es), [robhor@tel.uva.es](mailto:robhor@tel.uva.es)).

## Resumen

*Los videojuegos se han convertido en un elemento indispensable en el ámbito del entretenimiento. Estos tienen un gran potencial en los sistemas brain-computer interface (BCI), debido a la motivación que proporcionan. Para mejorar esta combinación, es conveniente centrarnos en la usabilidad y experiencia del usuario. Es por ello, que en este estudio se presenta la evaluación un videojuego BCI multijugador competitivo utilizando las propias señales cerebrales de los usuarios. El sistema ha sido probado en 22 sujetos de control. Se evaluó la eficiencia, eficacia y satisfacción mediante un análisis cuantitativo (precisión e ITR) y un análisis cualitativo (NASA-TLX y cuestionario de usabilidad). Tras la realización de las diferentes tareas se obtuvieron resultados preliminares los cuales mostraron que los usuarios no solo controlaron con éxito el sistema, alcanzando una precisión media de  $95.47\% \pm 9.01\%$  sino que también se mostraban satisfechos con su uso.*

## 1. Introducción

Los sistemas *brain-computer interface* (BCI) se definen como sistemas de comunicación que permiten al usuario interactuar con el medio a partir de la monitorización de su actividad cerebral, sin la intervención de músculos o nervios periféricos. Para interpretar las intenciones del usuario, en primer lugar, se registra su actividad cerebral mediante electroencefalografía, este suele ser el método más comúnmente usado debido su carácter no invasivo, su portabilidad y por ser de bajo coste [1].

Dado el gran avance en los últimos años del sector de los videojuegos, se estima que en la actualidad existen 18,1 millones de personas que juegan a videojuegos en España [2]. No obstante, la mayor cantidad de videojuegos están desarrollados de forma que se controlan mediante teclado, ratón o *joystick* dificultando así la accesibilidad de personas con grave discapacidad. Es por ello, que parece conveniente ahondar en la combinación videojuegos y sistemas BCI, con el objetivo de mejorar su calidad de vida e incrementar la independencia.

Hoy en día, el éxito o evaluación de sistemas BCI se suele centrar en la precisión y rendimiento. Sin embargo, aunque es un factor importante para la usabilidad del sistema, hay otros factores determinantes que también influyen y que puede afectar de manera positiva al funcionamiento general del sistema.

Es por ello, que es importante tener en cuenta un aspecto clave a la hora de trabajar con videojuegos BCI, la motivación [5] y no solo para sujetos sanos. Existen

estudios que sugieren que el juego competitivo mejora la atención y motivación de pacientes con ictus, siendo estos factores muy relevantes debido a las extensas terapias de rehabilitación. Por otro lado, otros experimentos hacen hincapié en el papel de evaluar la experiencia del usuario, así como la usabilidad y estabilidad de los sistemas BCI con la finalidad de mejorar los diseños y poder aumentar dicha motivación [6]. En resumen, la literatura científica sugiere que, a pesar de los avances en las capacidades de los sistemas BCI, existe una falta de usabilidad de estas por desarrolladores no expertos.

Un ejemplo de videojuego multijugador competitivo que se ha trasladado a sistemas BCI es el conocido juego “Conecta 4”, donde dos usuarios compiten para ganar con el objetivo de alinear 4 monedas del mismo color en horizontal, vertical o diagonal. Anteriormente, distintas versiones de este juego fueron llevadas a cabo en los estudios de Maby *et al.* [5] y Holz *et al.* [7] utilizando P300 y SMR, respectivamente. Las limitaciones de estos estudios recaen en los elevados tiempos de calibración y precisiones limitadas. Para suplir estas carencias, nos vamos a centrar en los potenciales evocados visuales modulados mediante código (c-VEPs), una señal de control caracterizada por alcanzar precisiones superiores con tiempos de calibración muy reducidos (e.g., >90% con 30s)

Tras realizar un análisis del estado de la técnica, no se ha encontrado ningún estudio que presente un videojuego BCI multijugador competitivo basado en c-VEPs, ni la evaluación de la usabilidad y cómo influye esta en el rendimiento del sistema. Por ello, en este estudio se ha tratado de proporcionar un videojuego BCI que cubra estas necesidades para los diferentes parámetros que evalúan la usabilidad y poder así aumentar dicha motivación. Por consiguiente, el objetivo de este trabajo es presentar la evaluación de la usabilidad de un videojuego multijugador utilizando c-VEPs así como la comparativa con estudios previos que utilizan distintas señales de control. La aplicación ha sido evaluada por 22 sujetos sanos durante una única sesión, donde realizaron tareas guiadas, tareas libres y tareas multijugador competitivas.

## 2. Sujetos y metodología

### 2.1. Sujetos

22 sujetos de control (28 años  $\pm$  2.60 años, 12 varones) participaron en los experimentos. Todos los sujetos dieron su consentimiento para participar y fueron informados previamente del objetivo del estudio.

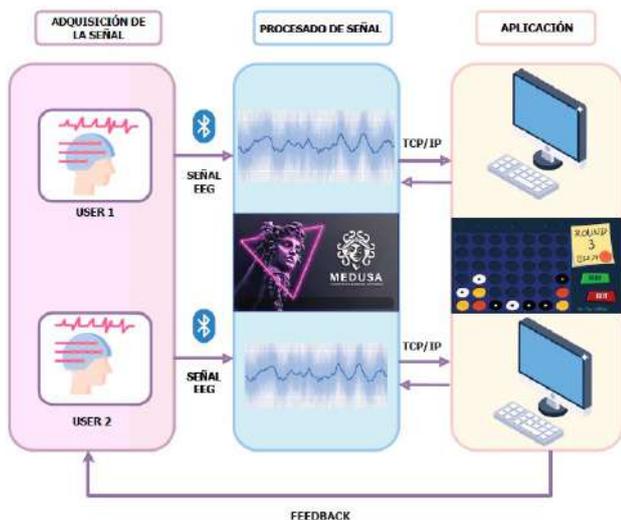


Figura 1. Arquitectura del sistema BCI.

## 2.2. Arquitectura del sistema BCI.

Tal y como muestra la Figura 1, el sistema BCI está formado por 3 etapas principales comunicadas entre sí: (1) adquisición del EEG; (2) procesado de la señal a través de MEDUSA [8], una plataforma de desarrollo de sistemas BCI ([www.medusabci.com](http://www.medusabci.com)); y (3) la aplicación, el videojuego desarrollado en Unity.

### 2.2.1 Adquisición

La señal EEG se registró mediante dos equipos g.Nautilus Pro (g.Tec, Guger Technologies, Austria), con una frecuencia de muestreo de 250 Hz. Esta señal se transmitía a través de Bluetooth al ordenador que ejecutaba la instancia del videojuego. Se ubicaron 8 electrodos en el cuero cabelludo en las posiciones Fz, Cz, Pz, P3, P4, PO7, PO8, Oz, con referencia el lóbulo de la oreja derecha (A2) y tierra sobre AFz.

### 2.2.2 Procesado de la señal

Los 2 jugadores deben colocar las monedas en una de las siete columnas del tablero. Para determinar en qué columna debe caer la moneda en cada turno, se hace uso del procesado de referencia para c-VEPs basado en paradigma circular, compuesto por un filtrado paso-banda y un análisis de correlación canónica para generar la plantilla principal para cada comando. Dicha plantilla representa la respuesta c-VEP del sujeto. La plantilla para el resto de los comandos se calcula mediante desplazamiento circular. Para más información, se puede consultar el estudio de Martínez-Cagigal et al. [6].

Los sujetos centran su atención en la celda del tablero donde desean situar la moneda, mientras cada una de las posiciones posibles se ilumina siguiendo una versión desfasada de una secuencia pseudo-aleatoria de máxima longitud ( $m$ -secuencia) de 63 bits. La frecuencia de actualización de pantalla es de 120 Hz, por lo que la duración de un ciclo completo de la secuencia corresponde a 0.525 s (i.e. 63/120) [6].

La matriz de selección está formada por 7 comandos, correspondientes a cada una de las columnas donde puede colocar la moneda. Los retardos se determinaron en

múltiplos de 9 muestras con el objetivo de distanciar lo máximo posible los retardos en la  $m$ -secuencia de 63 bits, facilitando su posterior decodificación.

La atención del usuario a dicho estímulo visual genera una respuesta cerebral específica que se ve reflejada en el EEG, permitiendo determinar la posición en la que el usuario quiere colocar la moneda mediante un procesado de señal basado en el paradigma de desplazamiento circular [6].

### 2.2.3 Aplicación

La aplicación consiste en una versión del famoso juego “Conecta 4” en la que la finalidad es alinear cuatro monedas del mismo color en horizontal, vertical o diagonal en un tablero de 6x7 celdas. Los dos jugadores colocan las monedas de manera alterna, en una de las siete columnas del tablero.

## 3. Protocolo experimental

El procedimiento para evaluar la aplicación se llevó a cabo por 22 sujetos de control mediante 7 tareas en una única sesión. Las tareas realizadas se dividieron en 3 bloques diferenciados: tareas guiadas, tareas libres y modo multijugador.

En primer lugar, de cara a tener un mayor número de selecciones y corroborar el correcto funcionamiento del sistema, un solo sujeto realizaba las tareas de manera individual, intentado seguir las figuras de la plantilla proporcionada. Se realizaron 4 tareas guiadas, incrementando el número de selecciones a realizar, donde el usuario debía colocar la moneda en función a la figura de la plantilla; y 3 tareas libres con un número de selecciones menor y con mayor autonomía. Por último, en el modo multijugador, dos sujetos competían entre sí en 3 partidas diferentes.

Previo a la realización de todas las tareas, se realizó un entrenamiento para determinar los patrones c-VEP del sujeto y poder identificar hacia donde miraba en cada momento. En dicho entrenamiento se le pidió que centrara su atención en la primera celda. Se registraron 10 *trials* con 10 ciclos, 100 ciclos en total (52.50 s).

Para calcular los parámetros establecidos en cuanto a la usabilidad: (1) eficiencia, se calculó la precisión para cada tarea de cada usuario y la tasa de información por minuto (ITR); (2) eficacia: se evaluó de manera cualitativa a través del cuestionario estandarizado NASA-TLX; y (3) satisfacción, estimada mediante un cuestionario SUS de 10 preguntas de aspecto positivo y negativo. El cuestionario puntuaba según la escala de Likert de 1 a 5 puntos. La última pregunta estaba enfocada a propuestas de mejoras de la aplicación.

## 4. Resultados y discusión

A raíz de los resultados obtenidos, se puede concluir que el sistema BCI propuesto funciona de manera favorable. Todos los sujetos respaldan estos resultados tanto de forma objetiva, mediante mediciones, como subjetiva, mediante cuestionarios (NASA-TLX y SUS).

SUJETO	TAREAS GUIADAS	TAREAS LIBRES	MULTIJUGADOR	TOTAL
S01	98.95 %	100.00 %	100.00 %	99.65 %±0.49 %
S02	96.24 %	94.10 %	95.23 %	95.19 %± 0.87 %
S03	91.20 %	94.10 %	90.50 %	91.93 %±1.55 %
S04	95.29 %	97.44 %	87.03 %	93.25 %±4.48 %
S05	100.00 %	100.00 %	85.00 %	95.00 %±7.07 %
S06	85.08 %	91.11 %	88.33 %	88.17 %±2.42 %
S07	92.87 %	94.87 %	91.53 %	93.09 %±1.37 %
S08	92.04 %	87.87 %	100.00 %	93.30 %±5.03 %
S09	97.06 %	100.00 %	100.00 %	99.02 %±1.38 %
S10	78.28 %	81.83 %	71.20 %	77.10 %±4.42 %
S11	100.00 %	100.00 %	100.00 %	100.00 % ±0.00 %
S12	100.00 %	100.00 %	95.23 %	98.41 %± 2.25 %
S13	100.00 %	100.00 %	86.57 %	95.52 %±6.33 %
S14	100.00 %	100.00 %	86.06 %	95.35 %± 6.57 %
S15	96.67 %	100.00 %	95.23 %	97.30 %±1.99 %
S16	95.76 %	96.67 %	85.94 %	92.79 %±4.85 %
S17	95.71 %	100.00 %	100.00 %	98.57 %±2.02 %
S18	82.45 %	92.64 %	100.00 %	91.69 %±7.19 %
S19	92.67 %	92.65 %	96.29 %	93.87 %±1.71 %
S20	92.67 %	93.26 %	96.66 %	97.32 %± 1.97 %
S21	100.00 %	100.00 %	77.81 %	92.60 %±10.46 %
S22	98.61 %	100.00 %	82.14 %	93.58 %± 8.11 %
MEDIA+STD	94.30 %±6.30 %	95.60 %±8.30 %	92.8 %±7.14 %	95.47 %±9.01 %
ITR+STD	27.58 bpm± 4.25 bpm	28.83 bpm±3.79 bpm	25.53 bpm± 5.40 bpm	27.07 bpm±5.32 bpm

**Tabla 1.** Resultados del análisis cuantitativo (precisión e ITR).

**4.1. Eficiencia**

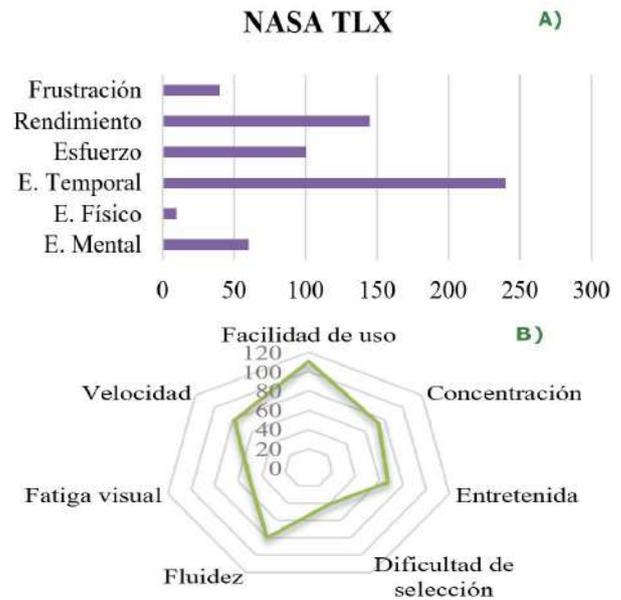
En la Tabla 1 se muestran las precisiones medias de cada sujeto para el conjunto de cada tipo de tarea, así como el total de todas estas y la ITR. La precisión media alcanzada para todas las tareas y de todos los sujetos fue de 95.47% ± 9.01%. Todos los sujetos fueron capaces de completar todas las tareas, y cabe resaltar que hubo sujetos que no tuvieron errores durante la elección de los comandos en diferentes tareas.

Por otro lado, se puede observar cómo existe cierto compromiso entre la precisión y la ITR. La precisión fue elevada para todas las tareas, siendo levemente inferior en el modo multijugador, esta diferencia se puede deber a la presión que genera competir contra otra persona, o al no tener tanta determinación para situar la moneda ya que solo se tiene 5,25 segundos para seleccionar la mejor opción y contrarrestar el movimiento del adversario. En cambio, la ITR máxima fue de 28,83 bpm para las tareas libres, esta se ve limitada puesto que, según la fórmula, esta depende directamente del número de comandos totales, establecido en este caso en 7, según el número de columnas donde el usuario podía situar la moneda.

**4.2. Eficacia**

En la Figura 2a), se muestran las puntuaciones medias obtenidas en el cuestionario NASA-TLX para todos los sujetos, donde se puede observar la importancia de cada uno de los factores como causantes de esa carga de trabajo y su importancia en cada tarea. La exigencia temporal es predominante. Las valoraciones de frustración, esfuerzo mental y esfuerzo físico son notablemente bajas. La puntuación global fue de 28.82 puntos, correspondiente a la media ponderada de las puntuaciones de los 6 parámetros.

El parámetro que más afectó al experimento fue la exigencia temporal, es decir, el usuario sentía cierta presión de tiempo debido al ritmo de selección de comandos. Hubo



**Figura 2.** Resultados del análisis cualitativo del videojuego BCI. A) Resultados del cuestionario NASA-TLX. B) Resultados del cuestionario SUS.

cierta discordancia en este parámetro, por otro lado, algunos usuarios tenían seguridad e indicaron que con menor tiempo también podrían haber realizado la selección, por lo que el tiempo establecido no les parecía el más adecuado.

En general, los sujetos destacaron que no tuvieron demasiada fatiga, por lo que esto permitía utilizar la aplicación durante un largo periodo de tiempo sin tener cansancio visual. Por otro lado, los usuarios se sentían seguros y relajados a medida que iban realizando las selecciones e indicaron que las tareas propuestas fueron fáciles y sencillas, y que se sentían satisfechos con el nivel de ejecución de estas.

**4.3. Satisfacción**

En la Figura 2b) se muestran los resultados cualitativos del cuestionario SUS proporcionado a los sujetos, reflejando que todos los sujetos fueron considerados satisfechos con la aplicación. Los sujetos indicaron que la aplicación desarrollada era fácil de usar y respondía con velocidad y fluidez. Se alcanzó un valor total de 80.60 sobre 100.

El aspecto mejor valorado fue el de interfaz de usuario, debido a ser intuitiva y accesible. Algunos sujetos indicaron que tuvieron dificultad para seleccionar comandos, esto estaba correlacionado con aquellos que alcanzaron una precisión levemente inferior. En contraposición, el aspecto peor valorado fue el tiempo de selección de comandos. De la misma manera, los resultados son consistentes entre los sujetos, lo que indica que dichas conclusiones están correctamente fundamentadas.

Como sugerencias de mejora se indicó incrementar el espacio entre iluminaciones para evitar posibles distracciones y ajustar el tiempo de selección a cada usuario, algunos sujetos necesitaban más tiempo para

pensar dónde querían colocar la moneda y otros necesitaban menos tiempo para concentrarse.

#### 5.4 Comparativa con otros estudios

A diferencia de los otros estudios BCI que desarrollaron un videojuego multijugador de tipo “Conecta 4” [5, 7] este alcanza precisiones muy elevadas a la par de valoraciones positivas en cuanto a la satisfacción.

En Maby *et al.* [7], se presentó una versión BCI del juego “Conecta 4” utilizando potenciales evocados P300. La evaluación experimental se realizó con 2 sujetos de control, obteniendo una precisión media del 83,30% e ITR de 37,00 bpm. Por otro lado, en Holz *et al.* [5] se desarrolló el videojuego “Connect 4” utilizando los SMR elicitados mediante imaginación motora de manos y pie. Se evaluó la aplicación en 4 usuarios con restricciones motoras severas. Se obtuvo una precisión media de 62,65% con ITR media de 0,53 bpm. Al constar de varias sesiones la carga de trabajo fue moderada, aunque la frustración contribuyó a la fatiga de algunos sujetos

Como se puede observar, la precisión de nuestro estudio ( $95.47\% \pm 9.01\%$ ) fue superior frente al 83,30% en Maby *et al.* [7] y 62,25% en Holz *et al.* [5]. Estos resultados sugieren que el tipo de señal más apropiada para sistemas de control es c-VEPs. En nuestro estudio se necesitó 52.50 segundos para calibrar el sistema BCI. Sin embargo, en el caso del entrenamiento para P300 se suele tardar entre 20 y 30 minutos, tiempo que se incrementa para el caso de SMR pudiendo durar horas o días, e incluso hay personas que nunca llegan a generarlos de manera satisfactoria.

Por otro lado, los videojuegos tienen un gran potencial para su uso en sistemas BCI, puesto que tienen como finalidad entretener y motivar a los usuarios. La motivación juega un papel relevante en el éxito de la interacción BCI. El número de sesiones se relaciona con la motivación y el estado del ánimo. En nuestro estudio y en Maby *et al.* [7] los sujetos pudieron alcanzar una mayor precisión después de una sola sesión, mientras que en Holz *et al.* [5] ningún usuario pudo usar SMR después de una sola sesión, teniendo que realizar 9 sesiones. A pesar de ello, no pudieron alcanzar una precisión del 70%, considerado el mínimo para obtener el control de un sistema BCI [6].

#### 5.5 Limitaciones y líneas futuras

A pesar de los resultados favorables y de la consecución con éxito del objetivo, este estudio no está exento de limitaciones. Si bien se ha demostrado la usabilidad del videojuego BCI desarrollado con sujetos de control, creemos que sería necesario evaluarlo también en personas con grave discapacidad motora, puesto que tradicionalmente han sido el objetivo de los sistemas BCI.

Por otro lado, como el parámetro con aspecto negativo fue la velocidad de selección, para superar esta limitación se contempla implementar técnicas de “early stopping”. Este algoritmo detectaría de forma dinámica el número de ciclos necesarios para emitir una selección, con lo cual no se

necesitaría esperar a que se mostraran los 10 ciclos establecidos, sino que se adaptaría a cada usuario.

## 5. Conclusiones

Este estudio se ha centrado en evaluar un videojuego BCI multijugador basado en c-VEPs en 22 sujetos de control. Para ello, se han llevado a cabo distintas tareas. Además, este estudio presenta ventajas frente a otros videojuegos BCI “Connect 4” anteriores. El videojuego obtuvo una eficiencia, eficacia y satisfacción elevada, justificada mediante el análisis cuantitativo (se alcanzó una precisión del  $95.47\% \pm 9.01\%$ ) y cualitativo correspondiente (resultados de los cuestionarios muy favorables).

## Agradecimientos

Este trabajo ha sido financiado por los proyectos RTC2019-007350-1, PID2020-115468RB-I00 y TED2021-129915B-I00 del Ministerio de Ciencia, Innovación/Agencia Estatal de Investigación y European Regional Development Fund (ERDF), así como por CIBER-BBN a través del Instituto de Salud Carlos III.

## Referencias

- [1] Wolpaw J. R., Birbaumer N., McFarland D. J., G. Pfurtscheller, and Vaughan T. M., (2002) “Brain-computer interfaces for communication and control.” *Clin. Neurophysiol.*, vol. 113, no. 6, pp. 767–91.
- [2] AEVI, Anuario de la industria del videojuego en España, 2021, pp 22.
- [3] Gürkök, Hayrettin; Nijholt, Anton; Poel, Mannes; Obbink, Michel, (2013). “Evaluating a multi-player brain-computer interface game: Challenge versus co-experience”. Entertainment Computing
- [4] Holz, E., Höhne, J., Staiger-Sälzer, P., Tangemann, M., y Kübler, A. (2013). “Brain-computer interface controlled gaming: Evaluation of usability by severely motor restricted end-users.” *Artificial Intelligence in Medicine*, 59(2), 111–120. g, 4(3), 195–203.
- [5] Holz, E., Höhne, J., Staiger-Sälzer, P., Tangemann, M., y Kübler, A. (2013). “Brain-computer interface controlled gaming: Evaluation of usability by severely motor restricted end-users.” *Artificial Intelligence in Medicine*, 59(2), 111–120.
- [6] Martínez-Cagigal V, Thielen J, Santamaría-Vázquez E, Pérez-Velasco S, Desain P, Hornero R., (2021) “Brain-computer interfaces based on code-modulated visual evoked potentials (c-VEP): a literature review.” *J Neural Eng.* Nov 26;18(6).
- [7] Maby, E., Perrin, M., Bertrand, O., Sanchez, G., y Mattout, J. (2012). “BCI Could Make Old Two-Player Games Even More Fun: A Proof of Concept with “Connect Four.”.” *Advances in Human-Computer Interaction*, 1–8.
- [8] Santamaría-Vázquez, E., Martínez-Cagigal V., and Hornero R., (2018) “MEDUSA: Una Nueva Herramienta Para El Desarrollo de Sistemas Brain-Computer Interface Basada En Python.” *Cognitive Area Networks* 5 (1): 87–92.
- [9] Martínez-Cagigal, V., Gómez-Pilar, J., Alvarez, D., & Hornero, R., (2016). “An asynchronous P300-based brain-computer interface web browser for severely disabled people.” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 25(8), 1332-1342.

# Rehabilitación de la marcha asistida por un sistema híbrido personalizable en dos casos de lesión medular incompleta

D. Herrera-Valenzuela<sup>1,2</sup>, J. Gil-Castillo<sup>3</sup>, J. Pina-de-Paz<sup>3</sup>, Á. Megía-García<sup>1</sup>, A. Gil-Agudo<sup>1</sup>, J.C. Moreno<sup>3</sup>, A.J. del-Ama<sup>4</sup>

<sup>1</sup> Unidad de Biomecánica y ayudas técnicas, Hospital Nacional de Paraplégicos, Toledo, España,

{dherrera, amegiag}@externas.sescam.jccm.es, amgila@sescam.jccm.es

<sup>2</sup> Escuela Internacional de Doctorado, Universidad Rey Juan Carlos, Madrid, España, ds.herrera.2020@alumnos.urjc.es

<sup>3</sup> Grupo de Neuro-rehabilitación, Instituto-Cajal, Consejo Superior de Investigaciones Científicas, Madrid, España,

{jgil, javier.pina}@cajal.csic.es, jc.moreno@csic.es

<sup>4</sup> Área de Tecnología Electrónica, Universidad Rey Juan Carlos, Madrid, España, antonio.delama@urjc.es

## Resumen

*Hoy en día la eficacia y aceptación de los exoesqueletos para la asistencia de la marcha es controvertida. Una de las razones es que son herramientas genéricas que se aplican a varias condiciones neurológicas y afectaciones de la marcha. Su combinación con estimulación eléctrica funcional en forma de un exoesqueleto híbrido permite combinar las ventajas de ambas tecnologías mitigando las respectivas desventajas, permitiendo adaptar el sistema y su asistencia al paciente. En este trabajo presentamos dos casos de estudio sobre la validación tecnológica de una neuroprótesis modular configurable con capacidad de hibridación con un exoesqueleto de miembro inferior. La adaptación de la asistencia permite obtener mejoras en la función de marcha en casos con diferentes necesidades funcionales. En particular, nuestro sistema permite configurar y asistir únicamente los grupos musculares y en las fases de la marcha según el patrón de marcha de cada individuo. En ambos casos se obtuvieron mejoras en el patrón de marcha, especialmente en la segunda sesión. La asistencia híbrida parece aportar mayores beneficios inmediatos a las lesiones medulares que tienen mayor grado de afectación de la marcha.*

## 1. Introducción

Las lesiones medulares tienen una incidencia anual que varía a nivel mundial entre 40 y 80 casos por millón de habitantes [1]. En España, la cifra está alrededor de los 30 casos por millón [2]. En estas lesiones, una de las consecuencias más comunes es la pérdida o afectación de la marcha, y también uno de los principales objetivos de rehabilitación [3]. Esta se ve afectada en distintos grados según el nivel y la severidad de la lesión, dando a lugar a una alta variabilidad de patrones de marcha dentro de la lesión medular [4]. Es por ello que, para lograr una rehabilitación eficiente, esta debe ser adaptada según las características particulares de cada lesión [5].

En las últimas décadas se ha visto un interés creciente en la rehabilitación de la marcha asistida por robots, ya que, el uso de exoesqueletos portables brinda más posibilidades para rehabilitación respecto a las herramientas tradicionales [6, 7], a la vez que permite obtener resultados equivalentes a los de la rehabilitación tradicional [8] y reducir las complicaciones derivadas de la sedestación prolongada [9]. A día de hoy, el desarrollo de estas tecnologías ha permitido que más de una docena de exoesqueletos portables de miembro inferior lleguen al mercado, sin embargo, su eficacia, aceptación y penetración en el mercado han sido limitadas porque son herramientas genéricas que buscan solucionar las necesidades de varias condiciones neurológicas. En otras palabras, carecen de un grado suficiente de personalización que les permita satisfacer los requerimientos individuales de cada paciente.

Una forma de potenciar la rehabilitación de la marcha asistida por robots es combinándola con neuroprótesis motoras (NP) para la aplicación de estimulación eléctrica funcional (FES). Estos sistemas han demostrado promover la rehabilitación de

la marcha en lesiones medulares incompletas, gracias a que fomentan la neuroplasticidad durante el periodo agudo y subagudo tras una lesión medular, permitiendo el reaprendizaje motor [10], así como beneficios secundarios derivados de la activación artificial de la musculatura [11]. Sin embargo, la aparición de fatiga muscular y la respuesta no lineal de la musculatura dificultan la utilización del FES para la asistencia de la marcha. La hibridación con un exoesqueleto permite la compensación de la fatiga muscular y mejorar el control del movimiento, aumentando tanto el tiempo de utilización como la calidad del movimiento generado [12].

En este trabajo presentamos dos casos de estudio sobre la validación tecnológica de una NP modular configurable con capacidad de hibridación con un exoesqueleto de miembro inferior. La hipótesis de este estudio es que la adaptación del sistema híbrido modular a las necesidades funcionales diversas de los pacientes, proporcionarán mejoras en la función de marcha. Para ello, se muestran los cambios en la cinemática de la marcha de dos pacientes con lesiones medulares incompletas de diferentes niveles y severidad, en dos sesiones realizadas con cada uno durante el periodo subagudo de rehabilitación intensiva durante su ingreso en un hospital de rehabilitación.

## 2. Metodología

El sistema híbrido utilizado está compuesto por una NP creada en el proyecto, compuesta de hasta cuatro módulos de estimulación independientes, con cuatro canales independientes cada uno, y una versión modificada del exoesqueleto ABLE (Able Human Motion S.L., Barcelona, Spain). Este último tiene actuadores únicamente en las rodillas, y articulaciones pasivas para cadera y tobillo, cuyo rango de movimiento está limitado por topes mecánicos dentro de rangos anatómicos. Tiene valores graduables para velocidad de transición de sedestación a bipedestación, ángulo de flexión de rodilla durante la oscilación, ratio flexión/extensión de la rodilla durante la oscilación y sensibilidad para la detección de la intención de paso. El exoesqueleto se usa con un andador adaptado a las dimensiones del robot. Por su parte, la NP asiste únicamente los grupos musculares necesarios en función de la fase de la marcha en la que se encuentra el individuo, gracias a la detección automatizada de eventos que se realiza con seis goniómetros Biometrics (Biometrics LTD, Newport, UK) ubicados en la cadera, rodilla y tobillo de cada pierna. Así pues, la personalización del sistema puede hacerse incluyendo solo el/los módulo(s) y canales necesarios de la NP, graduando los parámetros del exoesqueleto, y finalmente, utilizando la configuración que responda a las necesidades del paciente: sistema híbrido, solo NP o sólo exoesqueleto.

La configuración del sistema para cada caso se seleccionó según el grado de afectación de la marcha de cada paciente, la severidad y nivel neurológico de la lesión, el balance muscular

en miembros inferiores y una exploración visual sobre la marcha no asistida del sujeto, realizada por un fisioterapeuta. Para configurar la intensidad de la estimulación, se calibró cada canal por encima del umbral motor, siempre respetando el umbral de confort como límite máximo.

En este estudio se presentan los casos de dos pacientes con lesión medular, ingresados en el Hospital Nacional de Paraplégicos (HNP) para rehabilitación intensiva. En ambos sujetos, se realizó una primera sesión de familiarización y entrenamiento en el uso del exoesqueleto y del sistema híbrido, y 9 días después se realizó un estudio tridimensional de la marcha (3DGA) de los sujetos sin asistencia y con la configuración personalizada seleccionada para cada uno (sistema híbrido, solo NP o sólo exoesqueleto). En ambas sesiones, el registro de la cinemática de la marcha asistida con nuestro sistema se realizó con los goniómetros. En la segunda sesión, el registro de la marcha sin asistencia se realizó con un sistema de captura de movimiento VICON con 8 cámaras.

El primer caso es el de un hombre de 32 años con diagnóstico de lesión medular tipo D, de nivel L1, que al momento de la primera sesión llevaba 2 meses y 11 días tras una lesión medular de origen traumático. En la primera sesión, el sujeto tenía un WISCI II preferente de 16 (utilizaba dos bastones para desplazamientos de hasta 10 metros) y un WISCI II máximo de 20 (podía realizar desplazamientos de hasta 10 metros sin asistencia). En la segunda sesión, el sujeto tenía un WISCI II preferente de 19 (apoyo de un bastón para desplazamientos de 10 m) y un WISCI II máximo de 20.

El segundo caso corresponde a un hombre de 51 años con diagnóstico de lesión medular tipo C, de nivel T4, de etiología no traumática (cirugía de una hernia discal), ocurrida 1 mes y 27 días antes de realizar la primera sesión. En la primera sesión el sujeto tenía un WISCI II máximo de 5 (estaba comenzando su rehabilitación de la marcha en barras paralelas) y para la segunda sesión había aumentado a un WISCI II máximo de 8 (comenzaba su rehabilitación de la marcha con andador y asistencia de una persona).

Para el análisis de datos, se muestra el promedio de las curvas cinemáticas de las articulaciones de cadera, rodilla y tobillo para al menos tres ciclos de marcha de cada sujeto, normalizadas según el ciclo de marcha de los miembros inferiores derecho e izquierdo, de forma independiente. Se contrastan las curvas sin asistencia (registradas únicamente en la segunda sesión), con asistencia, y se presenta la zona de normalidad, obtenida de la base de datos de marcha de controles adultos del HNP. Adicionalmente, se muestra el *Movement Analysis Profile* (MAP) con los valores del *Gait Profile Score* (GPS) [13] para las mismas articulaciones en los escenarios con y sin asistencia, y el de los sujetos control, para tener valores cuantitativos específicos que reflejan la raíz del error cuadrático medio (RMSE) de las curvas cinemáticas. Finalmente, se presenta el valor de esfuerzo percibido por cada paciente al final de ambas sesiones usando la escala Modificada de Borg [14].

### 3. Resultados

#### Primer sujeto:

Durante la primera sesión, el primer sujeto recibió estimulación bilateral de 16 mA en los sóleos. Primero realizó caminata durante 6 minutos con la neuroprótesis, luego utilizó el exoesqueleto durante 14.5 min, tiempo en el cual consiguió realizar secuencias de varios pasos seguidos en los que el

exoesqueleto detectaba correctamente su intención de paso, y finalmente caminó durante 35.3 min con el sistema híbrido. Al final de la sesión el sujeto calificó el esfuerzo percibido con un 2 en la escala de Borg. Las figuras 1 y 2 muestran las curvas cinemáticas de esta sesión y el MAP, respectivamente.

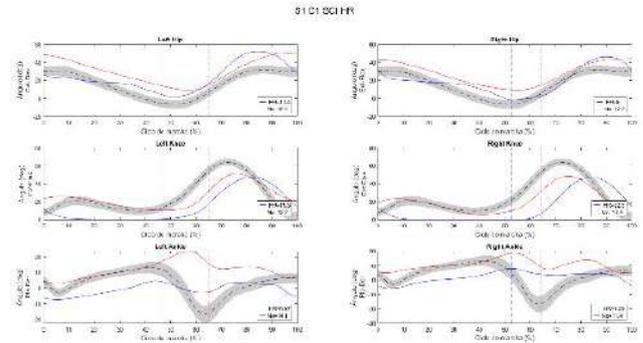


Figura 1. Cinemática de la sesión 1 del sujeto 1. Registro con asistencia híbrida (azul), marcha sin asistencia (roja), y promedio de normalidad (gris).

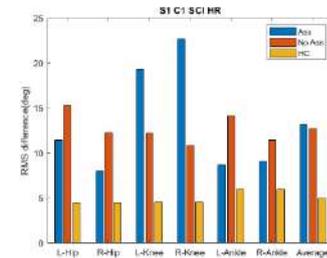


Figura 2. MAP del sujeto 1 la sesión 1. Asistencia híbrida (azul), sin asistencia (rojo), y sujetos control (amarillo).

En la segunda sesión, el sujeto recibió asistencia únicamente de la NP, ya que el uso del exoesqueleto fue considerado innecesario para el grado de afectación leve de la marcha del sujeto al momento del estudio. Primero realizó el 3DGA sin asistencia, en el que caminó durante 9.5 min, y posteriormente caminó 18.5 min con estimulación bilateral de 21 mA en los sóleos y gastrocnemios. Al final de la sesión el sujeto calificó el esfuerzo percibido con un 5 en la escala de Borg. En la parte izquierda de la figura 7 se puede ver al paciente con la NP. En las figuras 3 y 4 muestran las curvas cinemáticas de esta sesión y el MAP, respectivamente.

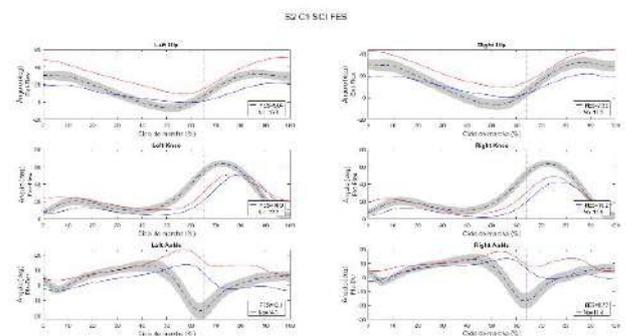
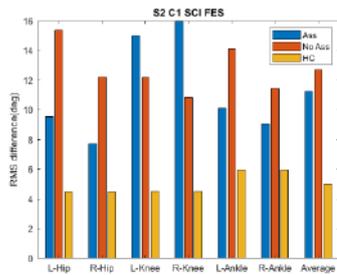


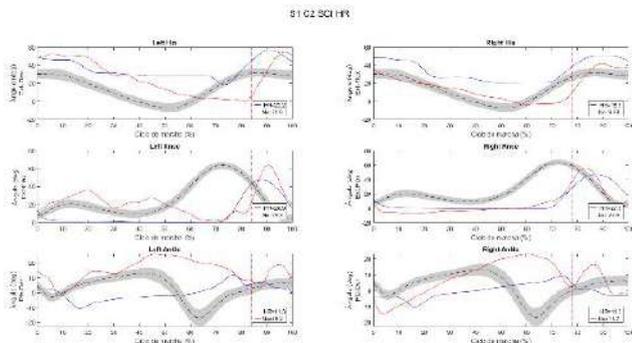
Figura 3. Cinemática de la sesión 2 del sujeto 1. Registro con asistencia híbrida (azul), marcha sin asistencia (roja), y promedio de normalidad (gris).



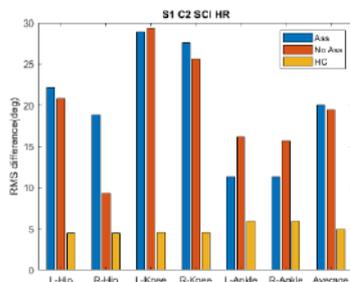
**Figura 4.** (Izquierda) MAP del sujeto 1 en la sesión 2. Asistencia híbrida (azul), sin asistencia (rojo), y sujetos control (amarillo).

Segundo sujeto:

Dado que al momento de realizar la primera sesión el sujeto no presentaba capacidad para caminar sin asistencia externa y sin soporte, se descartó el uso aislado de NP y solo se entrenó marcha con exoesqueleto y con asistencia híbrida en esta sesión. El sujeto caminó 29.3 minutos con el exoesqueleto, hasta ser capaz de realizar secuencias de pasos consecutivos con fluidez, donde el robot detectara la intención de paso, y posteriormente caminó 8.1 min con el sistema híbrido, recibiendo estimulación en los glúteos, de 32mA en el lado izquierdo y 40mA en el derecho, y en los músculos isquiotibiales de 24mA bilateral. El uso del sistema híbrido representa un aumento de su WISCI II máximo a 6 (camina 10 metros con caminador, ortesis y asistencia de una persona). Al final de la sesión, el sujeto calificó el esfuerzo percibido con un 3 en la escala de Borg. Las figuras 5 y 6 muestran las curvas cinemáticas de esta sesión y el MAP, respectivamente.



**Figura 5.** Cinemática de la sesión 1 del sujeto 2. Registro con asistencia híbrida (azul), marcha sin asistencia (roja), y promedio de normalidad (gris).



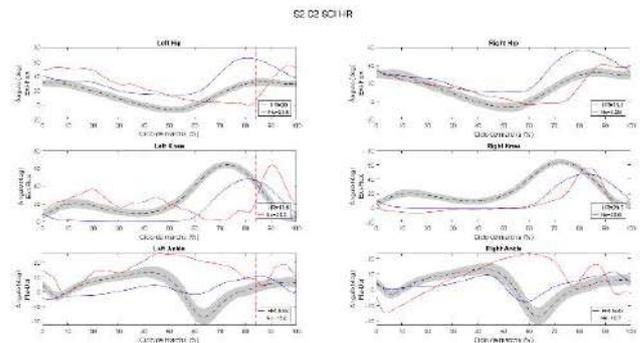
**Figura 6.** MAP del sujeto 2 en la sesión 1. Asistencia híbrida (azul), sin asistencia (rojo), y sujetos control (amarillo).

Para la segunda sesión, el segundo sujeto fue asistido únicamente con el sistema híbrido. Primero realizó el 3DGA

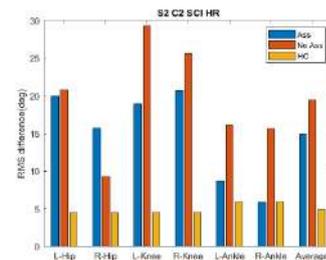
sin asistencia de nuestro sistema, en el que caminó durante 9.3 min con su andador convencional, y posteriormente caminó 59.6 min con el sistema híbrido, recibiendo estimulación en los tibiales anteriores, de 35 mA en el derecho, y 32 mA en el izquierdo, y en los isquiotibiales izquierdos de 47 mA y derechos de 50 mA. Al final de la sesión el sujeto calificó el esfuerzo percibido con un 3 en la escala de Borg. La figura 7 muestra al segundo sujeto usando el sistema híbrido, y las figuras 8 y 9 muestran las curvas cinemáticas de esta sesión y el MAP, respectivamente.



**Figura 7.** Sujetos en la segunda sesión usando nuestro sistema: a la izquierda, el primer sujeto asistido con la NP; a la derecha, el segundo sujeto asistido por el sistema híbrido.



**Figura 8.** Cinemática de la sesión 2 del sujeto 2. Registro con asistencia híbrida (azul), marcha sin asistencia (roja), y promedio de normalidad (gris).



**Figura 9.** MAP del sujeto 2 la sesión 2. Asistencia híbrida (azul), sin asistencia (rojo), y sujetos control (amarillo).

#### 4. Discusión

Los dos casos analizados presentan lesiones de nivel y alcance distintas (L1 ASIA D, T4 ASIA C), cuyas capacidades funcionales eran bastante diferentes (WISCI 20 y 5 respectivamente), por tanto, diferente afectación de marcha. En ambos casos el sistema fue adaptado y configurado para brindar niveles de asistencia específicos, logrando mejoras en la función de marcha. La capacidad de adaptación de la NP y la posibilidad de hibridarse con el exoesqueleto permitió configurar y asistir únicamente los grupos musculares que

requería cada sujeto y en las fases de la marcha que precisaban asistencia, según el patrón de marcha de cada individuo.

El índice GPS promedio disminuyó hacia valores más cercanos al patrón de marcha no patológicos, especialmente en la segunda sesión. Si bien el MAP de la primera sesión con asistencia presenta valores del GPS ligeramente mayores que sin asistencia, debe tenerse en cuenta que la sesión de captura de marcha sin asistencia se realizó en ambos casos en la segunda sesión, tras 9 días de la primera sesión. Estos días dieron lugar a una mejora funcional de los pacientes representado por un incremento en la escala WISCI-II de tres niveles en ambos casos. Por tanto, el sistema produjo mejoras funcionales inmediatas parecidas a las que se lograron con 9 días de rehabilitación. Mas allá de esta aproximación para cuantificar la mejoría en el patrón de marcha, el resultado importante es que el sistema y su adaptación al paciente, lograron mejoras en la función de marcha en dos pacientes con afectaciones distintas y en dos momentos distintos del proceso de rehabilitación para cada uno de ellos.

Respecto de la asistencia con NP personalizada frente al sistema híbrido, también personalizado, la asistencia híbrida parece aportar mayores beneficios inmediatos a las lesiones medulares que tienen mayor grado de afectación de la marcha, como es el caso del segundo sujeto.

La mejora del patrón de marcha usando exoesqueletos robóticos y/o neuroprótesis ha sido demostrada hace años. En este trabajo mostramos un sistema adaptable que mejora la asistencia en dos casos con requerimientos diferentes. Sin embargo, la comparación directa de estas mejoras con la literatura es complicada, dado que precisamente los déficits en la marcha de los pacientes son muy heterogéneos. Una manera de analizar la ventaja de nuestra metodología sería comparar el uso del sistema completo frente a la optimización, lo que ha quedado fuera de nuestro alcance.

Si bien se trata de dos casos de estudio, la metodología de personalización y las prestaciones de los sistemas no pueden ser generalizados todavía. Actualmente se está aumentando la muestra de este estudio.

## 5. Conclusiones

La adaptación de la asistencia con NP permite obtener mejoras en la función de marcha en casos con diferentes necesidades funcionales. En particular, nuestro sistema permite configurar y asistir únicamente los grupos musculares y en las fases de la marcha según el patrón de marcha de cada individuo. En ambos casos se obtuvieron mejoras en el patrón de marcha, especialmente en la segunda sesión. La asistencia híbrida parece aportar mayores beneficios inmediatos a las lesiones medulares que tienen mayor grado de afectación de la marcha.

## Agradecimientos

Este trabajo se ha realizado gracias a los proyectos financiados por la Agencia Estatal de Investigación: *Innovative robotic and sensing systems implementing robot-assisted walking therapies based on reinforcement-learning paradigms - NIMBLE* (Ref. PID2021-123657OB-C32), y el proyecto *Modular robotic and neuroprosthetic customizable systems for the assistance of pathological gait-TAILOR* (RTI2018-097290-B-C31) financiado por MCIN/AEI/10.13039/501100011033 y por “FEDER Una manera de hacer Europa”, por la “Unión Europea”.

## Referencias

- [1] W. H. Organization, «Spinal Cord Injury.» WHO, 19 November 2013. [En línea]. Available: <https://www.who.int/news-room/fact-sheets/detail/spinal-cord-injury>. [Último acceso: 18 July 2022].
- [2] J. Mazaira, F. Labanda, J. Romero y e. al., «Epidemiología de la lesión medular y otros aspectos,» *Rehabilitación*, vol. 32, nº 6, pp. 365-372, 1998.
- [3] L. Simpson, J. Eng, J. Hsieh, D. Wolfe y G. Program, «The health and life priorities of individuals with spinal cord injury: A systematic review,» *J. Neurotrauma*, vol. 29, nº 8, pp. 1548-1555, 2013.
- [4] D. Herrera-Valenzuela, I. Sinovas-Alonso, J. C. Moreno, Á. Gil-Agudo y A. del-Ama, «Derivation of the Gait Deviation Index for Spinal Cord Injury,» *Frontiers in Bioengineering and Biotechnology*, vol. 10, p. 10.3389/fbioe.2022.874074, 2022.
- [5] G. Stampacchia, M. Olivieri, A. Rustici, C. D'Avino, A. Gerini y S. Mazzoleni, «Gait rehabilitation in persons with spinal cord injury using innovative technologies: an observational study,» *Spinal Cord*, vol. 58, nº 9, pp. 988-997, 2020.
- [6] A. Esquenazi, M. Talaty y A. Jayaraman, «Powered Exoskeletons for Walking Assistance in Persons with Central Nervous System Injuries: A Narrative Review,» *PM R*, vol. 9, nº 1, p. 46–62, 2017.
- [7] B. Dobkin y P. Duncan, «Should body weight-supported treadmill training and robotic-assistive steppers for locomotor training trot back to the starting gate?,» *Neurorehabil. Neural Repair*, vol. 26, nº 4, pp. 308-317, 2012.
- [8] E. Cheung, T. Ng, K. Yu, R. Kwan y G. Cheing, «Robot-Assisted Training for People With Spinal Cord Injury: A Meta-Analysis,» *Arch. Phys. Med. Rehabil.*, vol. 98, nº 11, p. 2320–2331, 2017.
- [9] A. Rodríguez-Fernández, J. Lobo-Prat y J. Font-Llagunes, «Systematic review on wearable lower-limb exoskeletons for gait training in neuromuscular impairments,» *J Neuroeng Rehabil*, vol. 18, nº 1, p. 22, 2021.
- [10] S. Luo, H. Xu, Y. Zuo y e. al., «A Review of Functional Electrical Stimulation Treatment in Spinal Cord Injury,» *Neuromol Med*, vol. 22, p. 447–463, 2020.
- [11] E. Nightingale, J. Raymond, J. Middleton, J. Crosbie y G. Davis, «Benefits of FES gait in a spinal cord injured population,» *Spinal Cord*, vol. 45, nº 10, p. 646–657, 2007.
- [12] A. del-Ama, A. Gil-Agudo, J. Pons y J. Moreno, «Hybrid FES-robot cooperative control of ambulatory gait rehabilitation exoskeleton,» *J Neuroeng Rehabil*, vol. 11, nº 27, pp. doi:10.1186/1743-0003-11-27, 2014.
- [13] R. Baker, J. McGinley, M. Schwartz, S. Beynon, A. Rozumalski, K. Graham y O. Tirosh, «The Gait Profile Score and Movement Analysis Profile,» *Gait & Posture*, vol. 30, nº 3, pp. 265-269, 2009.
- [14] G. Borg, *Borg's Perceived Exertion and Pain Scales.*, Champaign, IL: Human Kinetics, 1998.

## Pósteres II

---

Jueves, 24 de noviembre de 2022

# Estudio de las especificaciones energéticas para un dispositivo cerebral implantado recargable inalámbricamente

M. Peñate Arrieta<sup>1</sup>, M. Jiménez Carrizosa<sup>1</sup>, R. Ramos Hortal<sup>1</sup>, R. Gil Simoes<sup>2</sup>

<sup>1</sup> Centro de Electrónica Industrial, Universidad Politécnica de Madrid, Madrid, España,

{m.penate}@alumnos.upm.es

{miguel.jimenezcarrizosa, regina.ramos}@upm.es

<sup>2</sup> Servicio de Neurocirugía, Hospital Universitario de la Princesa, Madrid, España

josercardo.gil@salud.madrid.org

## Resumen

*El glioblastoma multiforme (GBM) es el tumor más común y maligno entre las neoplasias de la glía y tiene una rápida evolución. Uno de los principales problemas del glioblastoma es la baja tasa relativa de supervivencia a 5 años debido. En la última década se ha desarrollado una técnica llamada “Tumor Treatment Fields” (TTF), la cual ha sido aprobada por la FDA para su tratamiento en el GBM. Los TTF consisten en la aplicación de campos eléctricos sobre el tumor, inhibiendo la reproducción de las células cancerígenas y no afectando a las sanas. La eficacia de dicho tratamiento está estrechamente relacionada con el tiempo de aplicación de los campos, siendo efectiva a partir de 18-20 horas diarias de aplicación. Actualmente, los TTF se aplican mediante electrodos colocados en el cuero cabelludo del paciente a través de un aparato casco, por lo que en la mayoría de los casos no se llega a completar el tiempo de aplicación diario necesario, además de que genera frecuentes complicaciones médicas en la piel. Este trabajo aboga por el desarrollo de un dispositivo implantado que genere los campos eléctricos in situ. Para ello, se estudian las necesidades energéticas de dicho dispositivo con el fin de dimensionar la batería del paciente. Se han obtenido resultados experimentales para una batería comercial de Litio-ion recargable de 130 mAh tanto en su fase de carga como en la de descarga, así como un estudio termográfico de la misma.*

## 1. Introducción

El glioblastoma multiforme (GBM) es un tipo de cáncer agresivo el cual se genera en el cerebro o medula espinal. Es un tumor de rápido crecimiento y aunque puede desarrollarse a cualquier edad, es más probable que se encuentre en personas adultas entre los 45 y 70 años [1]. El GBM, al igual que la gran mayoría de tumores cerebrales, no se expande más allá de las estructuras del sistema nervioso central. Este tipo de cáncer está incluido dentro de los gliomas denominados como gliomas de alto grado, siendo estos del grado III-IV. El GBM es el tumor más frecuente con una tasa del 80% de los gliomas detectados y teniendo una incidencia de 16-17 personas por cada 100.000 anualmente [2]. Las opciones tradicionales de tratamiento son: cirugía, radioterapia, quimioterapia, terapia con medicamentos dirigidos, ensayos clínicos, cuidados paliativos o una combinación de ellos.

El principal problema del GBM es que cuenta con una baja tasa de supervivencia a cinco años: 22% para personas de 22 a 44 años, 9% para personas entre los 45 y 54 años y del 6% para personas entre los 55 y los 64 años [3]. Estos datos

son muy poco esperanzadores y es por ello, por lo que se ha llevado a cabo alternativas a los tratamientos tradicionales para conseguir unos mejores resultados.

En el año 2015, la Administración de Alimentos y Medicamentos de los Estados Unidos (FDA) aprobó un dispositivo que aplicaba la técnica de “Tumor Treatment Fields” (TTF) para el tratamiento del GBM [4]. La técnica de TTF consiste en la generación de campos eléctricos alternos de frecuencia intermedia y baja intensidad. Los campos eléctricos que se generan inhiben la replicación de las células cancerosas, impidiendo la formación adecuada del huso mitótico durante la anafase y causando dislocación intracelular de macromoléculas y orgánulos durante la telofase tardía. Además, no perturba a las células sanas.

Actualmente, el único dispositivo comercial existente es el Novo TTF-100 A, también conocido como Optune. Éste utiliza campos de baja intensidad (1-3 V/cm) y de frecuencia intermedia (100-500 kHz) [5] aplicados mediante 4 transductores que se colocan en la piel de la cabeza del paciente previamente afeitada. Se ha demostrado que el uso de Optune con quimioterapia duplica la tasa de supervivencia a 5 años con respecto al uso de quimioterapia únicamente [5]. Algunas de las desventajas del dispositivo comercial son la necesidad de estar cargando con el material, que puede llegar a ser incómodo y pesado, el tener que afeitarse la cabeza para la colocación de los transductores, además de que puede provocar irritación en la piel.

En este trabajo se propone, de forma novedosa, el uso de los TTF de manera invasiva, es decir, realizando un prototipo que se implante sobre o debajo del hueso temporal del paciente. Con este tratamiento invasivo se puede evitar la necesidad de tener que estar cargando con cargadores portátiles pesados, la de rasurar la cabeza o la baja adherencia al tratamiento. En concreto, el principal objetivo de este trabajo es el estudio de las especificaciones energéticas de la batería que se empleará en el dispositivo implantado para poder aplicar el tratamiento de los TTF in situ.

## 2. Métodos

Para el cálculo de las especificaciones energéticas de la batería es necesario conocer o estimar la resistencia

eléctrica del glioblastoma, para ello se ha hecho un cálculo teórico y se ha comparado con otros estudios experimentales [6,7].

También se explicará el montaje realizado para poder medir las corrientes y tensiones de carga y descarga de la batería experimentada, además de la aplicación gráfica (GUI) desarrollada para la obtención de datos.

### 2.1. Resistencia y resistividad del GBM

Se procede a realizar el cálculo teórico de la resistencia eléctrica del GBM a través de su resistividad eléctrica, cuya expresión se muestra en la ecuación (1):

$$dR = \frac{\rho * dl}{A} \tag{1}$$

donde  $\rho$  es la resistividad del GBM y  $A$  el área transversal. Suponiendo que el GBM es esférico, se integra y se obtiene:

$$R = \frac{\rho}{\pi * r} * \log\left(\frac{\tan\left(\frac{\theta_1}{2}\right)}{\tan\left(\frac{\theta_0}{2}\right)}\right) \tag{2}$$

En [6] se muestra cómo las propiedades eléctricas de los tejidos tumorales difieren de las de los tejidos cerebrales sanos, obteniéndose una resistividad media para los meningiomas de  $530 \Omega \cdot \text{cm}$ , para los gliomas de grado bajo de  $160 \Omega \cdot \text{cm}$  y para los gliomas de grado alto de  $498 \Omega \cdot \text{cm}$ . En cuanto a las dimensiones medias de los GBM, que dependerán de cada caso, se han tomado los resultados obtenidos en [7], en el cual sólo se analizaron pacientes con un tamaño mínimo del tumor de 2,5 cm de diámetro. Con estos datos y suponiendo un  $\theta_0=5^\circ$  y un  $\theta_1=175^\circ$  se obtiene una resistencia de  $794,19 \Omega$ . Este valor difiere del obtenido de forma experimental en [6] que era de  $675 \Omega$ . Sin embargo, se aprecia que el orden de magnitud es el mismo.

### 2.2. Análisis de las topologías para generación de TTF

Se han analizado tres diferentes topologías para la generación de los TTF (con un electrodo, *half-bridge*, y *full-bridge*), así como los requerimientos energéticos de cada una de ellas para poder dimensionar la batería. Los análisis se han desarrollado en LTSpice, que es un programa para simular circuitos electrónicos. Para ello se ha empleado una fuente de tensión de 3.7 V que alimenta al circuito y una carga de  $675 \Omega$  para simular la resistencia del cerebro. Los resultados obtenidos se muestran en la sección 3.1.

### 2.3. Calibración de un sensor de corriente

Para poder medir de forma experimental la corriente se ha desarrollado y calibrado un sensor basado en el transductor de bucle cerrado LTS 6-NP que utiliza el efecto Hall para medir la corriente, y que detecta el flujo magnético causado por la corriente en el conductor a través de una bobina de detección.

### 2.4. Desarrollo de una GUI

Se ha desarrollado una interfaz en el software Visual Studio que recoge y muestra de manera visual el valor de la tensión de la batería que se está cargando/descargando,

la corriente que proporciona el sensor de corriente, la potencia instantánea y la energía en la batería.



Figura 1: Interfaz Gráfica

## 3. Resultados

A continuación, se muestran los resultados tanto de la simulación en LTSpice para las diferentes topologías propuestas para generar los TTF, así como los de la calibración del sensor de corriente y los resultados experimentales de carga y recarga de la batería.

### 3.1. Resultados de la simulación en LTSpice para el consumo energético de la batería

En la Tabla 1 se recogen los resultados de las simulaciones, y se observa cómo la primera topología es la que menos pérdidas tiene. Sin embargo, la topología *half-bridge* es la que menos potencia consume de la fuente (10.905 mW). Por otra parte, claramente la topología *full-bridge* es la que peores resultados ofrece desde el punto de vista energético.

Topología	Corriente entrada (mA)	Potencia entrada (mW)	Corriente carga (mA)	Potencia carga (mW)	Pérdidas (mW)
Un electrodo	3,014	11,52	3,014	10,902	0,25
Half-Bridge	2,947	10,905	2,718	10,049	0,85
Full-Bridge	22,521	83,326	24,636	19,875	63,451

Tabla 1. Resultados de corriente, potencia y rendimientos para las diferentes topologías

En la Tabla 2 se muestra la autonomía de la batería en horas, así como la energía consumida para dos escenarios: uno aplicando el tratamiento de TTF durante 18 horas y otro durante las 24 horas del día. En concordancia con los datos mostrados en la Tabla 1, la topología que proporciona más autonomía a la batería es la de *half-bridge*. Para el escenario de aplicación del tratamiento de 18 horas al día se tiene que se tendría que recargar la batería cada 56,11 horas, o lo que es lo mismo, casi cada dos días y medio.

Topología	Autonomía batería (h)	Energía 18 horas (Wh)	Tiempo entre recargas (h)	Energía 24 horas (Wh)	Tiempo entre recargas (h)
Un electrodo	41.75	0.207	53.75	0.276	41.75
Half-Bridge	44.11	0.196	56.11	0.262	44.11
Full-Bridge	5.77	1.499	5.77	1.999	5.77

Tabla 2. Energía y autonomía de la batería

### 3.2. Resultados de la calibración del sensor de corriente

Para la calibración del sensor de corriente se construyó un sencillo circuito con una fuente de tensión de 5 V que alimentaba a unas resistencias y se midió la corriente que pasaba por éstas, tal y como se refleja en la Figura 2.

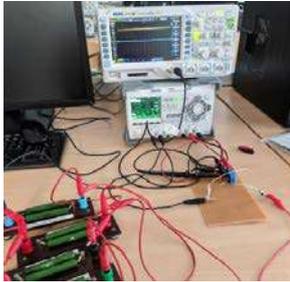


Figura 2 Montaje para calibración del sensor de corriente

Dibujando los valores obtenidos y aplicando una regresión lineal, se obtiene la gráfica que se muestra a continuación.

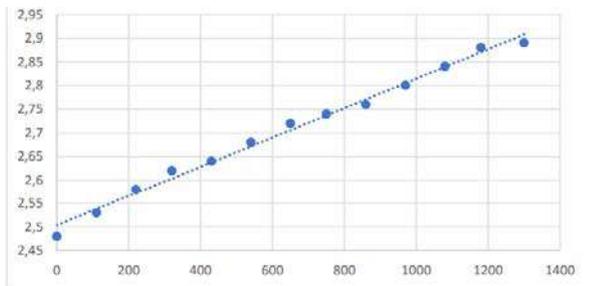


Figura 3. Corriente y tensión

Donde el eje x muestra los datos en miliamperios que lee el sensor y el eje y la tensión que hay en dicho punto. Se puede observar como la salida del sensor de corriente sube 0,03 V cada 100 mA.

Para obtener una mayor precisión de la salida que se obtiene en la Figura 3 (los datos están comprendidos entre 2,48V y 2,88V) se utiliza un amplificador operacional y se obtiene una salida entre 0V y 5V.

Entrada fuente alimentación (V)	Entrada amplificador (V)	Salida amplificador (V)	Consumo sensor (mA)
0	2,51	0,14	0
5	2,59	0,875	220
10	2,66	1,63	440
15	2,72	2,39	670
20	2,81	3,15	890
25	2,85	3,91	1110
30	3,11	4,67	1330

Tabla 3. Salida utilizando amplificador operacional

Obteniendo la relación de la corriente en miliamperios frente a la tensión de salida, se obtiene una relación lineal, cuya ecuación es:

$$y = 290,05x - 145,77 \quad (3)$$

### 3.3. Descarga de la batería GMB401524 utilizando la GUI

A continuación, se mostrarán los resultados obtenidos, donde se ha utilizado la GUI descrita previamente en la sección 2.4. La GUI recopila datos de corriente y tensión, con un tiempo de muestreo configurable. En este ejemplo es cada segundo. A partir de ellos calcula la potencia instantánea, la energía consumida por la batería y el estado de carga de la misma. Los datos pueden ser guardado por el usuario en formato csv o similar para su posterior tratamiento.

A continuación, se mostrarán los resultados para una resistencia equivalente de 666,66 Ω, simulando la resistencia del glioblastoma. En la figura 4 se muestra la gráfica de descarga experimental para la batería recargable de Litio-ión GMB401524 de la compañía GMB Battery. Dicha batería tiene una tensión nominal de 3,7 V y una capacidad de 130 mAh.

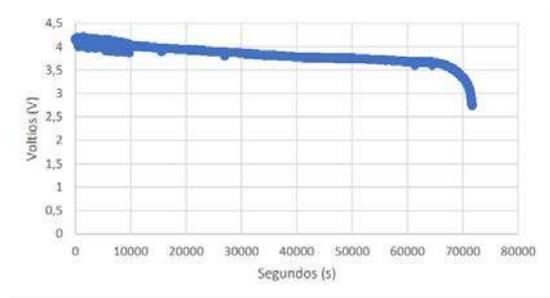


Figura 4. Curva de descarga de la batería

En la Figura 4 se puede observar cómo la descarga completa duró exactamente 71.684 segundos ó 19 horas y 54 minutos. También se puede apreciar cómo durante la mayor parte de la descarga la tensión de la misma decrece ligeramente pero de forma constante, manteniéndose, además, siempre en valores cercanos a su tensión nominal. Una vez que la tensión de la batería es de 3,66V, su descarga se produce de manera mucho más rápida (en torno a la hora 17). En cuanto al consumo energético, la corriente media que circula durante la descarga fue de 22,08 mA.

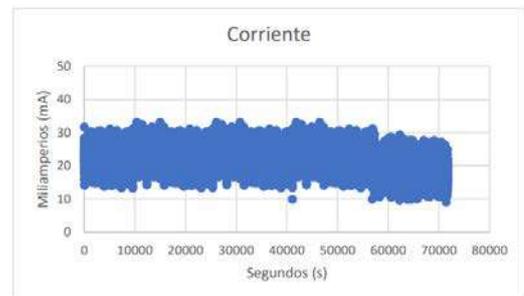


Figura 5. Corriente leída por el sensor de corriente

En la Figura 5 se muestra la potencia instantánea consumida, cuyo promedio fue de 0,08463 W. Por otra parte, la energía entregada por la batería fue de 1,685 Wh.

En la Figura 6 se muestra el análisis térmico de la descarga de la batería mediante una cámara termografía. Se puede concluir que la temperatura máxima no alcanza los 29°C.

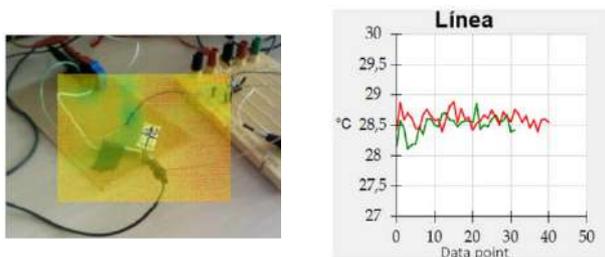


Figura 6. Análisis térmico de la descarga

### 3.4. Carga de la batería GMB401524 utilizando la GUI

Para la recarga de la batería se procedió con distintas ratios de carga C. Se realizaron una carga a 0,5C y otra a 1C. En las Figura 7 y 8 se muestran los resultados experimentales. Se puede observar cómo las gráficas de tensión y corriente son muy similares salvo en la escala de tiempos, las cuales están acorde con los tipos de carga. En ambas gráficas se puede ver como la batería se carga a corriente constante hasta que la tensión de la batería llega a su valor máximo (4,2V). Posteriormente, la corriente cae exponencialmente. La duración de la carga a 0,5C es de 2 horas y 19 minutos, y la de 1C de 1h y 22 minutos.

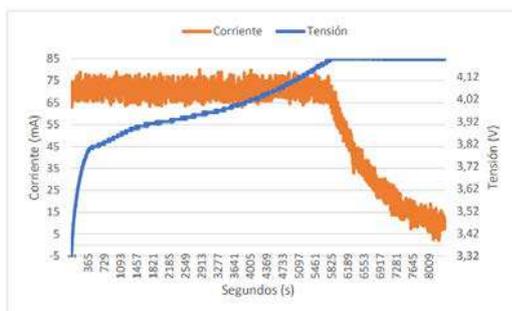


Figura 7. Carga de la batería 0,5C

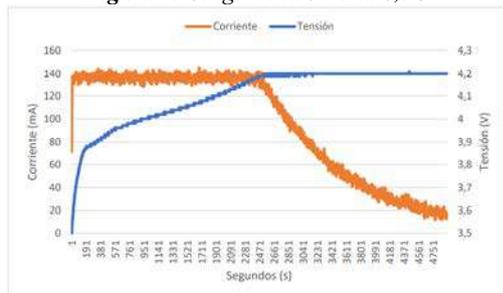


Figura 8. Carga de la batería 1C

## 4. Discusión

Por un parte, se ha realizado un análisis para saber cuáles son los requerimientos energéticos para aplicar los TTF en el GBM mediante un dispositivo implantado. Se ha obtenido de forma teórica el valor de la resistencia eléctrica del GBM, siendo ésta de 794.19  $\Omega$ . Dicho valor es de la misma magnitud a la de estudios experimentales (675  $\Omega$ ). Por otra parte, se ha visto cómo la mejor topología para la generación de los TTF desde el punto de vista de autonomía de la batería es la “half-bridge”, obteniéndose

que, para una aplicación del tratamiento de 18 horas al día, se necesitaría recargar la batería cada 2 días y medio aproximadamente.

También se han mostrado los resultados experimentales de la calibración del sensor de corriente para poder medir correctamente tanto la carga como la descarga de la batería a través de la GUI. Para una descarga a través de una resistencia de 666,66  $\Omega$ , la batería estuvo suministrando energía durante 19 horas y 54 minutos y entregó 1,69 Wh. Además, la temperatura no alcanzó los 29°C. Finalmente, para una carga nominal de 1C se necesitaron 1h y 22 minutos para una recarga completa de la batería.

## 5. Conclusiones

Los resultados obtenidos muestran que los requerimientos energéticos del TTF para un dispositivo implantado pueden ser compatibles con la tecnología actual de baterías recargables de pequeño tamaño. Se ha comprobado de manera experimental que se necesitan 1h y 22 minutos para recargarla, mientras que su descarga completa se alcanzó a las 19 horas y 54 minutos. También se ha comprobado que la temperatura no supera los 29°C, por lo que no supondrá ningún problema para el cerebro, ya que es a partir de los 40°C cuando el hipotálamo empieza a dejar de funcionar correctamente.

## Agradecimientos

Esta actividad de investigación ha sido financiada por el Programa “CEI Grants Program 2022” del Centro de Electrónica Industrial de la Universidad Politécnica de Madrid.

## Referencias

- [1] Rominiyi, O., Vanderlinden, A., Clenton, S.J. et al. “Tumor treating fields therapy for glioblastoma: current advances and future directions” *Br J Cancer* 124, 697–709 (2021). <https://doi.org/10.1038/s41416-020-01136-5>
- [2] Balana, C., J. Capellades, et al. (2007). “Clinical course of high-grade glioma patients with a “biopsy-only” surgical approach: a need for individualised treatment.” *Clin Transl Oncol* 9(12): 797-803.
- [3] Ostrom QT, et Al. “CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2013-2017”. *Neuro Oncol.* 2020 Oct 30;22(12 Suppl 2):iv1-iv96. doi: 10.1093/neuonc/noaa200.
- [4] Rivera F, et Al. “Tumor treating fields in combination with gemcitabine or gemcitabine plus nab-paclitaxel in pancreatic cancer: Results of the PANOVA phase 2 study. *Pancreatol.* 2019 Jan;19(1):64-72. doi: 10.1016/j.pan.2018.10.004. Epub 2018 Oct 17.
- [5] Novocure. Optune: instructions for use. 2019; Available from: [https://www.optune.com/Content/pdfs/Optune\\_IFU\\_8.5x11.pdf](https://www.optune.com/Content/pdfs/Optune_IFU_8.5x11.pdf). Última visita 15 Agosto 2022.
- [6] J. Latikka and H. Eskola, “The Resistivity of Human Brain Tumours In Vivo”, doi: 10.1007/s10439-018-02189-7.
- [7] V. Rajshekhar, “Continuous impedance monitoring during CT-guided stereotactic surgery: relative value in cystic and solid lesions,” *Br J Neurosurg*, vol. 6, no. 5, pp. 439–444, 1992, doi: 10.3109/02688699208995033.

# Deep Learning methodologies for brain image reconstruction in Positron Emission Tomography

N. Rufo Rafael de la Cruz<sup>1</sup>, J. E. Ortuño Fisac<sup>2,3</sup> and G. Kontaxakis<sup>3</sup>

<sup>1</sup> ETSI Telecomunicación, Universidad Politécnica de Madrid, Madrid, Spain, n.rufo@alumnos.upm.es

<sup>2</sup> CIBER de Bioingeniería, Biomateriales y Nanomedicina, Instituto de Salud Carlos III, Madrid, Spain

<sup>3</sup> Biomedical Image Technologies Group, Information Processing and Telecommunications Center, ETSI Telecomunicación, Universidad Politécnica de Madrid, Madrid, Spain, {je.ortuno, g.kontaxakis}@upm.es

## Abstract

*Positron emission tomography (PET) is a functional imaging modality that, by its nature, is very prone to noise. Most current image reconstruction methods in PET have dealt with noise, but not always satisfactorily. Recently, novel approaches have been proposed, using the properties of Deep Learning (DL) methodologies to address this problem. We have designed and implemented a DL-based image reconstruction method, directly obtaining images from sinograms using an encoder-decoder architecture. The proposed method was tested in geometric phantoms, brain phantoms, as well as real brain scans, and its performance was compared with the one obtained by analytical and iterative methods. Our results suggest that DL methodologies might constitute a valid alternative approach for PET image reconstruction, reducing the influence of noise compared to traditional methods and potentially achieving a more accurate image interpretation for diagnosis and treatment, in addition to a potential reduction of the injected radioactive dose.*

## 1. Introduction

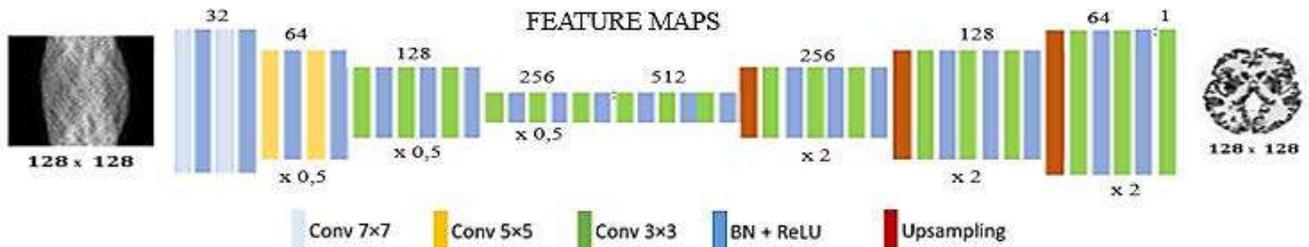
Positron emission tomography (PET) [1] is a nuclear medicine imaging modality that provides functional information about the metabolic activity of the living organism. It has been proven especially useful in diseases, such as cancer, where it has important role for both diagnosis and treatment [2]. Data collected during image acquisition are stored in the form of sinograms that cannot be interpreted directly by an observer and must be first reconstructed into images. However, a solution to the inverse problem is not straightforward due to the nature of this imaging modality, and thus, the reconstructed images are very prone to noise.

Image reconstruction methods that have been used over the years are analytical techniques, such as filtered backprojection (FBP) [3], which uses projection operators that do not model the physics and noise of a PET scanner, or iterative methods, such as Maximum Likelihood Expectation Maximization (ML-EM) [3], which can model the Poisson nature of the data and include a more complex physical model of the acquisition process. In addition, regularization terms can be added in the iterative process to consider a priori information. However, iterative methods have high computational demands, which increases the more elaborate and accurate the physical model of the system used is. Moreover, the arbitrariness in selecting the proper number of iterations

or in defining the most appropriate parameters involved in the specific algorithms used do not always contribute to obtaining the most accurate image as a final result. Hence, with the current growth of Deep Learning (DL) [4], very innovative methods have been recently proposed, that can learn the physical process of PET data generation and the noise model from the data provided for training, based on learning mappings from noisy data to high-quality images. Depending on the number, form, and type of these mappings, different architectures can be achieved. Of particular importance are the encoders and decoders that can represent abstract maps of the essential features of the data introduced in the encoder stage to be later transformed into images of outstanding quality in the decoder stage. Since noise is not a relevant feature in data, it is not included in the feature maps and, thus, not represented in the final image.

Currently, the main approaches to the use of DL in PET image reconstruction are: synthesis, analysis, unfolded, and direct methods [5]. The first three maintain the mechanisms of traditional iterative reconstruction methods, except for the use of DL for learning the elements that need prior image knowledge, avoiding any analytic, intuitive, or handcrafted component of the reconstruction process. Synthesis and analysis approaches learn the regularization step, and the unfolded methods extend the iterative methods in a succession of reconstruction operators, partially or completely interleaving them in repeated blocks with deep denoising operators.

The fourth approach leaves behind everything done so far with iterative methods and learns the direct mapping between sinogram and reconstructed image. These methods have the advantage not to require any prior information from the image, but demand a large amount of training data. Among these methods are: AUTOMAP [5], initially developed for magnetic resonance imaging (MRI), which learns to perform the image reconstruction in a way analogue to an inverse Radon transform, with fully connected layers before an autoencoder; DeepPET [6], which learns entirely on an encoder-decoder architecture; DirectPET [7], which again uses an encoder-decoder but incorporates at the end of the encoder phase the idea of the inverse Radon transform with fully connected layers. We study here the viability of direct methods by implementing a modified version of DeepPET.



**Figure 1.** Network architecture for our version of the DeepPET image reconstruction method.

## 2. Methodology

### 2.1. Deep Learning method

We implemented a modified version of DeepPET, where another image format is given and thus the abstraction level that the feature maps reach differs from the original, resulting in a reduction of the total number of layers from 31 to 22, which was found to produce better results and also reduces the memory requirements. We also changed the optimizer from Adam stochastic gradient descent to RMSProp from Keras, the Python DL Application Programming Interface (API), that we also found to produce better results. The mean squared error (MSE) was maintained as the loss function.

Fig. 1 depicts the resulting architecture, which consists of an encoding phase, where the sinogram input is compressed into feature maps through convolution layers, and a decoding phase, in which additional up-sampling layers are introduced to reach the final reconstructed image. After each convolution at encoding, a ReLU activation function and a batch normalization are used.

To further alleviate the massive memory requirements of this method, the network has been fed with a custom batch generator that loads GPU input data from the disk in a volatile and dynamic manner when needed.

### 2.2. Traditional methods

For comparison, the FBP with a Hamming filter and the ML-EM and its regularized version with a median filter, have been implemented using the ASTRA Toolbox library [8], as the most representative current image reconstruction methods in PET.

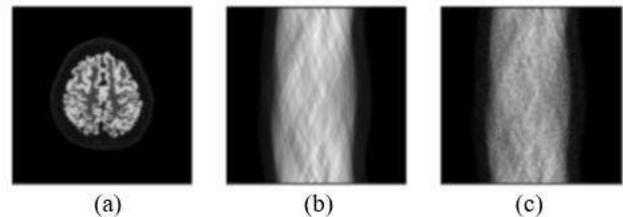
### 2.3. Datasets

To test the performance of the network the following datasets are used:

- Geometric phantom dataset: 100000 unique images, each of them with 20 ellipses of varying intensity, orientation and size simulated in MATLAB.
- Brain phantom dataset, containing the 4120 images (2D slices) from the 20 brain phantoms 3D volumes initially simulated for MRI and adapted to the PET domain, publicly available [9][10].
- Brain phantom lesions dataset: same as before, with hot and cold lesions introduced, publicly available [9][10].
- Real brain images dataset: 5230 PET images (2D slices) from 37 healthy patients 3D volumes, publicly available for research purposes [11].

For all datasets, 80% of the images are destined for training, 10% for validation, and 10% for testing, except for the geometric phantom, where only 100 images were kept for the tests.

For the brain datasets, data augmentation ( $\times 50$ ) was performed to increase the variability during the training. Sinograms were simulated using the ASTRA Toolbox library, with 2D ideal geometry and parallel projectors. Finally, random noise with 160 mean Poisson distribution is added to the sinograms to mimic the statistics of the PET acquisition. Neither attenuation nor scattering effects were included. An example is shown in Fig. 2.



**Figure 2.** (a) Sample from the brain phantom dataset, (b) ideal sinogram simulated, (c) Sinogram with Poisson noise.

## 3. Results and discussion

The experiments have been divided into three phases according to the dataset used and whose results have been evaluated with the structural similarity metric (SSIM), peak signal-to-noise ratio (PSNR) and normalized root mean square error (NRMSE) [12]. Besides, the working environment implemented (traditional methods, network model and training, inference module) has been uploaded to a GitHub repository to allow the community to perform their own experiments and explore this field, given that such code was not available before, available at: <https://github.com/NuriaRufo/DeepPET>.

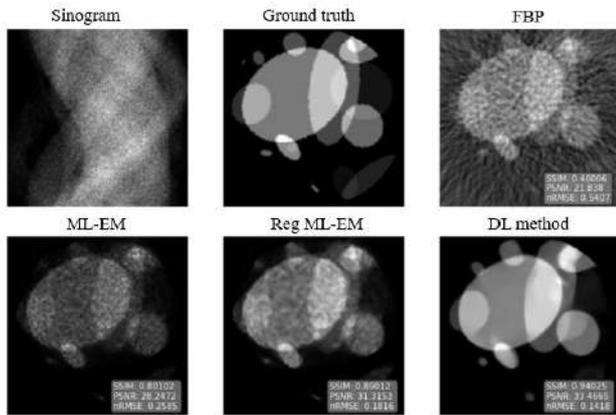
### 3.1. Geometric phantom dataset

Initially, a training of the network was conducted using the training set of geometric phantoms of 90000 images with size  $128 \times 128$ . The learning rate was set to  $10^{-4}$  and the batch size to 32. The GPU used was PNY QUADRO RTX-A6000 48GB GDDR6.

The mean error of images in the test set reconstructed with the different methods can be seen in Table 1. The DL method outperforms the traditional methods, which can be also qualitatively verified in Fig. 3.

Method	SSIM	PSNR	NRMSE
FBP	0,4769	21,6767	0,5606
ML-EM	0,8002	28,1906	0,2664
ML-EM regularized	0,8875	31,1608	0,1892
DL method	<b>0,9228</b>	<b>32,0470</b>	<b>0,1710</b>

**Table 1.** Average error for geometric phantom dataset reconstruction.



**Figure 3.** Reconstructed images with traditional methods and the direct DL method.

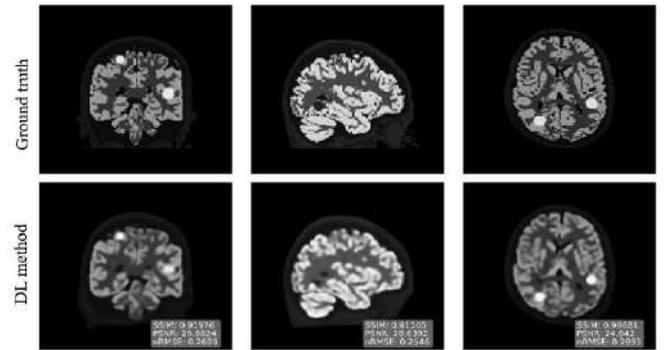
### 3.2. Brain phantom dataset

The first training was performed excluding lesions and considering only axial slices, to delimit the problem. The same hyperparameters and GPU of the previous experiment were maintained. Once the network was trained, the results showed that the DL method was able to reconstruct structures as delicate as the brain, although it did not surpass the quality of traditional methods.

To test the generalizability of the method, the model trained with the phantom without lesions and axial slices was confronted with the phantom test dataset with lesions. It was observed that the model could recognize lesions, although not with sufficiently high quality. Therefore, a new training was performed with the phantom training set with lesions, but still considering just the axial slices and where hyperparameters were maintained. The introduction of random lesions in the training set improved network performance by 4%, although it still did not outperform traditional methods.

In search of a more considerable amount of training data, sagittal and coronal slices were added to the axial ones for a new training with the same hyperparameters. However, the model trained with all slices was confronted with only axial slices from the test dataset to ensure homogeneity between brain phantom results. Fig. 4 shows how the qualitative results are finally favourable for the direct DL method. It is worth mentioning that the reconstruction times for this method are remarkably fast, as all images were reconstructed in approximately one second, compared to traditional reconstruction times that were 800 seconds using CPU and 20 seconds using GPU.

Despite the satisfactory results in single samples, the mean error yields a better reconstruction by the traditional



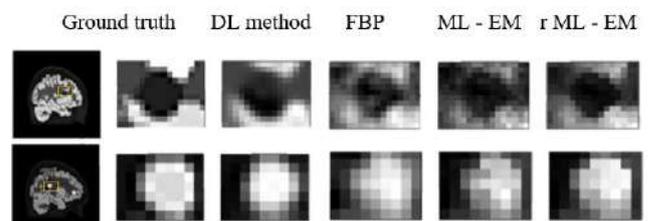
**Figure 4.** Three slices of a brain phantom test set reconstructed using the DL method (lower row), compared with the ground truth (upper row).

Method	SSIM	PSNR	NRMSE
FBP	0,3900	20,8424	0,6425
ML-EM	0,9223	27,6359	0,2934
ML-EM regularized	<b>0,9335</b>	<b>28,2540</b>	<b>0,2733</b>
DL method	0,8889	26,0090	0,4892
DL method (lesions)	0,9041	26,7298	0,4344
DL method, all cuts	0,9260	27,5362	0,3329

**Table 2.** Average error for the brain phantom test dataset containing lesion reconstruction and only axial slices.

ML-EM methods and its regularized version. However, a study of the distribution of these errors reveals that this is due to a few outliers with poor DL reconstructions, caused by the presence of structures adjacent to the brain, which are underrepresented in the training and, consequently, not learned well enough. Nevertheless, regarding the median in the error distribution, the DL method is placed ahead of the traditional methods. In addition, 74,4% of the cases were better reconstructed with the network than with the traditional methods.

A set of 5 random lesions was analysed in detail (Fig. 5). The DL method resulted in a better qualitative and quantitative reconstruction with average relative error at 0,2928, whereas this was 0,3443 for regularized ML-EM.



**Figure 5.** Reconstructed lesions using FBP, ML-EM, regularized ML-EM and the DL method.

### 3.3. Real brain images dataset

Regarding real brain PET images, the first experiment with the test dataset (all slices) was first performed with the best model trained with the brain phantom (trained with all slices). We observed that it tends to over-define the structures of the cerebellum, as it occurs normally in the brain phantom.

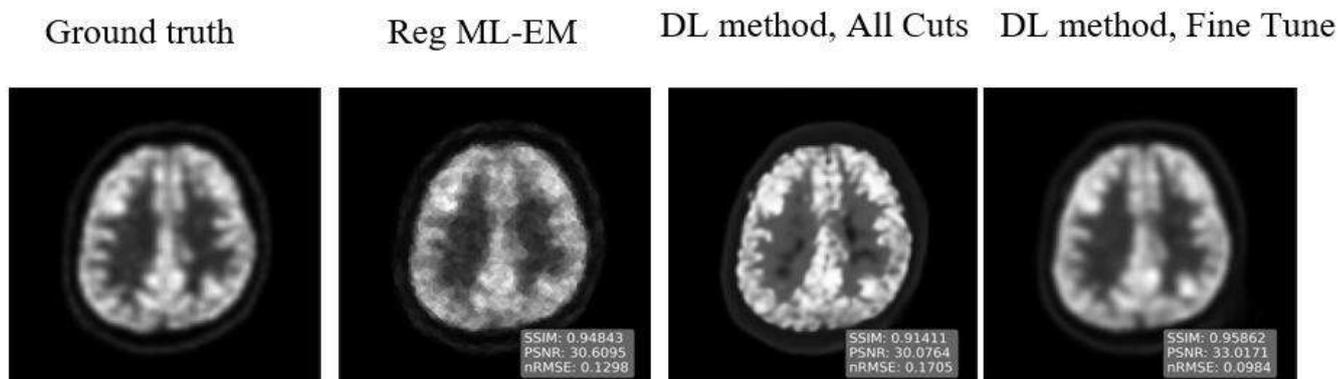


Figure 6. Reconstruction results for a sample of the real PET brain images dataset using the DL method, compared with the ground truth.

Method	SSIM	PSNR	NRMSE
FBP	0,4435	20,4825	0,5579
ML-EM	0,9241	30,1782	0,1900
ML-EM regularized	<b>0,9464</b>	31,9411	0,1530
DL method, All Cuts	0,8788	32,8847	0,2276
DL met. Fine Tune	0,9411	<b>32,8847</b>	<b>0,1489</b>

Table 3. Average error for real PET dataset image reconstruction.

Therefore, a new training was performed for fine-tuning, taking advantage of the prior knowledge of an already trained network to adapt it to the new problem. Based on the model trained with the phantom, training was performed with the real brain datasets to adapt the network to the domain of these images. The learning rate was lowered to 0,00005 and the number of epochs was set to only 8. In the training of one hour, the results of the tests yield that it both qualitatively (Figure 6) and quantitatively (Table 3) overcomes traditional methods, although the SSIM is again worsened by outliers containing structures other than the brain.

#### 4. Conclusions

The current study focused on the case of brain image reconstruction for PET. In view of the results, we conclude that direct DL image reconstruction methods constitute a powerful tool for PET image reconstruction from sinograms. It has been confirmed that the modified version of DeepPET is able to produce real PET image reconstructions with very satisfactory quality in terms of noise characteristics from noisy sinograms. Moreover, it is much faster and overcomes in image quality traditional FBP, ML-EM and regularised ML-EM algorithms in 100%, 90,3% and 85,4 % of the cases, respectively.

However, we have also confirmed the method's robustness, and therefore, the outperformance of the traditional methods is due to the homogeneity and size of the training dataset used. This, in turn, highlights the major disadvantage of direct methods, which is the vast memory requirements to hold and handle the minimum 250000 image pairs necessary to learn the whole mapping from sinogram to image and thus to obtain a functional and competitive model. Consequently, the balance may

tip towards the use of synthesis or unfolded methods that take the basis of iterative image reconstruction, which, as we have seen, has not always been trivial to overcome in terms of image quality, therefore combining Deep Learning with traditional methods may be a promising approach.

This work implies, however, the validation of direct reconstruction and shows the potential that DL can have in the field of medical image reconstruction, providing us with higher quality images compared with traditional methods that can result both in better diagnosis and treatment for cancer patients as well as in opening the possibility that less doses of radioactive material may be needed to perform the image acquisition.

#### References

- [1] M. E. Phelps, *PET: physics, instrumentation, and scanners*. Springer, 2006.
- [2] D. Delbeke, "Oncological applications of FDG PET imaging: Brain tumors, colorectal cancer lymphoma and melanoma," *J. Nucl. Med.*, 40(4): 591–603, 1999.
- [3] D.W. Wilson and B.M.W. Tsui, "Noise properties of filtered-backprojection and ML-EM reconstructed emission tomographic images," *IEEE Trans. Nucl. Sci.*, 40(4): 1198–1203, 1993.
- [4] R. Y. Choi, A. S. Coyner, J. Kalpathy-Cramer, *et al*, "Introduction to machine learning, neural networks, and deep learning," *Transl. Vis. Sci. Techn.*, 9(2), 2020.
- [5] A.J. Reader, G. Corda, *et al*, "Deep Learning for PET Image Reconstruction," *IEEE Trans. Rad. Plasma Med. Sci.*, 5(1): 1–25, 2020.
- [6] I. Häggström, C.R. Schmidtlein, *et al*, "DeepPET: A deep encoder-decoder network for directly solving the PET image reconstruction inverse problem," *Med. Imag. Anal.*, 54: 253–262, May 2019.
- [7] W. Whiteley, W. K. Luk, and J. Gregor, "DirectPET: full-size neural network PET reconstruction from sinogram data," *J. Med. Imag.*, 7(3): 32503, 2020.
- [8] W. van Aarle *et al.*, "The ASTRA Toolbox: A platform for advanced algorithm development in electron tomography," *Ultramicroscopy*, 157: 35–47, Oct. 2015.
- [9] C.A. Cocosco, V. Kollokian, R.K.-S. Kwan, and A.C. Evans, "BrainWeb: Online Interface to a 3D MRI Simulated Brain Database," *Neuroimage*, 5(4): S425, 1997.
- [10] A. Mehranian and A.J. Reader, "Model-Based Deep Learning PET Image Reconstruction Using Forward-Backward Splitting EM," *IEEE Trans. Rad. Plasma Med. Sci.*, 5(1): 54–64, 2021.
- [11] I. Mérida, *et al*, "CERMED-IDB-MRXFDG: a database of 37 normal adult human brain [18F]FDG PET, T1 and FLAIR MRI, and CT images available for research," *EJNMMI Res.*, 11(1):91, 2021.
- [12] Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli, "Image Quality Assessment: From Error Visibility to Structural Similarity," *IEEE Trans. Imag. Proc.*, 13(4): 600–612, April 2004.

# The Identification of Patients with Type 1 Diabetes Mellitus Subgroups Applying Cluster Analysis Techniques

F.J. Somolinos-Simón<sup>1</sup>, G. García-Sáez<sup>1,2</sup>, J. Tapia-Galisteo<sup>1,2</sup>, M.E. Hernando<sup>1,2</sup>

<sup>1</sup>Center for Biomedical Technology, ETSI de Telecomunicación, Universidad Politécnica de Madrid, Madrid, Spain

<sup>2</sup> CIBER-BBN: Networking Research Centre for Bioengineering, Biomaterials and Nanomedicine, Madrid, Spain

## Summary

*Patients with Type 1 Diabetes (T1D) require the administration of insulin to maintain glycaemic control. Currently, two modes of subcutaneous insulin delivery have gained wider acceptance: Multiple daily injections (MDI) and continuous subcutaneous insulin infusion (CSII). This research work aims to assess if patients with T1DM can be stratified according to four main determinants like age at diagnosis, diabetes duration, body mass index (BMI) and HbA1c level. Data have been pre-processed to clean or convert variables and a data-driven clustering analysis was performed using k-means method. Four clusters of T1DM patients have been identified in this research work. Each of these clusters is characterised by a higher value of one of the four variables used in the analysis with the exception of HbA1c level where no major differences were found between clusters. This preliminary research work shows that T1DM patients can be stratified in subgroups according to main determinants obtaining subgroups of patients characterised by a high value of one of the four variables used. Therefore, each variable stands out in each of the clusters obtained.*

## 1. Introduction

Diabetes Mellitus is a chronic disease that affects an increasing number of people worldwide. The condition is characterised by high blood glucose levels, which can lead to several complications at short and long term. In fact, Type 1 Diabetes mellitus (T1DM) and her complications are an important cause of mortality and morbidity in the world [1].

Clinicians need to adjust each patient's treatment to the evolving condition over time, considering the individual characteristics and the high number of treatment alternatives available. In T1DM patients is required the administration of insulin to maintain glycaemic control [2]. Currently, two modes of subcutaneous insulin delivery have gained wider acceptance: Multiple daily injections (MDI) and continuous subcutaneous insulin infusion (CSII) [3].

Randomized controlled trials (RCT) have shown that, when compared with MDI, CSII was associated with a slightly lower glycated haemoglobin (HbA1c) level [3].

The case study on diabetes of the H2020 HTx project ([htx-h2020.eu](http://htx-h2020.eu)) aims to link evidence from RCTs to Real World Data (RWD) and estimate the impact of health technology on specific subgroups of patients, as a first step to build prediction models to personalize treatment strategies.

This research work aims to assess if patients with T1DM can be stratified into subgroups based on a series of common characteristics. With the resulting clusters, disease progression, treatment, and complications can be compared in each subgroup. Moreover, this work will allow to extract patterns about the most effective and cost-effective treatment combinations according to individual patients' characteristics.

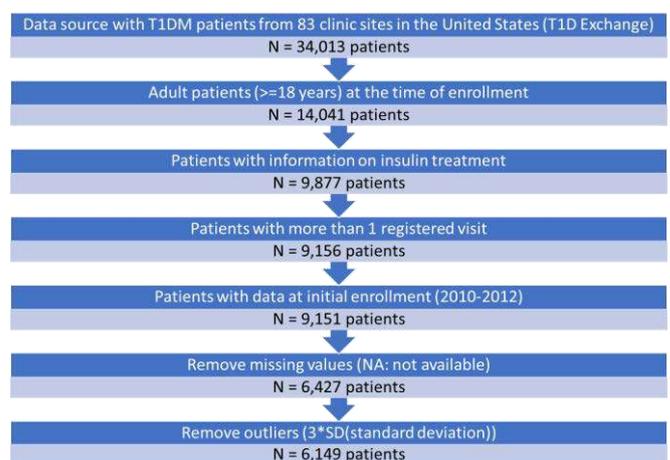
## 2. Methods

We used a longitudinal, prospective data repository with T1D patients from 83 clinics in the United States (T1D Exchange<sup>1</sup>).

As a first step, data have been pre-processed to clean or convert variables.

From the total of 34,013 patients, the following inclusion criteria were established: adult patients (>18 years) at the time of enrolment; patients with information on insulin treatment; and, patients with more than 1 registered visit (N= 9,156 patients).

Then, we have done a preprocessing and filtering of data removing missing values and outliers (mean  $\pm$  3\*SD (standard deviation)) (N = 6,149 patients) (see Figure 1).



**Figure 1.** Preprocessing and filtering of available data.

<sup>1</sup> The source of the data is the T1D Exchange, but the analyses, content and conclusions presented herein are solely the responsibility of the authors and have not been reviewed or approved by the T1D Exchange.

Once the final individuals for the study had been obtained, a characterisation of the sample has been carried out to find out the percentage of men and women, the predominant method of insulin administration, or the ethnicity, among other characteristics, such as HbA1c level, body mass index (BMI) or age at diagnosis (see Table 1).

DEMOGRAPHICS	
<b>Age at diagnosis, years</b>	
Aged 20 years or older	2,213 (35.99%)
<b>Sex</b>	
Women	3,224 (52.40%)
Men	2,925 (47.60%)
<b>Ethnicity</b>	
American indian/Alaskan native	17 (0.30%)
Asian	66 (1.10%)
Black non-hispanic	215 (3.50%)
Hispanic or latino	292 (4.70%)
Native Hawaiian/Other pacific islander	3 (0.10%)
White non-hispanic	5,440 (88.50%)
More than race	111 (1.80%)
CLINICAL FEATURES	
<b>Body mass index (BMI) (Kg/m<sup>2</sup>)</b>	
>=25 Kg/m <sup>2</sup>	3,749 (60.97%)
< 25 Kg/m <sup>2</sup>	2,400 (39.03%)
LABORATORY TESTS	
<b>HbA1c level (%)</b>	
>=7.5%	3,331 (54.17%)
< 7.5%	2,818 (45.83%)
<b>Insulin Delivery Method</b>	
Injections/pens	2,423 (39.40%)
Insulin pumps	3,654 (59.40%)
Pumps and Injections/pens	68 (1.10%)
Neither	4 (0.10%)
Data are n (%)	

**Table 1.** Baseline characteristics of selected individuals.

Table 1 shows that the sample is characterized by obese patients, most of them are white non-hispanic and they have been diagnosed very young.

Subsequently, a data-driven clustering analysis was performed using k-means method. After a review of the literature had been carried out, especially, Ahlqvist and colleagues [4], Bello-Chavolla and colleagues [5] and Dennis and colleagues [6], they have reported a data-driven cluster analysis based on variables using clinical features like BMI, age at diagnosis, diabetes duration, baseline HbA1c, glutamate decarboxylase antibodies and homeostatic model assessment 2 estimates of  $\beta$ -cell function and insulin resistance. In this project, clusters were based on four available variables in the

T1DExchange dataset: age at diagnosis, diabetes duration, body mass index (BMI) and glycosylated haemoglobin (HbA1c) level.

K-means method is based on each cluster is associated with a centroid (centre point) and each point is assigned to the cluster with the closest centroid. It is a clustering technique for which the number of clusters ( $k$ ) to be obtained must be specified previously. Therefore, average silhouette method has been used to determine the optimal number of clusters and this result was 4 clusters.

Cluster analysis was done on values centred to a mean value of 0 and an SD of 1. Therefore, the data were standardized to make variables comparable. K-means clustering was done with a  $k$  value of 4 using the  $k$ -means function in python 3.9.

### 3. Results

Four clusters of T1DM patients have been identified in this research work. An analysis has been made of the distribution of all the available variables, whether or not they were used to obtain the clusters. As it can be seen, each cluster is characterised by a high value of one of the four variables used with the exception of HbA1c level where no major differences were found between clusters.

Table 2 shows the mean values and standard deviation obtained for each of the variables in all clusters.

CLUSTER (N)	AGE AT DIAGNOSIS (YEARS)		DIABETES DURATION (YEARS)		HBA1C LEVEL (%)		BODY MASS INDEX (KG/M <sup>2</sup> )	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
<b>1 (2,244)</b>	11.90	5.99	12.00	6.65	8.13	1.37	24.24	2.80
<b>2 (1,312)</b>	<b>37.25</b>	9.14	12.19	8.70	7.31	1.08	26.21	3.91
<b>3 (1,108)</b>	15.51	9.48	20.55	10.32	8.21	1.24	<b>33.98</b>	3.30
<b>4 (1,485)</b>	13.34	7.98	<b>37.08</b>	8.42	7.30	0.93	26.24	3.42

\*SD: Standard deviation

**Table 2.** Mean values and standard deviation of the four main determinants in each of the clusters obtained.

Analysing the different clusters obtained, cluster 4 is characterised by a high value for the duration of diabetes, approximately 37 years, and a lower value for blood haemoglobin levels. This could mean the longer the diabetes duration, the longer the duration of treatment and the more controlled the haemoglobin level. Cluster 3 is characterised by a high BMI value of over 30 and is therefore a cluster in which patients are obese, while the rest are characterised by being overweight. Cluster 2 is characterised by a very late age at diagnosis, approximately 37 years. Finally, there are no major differences in haemoglobin levels between the clusters, however, cluster 3, characterised by BMI (obese patients), has the highest value (8.21%).

Figure 2 shows the graphs (boxplots) where it can be better observed how each variable stands out in each of the clusters obtained.

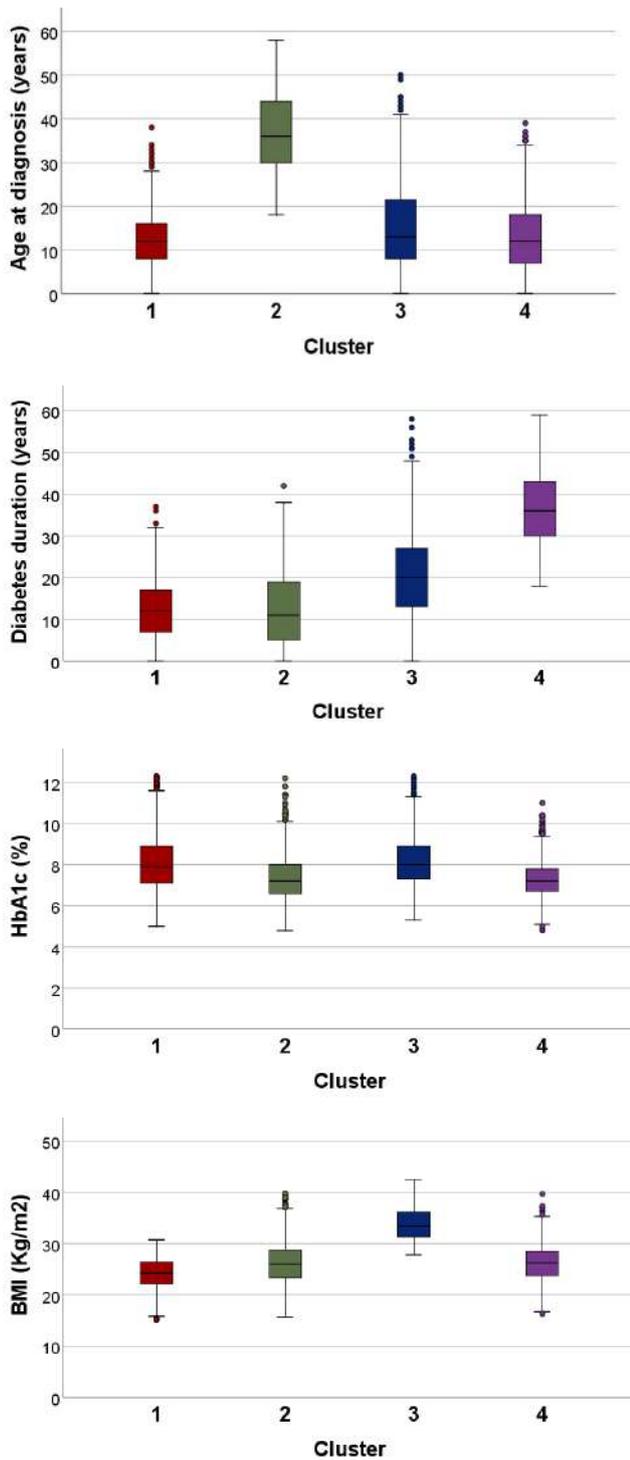


Figure 2. Box plots of the four main determinants in each of the clusters obtained.

Clusters of [4], [5], [6] were performed in patients with type 2 diabetes, however, we have obtained similar clusters characterized by age at diagnosis or BMI in patients with type 1 diabetes mellitus, what allows to stratify into subgroups based on common characteristics in order to study disease progression, treatment, and complications in each subgroup.

The insulin delivery method used in each patient was also analysed in order to understand the influence of the use of healthcare technology on insulin delivery.

The main insulin delivery methods used in T1D Exchange are injections/pens and insulin pumps. Figure 3 shows the percentage of use of these main insulin delivery methods in each of the clusters obtained.

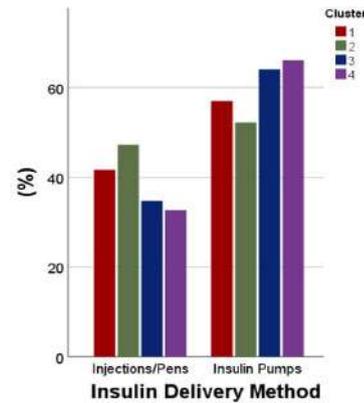


Figure 3. Insulin delivery methods in each of the clusters obtained.

It should be noted that the predominant method of insulin administration is the use of insulin pumps, by continuous subcutaneous insulin infusion (CSII), in all clusters. Nevertheless, there is a more noticeable difference between the main insulin delivery methods in clusters 3 and 4 which are characterised by more obese patients and longer duration of diabetes. Therefore, the method of insulin administration is very important for better control of T1DM depending on the characteristics of the patient. In this case, if they are obese and have had diabetes for a long time, the use of insulin pumps is recommended.

As for other patient characteristics such as gender or ethnicity, the percentage of men and women is similar in all clusters, with slightly more women than men in all clusters following a similar distribution as the total sample set (see Figure 4).

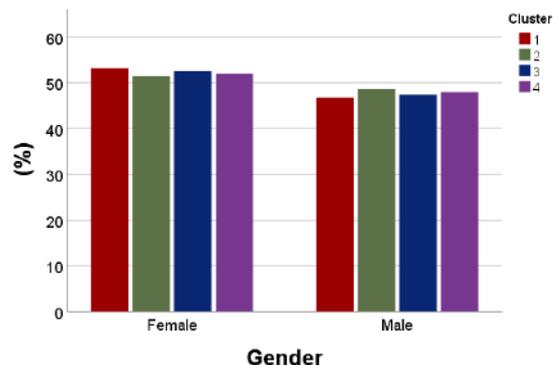


Figure 4. Gender distribution of patients in each of the clusters obtained.

Nonetheless, after analysing ethnicity, it can be seen that the majority of patients are White non-Hispanic. It is noteworthy that in clusters 1 and 3, the percentage of this type of ethnicity is slightly lower than the rest, while in these same clusters, characterised by a higher HbA1c level, there is a higher percentage of Blacks non-Hispanic and Hispanics or Latinos than in the rest of the groups (see Figure 5).

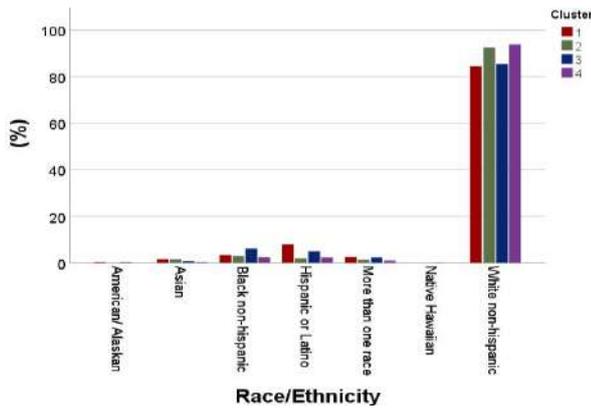


Figure 5. Ethnicity distribution of patients in each of the clusters obtained.

Finally, the study of complications and drugs associated to different clusters have been done in this work.

Figure 6 shows the percentage of patients who suffered different diseases at enrolment for each cluster. The most common are hypertension, depression and retinopathy in all clusters. Complications distribution is similar for each disease in each cluster. For example, the percentage of patients who suffered any disease is higher in cluster 4, due to diabetes duration is more elevated, therefore, these patients have been able to develop more complications because they have been with diabetes for a longer time. Although, in the case of depression, it should be noted that the majority group is cluster 3, obese people suffer more depression.

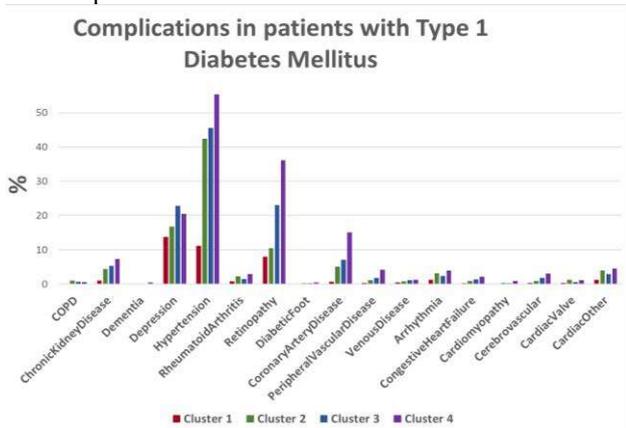


Figure 6. Complications distribution of patients in each of the clusters obtained.

Regarding medications, an analysis of insulin treatment in each of the clusters obtained has been done in this project. Figure 7 shows the percentage of patients who consumed different insulins at enrolment for each cluster. The most common are Insulin Lispro, Insulin Aspart and Insulin Glargine in all clusters. The first two insulins are rapid-acting analogs which result from changes to the amino acid structure of human insulin which lead to decrease in hexameric insulin formation after injection into the subcutaneous space. This leads to more rapid dissolution of insulin into monomers, more rapid insulin absorption into the bloodstream, and a shorter duration of action, whereas Insulin Glargine is a long-acting insulin which provides basal insulin coverage. It is worth mentioning that Insulin Lispro is less common in cluster 2,

characterised by older age at diagnosis, than in the other clusters, although there are no major differences between clusters.

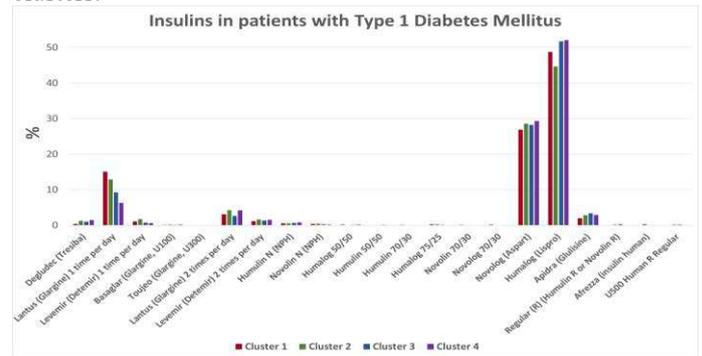


Figure 7. Insulins distribution of patients in each of the clusters obtained.

#### 4. Conclusions

This preliminary research work shows that T1DM patients can be stratified in subgroups according to main determinants. Future research includes:

- To study the evolution of the clusters throughout the different study visits and to analyse the different variables in these visits.
- To build a powerful tool considering the effectiveness of treatment in patient subgroups.

#### Acknowledgements

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 825162.

#### References

- [1] G. Danaei et al., Cardiovascular disease, chronic kidney disease, and diabetes mortality burden of cardiometabolic risk factors from 1980 to 2010: A comparative risk assessment. *The Lancet Diabetes and Endocrinology*, vol. 2, no. 8, pp. 634–647, 2014.
- [2] S. Clifford, M. Perez-Nieves, A. M. Skalicky, M. Reaney, and K. S. Coyne, A systematic literature review of methodologies used to assess medication adherence in patients with diabetes. *Curr Med Res Opin*, vol. 30, no. 6, pp. 1071-85, Jun 2014.
- [3] P. M. Rys, A. H. Ludwig-Slomczynska, K. Cyganek and M. T. Malecki, Continuous subcutaneous insulin infusion vs multiple daily injections in pregnant women with type 1 diabetes mellitus: a systematic review and meta-analysis of randomised controlled trials and observational studies. *Eur J Endocrinol*, vol. 178, no. 5, pp. 545-563, May. 2018.
- [4] E. Ahlqvist et al., “Novel subgroups of adult-onset diabetes and their association with outcomes: a data-driven cluster analysis of six variables,” *The Lancet Diabetes and Endocrinology*, vol. 6, no. 5, pp. 361–369, May 2018.
- [5] Bello-Chavolla OY, Bahena-López JP, Vargas-Vázquez A, et al. Clinical characterization of data-driven diabetes subgroups in Mexicans using a reproducible machine learning approach. *BMJ Open Diab Res Care* 2020;8:e001550. doi:10.1136/bmjdr-2020-001550.
- [6] J. M. Dennis, B. M. Shields, W. E. Henley, A. G. Jones, and A. T. Hattersley, “Disease progression and treatment response in data-driven subgroups of type 2 diabetes compared with models based on simple clinical features: an analysis using clinical trial data,” *The Lancet Diabetes and Endocrinology*, vol. 7, no. 6, pp. 442–451, Jun. 2019.

# Evaluación del potencial de las imágenes OCT maculares para el diagnóstico de la enfermedad de Parkinson

D. Romero-Bascones<sup>1</sup>, I. Gabilondo Cuellar<sup>1,2,3</sup>, M. Barrenechea Carrasco<sup>1</sup>, U. Ayala Fernández<sup>1</sup>

<sup>1</sup> Mondragon Unibertsitatea, Mondragón, España, {dromero, mbarrenetxea, uayala}@mondragon.edu

<sup>2</sup> Instituto de Investigación Sanitaria Biocruces Bizkaia, Barakaldo, España

<sup>3</sup> IkerBasque, Basque Foundation for Science, Bilbao, España

## Resumen

La enfermedad de Parkinson (EP) es una de las enfermedades neurodegenerativas más prevalentes y es clave desarrollar nuevos biomarcadores para su diagnóstico. La tomografía por coherencia óptica (OCT) ha permitido identificar cambios en la retina de los pacientes con EP. Pese a lo conocido hasta ahora, de cara a una aplicación clínica real, es necesario obtener más evidencias respecto a la robustez de las características y modelos de diagnóstico basados en OCT. Por ello, en este artículo evaluamos la repetibilidad test-retest y el rendimiento de un total de 785 características derivadas de OCT en una cohorte de 158 pacientes y 158 sujetos sanos. Para el diagnóstico se hace uso de modelos clásicos como son la regresión logística y random forest. Los resultados muestran que sólo el 29.9 % de las características evaluadas presenta una alta repetibilidad. En cuanto al diagnóstico, el máximo AUC alcanzado es de 0.64, lo cual evidencia que existen cambios en la retina derivados de la EP, pero con un potencial limitado en cuanto al diagnóstico preciso.

## 1. Introducción

La enfermedad de Parkinson (EP) es la segunda enfermedad neurodegenerativa más prevalente con un número de pacientes que se espera que crezca hasta los 14.2 millones de personas a nivel mundial en 2040 [1]. Para dar una mejor respuesta a la enfermedad es clave desarrollar nuevos biomarcadores que permitan diagnosticar y pronosticar el avance de la enfermedad. En este sentido, en los últimos años parte de la labor investigadora ha centrado sus esfuerzos en explorar la retina mediante la tomografía por coherencia óptica (OCT, *optical coherence tomography*) [2], una técnica que permite analizar la retina a escala micrométrica de una forma no invasiva. Los resultados de estos estudios evidencian que la retina, compuesta por neuronas, se ve afectada en enfermedades neurodegenerativas [3]. En el caso de la EP, la literatura actual apunta hacia una disminución del grosor de las capas inferiores de la retina [4]. Pese a dichos esfuerzos, la aplicación clínica real de la OCT para la EP está lejos de ser una realidad en parte por la falta de resultados robustos y reproducibles.

Por ello, en este artículo hemos explorado la capacidad de diagnóstico de la EP y la repetibilidad de 785 características de OCT maculares. El principal objetivo es determinar si, mediante la combinación de características que describen diferentes aspectos de la mácula, es posible mejorar el diagnóstico de la EP. En un segundo lugar, se busca determinar la repetibilidad de dichas características exploradas y estudiar cuáles funcionan mejor.

## 2. Métodos

### 2.1. Base de datos

Un total de 158 pacientes con EP y 158 sujetos sanos pareados por edad y sexo fueron incluidos en el estudio (ver Tabla 1). Todos los sujetos dieron su consentimiento escrito para participar en el estudio. Los sujetos con EP fueron diagnosticados según los criterios de UK Brain Bank.

Grupo	Sujetos	Ojos	Edad*	% varones
Parkinson	158	302	64.7 ± 8.6 [41, 80]	64.6
Control	158	296	64.2 ± 8.7 [41, 87]	64.6

\* Media ± desviación estándar [rango de edad]

**Tabla 1.** Características demográficas de la cohorte

Ambos ojos de cada sujeto fueron examinados mediante un escáner OCT de dominio spectral Spectralis (Heidelberg Engineering, Heidelberg, Alemania) siguiendo un protocolo de adquisición macular tipo ráster con 25 B scans y 512 A-scans cubriendo una zona de aproximadamente 3 x 3 mm<sup>2</sup>.

Para poder evaluar la repetibilidad de las características extraídas 10 sujetos sanos fueron examinados en dos ocasiones distintas siguiendo el mismo protocolo de adquisición descrito previamente.

Las imágenes fueron revisadas y aquellos ojos con patologías oftalmológicas con afección macular (drusas, degeneración macular entre otras) fueron excluidos. Asimismo, no se incluyeron aquellos ojos con más de 4 dioptrías de miopía. Tras la exclusión se incluyeron los ojos descritos en la Tabla 1. Las capas de la retina fueron segmentadas mediante el software propietario del escáner Spectralis. La segmentación fue revisada y los errores fueron corregidos manualmente.

Los datos fueron exportados a MATLAB para su posterior análisis. Respecto al preprocesado, la fovea fue detectada automáticamente como el punto de espesor total mínimo tras suavizar el mapa de espesor bidimensional [5]. Los ojos izquierdos fueron volteados para hacer coincidir el sector nasal de ambos ojos.

## 2.2. Extracción de características

Con el objetivo de estudiar diferentes aspectos de la retina se extrajeron un total de 785 características de las siguientes categorías:

### Espesor sectorizado

Respecto a los espesores, se estudió el espesor total de la retina (ET), las capas ganglionares y plexiforme interna (GCIPL, *ganglion cell and inner plexiform layer*), la capa nuclear interna (INL, *inner nuclear layer*), las capas exteriores nuclear y plexiforme y el complejo delimitado por la membrana externa limitante y la membrana de Bruch. Para cada capa, se computó la media y la desviación estándar en toda la mácula (radio < 3 mm), los sectores convencionales ETDRS (*Early Treatment Diabetic Retinopathy Study*) y 5 anillos concéntricos de 1 mm de diámetro. Estas características han sido el principal objeto de estudio hasta la fecha [4].

### Textura del espesor

Además, se calcularon también diversas características de textura: dimensión fractal, lacunaridad, 24 características derivadas de la matriz de coocurrencia de nivel de grises y 7 del histograma obtenido mediante la técnica *local binary patterns*. Las características de textura no son convencionales, pero pueden aportar información sobre cambios de segundo orden en la retina [6,7].

### Geometría foveal

La morfología foveal fue analizada en base a los mapas de espesor ET y GCIPL. Cada uno de ellos fue transformado a un patrón radial con 24 direcciones angulares mediante interpolación. Posteriormente, se computaron un total de 14 características geométricas (pendiente media, pendiente máxima, altura, radio, profundidad, radio y altura en el punto de máxima pendiente, espesor en el centro de la fovea, área foveal, volumen foveal, área y perímetro del disco en el punto de máxima altura y máxima pendiente). Dichas características se promediaron a lo largo de las 24 direcciones angulares para obtener un solo valor por ojo. Las características foveales no han sido demasiado exploradas en la literatura, pero permiten describir cambios macroestructurales en la fovea.

### Modelado matemático de la fovea

De forma similar, se aplicaron 5 modelos matemáticos diferentes para modelar la superficie foveal [8-12]. Tras el ajuste, los coeficientes de dichos modelos fueron considerados como características. Un enfoque de este tipo ha sido empleado en Parkinson con buenos resultados [9].

### Reflectancia

Finalmente, se evaluó la reflectancia (intensidad de imagen) del ET y de la capa GCIPL. Para ello, se obtuvieron en primer lugar mapas de reflectancia bidimensionales promediando la intensidad de los vóxeles de cada capa de interés a nivel de A-Scan. A partir de este punto, se calcularon las mismas características sectoriales y de textura computadas para los mapas de espesor. Si bien ha habido estudios investigando la reflectancia [13,14] no es una característica habitualmente utilizada. La inclusión

de estas características se fundamenta en la hipótesis de que cambios estructurales en las capas podrían detectarse antes en la reflectancia que en una reducción del espesor.

## 2.3. Repetibilidad test-retest

Como paso de validación técnica, se evaluó la repetibilidad test-retest de cada una de las características extraídas en base al índice de correlación intraclase “*agreement*” (ICC-A). Para ello, se hizo uso del subgrupo de 10 sujetos sanos en los que se adquirieron dos imágenes consecutivas siguiendo el mismo protocolo.

## 2.4. Modelos de diagnóstico

En primer lugar, se llevó a cabo un análisis univariante en el que se evaluó el potencial diagnóstico de cada una de las características. Se hizo uso de un modelo de regresión logística y una validación cruzada con 5  *folds*. Por cada característica, se evaluó el área bajo la curva (AUC, *area under the curve*), precisión (% de casos diagnosticados correctamente), sensibilidad (% de casos con EP diagnosticados correctamente) y especificidad (% de verdaderos casos de EP del total de casos diagnosticados como EP) del modelo de diagnóstico entrenado.

En un segundo enfoque, se entrenó un modelo de diagnóstico tipo *random forest* donde se combinaron múltiples características. La estrategia empleada para entrenar, ajustar y evaluar el modelo se muestra en la Figura 1. Previo al ajuste de los modelos, se hizo una selección no supervisada de características eliminando aquellas redundantes (correlación de Pearson  $r > 0.8$ ).

A partir de este punto, se empleó una estrategia con dos validaciones cruzadas anidadas. La exterior (*outer CV* de 5  *folds*) fue empleada para evaluar el modelo mientras que la interior (*inner CV* con 10  *folds*) se usó para ajustar los hiperparámetros del modelo. Dado que ambos ojos de cada sujeto no son estadísticamente independientes en todos los pasos se incluyeron ambos ojos de cada sujeto en el mismo  *fold* y nunca se usaron ojos distintos para validación y entrenamiento (p. ej. nunca se usó el ojo izquierdo para entrenamiento y el derecho para validación).

Los pasos seguidos para el ajuste del modelo son los siguientes:

1. Selección de características (supervisada): se escogieron las 50 mejores características en base al algoritmo de mínima-redundancia-máxima-relevancia [15].
2. Ajuste de modelo: se usó un esquema de validación cruzada con 10  *folds* para ajustar los siguientes hiperparámetros: número de árboles, número de características consideradas en cada punto de decisión y el tamaño mínimo de nodo.

Finalmente, se evaluó la capacidad de diagnóstico del modelo usando los datos separados como test en la validación cruzada exterior.

El proceso descrito se un total de 100 veces con diferentes particiones de los datos (distintos  *folds*) para reducir la variabilidad de los resultados y obtener intervalos de confianza del 95%.

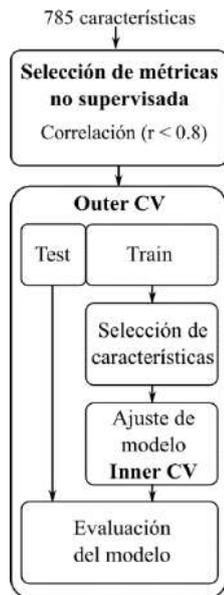


Figura 1. Estrategia de entrenamiento y evaluación del modelo multivariante de diagnóstico

### 3. Resultados

#### 3.1. Repetibilidad test-retest

La Figura 2 muestra la distribución de los valores de ICC para las características estudiadas para las tres categorías analizadas. El análisis reveló una gran dispersión en la repetibilidad de las características, con un 39.0 % de características con un ICC menor a 0.4 y un 29.9 % por encima de 0.8.

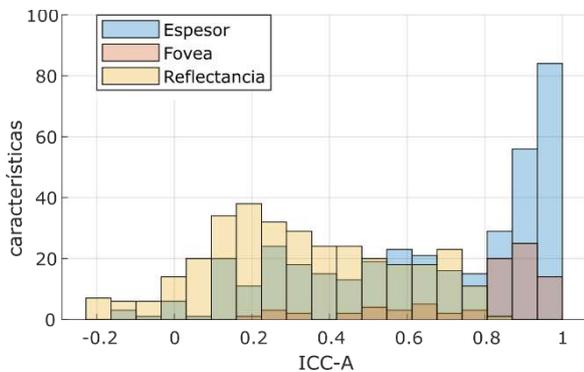


Figura 2. Distribución ICC-A (test-retest)

Resulta evidente a su vez que las características de reflectancia muestran en general una repetibilidad menor.

#### 3.2. Rendimiento diagnóstico

Como se puede observar en la Figura 3, la precisión y el AUC máximos fueron 61.1 % y 0.64, respectivamente. Tal y como se muestra en la Tabla 2, entre las características con mayor rendimiento destacan aquellas relacionadas con reflectancia media cuya repetibilidad es aceptable.

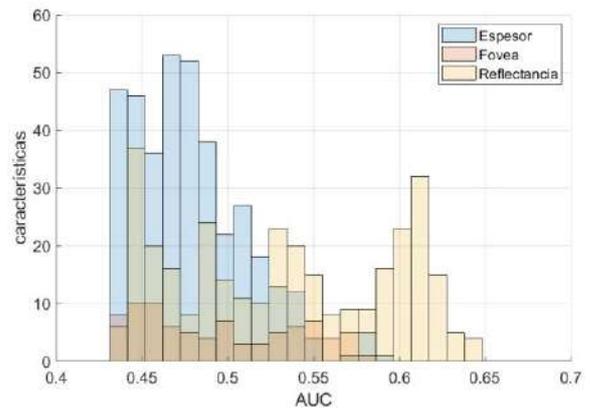


Figura 3. Distribución de la precisión y el AUC de las características evaluadas de forma univariante

Tipo	Capa	Característica	Prec.	AUC	Sen.	Esp.
Ref.	INL	Media anillo 2.5-3 mm	61.1	0.64	59.9	62.5
Ref.	ET	Media outer-superior	60.9	0.64	61.3	60.6
Ref.	ET	Media macula	60.4	0.63	60.6	60.2
Ref.	ET	Media anillo 2.5 -3 mm	60.4	0.64	59.6	61.1

\*Abreviaturas: reflectancia (Ref.), precisión (Prec.), sensibilidad (Sen.) y especificidad (Esp.)

Tabla 2. Características con mayor rendimiento

Respecto al análisis multivariante, la selección de características no supervisada descartó 504 redundantes ( $r > 0.8$ ) resultando en un total de 281 empleadas para entrenar el modelo tipo *random forest*. Como resultado del entrenamiento, dicho modelo obtuvo una precisión del 61.0 % [58.7, 63.7], no mejorando la capacidad de diagnóstico de las mejores características empleadas individualmente.

### 4. Discusión

En este artículo, se explora la capacidad diagnóstica de las características OCT de retina para la EP. Con el fin de caracterizar los diferentes aspectos de la retina se calcularon características de diferentes familias y complejidad.

La menor repetibilidad de las características de reflectancia puede venir derivada de la dificultad en la caracterización de ésta, muy afectada por factores como la iluminación o la dilatación pupilar. Del mismo modo, características complejas de textura como la dimensión fractal son probablemente muy sensibles a pequeños cambios en el centrado de la imagen o la segmentación.

Respecto al diagnóstico, la distribución univariante evidencia que existen numerosas características con bajo o nulo potencial para el diagnóstico. Asimismo, también se observa que existe cierto potencial para el diagnóstico

(AUC=0.66) aunque limitado. Que las características con mejor rendimiento sean en su mayoría relacionadas con la reflectancia podría estar relacionado con una peor calidad de imagen en pacientes con EP. Cabe destacar que, pese a que muchas de las características de reflectancia tienen una baja repetibilidad, aquellas que mejor rendimiento ofrecieron mostraron unos valores de ICC aceptables (ICC > 0.55)

Si bien los resultados obtenidos concuerdan con algunos estudios previos [16,17], se quedan lejos del rendimiento máximo reportado en la literatura (AUC ~0.9) [18]. Cabe destacar que en este último estudio se desarrolló un modelo basado en regresión logística y métricas de espesor convencionales que también han sido incluidas en el presente trabajo. Esta discrepancia puede venir dada por diferentes motivos. En primer lugar, en este estudio nos hemos enfocado en imágenes OCT maculares y no hemos incluido características relacionadas con la capa de fibras nerviosas alrededor del nervio óptico (OCT peripapilar). Estas métricas, sí incluidas en [18], podrían mejorar la precisión del diagnóstico. Por otro lado, los distintos resultados también podrían deberse a diferencias en el estadio de la enfermedad EP. Así, mientras en [18] la duración media de la enfermedad desde su diagnóstico fue de 7.6 años, en nuestro estudio es de 6 años.

En cualquier caso, un metaanálisis de 11 estudios concluyó que las diferencias en el espesor de la capa GCIPL entre sujetos sanos y pacientes de EP existe, pero es relativamente pequeña (d de Cohen = -0.4) [4]. Un efecto de este tamaño permitiría alcanzar una capacidad diagnóstica de un AUC del 0.61 (según resultados de simulaciones), más en línea con los resultados obtenidos.

Cabe mencionar las limitaciones de este estudio. En primer lugar, aunque la lista de características es extensa, puede que otras características no evaluadas pudieran mejorar el rendimiento. De forma similar, enfoques basados en aprendizaje profundo podrían también mejorar sustancialmente el diagnóstico. En este trabajo optamos por modelos clásicos debido a su simplicidad y al reducido tamaño muestral.

## Agradecimientos

Agradecemos a los participantes su participación en el estudio.

## Referencias

- [1] Dorsey ER, Bloem BR. The Parkinson pandemic—a call to action. *JAMA neurology*, vol 75, 2018, pp 9-10.
- [2] Huang D et al. Optical Coherence Tomography. *Science*, vol. 254, sup 5035, 1991, pp 1178-1181.
- [3] London A, Benhar I, Schwartz, M. The retina as a window to the brain—from eye research to CNS disorders. *Nature Reviews Neurology*, vol 9, 2013, pp 44-53.
- [4] Chrysou A, Jansonius NM, van Laar T. Retinal layers in Parkinson's disease: a meta-analysis of spectral-domain optical coherence tomography studies. *Parkinsonism & related disorders*, vol 64, 2019, pp 40-49.
- [5] Romero-Bascones D et al. Foveal pit morphology characterization: a quantitative analysis of the key methodological Steps. *Entropy*, vol 23, 2021.
- [6] Jáñez-García L, Bachtoula O, Salobrar-García E et al. Roughness of retinal layers in Alzheimer's disease. *Sci Rep*, vol 11, 2021, pp 11804.
- [7] Tazarjani HD, Amini Z, Kafieh R, Ashtari F, Sadeghi E. Retinal OCT Texture Analysis for Differentiating Healthy Controls from Multiple Sclerosis (MS) with/without Optic Neuritis. *Biomed Res Int.*, 2021.
- [8] Dubis A, McAllister J, Carroll J. Reconstructing foveal pit morphology from optical coherence tomography imaging. *British Journal of Ophthalmology*, vol 93, sup 9, 2009, pp 1223-1227.
- [9] Ding Y et al. Application of an OCT data-based mathematical model of the foveal pit in Parkinson disease. *Journal of Neural Transmission*, vol 121, sup 11, 2014, pp 1367-1376.
- [10] Scheibe P et al. Parametric model for the 3D reconstruction of individual fovea shape from OCT data. *Experimental Eye Research*, vol 119, 2014, pp 19–26.
- [11] Yadav S et al. CuBe: parametric modeling of 3D foveal shape using cubic Bézier. *Biomedical Optics Express*, vol 8, sup 9, 2017, pp 4181-4199.
- [12] Breher K, Agarwala R, Leube A, Wahl S. Direct modeling of foveal pit morphology from distortion-corrected OCT images. *Biomedical Optics Express*, vol 10, sup 9, 2019, pp 4815-4824.
- [13] Varga BE, Gao W, Laurik KL, Tátrai E, Simó M, Somfai GM, et al. Investigating Tissue Optical Properties and Texture Descriptors of the Retina in Patients with Multiple Sclerosis. *PLoS ONE*, vol 10, 2015.
- [14] Visser F, Vermeer KA, Ghafaryasl B et al. In vivo exploration of retinal nerve fiber layer morphology in Parkinson's disease patients. *J Neural Transm*, vol 125, 2018, pp 931–936.
- [15] Ding C, Peng H. Minimum redundancy feature selection from microarray gene expression data. *Journal of Bioinformatics and Computational Biology*, vol 3, 2005, pp 185–205.
- [16] Pinkhardt EH et al. The intrinsically restructured fovea is correlated with contrast sensitivity loss in Parkinson's disease. *J Neural Transm*, vol 127, 2020, pp 1275-1283.
- [17] Zou J, Liu K, Li F, Xu Y, Shen L, Xu H. Combination of optical coherence tomography (OCT) and OCT angiography increases diagnostic efficacy of Parkinson's disease. *Quant Imaging Med Surg*, vol 10, 2020, pp 1930-1939.
- [18] Garcia-Martin et al. Distribution of retinal layer atrophy in patients with Parkinson disease and association with disease severity and duration. *Am J Ophthalmol*, vol 157, 2014, pp 470-478.

# POWERUP: Design and preliminary evaluation of a 3D-printed upper-limb exoskeleton for children with cerebral palsy

C. Sánchez López de Pablo<sup>1</sup>, E. Urendes Jiménez<sup>1</sup>, S. del Riego Gómez<sup>1</sup>, A. Rojo Agustí, V. Costa Cortez<sup>1</sup>, R. Raya López<sup>1</sup>

<sup>1</sup> Departamento de Tecnologías de la Información, Universidad San Pablo-CEU, CEU Universities, Madrid, España, {cristina.sanchezlopezpablo, eloyjose.urendesjimenez, vanina.costacortez, rafael.rayalopez}@ceu.es {s.riego, ana.rojoagusti}@usp.ceu.es

## Abstract

*Cerebral Palsy (CP) is a disorder caused by a lesion in the immature brain. The most common impairments are related with movement and posture affecting both lower and upper limbs. Robotic devices are becoming popular as effective tools to improve or complement the rehabilitation process. This paper presents the POWERUP, an exoskeleton made of 3d-printed materials designed for upper limb rehabilitation of children with CP and similar disorders. We conduct a preliminary evaluation of its usability with 7 users with CP measuring the task success rate and the level of satisfaction. Finally, we conclude that the design fulfills the requirements to propose an affordable tool that could be easily took to the clinical practice in daily care centers to test the rehabilitation impact.*

## 1. Introduction

Cerebral Palsy (CP) is a neurological disorder caused by a brain injury at early ages, that mainly affects the mobility and posture of the child, limiting their autonomy. The injury may occur during gestation, childbirth, or the first three years of a child's life. Cerebral palsy is a permanent disorder, implying that the neurological lesion is irreversible and persists throughout life. However, early interventions are crucial to maximize the cognitive and physical skills. The incidence is 2-3 cases per 1000 births, [1].

There are four main types of cerebral palsy in terms of the individual's tone and posture. According to the brain region affected, CP is frequently divided into spastic, dyskinetic, ataxic and mixed. According to the body part affected, CP usually is classified as diplegia (lower limbs), hemiplegia (one body side) and quadriplegia (four limbs). Therefore, CP can affect essential functions of lower limbs: walking and running, and upper limbs: reaching and grasping.

Treatments for CP may vary from physical interventions (e.g. physical and occupational therapy, orthotics), to oral medication, neurosurgical interventions, and traditional orthopedic interventions. Some interventions with emergent level of evidence for upper limb are Robotics [2], Virtual Reality and Gaming [3].

Robotic rehabilitation proposes novel methods of physical intervention for neuromotor recovery. Robotic solutions

usually have in common: intensive training, weight support, playful scenarios with goal-oriented tasks and biomechanical metrics to assess the therapy progress. The common characteristic of weight-support has shown positive effects in reducing the abnormal coupling of shoulder abductors and elbow flexors, [4]. Some authors found that partial weight support led to greater improvements respect to full weight support for stroke survivors, [5]. Although first studies were focused on stroke survivors [6], the investigation over other populations as CP is becoming more studied.

Armeo™ (Hocoma) is one of the most popular devices for upper limb rehabilitation. It is a passive exoskeleton offering weight support through an adjustable spring mechanism, based on the T-WREX exoskeleton (Training-Wilmington Robotic Exoskeleton). Despite studies using Armeo™ with stroke population, only a few studies have tested it with children with CP, [7]. Chen et al. [2] evaluated the effectiveness of robotic devices for rehabilitation of upper limb in children with CP. They evaluated the InMotion2, the NJIT-RAVR, CosmoBot and REAPlan.

Some authors, [8], claim that more research is needed in these areas. These studies coincide that there is a research-practice gap that brings to certain degree of controversy. The lack of usability and not demonstrated cost-effectiveness of robotic-based interventions are the main barriers to integrate them in daily rehabilitation routines.

This work presents the design and evaluation of an upper limb exoskeleton made of 3D-printed materials, called POWERUP. The objective of this project was to create a usable and affordable solution for the use in the clinical practice, replicating the essential elements of similar systems that were demonstrated to be effective for the rehabilitation of upper limb.

## 2. The POWERUP platform

The POWERUP platform consists of an orthosis and a gamification environment for promoting the upper limb rehabilitation of children with CP and similar disorders. The POWERUP setup consists of an upper-limb

exoskeleton with an inertial sensor (ENLAZA sensor from Werium™) beneath the hand piece that controls a series of videogames. The goal of this setup is to create a usable environment in which the user can perform movements that serve as rehabilitation for limited motor functions and for the training and improvement of their range of motion in daily functional movements.

### 2.1. The 3d-printed exoskeleton

The POWERUP exoskeleton is a four-degrees-of-freedom (DoF) upper limb orthosis without electromechanical actuators. The present wearable prototype is designed for children aged 6 to 13 years old (arm + forearm length between 40 and 60 cm). Its function is to allow elbow and shoulder movements by providing stability, and to assist or resist flexion and extension of these joints by keeping the wrist and hand in the neutral position. The device is completely manufactured by 3D printing with PLA filament (a thermoplastic monomer derived from renewable and organic sources), which makes it a lightweight, low-cost, and easily reproducible solution.

The device can be adjusted in length and height to fit the physical characteristics of the children. POWERUP comprises four structural modules (shoulder/back, arm, forearm, and hand) (Figure 1). The use of bars, anchors, and elastic straps, allow the physiotherapist to easily adjust (width, height, and length) the device to the user's anthropometry. Furthermore, the modular design allows for simple and quick assembly.

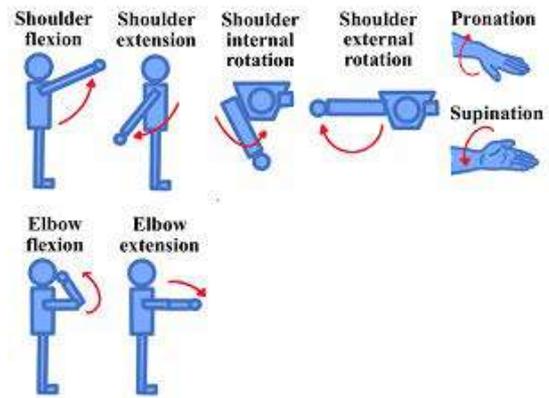
The exoskeleton provides the following 4 DoF (see Figure 2):

- Flexoextension and internal/external rotation of the shoulder
- Pronosupination of the wrist
- Flexoextension of the elbow

The hand module has a horizontal surface on which the hand rests and the wrist is fastened with straps, resting in a neutral position in the coronal plane. This surface rotates within a mechanism of concentric sliding rings, allowing pronation and supination of the wrist.



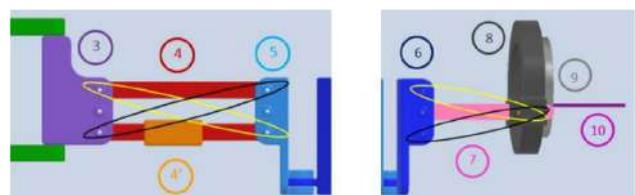
**Figure 1.** 3D-printed based passive upper-limb exoskeleton assembly. Segments: shoulder/back in blue, arm in red, forearm in orange, and hand in green



**Figure 2.** Degrees of freedom of POWERUP

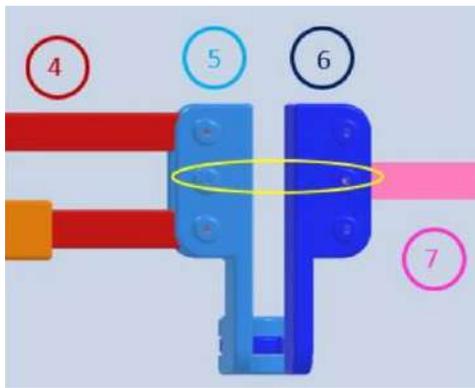
POWERUP has a dual function mechanism for elbow flexion and extension movement: assistance and resistance. The assistance mode provides anti-gravity weight support, assisting children in lifting and maintaining the weight of their arm, thereby improving flexion movement. As a resistive movement, the resistance mode opposes flexion, so children must exert extra effort to perform any activity that involves bringing the arm closer to the trunk. The goal of this mode is to provide a pro-gravity effect that promotes muscle tone gain or maintenance. A physiotherapist is responsible for and supervises the therapeutic prescription of the use of these modes.

This assistance-resistance method is implemented in the exoskeleton through lugs and elastic rubbers. The assistance or resistance mode is selected depending on which lugs the band is placed on. The exoskeleton can carry elastic rubbers to assist or resist shoulder flexion and extension (from piece 3 to piece 5 in Figure 3), as well as to assist or resist elbow flexion and extension (piece 6 to 9 in Figure 3). Elastic bands can also be used to prevent the elbow internal flexion that occurs naturally in many children with upper-limb motor disabilities (Figure 4).



**Figure 3.** Elastic rubbers positions in arm (left side) and forearm (right side). Assistance mechanism in yellow and resistance mechanism in black

The main purpose of the elastic rubbers is to allow the exoskeleton to evolve in tandem with the child's arm strength. If the user lacks the muscle strength to lift their arm during the first session, several rubber bands can be used to help. As the child requires less assistance, the number of rubber bands used may decrease as the sessions progress.



**Figure 4.** Elastic rubbers position for blocking elbow internal flexion

Once the user can lift and move their arm completely without assistance, elastic rubbers can be used in the resistance position. The system is composed of three segments (shoulder, arm, and elbow), whose modular design allows for quick and intuitive assembly. All of them can be easily adjusted to the child's anthropometry by the physiotherapist.

The POWERUP exoskeleton is anchored to an external metal rolling frame (see Figure 5) This structure is easily moveable and can be adjusted to the user's shoulder height. The wheels are locked once in a convenient position to prevent unintentional displacement during training sessions.



**Figure 5.** General view of POWERUP exoskeleton

## 2.2. The videogame

The Unity3D™ engine was used to create the POWERUP videogames. The goal of all videogames is to guide the avatar through targets that appear at various points on the screen, forcing the patient to perform elbow and shoulder rotational, flexion, and extension movements (Figure 6).

In each of the videogame's six possible scenarios, the child acts as the pilot of a vehicle (airplane, rocket, drone) or avatar (fairy, bird, basketball) that moves at a constant speed through a virtual scenario. When playing the game, targets appear at varying distances and heights from the avatar. The range of motion, that should be determined by the clinician in the configuration screen of the game, defines the positioning of these targets.



**Figure 6.** POWERUP videogames

The initial configuration of difficulty determines how quickly the targets appear. The number of targets of each session can also be selected by the clinician. This configuration of parameters makes it possible to set up a comfortable work area for the child, considering the specific situation of each patient. In addition, the game includes acoustic feedback, producing a sound when the child hits a target. This effect can help children with some degree of visual impairment increasing their motivation.

The videogame includes a database that stores the inertial sensor data from each exercise. When the game is finished, the videogame displays a new interface with a graph showing the angular evolution of the sensor and whether the patient has met the objectives.

## 3. Clinical experimentation

The clinical validation sessions were held at IRF La Salle (Madrid, Spain) during this first phase of clinical trials.

### 3.1. Participants

Participants were recruited by the physiotherapists at IRF La Salle according to the following inclusion criteria: children aged 6 to 13 year and with appropriate cognitive level to understand and perform the exercises and with some upper limb motor impairments.

Ten children volunteered to participate in the preliminary test using the videogame, but only seven children completed the testing sessions. Two of them were too small for the exoskeleton, and the third one chose not to continue due to the distance between their home and the center. Table 1 shows the characteristics of the children.

Subject	Age	Impairment	Gender
1	8	SMA	M
2	7	SMA	M
3	8	CP (Hemiparesis)	F
4	10	CP (Hemiparesis)	F
5	13	CP (Tetraparesis)	M
6	7	CP (Tetraparesis)	M
7	7	CP (Hemiparesis)	F

**Table 1.** Description of participants

### 3.2. Training sessions

Each clinical experimentation consisted of eight 50-minute sessions. They took place once or twice a week, depending on the child's physiotherapist schedule. The child sat comfortably in a chair or bench with a backrest but no armrests or wheels at the start of each session. A small platform was placed beneath them if their feet did not touch the ground, allowing their knees to be at 90°. Before putting on the exoskeleton, wrist orthosis, foot and ankle orthosis, and corset orthosis were placed if required to ensure the child was comfortable and in the correct position from the beginning of the session. The chair or bench was placed 1.5 meters away from a table with a tablet running the videogame.

### 3.3. Outcome measures

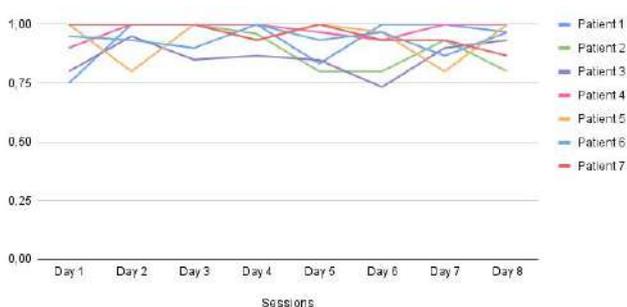
The metric used to evaluate the usability of the platform POWERUP was the success rate, that is the ratio of the number of targets hit by the patient over the total number of targets in the exercise. Additionally, the subjective opinion of the subjects was collected using the following satisfaction survey (0-5 points each question):

How satisfied are you with...

- The exoskeleton dimensions?
- The exoskeleton's weight?
- The ease with which the device's components can be adjusted?
- The device's safety and the possibility that the device will not harm you?
- The device's ease of use?
- The device's comfort?

## 4. Results and discussion

Throughout almost all eight sessions, patients achieved a hit ratio of 75% in at least one of the games performed (see Figure 7).



**Figure 7.** Success rate for each user and session

This value indicates that using this system gives the user the perception of control and the necessary movement assistance from the beginning, allowing them to exercise functional movements based on the maximum movement ranges of each joint. As the videogame difficulty and the number of elastic bands are adjusted to the child's needs, the maximum hit ratio remains constant, which gives the child a positive outcome as they see that even with increased difficulty, they can still hit most of the targets.

Although there is not a learning curve looking at the results, it is possible to see a generalized success for all users, which confirms the viability of using the system.

Regarding the satisfaction survey, the overall satisfaction result was 4.13, indicating the children's overall satisfaction. The exoskeleton's ease of use of the exoskeleton and its safety received the highest ratings, with an overall average of 4.4 and 4.5, respectively.

Finally, it is important to highlight that, in order to scale the presented solution to be able to generalize its use in clinical practice, it is important to (1) conduct a complementary clinical validation, with a larger number of patients and oriented to study their motor learning, generalization and retention of movements and functional tasks, (2) create new videogames more flexible and easy to adapt to the needs of each user and, (3) have an expert system, with a server in the cloud, that makes it possible to record and process the metrics and variables of each patient, for a better clinical follow-up of their evolution.

### Acknowledgement

This research was funded by MCIN/ AEI/ 10.13039/501100011033/ FEDER, UE, grant numbers RTI2018-097122-A-I00 and PID2021-127096OB-I00.

### Referencias

- [1] Oskoui M, Coutinho F, Dykeman J, Jetté N, Pringsheim T. An update on the prevalence of cerebral palsy: a systematic review and meta-analysis. *Dev Med Child Neurol.* 2013;55(6):509–19.
- [2] Chen Y-P, Howard AM. Effects of robotic therapy on upper-extremity function in children with cerebral palsy: A systematic review. *Dev Neurorehabil.* 2016;19(1):64–71.
- [3] Ravi DK, Kumar N, Singhi P. Effectiveness of virtual reality rehabilitation for children and adolescents with cerebral palsy: an updated evidence-based systematic review. *Physiotherapy.* 2017;103(3):245–58.
- [4] Coscia M, Cheung VCK, Tropea P, Koenig A, Monaco V, Bennis C, et al. The effect of arm weight support on upper limb muscle synergies during reaching movements. *J Neuroeng Rehabil.* 2014;11(1):22.
- [5] Ellis MD, Sukal-Moulton TM, Dewald JPA. Impairment-based 3-D robotic intervention improves upper extremity work area in chronic stroke: Targeting abnormal joint torque coupling with progressive shoulder abduction loading. *IEEE Trans Robot.* 2009;25(3):549–55.
- [6] El-Shamy SM. Efficacy of armo® robotic therapy versus conventional therapy on upper limb function in children with hemiplegic cerebral palsy. *Am J Phys Med Rehabil.* 2018;97(3):164–9.
- [7] Keller JW, van Hedel HJA. Weight-supported training of the upper extremity in children with cerebral palsy: a motor learning study. *J Neuroeng Rehabil.* 2017;14(1).
- [8] Novak I, Morgan C, Fahey M, Finch-Edmondson M, Galea C, Hines A, et al. State of the evidence traffic lights 2019: Systematic review of interventions for preventing and treating children with cerebral palsy. *Curr Neurol Neurosci Rep.* 2020;20(2):3.

# 5 years of the Master of Technological Innovation in Health (EIT Health): A critical analysis of its strengths and weaknesses

I. Oropesa<sup>1</sup>, G. Sou<sup>2,3</sup>, P. Carneiro<sup>4,5</sup>, J. Sanches<sup>5</sup>, M. J. Silva<sup>6</sup>,  
R. Rodríguez<sup>7</sup>, N. Vidal<sup>7</sup>, S. Kotzki<sup>8</sup>, E.J. Gómez<sup>1,9</sup>

<sup>1</sup> Grupo de Bioingeniería y Telemedicina, ETSI Telecomunicación, Centro de Tecnología Biomédica, Universidad Politécnica de Madrid, Madrid, España, {i.oropesa, enriquejavier.gomez}@upm.es

<sup>2</sup> CNRS, Laboratoire de Génie Electrique et Electronique de Paris, Sorbonne Université, 75005 Paris, France

<sup>3</sup> Université Paris-Saclay, Centrale Supélec, 91192, Gif-sur-Yvette, France, gerard.sou@sorbonne-universite.fr

<sup>4</sup> Instituto de Investigação e Inovação em Saúde & Institute of Molecular Pathology and Immunology of the University of Porto, Department of Pathology, University of Porto, Porto, Portugal, pcarneiro@i3s.up.pt

<sup>5</sup> Department of Bioengineering, Institute for Systems and Robotics, Instituto Superior Técnico, University of Lisbon, 1049-001, Portugal, jmrs@tecnico.ulisboa.pt

<sup>6</sup> INESC-ID, Instituto Superior Técnico, University of Lisbon, Lisbon, Portugal, mario.gaspar.silva@tecnico.ulisboa.pt

<sup>7</sup> Departamento de Ingeniería Electrónica y Biomédica, Universidad de Barcelona, Barcelona, España, {romen.rodriguez, nvidal}@ub.edu

<sup>8</sup> Grenoble-Alpes University, Grenoble 38000, France, sylvain.kotzki@univ-grenoble-alpes.fr

<sup>9</sup> Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina, Madrid, España

## Abstract

*The Master of Technological Innovation in Health (MTIH) is an EIT Health-sponsored and EIT-labelled programme oriented to training innovation and entrepreneurship (I&E) skills applied to biomedical engineering (BME). Training is based on a learning-by-doing approach to solve real-life problems. Students must carry out an international mobility, and a non-academic internship towards the completion of their master thesis. In this study we analyse the first five years of the programme. Analysis is based on Key Performance Indicators (KPIs) and a student survey. KPIs show the programme runs in good health, with high demand and a sustained turn-out of graduates. The survey shows overall student satisfaction, especially when it comes to internships; although opportunity for improvement remains in the programme administration, coordination of mobilities and in opportunities to put skills into practice.*

## 1. Introduction

In 2015, the members of the EIT Health Working Group on Technological Innovation in Health created a Consortium to develop a common Master programme, comprised of five European universities: Sorbonne Université (SU), Universidad Politécnica de Madrid (UPM), Universidade da Lisboa (UL), Universitat de Barcelona (UB) and Université Grenoble Alpes. It was supported by non-academic players such as MEDICEN, Medtronic Portugal and GE Healthcare Iberia.

The Master programme was designed to address the needs and opportunities of the biomedical engineering (BME) market, including (1) an increased societal, institutional and market demand for a well-trained BME workforce; (2) a receptive environment with funding opportunities for the creation of start-ups and the consolidation of innovative small and medium-sized enterprises (SMEs) [1]; (3) an increasing demand driven by an ageing

population, efficient health expenditure and safe and reliable health services; (4) an expanding field of technological opportunities; (5) a highly innovative field, with the highest number of generated patent applications, and one with a short product lifecycle; and (6) an increasing interest in BME education.

Thus, the Master of Technological Innovation in Health (MTIH) was born to support the emergence of a new generation of MedTech talents, products and services to solve health problems and to improve European citizens' health. The goals of the programme can be summed as:

1. Train students in technologies applied to healthcare,
2. Develop students' innovation and entrepreneurship (I&E) initiative,
3. Assure students' international mobility in both academic and industrial context,
4. Support students in their entrepreneurial journey (start-up creation).

The first student cohort was admitted in September 2017. In 2018, the programme obtained the EIT Label, a certificate of quality and excellence for educational programmes focused on innovation, entrepreneurship, creativity and leadership. The Label is awarded to degree programmes according to five quality criteria:

- EIT Overarching Learning Outcomes (OLOs) [2]
- Robust entrepreneurship education
- Integrated, innovative 'learning-by-doing' curricula
- Mobility
- Outreach strategy and access policy

The EIT Label enabled funding from the EIT Health Knowledge Innovation Community (KIC) from 2018 to

2021, which was mostly used to fund student mobility. The Label was awarded for a period of four years (2018-2022) and is currently being re-evaluated by EIT. In this paper we analyse the experience from the first 5 years of MTIH, including the year prior to obtaining the Label.

## 2. Programme description

MTIH is a joint programme embedded within local master programmes of the 5 Consortium universities: Engineering for Health at SU, Biomedical Engineering at Instituto Superior Técnico-UL and UPM, Innovation and Entrepreneurship in Biomedical Engineering at UB and Health Engineering and BioHealth at UGA.

The programme defines a student track comprised of the following elements:

- A mandatory I&E pillar (30 ECTS) providing students skills to design and implement innovative business solutions to specific health issues.
- A mandatory internship (15 ECTS) in a non-academic setting while working towards their Master thesis which is co-supervised by academic and non-academic advisors. Students must be able to secure their own the internships.
- A mandatory international mobility semester (30 ECTS) at a partner university.
- A mandatory biomedical engineering pillar: Medical Devices (SU), Biomedical Engineering (UL), ICT for Healthcare (UPM), I&E in Biomedical Engineering (UB), Health Engineering (UGA).

MTIH does not treat I&E content as a separate focus of the curriculum but ensures that the I&E added value is embedded in technical courses. The programme is designed as an innovative learning-by-doing, student-centred multidisciplinary program. Students work in small groups, conduct case studies and develop projects. They interact with stakeholders, identifying problems and proposing solutions, from the idea to the business plan. This way, MTIH delivers skills that foster students to make value and sustainability judgments, boost creativity, I&E, intellectual transformation and leadership.

Non-academic partners are strongly involved in the teaching activities. They are active in learning activities such as case studies, seminars, and giving talks related to their field of activity. They are also involved in mentoring/coaching activities. MTIH learning-by-doing methodology means that students have direct contact with non-academic stakeholders, both to define the needs and to learn the necessary I&E skills. This helps them build a contact network for their future careers.

The programme is organized over 2 years. Partner universities have defined MTIH student tracks composed of 2/3 academic semesters (30 ECTS each) + 1 semester for the Master thesis (30 ECTS). The discrepancy in semesters is due to the fact that in Spain the Bachelor degree is obtained after 4 full years of studies (240 ECTS), one more year than in France and Portugal (180 ECTS). Spanish students enter the master programme with heavier technical skills than students from France or

Portugal. Thus, MTIH in Spain can be shorter (1.5 years, 90 ECTS) than in France or Portugal (2 years, 120 ECTS) for an equivalent set of competencies and knowledge reached at the end.

At the end of the programme, students receive both their master's degree title (obtained locally upon completing the home university programme) and a Certificate of Completion issued by EIT Health, which testifies that the student has completed the training towards the acquisition of the competences described by the EIT OLOs.

## 3. Quality indicators

To analyse the quality of the programme, we will consider two different information sources.

### 3.1. Key Performance Indicators

Objective analysis is based on Key Performance Indicators (KPIs) co-defined with EIT and EIT Health:

- Programme attractiveness: number of applicants to MTIH / available spots (%).
- Start-ups: n° of start-ups created by MTIH students.
- Programme exchange: number of exchanges between sectors (industry, research, education, healthcare), including internships, study visits, etc.
- Number of students enrolled.
- Number of students trained in I&E by KIC, including non-MTIH students.
- Number of MTIH graduates.

KPIs are analysed yearly by the KIC and benchmarked against expected targets. In this study we analyse the results of the last natural year (2021), as well as the total accumulation of the past 5 years.

### 3.2. Student survey

In 2021 we designed a survey for graduates to measure student satisfaction with the programme. The survey is structured into the following structure:

- General questions: year of enrolment, graduation and employment status.
- Learning: overall rating of the programme (rated 0-10) and self-perception questions on skill acquisition.
- Remembrance: Likelihood of using skills, rated 0-10.
- Motivation: The will to pursue further EIT Health activities and/or to recommend the programme.
- After training: Confidence in putting into practice the skills acquired in the programme.

The survey was submitted by Consortium participants to former programme graduates and is currently being sent to students when they complete the programme.

## 4. Results and discussion

### 4.1. Key Performance Indicators

Table 1 provides an overview of the KPIs achieved after 5 years (2017-2022). Figure 1 reflects the evolution of

student-related KPIs. Overall, the programme has remained successful and met target values set by the KIC.

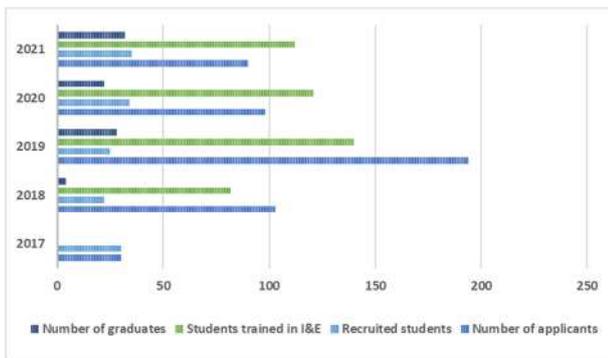
Programme attractiveness has remained high and above baseline values. Demand for MTIH peaked in 2019. In 2020, due to the COVID-19 pandemic, the demand decreased, given the uncertainty related with mobilities. This uncertainty carried over into 2021, coupled with the fact that students moving in 2022 no longer received funding from EIT Health, but depended on substantially lower Erasmus grants. Despite this, the number of new recruits has steadily increased in the last 4 years.

**Table 1.** KPI reporting

KPI	Target (2021)	Achieved value (2021)	5-year total
Programme attractiveness (%)	150	257	378*
Start-ups created	1	0	2
Programme exchange	60	46	199
Students enrolled	-	35	146
Students trained in I&E	60	116	455
Number of graduates	33	32	86

\* Reported as average

The success of the programme can be measured both by its outcomes and by its outreach. In terms of outcomes, by December 2021 the programme had graduated 86 students. In terms of outreach, the programme has managed to provide training in I&E applied to BME to over 450 students. This number includes students from local master programmes who take courses on this topic.



**Figure 1.** MTIH Student evolution

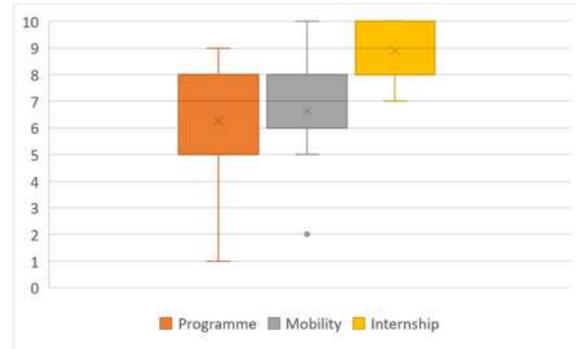
EIT Health strives to empower innovators to take action to put innovative products and services into the market. As such, creation of start-ups by KIC-trained engineers is an essential KPI. In this sense, the fact that 2 MTIH students have founded their own MedTech companies is an important success story for the programme.

#### 4.2. Student survey

To this day, 16 students have answered the survey. Five responses were removed on account of belonging to students who had not graduated yet, which might have been submitted by error. Of the rest, 1 student graduated in 2018, 2 in 2019, 5 in 2021 and 3 in 2022.

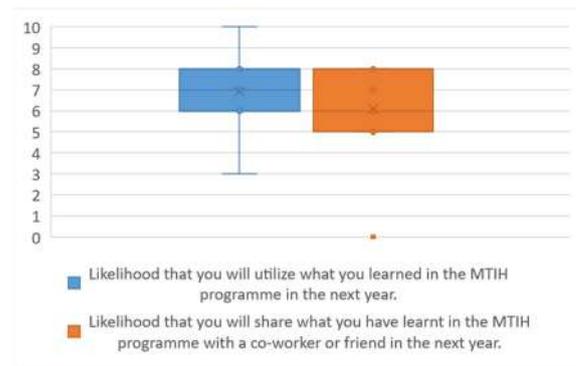
Overall, students rate MTIH positively at 6.3±2.6 (Figure 2). Dissatisfaction mostly comes from administrative

issues, such as the late release of mobility funds by EIT, or communication / management problems, reflecting the fact that MTIH is run by teaching staff with no additional resources. Moreover, a significant number of respondents were affected by the COVID-19 pandemic during their mobility, which had an impact on their evaluation. The aspect best rated by students are the internships in non-academic settings (8.9±0.9). Moreover, 72% of the students would further recommend the programme.



**Figure 2.** Distribution of rates to common programme aspects

In terms of learning (Figure 4, top row), 82% of students feel they have at least a solid understanding of I&E concepts. The same percentage have a similar view of competences acquired in mobility. To some extent, all students were able to put competences into practice during the internship, although 9% feel that this was not very often.



**Figure 3.** Remembrance questions' distributions

The likelihood of putting skills into practice in the next year was rated 6.9±1.8, which suggests that not all students are in a position to or are willing to focus on I&E (Figure 3). When analyzing after-learning effects (Figure 4, bottom row), 64% of students show self-confidence in putting their I&E skills into practice, and 27% feel ready to build their own company. To raise these numbers, students should require more opportunities and autonomy to put into practice their skills, for instance as part of their internship. Another aspect may be related to a lack of training in regulatory aspects, which may deter students to take the step of becoming entrepreneurs. In this sense, students could be encouraged to further take part in KIC acceleration activities to push forward their ideas. Most of them (91%) are keen to check out new KIC activities, which suggests a motivation for continuous improvement in I&E applied skills.

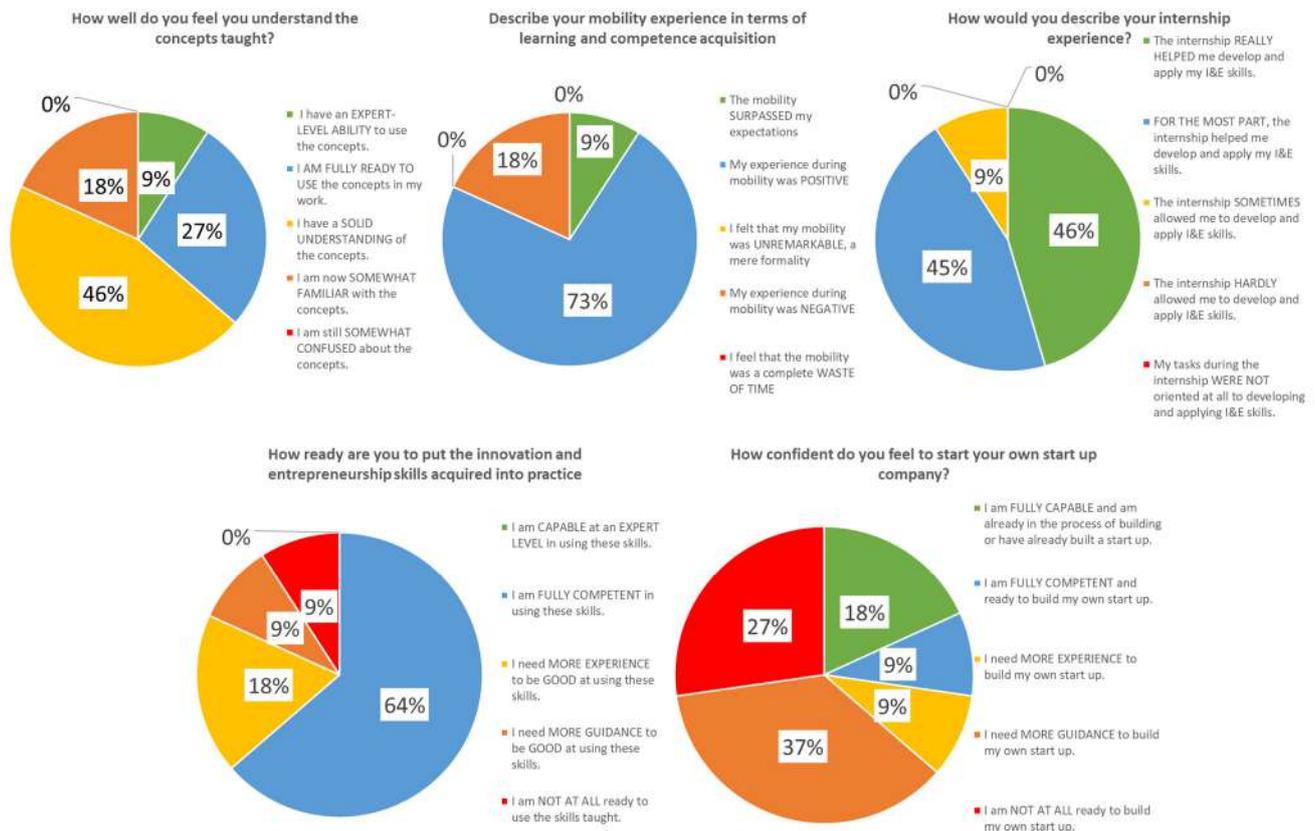


Figure 4. Selected pie charts. Top row: Learning-related questions. Bottom row: After training questions.

#### 4.3. Study limitations

There were several main limitations in this study. Concerning KPIs, the lack of a robust tracking system, combined with the fact that reporting in EIT Health is carried out annually instead of per academic year, may result in small deviations in the numbers reported. This issue is being addressed with the KIC, as EIT has defined its Impact Framework to measure programme outcomes for 2022-2027, which should ease the process in the future. The second limitation concerns the low number of survey respondents. The survey was only developed after the 4<sup>th</sup> year, and even though it was sent retroactively, the response rate was expectably low. We expect this to change in the future as more graduates answer the survey. New alumni surveys will also be planned to track their career progression.

### 5. Conclusions

After five years, the MTIH remains an attractive offer for BME students seeking to train I&E skills, with a high demand and a sustained turn-out of graduates. The main strength of the programme relies on its internship opportunities. Some weaknesses can be identified in the mobility organization, as well as in the need to further promote an entrepreneurial mindset in the students. To this end, the KIC is furthering ReCog, a programme to support established degrees and help students consolidate their career pathways towards entrepreneurship. An example of this is the future WorkInHealth Foundation

initiative which will provide opportunities to match talent with employers [3]. The programme is currently in the process of renewing the EIT Label, which will be awarded for an unlimited period if evaluated positively.

### Funding & acknowledgements

MTIH is part of an activity that has received funding from EIT Health. EIT Health is supported by the European Institute of Innovation and Technology (EIT), a body of the European Union that receives support from the European Union's Horizon 2020 Research and Innovation programme.

The authors would like to thank all local University staff and all KIC members supporting the programme.

### In memoriam

The authors would like to dedicate this paper to the memory of Prof. Philippe Sabatier (1953-2021), and his invaluable contribution in setting up and running MTIH.

### References

- [1] Maria IE, Dirk J. Biomedical engineering and care services. European Economic and Social Committee, 2015 (CCMI/128-EESC-0000).
- [2] EIT Learning Outcomes. <https://eit.europa.eu/our-activities/education/eit-learning-outcomes> (Consulted: July 2022)
- [3] WorkInHealth Foundation. <https://eithealth.eu/what-we-do/workinhealth-foundation/> (Consulted: July 2022)

# Efecto de la distancia sonda – muestra en la aplicación de la espectroscopia de reflectancia difusa

S. Buendia Aviles<sup>1,2</sup>, M. Cunill Rodríguez<sup>1</sup>, J. A. Delgado Atencio<sup>1</sup>, M. Muñoz Pérez<sup>1</sup>, J. L. Arce Diego<sup>2</sup>, F. Fanjul Vélez<sup>2</sup>

<sup>1</sup>Grupo de Óptica Biomédica, Universidad Politécnica de Tulancingo, Tulancingo de Bravo, Hidalgo, México, {sonia.buendia.1731023, margarita.cunill, jose.alberto, magda.munoz}@upt.edu.mx

<sup>2</sup>Grupo de Ingeniería Biomédica, Departamento TEISA, Universidad de Cantabria, Av. De los Castros 46, Santander, España, {arcedj, fanjulf}@unican.es, sonia.buendia@alumnos.unican.es

## Resumen

*La espectroscopia de reflectancia difusa (ERD) es una de las técnicas ópticas mínimamente o no invasivas empleadas por muchos autores por su facilidad de aplicación a nivel clínico. Entre sus aplicaciones se encuentra el diagnóstico de patologías diversas, en especial la detección de lesiones tumorales. En este trabajo se investiga la influencia de la distancia sonda-medio turbio en la forma y valores del espectro de reflectancia difusa registrado con una sonda comercial de fibras ópticas en fantomas de plastilina de tres colores: rojo, azul y verde. El montaje óptico experimental implementado estuvo conformado por una fuente de luz, un espectrómetro de fibra óptica y una sonda de fibras ópticas colocada ligeramente en contacto y perpendicular a la superficie de los fantomas. La distancia sonda-fantoma se varió mediante un mecanismo de desplazamiento micrométrico espacial, y se seleccionaron específicamente tres distancias: 1, 2 y 5 mm. Este estudio tiene un impacto directo en la aplicación clínica de la espectroscopia de reflectancia difusa, ya que en el entorno clínico es habitual que se produzcan variaciones en las distancias de aplicación, lo que puede tener una gran influencia en el resultado diagnóstico de la técnica.*

## 1. Introducción

El estudio y aplicación de las metodologías ópticas no invasivas ha sido de gran interés en las últimas décadas especialmente por su amplio campo de aplicación. En particular la espectroscopia de reflectancia difusa (ERD) implementada mediante el uso de sondas ópticas ha sido utilizada en variadas aplicaciones biomédicas [1–5], ya que es relativamente sencilla y de menor complejidad para su implementación. La ERD se enfoca en el análisis de los espectros de reflectancia que se obtienen a través de la radiación electromagnética reflejada por una superficie como producto de la luz que incide sobre un medio, después de propagarse a través de este [6, 7]. Normalmente esta reflectancia puede ser de dos tipos: reflectancia especular o reflectancia difusa.

A partir del procesamiento de la información contenida en la luz reflejada difusamente de una superficie o tejido es posible inferir el estado y composición de una muestra, por esta razón, es importante registrar correctamente los espectros para evitar perder información de las muestras o en caso contrario, limitar la introducción de información ajena a las propiedades de la muestra. Dado que el cálculo del porcentaje de reflectancia (%R) se basa en una relación

entre un estándar y la medición de la muestra, ambos a una distancia específica, la más mínima modificación haría que todo el espectro cambiara. Es por ello por lo que la aplicación de la ERD implica diversas consideraciones y efectos a fin de garantizar la veracidad de los datos obtenidos. Por ejemplo, el efecto de la presión ejercida en tejidos *in-vivo* [8–11] y *ex-vivo* [12, 13], el ángulo de incidencia de la sonda [14–16], la distancia sonda-muestra [15, 17–20], la eficiencia de recolección en dependencia de la geometría de la sonda, y la distancia de detección ( $r$ ) [21–24], pueden modificar significativamente el espectro adquirido. El análisis de la influencia de la distancia sonda-muestra ha sido abordado para intervalos de distancias (10–1000  $\mu\text{m}$ ) [14–16] donde es más frecuente la incidencia de este tipo de error durante la toma de espectros por parte de un operador humano, y menos frecuentemente reportado para distancias superiores a 1 mm. En este trabajo se evalúa el efecto de la distancia sonda-muestra para tres distancias sonda-fantoma: 1, 2 y 5 mm, categorizadas como distancias notables al ser más perceptibles, por lo que se espera sean de menor frecuencia en la práctica, pero determinantes para delimitar el rango máximo permitido de la distancia sonda-muestra durante la toma de espectros para muestras inertes y en base al montaje propuesto.

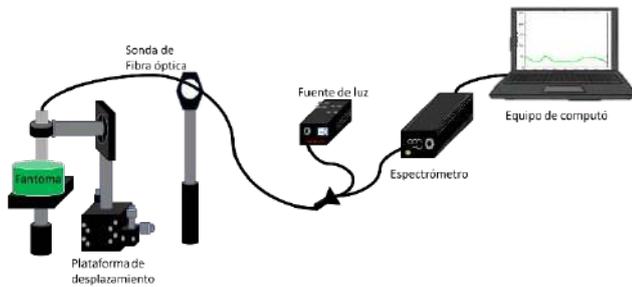
## 2. Materiales y métodos

### 2.1 Montaje experimental para el registro de los espectros

La Fig. 1 muestra el montaje óptico experimental para llevar a cabo las mediciones de los espectros de reflectancia difusa, los cuales se registran mediante un espectrómetro de fibra óptica (BLK-CXR-SR-50, StellarNet Inc., USA) al cual se le acopló, en su puerto óptico de entrada, uno de los extremos de una sonda bifurcada de fibras ópticas (R600-8-UV/VIS-SR, StellarNet Inc., USA). El otro extremo de la sonda se acopló a una fuente de luz de deuterio-tungsteno (Hamamatsu L10671, Hamamatsu Photonics, Japan) que presenta una emisión espectral en la región UV-visible-IR (200–1600 nm). El extremo común de la sonda se colocó en ligero contacto con la superficie de los fantomas de plastilina durante las mediciones espectrales. Para ello los fantomas se posicionaron debajo del extremo común de la sonda mediante una plataforma plana metálica. La

distancia sonda-superficie de los fantasmas se varió por medio de una plataforma micrométrica con desplazamientos XYZ (RB13M/M, Thorlabs, Newton, NJ, USA).

Los detalles cuantitativos y de la geometría de la sonda de fibras ópticas utilizada, así como el procedimiento y características de los fantasmas implementados, corresponden a los reportados en [19]. Para ello se emplearon tres fantasmas de 180 gramos elaborados a base de plastilina azul, verde y roja en forma de cilindros de 6 cm x 3.5 cm.



**Figura 1.** Instalación experimental implementada en la toma de espectros de reflectancia difusa en fantasmas de plastilina.

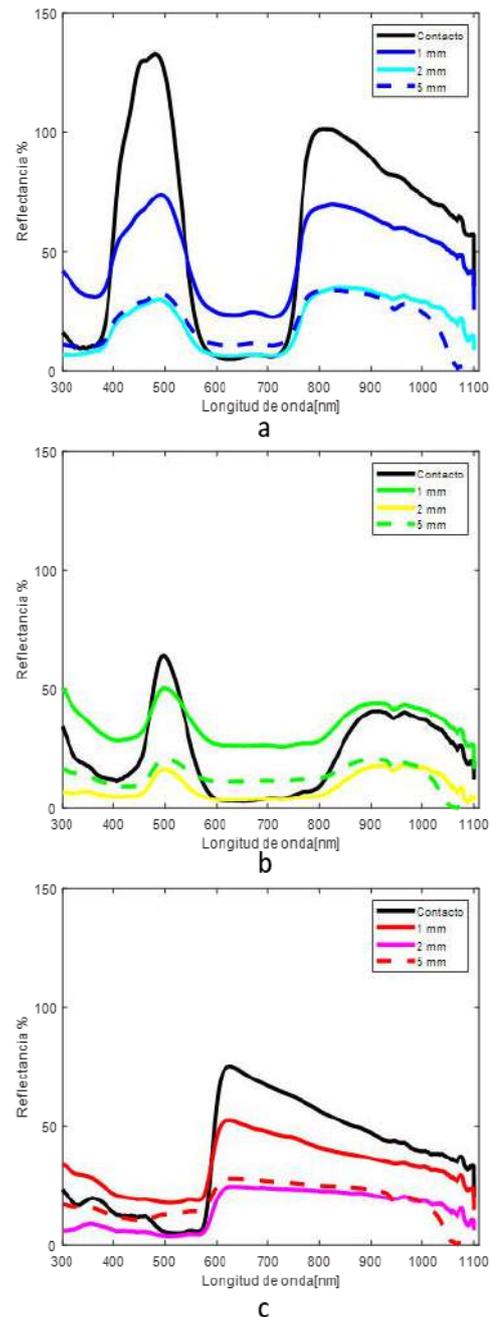
## 2.2 Registro de los espectros de reflectancia difusa

Los espectros de reflectancia difusa fueron adquiridos con un tiempo de integración de 500 milisegundos, 5 lecturas promediadas y 4 píxeles de suavizado. Los parámetros de adquisición se mantuvieron fijos en todas las mediciones. En total se efectuaron 4 mediciones para cada distancia establecida y fueron promediadas para su presentación gráfica. Primero se capturaron las curvas espectrales en contacto con la superficie de cada plastilina, colocando la sonda gentilmente en el centro de ellas y posteriormente la sonda se desplazó verticalmente hacia arriba manipulando el mecanismo de desplazamiento lineal a lo largo del eje Z de la plataforma micrométrica estableciendo los valores 1, 2 y 5 mm.

## 3. Resultados y discusión

En las Fig. 2 aparece el resultado de los espectros de reflectancia difusa obtenidos experimentalmente para los tres tipos de fantasmas (azul, verde y rojo) cuando la distancia sonda-fantomas se estableció en los tres valores especificados previamente (1, 2 y 5 mm), además del espectro promedio obtenido para las mediciones realizadas en contacto. Se observa que la firma espectral (la forma y posición característica) de las curvas de reflectancia difusa para un mismo tipo de fantoma es muy similar para las tres distancias sonda-fantoma utilizadas. Es importante resaltar que la forma de estas curvas es semejante a las obtenidas en [19] para otro conjunto de distancias sonda-fantomas (10, 20, 30, 50, 100, 500  $\mu$ m), motivo por el cual es posible observar las bandas espectrales características de cada uno de los fantasmas. Como puede apreciarse los máximos de reflexión difusa de los tres tipos de fantasmas concuerdan

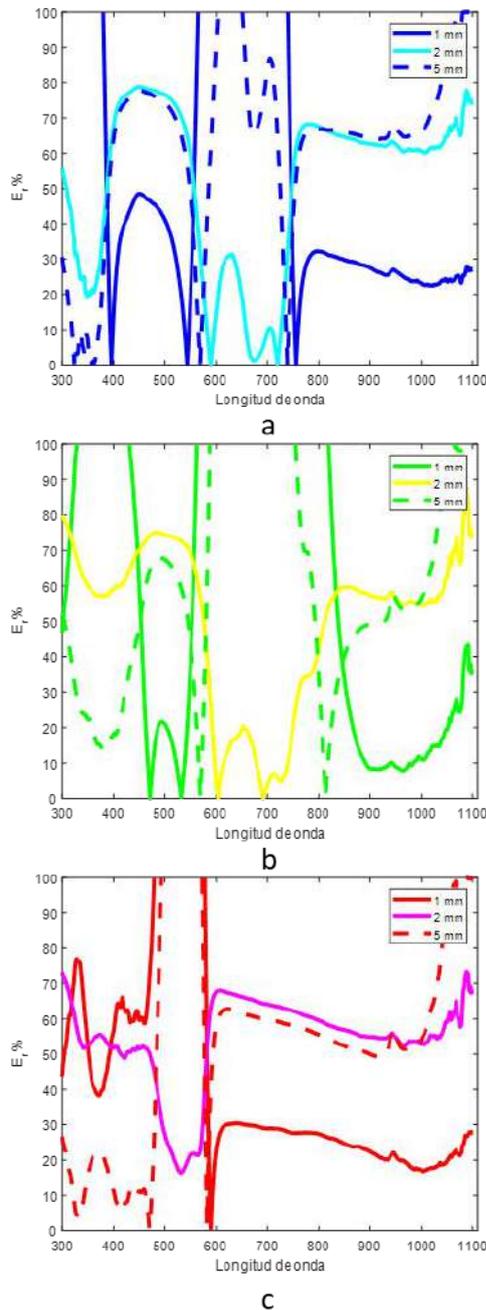
bastante bien con los máximos de las curvas espectrales de reflexión de pigmentos de color azul, verde y rojo [25].



**Figura 2.** Curvas espectrales obtenidas a las distancias sonda-muestra de 1, 2 y 5 mm para tres fantasmas de plastilina: a) espectros correspondientes a la plastilina azul, b) espectros correspondientes a la plastilina verde y c) espectros correspondientes a la plastilina roja.

Si bien la distancia sonda-muestra no afectó la forma de los espectros, no fue así con el porcentaje de reflectancia. Como se puede apreciar la mayoría de los casos registró una menor reflectancia en comparación a la registrada en contacto con la muestra, especialmente en las bandas características para cada plastilina [19]. De manera semejante se observó generalmente un menor porcentaje de reflectancia difusa (%R) para una separación sonda fantoma de 2 mm, contrario a lo esperado, dado que se podría suponer una disminución del %R conforme al

incremento en la distancia sonda-fantoma. No obstante, el %R obtenido a 5 mm fue siempre ligeramente superior al registrado a 2 mm, pero inferior respecto de las mediciones realizadas en contacto y a 1 mm. Esto podría ser debido a las propiedades de esparcimiento de estas muestras y su relación con la apertura numérica de las fibras ópticas.



**Figura 3.** Gráficas correspondientes al cálculo del error relativo porcentual de las mediciones realizadas a una distancia sonda-muestra de 1, 2 y 5 mm respecto a las curvas espectrales obtenidas en contacto: a)  $E_r$  (%) para la plastilina azul, b)  $E_r$  (%) para la plastilina verde y c)  $E_r$  (%) para la plastilina roja.

Con la finalidad de corroborar los resultados anteriores se realizó el cálculo del error relativo porcentual ( $E_r$  (%)). Este análisis confirma que a mayor distancia entre la muestra y la superficie de medición el porcentaje de error relativo porcentual ( $E_r$  (%)) aumenta significativamente, aunque existe una mínima diferencia entre los valores

obtenidos a 2 y 5 mm, siendo ligeramente inferior el  $E_r$  (%) para 5 mm. Estos resultados se corresponden con el análisis de las gráficas de las curvas espectrales previamente descrito. Por lo tanto, en este estudio se obtuvieron errores relativos porcentuales de magnitudes elevadas respecto a las mediciones realizadas en contacto. Un aspecto relevante es que para las tres muestras se obtuvo un menor error relativo porcentual para la región del infrarrojo cercano respecto de la región visible, sin embargo, estos aún continúan siendo valores considerables.

Por otra parte, se obtuvo que los valores del  $E_r$  (%) fueron siempre menores para la plastilina roja respecto a los otros dos fantasmas, tal y como es posible observar en la Figura 3-c), cuya banda espectral característica concuerda con la región de las longitudes correspondientes al infrarrojo cercano del espectro electromagnético.

En algunos trabajos previos se identificó la existencia de un crecimiento monótonico del %R para distancias sonda-piel entre 0 y 1000  $\mu\text{m}$ . Sin embargo, este incremento no es igual para todas las regiones anatómicas de medición [18], mientras que otros autores coinciden que este crecimiento es perceptible solo para mediciones de distancia sonda-tejido inferiores a 500  $\mu\text{m}$ , y decrecen para distancias superiores [20]. De manera semejante los resultados obtenidos en este trabajo en comparación a los previamente obtenidos y reportados en [19] muestran que existe un crecimiento significativo del %R en las mediciones espectrales realizadas sobre fantasmas de plastilina cuando se varía la distancia sonda-fantoma entre 0 y 50  $\mu\text{m}$ , siendo a 50  $\mu\text{m}$  donde se obtuvo el máximo %R para todas las muestras de plastilina usadas. Sin embargo, para distancias inferiores a 30  $\mu\text{m}$  y superiores a 1 mm el %R que se obtiene para la plastilina roja es inferior a la obtenida en contacto, comportamiento que no ocurre en las plastilinas verde y azul.

Los resultados presentados permiten corroborar que, si bien la distancia sonda-muestra para superficies inertes no altera la firma espectral de la muestra, el porcentaje de reflectancia difusa obtenido es susceptible a estas variaciones, tal como se reporta para rangos de distancias sonda-muestra inferiores en muestras inertes [19]. En consecuencia, los resultados de este trabajo contribuyen a ampliar el rango de distancias bajo análisis, elemento fundamental en la práctica clínica con objeto de obtener medidas que permitan realizar un diagnóstico adecuado basado en espectros de reflectancia.

#### 4. Conclusiones

En este estudio los resultados muestran que la distancia entre la sonda multi-fibra usada en nuestro montaje experimental garantiza que se mantenga la firma espectral de los fantasmas de plastilina azul, verde y roja implementados en este trabajo. Sin embargo, los errores introducidos por un operador para la distancia sonda-muestra para este tipo de mediciones deben permanecer por debajo de 1 mm, siendo el mínimo posible dado que el porcentaje de reflectancia difusa para las bandas relevantes de información se pueden ver comprometidas para distancias superiores, especialmente cuando se traslada a aplicaciones in-vivo para la discriminación de tejidos.

Estos resultados son particularmente interesantes en fantasmas de plastilina de color rojo, al ser los fantasmas con mejor respuesta espectral, ya que su firma espectral tiene un comportamiento más parecido a la hemoglobina en la región visible del espectro. Como trabajo futuro se plantea reproducir este estudio para evaluar los resultados sobre tejido *in vivo*.

## Agradecimientos

Este Trabajo ha sido parcialmente financiado por el proyecto del Plan Nacional de I+D+i “Cribado diagnóstico de microorganismos mediante microscopia avanzada e inteligencia artificial en patologías humana” (PID2021-127691OB-I00), del Ministerio de Ciencia e Innovación, cofinanciado con fondos FEDER.

## Referencias

- [1] F. Fanjul-Vélez, S. Pampín-Suárez, y J. L. Arce-Diego, “Application of Classification Algorithms to Diffuse Reflectance Spectroscopy Measurements for Ex Vivo Characterization of Biological Tissues”, *Entropy*, vol. 22, núm. 7, Art. núm. 7, jul. 2020, doi: 10.3390/e22070736.
- [2] G. Zonios y A. Dimou, “Modeling diffuse reflectance from semi-infinite turbid media: application to the study of skin optical properties”, *Opt. Express*, vol. 14, núm. 19, pp. 8661–8674, sep. 2006, doi: 10.1364/OE.14.008661.
- [3] S. F. Malin, T. L. Ruchti, T. B. Blank, S. N. Thennadil, y S. L. Monfre, “Noninvasive Prediction of Glucose by Near-Infrared Diffuse Reflectance Spectroscopy”, *Clin. Chem.*, vol. 45, núm. 9, pp. 1651–1658, sep. 1999, doi: 10.1093/clinchem/45.9.1651.
- [4] D. C. G. de Veld, M. Skurichina, M. J. H. Witjes, R. P. W. Duin, H. J. C. M. Sterenborg, y J. L. N. Roodenburg, “Autofluorescence and diffuse reflectance spectroscopy for oral oncology”, *Lasers Surg. Med.*, vol. 36, núm. 5, pp. 356–364, 2005, doi: 10.1002/lsm.20122.
- [5] F. Fanjul-Vélez, L. Arévalo-Díaz, y J. L. Arce-Diego, “Intra-class variability in diffuse reflectance spectroscopy: application to porcine adipose tissue”, *Biomed. Opt. Express*, vol. 9, núm. 5, pp. 2297–2303, may 2018, doi: 10.1364/BOE.9.002297.
- [6] H. Yavari, “Diffuse Reflectance Spectroscopy: Using Multivariate analysis method for determination of tissue optical properties”, 2016, Consultado: el 1 de septiembre de 2022. [En línea]. Disponible en: <http://lup.lub.lu.se/student-papers/record/8880115>
- [7] G. Kortüm, *Reflectance Spectroscopy: Principles, Methods, Applications*. Springer Science & Business Media, 2012.
- [8] Y. Ti y W.-C. Lin, “Effects of probe contact pressure on in vivo optical spectroscopy”, *Opt. Express*, vol. 16, pp. 4250–62, abr. 2008, doi: 10.1364/OE.16.004250.
- [9] J. A. Delgado Atencio *et al.*, “Influence of probe pressure on human skin diffuse reflectance spectroscopy measurements”, *Opt. Mem. Neural Netw.*, vol. 18, pp. 6–14, mar. 2009, doi: 10.3103/S1060992X09010020.
- [10] S. Hariri Tabrizi y A. Shakibaei, “The Effect of Probe Pressure on In Vivo Single Fiber Reflectance Spectroscopy”, *J. Lasers Med. Sci.*, vol. 7, pp. 233–237, oct. 2016, doi: 10.15171/jlms.2016.41.
- [11] A. P. Popov, A. V. Bykov, y I. V. Meglinski, “Influence of probe pressure on diffuse reflectance spectra of human skin measured in vivo”, *J. Biomed. Opt.*, vol. 22, núm. 11, p. 110504, nov. 2017, doi: 10.1117/1.JBO.22.11.110504.
- [12] M. Bregar, B. Cugmas, F. Pernuš, B. Likar, y M. Bürmen, “A study on the properties of contact pressure induced by manually operated diffuse reflectance fiber optic probes”, *Prog. Biomed. Opt. Imaging - Proc. SPIE*, vol. 9327, mar. 2015, doi: 10.1117/12.2077190.
- [13] X. U. Zhang, D. J. Faber, T. G. V. Leeuwen, y H. J. C. M. Sterenborg, “Effect of probe pressure on skin tissue optical properties measurement using multi-diameter single fiber reflectance spectroscopy”, *J. Phys. Photonics*, vol. 2, núm. 3, p. 034008, jun. 2020, doi: 10.1088/2515-7647/ab9071.
- [14] A. I. Mundo *et al.*, “Diffuse reflectance spectroscopy to monitor murine colorectal tumor progression and therapeutic response (Erratum)”, *J. Biomed. Opt.*, vol. 25, núm. 4, p. 049803, abr. 2020, doi: 10.1117/1.JBO.25.4.049803.
- [15] B. Wang, S. Fan, L. Li, y C. Wang, “Study of probe-sample distance for biomedical spectra measurement”, *Biomed. Eng. Online*, vol. 10, p. 95, nov. 2011, doi: 10.1186/1475-925X-10-95.
- [16] S. Buendía Aviles, M. Cunill Rodríguez, J. A. Delgado Atencio, J. L. A. Diego, y F. Velez Fanjul, “Análisis diagnóstico de afecciones cutáneas mediante el uso de Espectroscopia de Reflectancia Difusa”, 2019.
- [17] T. Papaioannou *et al.*, “Effects of Fiber-Optic Probe Design and Probe-to-Target Distance on Diffuse Reflectance Measurements of Turbid Media: An Experimental and Computational Study at 337 nm”, *Appl. Opt.*, vol. 43, pp. 2846–60, jun. 2004, doi: 10.1364/AO.43.002846.
- [18] A. E. Martínez Rodríguez, “Estudio experimental de la influencia de la distancia sonda-piel en la métrica de los espectros de reflexión difusa”, p. 90.
- [19] S. Buendía Aviles, “Impacto de la distancia sonda-muestra en la aplicación de la espectroscopia de reflectancia difusa”, presentado en 3er Congreso Nacional de Cuerpos Académicos de Investigación y Posgrado, Universidad Politécnica del Estado de Morelos, el 6 de julio de 2022. [En línea]. Disponible en: <https://congreso.upemor.edu.mx/>
- [20] A. I. Mundo *et al.*, “Diffuse reflectance spectroscopy to monitor murine colorectal tumor progression and therapeutic response”, *J. Biomed. Opt.*, vol. 25, núm. 3, p. 035002, mar. 2020, doi: 10.1117/1.JBO.25.3.035002.
- [21] P. R. Bargo, S. A. Prahl, y S. L. Jacques, “Collection efficiency of a single optical fiber in turbid media”, *Appl. Opt.*, vol. 42, núm. 16, pp. 3187–3197, jun. 2003, doi: 10.1364/AO.42.003187.
- [22] R. Hennessy, W. Goth, M. Sharma, M. Markey, y J. Tunnell, “Effect of probe geometry and optical properties on the sampling depth for diffuse reflectance spectroscopy”, *J. Biomed. Opt.*, vol. 19, p. 107002, oct. 2014, doi: 10.1117/1.JBO.19.10.107002.
- [23] I. Fridolin y L.-G. Lindberg, “Optical non-invasive technique for vessel imaging: I. Experimental results”, *Phys. Med. Biol.*, vol. 45, pp. 3765–78, ene. 2001, doi: 10.1088/0031-9155/45/12/318.
- [24] C. Reble, I. Gersonde, S. Schanzer, M. C. Meinke, J. Helfmann, y J. Lademann, “Evaluation of detection distance-dependent reflectance spectroscopy for the determination of the sun protection factor using pig ear skin”, *J. Biophotonics*, vol. 11, núm. 1, p. e201600257, 2018, doi: 10.1002/jbio.201600257.
- [25] E. Hecht, *Optica*. Addison-Wesley Iberoamericana Espana, S.A., 2001. [En línea]. Disponible en: <https://books.google.com.mx/books?id=XdCaQQAACA-AJ>

# Intra-variability in IB4- and IB4+ DRG neurons: a population of models study

V. Puche-García<sup>1</sup>, N. López-Sánchez<sup>1</sup>, E. Villalba-Riquelme<sup>2</sup>, S. Hingorani-Jai Prakash<sup>3</sup>, A. Ferrer-Montiel<sup>2</sup>, V. Moreno-Manzano<sup>3</sup> and J.M. Ferrero<sup>1</sup>

<sup>1</sup> Centro de Investigación e Innovación en Bioingeniería (Ci2B), Universitat Politècnica de València, España

<sup>2</sup> Instituto de Investigación, Desarrollo e Innovación en Biotecnología Sanitaria de Elche (IDI BE), Universitat Miguel Hernández, España

<sup>3</sup> Laboratorio en Regeneración Tisular y Neuronal, Centro de Investigación Príncipe Felipe, España

[cferrero@eln.upv.es](mailto:cferrero@eln.upv.es)

## Abstract

The neural action potential (AP) depends on the level of expression of the different channel families, decisively influencing neuronal excitability. Since the experimental study of these phenomena has considerable limitations, the contributions of computational models are increasingly useful. In this study, a population of models of cultured dorsal root ganglion (DRG) neurons has been developed allowing the identification of parameter sets that give rise to APs congruent with the experimental ones regarding two types of DRG neurons: IB4- and IB4+. The results show a high degree of consistency with the experimental results reproducing the greater excitability of IB4- DRG neurons over IB4+ neurons. The intra-species APs have been successfully simulated by considering the intrinsic variability of the expression levels of the different families of channels. This methodology allows the determination of the correct ion channel expression levels to model the neuron AP.

## 1. Introduction

*In-vivo* and *in-vitro* assays benefit from complementary *in-silico* experiments when studying neuron excitability and underlying complex phenomena. This approach allows computationally simulating the physiological behavior of excitable cells, describing their activity through a system of mathematical equations [1]. Moreover, *in-silico* experiments help to determine the cause-effect relationships between the changes in a set of specific parameters of the model and the resulting effects.

Building a representative population of models (PoM) of excitable cells enables to simulate and analyze the variability of experimental measurements. Therefore, the high physiological variability exhibited in internal processes at every level can be represented. In the biological context, this uncertainty is manifested in changes in the values of representative biomarkers within certain ranges. Even with homogeneously cultured cells, these exhibit a significant variation between the different cells of the population, generally due to an intrinsic level of randomness.

The study of the variability in dorsal root ganglion (DRG) neurons enables to better understand the mechanisms involved in the normal and pathological behavior of the nervous system. This type of neuron plays an important role in the onset and maintenance of neuropathic pain [2,

3]. As a result of the high incidence of neurological injuries [4] and the current limited strategies to treat them [5], tissue regeneration is becoming an innovative field dedicated to discovering new ways to restore functionality [6]. Improving neuronal excitability is crucial to this point since the greater the probability of restoring the loss of functionality depends on the excitability of cells.

In this work, a PoM, commonly used in cardiac *in-silico* studies [7, 8], and its corresponding sensitivity analysis have been developed. The PoM reproduces the intrinsic variability of DRG neurons and allows the assessment of the effect of different levels of ion channel expression on the action potential (AP) features. Additionally, the developed PoM allows the identification of two DRG neuron subsets: IB4- and IB4+. These two types differ in excitability and AP features [9, 10, 11]. Using the models, a statistical analysis has been carried out on the influence of the type of neuron and the expression level of ion channels on neuronal excitability and AP characteristics.

## 2. Methods

### 2.1. Experimental dataset

Our dataset consisted of microelectrode recordings of isolated DRG preparations. The electrophysiology of these cultured neurons is defined by the varying expression of four families of Nav channels (Nav1.6, Nav1.7, Nav1.8, Nav1.9) and four families of Kv channels (Kv1, Kv2, Kv3, Kv4). Whole-cell recordings were performed through current clamp protocol at room temperature. The extracellular solution contained (in mM): 140 NaCl, 4 KCl, 2 CaCl<sub>2</sub>, 2 MgCl<sub>2</sub>, 10 HEPES, 5 Glucose, 20 Mannitol adjusted to pH 7.4. The internal solution contained (in mM): 144 KCl, 2 MgCl<sub>2</sub>, 10 HEPES, 5 EGTA adjusted to pH 7.2. The experimental recordings yield physiological ranges of AP biomarkers for each type of neuron (IB4- and IB4+).

We used eight biomarkers to quantify the features of the DRG neuron APs: overshoot, amplitude, Resting Membrane Potential (RMP), time-to-peak, After-Hyper Polarization amplitude (AHP), time to reach half recovery to rest, width (measured as AP duration at half amplitude) and membrane potential at which depolarization speed is ≥

10 V/s ( $V_{th}$ ). Also, the rheobase current was obtained for each neuron.

## 2.2. DRG neuron model

For the formulation of ionic currents, the existing model of murine DRG neuron AP from Zheng et al. 2019 [12] was modified. The model was adapted to fit the conditions of our laboratory experiments. The model was implemented in MATLAB for a 0-dimensional neuron patch and solved using the explicit Euler method.

## 2.3. Population of models

To account for experimental variability, a population of 1000 models was constructed by randomly varying the maximum conductances of  $N_{AV}$  and  $K_V$  channels within a pre-established range of  $\pm 60\%$  from the control value, except for  $N_{AV1.6}$  that was varied from 0 to 6 mS/cm<sup>2</sup>.

As for the calibration, several filters were implemented during the construction of the PoM to distinguish whether the model behaves in an electrophysiological realistic way. Specifically, models that exhibited spontaneous activity, or were not able to trigger consecutive APs when periodically stimulated, were not included in the final population.

A model was considered to be “IB4-/IB4+ congruent” depending on whether its biomarkers fall within the physiological ranges of the corresponding IB4-/IB4+ cultured DRG neurons. As a result, three population sets were obtained: DRG neuron models IB4- congruent, IB4+ congruent, and non-congruent models. APs belonging to this last category were discarded since they did not faithfully represent the experimental APs for any group.

Two different stimulation protocols were implemented to reproduce our experimental recordings. A “long” stimulation protocol consisted of a single stimulus lasting 1000 ms and was used to determine the rheobase current. On the other hand, a “short” stimulation protocol was defined to obtain the remaining biomarkers by applying five consecutive stimuli 10 ms in duration separated by a stimulation period of 250 ms.

## 2.4. Statistical Methods

To study how the variability in the conductances of  $N_{AV}$  and  $K_V$  channels distinctly affect IB4- and IB4+ neurons, the partial correlation coefficients (PPC) were calculated. This method was chosen since for each biomarker the effects of the rest of the inputs on the biomarker are adjusted, thereby enabling to quantify the correlation between the varied conductances and the described biomarkers.

The PCC between  $x$  and  $y$  is defined as the correlation coefficient between the residuals as follows:

$$PPC(x, y, z_i) = \frac{Cov(r_x, r_y)}{Var(r_x) \cdot Var(r_y)} \quad (1)$$

Biomarker-related differences between species (IB4- and IB4+) were assessed by one-way ANOVA considering  $p < 0.05$  as statistically significant.

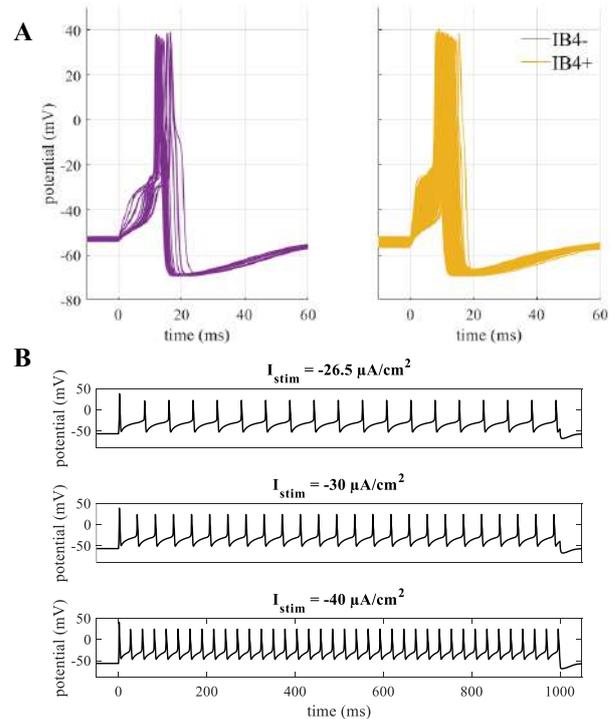
## 3. Results and discussion

### 3.1. Modelling IB4- and IB4+ DRG neurons

Both species (IB4- and IB4+) were successfully simulated and AP morphologies were validated according to [12] and were consistent with experimental recordings.

As previously stated, the PoM is composed of 1000 models of which 10 were found to be congruent with IB4- neurons and 162 with IB4+ neurons, while 15 additional models were found to be congruent with both types of neurons. Therefore, the following sets were obtained from the PoM (Fig. 1-A): 25 congruent with DRG IB4- neurons, 177 congruent with DRG IB4+ neurons and the remaining 813 models were not consistent with any of the previous sets.

In addition, modulation of excitability as a function of stimulus strength was correctly reproduced. For long simulations (1000 ms), the increase in the firing frequency was correctly predicted, so that it increases with the increase in stimulation current, as shown in Fig. 1-B for control values of channel conductances. For stimulus currents of  $-26.5$ ,  $-30$ , and  $-40 \mu\text{A}/\text{cm}^2$ , the firing frequency is 18.20, 24.40, and 37.04 Hz, respectively, showing that the greater the stimulus strength, the higher the frequency (as experimentally observed).



**Figure 1.** A) APs of DRG IB4- ( $n=25$ ) and IB4+ ( $n=177$ ) neuron models obtained from the PoM. B) APs trains for the long stimulation protocol simulated under control values for the channel conductances.

### 3.2. Study of the conductance variability between IB4- and IB4+ DRG neurons

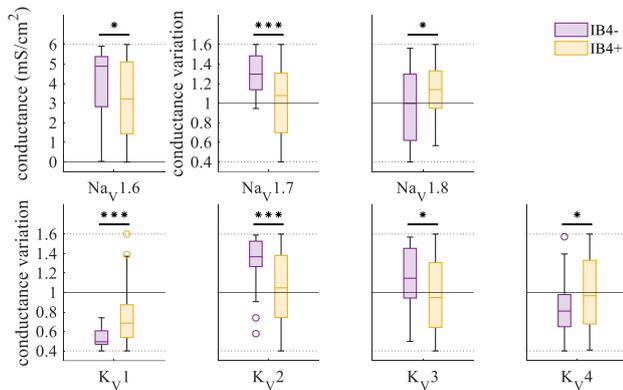
The distribution of the variation of conductances is displayed in Fig. 2. Significant differences in conductance variation between IB4- and IB4+ are observed. Thus, electrophysiological differences between the two groups

can be modelled by varying the conductances of the channel families.

From both species, high values of  $G_{Kv1}$  conductance are not suitable for the building of congruent models, being more significant in the IB4- case. Therefore, the random range could be upwardly limited to the control value, since no consistent values were obtained above it.

In the IB4- case, low values of  $G_{Nav1.7}$  and  $G_{Kv2}$  did not fit well to obtain congruent models. On the other hand, the conductances of the IB4+ congruent models show distributions that cover the entire range of random values except for  $G_{Kv1}$  and  $G_{Nav1.8}$ . Thus, the ranges from the other conductances could be extended.

The PoM has made it possible to identify sets of parameters (maximum conductances of ionic currents) consistent with the experimental results obtained in both IB4- and IB4+ DRG neurons. In addition, the intra-species APs have been successfully simulated by considering the intrinsic variability of the expression levels of the different populations of channels.



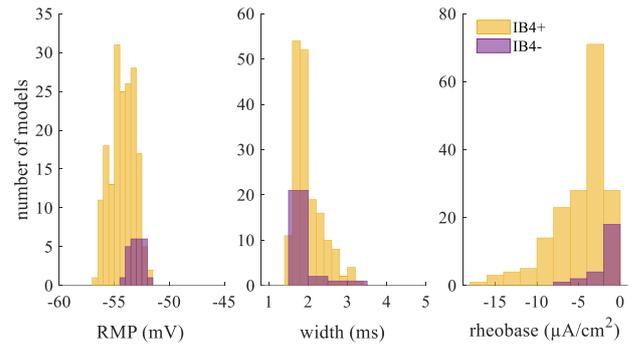
**Figure 2.** Distribution of the conductances for the sets of simulated DRG neurons for IB4- (purple) and IB4+ (yellow). The dashed lines delimit the chosen ranges for the variability of the conductances and the solid black line marks the control value for these. Current conductances that show statistically significant differences are indicated as Mann-Whitney U test: \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ .

Therefore, the electrophysiological differences between the two types of neurons can be computationally modelled from the variation of parameters that determine their bioelectric activity since this variability influences the biomarkers measured for each specie.

### 3.3. Study of differences in biomarkers

As described in the literature [9, 10, 11], APs of DRG neurons show differences related to the species IB4- and IB4+. Thus, an ANOVA analysis was carried out to study if a significant difference between IB4- and IB4+ DRG neuron models regarding the AP duration (width), the rheobase current and the RMP exists. These biomarkers were chosen since the main differences between the mentioned species consist of IB4+ neurons having longer AP durations, a higher threshold for AP firing (rheobase current) and lower activity (lower RMP). A comparison of the distribution of these biomarkers between species is shown in Fig. 3.

The ANOVA analysis shows a  $p$ -value =  $3.68e-08$  for RMP,  $p$ -value = 0.113 for width and  $p$ -value =  $1.92e-05$  for rheobase current. Therefore, IB4- and IB4+ DRG neuron models can be considered significantly different based on rheobase current and RMP, but not AP width. Even so, the distribution of AP width for IB4+ models covers a wider range for higher durations being consistent with both literature [9, 10, 11] and our experimental data. Further simulations would be needed to faithfully represent the AP width observed in experimental recordings and re-examine if there is a significant difference regarding this biomarker.

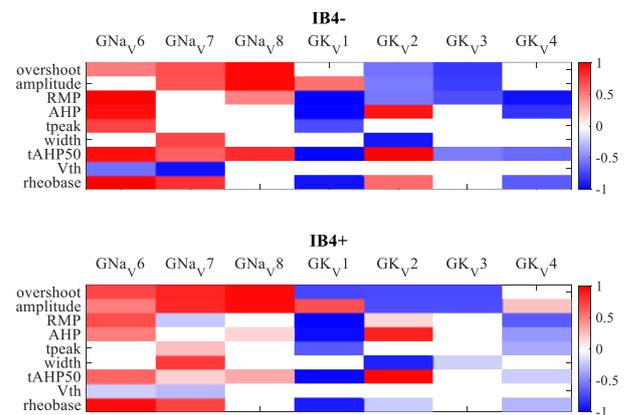


**Figure 3.** Distribution of main biomarkers (RMP, width, rheobase) for the simulated DRG neuron sets of IB4- (purple) and IB4+ (yellow) species.

Regarding neuronal excitability, the difference between IB4- and IB4+ neurons has been correctly estimated, with the rheobase being lower in the IB4- models. This result is consistent with the experimental results.

### 3.4. Sensitivity analysis on conductance level and biomarkers

Sensitivity analysis results are shown in Fig. 4, where the significant partial correlation coefficients for each variable-biomarker pair ( $p < 0.05$ ) are shown. Most of the biomarkers showed a significant partial correlation with different inputs.



**Figure 4.** Representation of the partial correlation coefficients for each variable-biomarker pair. The value of the coefficient is shown by a colour scale showing the positive correlations in red and the negative ones in blue. Pairs without significant correlation ( $p$ -value  $> 0.05$ ) are shown in white.

For both IB4- and IB4+ cases, a high positive correlation was found between the overshoot and amplitude

biomarkers and the conductances  $G_{\text{Nav}1.7}$  and  $G_{\text{Nav}1.8}$ . This is consistent since these channel families are related to depolarization. On the other hand,  $G_{\text{Nav}1.6}$  also correlated positively for both species with the threshold current for the long stimulation protocol (rheobase). This indicates that, at higher conductances of this family, the threshold current necessary for firing APs becomes less negative, thus increasing cell excitability.

For the  $K_V$  channel families, the  $G_{Kv1}$  conductance showed a negative correlation for several biomarkers including RMP, AHP, tpeak and tAHP50.

In addition,  $G_{Kv2}$  was negatively correlated to AP width for both cases, so at higher levels of this conductance, the duration of AP will be shorter. Due to its significant correlation, a small increase in conductance could drastically decrease the duration of the AP. This may explain the excessively short AP durations obtained in the simulated models compared to the recordings in cultured DRG neurons.

Therefore, it is shown that the variability in the conductances can generate models consistent with the experimental data. The variability influences the studied biomarkers to a lesser or greater extent. In any case, the ranges of the variables continue to be decisive in obtaining simulated models of DRG neurons, since they determine the electrical properties of the model. Thanks to the PoM, it has been possible both to model the electrical properties of cultured IB4- and IB4+ DRG neurons and to simulate the characteristic variability of experimental records.

### 3.5. Limitations and future perspectives

This study accounts the response of a 0D neuron patch. Extending the model to a 3D morphology of DRG neurons and implementing networks of synaptic-connected neurons would enable to further examine the characteristic excitability and conduction properties of these type of neurons.

Simulations have underestimated the duration of the AP, so the inclusion of the  $\text{Nav}1.9$  channel would contribute to obtaining longer durations AP and longer depolarization times since it has slower kinetics than the rest of the Nav channels [10].

## 4. Conclusions

In this study, a mathematical model has been developed for the computational simulation of the electrophysiological behaviour of cultured DRG neurons. The PoM approach allows to reproduce the intrinsic variability of the expression levels of the different ionic channels, to study its influence on neuronal excitability and to build two different populations of models that faithfully represent both types of neurons, being able to reproduce the variability present in experiments with cultured neurons in the laboratory. We prove that differences in electrophysiological features and excitability between IB4- and IB4+ neurons result from the specific variation of channel conductances. As the experimental recordings, IB4- DRG models show greater excitability than IB4+ neurons. Further computational research would be needed

to help design and improve strategies oriented to heal neuronal excitability-related pathologies giving a complementary approach to experimental studies.

## References

- [1] Hodgkin AL, Huxley AF. A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of physiology*, vol 117, sup 4, 1952, pp 500-544
- [2] Brown A, Weaver LC. The dark side of neuroplasticity. *Experimental neurology*, vol 235, sup 1, 2012, pp 133-141
- [3] Eldahan KC, Rabchevsky AG. Autonomic dysreflexia after spinal cord injury: systemic pathophysiology and methods of management. *Autonomic Neuroscience*, vol 209, 2018, pp 59-70
- [4] Website from World Health Organization (WHO). <https://www.who.int/news-room/fact-sheets/detail/spinal-cord-injury> [Accessed: Septemeber 2021]
- [5] Van Hecke O, Torrance N, Smith BH. Chronic pain epidemiology and its clinical relevance. *British journal of anaesthesia*, vol 111, sup 1, 2013, pp 13-18
- [6] Waxman SG, Zamponi GW. Regulating excitability of peripheral afferents: emerging ion channel targets. *Nature neuroscience*, vol 17, sup 2, 2014, pp 153–163
- [7] Vagos MR, Arevalo H, de Oliveira BL, Sundnes J, Maleckar MM. A computational framework for testing arrhythmia marker sensitivities to model parameters in functionally calibrated populations of atrial cells. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol 27, sup 9, 2017, pp 093941
- [8] Muszkiewicz A, Britton OJ, Gemmel P, Passini E, Sánchez C, Zhou X, Carusi A, Quinn TA, Burrage K, Bueno-Orovio A, Rodriguez B. Variability in cardiac electrophysiology: using experimentally-calibrated populations of models to move beyond the single virtual physiological human paradigm. *Progress in biophysics and molecular biology*, vol 120, sup 1-3, 2016, pp 115-127
- [9] Stucky CL. IB4-Positive Neurons, Role in Inflammatory Pain. *Encyclopedia of Pain*, Springer Berlin Heidelberg, 2007, pp 952-955
- [10] Fang X, Djouhri L, McMullan S, Berry C, Waxman SG, Okuse K, Lawson SN. Intense isolectin-B4 binding in rat dorsal root ganglion neurons distinguishes C-fiber nociceptors with broad action potentials and high Nav1. 9 expression. *Journal of Neuroscience*, vol 26, sup 27, 2006, pp 7281-7292
- [11] Wu B, Su X, Zhang W, Zhang Y, Feng X, Ji Y, Tan Z. Oxaliplatin Depolarizes the IB4–Dorsal Root Ganglion Neurons to Drive the Development of Neuropathic Pain Through TRPM8 in Mice. *Frontiers in Molecular Neuroscience*, 2021, pp 108
- [12] Zheng Y, Liu P, Bai L, Trimmer JS, Bean BP, Ginty DD. Deep sequencing of somatosensory neurons reveals molecular determinants of intrinsic physiological properties. *Neuron*, vol 103, sup 4, 2019, pp 598-616
- [13] Kandel ER, Schwartz JH, Jessell TM, Siegelbaum S, Hudspeth AJ, Mack S. Principles of neural science (5a), New York: McGraw-hill, 2000
- [14] Marino S, Hogue IB, Ray CJ, Kirschner DE. A methodology for performing global uncertainty and sensitivity analysis in systems biology. *Journal of theoretical biology*, vol 254, sup 1, 2008, pp 178–196

# La Complejidad Preoperatoria de la Distribución Espectral de las Ondas Fibrilatorias de Superficie Predice el Resultado de la Ablación por Catéter de Fibrilación Auricular Persistente

P. Escribano Cano<sup>1</sup>, J. Ródenas García<sup>1</sup>, M. García Teruel<sup>1</sup>, M.A. Arias Palomares<sup>2</sup>,  
J.J. Rieta Ibáñez<sup>3</sup>, R. Alcaraz Martínez<sup>1</sup>

<sup>1</sup> Grupo de Investigación en Ingeniería Electrónica, Biomédica y de Telecomunicación, DIEEAC, Universidad de Castilla-La Mancha, Albacete, España, {pilar.escribano, juan.rodenas, manuel.garcia, raul.alcaraz}@uclm.es

<sup>2</sup> Unidad de Arritmias, Hospital Virgen de la Salud, Toledo, España

<sup>3</sup> BioMIT.org, Departamento de Ingeniería Electrónica, Universitat Politècnica de Valencia, Valencia, España

## Resumen

*Como la arritmia más diagnosticada, la fibrilación auricular (FA) es tratada principalmente mediante el procedimiento de ablación por catéter (AC), a pesar de su alta tasa de recurrencia en pacientes de FA persistente. La predicción preoperatoria de su resultado ayudaría a seleccionar aquellos pacientes que realmente se beneficiarán del tratamiento. Esta problemática alentó el estudio de la complejidad  $C_0$  como una nueva métrica de la distribución espectral de potencia (DEP) de las ondas fibrilatorias (ondas-f). La complejidad  $C_0$  es una ratio que divide la DEP en dos componentes divergentes mediante un umbral  $\theta$ , que se obtuvo multiplicando el valor medio de la DEP por un factor,  $\alpha$ , que oscila entre 1,5 y 2,5. Todos los índices de complejidad  $C_0$ , calculados a partir de las ondas-f de 74 pacientes, mostraron diferencias estadísticamente significativas entre aquellos que mantuvieron el ritmo sinusal y los que recayeron a FA tras 9 meses de seguimiento. Se obtuvieron valores de sensibilidad (Se), especificidad (Es) y exactitud (Ex) de aproximadamente un 70%, lo que supone un rendimiento superior a los de los predictores comunes, como la frecuencia dominante (FD) y amplitud de las ondas-f normalizada (AOFn). Además, la combinación de la complejidad  $C_0$ , con  $\alpha = 2$ , y la FD a través de un árbol de decisión mejoró la clasificación hasta valores de Se, Es y Ex del 75,33%, 77,33% y 76,58%, respectivamente. La DEP de las ondas-f contiene información relevante para anticipar el resultado de la AC de FA persistente.*

## 1. Introducción

La arritmia supraventricular de fibrilación auricular (FA) se ha convertido en uno de los problemas de salud pública más importantes [1], alcanzando una prevalencia de entre 2 y 4% en adultos, a la cual se prevé que siga en aumento [2]. Los síntomas relacionados con la arritmia, como palpitaciones y dificultad respiratoria, entre otros, afectan severamente a la calidad de vida de los pacientes, y lo que es más importante, la FA es el factor de riesgo más común de sufrir un accidente cerebrovascular isquémico [3]. Además, la arritmia está asociada a un proceso de remodelación del sustrato auricular que favorece su perpetuación [4], de manera que se recomienda su detección y tratamiento temprano [5].

La ablación por catéter (AC) es un tratamiento para pacientes de FA muy común y adecuado para terminar con

los síntomas relacionados con la arritmia, puesto que su principal objetivo es restaurar el ritmo sinusal (RS) [2]. Sin embargo, los pacientes de FA persistente, aquellos que padecen episodios de la arritmia que no terminan por sí solos en una semana y requieren ser intervenidos externamente para ponerles fin, presentan una alta tasa de recurrencia a medio plazo (alrededor del 35% en el primer año) a pesar de aplicar líneas adicionales de ablación o procedimientos repetidos [6]. Esta situación ha generado interés clínico en la predicción preoperatoria del resultado del procedimiento de AC para la selección óptima de los pacientes que se beneficiarían del tratamiento y, por lo tanto, evitar riesgos innecesarios al resto de pacientes, disminuyendo la tasa de hospitalizaciones y procedimientos repetidos, y alcanzando una reducción de la carga y el coste asociado al sistema sanitario para el tratamiento de la FA [7].

Hasta el momento, los predictores clínicos estudiados en trabajos previos solo han proporcionado resultados limitados y, en ocasiones, contradictorios [8]. De forma alternativa, métricas extraídas de las ondas fibrilatorias (ondas-f) del electrocardiograma (ECG) de superficie, como la frecuencia dominante (FD) [9] y la amplitud de las ondas-f (AOF) [10], también han sido analizadas como predictores. Aunque han demostrado una habilidad prometedora para la anticipación del resultado del procedimiento de ablación, esta es todavía insuficiente para su aplicación clínica. Por ello, el presente trabajo tiene como objetivo principal el estudio de una novedosa métrica espectral de las ondas-f para la mejora de la predicción preoperatoria del resultado de la AC en pacientes de FA persistente.

## 2. Materiales y métodos

### 2.1. Población de estudio

Este estudio involucra a 74 pacientes de FA persistente (21 mujeres y 53 hombres) con una edad media de 65 años (entre 39 y 82 años). Todos ellos fueron sometidos, bajo indicaciones clínicas estándar, a su primera intervención de AC de radiofrecuencia en el Hospital Universitario de Toledo.

## 2.2. Protocolo de ablación y seguimiento

La AC es una intervención que evita la operación a corazón abierto mediante la introducción de los catéteres a través del acceso venoso femoral mientras el paciente está sedado y sometido a medicación anticoagulante. Se basa principalmente en la técnica de aislamiento de las venas pulmonares, que tiene como objetivo acabar con la actividad eléctrica anormal desencadenada en esa región, a la que se accede por punción transeptal. El procedimiento se llevó a cabo mediante la aplicación de lesiones ablativas punto por punto con un catéter de radiofrecuencia para crear una línea circunferencial contigua alrededor de cada vena pulmonar (VP), cuya ubicación se determinó mediante un catéter de mapeo [11]. El final del procedimiento se logró tras aislar correctamente todas las VPs o después de restaurar el RS por cardioversión eléctrica si aún persistía la FA al final del procedimiento. Todas las intervenciones se completaron con éxito y los pacientes no sufrieron ninguna complicación. Tras un seguimiento de 9 meses, durante el cual los pacientes recibieron tratamiento anticoagulante y antiarrítmico a juicio clínico, se obtuvo una recaída a FA de 19 pacientes, mientras que los 55 restantes mantuvieron el RS.

## 2.3. Adquisición y procesamiento de la señal de ECG

El ritmo cardiaco de los pacientes se registró de forma preoperatoria y no invasiva a través de la señal superficial del ECG de 12 derivaciones con una duración entre 6 segundos y 5 minutos. Estas señales, junto con la información clínica posterior al periodo de seguimiento, constituyen la base de datos del estudio. La adquisición de señales se realizó con una frecuencia de muestreo de 977 Hz y una resolución de 16 bits. Posteriormente, se seleccionó la derivación V1 para la extracción de las ondas-f por ser en la que presentan mayor amplitud con respecto a la actividad ventricular [12]. Previamente, esta señal se preprocesó mediante técnicas de filtrado comunes y basadas en la transformada wavelet para atenuar las componentes de ruido de alta frecuencia, la interferencia de la red eléctrica y el desvío de la línea base [13,14]. Posteriormente, las ondas-f fueron extraídas para su caracterización mediante un algoritmo basado en el método de cancelación del valor singular adaptativo de los complejos QRST [15].

## 2.4. Caracterización de las ondas-f

Dada la disparidad en la longitud de las señales del ECG adquiridas de los pacientes, las ondas-f extraídas fueron segmentadas en intervalos no solapados de 6 segundos. Se limitó a 5 el máximo número de intervalos por paciente y se promediaron los parámetros derivados de los segmentos de cada ECG para realizar un análisis enfocado en el paciente y evitar la obtención de resultados sesgados.

En cuanto a la caracterización de las señales de ondas-f, se tomaron la FD y la AOF normalizada (AOFn) como métricas de referencia, puesto que han sido ampliamente abordadas en trabajos previos. Las métricas espectrales fueron calculadas a partir de la densidad espectral de potencia (DEP) de cada segmento, que se estimó a través del periodograma de Welch con una ventana de Hamming

de 4000 puntos, una superposición de 3000 puntos entre ventanas adyacentes y una resolución espectral de 0,1 Hz. En particular, la FD se estimó como aquella para la que se obtuvo la máxima amplitud del espectro de DEP en el rango frecuencial de 3 a 25 Hz. [9]. Por otro lado, la AOF fue estimada como el valor cuadrático medio de las ondas-f en el dominio del tiempo [16], la cual se normalizó expresándola como un porcentaje de la magnitud del pico R para minimizar la influencia de efectos espurios como la conductividad de la piel, que alteran la amplitud del ECG.

Además, la complejidad  $C_0$  fue una métrica novedosa introducida como medida no lineal de las ondas-f que divide la DEP de la señal en dos componentes divergentes, las partes regular e irregular [17,18]. Su cómputo se llevó a cabo después de normalizar la DEP sobre la potencia total en la banda de frecuencia de 3 a 25 Hz, como se muestra a continuación, para, así, obtener una función de probabilidad de área unitaria.

$$DEP_n(f) = \frac{DEP(f)}{\sum_{f_i}^{f_s} DEP(f)}, \quad (1)$$

donde  $f_i \leq f \leq f_s$ ;  $f_i$  y  $f_s$  son los límites inferior y superior de la banda de frecuencia, 3 y 25 Hz, respectivamente. Se evaluaron diferentes umbrales  $\theta$  de separación que se obtuvieron multiplicando el valor medio de la DEP por un factor,  $\alpha$ , comprendido entre 1,5 y 2,5. De esta manera, la parte irregular de la división se obtuvo como las componentes de la DEP con un valor inferior al umbral calculado, siendo la región superior la parte regular. Concretamente, los valores superiores a  $\theta$  fueron rechazados sustituyéndolos por cero y, por tanto, obteniendo la parte irregular:

$$DEP'_n(f) = \begin{cases} DEP_n(f), & DEP_n(f) \leq \theta \\ 0, & DEP_n(f) > \theta \end{cases} \quad (2)$$

Finalmente, la complejidad  $C_0$  se calculó como la ratio de la suma total de las componentes de DEP de la parte irregular entre la suma total de la DEP [17,18]:

$$C_0 = \frac{\sum_{f_i}^{f_s} DEP'_n(f)}{\sum_{f_i}^{f_s} DEP_n(f)} \quad (3)$$

Por lo tanto, este índice da como resultado un número real entre 0 y 1, que será más alto cuanto mayor sea el predominio de la parte irregular de la señal [17,18].

## 2.5. Análisis estadístico

Todas las métricas analizadas fueron expresadas a través de su mediana y rango intercuartílico (RI) para los pacientes que mantuvieron el RS y aquellos que recayeron a FA. Por otro lado, tras verificar la no-normalidad y homocedasticidad en la distribución de los datos mediante la prueba de Lilliefors y la prueba de Levene, respectivamente, se utilizó la prueba no paramétrica U de Mann-Whitney para evaluar la diferenciación estadística entre ambos grupos de pacientes. Para ello, se estableció un valor de significación estadística de  $p < 0,05$  como estadísticamente significativo.

El rendimiento de predicción de cada índice para clasificar a los pacientes entre aquellos que recaen a FA o mantienen

el RS se estimó mediante el promediado de 100 pruebas de validación de exclusión con 2/3 de la base de datos para entrenamiento y 1/3 para validación. En la fase de entrenamiento, la curva característica operativa del receptor (ROC) se empleó para estimar el umbral óptimo para discernir entre los dos grupos de pacientes. Esta herramienta traza la fracción de verdaderos positivos (sensibilidad), que son los positivos correctamente clasificados, frente a la fracción de falsos positivos (1 – especificidad), que son las muestras positivas clasificadas como negativas, para varios umbrales establecidos. La sensibilidad (Se) y la especificidad (Es) se consideró como la proporción de pacientes que recayeron a FA y que mantuvieron el RS correctamente clasificados, respectivamente. El umbral óptimo de clasificación fue elegido como aquel que proporcionaba el mejor balance entre Se y Es [19], aunque ello pudiese suponer no alcanzar la máxima exactitud (Ex), que es el porcentaje total de pacientes correctamente clasificados. Finalmente, el área bajo la curva ROC (AUC) también se calculó como una medida agregada del rendimiento de una variable en todos los umbrales de clasificación posibles.

Para explorar información complementaria entre características individuales y mejorar la predicción del resultado de la AC, también se realizó un análisis multivariante basado en un modelo de árbol de decisión de 5 nodos. Se empleó un algoritmo de selección de características secuencial progresivo que minimiza el error de clasificación y retorna un subconjunto de métricas óptimo para un modelo de clasificación. Este algoritmo tomó la complejidad  $C_0$ , con  $\alpha = 2$ , y la FD, como las mejores métricas para construir el modelo de predicción, cuyo rendimiento de clasificación se evaluó mediante 100 ciclos, al igual que los índices individuales.

### 3. Resultados

La tabla 1 muestra los valores obtenidos de los índices calculados para los dos grupos de pacientes. Se puede apreciar que todos ellos proporcionaron diferencias estadísticamente significativas, obteniendo  $p$ -valores inferiores a 0,05, excepto en el caso de la AOFn. Los resultados de la FD y la AOFn siguieron las mismas tendencias que las recogidas en trabajos previos, con valores superiores de la FD e inferiores de la AOFn para los pacientes que recayeron a FA en comparación con los que mantuvieron el RS. Por otro lado, los pacientes para los que el tratamiento de AC fue exitoso en el horizonte temporal del estudio mostraron valores de la complejidad  $C_0$  más altos para todos los umbrales considerados.

Por otro lado, el rendimiento de clasificación de los parámetros individuales se muestra en la Tabla 2. Se aprecia una mejor capacidad discriminadora de la complejidad  $C_0$  respecto a los parámetros de referencia, con una mayor Ex para  $\alpha = 1,75$  y una mayor AROC para  $\alpha = 2$ . No obstante, la combinación de la FD y la complejidad  $C_0$  para  $\alpha=2$  mediante un árbol de decisión de 5 nodos alcanzó mejoras de hasta un 23%, dado que se obtuvieron valores de Se, Es, Ex y AROC del 75,33%, 77,33%, 76,58% y 81,00%, respectivamente.

Índice	Grupo de pacientes		$p$ -valor
	Mantiene RS Mediana (RI)	Recae a FA Mediana (RI)	
FD (Hz)	5,35 (1,44)	5,68 (1,47)	0,036
AOFn (%)	6,41 (7,21)	5,58 (4,10)	0,316
$C_0$ 1,50 (%)	26,41 (7,24)	23,25 (3,20)	0,004
$C_0$ 1,750 (%)	29,15 (6,98)	25,44 (2,97)	0,002
$C_0$ 20 (%)	31,79 (8,49)	26,72 (3,13)	<0,001
$C_0$ 2,250 (%)	33,70 (9,33)	28,60 (4,26)	0,003
$C_0$ 2,50 (%)	35,94 (10,20)	31,25 (3,34)	0,004

**Tabla 1.** Resultados y diferenciación estadística de las métricas analizadas de ambos grupos de pacientes.

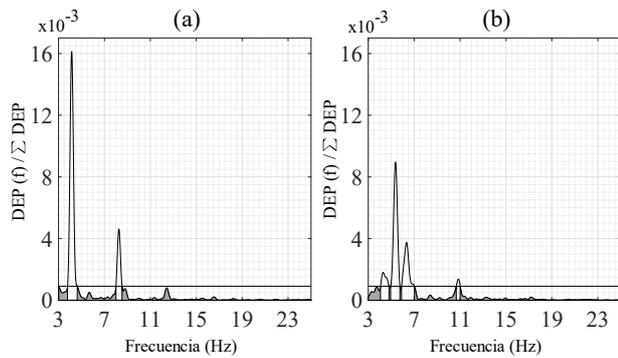
Índice	Se (%)	Es (%)	Ex (%)	AROC (%)
FD	63,16	61,82	62,16	66,22
AOFn	52,63	54,55	54,05	57,80
$C_0$ 1,50	68,42	69,09	68,92	72,44
$C_0$ 1,750	73,68	72,73	72,97	74,45
$C_0$ 20	68,42	70,91	70,27	75,50
$C_0$ 2,250	68,42	70,91	70,27	72,73
$C_0$ 2,50	68,42	69,09	68,92	72,15

**Tabla 2.** Resultados de clasificación de las métricas individuales analizadas.

### 4. Discusión y conclusiones

La alta tasa de recurrencia a medio plazo de los pacientes de FA persistente tras someterse al procedimiento de AC supone una gran carga para el sistema sanitario, lo que ha generado interés clínico en la predicción preoperatoria de su resultado debido a que continúa siendo el tratamiento de primera línea para la FA [1]. Conocer con precisión la probabilidad de éxito del tratamiento antes de la intervención permitiría la selección óptima de los pacientes, evitándoles riesgos innecesarios y minimizando los costes sanitarios asociados al tratamiento de la FA, entre otros beneficios [7]. Con este propósito, la complejidad  $C_0$  de las ondas-f se ha introducido como un nuevo predictor potencial en el presente trabajo [17,18]. De hecho, independientemente del umbral seleccionado, exhibió una mejor correlación con el resultado de la AC que la FD y la AOFn, que son predictores comunes en este campo. Además, su combinación para  $\alpha = 2$  con la FD a través de un árbol de decisión de 5 nodos mejoró la precisión diagnóstica hasta en un 23% con respecto a los índices individuales.

Los resultados proporcionados por la FD y la AOFn fueron consistentes con estudios previos, que observaron que una FD más alta y una AOFn más pequeña se correlacionaron con una mayor probabilidad de recurrencia de FA después de la AC [9,10]. Los resultados de los índices de complejidad  $C_0$  podrían estar asociados con el grado de organización de la actividad fibrilatoria debido a que es una métrica que evalúa globalmente todo el espectro de las ondas-f, teniendo en cuenta tanto la FD, como su contenido armónico. En este sentido, trabajos previos han observado que los pacientes con una actividad fibrilatoria más organizada, debido a la presencia de un mayor peso armónico en comparación con el resto de los componentes de frecuencia, tienen una mayor probabilidad de responder



**Figura 1.** Ejemplo de división de la DEP para el cómputo de la complejidad  $C_0$  en un paciente que mantuvo el RS (a) y un paciente que recayó a FA (b).

con éxito a cualquier tratamiento de la FA, incluido al de AC [9]. Por ello, como se muestra en el ejemplo de la Figura 1 (a), una componente de FD y armónicos fuertes en pacientes que mantuvieron el RS proporcionaron una parte irregular de la distribución de DEP más elevada en comparación con los pacientes que recayeron a FA y, por lo tanto, un resultado más alto de la ratio de complejidad  $C_0$ . Por el contrario, los pacientes que recayeron a FA (Figura 1 (b)) presentaron una distribución de DEP más desorganizada con diferentes componentes de baja frecuencia y no una frecuencia principal clara, lo que conllevó a la obtención de un resultado del índice de complejidad  $C_0$  más pequeño.

Todos estos resultados y observaciones sugieren que la distribución espectral de las ondas-f puede proporcionar información relevante para la predicción preoperatoria del resultado de la AC en pacientes de FA persistente. Además, se destaca la funcionalidad de la novedosa métrica  $C_0$  como un estimador global de la organización de las ondas-f. No obstante, estos hallazgos deberían ser corroborados en el futuro con una base de datos más amplia.

## Agradecimientos

Esta investigación ha recibido apoyo financiero de las subvenciones públicas PID2021-00X128525-IV0 y PID2021-123804OB-I00 del Gobierno de España, 10.13039/501100011033 junto con el Fondo Europeo de Desarrollo Regional (UE), SBPLY/17/180501/000411 y SBPLY/21/180501/000186 de la Junta de Comunidades de Castilla La Mancha, y AICO/2021/286 de la Generalitat Valenciana. Además, Pilar Escribano es beneficiaria de un contrato predoctoral 2020 PREDUCLM-15540, cofinanciado por el programa operativo del Fondo Social Europeo (FSE) 2014-2020 de Castilla-La Mancha.

## Referencias

[1] Zoni-Berisso M, Lercari F, Carazza T, Domenicucci S. Epidemiology of atrial fibrillation: European perspective. *Clinical Epidemiology* 2014;6:213-220.

[2] Hindricks G, Potpara T, et al. 2020 ESC guidelines for the diagnosis and management of atrial fibrillation developed in collaboration with the European Association for Cardio-Thoracic Surgery (EACTS). *European Heart Journal* Feb 2021;42(5):373-498.

[3] Warmus P, Niedziela N, et al. Assessment of the manifestations of atrial fibrillation in patients with acute

cerebral stroke - a single-center study based on 998 patients. *Neurological Research* Jun 2020;42(6):471-476.

- [4] Schotten U, Dobrev D, et al. Current controversies in determining the main mechanisms of atrial fibrillation. *Journal of Internal Medicine* May 2016;279(5):428-38.
- [5] Nattel S, Guasch E, et al. Early management of atrial fibrillation to prevent cardiovascular complications. *European Heart Journal* Jun 2014 ;35(22):1448-56.
- [6] Schmidt B, Brugada J, et al. Ablation strategies for different types of atrial fibrillation in Europe: results of the ESC/EORP EHRA atrial fibrillation ablation long-term registry. *Europace* 2020;22(4):558-566.
- [7] Walsh K, Marchlinski F. Catheter ablation for atrial fibrillation: current patient selection and outcomes. *Expert Review of Cardiovascular Therapy* 2018;16(9):679-692.
- [8] Dretzke J, Chuchu N, et al. Predicting recurrent atrial fibrillation after catheter ablation: a systematic review of prognostic models. *Europace* May 2020;22(5):748-760.
- [9] Alcaraz R, Hornero F, Rieta JJ. Electrocardiographic spectral features for long-term outcome prognosis of atrial fibrillation catheter ablation. *Annals of Biomedical Engineering* 2016;44(11):3307-3318.
- [10] Nault I, Lellouche N, et al. Clinical value of fibrillatory wave amplitude on surface ECG in patients with persistent atrial fibrillation. *Journal of Interventional Cardiac Electrophysiology* Oct 2009; 26(1):11-9.
- [11] Morin DP, Bernard ML, et al. The state of the art: Atrial fibrillation epidemiology, prevention and treatment. *Mayo Clinic Proceedings* 2016;91(12):1778-1810.
- [12] Petrénas A, Marozas V, Sörnmo L. Lead systems and recording devices. In *Atrial Fibrillation from an Engineering Perspective*. Springer, 2018; 25-48.
- [13] Sörnmo L, Laguna P. The electrocardiogram. a brief background. In *Bioelectrical Signal Processing in Cardiac and Neurological Applications*. Elsevier, 2005; 411-452.
- [14] García M, Martínez M, Rodenas, et al. A novel wavelet-based filtering strategy to remove powerline interference from electrocardiograms with atrial fibrillation. *Physiological Measurement* Nov 2018;39(11):115006.
- [15] Alcaraz R, Rieta JJ. Adaptive singular value cancelation of ventricular activity in single-lead atrial fibrillation electrocardiograms. *Physiological Measurement* Dec 2008;29(12):1351-69.
- [16] Alcaraz R, Hornero F, Rieta JJ. Noninvasive time and frequency predictors of long-standing atrial fibrillation early recurrence after electrical cardioversion. *Pacing and Clinical Electrophysiology* Oct 2011;34(10):1241-50.
- [17] Lu Y, Jiang D, et al. Predict the neurological recovery under hypothermia after cardiac arrest using  $C_0$  complexity measure of EEG signals. 30<sup>th</sup> Annual International Conference of the IEEE Engineering in Medicine and Biology Society Aug 2008;2133-6.
- [18] Zhou Y, Xie L, et al. The Study of  $C_0$  Complexity on Epileptic Absence Seizure. 7<sup>th</sup> Asian-Pacific Conference on Medical and Biological Engineering 2008;19:420-425.
- [19] Habibzadeh, F., Habibzadeh, P., et al. On determining the most appropriate test cut-off value: the case of tests with continuous results. *Biochem Med* 2016; 26:297-307.

# Análisis de Biomarcadores Electrocardiográficos en Pacientes con Cardiopatía Arritmogénica de Ventrículo Izquierdo

M. Bragulat Arévalo<sup>1</sup>, P. Gomis Roman<sup>1</sup>, E. Arbelo Lainez<sup>2</sup>, P. Jordà Burgos<sup>2</sup>, F. Palmieri<sup>1</sup>

<sup>1</sup> Centro de Recerca en Ingeniería Biomédica, UPC, Barcelona, España, mireiabragulat@gmail.com  
{pedro.gomis, flavio.palmieri}@upc.edu

<sup>2</sup> Departamento de cardiología, Hospital Clinic, Barcelona, España, [arbelo@clinic.cat](mailto:arbelo@clinic.cat), [paloma.jorda.b@gmail.com](mailto:paloma.jorda.b@gmail.com)

## Resumen

*La cardiopatía arritmogénica de ventrículo izquierdo (ALVC) es una afección cardíaca que conlleva un crecimiento del tejido fibrograso en este ventrículo, que sin detección temprana podría provocar muerte súbita. La caracterización de la ALVC en las pruebas electrocardiográficas (ECGs) es reducida. En este trabajo, evaluamos el comportamiento de varios biomarcadores relacionados con el complejo QRS en señales ECG, bajo la hipótesis de que el ALVC pueda afectar a la fase de despolarización del corazón. Dichos biomarcadores se extrajeron de ECGs de 28 pacientes que padecen ALVC para caracterizar diferencias por sexo y diagnóstico de hipertensión. Además, se analizó un subgrupo de 11 pacientes para caracterizar los cambios entre registros con diferencias de 2 y 4 años. Al evaluar los cambios de la amplitud, valor absoluto del área, energía y potencia del QRS se observa cómo disminuyen a lo largo del tiempo, siendo esta disminución significativa ( $p < 0.05$ ) al cabo de 4 años. No se encontraron diferencias entre hombres y mujeres o hipertensos y no. Se concluye que las métricas relacionadas con la amplitud, área y potencia del QRS pueden ser un buen predictor del desarrollo de la ALVC.*

## 1. Introducción

Las cardiopatías arritmogénicas (AVC) son enfermedades genéticas de prevalencia muy baja. Provocan que en el miocardio ventricular se genere tejido fibrograso, que conlleva al sufrimiento de arritmias que pueden llegar a provocar la muerte. Existen diferentes tipos de AVC en función del ventrículo que queda mayormente afectado. En el caso de este estudio, se ha centrado la atención en la cardiopatía arritmogénica de ventrículo izquierdo (ALVC). Pacientes con ALVC tienen especial interés ya que anteriormente eran mal diagnosticados con cardiopatía arritmogénica de ventrículo derecho (ARVC). Por lo tanto, la información de diagnóstico y de caracterización de ALVC es reducida. Actualmente, todos estos conocimientos se encuentran en los criterios de Padua de 2020 [1-3], destacando las técnicas de imágenes de los tejidos ventriculares por resonancia magnética como las más exitosas en la evaluación de ALVC. Sin embargo, el uso de electrocardiografía para evaluar la despolarización cardíaca es considerada una alternativa menos costosa en la evaluación de ALVC.

La hipótesis que se espera en los resultados es que, al crecer el grueso de la pared del ventrículo con la evolución de la enfermedad, se observará una disminución en la amplitud, valor absoluto del área, energía y potencia del QRS. La

explicación clínica es que al corazón le costará más realizar el bombeo de la sangre a menor capacidad espacial del ventrículo.

Este estudio pretende encontrar características de la enfermedad y su evolución en registros electrocardiográficos (ECGs). Por una parte, se busca encontrar diferencias significativas de la evolución de la enfermedad en un mismo grupo de paciente a lo largo de los años. Se analizan los biomarcadores en registros con diferencia de 2 y 4 años respecto al primer registro de la historia clínica. Por otra parte, se analizan los primeros registros de los pacientes diferenciados por sexo y diagnóstico de hipertensión de los pacientes.

El objetivo del estudio es evaluar la evolución de biomarcadores relacionados con la despolarización de los ventrículos a lo largo de la enfermedad ya que podrían ser indicadores robustos de empeoramiento de los pacientes con ALVC. De esta forma, se pretende ampliar la posibilidad de caracterizar dicha patología y mejorar los conocimientos sobre la evolución de la enfermedad a través de los ECGs.

Por tanto, en caso de éxito, este estudio podría establecer las bases para dar lugar a una nueva herramienta para los médicos a la hora de predecir la evolución de la AVLC, mediante una prueba más accesible y veloz que las de diagnóstico actual según los criterios de Padua.

## 2. Material

El *Hospital Clinic de Barcelona* proporcionó una base de datos de 28 pacientes diagnosticados con ALVC. Cada paciente consta de un número de ECG de 12 derivaciones de entre 20 y 40 segundos (con frecuencia de muestreo de 1000 Hz) obtenidos entre 2011 y 2021.

Todos los registros se procesaron y analizaron en Matlab.

## 3. Metodología

### 3.1. Pre-procesado

El pre-procesado inicial del ECG incluyó un filtrado paso bajo con frecuencia de corte a 40 Hz para eliminar el ruido muscular y la interferencia de red, seguido de un ajuste polinomial para atenuar la línea base. Los complejos QRS, la onda T y la onda P se detectaron y delinearono utilizando un delineador basado en la transformada wavelet aplicado en las 12 derivaciones. Luego se aplicó una regla de selección por multiderivación sobre las anotaciones

detectadas previamente. Esta regla consiste en ordenar las anotaciones y seleccionar como inicio (final) de una onda la primera (última) anotación cuyos  $k$  vecinos más cercanos se encuentren dentro de un intervalo de  $\delta$  ms. Finalmente, se realizó una inspección visual de los valores para confirmar su correcta detección de los puntos por el delineado y descartar todos los latidos ectópicos. [4, 5]

### 3.2. Biomarcadores

A partir del delineado obtenido tanto por cada derivación (SL) como aplicando la regla de selección por multiderivación (SLR), se extrajeron los siguientes biomarcadores: el voltaje pico a pico (mv), la duración de complejo QRS y el intervalo QT.

- Intervalo QT corregido (s):

$$QT_c = \frac{QT}{\sqrt{\frac{RR}{1s}}} \quad (1)$$

Siendo:

- o  $QT$  es el intervalo QT (s)
- o  $RR$  es el intervalo R-R (s)

Los intervalos QT y QTc se convirtieron luego a ms. Además, se calcularon [6]:

- Área en valor absoluto del QRS ( $mV \cdot s$ ):

$$AreaAbs_{QRS} = \frac{1}{f_s} \sum_{n=1}^N |QRS(n)| \quad (2)$$

Siendo:

- o  $QRS(n)$  es el complejo QRS de la señal ECG.
- o  $N$  son el número de muestras de la señal.
- o  $f_s$  es la frecuencia de muestreo de la señal.

- Energía del QRS ( $mV^2 \cdot s$ ):

$$Energía = \frac{1}{f_s} \sum_{n=1}^N |QRS(n)|^2 \quad (3)$$

- Potencia media del QRS ( $mV^2$ ):

$$P_m = \frac{1}{N} \sum_{n=1}^N |QRS(n)|^2 \quad (4)$$

Para unificar los datos extraídos en cada paciente y poder realizar posteriormente los estudios estadísticos, se calculó la mediana de todos los latidos para cada registro. De esta forma, se evitó que los valores atípicos afectaran la representación de las medidas.

### 3.3. Estudios estadísticos

Para los estudios estadísticos se realizó un análisis previo de la población. En todos los casos, la población no se distribuyó normalmente, por lo que se aplicaron pruebas no paramétricas.

Para el análisis de la diferencia temporal entre registros, se aplicó la prueba Wilcoxon. Dado que no todos los pacientes de la base de datos constaban de registros con diferencia de 2 y/o 4 años respecto al primer registro, esta quedó reducida para este análisis a 11 pacientes. [7]

Se aplicó la prueba Mann-Whitney U en el estudio para comparar el primer registro de todos los pacientes diferenciados en función del sexo y de su diagnóstico de hipertensión. En este caso, la base de datos utilizada se compuso de los 28 pacientes, aunque no estaba equilibrada, habiendo 8 mujeres frente a 20 hombres; y 6 hipertensos, frente a 22 no hipertensos. [7]

Los resultados se muestran en términos de mediana y 1r y 3r cuartil y mediante boxplots. P valores  $< 0,05$  se consideran significativos.

## 4. Resultados y discusiones

En este trabajo se ha analizado el comportamiento de varios biomarcadores basados en el análisis del QRS para conseguir una caracterización de ALVC en ECGs de corta duración.

Las Tablas 4.1 y 4.2 muestran los resultados de los estudios a 2 y 4 años, respectivamente, para aquellas derivaciones con P valor significativo.

		ECG_0 (N = 11)	ECG_2 (N = 11)	P valor	ECG_0 (N = 11)	ECG_4 (N = 11)	P valor
Voltaje pico a pico (mV)	aVF	0.45 [0.34-0.63]	0.47 [0.2-0.56]	0,042	0.49 [0.43 - 0.63]	0.4 [0.22 - 0.49]	0,012
	V1	1.09 [0.64-1.71]	0.95 [0.62-1.84]	0,105	1.02 [0.54 - 1.58]	0.87 [0.39 - 1.04]	0,007
	V2	1.39 [0.88-2.48]	1.07 [0.92-2.6]	0,700	1.54 [1.06 - 1.81]	1.16 [0.83 - 1.39]	0,001
Duración del QRS (ms)	V1	121 [106-140.3]	121 [110-156]	0,510	120 [105 - 124]	116 [105 - 129]	0,014
	V2	120 [106-132.3]	117 [109-142]	0,403	118 [109 - 122]	120 [108 - 130]	0,031
Intervalo QT (ms)	V1	435 [392-485]	451 [418-460]	0,695	413 [390.5 - 434.2]	438 [404.5 - 491]	0,027
	V3	445 [398-470]	438.5 [392.5-469]	0,365	392 [372.5 - 429.5]	426 [389 - 480]	0,024
Intervalo QT corregido (ms)	V2	428 [400-468]	415 [398-465]	0,966	424 [387 - 457]	428 [395 - 461]	0,012
Área del QRS en valor absoluto ( $mV \cdot s$ )	V1	42.32 [26.8 - 65.4]	37.25 [17.8 - 61.3]	0,275	31.31 [19.1 - 53.4]	30 [11.1 - 36.7]	0,009
	V2	55.84 [38.5 - 83.3]	57.54 [32.3 - 88.3]	0,695	45.55 [32.3 - 61.9]	38.65 [23.1 - 50.6]	0,001
Energía del QRS ( $mV^2 \cdot s \cdot 1000$ )	V1	26.2 [9.4 - 84.7]	22.93 [4.6 - 58.2]	0,432	16.38 [5.8 - 44.3]	13.42 [2.2 - 23.1]	0,007
	V2	44.33 [17.6 - 115.4]	28.2 [16.6 - 132.9]	0,557	32.5 [18.5 - 74.9]	23.88 [7.2 - 42.5]	0,001
Potencia media QRS ( $mV^2 \cdot 1000$ )	V1	178.6 [57.1 - 453.3]	177.8 [44.7 - 495.6]	0,432	135.9 [50.2 - 373.4]	116.7 [18.4 - 200]	0,002
	V2	360.6 [134.6 - 915.7]	212.2 [141.2 - 1097.8]	0,465	288.8 [158 - 665.2]	197.7 [63.6 - 335.6]	0,001

Tabla 4.1. Mediana [1r y 3r cuartil] y P valor, de los biomarcadores en SL entre 0 y 2 años (izquierda) y 0 y 4 años (derecha)

		ECG_0 (N = 11)	ECG_2 (N = 11)	P valor	ECG_0 (N = 11)	ECG_4 (N = 11)	P valor
Voltaje pico a pico (mV)	aVF	0.49 [0.37 - 0.68]	0.48 [0.2 - 0.55]	0,064	0.49 [0.44 - 0.63]	0.45 [0.23 - 0.49]	0,009
	V1	1.16 [0.63 - 1.78]	0.93 [0.59 - 1.83]	0,105	1.01 [0.54 - 1.43]	0.87 [0.38 - 1.04]	0,007
	V2	1.42 [0.86 - 2.49]	1.04 [0.93 - 2.64]	0,846	1.54 [1.06 - 1.81]	1.17 [0.83 - 1.46]	0,001
Duración del QRS (ms)		133 [125 - 173]	130 [123 - 174.75]	0,797	127 [118 - 135]	129 [116 - 139]	0,151
Intervalo QT (ms)		448.5 [405 - 493]	468 [435.25 - 483]	0,557	418 [390 - 433]	438 [411 - 488]	0,380
Intervalo QT corregido (ms)		442 [421 - 479]	441 [426 - 484]	0,492	415 [401 - 440]	426 [406 - 453]	0,677
Área del QRS en valor absoluto (mV · s)	V1	42.3 [26.8 - 65.4]	37.2 [17.8 - 61.3]	0,275	32.3 [19.1 - 50.2]	29.7 [11.5 - 37.1]	0,003
	V2	55.8 [38.6 - 83.4]	57.5 [32.4 - 88.3]	0,695	46.3 [32.7 - 60.5]	38.3 [22.9 - 49.5]	0,001
Energía del QRS (mV <sup>2</sup> · s · 1000)	V1	26.2 [9.4 - 8.7]	22.9 [4.6 - 58.2]	0,432	16.23 [5.84 - 41.7]	13.4 [2.2 - 23]	0,007
	V2	44.3 [17.6 - 115.4]	28.2 [16.7 - 132.9]	0,557	32.2 [18.5 - 74.7]	23.4 [7.2 - 41.3]	0,001
Potencia media QRS (mV <sup>2</sup> · 1000)	V1	194.1 [49.3 - 766.8]	158 [37.1 - 463.9]	0,375	117.4 [44.4 - 332.2]	112.7 [17.7 - 178.4]	0,005
	V2	345.0 [87.5 - 880.5]	189 [136.8 - 1055]	0,492	268.4 [147.6 - 697]	168.7 [58.7 - 309.2]	0,001

Tabla 4.2. Mediana [1r y 3r cuartil] y P valor, de los biomarcadores en SLR entre 0 y 2 años (izquierda) y 0 y 4 años (derecha)

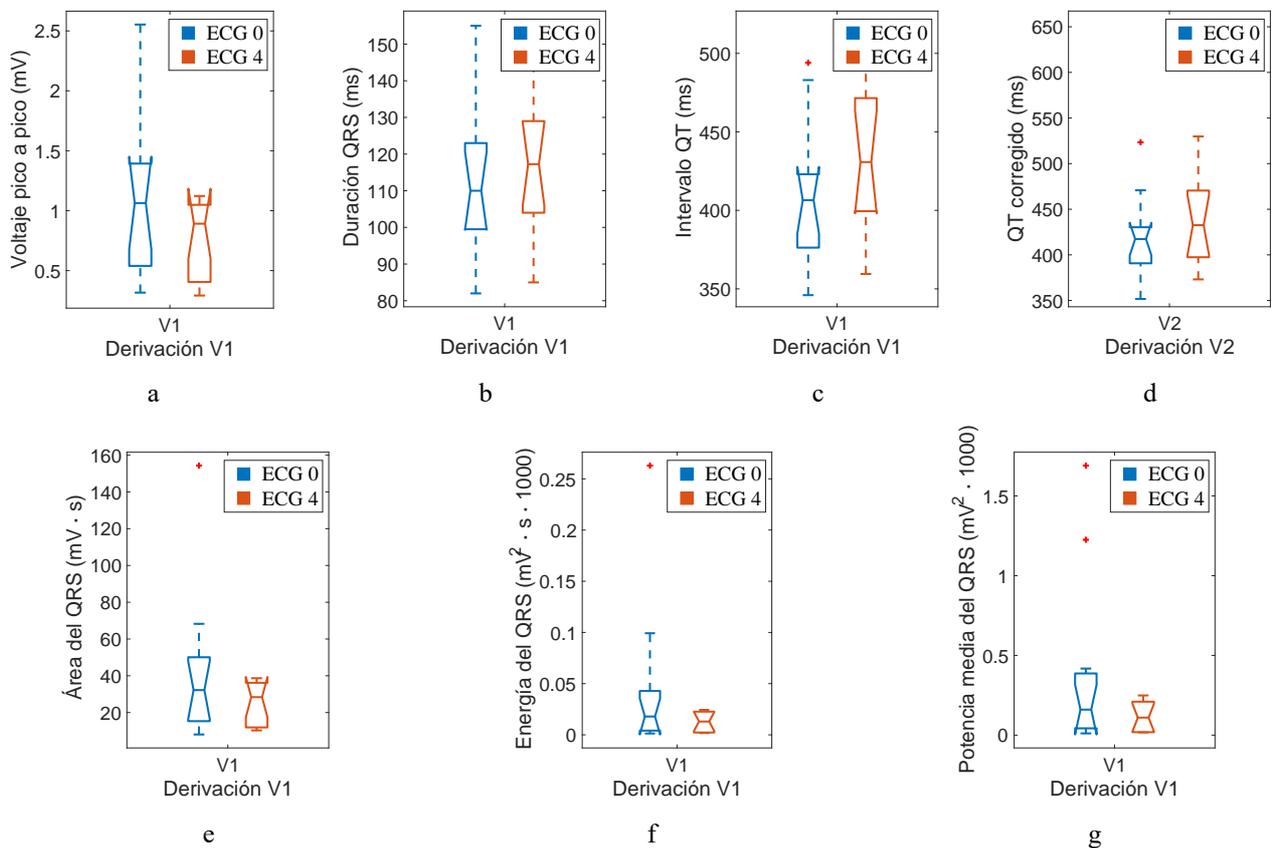


Figura 4.1. Distribución del voltaje pico a pico (a), duración del QRS (b), intervalo QT (c), intervalo QT corregido (d), área del QRS (e), energía del QRS (f) y la potencia media del QRS (g) entre 0 y 4 años en SL.

Al comparar los valores de las medidas en Tabla 4.1 y 4.2, y observando la Figura 4.1 se encuentra que existe una disminución de la amplitud del voltaje pico a pico del QRS, así como del valor absoluto del área, de la energía y la potencia del QRS tanto en SL como en SLR, al cabo de 4 años. Estos resultados se presuponen que tienen relación

con la evolución de la enfermedad, ya que la ALVC genera tejido fibroso en el ventrículo izquierdo, lo que provoca una pérdida de fuerza en el impulso cardíaco.

En las comparaciones de los valores del estudio temporal de 2 años de diferencia, también se observan la disminución de las medianas para estos biomarcadores. Sin

embargo, los cambios no son significativos. Por lo tanto, el factor temporal forma un papel importante en el desarrollo de la enfermedad y su afección al corazón.

Por otro lado, se encontraron aumentos significativos de las medianas en la duración del QRS, el intervalo QT y el corregido en el análisis a 4 años tanto en SL como en SLR. Esto puede indicar que ALVC genera una disminución de la capacidad de realizar el latido que provoca una duración más larga de este.

En la comparación entre el uso de las dos metodologías utilizadas para calcular las marcas (SL y SLR) no se han observado diferencias significativas. Esto podría ser una consecuencia del uso la mediana al analizar los biomarcadores de los registros, en lugar de la media.

Por último, en los análisis por diferencias intergrupales de sexo y diagnóstico de hipertensión no se ha encontrado ningún resultado de interés, para ninguno de los biomarcadores analizados, probablemente debido al reducido número de mujeres y hipertensos respecto al otro grupo de análisis.

Las principales limitaciones del estudio son el reducido número de pacientes (11 registros por cada grupo al evaluar los ECGs al cabo de 2 y 4 años respecto el primero), el desbalance entre sexo (N=8 mujeres y N=20 hombres) así como en el número de hipertensos (N=6) respecto a no hipertensos (N = 22).

## 5. Conclusiones

En este estudio se ha analizado ECGs de pacientes con AVLC, extrayendo biomarcadores relacionados con la despolarización de los ventrículos, ya que dicha patología disminuye la capacidad de bombear de los ventrículos.

En conclusión, los biomarcadores relacionados con el complejo QRS (voltaje, área, energía y potencia media del QRS) resultan un buen indicador del desarrollo de ALVC. Por lo tanto, se recomienda la observación de estos biomarcadores comparando su medida entre pacientes con ALVC y pacientes sanos para identificar si aparte de caracterizar la enfermedad, pueden resultar en una fuente fiable para el diagnóstico de la enfermedad.

## Agradecimientos

F. Palmeri desea agradecer el apoyo de Product & Technology, AGAUR, Generalitat de Catalunya DI001-2018.

## Referencias

- [1] D. Corrado y C. Basso, «Arrhythmogenic left ventricular cardiomyopathy,» *Heart*, vol. 108, nº 9, pp. 733-743, 2021.
- [2] D. Colorado, M. Perrazolo Marra, A. Zorzi, G. Beffagna, A. Cipriani, M. De Lazzari, F. Migliore, K. Pilichou, A. Rampazzo, I. Rigato, S. Rizzo, G. Thiene y A. Anastasakis, «Diagnosis of arrhythmogenic cardiomyopathy: The Padua criteria,» *International Journal of Cardiology*, vol. 319, pp. 106-114, 2020.
- [3] D. Corrado, A. Zorzi, A. Cipriani, B. Bauce, R. Bariani, G. Beffagna, M. De Lazzari, F. Migliore, P. Kalliopi, A. Rampazzo, I. Rigato, S. Rizzo, G. Thiene y P. Ma, «Evolving Diagnostic Criteria for Arrhythmogenic

Cardiomyopathy,» *Journal of the American Heart Association*, vol. 10, no 18, p. e021987. 2021.

- [4] P. Laguna y L. Sörnmo, *Bioelectrical Signal Processing in Cardiac and Neurological Applications*, Elsevier/Academic Press, 2005.
- [5] J. Pablo Martínez, R. Almeida, S. Olmos, A. P. Rocha y L. Paula, «A Wavelet-Based ECG Delineator: Evaluation on Standard Databases,» *IEEE Transactions on Biomedical Engineering*, pp. 570-581, 2004.
- [6] P. Gomis, «Señales en tiempo discreto de valores discretos,» de *Estimación Espectral de Señales Biomédicas Métodos Clásicos (FFT) y Paramétricos: Aplicaciones Prácticas con Matlab*, UPC commons, 2007, pp. 4-9. [https://upcommons.upc.edu/bitstream/handle/2117/330688/tutorial\\_estimacion\\_espectral-5308.pdf](https://upcommons.upc.edu/bitstream/handle/2117/330688/tutorial_estimacion_espectral-5308.pdf)
- [7] K. Kirk y A. MacSween, «SPSS Workbook,» de *Statistical Tests Using Spss Software*, Edinburgh, Queen Margaret University College, 2003.

# Desarrollo de un modelo de colon para formación en cirugía laparoscópica a partir de técnicas de procesado semiautomático de imagen médica e impresión 3D

Carlos Lobato Gómez, Èlia Pedregosa, Juan A. Sánchez-Margallo\*, David. Patrocinio Caballero, Francisco M. Sánchez-Margallo

Centro de Cirugía de Mínima Invasión Jesús Usón, Cáceres, España,  
{clobato, epedregosa, jasanchez, dpatrocinio, msanchez}@ccmijesususon.com

\*Corresponding author

## Resumen

*El aprendizaje de técnicas laparoscópicas necesita de simuladores más realistas para favorecer la adquisición de habilidades quirúrgicas y completar los programas de aprendizaje tradicionales. Los simuladores quirúrgicos actuales son útiles para los cirujanos noveles, pero tienen un uso limitado en técnicas quirúrgicas más complejas. La finalidad de este estudio es mejorar las técnicas de elaboración de simuladores para cirugía laparoscópica que permitan obtener resultados más realistas y que estos puedan contribuir a la mejora del aprendizaje de los profesionales sanitarios. En concreto, en este trabajo se presenta la metodología para la creación de simuladores quirúrgicos colorrectales. Partiremos de un método para la segmentación semi-automática del colon, a partir de un estudio DICOM, y que permita obtener modelos de colon de una manera más eficiente, para pasar a detallar la metodología para la elaboración del simulador de colon correspondiente mediante el uso de diversas técnicas de modelado e impresión 3D.*

## 1. Introducción.

La cirugía laparoscópica es una disciplina quirúrgica técnicamente muy demandante para el cirujano, la cual requiere de un amplio periodo de formación hasta alcanzar las competencias quirúrgicas necesarias para su aplicación en la práctica clínica real [1]. Es por ello los continuos esfuerzos por desarrollar nuevas herramientas y modelos de formación que permitan reducir el uso de animales de experimentación y cadáveres y que faciliten la práctica quirúrgica para favorecer la adquisición de habilidades psicomotoras laparoscópicas previo a la práctica quirúrgica real, como es en el uso de simuladores de formación (box trainers).

El cáncer colorrectal (CCR) es la tercera neoplasia maligna más comúnmente diagnosticada en el mundo y la cuarta causa principal de muerte por cáncer [2]. En España supone más de 15.000 fallecimientos anuales (Sociedad Española de Oncología Médica -SEOM-). Aproximadamente el 80% de los casos recién diagnosticados de cáncer de colon requieren cirugía. La cirugía sigue siendo el tratamiento de elección para la mayoría de los casos de CCR, tanto como tratamiento con intención curativa como terapia paliativa [3]. En el caso de la cirugía mínimamente invasiva con abordaje laparoscópico, esta cirugía se lleva a cabo por medio de incisiones en la pared abdominal a través de las cuales se insertan los trocares para poder manejar los

instrumentos quirúrgicos y la óptica laparoscópica. Los estudios confirman cada vez más que las resecciones laparoscópicas son seguras, oncológicamente equivalentes a la cirugía abierta y tienen mejores resultados a corto plazo [4]. Sin embargo, esta técnica es técnicamente muy compleja y presenta una curva de aprendizaje pronunciada. En la actualidad, existen escasos modelos artificiales realistas para la práctica laparoscópica de cirugía del colon, que faciliten el entrenamiento de los profesionales sanitarios en estas técnicas quirúrgicas previo a su aplicación a la práctica clínica real [5].

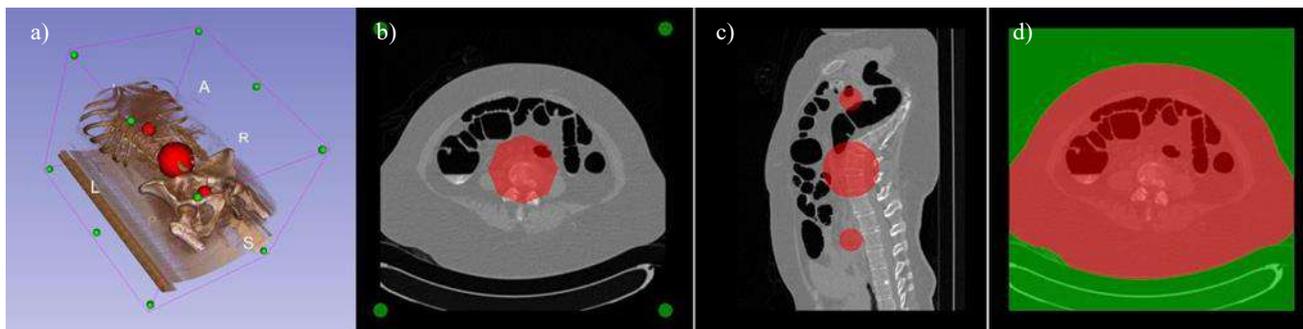
En este trabajo presentamos los primeros pasos para la aplicación de diferentes métodos para el diseño y desarrollo de un modelo de colon realista basado en técnicas de análisis de imagen médica combinadas con técnicas actuales de impresión 3D y el uso de nuevos materiales.

## 2. Materiales y métodos.

Para el desarrollo del modelo de colon para su uso en formación en cirugía laparoscópica, se ha seguido el siguiente flujo de trabajo: (1) Obtención del estudio DICOM; (2) Segmentación del modelo; (3) Edición y preparación del modelo; y (4) Técnicas de modelado e impresión 3D

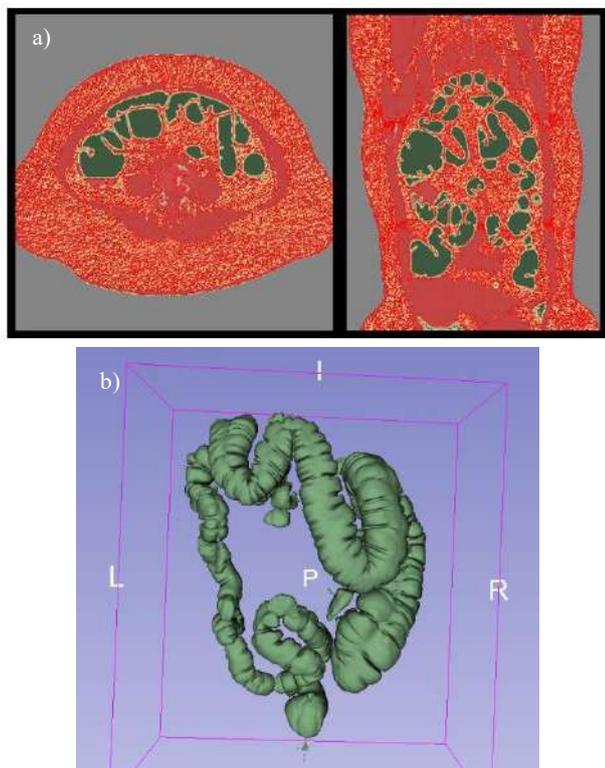
### 2.1. Segmentación del modelo.

La segmentación de imágenes médicas es un proceso fundamentalmente manual que requiere de un trabajo meticuloso y de larga duración. Para mejorar la eficiencia del proceso de modelado se ha desarrollado una herramienta de segmentación semi-automática centrada en la estructura anatómica del colon. Para ello, se ha utilizado el software 3DSlicer ([www.slicer.org](http://www.slicer.org)), un programa de código abierto para la visualización y procesamiento de imágenes médicas. Este software permite incorporar herramientas de elaboración propias a través de módulos programables en Python. El método de segmentación presentado se ha desarrollado siguiendo la línea de estudios anteriores [6], en los cuales se han utilizado diversos algoritmos de procesado de imagen para obtener un modelo 3D del colon a partir de estudios de Tomografía Axial Computerizada (TAC) del paciente. Este estudio materializa distintas técnicas presentadas en el estado del arte en una herramienta de software libre a mediada para su aplicación en cirugía colorrectal, la cual podrá



**Figura 1.** a) Se muestra una vista en 3D del TAC del paciente, donde las semillas verdes se sitúan en los bordes de la matriz y las semillas rojas se sitúan dentro del cuerpo del paciente. b) Vista axial del TAC con las semillas. c) Vista sagital del TAC del paciente. d) Se muestra en vista axial el TAC del paciente con el procesado del algoritmo de Crecimiento de regiones.

descargarse de manera gratuita desde el repositorio de 3DSlicer. Para constatar el correcto funcionamiento, se han utilizado 20 modelos de distintos pacientes obtenidos de The Cancer Imaging Archive, una base de datos pública proporcionada por el National Cancer Institute (NCI).



**Figura 2.** Proceso de segmentación. a) Vistas axial y coronal en las que se muestra (en verde) la segmentación del gas. b) Resultado del modelo 3D del colon segmentado.

La problemática inicial que presenta el procesado del TAC de abdomen, es la fragmentación de la zona anatómica del colon. La presencia de obstrucciones causadas por los movimientos peristálticos o las heces dificulta la segmentación y la distribución del gas a lo largo del mismo no suele ser uniforme. Es, por lo tanto, recomendable que el paciente se someta a un tratamiento, el cual consiste en una dieta blanda para evitar la acumulación de heces y la toma de un contraste yodado, haciendo que cualquier líquido remanente quede radioopaco para facilitar su posterior procesado. Por otro lado, para poder visualizar

correctamente el colon dentro del TAC y mejorar el automatismo de la segmentación, será necesario realizar una insuflación de gas en el colon del paciente durante la adquisición de imágenes. Para ello, se emplea CO<sub>2</sub> ya que ha demostrado ser menos molesto y con menos riesgo para el paciente [7]. Con todo ello, se consigue facilitar el proceso de segmentación del colon.

La presencia de gas dentro del colon permite facilitar la diferenciación de la luz (interior) del colon de sus paredes. Hay que tener en cuenta que el tejido del colon tiene una densidad muy parecida al de los tejidos adyacentes y dificulta el filtrado, por lo que se aplica un umbral a aproximadamente  $-800$  Unidades Hounsfield para descartar los vóxeles con radiodensidades superiores. Otro de los aspectos a tener en cuenta para automatizar este proceso de segmentación es que hay distintas cavidades en el abdomen las cuales contienen gas, a parte del colon, incluyendo el aire exterior al paciente.

Para eliminar el aire exterior se ha empleado un algoritmo tipo Crecimiento de Regiones, donde se colocan semillas en el cuerpo del paciente y en la zona del aire que lo rodea. Mediante este algoritmo de crecimiento las semillas van inundando los vóxeles colindantes cuando comparten similitud hasta encontrarse con las semillas contrarias. En los márgenes de la matriz se colocan doce semillas verdes (Figura 1b). En Figura 1c se puede apreciar el mismo paciente desde la vista sagital, a lo largo del mismo se han colocado tres semillas rojas. En la Figura 1d se puede apreciar cómo las semillas han crecido hasta ocupar todo el volumen disponible. Una vez diferenciados ambos espacios (aire exterior y paciente), se puede excluir el volumen de aire exterior utilizando el segmento creado con las semillas verdes para enmascararlo.

Para eliminar de la segmentación el resto de cavidades que puedan contener gas se ha programado un proceso con interrupciones, donde se le pide al usuario que coloque manualmente las semillas de crecimiento tanto en los pulmones, como en el estómago o en el intestino delgado, en caso de que el CO<sub>2</sub> ha conseguido traspasar más allá del colon.

## 2.2. Edición y preparación del modelo.

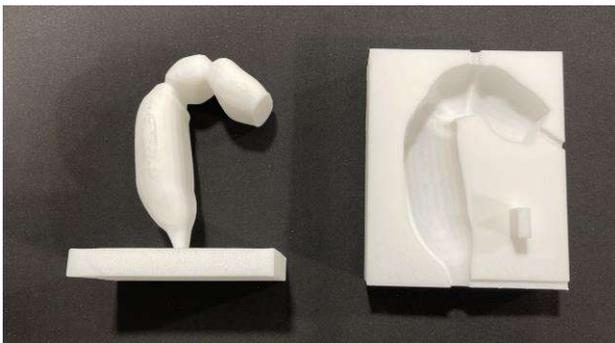
Para poder crear una versión imprimible del modelo, se usa el programa informático Blender, el cual es un software multiplataforma dedicado especialmente al modelado,

renderizado, la animación y la creación de gráficos 3D. Este software se ha utilizado para editar el modelo segmentado del colon obtenido 3D Slicer, el cual ha sido exportado previamente en una extensión editable. Se posprocesó el modelo para poder quitar imperfecciones y terminar de adecuar este para la futura impresión en 3D.

### 2.3. Técnicas de modelado e impresión.

Una vez terminada la preparación del modelo, el siguiente paso es poner las diferentes técnicas de impresión y modelado a punto. Se ha trabajado paralelamente en dos técnicas diferentes: modelado por deposición fundida (FDM) y técnica de moldes por pincelado. Los resultados obtenidos por cada técnica han sido analizados y comparados para elegir la más adecuada al objetivo del estudio.

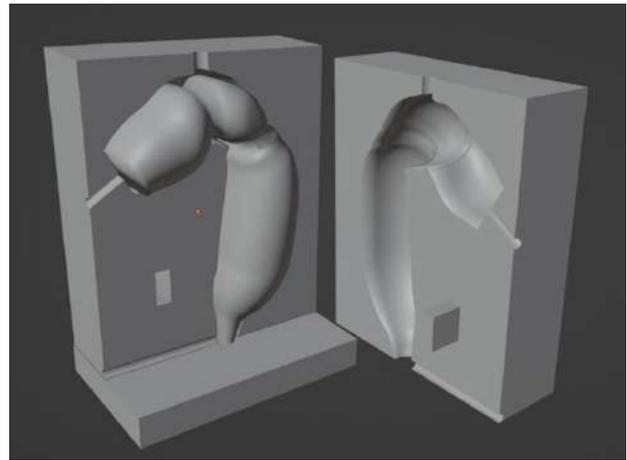
La tecnología de FDM permite crear modelos, prototipos y piezas finales en materiales termoplásticos de alto rendimiento. En nuestro caso, esta técnica se ha utilizado para imprimir un molde de la zona anatómica del recto y parte del colon, los cuales han sido obtenidos en paso previos a partir de un estudio de TAC del paciente (Figura 3). Este molde en bloque ha sido impreso en un polímero termoplástico basado en una mezcla de poliestireno y caucho de polibutadieno (HIPS), caracterizado por ser un material duro y resistente a los impactos (Figura 4).



**Figura 4.** Molde en bloque de colón impreso en HIPS.

El molde en bloque está compuesto por dos partes, la figura del recto y parte del colon, y sus dos negativos. Esta técnica nos permitirá reproducir posteriormente modelos de silicona que tengan detalles en todas sus caras. El conjunto de los dos negativos del modelo crea una caja alrededor, asegurándose de dejar una separación mínima entre ambos objetos para rellenarlo posteriormente de silicona. Es necesario que el modelo se fije a la base de la caja, para que se mueva lo menos posible durante el vertido de la silicona. Para ello, se diseñó una pestaña extra para una mayor fijación entre piezas y varios orificios de salida para la silicona sobrante. Una vez depositada la silicona en el interior del modelo por el orificio de entrada, y esta se haya curado correctamente, la caja fue desmontada para poder separar adecuadamente la silicona del modelo interno.

Respecto a las pruebas con la técnica de moldes por pincelado, su proceso de fabricación tiene similitudes con la técnica anterior, ya que el molde empleado se creó con la misma técnica de impresión (FDM) y material (HIPS)



**Figura 3.** Representación 3D en Blender de las partes del modelo en bloque del colon.

que para el molde en bloque. Una vez obtenido el molde, se recubrió con una fina película de vaselina y se aplicaron capas de silicona a través de pinceladas sobre el modelo hasta alcanzar el grosor esperado.

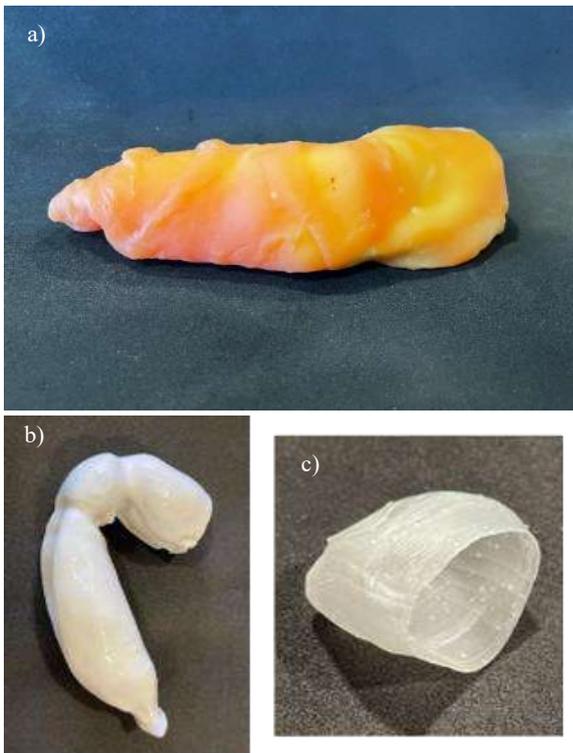
En ambos experimentos, se usaron las siliconas de la serie Ecoflex 0030 (Figura 5a) y de la serie Roma de Sumbeart (Figura 5b). Para conseguir una textura más realista del modelo, a la serie Ecoflex se le añadió Slacker, un aditivo que provoca un cambio de viscosidad en la textura final del modelo.

Paralelamente a la tecnología FDM, se realizaron también pruebas con la tecnología de impresión en 3D por estereolitografía (SLA). Para llevar a cabo las primeras pruebas, se empleó una pequeña parte del modelo de colon en 3D segmentado. En este caso, el tipo de material empleado ha sido resina Elastic 50 de Formlabs. Esta resina está diseñada para aplicaciones que requieran elongación, resistiendo a ciclos repetidos de deformación sin desgarro. Se trata de un material transparente, por lo que resulta muy adecuado para el desarrollo de modelos médicos para la simulación o formación. Una vez finalizada la impresión con resina, fue necesario hacer un proceso de poscuración para obtener mejores resultados. Para ello, se sumergió la pieza en un baño de isopropanol durante 10 minutos y, finalizado este proceso, se sometió a la pieza a una curación bajo temperatura y luz ultravioleta (Figura 5c).

## 3. Resultados y discusión.

El método de segmentación semi-automático proporciona unos resultados aceptables. Sin embargo, una vez finalizado el proceso aún quedan pequeños fragmentos no pertenecientes al colon que no consiguen eliminarse, por lo que es necesario un posprocesado manual.

Se ha medido el tiempo de cómputo para 20 modelos distintos, utilizando un procesador 12th Gen Intel(R) Core (TM) i7-12700H 2.30 GHz y 16 Gb de RAM. En la Figura 6 se compara con el tiempo empleado para obtener el mismo resultado segmentando de forma manual.

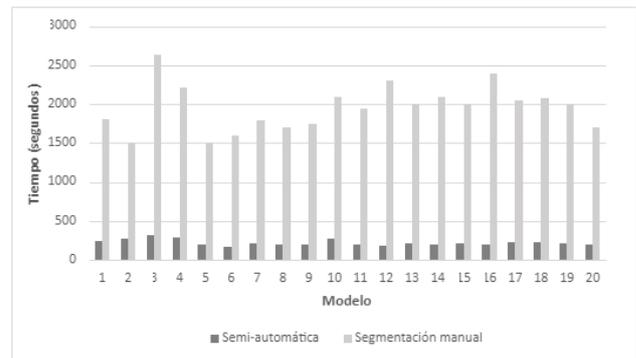


**Figura 5.** Resultado de los modelos. a) Se muestra el modelo de silicona Ecoflex. b) Resultado de utilizar la silicona Roma. c) Resultado en resina elástica impresa.

Actualmente, el procesado únicamente tiene en cuenta el gas contenido en el colon, por lo que los restos de líquido y material fecal que quedan retenidos no son incluidos, formando imperfecciones en el segmento final. Esto se corregirá como trabajo futuro.

En cuanto al proceso de fabricación realizado con la técnica en bloque, se descarta seguir el estudio con este método, ya que el proceso de diseño es poco eficiente y presenta problemas, dado que la silicona tiende a escaparse por las ranuras. Por otra parte, los resultados obtenidos a partir de la técnica por pincelado han sido mejores. Como se puede observar en la Figura 5a, la forma del modelo de silicona Ecoflex no se ha conseguido tan bien como con la silicona Roma (Figura 5b), pero presenta una textura más realista y que podría semejarse a la de un colon real. Es por eso, que se seguirá con el estudio utilizando la silicona de la serie Ecoflex, aplicando la técnica del *rotocasting* [8] para conseguir un grosor más homogéneo. Los resultados obtenidos a partir de la impresión por SLA, han sido satisfactorios hasta el momento (Figura 5c). Para futuros estudios, se continuará trabajando para obtener un modelo de mayor tamaño del que actualmente se ha logrado.

En este trabajo se ha comprobado la viabilidad de diferentes técnicas para reproducir un simulador quirúrgico laparoscópico. Para ello, se ha diseñado una herramienta de procesado de imagen semi-automática, la cual ha demostrado que permite obtener modelos de colon adaptados al paciente de forma eficiente. Con las diferentes técnicas de modelado utilizadas se han conseguido resultados prometedores, por lo que se seguirá trabajando en estos procedimientos para conseguir recrear un simulador al completo.



**Figura 6.** Comparación en segundos del tiempo empleado en la segmentación semi-automática y la segmentación manual.

## Agradecimientos

Este trabajo ha sido parcialmente financiado por la Junta de Extremadura, el Fondo Europeo de Desarrollo Regional (FEDER) y el Programa Erasmus+ de la Unión Europea (TA18023, IB20189, 621668-EPP-1-2020-1-ES-EPPKA2-KA).

## Referencias

- [1] Sánchez-Margallo FM, Sánchez-Margallo JA, Usón J. Addressing the Challenges of Laparoscopic Surgery. In: Sánchez-Margallo FM, Sánchez-Margallo JA, editors. Recent Advances in Laparoscopic Surgery. IntechOpen; 2019. p. 1–5.
- [2] Arnold M, Sierra MS, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global patterns and trends in colorectal cancer incidence and mortality. Gut. 2017 Apr;66(4):683–691.
- [3] George J, Rockall T. Surgery for colorectal cancer. Surg. 2020 Jan;38(1):32–7.
- [4] Jayne DG, Thorpe HC, Copeland J, Quirke P, Brown JM, Guillou PJ. Five-year follow-up of the Medical Research Council CLASICC trial of laparoscopically assisted versus open surgery for colorectal cancer. Br J Surg 2010; 97: 1638e45.
- [5] Sánchez-Margallo F, Durán-Rey D, González-Portillo M, López-Agudelo I, Sánchez-Margallo JA. Educational models for training in minimally invasive colorectal surgery. State Soc. 2021;2021(1):115–40.
- [6] Wyatt, C. L., Ge, Y., & Vining, D. J. (2000). Automatic segmentation of the colon for virtual colonoscopy. Computerized medical imaging and graphics, 24(1), 1-9.
- [7] Herráiz, M. (2013). Colonoscopia con insuflación de dióxido de carbono: ¿lujo o necesidad?. Gastroenterología y Hepatología, 36(1), 43-47.
- [8] Jaycon Systems. (2017, agosto 1). *All about rotocasting: What you need to know*. Jaycon Systems | Product Design, PCB & Injection Molding.

# Simulador clínico para el entrenamiento de la técnica de ecografía torácica y tratamiento de neumotórax y derrame pleural en neonatos

A. Rojas García<sup>1</sup>, M. Otero Arteseros<sup>1</sup>, D. Moreno Blanco<sup>1</sup>, F. J. Rubio<sup>2</sup>, H. Peinado<sup>3</sup>, D. Elorza Fernández<sup>3</sup>, E. J. Gómez<sup>1,4</sup>, M. Quintana<sup>2</sup>, P. Sánchez Gonzalez<sup>1,4</sup>

<sup>1</sup> Grupo de Bioingeniería y Telemedicina, ETSI Telecomunicación, Centro de Tecnología Biomédica, Universidad Politécnica de Madrid, España, {m.oteroa, adriana.rojasg}@alumnos.upm.es {diego.morenob, p.sanchez, enriquejavier.gomez}@upm.es

<sup>2</sup> Fundación para la Investigación Biomédica del Hospital Universitario La Paz de Madrid

<sup>3</sup> Servicio de Neonatología, Hospital Universitario La Paz de Madrid, 28046 Madrid, España

<sup>4</sup> Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina, Madrid, España

## Resumen

*Uno de los retos fundamentales del ejercicio de la medicina es sin duda la seguridad. Seguridad entendida como la reducción de riesgos clínicos y la iatrogenia. Para ello, es evidente que el entrenamiento y la práctica de habilidades técnicas se convierten en herramienta fundamental para el dominio del ejercicio profesional. Derivada de esta necesidad, en los últimos años se ha fomentado considerablemente el empleo de simuladores en ámbito clínico, en escenarios tanto de entrenamiento como de aprendizaje. Los simuladores permiten crear escenarios reales velando por la seguridad del paciente, permitiendo la realización de múltiples procedimientos y obteniendo una retroalimentación del ejercicio, colaborando con la calidad del aprendizaje. En este trabajo se ha desarrollado un sistema de control y monitorización neonatal que permite la evaluación clínica de la técnica de diagnóstico por imagen de la ecografía pulmonar, así como dos mecanismos de resolución de patologías pulmonares, el neumotórax y el derrame pleural. Para el desarrollo del dispositivo se ha diseñado un tórax neonatal, así como las estructuras que lo complementarán interna y externamente. Dichos diseños anatómicos se encuentran en estrecha relación con la estructura electrónica que permite el funcionamiento del sistema y la simulación de la ecografía, derrame y neumotórax neonatal. Asimismo, se ha diseñado el software que controla la electrónica y permite la completa integración de los dispositivos. El simulador desarrollado, SIMUNEO, es un sistema portable, fácil de utilizar y de bajo consumo, características que lo hacen atractivo para introducirlo al mercado.*

## 1. Introducción

El neumotórax es una patología pulmonar desencadenada por la entrada de aire en la cavidad torácica en el espacio pleural situado entre las pleuras parietal y visceral [1]. La presencia de aire en el espacio pleural se debe a la aparición de uno de estos tres eventos: comunicación entre la pleura y los espacios alveolares, comunicación directa o indirecta entre la atmósfera y el espacio pleural o la presencia de gas desencadenando la aparición de organismos en el espacio pleural [2]. El neumotórax se divide en traumático, iatrogénico, catamenial y espontáneo subdividiéndose este último a su vez en primario y secundario [1].

Los neumotórax neonatales son una afección relativamente frecuente. Su frecuencia oscila alrededor del 1-2% de los

recién nacidos a término y del 6% en los recién nacidos pretérmino [3].

La acumulación patológica de líquido en el espacio pleural se denomina derrame pleural (DP). El líquido pleural (LP) se encuentra situado en espacio pleural lubricando las paredes que lo conforman, la pleura parietal que recubre la pared torácica y la visceral que recubre el pulmón. El análisis del líquido pleural ofrece una orientación en la naturaleza del derrame pleural, permitiendo la distinción entre hemotórax (acumulación de sangre), quilotórax (acumulación líquido linfático) o empiema (acumulación de pus) [4]. Es importante saber diferenciar dos tipos de derrames, trasudados y exudados, para poder detectar y conocer las causas principales o secundarias y así realizar un correcto diagnóstico de la enfermedad.

La aparición de líquido en el espacio pleural neonatal sigue las mismas reglas que se describen para pacientes adultos [5]. Ante un neonato que presente signos clínicos y radiológicos de derrame pleural, de entre todos los tipos posibles de derrames, se debe pensar en la posibilidad de un quilotórax, por lo que es indispensable evacuar el líquido para su estudio.

Las técnicas de diagnóstico por imagen más empleadas comúnmente para el diagnóstico de patologías respiratorias son TC, angiografía, radiografía de tórax, resonancia magnética, ecografía, entre otras [6].

La primera vez que la ecografía pulmonar se empleó como herramienta de diagnóstico en el síndrome de distrés respiratorio del recién nacido se remonta a los años 90. A diferencia que como ocurre en adultos, las sondas lineales se adaptan mejor a la pared torácica tan fina y los pulmones pequeños que tienen los neonatos [7].

Debido a su frágil condición, en neonatos suelen examinarse exclusivamente dos zonas, el tórax anterior y lateral en posición supina debido a que muchos de ellos no toleran el cambio de posición por inestabilidad cardiorrespiratoria. Para estandarizar la técnica, el tórax es dividido en tres regiones: anterior, lateral y posterior, utilizando las líneas axilares posteriores como límites.

El objetivo principal de este trabajo ha sido desarrollar un simulador que los profesionales puedan emplear con fines educativos formando desde cero a los estudiantes para el manejo del ecógrafo, así como las técnicas de neumotórax y derrame pleural, sin poner en riesgo la seguridad del paciente.

## 2. Diseño

El diseño del simulador SIMUNEO se ha realizado en un proceso colaborativo entre los ingenieros del Grupo de Bioingeniería y Telemedicina y los profesionales del Centro Avanzado de Simulación y Entrenamiento Clínico del Instituto de Investigación del Hospital Universitario de La Paz en Madrid y el servicio de neonatología.

El sistema diseñado tiene la arquitectura que se puede observar en la Figura 1. El sistema consta de una aplicación web a modo de interfaz y un microcontrolador que gestiona todos los sensores y elementos electrónicos de la parte física. Estos dos elementos están conectados a través de un servidor REST y todos los datos de interés son almacenados en una base de datos.

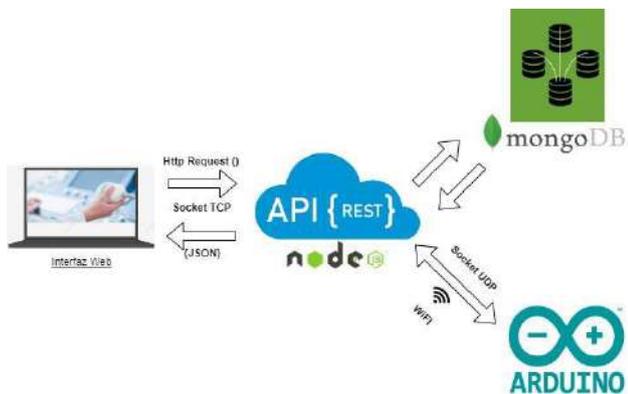


Figura 1. Diagrama de arquitectura del simulador

El diagrama de flujo, que se puede observar en la Figura 2, muestra el funcionamiento completo del simulador.

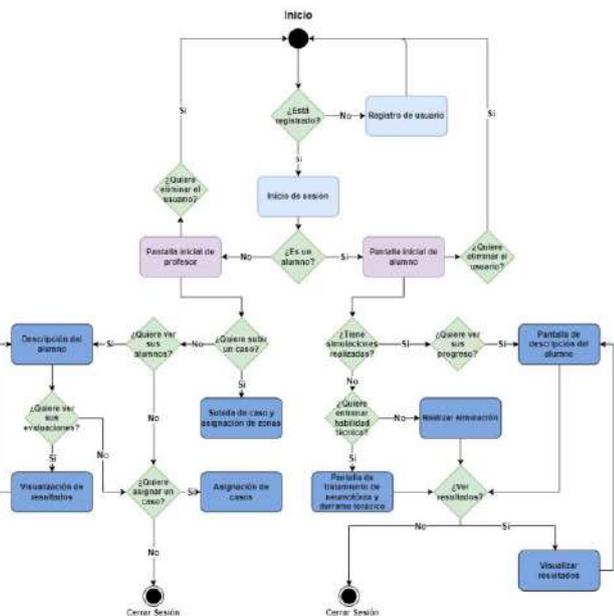


Figura 2. Diagrama de flujo del simulador

### 2.1. Diseño de las estructuras anatómicas

La estructura anatómica se ha diseñado a partir de la reconstrucción e impresión 3D de unas costillas de un neonato segmentadas a partir de una imagen de resonancia magnética. El modelo resultante de la segmentación se ha ajustado en base al criterio de los clínicos para mejorar la experiencia de aprendizaje. Además, se ha diseñado una piel sintética a partir de siliconas que aporte un mayor realismo al simulador. También se ha estudiado en esta fase de diseño el uso de otros materiales para simular estructuras anatómicas importantes en este procedimiento, como puede ser la pleura.

### 2.2. Diseño de la electrónica y red de sensores

Para los componentes electrónicos se han diseñado dos circuitos. El primero, que se observa en la Figura 3, se han colocado 6 sensores reed para detectar la posición de la sonda ecográfica que lleva un imán. Esto permite simular las diferentes zonas ecográficas.

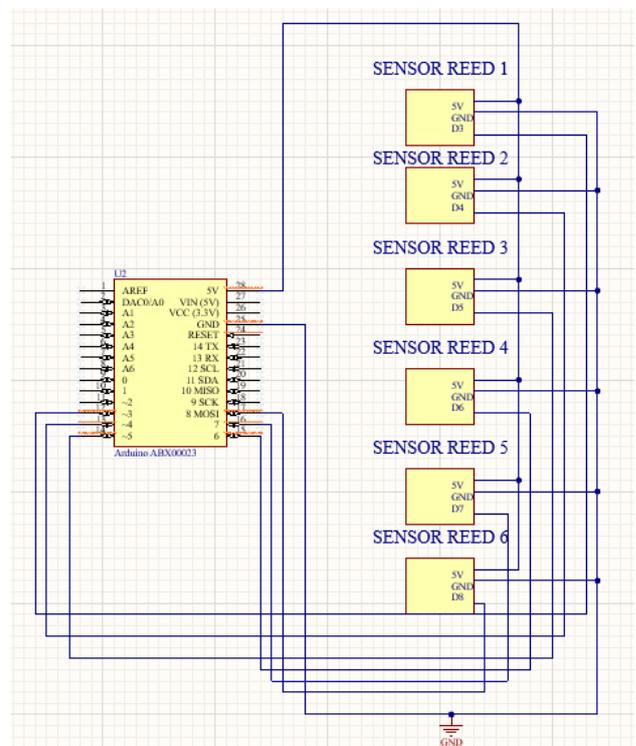


Figura 3. Diagrama de la red de sensores para la simulación ecográfica

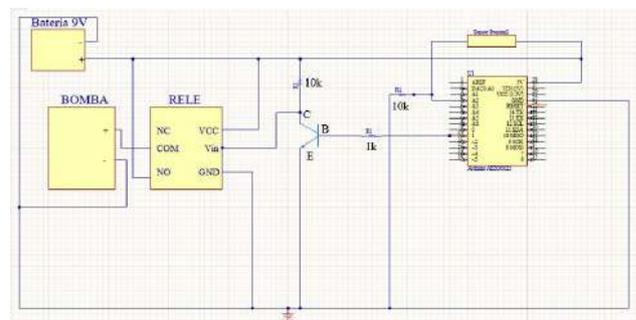


Figura 4. Diagrama de los sistemas de simulación de neumotórax y derrame pleural

El segundo circuito es el encargado de las simulaciones de neumotórax y derrame pleural. En la Figura 4 se observan las interconexiones de los sensores de presión con los relés que activan las bombas de aire y agua en cada caso. Estos relés son alimentados mediante una batería externa

### 2.3. Diseño de la aplicación web y componentes software

Para el diseño de la aplicación web, se ha desarrollado un mockup que ha sido validado con el equipo clínico. En la Figura 5 puede observarse la pantalla de inicio diseñada en el mockup a modo de ejemplo. En este mockup se han incluido todas las pantallas previstas en la aplicación web.

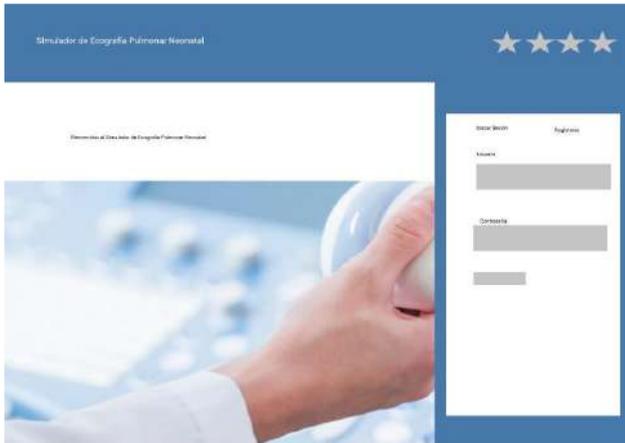


Figura 5. Pantalla de inicio de sesión en el mockup diseñado

### 3. Resultados

En la Figura 6 se puede observar el sistema final desarrollado. En el se observan las costillas impresas en 3D con todos los componentes electrónicos integrados. Además, también se han creado cajas impresas en 3D para almacenar todos los componentes y evitar que el agua utilizada para simular los derrames pueda alcanzar los componentes electrónicos.

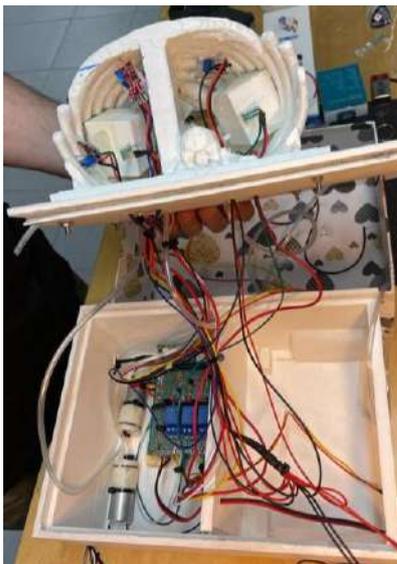


Figura 6. Sistema final desarrollado

En la Figura 7 se puede observar el resultado final una vez cubierto con la piel artificial creada a partir de las siliconas.

Esto dota al sistema de un mayor realismo. Además, en esta figura también se observa la sonda ecográfica utilizada a la que se ha añadido un imán para activar los sensores específicos.



Figura 7. Vista final del sistema con la piel artificial

Finalmente, en la Figura 8 se puede observar el uso real del sistema durante una sesión de validación con los profesionales clínicos.



Figura 8. Sistema final utilizado por un participante

La Figura 9 y la Figura 10 muestran capturas reales de la aplicación web. En concreto muestran las pantallas referentes a la resolución de caso mediante la ecografía y a la visualización de resultados.

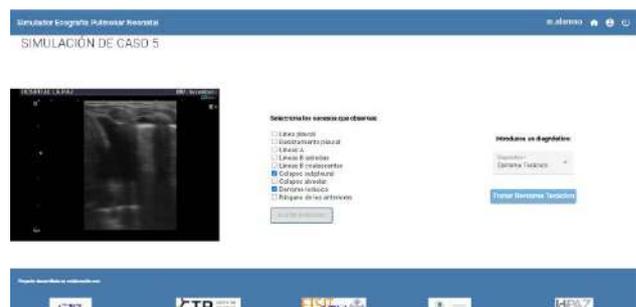


Figura 9. Pantalla de resolución de caso en la aplicación web final

En la resolución de los casos, el estudiante deberá recorrer las distintas zonas ecográficas definidas intentando diagnosticar al paciente. Una vez emitido el diagnóstico, en caso de diagnosticarse un derrame pleural o un neumotórax, se procederá a la simulación del tratamiento de este. Una vez finalizado el caso, tanto el alumno como el tutor responsable podrán revisar la realización del caso y ver los aciertos y errores cometidos.



**Figura 10.** Pantalla de visualización de resultados en la aplicación web final.

En esta pantalla de visualización de resultados, se detallan la descripción del caso que se acaba de simular, los sucesos elegidos correctos e incorrectos, el resultado del diagnóstico establecido y la nota generada automáticamente. El docente también puede acceder a esta pantalla, observando los fallos que más comete el alumno y realizando un seguimiento de su aprendizaje, con el fin de corregir posibles fallos de conceptos de diagnóstico o de patrones ecográficos y poder mejorar la calidad y eficiencia del servicio.

Los resultados y métricas definidos y evaluados se basan en los siguientes factores:

- Sucesos ecográficos correctos e incorrectos: El número de sucesos marcados por el alumno que son correctos menos un cuarto del número de sucesos marcados por el alumno que incorrectos, entre el número de sucesos que se observan en el caso. El resultado de esta operación se multiplica por 10 para obtener una nota numérica del 1-10.
- Diagnóstico: Si el diagnóstico escogido por el alumno es el mismo que el del caso, se suma 2 a la nota obtenida en la comparación de sucesos ecográficos, pudiendo obtener como máximo una nota de 10. Si es distinto, se restan 2 puntos.
- Tratamiento incorrecto: El alumno puede creer oportuno tratar una de las patologías respiratorias que ofrece el simulador. Sin embargo, esta elección puede no concordar con el tipo que presenta el caso. Dicho de otra forma, si el caso simulado es un caso de tipo neumotórax o tipo otro y el alumno marca como opción “derrame torácico” y trata este caso como tal, se restarán 3 puntos a la nota calculada en la comparación de diagnósticos. Si el caso es de tipo neumotórax y el alumno decide tratar el neumotórax, la nota no se verá afectada.

#### 4. Discusión y Conclusiones

Se ha diseñado y desarrollado un prototipo de simulador neonatal con finalidad educativa para velar por la seguridad de los pacientes y complementar la formación de los estudiantes enseñando y practicando el manejo del ecógrafo y la realización de las técnicas de neumotórax y de derrame. El dispositivo está en proceso de validación por los profesionales sanitarios del servicio de neonatología del Hospital Universitario La Paz de Madrid.

SIMUNEO permite familiarizarse con el manejo del ecógrafo y con la resolución de las técnicas de neumotórax y derrame pleural. En la búsqueda de simuladores similares, comerciales o experimentales, realizada hasta la fecha no se ha encontrado ningún simulador que permita realizar la evaluación y las técnicas que ofrece SIMUNEO. Todos los simuladores neonatales que ofrecen la resolución de estas técnicas son de alta fidelidad y de cuerpo completo, lo que encarece considerablemente el producto.

La posibilidad de aportar realimentación al ejercicio es otra de las claves a comparar con los sistemas encontrados, ya que, al ser un sistema con propósitos pedagógicos, resulta imprescindible establecer unas métricas que evalúen el ejercicio. Los materiales de los que está compuesto el sistema, así como no ser un sistema de cuerpo completo hacen que el sistema que se ha desarrollado en este trabajo sea de bajo coste, lo cual es un punto fuerte para considerar a la hora realizar mejoras futuras. Destacar también que La aplicación presenta una transmisión fluida de datos entre los distintos componentes y contiene todos los módulos requeridos por una plataforma de docencia y entrenamiento simulado.

Además de la validación clínica, también se ha diseñado un plan de Control de Calidad del Sistema. Los aspectos que se deben comprobar deberán cumplirse un total de 10 veces para que el control sea satisfactorio al 100%.

#### Referencias

- [1] J. C. Valle Sagastume, “Neumotórax Espontáneo,” *Revista Médica de Costa Rica y Centroamérica*, vol. LXXI, no. 617, pp. 723–727, 2015.
- [2] M. Noppen and T. de Keukeleire, “Pneumothorax,” *Respiration*, vol. 76, no. 2, pp. 121–127, Aug. 2008, doi: 10.1159/000135932.
- [3] L. Fernandez, “Tratamiento de los escapes aéreos en Neonatología Related papers”.
- [4] V. Villena Garrido *et al.*, “Diagnosis and Treatment of Pleural Effusion,” *Archivos de Bronconeumología (English Edition)*, vol. 42, no. 7, pp. 349–372, Jul. 2006, doi: 10.1016/S1579-2129(06)60545-4.
- [5] D. Kurepa, N. Zaghoul, L. Watkins, and J. Liu, “Neonatal lung ultrasound exam guidelines,” *Journal of Perinatology* 2018 38:1, vol. 38, no. 1, pp. 11–22, Nov. 2017, doi: 10.1038/jp.2017.140.
- [6] A. J. E. Moore, J. Wachsmann, M. R. Chamarthy, L. Panjikaran, Y. Tanabe, and P. Rajiah, “Imaging of acute pulmonary embolism: an update,” *Cardiovasc Diagn Ther*, vol. 8, no. 3, p. 225, Jun. 2018, doi: 10.21037/CDT.2017.12.01.
- [7] I. Corsini, N. Parri, B. Ficial, and C. Dani, “Lung ultrasound in the neonatal intensive care unit: Review of the literature and future perspectives,” *Pediatr Pulmonol*, vol. 55, no. 7, pp. 1550–1562, Jul. 2020, doi: 10.1002/PPUL.24792.

# Development of a software for radiopharmaceutical management in nuclear medicine

B. Martínez Ruiz <sup>1</sup>, R. Tudela Fernández<sup>3</sup>, M. Santos Virosta<sup>2</sup>, I. Romero Zayas<sup>2</sup>, F. Campos Añón<sup>2</sup>, R. Sala Llonch<sup>1,3,4</sup>, A. Niñerola Baizán<sup>1,2,3</sup>

<sup>1</sup> Department of Biomedicine, Faculty of Medicine, University of Barcelona, Barcelona, Spain {bmartiru10@alumnes.ub.edu},

<sup>2</sup> Nuclear Medicine Department, Hospital Clínic, Barcelona, Spain {msantosv, icromero, frcampos, ninerola@clinic.cat}

<sup>3</sup>CIBER-BBN, ISCIII, Barcelona, Spain, <sup>4</sup>Institute of Neurosciences, Barcelona, Spain {raul.tudela, roser.sala@ub.edu}

## Abstract

Radiopharmaceuticals are radioactive drugs mostly injected intravenously into patients for diagnostic imaging and therapeutic treatments in nuclear medicine. To achieve an accurate interpretation of the results and a successful therapeutic effect, these medicines must be strictly supervised by the radiopharmacy department from the moment that the raw material arrives at the hospital until it is administrated to the patient.

This work aims to develop a user-friendly and efficient software that integrates all processes involved in the management of radiopharmaceuticals in the Radiopharmacy Unit at the Hospital Clínic of Barcelona. To develop the proposed application, several programming options were evaluated. The final solution was implemented in an open-source Python environment, Anaconda-Navigator. We used the packages Tkinter and SQLite to create the graphical interface and the database communication tools, respectively. Our prototype was correctly implemented at the clinical setting, and we received feedback from three independent users. The mean of the obtained ratings was 3.6 out of 4.

Although being a prototype, our solution was positively ranked, being user-friendly and efficient for performing fast and easy traceability of each radiopharmaceutical dose. Its implementation in the Radiopharmacy Unit will be beneficial for radiopharmaceutical management, as duplications and unnecessary procedures will be removed. In addition, some improvements are suggested to reach its full potential.

## 1. Introduction

Nuclear medicine is a medical discipline that uses radioactive substances, known as radiopharmaceuticals, for diagnostic imaging and therapeutic treatments. They are administrated by intravenous injection, orally or by inhalation. Radiopharmaceuticals emit gamma or positron radiation which decays over time and is detected using specific equipment. Therefore, the radiopharmacist is responsible for ensuring their correct preparation and guaranteeing their safety and efficiency [1]. This is of paramount importance as the accurate interpretation of the obtained images, or the delivery of the right therapeutic dose depends on that [2].

A radiopharmacy workflow, which is represented in Figure 1, starts when a patient needs a diagnostic or therapeutic imaging test and, automatically, a medical prescription for a radioactive tracer is ordered. Then, a radiopharmacist

looks at the stock and executes an order based on the products needed for the exam. These products can be classified into 3 groups depending on their properties: ready-for-use radiopharmaceuticals, generators, and non-radioactive kits. Some days later, the products arrive at the hospital, and they are registered.

Once the patient arrives at the Nuclear Medicine Department, a nurse orders the radiopharmaceutical dose. Depending on the exam and the product type, the medicinal radiocompounds elaboration goes through one of the following paths [3]:

- *Ready-for-use radiopharmaceuticals*: dose measurement and dispensation are required.
- *Extemporaneously prepared radiopharmaceuticals*. These are prepared in aseptic conditions. They are composed of a kit and a <sup>99m</sup>Tc elution that are previously extracted from a <sup>99</sup>Molybdenum (<sup>99</sup>Mo) generator.
- *Autologous radiopharmaceuticals*. Blood cells, such as leukocytes, erythrocytes and platelets, are labelled with a radionuclide (<sup>99m</sup>-Technetium or <sup>111</sup>-Indium).

Finally, a single radiopharmaceutical dose is dispensed which will be administrated to the patient by a nurse.

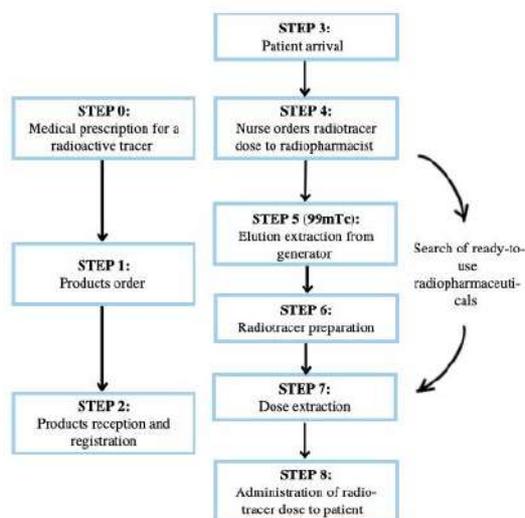


Figure 1. Radiopharmacy Unit's workflow

In the recent years, the amount of data involved in the procedures mentioned above has significantly increased. Therefore, an adequate data management protocol and associated policies will result in benefits for patients, medical research, healthcare systems, and providers [4]. Despite the evolution in technology and the digital transformation of the majority of fields, paper documentation is still present at many hospital departments, including the Radiopharmacy unit at the Hospital Clinic of Barcelona (RUHCB), which is the main focus of this work. Starting from the moment the raw material of radiopharmaceuticals arrives at the radiopharmacy until the patient's administration, radiopharmacists must write down all the operations involved according to the Royal Decree-Law 1345/2007 [5].

The main objective of this project is to integrate all the processes explained above in a centralized and electronic way. It consists of the development of a software for the RUHCB. Although there are some applications available in the market, that partly provide these functions, none of them fulfils all required specifications at a reasonable cost.

## 2. Material and methods

### 2.1. Available data at the Radiopharmacy Unit

Nowadays, the RUCHB stores all the information manipulated throughout the process in paper and in different Excel documents, classified into four main folders: product orders, product receptions, radiolabelling and prescription. All these files are annually generated and saved in the hospital cloud services.

Depending on the radiopharmacy stock, orders are executed and recorded in a file. This process is based on an Excel macro which fills out automatically the stock table and its corresponding order report, which will be sent to the supplier company once an order form is completed. Regarding the product arrival registration, the data of each different kit, ready-for-use radiopharmaceutical, and 99Mo generator is introduced manually in different sheets without any direct relation with its corresponding purchase.

Once radiopharmaceutical preparations are required, technicians write down all information involved such as elution used, activity, or volume, among others, in a textbook as it is more accessible and quicker. Then, it is transcribed into its corresponding Excel document as soon as possible, typically hours or days. There are four different files for each type of preparation which are distinguished by their composition. One document corresponds to radiopharmaceuticals whose radionuclide is 99m-Technetium, and the others are for autologous radiopharmaceuticals as they are made of different cellular components such as leukocytes, erythrocytes or platelets, or distinct radionuclides: 99m-Technetium or 111-Indium.

Finally, patients' doses are dispensed and identified with labels that are previously printed. It's important to highlight that the dispensation step is not systematically

registered, making it difficult to execute a proper dose traceability.

### 2.2. Software development

Before starting to develop the application, we performed a deep study of the different options for each part. The idea was to create a software from scratch, without using specialized software. This would give us more freedom in the fields of customization and external implementations. Therefore, the software was developed in Python, using the Spyder environment within Anaconda-Navigator. The SQLite3 and Tkinter libraries were used for building and manipulating a database and designing a graphical interface, respectively.

Concerning the database management system, data is organized as independent or dependent tables in which end-users can extract, modify, add, and delete entries. The identification of these tables is highly relevant to understand how the system communicates with the database. Based on the available documents, these tables were classified into two main groups: master and transaction tables. The first ones hold general information about products, users or medical exams, and the others are filled with information of the different tasks performed while using the application. These tables are connected by foreign keys which are unique identifications from one row in a relational table to another row in a table [6].

As this software will be used by several medical professionals, such as technicians, radiopharmacists and nurses, the graphical interface should be user-friendly and intuitive. It is composed of 11 blocks and each one is destined to a specific task in the RUCHB's workflow.

To create all these front-ends, the most common widgets applied are: labels for showing static texts, entries to introduce data and buttons to head the user to another window or to execute a specific function such as generating reports and summaries, or updating displayed information. In addition, comboboxes were utilized to drop down a list of items which were previously selected using SQL commands. Finally, we used the calendar widget, treeviews to display tabulated information, and message boxes, as visualization tools.

Regarding the back-ends, they are based on SQL statements which allow us to create database tables, insert or delete records, extract specific data from a table and modify existing registers. In addition, several numeric parameters are automatically computed according to mathematical equations and considering data which have been previously introduced. For example, the remaining radioactivity of 99Mo generators is calculated based on the radioactivity decay formula, Equation 1, and its corresponding elutions executed. This value is important to be aware if there is enough activity left to perform a new elution.

$$N = N_0 * e^{\left(\frac{-\ln(2)}{\lambda}\right)*t}$$

*Equation 1. Radioactivity decay law*

Where  $N_0$  and  $N$  represent the activity at two different moments,  $t$  is the time variable and  $\lambda$  corresponds to the radioactivity decay constant, also known as half-life, which is unique for each radionuclide.

### 2.3. Software validation

During the development of the software, every process was tested based on the radiopharmacy current documents. Processes were also evaluated by the end-users of the application in monthly meetings. After each session, a list of improvements was collected to adapt the software for the specific needs.

Additionally, once the prototype was completed, three radiopharmacists of this RUHCB used the software in addition to their own registration system for a week. In the end, they filled out a questionnaire which was designed to evaluate the general characteristics, as well as the usefulness and aesthetics. The items included in this questionnaire were: automatization, usefulness, satisfaction, comprehension and easiness to manipulate.

### 2.4. Second version of the software

The first implementation of the software was developed to run locally, and it was a single-user application. These limitations, and other standardization requirements, were accounted for in a second version of the software.

First, the translation of the local application into a web service meant several significant modifications since the Python package for the graphical user interface development isn't compatible with the web version. Therefore, both Python libraries, Tkinter and SQLite3, were replaced by another programming language, HTML, and a better Python library, Django. Notably, the new package (Django) offers a ready-to-use graphical interface for the administrator user to get visual access to the database and can modify it.

As regards the standardization of the protocol, we applied the instructions of chapter IV provided by the order SND/939/2022 published on 4th October 2022. This section is dedicated to automatic systems implemented in radiopharmacy [7].

In addition, a connection between this software and the SAP platform, the information system of the Hospital Clínic, has been successfully established, based on the Health Leven Seven standards. These protocols describe how data is packaged and translated between systems [8].

Finally, this software was uploaded to an Ubuntu server which is integrated into the Hospital Clínic server meaning that all the information is safer from external threats.

## 3. Results

Our proposed software integrates all the procedures involved in the RUHCB's workflow and provides simple and recognizable user interfaces for ease of use.

Different from the current management system, this application can compute an approximated value of the remaining radioactivity of  $^{99m}\text{Tc}$  generators,  $^{99m}\text{Tc}$  elutions and multi-dose radiopharmaceuticals. This calculation is based on the equation of the radioactivity decay and their respective doses dispensed. In addition, it allows end-users to be able to obtain complete traceability of single-doses, kits, elutions, preparations and patients.

In addition, our software provides a stock control of the products received and, the elutions and preparations performed which eases the process of orders and dispensation. In the start screen, the administrator user has access to the software database to make changes when needed.

As it was mentioned before, our application can be split into several sections. In Figure 2, we can observe the procedure performed related to  $^{99m}\text{Tc}$  elutions. The first screenshot corresponds to a form in which some parameters are automatically calculated once the above data is introduced. The image at the bottom represents the elutions stock where end-users can verify that the previous information has been saved and they can also consult about the remaining radioactivity.

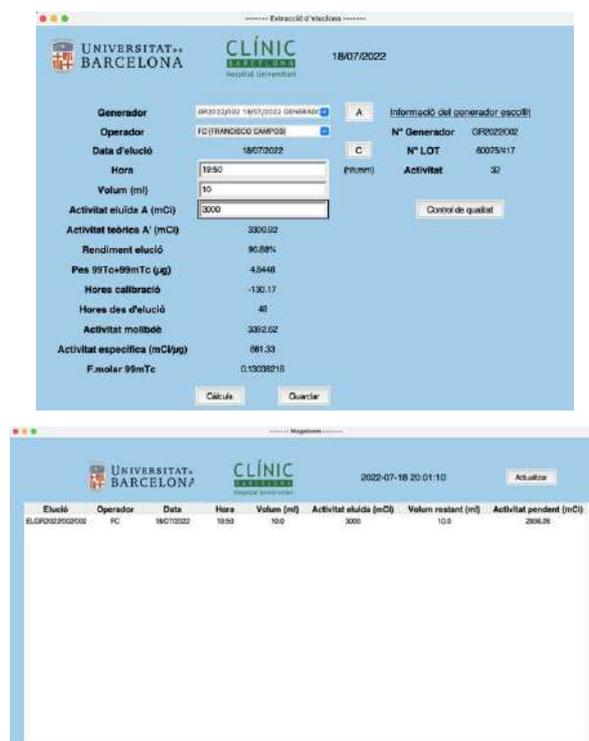


Figure 2. Screenshots of the graphical interface of the first software version. Example of calculation for  $^{99m}\text{Tc}$  decay.

Regarding the end-users opinion, their feedback was very optimistic and useful for future modifications. The results, obtained from the opinion questionnaire were 2.7/4 for satisfaction and efficiency, and 3.6/4 for usefulness and easiness. With respect to the exams for each window, design features were rated with a 3.4/4 and automatization with a score of 3.25/4. In addition, they mentioned that the software is on the right track, holding promise for the final implementation, which will be carried out soon.

## 4. Discussion

Our approach can lead to important improvements in clinical settings. It will allow accurate radiopharmaceutical management, as duplication and unnecessary procedures will be removed, and human errors will be less susceptible to happen. Furthermore, the time invested in those actions can be dedicated to more relevant processes to ensure patient safety and radiopharmaceutical efficiency.

Although this project has been shaped to the RUCHB, it can be implemented in other radiopharmacy of Spain as they all follow the same protocol in terms of radiopharmaceutical management.

On the other hand, further implementations have been suggested while the software was being developed. For example, the installation of a label printing so that, the label will be released once the dose is dispensed. Besides, as the first version of this software runs locally and natively on the operating system of a computer, another improvement was to translate it into a web application. As a result, the resulting software could be used simultaneously by several users without decreasing safety.

The translation process of this local application into a web service meant several significant modifications since the Python package for the graphical user interface development, Tkinter, isn't compatible with the web version. Besides, SQLite3 was replaced by Django because this new Python package offers a ready-to-use graphical interface for the administrator user in which he is able to get visual access to the database and also can modify it.

Although the second software version hasn't been tested in a real scenario, it shows signs of huge potential concerning the previous edition. Besides the benefits mentioned before, it offers more automatization, user restrictions in terms of accessibility to some application functions or pages and better communication with the database, to quote a few. Nowadays, it is still been retouched and adapted to the RUCHB's demands.

Additionally, as paper-based information is still so common in healthcare, this end-of-degree project is aimed to encourage other hospital departments to introduce centralized electronic management promoting effectiveness, timeliness, and communication, among others.

## 5. Conclusion

The principal aim of this project was to develop a user-friendly and efficient software which puts together all processes involved in the RUCHB in a centralized electronic system.

Therefore, the steps within the workflow were studied in detail. This allowed us to choose the programming solution that best fitted our needs. Then, the graphical interface was designed based on the radiopharmacists' opinions and

demands, and also by choosing a design that provides good usefulness and performance.

To sum up, the incorporation of the software will be useful for the radiopharmaceutical management and single-dose traceability.

## Acknowledgements

First and foremost, I would like to express my sincere gratitude to all the clinical professional from Hospital Clinic who have given feedback during the process, as well as professors from University of Barcelona who provided help during the project development.

Furthermore, I would like to acknowledge the Radiopharmacy Unit of the Hospital Clinic for allowing me to use their installations and to execute some tests. Especially, Francisco Campos, Inmaculada Romero and Manuel Santos for their time and patience to explain to me how their department works and finally, evaluating the prototype.

Last but not least, I would like to thank my family and friends for their dedication. Without them and their unconditional support, it wouldn't have been the same.

## References

- [1] Website of the British Nuclear Medicine Society. <https://www.bnms.org.uk/page/WhatdoesaRadiopharmacistdo> (Retrieved: June 2022)
- [2] Website of the Human Health Campus. <https://humanhealth.iaea.org/HHW/Radiopharmacy/VirRad/Documentation/index.html>. (Retrieved: June 2022)
- [3] Or, B., Tests, M., Products, M., Tests, P., Documentation, C., & Dosimetry, R. (1991). *Medicinal Products Clinical Documentation*.
- [4] Dash, S., Shakyawar, S. K., Sharma, M., & Kaushik, S. (2019). Big data in healthcare: management, analysis and future prospects. *Journal of Big Data*, 6(1). <https://doi.org/10.1186/s40537-019-0217-0>
- [5] España. (2007). Real Decreto-ley 1345/2007, de 11 de octubre de 2007, por el que se regula el procedimiento de autorización, registro y condiciones de dispensación de los medicamentos de uso humano fabricados industrialmente. *Boletín Oficial Del Estado*, 45652–45698
- [6] Website of the Full Stack Python. <https://www.fullstackpython.com/databases.html> (Retrieved: June 2022)
- [7] España (2022). Orden SND/939/2022, de 29 de septiembre, por la que se aprueban las normas de correcta preparación extemporánea de radiofármacos. *Boletín Oficial Del Estado*, 135925 – 135945
- [8] Website of the Health Leven Seven International. <https://www.hl7.org/implement/standards/> (Retrieved: October 2022)

# Premio José María Ferrero Corral

---

Jueves, 24 de noviembre de 2022

# Análisis computacional de las características ideales de fármacos en isquemia miocárdica aguda

A. Loidi<sup>1</sup>, J.M. Ferrero<sup>1</sup>

<sup>1</sup> Centro de Investigación e Innovación en Bioingeniería (CI2B), Universitat Politècnica de València, Valencia, España, aloiyar@etsii.upv.es, cferrero@ci2b.upv.es

## Resumen

Se ha demostrado experimentalmente que el aumento de la concentración extracelular de potasio (hiperkalemia) y la existencia de alternantes eléctricos en el potencial de acción son altamente proarrítmicos durante la fase aguda de la isquemia miocárdica (IAM), la cual es una de las principales causas de muerte a nivel mundial. Sin embargo, no existe ningún fármaco específicamente diseñado para hacer frente a las consecuencias electrofisiológicas y arritmogénicas de la patología isquémica de manera efectiva. En este trabajo, se ha utilizado la simulación computacional para definir las características de un fármaco "ideal" que, afectando a varias familias de canales iónicos celulares, proteja al miocardio frente a las arritmias de origen isquémico. Se ha utilizado el modelo computacional validado de O'Hara-Rudy modificado para poder simular las condiciones de la IAM y el efecto del fármaco, que se introdujo utilizando el modelo simple "Pore Block Model". Tras definir una serie de biomarcadores y condiciones a cumplir por la población de modelos de fármacos, los resultados demuestran que de la población inicial (N=10.000) únicamente 10 modelos de fármacos consiguen superar todos los criterios establecidos. De estos últimos, destacan dos por su capacidad de establecer condiciones hiperkalemicas y eléctricas compatibles con una tendencia antiarrítmica en la totalidad de biomarcadores definidos. En conclusión, el trabajo demuestra la viabilidad de los experimentos in-silico para definir un fármaco "ideal" que sea capaz de minimizar la vulnerabilidad a las arritmias en la IAM.

## 1. Introducción

Se sabe que uno de los tres componentes principales responsables de desencadenar todas las demás alteraciones tras la oclusión coronaria durante la IAM es el aumento de la concentración extracelular de potasio ( $[K^+]_e$ ) en la vecindad de los cardiomiocitos [1,2]. Esta acumulación, comúnmente denominada hiperkalemia, es altamente proarrítmica porque conduce a una mayor probabilidad de bloqueos unidireccionales y las consiguientes arritmias reentrantes [3,4]. La evolución temporal de la  $[K^+]_e$  en la IAM se caracteriza por producirse en tres fases [5-8]. La primera corresponde a un aumento muy rápido de la  $[K^+]_e$  inmediatamente después de la oclusión coronaria y dura aproximadamente 5-7 minutos. A ésta le sigue una fase de meseta. Por último, la tercera fase corresponde a un segundo aumento más lento de la  $[K^+]_e$ .

Un examen más detallado de la influencia de la hiperkalemia y de los otros dos componentes principales (acidosis y anoxia) en los cambios dinámicos de la morfología del potencial de acción (PA) muestra alternantes eléctricos en las células cardíacas. Los

alternantes comienzan a aparecer aproximadamente a partir del tercer o cuarto minuto de isquemia, es decir, en la fase aguda, y desaparecen hacia el octavo minuto. Los alternantes corresponden a PA consecutivos que difieren significativamente tanto en amplitud como en duración [9]. Esta fase de alternancia promueve, junto con la hiperkalemia, un escenario completamente heterogéneo durante el proceso del IAM y se establece como un factor de arritmias cardíacas potencialmente mortales.

El objetivo de este trabajo es definir las características de un fármaco "ideal" con efecto multicanal, utilizando el modelo "Pore Block Model" [10], que contrarreste estas dos condiciones proarrítmicas en la IAM mediante el uso de modelos computacionales.

## 2. Material y métodos

En las simulaciones se utilizó un modelo modificado de O'Hara-Rudy para poder tener en cuenta la isquemia aguda. Para simular los efectos de la IAM a nivel de la membrana del cardiomiocito, se implementan las modificaciones de [11]. Explicándolas brevemente, en primer lugar, se añade la corriente de potasio sensible a ATP ( $I_{KATP}$ ) al modelo, ya que no se había considerado originalmente, y se formula adaptando el modelo de Ferrero [12] a los cardiomiocitos humanos utilizando los datos de Babenko [13], que modifica la conductancia máxima y la sensibilidad al  $[ATP]_i$  y al  $[ADP]_i$ . En segundo lugar, los efectos sobre las bombas iónicas de  $[ATP]_i$  y  $[ADP]_i$  se modelan considerando los datos de [14] y [15], que introducen diferentes factores de escala que afectan a la bomba de sodio/potasio ( $I_{NaK}$ ), a la bomba de calcio sarcolemal ( $I_{pCa}$ ) y a la bomba SERCA ( $I_{up}$ ). En tercer lugar, [16] introducen en el modelo los efectos de la acidosis intracelular y extracelular aplicando diferentes factores multiplicativos sobre los canales pH-dependientes como las corrientes  $I_{Na}$ ,  $I_{CaL}$  e  $I_{NaK}$  y sobre el intercambiador sodio/calcio ( $I_{Ca-Na}$ ). Además, los efectos de la lisofosfatidilcolina (LPC) sobre la  $I_{Na}$  y la  $I_{NaL}$  se han modelado utilizando datos de [17]. Por último, sobre este modelo se han realizado las modificaciones necesarias para incorporar el efecto de los fármacos en la corriente de cada familia de canales,  $I_S$ , como

$$I_S = G_S \cdot f_f([F]) \cdot f_v(V_m) \cdot f_l([L]) \cdot (V_m - E) \quad \text{Ec. 1}$$

donde  $G_{S,max}$  es la conductancia máxima de la familia,  $f_f$ ,  $f_v$  y  $f_l$  son los factores correspondientes a las puertas fármaco, voltaje y ligando dependientes respectivamente,

$V_m$  es el potencial de membrana y  $E$  es el potencial de equilibrio correspondiente al ion que atraviesa el canal.

Para definir una población de fármacos candidatos, las corrientes iónicas seleccionadas se multiplicaron por factores ( $f_f$ ) aleatorios entre 0 y 2 uniformemente distribuidos para imitar el efecto de inhibición/excitación de los fármacos sobre los canales iónicos. Se consideró una población realista de fármacos en la que sólo las corrientes  $I_{Kr}$ ,  $I_{NaL}$ ,  $I_{CaL}$ ,  $I_{Na}$  e  $I_{Ks}$  se veían afectadas por el fármaco. De esta manera, se ha generado una población de 10.000 modelos de fármacos, donde para cada modelo se simularon cinco minutos de normoxia seguidos de 30 minutos de isquemia progresiva sobre el modelo de célula aislada, concretamente un cardiomiocito del endocardio ventricular. Las entradas del modelo para su inicialización están basadas en datos experimentales y simulados tomados de diferentes estudios [15,17-20]. Las características del tren periódico de estímulos necesario para la generación de los PA de las simulaciones han consistido en pulsos de 1 milisegundo de duración y una amplitud dos veces el umbral diastólico normóxico con un periodo de estímulo de 1 Hz.

Para realizar el análisis de los resultados, se han tenido en cuenta varios biomarcadores (BM) definidos sobre las variables de estado del modelo correspondientes a la evolución temporal de la  $[K^+]_e$  y del potencial de membrana. Los principales BM utilizados para la clasificación de fármacos ideales se muestran en la Figura 1 sobre las curvas simuladas de la evolución temporal de la  $[K^+]_e$  (1A) y el APD90 (1B) bajo condiciones de control. Los BM de la Figura 1A se corresponden con el valor de la  $[K^+]_e$  en la meseta, el tiempo que tarda en alcanzarla y el tiempo que transcurre la célula expuesta a concentraciones comprendidas entre los 8 y los 13 mM de potasio extracelular, la cual se considera la ventana de concentraciones más proarrítmica. Los BM de la Figura 1B se corresponden con el inicio, el final y la duración del periodo de alternantes.

No todos los modelos generados son válidos ya que pueden no ser fieles representaciones de situaciones reales. Para poder validar cada uno de estos modelos, se filtran de manera que estos deben cumplir en normoxia las condiciones de poseer despolarizaciones positivas, potenciales de reposo de  $\pm 3$  mV respecto a los de control ( $V_{rep\ control} = -87.53$  mV) y valores de APD90 de  $\pm 10\%$  respecto a los de control ( $APD90_{control} = 270$  ms). Estos valores definidos para las restricciones son criterios más estrictos que el rango de valores fisiológicos encontrados en la literatura como los indicados en [21]. Esto implica que los resultados obtenidos cumplen con mayor exigencia las necesidades fisiológicas en normoxia. Posteriormente, se busca dentro de los modelos de fármacos válidos aquellos en los que el BM de la duración del periodo de alternantes sea igual a 0, es decir, se elimine este periodo proarrítmico. Finalmente, debido a que la mayoría de los fármacos comerciales tienen un efecto inhibitorio sobre los canales iónicos, se aplica el último filtro sobre la población y se busca si de los modelos de fármacos obtenidos alguno cumple con factores inhibitorios sobre los cinco canales modificados.

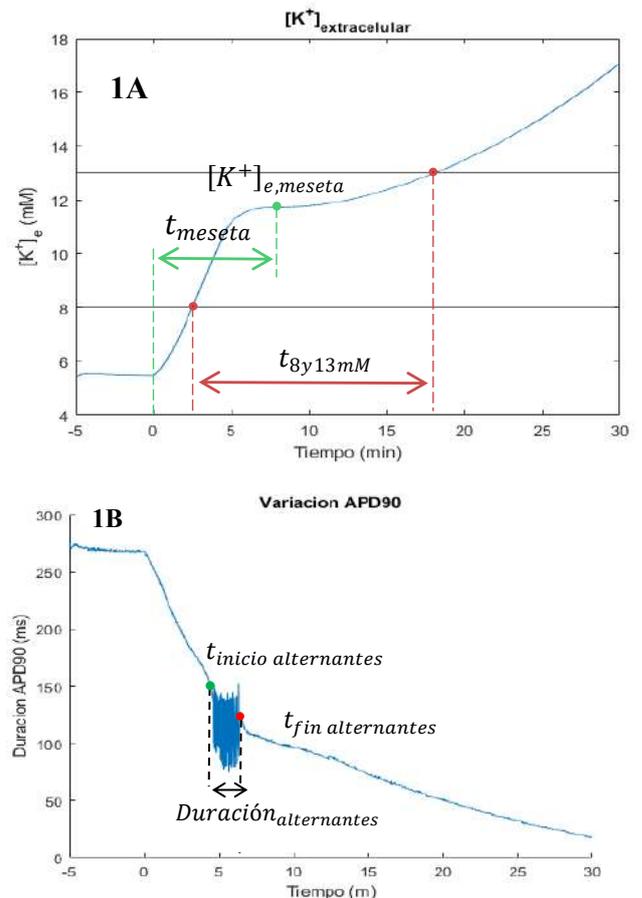


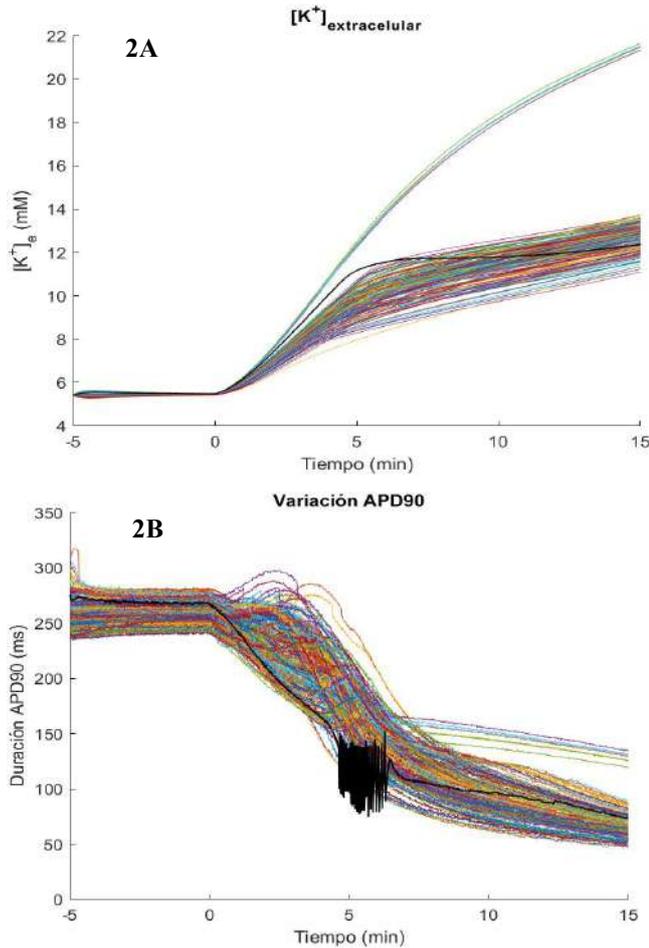
Figura 1. Biomarcadores definidos sobre las curvas de la evolución temporal de la  $[K^+]_e$  (1A) y el APD90 (1B)

### 3. Resultados y discusión

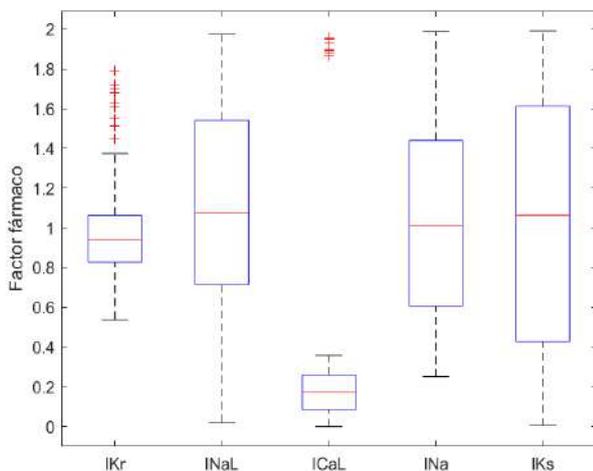
De la población de 10.000 modelos de fármacos iniciales, únicamente 1.661 consiguieron pasar el filtro de condiciones de normoxia. De estos últimos, 184 son los casos en los que se consiguieron eliminar por completo el periodo de alternantes y los cuales se muestran sobre la Figura 2 donde las curvas de control vienen definidas en color negro. Con el fin de observar si existe alguna tendencia o requisito que estas combinaciones deban cumplir, se realiza una representación de diagramas de cajas para cada uno de los factores de manera individual (Figura 3) y se puede ver como para que un fármaco supere el filtro de normoxia y de eliminación de alternantes debe existir un claro efecto inhibitorio sobre la corriente del canal  $I_{CaL}$ , mientras que el resto de las corrientes aceptan un rango mayor de posibilidades. Finalmente, se identificaron aquellos fármacos que contenían únicamente combinaciones de factores inhibitorios para los cinco canales, que es el caso más común en los fármacos existentes. En la Figura 4 se muestran todas las combinaciones de factores de los 184 modelos de fármacos, marcadas de color verde aquellos casos en los que únicamente existe una combinación inhibitoria para todos los canales afectados.

Se identificaron un total de 10 candidatos que contenían una combinación total de factores inhibitorios y, por lo tanto, que superaban todos los filtros establecidos. Las combinaciones de factores de estos 10 candidatos se

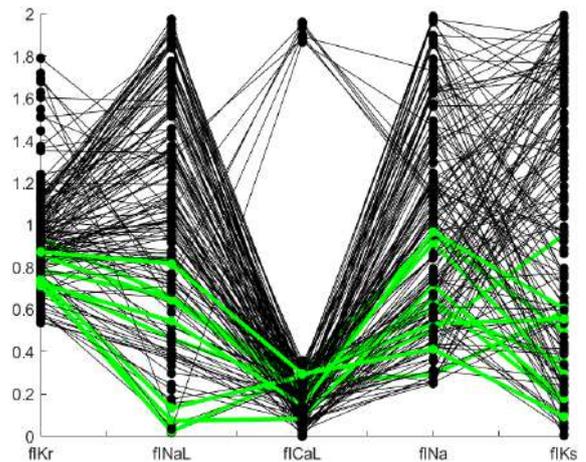
recogen en la Tabla 1 junto con el valor del BM  $t_{8y13mM}$ . El resto de BM no se han recogido debido a que, en los 10 modelos de fármacos obtenidos, el periodo de alternantes se ha eliminado y, además, se ha conseguido eliminar la fase de meseta de la evolución temporal de la  $[K^+]_e$ , lo cual se consideró como una tendencia antiarrítmica.



**Figura 2.** Resultado de las curvas de la evolución temporal de la  $[K^+]_e$  (2A) y el APD90 (2B) de la población de modelos de fármacos ( $N=184$ ) que supera el filtro de condiciones de normoxia y elimina el periodo de alternantes.



**Figura 3.** Diagrama de cajas para cada uno de los factores



**Figura 4.** Combinaciones de factores. Marcadas de color verde aquellas que poseen factores inhibitorios sobre los cinco canales iónicos.

N.º	$f_{IKr}$	$f_{INaL}$	$f_{ICaL}$	$f_{INa}$	$f_{IKs}$	BM $t_{8y13mM}$
1	0,85	0,83	0,26	0,29	0,59	16,22
2	0,72	0,02	0,30	0,95	0,49	<b>15,52</b>
3	0,93	0,66	0,26	0,61	0,35	<b>15,48</b>
4	0,79	0,66	0,24	0,65	0,16	<b>15,35</b>
5	0,74	0,46	0,24	0,50	0,95	16,25
6	0,85	0,64	0,21	0,91	0,29	<b>14,25</b>
7	0,71	0,07	0,08	0,72	0,18	<b>14,45</b>
8	0,74	0,13	0,28	0,53	0,55	17,12
9	0,72	0,54	0,14	0,97	0,60	15,80
10	0,87	0,82	0,29	0,41	0,09	16,09

**Tabla 1.** Combinaciones de factores que cumplen con todos los filtros definidos junto con su correspondiente valor del BM  $t_{8y13mM}$

Se puede ver como las combinaciones 2, 3, 4, 6 y 7 consiguen también reducir el valor del BM  $t_{8y13mM}$ , lo cual se considera una tendencia antiarrítmica. Por lo tanto, estas 5 combinaciones poseen una tendencia antiarrítmica en la totalidad de BM definidos para la IAM. Debe destacarse la actuación de las combinaciones 6 y 7 que reducen por más de un minuto el valor del BM  $t_{8y13mM}$  respecto al control ( $t_{8y13mM,Control} = 15,57$  min).

Las limitaciones que presenta el actual trabajo se corresponden con el hecho de que el modelo de fármaco implementado es un modelo simple basado en el “*Pore Block Model*” y podría formularse basándose en estados del canal. Además, las simulaciones están realizadas sobre un modelo de célula aislada y no de tejido. Aún con estas limitaciones, los resultados suponen una clasificación rápida y preliminar de fármacos seguros y eficaces en la patología isquémica que pretenden ser validadas a través de simulaciones de modelos de tejido tridimensionales

## 4. Conclusiones

Nuestros resultados demuestran la viabilidad de utilizar la simulación de PA para definir un fármaco "ideal" que reduzca los valores de los BM isquémicos definidos y, por tanto, minimice la vulnerabilidad arritmogénica en la isquemia aguda de miocardio.

Los rangos de combinaciones de factores de fármacos obtenidos pueden servir de guía para llevar a cabo el diseño de una molécula que corresponda al fármaco "ideal".

## Agradecimientos

Este proyecto ha recibido financiación del programa de investigación e innovación de la Unión Europea Horizonte 2020 de la Unión Europea bajo el acuerdo de subvención N.º 101016496 (SimCardioTest) y del Programas Estatales de Generación de Conocimiento y Fortalecimiento Científico y Tecnológico del Sistema de I+D+i y de I+D+i Orientada a los Retos de la Sociedad (PID2019-104356RB-C41).

## Referencias

- [1] H. Moréna, M. J. Janse, J. W. Fiolet, W. J. Krieger, H. Crijns y D. Durrer, "Comparison of the effects of regional ischemia, hypoxia, hyperkalemia, and acidosis on intracellular and extracellular potentials and metabolism in the isolated porcine heart.", *Circulation Research*, vol. 46, sup. 5, 1980, pp. 634–646.
- [2] I. Kodama, A. Wilde, M. Janse, D. Durrer y K. Yamada, "Combined effects of hypoxia, hyperkalemia and acidosis on membrane action potential and excitability of guinea-pig ventricular muscle", *Journal of Molecular and Cellular Cardiology*, vol. 16, sup. 3, 1984, pp. 247–259.
- [3] M. J. Janse et al., "Flow of "injury" current and patterns of excitation during early ventricular arrhythmias in acute regional myocardial ischemia in isolated porcine and canine hearts. Evidence for two different arrhythmogenic mechanisms.", *Circulation Research*, vol. 47, sup. 2, 1980 pp. 151–165.
- [4] J. N. Weiss, Z. Qu y K. Shivkumar, "Electrophysiology of Hypokalemia and Hyperkalemia", *Circulation: Arrhythmia and Electrophysiology*, vol. 10, sup. 3, 2017.
- [5] A. Kleber, "Extracellular potassium accumulation in acute myocardial ischemia\*", *Journal of Molecular and Cellular Cardiology*, vol. 16, sup. 5, 1984, pp. 389–394.
- [6] G. Aksnes, "Why do ischemic and hypoxic myocardium lose potassium?", *Journal of Molecular and Cellular Cardiology*, vol. 24, n.º 3, 1992, pp. 323–331.
- [7] J. Weiss y K. I. Shine, "[K<sup>+</sup>]<sub>o</sub> accumulation and electrophysiological alterations during early myocardial ischemia", *American Journal of Physiology-Heart and Circulatory Physiology*, vol. 243, n.º 2, 1982, pp. H318–H327.
- [8] R. Friedrich, H. Hirche, U. Kebbel, V. Zylka y R. Bissig, "Changes of extracellular Na<sup>+</sup>, K<sup>+</sup>, Ca<sup>2+</sup> and H<sup>+</sup> of the ischemic myocardium in pigs", *Basic Research in Cardiology*, vol. 76, n.º 4, 1981, pp. 453–456.
- [9] A. G. Kléber, M. J. Janse, F. J. van Capelle y D. Durrer, "Mechanism and time course of S-T and T-Q segment changes during acute regional myocardial ischemia in the pig heart determined by extracellular and intracellular recordings.", *Circulation Research*, vol. 42, n.º 5, 1978, pp. 603–613.
- [10] G. R. Mirams, Y. Cui, A. Sher, M. Fink, J. Cooper, B. M. Heath, N. C. McMahon, D. J. Gavaghan y D. Noble, "Simulation of Multiple Ion Channel Block Provides Improved Early Prediction of Compounds' Clinical Torsadogenic Risk. *Cardiovasc. Res.*, vol. 91, 2011, pp. 53–61.
- [11] E. F. Carpio, J. F. Gomez, J. F. Rodríguez-Matas, B. Trenor y J. M. Ferrero, "Analysis of vulnerability to reentry in acute myocardial ischemia using a realistic human heart model", *Computers in Biology and Medicine*, vol. 141, 2022, p. 105038.
- [12] J. M. Ferrero, J. Sáiz, J. M. Ferrero y N. V. Thakor, "Simulation of Action Potentials From Metabolically Impaired Cardiac Myocytes", *Circulation Research*, vol. 79, n.º 2, 1996, pp. 208–221.
- [13] A. P. Babenko, G. Gonzalez, L. Aguilar-Bryan y J. Bryan, "Reconstituted Human Cardiac K ATP Channels", *Circulation Research*, vol. 83, n.º 11, 1998, pp. 1132–1143.
- [14] S. Cortassa et al., "A Computational Model Integrating Electrophysiology, Contraction, and Mitochondrial Bioenergetics in the Ventricular Myocyte", *Biophysical Journal*, vol. 91, n.º 4, 2006, pp. 1564–1589.
- [15] J. R. Terkildsen, E. J. Crampin y N. P. Smith, "The balance between inactivation and activation of the Na<sup>+</sup>-K<sup>+</sup> pump underlies the triphasic accumulation of extracellular K<sup>+</sup> during myocardial ischemia", *American Journal of Physiology-Heart and Circulatory Physiology*, vol. 293, n.º 5, 2007, pp. H3036–H3045.
- [16] A. Gonzalez, J.M. Ferrero y J.F. Rodríguez-Matas, "Desarrollo de un software para el estudio, mediante simulación computacional, de las causas de la hiperkalemia durante la isquemia miocárdica aguda", *Universitat Politècnica de València*, 2019.
- [17] M. Gautier, H. Zhang y I. M. Fearon, "Peroxynitrite formation mediates LPC-induced augmentation of cardiac late sodium currents", *Journal of Molecular and Cellular Cardiology*, vol. 44, n.º 2, 2008, pp. 241–251.
- [18] K. Sakamoto, M. Ishikawa, K. Koga, T. Urushidani y T. Nagao, "Energy Preserving Effect of l-cis Diltiazem in Isolated Ischemic and Reperfused Guinea Pig Hearts. A 31P-NMR Study.", *The Japanese Journal of Pharmacology*, vol. 83, n.º 3, 2000, pp. 225–232.
- [19] J. N. Weiss, N. Venkatesh y S. T. Lamp, "ATP-sensitive K<sup>+</sup> channels and cellular K<sup>+</sup> loss in hypoxic and ischaemic mammalian ventricle.", *The Journal of Physiology*, vol. 447, n.º 1, 1992, pp. 649–673.
- [20] Daleau, P, "Lysophosphatidylcholine, a metabolite which accumulates early in myocardium during ischemia, reduces gap junctional coupling in cardiac cells.", *Journal of molecular and cellular cardiology*, vol. 31, n.º 7, 1999, pp. 1391-401.
- [21] E. Passini, C. Trovato, P. Morissette, F. Sannajust, A. Bueno-Orovio y B. Rodriguez, "Drug-induced shortening of the electromechanical window is an effective biomarker for in silico prediction of clinical risk of arrhythmias", *British Journal of Pharmacology*, vol. 176, n.º 19, 2019, pp. 3819–3833.

# Acoplamiento cruzado amplitud-amplitud: un nuevo método para caracterizar la actividad neuronal y las alteraciones provocadas por la enfermedad de Alzheimer

M. Revilla-Vallejo <sup>1,2</sup>, K. Kamarova <sup>1</sup>, C. Gómez <sup>1,2</sup>, R. Hornero <sup>1,2,5</sup>, Yoshihito Shigihara <sup>3</sup>, Hideyuki Hoshi <sup>3,4</sup>, J. Poza <sup>1,2,5</sup>

<sup>1</sup> Grupo de Ingeniería Biomédica, Universidad de Valladolid, Valladolid, España, marcos.revilla@gib.tel.uva.es.

<sup>2</sup> Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina, (CIBER-BBN), Valladolid, España.

<sup>3</sup> Precision Medicine Centre, Hospital Hokuto, Obihiro, Japón

<sup>4</sup> Medical Imaging Business Centre, Healthcare Business Group, RICOH Company, Ltd., Japón

<sup>5</sup> IMUVA, Instituto de Investigación en Matemáticas, Universidad de Valladolid, Valladolid, España.

## Resumen

*La actividad electromagnética cerebral se ha analizado tradicionalmente en cada una de las seis bandas de frecuencia convencionales. De este modo, en los análisis de conectividad se obtienen seis redes neuronales funcionales, con resultados en ocasiones contradictorios. La construcción de un modelo de red multicapa permite una visión conjunta del cerebro y posibilita análisis agregados de las alteraciones que trastornos neurodegenerativos, como la enfermedad de Alzheimer (EA), provocan en la actividad neuronal. El objetivo del presente trabajo es plantear una nueva forma de construcción de la red neuronal funcional multicapa y evaluar cómo esta se ve alterada con la EA. Para ello, se emplearon dos bases de datos, la primera, de actividad electroencefalográfica con 251 sujetos y, la segunda, de actividad magnetoencefalográfica con 188 sujetos. Ambas bases de datos constaban de sujetos cognitivamente sanos, pacientes con deterioro cognitivo leve por EA y pacientes con demencia por EA. Se utilizó el acoplamiento cruzado amplitud-amplitud (AAC) para construir la red multicapa, basado en la medida de conectividad funcional de correlación entre la amplitud de las envolventes (AEC). La red AAC obtenida muestra diferencias entre el grupo no patológico y los grupos patológicos dando cuenta de una alteración de la red cerebral global, especialmente en las interacciones inter-banda (acoplamiento cruzado). Estas alteraciones en AAC sugieren la pérdida de facultades cognitivas característica de la EA, como el pensamiento activo o la atención.*

## 1. Introducción

La electroencefalografía (EEG) y la magnetoencefalografía (MEG) son dos técnicas de registro de la actividad electromagnética cerebral [1]. La actividad neuronal ha sido analizada utilizando diferentes aproximaciones para caracterizar los procesos cerebrales y cómo esos se modifican con diferentes trastornos neurodegenerativos [2]. La enfermedad de Alzheimer (EA) es la patología neurodegenerativa con mayor prevalencia y la causa principal de demencia. En la actualidad se estima que aproximadamente 55 millones de personas la padecen [3]. Estudios previos han demostrado que la EA modifica la actividad electromagnética cerebral, provocando, entre otras alteraciones, el enlentecimiento de las señales EEG y MEG y la desconexión progresiva entre las diferentes regiones neuronales [2, 3].

Tradicionalmente, la actividad neuronal registrada se ha analizado de forma desagregada por bandas de frecuencia, concretamente, en las seis bandas de frecuencia convencionales ( $\delta$ , delta: 1-4 Hz;  $\theta$ , zeta: 4-8 Hz;  $\alpha$ , alfa: 8-13 Hz;  $\beta$ -1, beta-1: 13-18 Hz;  $\beta$ -2, beta-2: 18-30 Hz y  $\gamma$ , gamma, > 30 Hz) [1]. A partir de la actividad en cada una de las bandas de frecuencia, una de las aproximaciones más frecuentemente utilizadas es el cálculo de la conectividad funcional entre las diferentes regiones cerebrales para evaluar las relaciones temporales que existen entre la actividad neuronal procedente de estas regiones [4]. El resultado es una red neuronal funcional formada por nodos, correspondientes a cada una de las regiones cerebrales, y por enlaces entre ellas, correspondientes a las magnitudes de las relaciones temporales. El análisis desagregado por bandas de frecuencia da lugar a seis redes neuronales funcionales para describir el comportamiento de la red cerebral global. Estudios recientes han sugerido la necesidad de armonizar la información de estas seis redes funcionales en una única; de modo que sea posible analizar la actividad neuronal del cerebro desde un punto de vista agregado (red multicapa) y facilitar la interpretación de las alteraciones que las enfermedades neurodegenerativas provocan en la misma [5, 6].

La hipótesis de este trabajo es que la EA provoca alteraciones en la actividad neuronal que pueden ser analizadas mediante una red multicapa. En consecuencia, el objetivo principal es plantear una nueva forma de construcción de la red neuronal funcional multicapa y evaluar cómo esta es alterada a lo largo del continuo de la EA.

## 2. Materiales

### 2.1. Sujetos

En el presente trabajo se han empleado dos bases de datos, una de registros EEG y otra de registros MEG. La primera (BD-EEG), registrada en un proyecto multicéntrico entre España y Portugal consta de 251 sujetos: 51 sujetos de control, 51 pacientes con deterioro cognitivo leve (DCL) por EA y 149 pacientes con demencia debida a EA. La segunda (BD-MEG), registrada en el Hospital Hokuto en Japón, consta de 188 sujetos: 38 sujetos de control, 50 pacientes con DCL por EA y 100 pacientes con demencia debida a EA. Los diferentes gru-

pos se determinaron siguiendo los criterios diagnósticos NIA-AA [7]. Todos los sujetos, o bien sus familiares, dieron su consentimiento informado para participar en el estudio, diseñado según el Código Ético de la Asociación Médica Mundial (Declaración de Helsinki). Los comités correspondientes aprobaron el protocolo de trabajo.

La Tabla 1 resume los datos sociodemográficos de las dos bases de datos. En ambos casos no se encontraron diferencias estadísticamente significativas en la distribución de sexos ( $p$ -valor  $> 0.05$ , test Chi-cuadrado). Por otra parte, las dos bases de datos muestran diferencias estadísticamente significativas entre grupos en la edad y en el *Mini-Mental State Examination* (MMSE) ( $p$ -valor  $< 0.05$ , test de Kruskal-Wallis). La edad del grupo de pacientes con DCL es significativamente mayor en la BD-EEG, mientras que en la BD-MEG, la edad de los pacientes con EA es mayor que la de los otros dos grupos. Por su parte, los valores de MMSE muestran diferencias entre los grupos en las dos bases de datos, reflejando el deterioro cognitivo consecuencia de la progresión de la EA.

## 2.2. Registros EEG y MEG

La actividad EEG se registró con un sistema de 19 canales, con los electrodos posicionados según el sistema internacional 10-20, y empleando un electroencefalógrafo Nihon Kohden Neurofax JE-921A con frecuencia de muestreo de 500 Hz. Por otro lado, la actividad electromagnética de la BD-MEG se registró empleando un sistema MEG de 160 canales (MEG Vision PQ1160C, Yokogawa Electric) y una frecuencia de muestreo de 1000 Hz. En ambos casos, se solicitó a los sujetos que permanecieran sentados, en reposo y con los ojos cerrados, obteniendo un total de 5 minutos de registro. Las señales obtenidas en las dos bases de datos fueron preprocesadas de la siguiente forma [8]: (i) filtrado digital con un filtro FIR (*Finite Impulse Response*) en la banda 1-70 Hz; (ii) filtrado de ranura para eliminar la interferencia de 50 Hz; (iii) análisis de componentes independientes (ICA) para la detección y eliminación de artefactos; y (iv) segmentación de épocas de 5 s y selección de aquellas libres de artefactos mediante inspección visual. Adicionalmente, en la BD-MEG se empleó el algoritmo SOUND como una etapa adicional de eliminación de artefactos [9].

## 3. Métodos

### 3.1. Reconstrucción de las fuentes cerebrales

La actividad cerebral procedente de los 19 electrodos EEG, así como de los 160 canales MEG, fue proyectada a nivel de fuente, obteniendo 68 señales correspondientes a las regiones corticales en las que se divide el volumen cerebral según el atlas *Desikan-Killiany*. El algoritmo de localización de fuentes empleado fue sLORETA (*Standardized Low - Resolution Brain Electromagnetic Tomography*). Estas 68 señales a nivel de fuente se emplearon como punto de partida para el análisis posterior.

### 3.2. Acoplamiento cruzado de frecuencia: amplitud-amplitud

La construcción de la red neuronal funcional se ha rea-

	Controles	DCL	EA
<b>BD-EEG</b>			
<i>N</i>	51	51	149
Edad (años)	80.14 [7.09]	85.53 [7.25]	80.57 [7.61]
Sexo (H:M)	26:25	15:36	35:114
MMSE	28.82 [1.13]	23.33 [2.84]	12.92 [8.72]
<b>BD-MEG</b>			
<i>N</i>	38	50	100
Edad (años)	73.53 [5.98]	76.76 [6.16]	81.05 [6.81]
Sexo (H:M)	19:18	11:39	43:57
MMSE	28.77 [1.76]	26.16 [2.59]	18.16 [5.67]

**Tabla 1.** Datos sociodemográficos y clínicos de las bases de datos empleadas; *N*: número de sujetos; *H*: hombre; *M*: mujer; *MMSE*: *Mini-Mental State Examination* (rango: [0, 30]). *Edad*, *MMSE*: media [desviación estándar].

lizado tradicionalmente empleando la actividad electromagnética cerebral filtrada en el rango de frecuencias de las bandas convencionales [4]. De esta forma, se obtiene una red neuronal funcional (con 68 nodos y las conexiones correspondientes) para cada una de esas bandas de frecuencia. En este dominio, una de las medidas de conectividad más robusta a las interacciones espúreas es la correlación entre la amplitud de las envolventes (*amplitude envelope correlation*, AEC) ortogonalizada [10, 11]. La AEC estima la correlación entre las amplitudes de las envolventes de las señales procedentes de las diferentes regiones cerebrales en la banda de frecuencia en cuestión [11]. El paso adicional de ortogonalización reduce el acoplamiento espúreo a través de una regresión lineal. De esta forma, se eliminan efectos no deseados derivados de la conducción de volumen, entre otros [12].

En este estudio, se planteó el acoplamiento cruzado de frecuencia amplitud-amplitud (*amplitude-to-amplitude coupling*, AAC) como una extensión de la AEC ortogonalizada para contemplar las interacciones multicapa [5, 6, 13]. La AAC se basó en la AEC ortogonalizada filtrando de forma apropiada las señales consideradas. El procedimiento consistió en filtrar la envolvente de las señales de las bandas de frecuencia superiores en el rango de frecuencias de las bandas de frecuencia más bajas y, posteriormente, calcular la AEC ortogonalizada entre estas señales [13]. Por ejemplo, para la interacción de las bandas  $\delta$  y  $\theta$ , la envolvente de la señal de la banda  $\theta$  se filtró en el rango 1-4 Hz y, a continuación, se calculó la AEC entre la señal  $\delta$  original y la señal  $\theta$  resultante del filtrado. De este modo, se obtuvo una red neuronal funcional de 408 nodos (68 nodos x 6 bandas de frecuencia) formada por 21 subredes correspondientes a las interacciones intra-banda (AEC) e inter-banda (AAC) para cada uno de las épocas y sujetos de las dos bases de datos. Las subredes de la diagonal principal se correspondían con las interacciones intra-banda convencionales, mientras que el resto, con las interacciones inter-banda.

Los cambios entre las redes AAC obtenidas para cada uno de los grupos se caracterizaron a través de la matriz diferencia entre las mismas. En estas matrices solo se

representaron aquellos valores con diferencias estadísticamente significativas en cada una de las comparaciones entre los grupos objeto de estudio ( $p < .05$ , test U de Mann-Whitney).

#### 4. Resultados y discusión

En primer lugar se evaluaron los factores potenciales de confusión en los resultados obtenidos. Para ello se analizó la existencia de correlaciones estadísticamente significativas entre la red AAC promediada por épocas para cada sujeto y la edad. Los resultados mostraron que la edad no se encontraba estadísticamente relacionada con las redes AAC promediadas por épocas ( $p > .05$ , test bivariado de correlación de Spearman, corrección FDR).

En la Figura 1 se muestran las matrices AAC promedio por épocas y sujetos para cada uno de los tres grupos considerados (controles, DCL y EA), así como las matrices diferencia entre estos. Las matrices AAC se muestran en escala logarítmica. En las matrices diferencia, en color rojo se muestran los incrementos de conectividad significativos del primer grupo con respecto al segundo, mientras que en color azul las disminuciones significativas ( $p < .05$ , test U de Mann-Whitney).

Los resultados muestran un elevado grado de replicación entre las dos bases de datos y técnicas de registro. Las mayores diferencias se encuentran entre el grupo no patológico y los grupos patológicos, especialmente en la transición entre controles y pacientes con EA. En las subredes de la diagonal principal (subredes AEC) es posible apreciar la dicotomía entre bandas altas y bajas de frecuencia observada previamente en la literatura. La conectividad aumenta principalmente en las bandas bajas a medida que la EA progresa, mientras que en las bandas altas disminuye [2]. Analizando a continuación las interacciones cruzadas entre bandas de frecuencia, en primer lugar se observa como el acoplamiento entre la banda  $\alpha$  y las bandas  $\beta$ -1,  $\beta$ -2 y  $\gamma$  es mayor en el grupo de controles que en el de EA para ambas bases de datos [14]. Esta mayor conectividad puede estar relacionada con una menor actividad en el estado de reposo, del que son característicos los ritmos  $\alpha$  en comparación con los estados de pensamiento o atención asociados con los ritmos  $\beta$  y  $\gamma$ . A lo largo de la EA se aprecia un claro deterioro en el pensamiento activo y la atención, siendo predominantes los estados de reposo [3]. Del mismo modo, se observa una menor conectividad entre la banda  $\delta$  y las bandas  $\beta$ -1 y  $\beta$ -2, mostrando esa mayor actividad en el estado de somnolencia frente a los estados de pensamiento activo o atención. Por otra parte, las diferencias negativas en la comparación Control-EA entre las bandas  $\beta$  y  $\gamma$  indican una mayor actividad en las bandas altas de frecuencia a medida que la EA progresa. La banda  $\gamma$  se ha relacionado tradicionalmente con elevada actividad mental en entornos de pánico o estrés [1]. La EA y sus alteraciones características pueden justificar esta elevada actividad al alterar los mecanismos básicos de funcionamiento del cerebro y, por tanto, someter a un elevado grado de estrés al mismo [3]. En la misma línea de razonamiento se encuentran las diferencias entre C

y DCL, pudiendo resultar estas fundamentales para un diagnóstico precoz de la enfermedad.

Este trabajo consta de tres limitaciones principales. En primer lugar el elevado coste computacional de la aproximación de red multicapa. Como línea futura se plantea optimizar el proceso de filtrado y cálculo de la AEC para reducir el tiempo de computación de red AAC completa. En segundo lugar, se ha empleado una plantilla para la localización de fuentes con una resolución espacial reducida. El empleo de un mayor número de fuentes permitiría una mayor resolución espacial y, por tanto, la realización de un mapeo de las alteraciones observadas con las diferentes regiones cerebrales. Si bien, esto incrementaría el coste computacional de la aproximación. En último lugar, la necesidad de resumir la información de la red multicapa mediante la aplicación de parámetros derivados de la Teoría de Grafos.

#### 5. Conclusiones

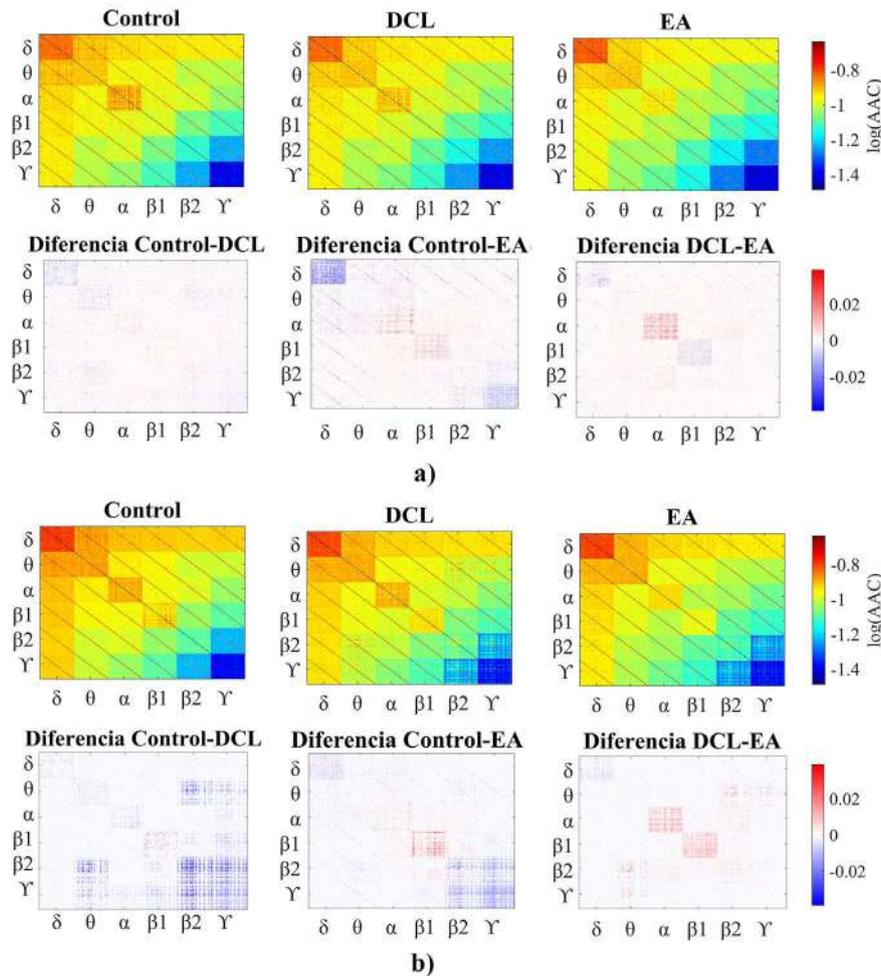
La metodología propuesta de caracterización de la actividad neuronal mediante la AAC es replicable, con un alto grado de alineación entre las dos bases de datos empleadas, a pesar del uso de técnicas de registro diferentes (EEG y MEG). Además, los resultados observados muestran que esta nueva red multicapa es capaz de describir la actividad neuronal en su conjunto, proporcionando una visión agregada del funcionamiento cerebral. Las redes AAC obtenidas sugieren la importancia de la comunicación inter-banda en el funcionamiento cerebral. Por último, las alteraciones observadas en esta red multicapa podrían estar relacionadas con el deterioro de ciertas funciones cognitivas consecuencia de la progresión de la EA.

#### Agradecimientos

Este estudio ha sido financiado por el ‘Ministerio de Ciencia e Innovación - Agencia Estatal de Investigación’ y el Fondo Europeo de Desarrollo Regional (FEDER) bajo el proyecto PGC2018-098214-A-I00, por la Comisión Europea y el FEDER bajo el proyecto ‘Análisis y correlación entre la epigenética y la actividad cerebral para evaluar el riesgo de migraña crónica y episódica en mujeres’ (‘Programa de Cooperación Interreg V-A España-Portugal POCTEP 2014-2020’), y por el CIBER en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN) a través del Instituto de Salud Carlos III co-financiado con fondos FEDER. M. Revilla-Vallejo disfruta de un contrato predoctoral financiado por el Fondo Social Europeo, y la Junta de Castilla y León, a través de la Consejería de Educación.

#### Referencias

- [1] N. Kulkarni *et al.* *Electroencephalogram and Its Use in Clinical Neuroscience. EEG-Based Diagnosis of Alzheimer Disease*. Academic Press, 2018.
- [2] L. Koelewijn *et al.* Alzheimer’s disease disrupts alpha and beta-band resting-state oscillatory network connectivity. *Clin. Neurophysiol.*, 128(11):2347–2357, 2017.



**Figura 1.** a) Base de datos de EEG; b) Base de datos de MEG. Redes multicapa AAC promedio en la parte superior y matrices de diferencia en la parte inferior para controles, DCL y EA. Las redes multicapa AAC se representan en escala logarítmica. En las matrices diferencia, en color rojo se muestran los incrementos de conectividad del primer grupo con respecto al segundo y en color azul los decrementos. Solo se representan aquellos valores de la matriz diferencia en los que existen diferencias estadísticamente significativas ( $p < .05$ , test U de Mann-Whitney).

- [3] S. Gauthier *et al.* World Alzheimer Report 2021. Journey through the diagnosis of dementia. *Alzheimer's Disease International*, 2021.
- [4] M. Rubinov and O. Sporns. Complex network measures of brain connectivity: Uses and interpretations. *NeuroImage*, 52(3):1059–1069, 2010.
- [5] S. J. Ruiz-Gómez *et al.* A new method to build multiplex networks using canonical correlation analysis for the characterization of the Alzheimer's disease continuum. *Journal of Neural Engineering*, 18(2):1–12, 2021.
- [6] P. Tewarie *et al.* Interlayer connectivity reconstruction for multilayer brain networks using phase oscillator models. *New Journal of Physics*, 23(6):1–16, 2021.
- [7] J.R.Jr. Clifford *et al.* NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. *Alzheimers Dement.*, 14(4):535–562, 2018.
- [8] A. Maturana-Candelas *et al.* EEG Characterization of the Alzheimer's disease continuum by means of multiscale entropies. *Entropy*, 21(6):544, 2019.
- [9] V. Rodríguez-González *et al.* Consistency of local activation parameters at sensor- And source-level in neural signals. *Journal of Neural Engineering*, 17(5):056020, 2020.
- [10] Z. Liu *et al.* Large-scale spontaneous fluctuations and correlations in brain electrical activity observed with magnetoencephalography. *Neuroimage*, 51(1):102–111, 2010.
- [11] C.T. Briels *et al.* Reproducibility of EEG functional connectivity in Alzheimer's disease. *Alzheimer's Research and Therapy*, 12(1):1–14, 2020.
- [12] G. C. O'Neill *et al.* Dynamics of large-scale electrophysiological networks: A technical review. *NeuroImage*, 180(Pt B):559–576, 2018.
- [13] M. Siems and M. Siegel. Dissociated neuronal phase- and amplitude-coupling patterns in the human brain. *NeuroImage*, 209:1–12, 2020.
- [14] C. Lihui *et al.* Reconstruction of functional brain network in Alzheimer's disease via cross-frequency phase synchronization. *Neurocomputing*, 314:490–500, 11 2018.

# Efecto de la Congelación de Capas en Redes Neuronales Convolucionales Entrenadas para la Evaluación de la Calidad del ECG

A. Huerta Herraiz<sup>1</sup>, A. Martínez Rodrigo<sup>1</sup>, J.J. Rieta Ibáñez<sup>2</sup>, R. Alcaraz Martínez<sup>1</sup>

<sup>1</sup> Grupo de Investigación en Ingeniería Electrónica, Biomédica y de Telecomunicación, Universidad de Castilla-La Mancha, Cuenca, España, {alvaro.huerta, arturo.martinez, raul.alcaraz}@uclm.es

<sup>2</sup> BioMIT.org, Dep. de Ingeniería Electrónica, Universitat Politècnica de Valencia, Valencia, España, jjrieta@upv.es

## Resumen

*La utilización de dispositivos portátiles de registro del electrocardiograma (ECG) para el cribado y el seguimiento a largo plazo de enfermedades cardiovasculares es cada vez más común. Sin embargo, debido a las elevadas alteraciones que aparecen en la señal, tales como los artefactos, la evaluación de la calidad es crucial a la hora de realizar un diagnóstico correcto. Actualmente, para este propósito la tendencia es utilizar redes neuronales convolucionales preentrenadas debido a su buen rendimiento. Es posible optimizar la fase de ajuste final mediante la congelación de capas, preservando su conocimiento inicial, lo que permite aprender características más genéricas. Por lo tanto, este trabajo tiene como objetivo analizar cómo afecta la congelación de capas de una red preentrenada cuando se utiliza para evaluar la calidad del ECG. La red utilizada es conocida como AlexNet, a la cual se le introdujeron escalogramas como entrada y se congelaron progresivamente las capas de 1 a 7. El proceso de validación se llevó a cabo con dos bases de datos distintas, compuestas por casi 70.000 fragmentos de ECG de 5 segundos de duración. Después de 10 iteraciones de validación, no se apreciaron diferencias sustanciales en la clasificación, sin embargo, se observó un ligero aumento de la exactitud a medida que se incrementaba el número de capas congeladas. Así pues, a la vista de los resultados se recomienda congelar todas las capas excepto la última durante el ajuste final de AlexNet a la hora de evaluar la calidad del ECG, optimizando recursos sin bajar el rendimiento.*

## 1. Introducción

Hoy en día, la manera más habitual de monitorizar y diagnosticar muchas enfermedades cardíacas es el electrocardiograma (ECG) en reposo. Sin embargo, puede ser insuficiente en aquellos casos en los que los trastornos cardíacos sean de naturaleza intermitente y aleatoria, haciéndose necesario una monitorización continua [1]. En el mercado proliferan nuevos dispositivos portátiles y vestibles que permiten una monitorización a muy largo plazo, abarcando incluso varios meses [1]. Así pues, se espera contribuir a corto plazo en la mejora del diagnóstico de enfermedades con alteraciones del ritmo cardíaco, tales como la fibrila-

ción auricular paroxística (FA). En esta arritmia los episodios iniciales son en su mayoría asintomáticos y sólo duran unos pocos segundos o minutos [2].

Sin embargo, a pesar de las enormes posibilidades de monitorización de estos nuevos dispositivos, la calidad general del registro de ECG se ve muy mermada debido a que se obtiene durante las actividades cotidianas de los pacientes, tales como andar, trabajar, etc. [3]. Así pues, la enorme longitud del registro adquirido hace inviable la visualización de la grabación completa del ECG por parte de un profesional y, por lo tanto, se hace necesaria la automatización del proceso de identificación de los tramos afectados por perturbaciones asociadas a los sistemas vestibles o portátiles. Estas perturbaciones son ampliamente conocidas y se corresponden principalmente con artefactos de movimiento, ruido impulsivo, y mal contacto de electrodos. Además, también es común que la señal de ECG se vea afectada por otros ruidos ampliamente estudiados, como interferencias de la red eléctrica, ruido muscular, o alteraciones en la línea base. Todos estos factores afectan a la calidad general del registro de ECG de larga duración, por lo que tienen que ser evaluados como un paso previo a su análisis e interpretación para el diagnóstico del paciente.

En los últimos años, se ha observado en el estado del arte un gran número de trabajos que han tratado de evaluar automáticamente la calidad del ECG de larga duración. De hecho, se pueden encontrar desde sistemas relativamente sencillos basados en reglas de decisión hasta algoritmos bastante complejos que usan herramientas propias de la inteligencia artificial. En cuanto a los primeros métodos publicados que obtuvieron resultados aceptables, están basados en la extracción de puntos fiduciales y características morfológicas del registro de ECG, las cuales se combinan con técnicas comunes de aprendizaje automático, tales como máquinas de vectores de soporte o árboles de decisión [4]. Sin embargo, los nuevos algoritmos de aprendizaje profundo han superado la capacidad de generalización de estos métodos a la hora de identificar intervalos de ECG de alta o baja calidad [5]. Otra característica muy importante es que estos métodos son capaces de admitir directamente la señal de ECG en bruto, sin requerir etapas

adicionales de preprocesado, ni la selección y extracción, manual o automática, de variables [5].

El principal problema que presentan los algoritmos de aprendizaje profundo es la gran cantidad de datos necesarios para realizar un entrenamiento adecuado. Sin embargo, es posible utilizar una técnica que se conoce como *transferencia de aprendizaje*, mediante la cual se transfiere el conocimiento adquirido por redes neuronales convolucionales (RNC) previamente entrenadas con un objetivo para ser utilizadas en una nueva y diferente tarea [6]. Para ello, a la hora de realizar el ajuste final de la red, existe la posibilidad de entrenar sólo algunas capas, congelando las restantes y conservando su conocimiento inicial. Así pues, el objetivo principal del presente trabajo es analizar cómo afecta la congelación de un número variable de capas al rendimiento de una conocida RNC a la hora de evaluar la calidad del ECG. Actualmente, y cada vez más, existen arquitecturas de RNC preentrenadas que están disponibles para ser utilizadas mediante la técnica de transferencia de aprendizaje. Muchas de ellas han sido originalmente entrenadas para discernir entre más de 1.000 tipos diferentes de imágenes [7]. Para el presente trabajo se tomó la arquitectura AlexNet como base, debido a su gran popularidad en aplicaciones que emplean el registro de ECG [5].

## 2. Métodos

### 2.1. Algoritmos de evaluación de la calidad del ECG

Este trabajo está basado en una publicación previa que utiliza aprendizaje profundo para evaluar la calidad del ECG y clasificarlo en las categorías de alta y baja calidad [5]. Dado que la arquitectura AlexNet se ha diseñado para recibir imágenes bidimensionales como entrada [8], las señales de ECG se fragmentaron en segmentos de 5 segundos de longitud y, posteriormente, cada segmento se convirtió a una imagen a través de la transformada Wavelet continua (TWC) [9]. La imagen generada se conoce como escalograma y ha sido ampliamente utilizada en muchos algoritmos de aprendizaje profundo introducidos en trabajos previos relacionados con el ECG [9]. Esto es debido a su buen rendimiento, ya que resalta de manera notable las características de las señales no estacionarias como es el ECG. Para realizar la transformación se utilizó una función *Morlet* como Wavelet madre y se aplicó un número de escalas de 48 voces por octava. Todos los detalles se pueden encontrar en el trabajo [5].

### 2.2. Arquitectura de la RNC AlexNet

Como se ha comentado previamente, existen muchas redes pre-entrenadas que están disponibles para su uso en aplicaciones muy diversas. En este trabajo, se usó la red AlexNet, la cual fue inicialmente entrenada para distinguir entre más de 1.000 clases de imágenes diferentes [8]. Esta red se hizo muy popular gracias a un concurso llamado *ImageNet Challenge* [10], y su arquitectura está compuesta por ocho capas con capacidad de aprendizaje, donde cinco son capas convolucionales y las otras tres son capas totalmente conectadas [8]. La primera capa convolucional tiene una dimensión espacial de  $11 \times 11 \times 3$  con 96 núcleos.

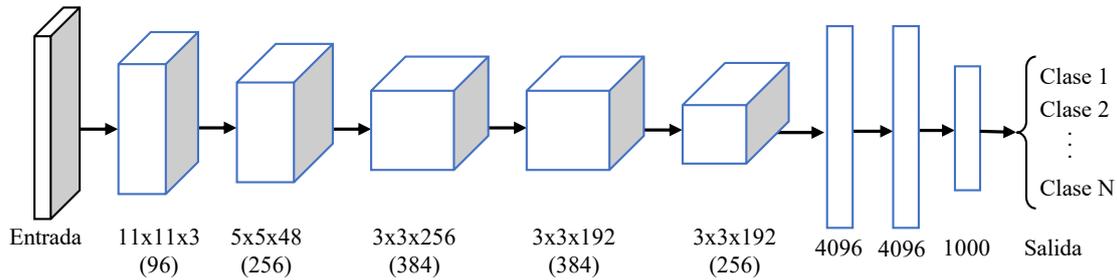
La segunda capa convolucional reduce ligeramente la entrada de la primera y aplica filtros de tamaño  $5 \times 5 \times 3$  con 256 núcleos. Las tres últimas capas convolucionales se conectan directamente sin aplicar funciones de normalización ni de agrupación. La tercera capa convolucional está compuesta por 384 núcleos con un tamaño de  $3 \times 3 \times 256$ . La cuarta capa convolucional contiene 384 núcleos de dimensiones  $3 \times 3 \times 192$ , mientras que la última capa convolucional aplica 256 núcleos de tamaño  $3 \times 3 \times 192$ . Cada una de las capas totalmente conectadas contiene 4.096 neuronas, y la salida de la última se conecta a un clasificador basado en la función *softmax* capaz de asignar probabilidades a 1.000 clases diferentes. El número de parámetros con capacidad de aprendizaje en estas capas, los cuales se congelaron durante el proceso de ajuste final, se detallan en la Tabla 1. La arquitectura de AlexNet se completa con varias capas y funciones de activación. De hecho, contiene funciones de activación junto a todas las capas con capacidad de aprendizaje para reducir las proporciones espaciales del mapa de características, y también presenta tres capas de agrupación después de las capas convolucionales primera, segunda y quinta. Antes y después de la primera capa totalmente conectada se incluyen dos funciones de regularización y *dropout* con el fin de evitar el sobre-entrenamiento de la red. La descripción completa de la arquitectura de AlexNet puede encontrarse en [8]. La Figura 1 muestra un esquema simplificado de dicha arquitectura, en el que las capas con capacidad de aprendizaje están en azul.

### 2.3. Desarrollo del experimento

Con el fin de obtener unos resultados que no estuviesen sesgados, se utilizaron dos bases de datos: una base de datos propia (PDB) y una base de datos pública perteneciente al PhysioNet/CinC Challenge 2017 (PC2017DB) [11]. Se escogieron estas dos debido a que ambas contienen registros de ECG con morfologías muy diferentes, que fueron adquiridos de maneras muy distintas. Para realizar los procesos de entrenamiento y ajuste final de la red AlexNet se utilizó la PDB. Esta base de datos fue adquirida utilizando un dispositivo portátil (Nuubo e-minder), el cual registra el ECG con una frecuencia de muestreo de 250 Hz y 12 bits de resolución. El proceso de etiquetado de los diferentes fragmentos de alta y baja calidad fue realizado por dos expertos, de acuerdo con el siguiente criterio: se entendieron como señales ruidosas aquellas en las que no se podían distinguir claramente los picos R, mientras que en otro caso se consideraban de alta calidad. En total, este conjunto de datos finalmente se compuso de 20.000 intervalos de ECG de 5 segundos de duración, distribuidos de manera balanceada entre los dos grupos. Hay que destacar que la clase de alta calidad estaba compuesta tanto por episodios de ritmo sinusal normal (RSN) como de FA. Para la validación se utilizó la PC2017DB, que es muy común dentro del estado del arte y está disponible de forma libre. Esta base de datos se registró mediante un dispositivo portátil (AliveCor) que mide la actividad eléctrica en las yemas de los dedos, con una frecuencia de muestreo de 300 Hz y una resolución de 16 bits. La principal ventaja de PC2017DB es que se encuentra anotada, presentando cuatro categorías diferen-

Capa	Conv. 1	Conv. 2	Conv. 3	Conv. 4	Conv. 5	T.C. 1	T.C. 2	T.C. 3	Total
Parámetros	34.944	307.456	885.120	663.936	442.624	37.752,832	16.781.312	4.097.000	60.965.224

**Tabla 1.** Número de parámetros con capacidad de aprendizaje en las capas convolucionales (Conv.) y totalmente conectadas (T.C.) pertenecientes al modelo AlexNet.



**Figura 1.** Representación simplificada de la arquitectura de AlexNet. Las capas con parámetros con capacidad de aprendizaje están representadas en azul.

tes, como FA, ritmo sinusal, otros ritmos y ruido. Después de segmentar todos los registros, se obtuvieron 48.607 segmentos de 5 segundos distribuidos de manera muy desigual entre los dos grupos, perteneciendo 47.439 a la clase de alta calidad y 1.168 a la de baja calidad.

Tradicionalmente, en el proceso de ajuste final de una RNC pre-entrenada se sustituye la última capa con capacidad de aprendizaje y se adapta a la nueva tarea que se pretende abordar. En el presente trabajo, la última capa de AlexNet, la cual se diseñó para ser capaz de clasificar hasta 1.000 clases diferentes, se modificó para que sólo distinguiese entre dos clases: segmentos de ECG de alta y de baja calidad. Lo que se pretende en el presente trabajo es evaluar la capacidad de aprendizaje de las distintas capas y, para ello, se realizaron ocho experimentos diferentes congelando progresivamente desde la primera hasta la séptima de sus ocho capas con capacidad de aprendizaje. Así pues, en primer lugar, no se congeló ninguna de las capas y, por lo tanto, todos los parámetros de las mismas adquirieron nuevo conocimiento. A continuación, sólo se congelaron los parámetros con capacidad de aprendizaje pertenecientes a la primera capa convolucional, conservando así su conocimiento inicial. Este procedimiento se repitió congelando las dos primeras capas, las tres primeras capas y así sucesivamente hasta que se congelaron las siete primeras capas. Como era de esperar, los primeros experimentos necesitaron más tiempo y recursos computacionales que los últimos, ya que era necesario entrenar un mayor número de parámetros durante el proceso de ajuste final.

#### 2.4. Análisis del rendimiento

Los ocho experimentos conducidos para el ajuste final de la RNC utilizaron la base de datos PDB para el entrenamiento (basado en un tamaño de lote de 32 imágenes durante 10 épocas), así como la base de datos PC2017DB para la validación. El enorme desequilibrio entre clases en este último caso (el 98 % de los segmentos pertenecían a la clase de alta calidad y el 2 % restante a la de baja calidad), no fue un factor de confusión importante ya que estos datos sólo se utilizaron en fase de validación, una vez que la RNC ya estaba entrada sobre una base de datos totalmente ba-

lanceada. Los resultados se evaluaron utilizando medidas estadísticas clásicas, tales como la sensibilidad ( $Se$ ), la especificidad ( $Es$ ) y la exactitud ( $Ex$ ). La  $Se$  se consideró como la tasa de segmentos de ECG de alta calidad correctamente identificados, mientras que por el contrario la  $Es$  fue el número de segmentos de baja calidad correctamente detectados. Por último, la  $Ex$  fue el número total de todos los intervalos de ECG clasificados correctamente. Además, para obtener una validación robusta, cada uno de los ocho experimentos se repitió 10 veces y se obtuvo el promedio de los resultados de clasificación con la base de datos de validación. Por lo tanto, en la siguiente sección se detallarán los resultados en media  $\pm$  desviación estándar para estas 10 iteraciones.

### 3. Resultados

La Tabla 2 presenta los resultados de clasificación obtenidos para cada uno de los 8 modelos entrenados y validados como se ha descrito previamente. Como se puede observar, los valores de  $Ex$ ,  $Se$  y  $Es$  son muy similares en todos los casos. De hecho, tras realizar un test de McNemar no se observaron diferencias estadísticamente significativas entre los resultados de ninguno de los ocho modelos bajo estudio. Además, también cabe resaltar que, a pesar de que el experimento se repitió 10 veces, los valores de la desviación estándar apenas variaron entre el 1 y el 3 %. Asimismo, la  $Se$  y la  $Es$  también presentaron valores muy balanceados, incluso a pesar de la gran diferencia existente entre el número de segmentos de ECG pertenecientes a cada clase en la base de datos PC2017DB.

### 4. Discusión

A la vista de los resultados obtenidos, la congelación de las capas durante el ajuste final de AlexNet para la evaluación de la calidad del ECG no mejora excesivamente el rendimiento de clasificación. A pesar de este resultado, este estudio se considera interesante debido a la mejora sustancial de los tiempos de entrenamiento y la optimización de recursos computacionales, cuando se ajusta al mínimo los parámetros de aprendizaje que se tienen que

Capa	Se	Es	Ex
0 (Original)	0.881±0,009	0.838±0,09	0.875±0,009
1	0.883±0,014	0.841±0,024	0.873±0,013
1 a 2	0.885±0,013	0.834±0,030	0.878±0,011
1 a 3	0.878±0,010	0.850±0,017	0.878±0,009
1 a 4	0.877±0,019	0.849±0,036	0.876±0,017
1 a 5	0.884±0,008	0.837±0,018	0.883±0,007
1 a 6	0.882±0,008	0.845±0,017	0.881±0,008
1 a 7	0.887±0,009	0.837±0,018	0.885±0,009

**Tabla 2.** Resultados de clasificación obtenidos en la fase de validación (sobre la base de datos PC2017DB) después de congelar progresivamente cada una de las capas con capacidad de aprendizaje del modelo AlexNet.

re-entrenar. En este sentido, el presente trabajo es pionero si se tiene en cuenta que se ha enfocado a la evaluación de la calidad del ECG, ya que los autores no conocen estudios previos que evalúen el impacto de la congelación de capas en algoritmos de aprendizaje profundo utilizados para tal fin. Sin embargo, hay que tener en cuenta que este estudio podría no ser directamente extrapolable a otros algoritmos basados en RNCs diferentes a AlexNet. Esto es debido a que, por su construcción, AlexNet presenta una estructura secuencial de capas, mientras que otras arquitecturas pueden presentar ramas paralelas, como GoogleNet, y la congelación de los parámetros con capacidad de aprendizaje podría tener un efecto diferente [12].

No obstante, cabe destacar que los resultados obtenidos en el presente trabajo se encuentran en línea con los aportados por otros anteriores en cuanto a congelación de capas, aunque en contextos muy diferentes de la evaluación de la calidad del ECG. Por ejemplo, en el trabajo de Xiao et. al. se propuso un nuevo enfoque para calcular de manera inteligente el número de capas a congelar en el caso de la clasificación de imágenes naturales [13]. En dicho trabajo, el algoritmo fue verificado con varias RNCs, incluyendo VGG, ResNet, y DenseNet, proporcionando en todos los casos una mejora sustancial en el entrenamiento, sin bajar el rendimiento de clasificación. Por otro lado, Chudzick et. al. [14], desarrollaron un trabajo en el que estudiaban el impacto de la congelación de capas para la detección de microaneurismas a partir de fotografías del fondo del ojo mediante una RNC previamente propuesta y pre-entrenada por ellos mismos. Los resultados que obtuvieron mostraron una tendencia constante en los parámetros de clasificación cuando se congelaban muchas de las capas de la red, pero reduciendo considerablemente el tiempo de entrenamiento.

## 5. Conclusiones

Se ha observado como la congelación de capas en el ajuste final de la red AlexNet para la evaluación de la calidad del ECG de larga duración presentan ventajas interesantes. Así pues, a pesar de que el rendimiento de la clasificación de diferentes segmentos de ECG de alta y baja calidad no mejora significativamente, sí que sirve para optimizar el tiempo de entrenamiento y los recursos computacionales. En futuros trabajos se abordará el efecto de la congelación de capas en otras RNCs, así como en otras aplicaciones basadas en el registro del ECG.

## Agradecimientos

Este trabajo ha sido financiado por los proyectos de investigación competitiva PID2021-00X128525-IV0, PID2021-123804OB-I00 y TED2021-130935B-I00 del Gobierno de España 10.13039/ 501100011033 conjuntamente con el Fondo Europeo de Desarrollo Regional (EU), SBPLY/17/180501/000411 y SBPLY/21/180501/000186 de la Junta de Comunidades de Castilla-La Mancha, y AI-CO/2021/286 de la Generalitat Valenciana, además de la compañía Daiichi Sankyo S.A.U.

## Referencias

- [1] Sana F, Isselbacher EM, Singh JP, Heist EK, Pathik B, Ar-moundas AA. Wearable devices for ambulatory cardiac monitoring: JACC state-of-the-art review. *Journal of the American College of Cardiology* 2020;75(13):1582–1592.
- [2] Stachon P, Ahrens I, Faber T, Bode C, Zirlirk A. Asymptomatic atrial fibrillation and risk of stroke. *Panminerva Med Dec* 2015;57(4):211–5.
- [3] Smital L, Haider CR, Vitek M, et. al. Real-time quality assessment of long-term ECG signals recorded by wearables in free-living conditions. *IEEE Trans Biomed Eng* 2020; 67(10):2721–2734.
- [4] Satija U, Ramkumar B, Manikandan MS. An automated ECG quality assessment method for unsupervised diagnostic systems. *Biocybern Biomed Eng* 2018;38:54–70.
- [5] Huerta Herraiz Á, Martínez-Rodrigo A, Bertomeu-González V, Quesada A, Rieta JJ, Alcaraz R. A deep learning approach for featureless robust quality assessment of intermittent atrial fibrillation recordings from portable and wearable devices. *Entropy* 2020;22(7):733.
- [6] Weiss K, Khoshgoftaar TM, Wang D. A survey of transfer learning. *Journal of Big data* 2016;3(1):1–40.
- [7] Krishna ST, Kalluri HK. Deep learning and transfer learning approaches for image classification. *Int Journal of Recent Technology and Engineering* 2019;7:427–432.
- [8] Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. In *Advances in Neural Inform. Processing Systems*. 2012; 1097–1105.
- [9] Hong S, Zhou Y, Shang J, Xiao C, Sun J. Opportunities and challenges of deep learning methods for ECG data: A systematic review. *Comput Biol Med* 2020;122:103801.
- [10] Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, Huang Z, Karpathy A, Khosla A, Bernstein M, Berg AC, Fei-Fei L. ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision IJCV* 2015;115(3):211–252.
- [11] Clifford GD, Liu C, Moody B, Lehman LWH, Silva I, Li Q, Johnson AE, Mark RG. AF classification from a short single lead ECG recording: The PhysioNet/Computing in Cardiology Challenge 2017. *Comput Cardiol* 2017;44.
- [12] Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A. Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015; 1–9.
- [13] Xiao X, Mudiyansele TB, Ji C, Hu J, Pan Y. Fast deep learning training through intelligently freezing layers. In *International Conference on Internet of Things and IEEE Green Computing and Communications and IEEE Cyber, Physical and Social Computing and IEEE Smart Data*. 2019; 1225–1232.
- [14] Chudzick P, Majumdar S, Caliva F, Al-Diri B, Hunter A. Microaneurysm detection using deep learning and interleaved freezing. In *Medical Imaging 2018: Image Processing*, volume 10574. 2018; 379–387.

# Assessment of common information in surface electromyography recordings with adhesive electrodes and an intravaginal probe

J. M. Mira-Tomas<sup>1</sup>, M. Albaladejo-Belmonte<sup>1</sup>, E. Guijarro<sup>1</sup>, P. Villa-Muñoz<sup>2</sup>, F. J. Nohales-Alfonso<sup>2</sup>, J. Alberola-Rubio<sup>3</sup>, J. Garcia-Casado<sup>1\*</sup>

<sup>1</sup> Centro de Investigación e Innovación en Bioingeniería (CI2B), Universitat Politècnica de València, Valencia, Spain, {juamito, moalbel, eguijaro, jgarcia}@ci2b.upv.es

<sup>2</sup> Servicio de Obstetricia y Ginecología, HUP La Fe de Valencia, Spain {vimupau, fnohales}@gmail.com

<sup>3</sup> Sonda Devices S.L., Valencia, Spain. palberola@sondadevices.com

## Abstract

*Dysfunctions in pelvic floor muscles (PFM) are a highly prevalent group of pathologies which critically alter daily life activities. Surface electromyography (sEMG) of PFM has emerged as a potential evaluative tool but recordings are affected by interferences of nearby muscle groups. The aim of this study is to assess the association between internal and external sEMG signals in order to better identify the different PFM sources. Four monopolar and two bipolar signals were recorded from the perineum (external recordings), and two monopolar signals with an intravaginal probe in 32 women with vulvodynia. Cross-correlation (CC) and normalized mutual information (NMI) were computed between signal pairs to assess their common information. External monopolar electrodes showed high CC (82 %) and NMI (19 %) among them, probably due to common mode interferences and volume conduction. External monopolar signals on posterior PFM have relevant common information with probe signals of the same side (CC: 55 %, NMI:6.5 %), suggesting that deep PFM activity could be monitored without using intravaginal probes, which are painful for some patients. Little common information was found between intravaginal probe and external bipolar signals, suggesting that deep and surface activity can be assessed separately with these recordings set-ups. These findings can be relevant for the recording and interpretation of sEMG signals when assessing PFM electrophysiological condition in the clinical management of PFM disorders.*

## 1. Introduction

Pelvic floor muscles (PFM) are a group of muscles and ligaments that lay at the base of the abdominal cavity. They are responsible for supporting pelvic organs (bladder, rectus, and uterus/prostate) in right position. Moreover, they play a key role in providing stability in standing, as well as in sphincteric and sexual functions [1]. They are classified in 3 layers depending on their depth, so activity from deep muscles is different from the surface ones.

Dysfunctions of PFM include urinary incontinence, pelvic organs prolapse, anal incontinence, sensory abnormalities of the lower urinary tract, defecatory dysfunction and chronic pain syndromes related to the pelvic floor. They may significantly affect common daily living tasks of patients and have a high economic impact, as in the case of urinary incontinence, which accounts for at least 2% of the

health budget in developed countries [2]. The high estimated probability (23.7 % in USA population [3]) of suffering from one or more of these pelvic floor disorders makes it necessary to develop accurate and quick diagnostic tools and treatments approaches. Diagnosis may include from an initial manual examination to some extra tests such as defecating proctogram, uroflow test and surface electromyography (sEMG) [4].

Electrophysiological evaluation with sEMG can be a powerful tool to objectively assess PFM condition. Although some studies in the field have recorded PFM sEMG with self-adhesive surface electrodes, their use has been questioned because of their high susceptibility to crosstalk from neighboring muscular groups, which is often due to volume conduction of electrical activity [5]. On the other hand, intracavitary probes are used for mainly recording deep PFM activity, although they can also be affected by crosstalk and can cause pain and discomfort to the patients.

The aim of this study was thus to quantify the common information between PFM sEMG signals internally and externally recorded and to assess the origin and implication of this shared information. The starting hypothesis is that external sEMG signals also contain information from deep PFM and therefore the use of an intravaginal probe could be eliminated. To do this, sEMG signals were simultaneously acquired with an intravaginal probe and self-adhesive electrodes on the perineum surface.

## 2. Materials and methods

### 2.1. Database

A cohort of 32 female patients diagnosed with chronic pelvic pain associated with vulvodynia participated in a prospective follow-up study performed at the Hospital Universitari i Politènic La Fe (Valencia, Spain), which met the Helsinki Declaration.

Simultaneous sEMG recording with external adhesive electrode and intravaginal probe was performed for each patient. Two pairs of disposable Ag/AgCl electrodes (Red Dot 2660-5, 3M, St. Paul, MN, USA) were placed on both sides of the vulva over the perineum and two additional

electrodes (ground: GND, reference: REF) were attached on both iliac spines. These regions were previously exfoliated with an abrasive gel (Nuprep 114g, Weaver and Company, Aurora, CO, USA) to reduce skin-electrode impedance. An intravaginal probe with recording poles on both sides (Periform®+, Neen, Sutton-in-Ashfield, Nottinghamshire, UK) and lubricated with conductive gel was also used to record deep PFM activity of left and right sides.

A multipurpose biomedical signal amplifier (Grass 15LT+4 Grass 15A94, Grass Instruments, West Warwick, RI, USA) was used to record four monopolar signals (M1, M2, M3, M4) from the perineum surface and two monopolar signals (P1, P2) from the intracavitary probe. The device was configured with a band-pass filter between 3 and 1000 Hz and signals were digitalized at a rate of 10 kHz with 16 bits. Two bipolar signals were additionally computed as the difference of the two external monopolar signals of the same side (B1 = M1 – M3 (right), B2 = M2 – M4 (left)).

During the recording, patients were in a dorsal lithotomy position and were asked to follow a protocol of PFM voluntary contractions designed by clinicians. Each sEMG recording consisted of 5 maximum voluntary contractions of 5 s separated by 10 s of maximum relaxation. The signal segment recorded from 1 s before the first contraction to 1 s after the last one was annotated for subsequent analysis. Signals were digitally filtered to attenuate frequency components out of [30, 450] Hz bandwidth and the power line interference (50 Hz), as in [6]. An example of the 8 sEMG signals of one patient is represented in Figure 1.

## 2.2. Similarity metrics

Common information between sEMG signals was assessed according to two different similarity metrics: cross-correlation (CC) and normalized mutual information (NMI). Signals were previously standardized to have zero mean and unit variance.

Cross-correlation is one of the common indicators for assessing crosstalk in electromyography [7]. It quantifies

the magnitude of any common component contained in two signals, following the assumption that the shapes of the waveforms from both muscles under consideration are the same. Since we aim to quantify instantaneous common information, we used Pearson's correlation coefficient that is a simple and normalized (range -1 to 1) indicator:

$$CC = \frac{cov(X, Y)}{\sigma_X \sigma_Y} \quad (1)$$

, where  $cov(X, Y)$  is the covariance of signals  $x[n]$  and  $y[n]$ , and  $\sigma_X, \sigma_Y$  are their standard deviation.

Normalized mutual information in probability and information theory expresses the mutual dependence between two random variables and has been used in several sEMG applications[8]. Based on Shannon entropy mutual information is computed as in (2):

$$MI(X; Y) = \sum_{x_i y_j} P_{X,Y}(x_i, y_j) \log_2 \frac{P_{X,Y}(x_i, y_j)}{P_X(x_i) P_Y(y_j)} \quad (2)$$

, where  $P_X$  and  $P_Y$  are the probability distributions of  $x[n]$  and  $y[n]$ , and  $(P_{X,Y})$  their joint probability distribution.  $P_X, P_Y$  and  $P_{X,Y}$  were obtained by computing the histograms of  $x[n]$  and  $y[n]$  and their joint histogram, respectively, for  $b$  bins and dividing them by the number of signal samples ( $N$ ). Value of  $b$  was determined by the Rice's rule:

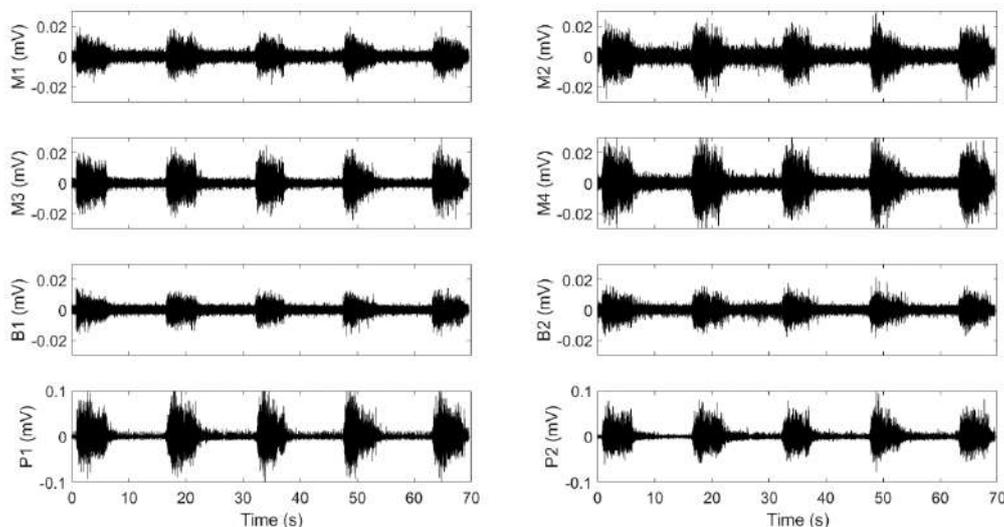
$$b = \lfloor 2^{\frac{3}{\sqrt{N}}} \rfloor \quad (3)$$

To set the range of  $MI(X; Y)$  from 0 to 1, it is normalized according to equation (4):

$$NMI(X; Y) = \frac{MI(X; Y)}{\sqrt{H(X)H(Y)}} \quad (4)$$

, where  $H(\cdot)$  is the Shannon entropy estimation for an individual signal, computed as follows:

$$H(X) = - \sum_{x_i} P_X(x_i) \log_2(P_X(x_i)) \quad (5)$$



**Figure 1.** Conditioned signals including 5 voluntary contractions and 4 inter-contractions resting periods. M1, M3, B1 and P1 correspond to right PFM side and M2, M4, B2 and P2 to the left PFM side

### 3. Results

Table 1 shows the mean  $\pm$  standard deviation of the parameters assessing common information (CC: lower triangular matrix, NMI: upper triangular matrix) over the whole population. Relevantly coupled signal pairs for both CC and NMI have been selected based on a threshold criterion. Relevant pairs were considered those whose absolute values were cumulatively added (from highest to lowest) until they reached the 50% threshold of the total sum of all absolute values of the same metric. Relevantly coupled signal pairs are shown in gray in Table 1, and are represented in Figures 2 and 3, with different color according to the type of connection (brown: left-sided, blue: right-sided, green: inter-sided).

The highest CC and NMI values were obtained between ipsilateral and contralateral sEMG signals of external monopolar electrodes. Particularly, values between M1 vs. M2 (anterior area of PFM) and M4 vs. M2 (left area of PFM) were higher than those between M3 vs. M4 (posterior area of PFM) and M1 vs. M3 (right area of PFM). On the other hand, contralateral bipolar and probe signals (B1 vs. B2, P1 vs. P2, respectively) showed low CC and NMI values.

Both probe and bipolar signals showed higher CC and NMI values with posterior than anterior external monopolar signals of their same side. However, common information between bipolar surface signals and probe signals was low and their CC and NMI were below the relevance threshold (CC: 42 %, NMI: 6.0 %).

### 4. Discussion

In the present study common information in sEMG signals recorded with self-adhesive electrodes and intravaginal probes were assessed according to CC and NMI, which

provided similar results. Some studies have previously assessed correlation between two bipolar configurations (ipsilateral and contralateral recordings) of sEMG with the same vaginal probe [10] and between PFM sEMG and intravaginal pressure computing Pearson’s coefficient of mean scores of contractions [11]. However, no study has so far quantified the common information between deep and superficial PFM activity recorded simultaneously with sEMG electrodes. Monopolar signals externally recorded (M1, M2, M3, M4) showed the highest common information rates, which was especially remarkable in electrodes located in contralateral PFM sides. One reason could be that motoneurons of different motor units are synchronized: although PFM sides have their own nerve supply, they cannot be contracted independently, which suggests that same pre-synaptic signal may be exciting their motoneurons simultaneously leading to similar waveforms [12]. However, this would also lead to significant common information between P1 and P2, that was not found. The most plausible reason is that these high common information rates were a result of crosstalk, i.e. electrical potentials detected not only from the muscle under the electrode, but also from muscles further away due to volume conduction.

Crosstalk potentials would be more similar in nearby electrodes, since signals’ characteristics change as they travel through the body tissue [13], what would also justify smaller common information in diagonal (more distant) monopolar pairs.

An instrumental factor could also have influenced high CC and NMI values obtained between external monopolar signals. Reference electrode was placed over the patient’s right ischiatic spine, what may not be far enough for not sensing part of the activity of right posterior PFM.

	NMI (%)							
	M1	M2	M3	M4	B1	B2	P1	P2
M1		12.0 $\pm$ 5.0	10.0 $\pm$ 5.5	8.2 $\pm$ 3.9	5.9 $\pm$ 2.9	2.4 $\pm$ 1.5	4.6 $\pm$ 3.6	4.3 $\pm$ 2.3
M2	70 $\pm$ 11		6.6 $\pm$ 3.2	19.0 $\pm$ 6.9	3.5 $\pm$ 1.6	3.4 $\pm$ 1.8	4.2 $\pm$ 3.1	6.0 $\pm$ 2.9
M3	64 $\pm$ 16	44 $\pm$ 17		7.9 $\pm$ 3.5	6.5 $\pm$ 3.8	2.6 $\pm$ 1.8	9.0 $\pm$ 4.9	4.9 $\pm$ 2.4
M4	56 $\pm$ 15	82 $\pm$ 14	50 $\pm$ 15		2.6 $\pm$ 1.3	5.7 $\pm$ 3.1	4.8 $\pm$ 3.5	8.1 $\pm$ 3.4
B1	35 $\pm$ 27	26 $\pm$ 18	-45 $\pm$ 24	4 $\pm$ 20		3.7 $\pm$ 1.9	3.8 $\pm$ 2.2	2.4 $\pm$ 1.4
B2	14 $\pm$ 16	14 $\pm$ 20	-15 $\pm$ 18	-42 $\pm$ 18	36 $\pm$ 13		2.5 $\pm$ 1.5	3.6 $\pm$ 1.9
P1	31 $\pm$ 18	18 $\pm$ 18	55 $\pm$ 16	22 $\pm$ 17	-31 $\pm$ 18	-10 $\pm$ 13		4.8 $\pm$ 3.4
P2	32 $\pm$ 13	40 $\pm$ 13	34 $\pm$ 13	50 $\pm$ 12	-4 $\pm$ 12	-26 $\pm$ 17	24 $\pm$ 13	

Table 1. Lower triangular matrix represents CC mean  $\pm$  standard deviation values of the population between each pair of signals. Upper triangular matrix shows the same information for NMI.

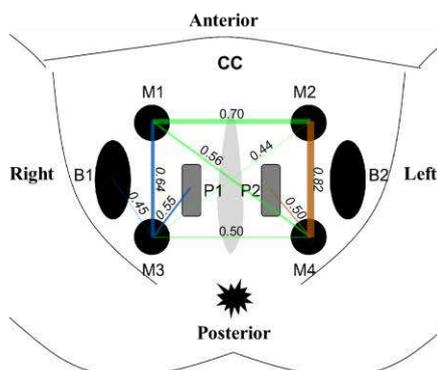


Figure 2. Relevant cross-correlation values between channel pairs

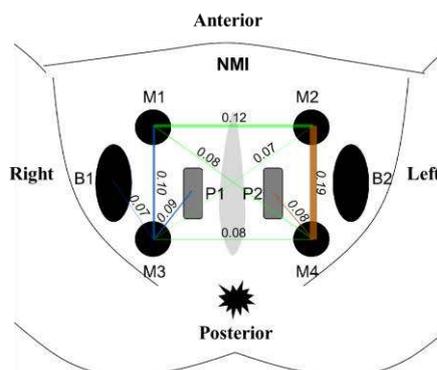


Figure 3. Relevant normalized mutual information values between channel pairs

This would partially cancel common information with M3 (closest electrode) and would explain why CC and NMI between anterior (M1 vs. M2) and left-sided external signals (M4 vs. M2) was greater than the ones for right side (M1 vs. M3) and posterior (M3 vs. M4).

Unlike external monopolar signals, contralateral probe and external bipolar recordings (independently assessed) shared minimum information. In view of the above, the reason in the first case would be that crosstalk had a significantly lower power than PFM activity at the intracavitary recording site (electrodes close to activity source); while in the second case it would be that common mode interferences on perineum surface were mostly cancelled in bipolar signals when subtracting monopolar signals. It is also noteworthy that minimum common information was found in probe-bipolar signal pairs.

On the other hand, signals recorded by intravaginal probe electrodes exhibited high common information with those of posterior monopolar electrodes of their corresponding side. This could be related to the fact that from an anatomical point of view, posterior external electrodes are closer to deep PFM than the anterior ones [14].

Bipolar channels also showed a high degree of common information with posterior monopolar electrodes of the same side. This was expected since bipolar signals were obtained from same-sided monopolar signals. The reason why similarity was higher with posterior than with anterior monopolar electrodes could be that signal-to-noise ratio of the first ones is generally higher since they are closer to the deep PFM [15] which generate the majority of the electrical activity that is associated with PFM as a whole.

## 5. Conclusions

The present study assessed common information in surface electromyography recordings performed with self-adhesive electrodes and intravaginal probes. Monopolar signals recorded from the perineum surface showed the highest level of common information among them, probably associated to crosstalk rather than to common activation patterns. On the other hand, the minimum shared information was found between external bipolar signals and recordings of intravaginal probes, suggesting that they could be used by clinicians to separately assess the activity of superficial and deep PFM.

A relevant degree of common activity was also found between posterior monopolar signals from the perineum and probe recordings on the same side, suggesting that further studies could focus on retrieving the activity of deep PFM from external monopolar sEMG recordings to avoid the use of intracavitary probes, which can cause pain in patients suffering from chronic pain syndromes.

## Acknowledgements

This study was funded by Generalitat Valenciana in Programa para la promoción de I+D+i ACIF/2021/012, AICO/2021/126; and by private contracts from Merz Pharmaceuticals GmbH S.L.

## References

- [1] Quaghebeur J, Petros P, Wyndaele JJ, De Wachter S. Pelvic floor function, dysfunction, and treatment, *Eur. J. Obstet. Gynecol. Reprod. Biol.*, vol. 265, 2021, pp. 143-9, doi: 10.1016/j.ejogrb.2021.08.026.
- [2] Hu T. Impact of Urinary Incontinence on Health-Care Costs, *J. Am. Geriatr. Soc.*, vol. 38, sup 3, 1990, pp. 292-5, doi: 10.1111/j.1532-5415.1990.tb03507.x.
- [3] Nygaard I, et al. Prevalence of Symptomatic Pelvic Floor Disorders in US Women, *JAMA*, vol. 300, sup 11, 2008, pp. 1311-6, doi: 10.1001/jama.300.11.1311.
- [4] Pelvic Floor Dysfunction: Symptoms, Causes & Treatment», *Cleveland Clinic*. <https://my.clevelandclinic.org/health/diseases/14459-pelvic-floor-dysfunction> (Accessed: June 2022).
- [5] Flury N, Koenig I, Radlinger L. Crosstalk considerations in studies evaluating pelvic floor muscles using surface electromyography in women: a scoping review, *Arch. Gynecol. Obstet.*, vol. 295, sup 4, 2017, pp. 799-809, doi: 10.1007/s00404-017-4300-5.
- [6] Albaladejo-Belmonte M, Tarazona-Motes M, Nohales-Alfonso FJ, De-Arriba M, Alberola-Rubio J, Garcia-Casado J. Characterization of Pelvic Floor Activity in Healthy Subjects and with Chronic Pelvic Pain: Diagnostic Potential of Surface Electromyography, *Sensors*, vol. 21, sup 6, 2021, doi: 10.3390/s21062225.
- [7] Winter DA, Fuglevand AJ, Archer SE. Crosstalk in surface electromyography: Theoretical and practical estimates, *J. Electromyogr. Kinesiol.*, vol. 4, sup 1, 1994, pp. 15-26, doi: 10.1016/1050-6411(94)90023-X.
- [8] Bingham A, Arjunan SP, Jelfs B, Kumar DK. Normalised Mutual Information of High-Density Surface Electromyography during Muscle Fatigue, *Entropy*, vol. 19, sup 12, 2017, doi: 10.3390/e19120697.
- [9] Ballmer C, et al. Electromyography of pelvic floor muscles with true differential versus faux differential electrode configuration, *Int. Urogynecology J.*, vol. 31, sup 10, 2020, pp. 2051-9, doi: 10.1007/s00192-020-04225-4.
- [10] Madill SJ, McLean L. Quantification of abdominal and pelvic floor muscle synergies in response to voluntary pelvic floor muscle contractions, *J. Electromyogr. Kinesiol.*, vol. 18, sup 6, 2008, pp. 955-964, doi: 10.1016/j.jelekin.2007.05.001.
- [11] Workman DE, Cassisi JE, Dougherty MC. Validation of surface EMG as a measure of intravaginal and intra-abdominal activity: Implications for biofeedback-assisted Kegel exercises, *Psychophysiology*, vol. 30, sup 1, 2007 pp. 120-5, doi: 10.1111/j.1469-8986.1993.tb03210.x.
- [12] Enck P, Vodusek DB. Electromyography of pelvic floor muscles, *J. Electromyogr. Kinesiol. Off. J. Int. Soc. Electrophysiol. Kinesiol.*, vol. 16, sup 6, 2006, pp. 568-577, doi: 10.1016/j.jelekin.2006.08.007.
- [13] van Vugt JPP, van Dijk JG. A convenient method to reduce crosstalk in surface EMG, *Clin. Neurophysiol.*, vol. 112, sup 4, 2001, pp. 583-592, doi: 10.1016/S1388-2457(01)00482-5.
- [14] Katya Carrillo G, Antonella Sanguinet M. Anatomía del piso pélvico, *Rev. Médica Clínica Las Condes*, vol. 24, sup 2, 2013, pp. 185-9, doi: 10.1016/S0716-8640(13)70148-2.
- [15] Day S. Important factors in surface EMG measurement. *Bortec Biomedical Ltd publishers*, 2002, pp. 1-17

# Alteraciones patológicas de la estructura neurofisiológica en los diferentes subtipos de migraña en mujeres

V. Gutiérrez-de Pablo<sup>1,2</sup>, J. Gómez-Pilar<sup>1,2</sup>, V. Rodríguez-González<sup>1,2</sup>, J. Poza<sup>1,2,3</sup>, A. L. Guerrero<sup>4</sup>, M. Alves-Ferreira<sup>5</sup>, N. Pinto<sup>6</sup>, R. Hornero<sup>1,2,3</sup>, C. Gómez<sup>1,2</sup>

<sup>1</sup> Grupo de Ingeniería Biomédica, Universidad de Valladolid, Valladolid, España

<sup>2</sup> Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), España

<sup>3</sup> IMUVA, Instituto de Investigación en Matemáticas de la Universidad de Valladolid, Valladolid, España

<sup>4</sup> Unidad de Cefaleas, Servicio de Neurología, Hospital Clínico Universitario de Valladolid, España

<sup>5</sup> UnIGENe, IBMC – Institute for Molecular and Cell Biology; i3S, Porto, Portugal

<sup>6</sup> Instituto de Patologia e Imunologia Molecular da Universidade do Porto (IPATIMUP), Porto, Portugal

## Resumen

*La migraña puede ser clasificada en dos subtipos: migraña crónica (MC) y episódica (ME), en función de los días que se sufre dolor de cabeza al mes (más, o menos de 15 días). Estos dos subtipos de migraña requieren un tratamiento diferente, lo cual sugiere que el desarrollo de cada una también es distinta. El objetivo de este estudio es analizar las alteraciones en las relaciones de la estructura existente entre las diferentes jerarquías de la actividad neurofisiológica (i.e., activación local y sincronización) asociadas a los subtipos de la migraña. Para ello, se analizó la actividad electroencefalográfica (EEG) basal de 90 mujeres (30 controles, 30 pacientes con ME y 30 pacientes con MC). Se calcularon diferentes medidas ampliamente utilizadas para caracterizar los dos niveles neurofisiológicos y se cuantificaron las correlaciones existentes entre todas ellas para generar redes de asociación que permitan su análisis conjunto. Los resultados mostraron diferencias entre las redes de los tres grupos bajo estudio, con cambios estadísticamente significativos entre las controles, que actúan como grupo intermedio, y las pacientes con ME y MC. Además, se observó que la ME y la MC inducen modificaciones específicas en la estructura jerárquica electrofisiológica. Estos resultados podrían estar asociados a posibles alteraciones biológicas que afectan a los generadores neuronales a diferentes niveles, lo que mostraría una diferenciación objetiva entre ambos subtipos de migraña.*

## 1. Introducción

La migraña es una patología neurológica compleja caracterizada por un dolor de cabeza pulsátil e intenso, y otros síntomas que varían desde cansancio hasta parálisis de alguna parte del cuerpo [1]. Esta patología puede ir acompañada de aura, es decir, aparición de diferente sintomatología, incluyendo alteraciones visuales, auditivas y motoras, las cuales pueden ir acompañadas o precedidas de dolor [1]. Por otro lado, la exposición a luces, ruidos o situaciones de alto estrés pueden ser efectos disparadores de ataques de esta patología [2]. Además, es ampliamente aceptado que la migraña se ve influida por factores genéticos y hormonales [1], [3], [4]. Es por ello que es una de las patologías neurológicas más prevalentes, que afecta a un 14% de la población adulta en países desarrollados, incrementándose esta hasta un 18% en el caso de las mujeres [5]. Dependiendo de la frecuencia de aparición de los ataques de migraña, se distingue entre migraña episódica (ME), si el paciente sufre ataques con una frecuencia inferior a 15 días al mes, y migraña crónica (MC), si el paciente sufre ataques con una frecuencia igual o superior a 15 días al mes y al menos 8 de estos días cumplen el criterio de dolor migrañoso de intensidad moderada a alta [6].

A nivel fisiológico, la migraña provoca diversas alteraciones que pueden observarse mediante resonancia magnética funcional (fMRI), magnetoencefalografía (MEG) o electroencefalografía (EEG) [7]–[9]. En este estudio, se ha empleado EEG debido a su bajo coste, portabilidad, alta resolución temporal y no invasividad [9]. Por todas estas ventajas, es una técnica ampliamente utilizada en los ámbitos científico y clínico [2], [9].

Diversos estudios han analizado el efecto que la migraña provoca sobre los diferentes niveles neurofisiológicos de la actividad cerebral. Concretamente, se diferencian dos niveles: activación local y sincronización. El nivel de activación local se caracteriza en términos de las propiedades espectrales, no lineales y de dinámica, en cada región de interés (ROI). El nivel de sincronización se describe mediante la conectividad funcional (CF) entre pares de ROIs para cada banda de frecuencia, varias propiedades de red y la estructura multicapa de la CF. A nivel de activación local, se ha observado un cambio en las propiedades espectrales de la actividad EEG en pacientes con migraña al compararlos con sujetos de control [9]. Otro estudio reveló que la migraña está asociada con un incremento de la entropía de la señal en etapas previas al inicio de un ataque, en comparación a los periodos entre ataques [10]. A nivel de sincronización, se ha observado un incremento de coherencia en fases previas al ataque de migraña; no obstante, durante el ataque o entre ataques, la coherencia es menor que la de sujetos de control [11]. Sin embargo, prácticamente no existen trabajos en los que se trate de diferenciar entre pacientes con ME y con MC, y en los que se estimen las diferencias que puedan existir entre estos subgrupos y pacientes sanos. Además, tampoco existen trabajos en los que se estudien de manera conjunta los diferentes niveles que componen la estructura electrofisiológica cerebral.

Bajo estas premisas, la hipótesis de partida de este trabajo es que la estructura neurofisiológica y las relaciones entre los niveles que la componen se ven alterados por la presencia de la migraña. El objetivo del trabajo es analizar la relación existente entre diferentes parámetros que caracterizan la actividad cerebral para identificar las huellas neurofisiológicas asociadas a la ME y a la MC.

## 2. Materiales y métodos

### 2.1. Participantes

En el estudio se incluyeron 90 mujeres entre 18 y 40 años, divididas en 30 controles (edad: 29 [26 35] años, mediana

[rango intercuartil]), 30 pacientes con ME (edad: 32.5 [29-36] años) y 30 pacientes con MC (edad: 33.5 [27-39] años). No existían diferencias significativas en edad entre los grupos ( $p > 0.05$ , test  $U$  de Mann-Whitney). Todas las participantes fueron mujeres, para evitar posibles sesgos debidos al sexo, a la prevalencia desbalanceada que posee la migraña y a la relación entre esta patología y factores hormonales. Las pacientes fueron reclutadas en su primera visita ambulatoria, libres de tratamiento, en la Unidad de Cefaleas del Hospital Clínico Universitario de Valladolid (España) y en el *Instituto de Investigação e Inovação em Saúde da Universidade do Porto* (Portugal). Todas las pacientes fueron diagnosticadas según los criterios del *International Classification of Headache Disorders* [6]. Se excluyeron del análisis a pacientes con otras patologías neurológicas o psiquiátricas, con historial de abuso de sustancias o antecedentes de traumas que pudieran afectar al EEG.

## 2.2. Registro de EEG y preprocesado

Los registros de EEG se registraron mediante un sistema de 32 canales (BrainVision®), colocados siguiendo el sistema internacional 10-10 y empleando el electrodo Cz como referencia. Se registraron 10 minutos de actividad EEG en estado de reposo, con las participantes relajadas y con los ojos cerrados. La frecuencia de muestreo empleada para la adquisición de los registros fue de 500 Hz.

Posteriormente, se preprocesaron los registros para eliminar ruido y artefactos que no correspondían a actividad cerebral [9]. En primer lugar, se aplicó un re-referenciado promedio para recuperar el electrodo Cz. En segundo lugar, se eliminó la media de la señal para eliminar la componente continua. Posteriormente, las señales se filtraron entre 0.4 y 98 Hz con un filtro de respuesta finita al impulso (FIR) de orden 2000 y ventana de Hamming. A continuación, se aplicó un filtro FIR de ranura a 50 Hz (orden 2000, ventana de Hamming) para eliminar la componente de red eléctrica. Después, mediante un análisis de componentes independientes (ICA), se eliminaron los artefactos asociados a parpadeos, ruido ocular y muscular. Por último, las señales se segmentaron en épocas de 5 segundos y, tras una inspección visual, se rechazaron aquellas con algún artefacto remanente.

## 2.3. Localización de fuentes

Una vez preprocesadas las señales EEG, se proyectaron a nivel de fuentes cerebrales de cara a reducir los efectos de conducción de volumen. Para ello, se empleó el algoritmo *standardized Low Resolution Brain Electromagnetic Tomography* (sLORETA) [12]. Se estimó la activación de 15000 fuentes neuronales, que fueron proyectadas en las 68 ROIs del atlas Desikan-Killiany.

## 2.4. Redes de asociación

Las redes de asociación permiten cuantificar e ilustrar gráficamente las relaciones directas (*i.e.*, los enlaces) entre los factores considerados (*i.e.*, los nodos), tomando los sujetos de cada grupo como las instancias a correlar. En este estudio, las variables empleadas permiten caracterizar los niveles neurofisiológicos que conforman la estructura jerárquica global de la actividad cerebral. Las medidas calculadas se indican a continuación:

1) *Análisis espectral*: potencia relativa (RP) en las bandas de frecuencia típicas: delta (0.4-4 Hz), zeta (4-8 Hz), alfa

(8-13 Hz), beta 1 (13-19 Hz), beta 2 (19-30 Hz) y gamma (30-98 Hz); frecuencia mediana (MF); frecuencia alfa individual (IAF); y entropía espectral (SE).

2) *Análisis no lineal*: complejidad de Lempel-Ziv (LZC); entropía muestral (SampEn); entropía difusa (FuzzyEn); y medida de la tendencia central (CTM).

3) *Análisis de dinámica*: actividad, complejidad y movilidad de Hjorth.

4) *Análisis de CF estática*: *phase lag index* (PLI) en las bandas de frecuencia típicas: delta, zeta, alfa, beta 1, beta 2 y gamma.

5) *Análisis de red*: coeficiente de agrupamiento (CIC); grado de red; longitud de camino (PL); *closeness centrality* (CC); entropía de grafo (GE); índice de *small world* (SW); *betweenness centrality* (BC); y *average node strength* (NS).

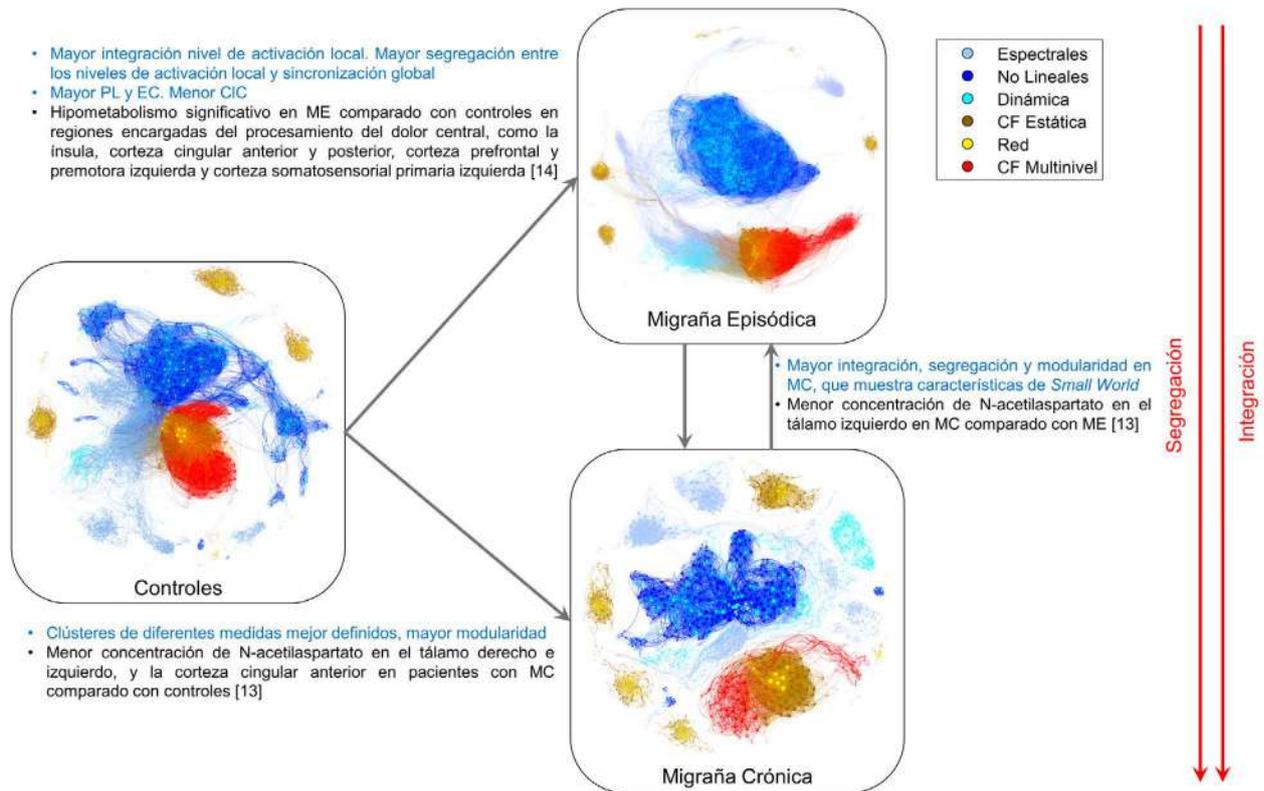
6) *Análisis de CF multinivel*: *overlapping weighted degree* (OWD) y coeficiente de participación.

Para la obtención del grado de relación entre variables, se calculó la correlación parcial de Spearman, corrigiendo el valor de correlación por la edad de los sujetos, para así eliminar el posible sesgo que esta pudiera introducir. Se generaron 50 redes para cada grupo mediante *bootstrapping*; estas redes fueron posteriormente umbralizadas (sólo se mantuvieron los enlaces con un valor superior a 0.8) para aumentar su estabilidad. A continuación, para eliminar el posible sesgo asociado a comparar redes de diferente tamaño, los valores de los enlaces fueron normalizados por la densidad de cada red. Por último, se calcularon varias medidas de la teoría de grafos que permiten resumir diferentes características de estas redes: grado de nodo promedio, PL, CIC, modularidad y centralidad.

Para representar las redes, se utilizó el algoritmo *Force Atlas 2*, que emplea el grado de nodo como variable que genera un efecto repulsor. Además, en estas redes la anchura del enlace simboliza el grado de correlación, teniendo mayor anchura cuanto mayor es el valor de esta. Para poder representar las redes mostradas en la Figura 1, se obtuvo el valor mediano de cada enlace, para así obtener la red mediana de las 50 iteraciones de *bootstrap*.

## 3. Resultados

En la Figura 1 se pueden apreciar las redes de asociación medianas obtenidas mediante *bootstrapping* correspondientes a los tres grupos bajo análisis: controles, ME y MC. En la red correspondiente a los controles se aprecia que las métricas no lineales y de dinámica están muy relacionadas, llegando a agruparse en clústeres muy próximos. Además, parte de las medidas espectrales también se relacionan con las anteriores, si bien aparecen vinculadas a otros pequeños clústeres. De todos estos, el más relevante es el que se asocia con las características englobadas en el nivel de sincronización (CF, red y CF multicapa), como se puede apreciar en el clúster central en el que priman los colores cálidos. Sin embargo, para los niveles bajo estudio, aunque gran parte de los nodos que los caracterizan aparecen integrados en clústeres grandes y heterogéneos, también aparecen algunas medidas independientes, formando clústeres más reducidos. Por otro lado, en la red correspondiente a las pacientes con ME, se observan unos clústeres más separados, habiendo una clara diferenciación entre un clúster de activación local y



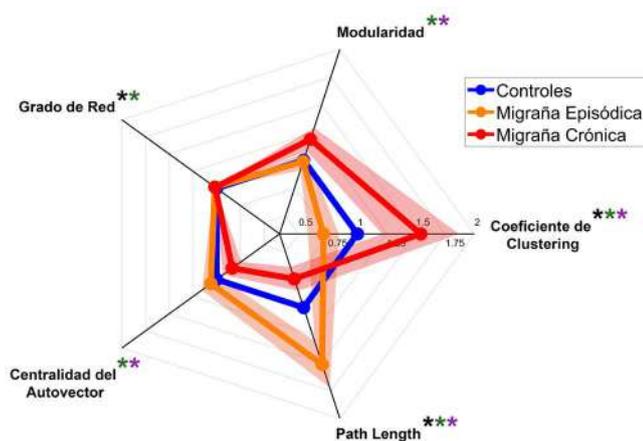
**Figura 1.** Redes de asociación generadas para cada grupo, con correlaciones significativas iguales o superiores a 0.8. En letra azul, resultados de este estudio. En letra negra, resultados biológicos de estudios previos.

otro de sincronización. Además, se aprecia también una mayor integración entre las medidas no lineales y de dinámica, así como una segregación entre estos dos tipos de medidas y las métricas espectrales. Sin embargo, las medidas espectrales aparecen más integradas con las métricas de sincronización, al igual que una parte de las medidas de dinámica. Finalmente, en la red correspondiente a las pacientes con MC, las únicas medidas que permanecen integradas son las del nivel de sincronización. En comparación con los otros dos grupos, la red de MC mantiene integrados los clústeres de los parámetros de CF, red y CF multicapa. Esta estabilidad no se aprecia para el nivel de activación local.

La Figura 2 cuantifica los cambios descritos mediante medidas de red. En ella se aprecian los valores medianos de los parámetros en cada red obtenida y su desviación típica. Se puede apreciar que, por un lado, hay diferencias significativas para el PL y el CIC entre los tres grupos bajo estudio ( $p$ -valor  $< 0.05$ , test  $U$  de Mann Whitney con corrección de Bonferroni). Sin embargo, para el resto de las medidas, las diferencias no son tan evidentes; en el caso de la modularidad y la centralidad del autovector, los sujetos de control ya no muestran diferencias con respecto a los pacientes con ME, y para el caso del grado de red, no se muestran diferencias entre ME y MC ( $p$ -valor  $< 0.05$ , test  $U$  de Mann Whitney con corrección de Bonferroni). Para todas las medidas de red, se aprecia que el grupo de controles parece establecerse como punto intermedio entre ambos subgrupos de migraña, proporcionando una diferenciación entre dichos subgrupos en su estructura electrofisiológica.

#### 4. Discusión

Las redes analizadas muestran que la estructura de relaciones entre niveles neurofisiológicos se ve modificada por la ME y la MC. Además, se puede observar que la pertenencia a un subtipo de migraña afecta de forma específica a la estructura neurofisiológica: la MC muestra un mayor nivel de segregación e integración (menor PL y mayor CIC) y la ME un menor nivel de ambas (mayor PL y menor CIC), en ambos casos respecto a los controles, lo que sugiere que la estructura neurofisiológica de la MC posee características de *Small World*. Estos cambios se deben principalmente al nivel de activación local, el cual sintetiza las relaciones entre los parámetros calculados a partir de la actividad generada por grupos neuronales localizados en ROIs individuales. Al analizar la estructura de estas redes mediante la teoría de grafos, podemos observar que los valores obtenidos concuerdan con las observaciones previamente mencionadas, apoyando las diferencias entre los dos subtipos de migraña. Además, la modularidad muestra que la red de la MC posee unos clústeres mejor definidos que en los otros grupos. Asimismo, la centralidad del autovector muestra una mayor centralidad en la ME, y menor en la MC, de nuevo mostrando al grupo de control como punto medio entre los dos estados de la migraña. Estudios previos indican que existe una mayor concentración de ácido N-acetil aspártato en los pacientes con MC en el tálamo y la corteza cingular anterior respecto a los controles, y en el tálamo izquierdo respecto a los pacientes con ME [13]. Por otro lado, se ha observado que los pacientes con ME sufren un hipometabolismo significativo en diversas regiones como la ínsula, la corteza cingular anterior y posterior, cortezas



**Figura 2.** Parámetros de red normalizados respecto a los controles. Los asteriscos representan diferencias entre controles y ME (negro); controles y MC (verde); y entre ME y MC (violeta).

prefrontal y premotora izquierda, y corteza somatosensorial primaria izquierda; todas ellas previamente asociadas al procesamiento central del dolor [14]. Estos cambios biológicos podrían afectar a cómo los osciladores neuronales generan su actividad a nivel local en la ME y en la MC, pudiendo modificar la estructura de relaciones neurofisiológicas.

Existen una serie de limitaciones en este estudio que han de ser destacadas. En primer lugar, las redes muestran los valores de correlación en valor absoluto, eliminando información sobre la tendencia de las correlaciones; sería interesante analizar por separado las correlaciones positivas y negativas. En segundo lugar, no se ha hecho una subdivisión en los pacientes con ME entre ME de alta frecuencia (*i.e.*, más de 10 días de dolor de cabeza al mes) y ME de baja frecuencia (*i.e.*, 10 o menos días de dolor de cabeza al mes); esto proporcionaría un matiz más a la evaluación, dotando al estudio de una mayor especificidad. Por último, no se han dividido a los pacientes en función del estado migrañoso en el que se encontraban (ictal, interictal, preictal o postictal); esto también permitiría analizar la reorganización de la estructura fisiológica en función de si se está sufriendo un ataque o el paciente está en una fase entre ataques.

## 5. Conclusiones

Como conclusión, se han observado diferencias en la estructura fisiológica de las pacientes con ME y MC en comparación con el grupo de control, debidas especialmente al nivel de activación local. Estos cambios parecen estar relacionados con diferentes alteraciones biológicas en regiones cerebrales individuales. Este estudio muestra no solo una diferenciación de la estructura neurofisiológica entre controles y la ME y MC, sino también entre ambos subgrupos de migraña.

## Agradecimientos

Esta investigación ha sido financiada por la Comisión Europea y el Fondo Europeo de Desarrollo Regional (FEDER) en el marco del proyecto “Análisis y correlación entre la epigenética y la actividad cerebral para evaluar el riesgo de migraña crónica y episódica en mujeres” (Programa de Cooperación Interreg V-A España- Portugal, POCTEP 2014–2020), por el ‘Ministerio de Ciencia e Innovación / Agencia Estatal de Investigación /10.13039/

501100011033/’ y FEDER ‘A way of making Europe’ bajo los proyectos PGC2018-098214-A-I00, PDC2021-120775-I00, PID2020-115468RB-I00 y RTC2019-007350-1, y por ‘CIBER en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN)’ a través del ‘Instituto de Salud Carlos III’. V. Gutiérrez-de Pablo y V. Rodríguez-González son beneficiarios de un contrato PIF-UVa de la Universidad de Valladolid.

## Referencias

- [1] V. Anttila *et al.*, *Genetics of migraine*, 1st ed., vol. 148. Elsevier B.V., 2018.
- [2] M. Björk *et al.*, “What initiates a migraine attack? Conclusions from four longitudinal studies of quantitative EEG and steady-state visual-evoked potentials in migraineurs” *Acta Neurol. Scand.*, vol. 124, no. SUPPL. 191, pp. 56–63, 2011, doi: 10.1111/j.1600-0404.2011.01545.x.
- [3] P. Gormley *et al.*, “Common Variant Burden Contributes to the Familial Aggregation of Migraine in 1,589 Families” *Neuron*, vol. 98, no. 4, pp. 743–753, May 2018, doi: 10.1016/j.neuron.2018.04.014.
- [4] J. L. Brandes, “The Influence of Estrogen on Migraine” *JAMA*, vol. 295, no. 15, p. 1824, Apr. 2006, doi: 10.1001/jama.295.15.1824.
- [5] Z. Katsarava *et al.*, “Defining the Differences Between Episodic Migraine and Chronic Migraine” *Curr. Pain Headache Rep.*, vol. 16, no. 1, pp. 86–92, Feb. 2012, doi: 10.1007/s11916-011-0233-z.
- [6] J. Olesen, “International Classification of Headache Disorders” *Lancet Neurol.*, vol. 17, no. 5, pp. 396–397, 2018, doi: 10.1016/S1474-4422(18)30085-1.
- [7] T. J. Schwedt *et al.*, “Functional MRI of migraine” *Lancet Neurol.*, vol. 14, no. 1, pp. 81–91, 2015, doi: 10.1016/S1474-4422(14)70193-0.
- [8] W. T. Chen *et al.*, “Peri-ictal normalization of visual cortex excitability in migraine: An MEG study” *Cephalalgia*, vol. 29, no. 11, pp. 1202–1211, 2009, doi: 10.1111/j.1468-2982.2009.01857.x.
- [9] J. Gomez-Pilar *et al.*, “Exploring EEG Spectral Patterns in Episodic and Chronic Migraine During the Interictal State: Determining Frequencies of Interest in the Resting State” *Pain Med.*, vol. 21, no. 12, pp. 3530–3538, 2020, doi: 10.1093/pm/pnaa117.
- [10] Z. Cao *et al.*, “Exploring resting-state EEG complexity before migraine attacks” *Cephalalgia*, vol. 38, no. 7, pp. 1296–1306, Jun. 2018, doi: 10.1177/0333102417733953.
- [11] Z. Cao *et al.*, “Resting-state EEG power and coherence vary between migraine phases,” *J. Headache Pain*, vol. 17, no. 1, 2016, doi: 10.1186/s10194-016-0697-7.
- [12] R. D. Pascual-Marqui, “Standardized low-resolution brain electromagnetic tomography (sLORETA): technical details” *Methods Find. Exp. Clin. Pharmacol.*, vol. 24 Suppl D, pp. 5–12, 2002, [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/12575463>.
- [13] D. M. Niddam *et al.*, “Neurochemical changes in the medial wall of the brain in chronic migraine” *Brain*, vol. 141, no. 2, pp. 377–390, 2018, doi: 10.1093/brain/awx331.
- [14] J. H. Kim *et al.*, “Interictal metabolic changes in episodic migraine: A voxel-based FDG-PET study” *Cephalalgia*, vol. 30, no. 1, pp. 53–61, 2010, doi: 10.1111/j.1468-2982.2009.01890.x.

# Glucose prediction based on deep learning models and objective physical activity data

C. Gallardo<sup>1,2</sup>, G. García-Sáez<sup>1,2</sup>, C. Pérez-Gandía<sup>1,2</sup>, Álvaro Gutiérrez<sup>1</sup>, M. Rigla<sup>3</sup>, M.E. Hernando<sup>1,2</sup>

<sup>1</sup> Departamento de Tecnología Fotónica y Bioingeniería, ETSI de Telecomunicación, Universidad Politécnica de Madrid, Madrid, Spain.

<sup>2</sup> Bioengineering and Telemedicine Group, Centro de Tecnología Biomédica, Universidad Politécnica de Madrid, Madrid, Spain; CIBER-BBN: Networking Research Centre for Bioengineering, Biomaterials and Nanomedicine, Madrid, Spain.

<sup>3</sup> Endocrinology and Nutrition Dpt., Parc Taulí Research and Innovation Institute Foundation (I3PT), Autonomous University of Barcelona, Sabadell, Spain

## Summary

*Type 1 diabetes is a chronic disease caused by a deficiency in the production of insulin by the pancreas. To control blood glucose levels, the administration of exogenous insulin is needed and patients have to manually fit the therapy to their diet and physical activity. Nowadays, artificial pancreas systems are an alternative treatment, but commercial devices do not consider physical activity as input to the control algorithms. In this paper, we explore the use of physical activity data and compare different glucose prediction models varying input data to select a good alternative to be integrated in an artificial pancreas algorithm. With a dataset of 666 samples from 12 type 1 diabetes patients provided by the Ohio university, more than 200 models have been trained using deep learning techniques and different input features like glucose levels, heart rate, insulin doses and carbohydrates intake have been mixed. Using data obtained with wearable monitors (objective), the results are better than using a bigger number of features, achieving an RMSE of 14.83 mg/dL. However, the best model was obtained using the biggest dataset of data obtained from test/patients estimations (non-objective) available with an RMSE of 14.48 mg/dL. The results obtained improve those achieved by other authors, but it would be necessary to have bigger datasets of objective physical activity data.*

## 1. Introduction

Type 1 diabetes is a chronic disease caused by a deficiency in the production of insulin by the pancreas, its most common characteristic is hyperglycemia, which means high values of glucose in the blood, and the treatment is focused on blood glucose control through exogenous insulin, diet and lifestyle to prevent complications. In 2017, the global available number of prevalent cases of type 1 diabetes were around 9 million people and its incidence rates are rising, being the cause of this growth unknown [1].

Artificial pancreas systems, that links glucose sensors and insulin pumps, are nowadays a real alternative for the treatment of type 1 diabetes. However, their benefits diminish during post-meal periods and associated to the practice of physical activity.

The presence of delays is one of the main problems in artificial pancreas systems, as they act in the subcutaneous tissue, measuring the glucose with a delay between 10 and 12 minutes, while administered insulin takes 30 minutes to take its effect. These delays penalize glucose control even more after meals or physical activity, making more

difficult anticipating to avoid an increase of normal blood glucose values [2].

To decide the insulin doses, patients use bolus calculators that almost have no advanced in the last 15 years. Bolus calculators do not keep in mind if the patient has performed physical activity before the last meal, so they are not able to counteract the extra caloric waste and the changes in insulin sensibility that physical activity produces and could affect few hours later. It is known that exercise enhances the action of insulin, achieving more effect with the same dose of insulin. Indeed, these systems rely on the patient's capacity for estimating the quantity of carbohydrates for each meal and the effect that insulin has in the body in each moment of the day.

For those patients with diabetes mellitus who are insulin-dependent, regular physical activity could help to improve metabolic control, decreasing glycemia concentrations and insulin needs [3]. However, hypoglycemia is a hazard after physical activity, and is one of the main barriers for patients to not be physically active. For this reason, it is fundamental considering some aspects like the duration and intensity of physical activity, the metabolic control of the patient, the insulin therapy pattern, and the diet [4].

The use of deep learning methodologies as a possible solution for glucose management has attracted increasing attention in the last years, including not only prediction of blood glucose levels, but also glycemic control, detection of harmful events like hyperglycaemia or hypoglycemia and even prediction of HbA1c [5].

It is necessary to control the possible appearance of hypoglycemia after physical activity practicing and a predictor that could help to keep track the impact of physical activity on glucose values could provide many benefits to patients with type 1 diabetes. For this reason, the main goal of this work is to improve the bolus calculators with deep learning models that help to predict glucose changes considering not only previous glucose data, but also physical activity information.

## 2. Methods

### 2.1. Dataset

The dataset used in this work includes clinical data from 12 patients with type 1 diabetes along a follow-up period of 8 weeks per patient. It was loaned by Ohio University, which was used in 2020 to set out a scientific challenge where

different researchers presented their best results obtained with the common dataset [6].

Patients were in insulin pump therapy with continuous glucose monitoring. They wore the Medtronic Enlite CGM sensor for continuous glucose measurement and insulin take [6]. The dataset contains glucose levels every 5 minutes, coming from the continuous glucose sensor, self-monitored capilar glucose levels periodically performed by the patient, insulin doses, carbohydrates estimation reported by the patient, sleep, stress, time estimation of the physical activity reported by the patient, physical activity data provided by smart bands and more information that is not useful for this work. Up to 6 of the 12 patients have heart rate records obtained from a smart band (the 2018 challenge dataset).

Our idea is to use objective physical activity data automatically registered by band devices, so data reported by the patient were discarded.

## 2.2. Preprocessing

First, data cleaning and filtering were carried out for a suitable input data quality. In the dataset, some gaps with missing data and outlier points can be found due to calibration errors or connection loss. To deal with missing data errors in glucose levels, those days with more than 40 missing points were discarded. Days with gaps of more than 18 missing points (one and a half hour of disconnection) were also discarded, as it is less accurate to make interpolation with long gaps and squared splines interpolation is applied. Not only in glucose levels but also in meal intake can be found some missing data. Only days with at least 3 different intakes were considered for data input.

After data cleaning, the dataset was reduced from the initial 666 days to the final 386 days with good quality data. A moving average filter with a sliding window of 5 samples was applied to smooth the signal and avoid peaks of data that may confuse the model. When the filter was applied, the preprocessing step finished, resulting with a calibrated and smoothed signal without gaps.

Meal intakes and insulin data were transformed from their impulse form, which represents the input moment, to a rectangle whose area is the value of the pulse. In this case, the length of the rectangle was 12 points, it means that the pulse value is divided by 12 and distributed in the samples corresponding with the next hour from the intake or insulin administration.

Finally, heart rate data were interpolated when a gap was found, but no filter was applied to smooth the heart rate signal.

## 2.3. Model choices

First decision to be made was the prediction horizon and the input length of the signal. A sliding window with 12 points was created for data ingestion (which means that an hour of previous data is necessary to make the predictions). A 30-minute prediction horizon is chosen to easier compare results to other related articles [5].

Once the decision was made, the sliding window was created, and standardization of the data was done. It was done due to the different ranges in the input data, because it could have affected in a different way to the model weights. The equation used for data standardization was a z-score. This process was applied not only to the training and validation data but also to testing data.

For the architecture of the model, LSTM layers were selected as hidden layers due to their good performance in time series forecasting. The number of hidden layers and nodes contained in them was variable and the activation function used in the LSTM layers was a hyperbolic function. LSTM layers were combined with a single full-connected layer with ReLU activation function, used as output layer that achieves the final regression output.

The loss function used for gradient descent was the RMSE and the optimizer used was Adam with  $\beta_1$  fixed to 0.9,  $\beta_2$  fixed to 0.999 and learning rate to 0.001.

Hyperparameters	Value range
Number of layers	{1, ..., 5}
Initial LSTM layer	{32, 64, 128, 256}
Middle LSTM layer/s	{32, 64, 128, 256}
Final LSTM layer	{32, 64, 128, 256}

*Table 1. Hyperparameters used for model training*

A grid search was done to select the suitable hyperparameters. The search was done searching between the hyperparameters showed in Table 1. Different approaches were followed in this hyperparameters search:

- Case 1: Use objective data, which means only data obtained through sensors (blood glucose from CGM and heart rate from smart band). The main problem of this approach is that only 193 days from the 386 preprocessed days contain this information, so the dataset size is reduced even more.
- Case 2: Use additional features, including non-objective data, like carbohydrates intakes, to compare how it could affect adding new features, with the same dataset size. Also, insulin doses data is used in this case for model training.
- Case 3: Use the three features that have been obtained in the same way for all the patients. Those features are glucose levels, insulin doses and carbohydrates intake. Models are trained using all the dataset and only the 193 samples used in previous cases to compare how the dataset size could affect to the results.

## 2.4. Evaluation

For the evaluation of the models, always the Ohio test dataset was used. This dataset was composed by more

samples of the different patients and were the same for all the participants. First, as done by K. Li *et al.* and J. Daniels *et al* [7-9], a baseline was set to compare our models with more basic algorithms. A support vector regression (SVR) algorithm was trained and tested obtaining a RMSE of 21.83 mg/dL with the 6 patients that contained objective physical activity information and 18.4 mg/dL using all the patients. SVR algorithm is a classic machine learning approach, which a priori should have worse results than our deep learning approach, for this reason those were considered the minimum scores that should be achieved.

With this approach, also Ohio challenge results could be used as baseline for the improvement of predictions. Test set was preselected by the Ohio challenge. Extrapolation for missing data was done and a filter using previous data was used.

### 2.5. Evaluation metrics

To test the performance of the models, the root mean squared error (RMSE), mean absolute error (MAE) and Pearson correlation were used as primary metrics to evaluate the accuracy. The RMSE is defined as:

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (x_i - \hat{x}_i)^2}{N}}, \quad (1)$$

where  $N$  is defined as the number of samples,  $x_i$  are the real values and  $\hat{x}_i$  are the predicted values. MAE equation is defined as:

$$MAE = \frac{\sum_{i=1}^N |\hat{x}_i - x_i|}{N}, \quad (2)$$

and finally, Pearson correlation is defined as:

$$PR = \frac{\sum_{i=1}^N (x_i - \mu) (\hat{x}_i - \hat{\mu})}{\sqrt{\sum_{i=1}^N (x_i - \mu)^2} \sqrt{\sum_{i=1}^N (\hat{x}_i - \hat{\mu})^2}}, \quad (3)$$

where  $x_i$  are the real values,  $\hat{x}_i$  are the predicted values,  $\mu$  is the mean of the real values and  $\hat{\mu}$  is the mean of the predicted values. The goal is to decrease as much as possible both RMSE and MAE, while the best score for Pearson correlation is 1. Also, Clarke Error Grid is calculated to evaluate the clinical accuracy [10]. It is a graph divided into 5 zones where each predicted value is labelled depending on the misestimation it has obtained:

- Zone A: Predictions in this zone lie within 20% of the real value when glucose levels are > 70 mg/dL and they are not more than 70 mg/dL when there is a hypoglycemic event
- Zone B: Predictions lie outside the 20% error, so the treatment will be incorrect but uncritical
- Zone C: Predictions in this region could result in unnecessary treatment
- Zone D: Predictions point to a harmful glycaemic event that has not been detected
- Zone E: If there is an action against those predicted points, it could lead to opposite corrective action being undertaken to treat a harmful glycaemic event

## 3. Results

Grid search with glucose levels from CGM and heart rate from smart band best models scores are represented in Table 2. The best model performance obtained a RMSE of 15.00 mg/dL. It already improves the baseline that had been set in 21.83 mg/dL using the 6 patients with objective physical activity data.

Layers	First Neurons	Middle Neurons	Last Neurons	RMSE (mg/dL)
5	64	32	32	15.00
3	256	32	32	15.01
5	64	256	64	15.12

**Table 2.** Case 1: Best scores for models trained with glucose and heart rate using Ohio 2018 dataset

Table 3 shows the Case 2 results, where insulin and carbohydrates intake features were added (a), and where only insulin was added (b). Despite being trained with more features, the best score is slightly worse when carbohydrates and insulin features are added. None of the metrics in the best model of Case 2 (a) have better scores than in the best model of Case 1. Thinking that carbohydrates intakes could be confusing the model, because it is an estimation done by the patient, a training without this feature was done, giving better results than all the previous models (see Table 3 (b)).

	Layers	First Neurons	Middle Neurons	Last Neurons	RMSE (mg/dL)	
<b>(a)</b>	<b>BG</b>	4	128	128	32	15.04
	<b>HR</b>	4	128	32	32	15.18
	<b>ID</b>	5	256	128	32	15.33
	<b>CI</b>					
<b>(b)</b>	<b>BG</b>	4	32	32	64	14.83
	<b>HR</b>	4	32	256	256	15.00
	<b>ID</b>	3	32	32	32	15.02

**Table 3.** Case 2: Three best scores for models trained with (a) blood glucose, heart rate, insulin doses and carbohydrates intake and with (b) blood glucose, heart rate and insulin doses using the patients with objective data

When training with the whole dataset (Case 3), a higher improvement in the score can be seen when the number of samples is increased, obtaining 0.52 mg/dL better RMSE than in Case 1 (see Table 4). To check if the model had better scores due to the different features or to the new amount of data used, also some models were trained with the set of data with objective physical activity.

Best scores obtained are being summarized in Table 5. The main baselines set with the SVR algorithm were successfully surpassed both in the one obtained with the 6 patients with objective data and in one trained with the whole Ohio dataset. Zone E (the worst possibility) and zone C have not been included because none of the models had points in those zones. So, the remaining percentage is for points in zone B.

Ohio dataset	Layers	First Neurons	Middle Neurons	Last Neurons	RMSE (mg/dL)
(a) 2018	1	64	0	0	14.48
	1	32	0	0	14.57
	2	128	0	256	14.61
(b) 2018	4	128	32	32	15.12
	5	64	128	64	15.20
	4	256	256	128	15.27

**Table 4.** Three best scores for models trained with glucose, insulin doses and carbohydrates intake using the (a) whole dataset and the (b) set of data with only objective data

Checking all the results we can see how the best scores come from the model trained with the biggest dataset. Also, it is interesting that the second best RMSE is obtained by the model trained with blood glucose levels, insulin doses and heart rate. Adding carbohydrates intake as feature confuse the models and it might be due to uncertainty in carbohydrates estimations done by the patients. With those results we can conclude that also using objective data could be important to improve blood glucose prediction models.

Case	RMSE (mg/dL)	MAE (mg/dL)	Pearson Score	Zone A (%CGE)	Zone D (%CGE)
1	15.00	10.29	96.98%	94.22%	0.76%
2 (a)	15.04	10.47	96.97%	94.30%	0.87%
2 (b)	14.83	10.40	97.04%	93.76%	0.93%
3 (a)	14.48	9.92	96.79%	94.77%	0.93%
3 (b)	15.12	10.50	96.95%	94.16%	1.01%
SVR (2018)	21.83	14.17	93.63%	88.92%	1.92%
SVR	18.40	13.14	94.67%	89.71%	0.83%

**Table 5.** Summary of the best scores in all the cases and comparison with SVR baseline values

The best challenge result was obtained by the J. Freiburghaus *et al.* network [11], achieving an RMSE of 17.45 mg/dL. Our best model improves the predictor performance to an RMSE of 14.48 mg/dL, using the same features (glucose levels, carbohydrates intakes and insulin) and same test dataset.

#### 4. Conclusion

Type 1 diabetes needs special treatment to control glucose values and to avoid hyperglycemia and hypoglycemia that can be caused by different facts, such as the unmatching between carbohydrates intake and insulin administration, emotional states, or physical activity.

This paper has shown that using objective data we could achieve better results than using several features obtained

by patients' estimations. Our models improve the results obtained in the 2020 Ohio challenge, but objective data models have not achieved the best results, probably due to the limited number of samples. It might be very useful, and it could be used as baseline for insulin estimation doses in artificial pancreas systems, helping to avoid harmful events in type 1 diabetes patients.

However, the problem is still the small quantity of objective data available. For the improvement of deep learning models, it is always necessary a big amount of data, as it has been also shown in this paper. On the Internet of Things era, it would be feasible to increment the amount of objective data available in future works for training and testing.

#### Acknowledgement

This work has been partly supported by the Spanish FIS grant from the Ministry of Health and Consumer Affairs "PROPHECY"- FIS PI21/01028, co-funded by FEDER.

#### References

- [1] Green, A., Hede, S.M., Patterson, C.C. *et al.* "Type 1 diabetes in 2017: global estimates of incident and prevalent cases in children and adults". *Diabetologia* **64**, 2741–2750 (2021).
- [2] P. Senior, I. Hramiak, "Fast-Acting Insulin Aspart and the Need for New Mealtime Insulin Analogues in Adults With Type 1 and Type 2 Diabetes: A Canadian Perspective", *Can J. Diabetes* **43**, 515-523 (2019).
- [3] "A. Kennedy, K. Nirantharakumar, M. Chimen, TT. Pang, K. Hemming, RC. Andrews, P. Narendran, "Does exercise improve glycaemic control in type 1 diabetes? A systematic review and meta-analysis", *PLoS One* (2013)
- [4] M.C. Riddell, I.W. Gallen, C.E. Smart *et al.* "Exercise management in type 1 diabetes: a consensus statement", *Lancet Diabetes Endocrinol*, 377-390 (2017).
- [5] T. Zhu, K. Li, P. Herrero, P. Georgiou, "Deep Learning for Diabetes: A Systematic Review", *IEEE Journal of Biomedical and Health Informatics*, 2744-2757 (2021)
- [6] C. Marling, R. Bunescu, "The OhioT1DM Dataset for Blood Glucose Level Prediction: Update 2020", 5<sup>th</sup> International Workshop on Knowledge Discovery in Healthcare Data (KDH), 77-81 (2020).
- [7] K. Li, C. Liu, T. Zhu, P. Herrero, and P. Georgiou, "Glunet: A deep learning framework for accurate glucose forecasting," *IEEE journal of biomedical and health informatics* (2019).
- [8] J. Daniels, P. Herrero, and P. Georgiou, "A multitask approach to personalized blood glucose prediction", *IEEE journal of biomedical and health informatics* (2022).
- [9] K. Li, J. Daniels, C. Liu, P. Herrero, and P. Georgiou, "Convolutional recurrent neural networks for glucose prediction" *IEEE journal of biomedical and health informatics* (2019).
- [10] W. L. Clarke, "The Original Clarke Error Grid Analysis (EGA)", *Diabetes technology and therapeutics* (2005).
- [11] J. Freiburghaus, A. Rizzotti-Kaddouri, F. Albertetti, "A Deep Learning Approach for Blood Glucose Prediction of Type 1 Diabetes", 5<sup>th</sup> International Workshop on Knowledge Discovery in Healthcare Data (KDH), 138-142 (2020).

# Sistemas de Ayuda a la Decisión

---

Viernes, 25 de noviembre de 2022

# Modelo de Deep Learning basado en redes convolucionales 1D para el diagnóstico de la apnea del sueño mediante señales respiratorias

M. Fernández-Poyatos<sup>1</sup>, G.C. Gutiérrez-Tobal<sup>1,2</sup>, D. Álvarez González<sup>1,2</sup>, D. Gozal<sup>3</sup>, R. Hornero<sup>1,2</sup>, B. Sahelices<sup>4</sup>

<sup>1</sup> Grupo de Ingeniería Biomédica, Universidad de Valladolid, Valladolid, España, marta.fernandez.poyatos@estudiantes.uva.es

<sup>2</sup> Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina, (CIBER-BBN), España, gonzalo.gutierrez@ciber-bbn.es

<sup>3</sup> Department of Child Health, and the Child Health Research Institute, The University of Missouri School of Medicine, Columbia, MO, USA

<sup>4</sup> Grupo de caracterización de materiales y dispositivos electrónicos, Universidad de Valladolid, Valladolid, España, benja@infor.uva.es

## Resumen

*La apnea obstructiva del sueño (AOS) es una enfermedad con una alta prevalencia en la población adulta. La prueba diagnóstica estándar es la polisomnografía nocturna (PSG), en la que se recogen hasta 32 señales fisiológicas del paciente. Sin embargo, es una prueba compleja y con un alto coste, por lo que resulta necesario encontrar alternativas diagnósticas más sencillas. Este estudio se centra en el uso de una arquitectura deep learning de redes convolucionales 1D para la predicción del grado de severidad de la apnea en general, y de la contribución de los episodios centrales en particular. Para ello se han empleado únicamente las dos señales de esfuerzo respiratorio: torácico y abdominal. El modelo con mayor rendimiento diagnóstico alcanzó un índice kappa de 4 clases de 0.544 (precisión del 92.3%) para la predicción del índice de apnea-hipopnea (IAH) con origen central. Los resultados sugieren que la arquitectura propuesta podría resultar de utilidad en la simplificación del diagnóstico y gestión de los enfermos de AOS.*

## 1. Introducción

La apnea obstructiva del sueño (AOS) es un trastorno caracterizado por episodios de ausencia completa (apnea) o parcial (hipopnea) del flujo aéreo en las vías respiratorias durante el sueño, con una duración mínima de 10 segundos [1], ocasionando un sueño fragmentado y no reparador. Está relacionado con múltiples patologías de tipo cardiovascular, pudiendo aumentar la mortalidad en pacientes con enfermedades pulmonares, producir complicaciones en procesos postoperatorios [2], somnolencia diurna [3] o aumentar el riesgo de aparición de diabetes de tipo 2 [3]. La prevalencia de la apnea llega a afectar hasta al 38% de la población adulta [4], aunque es una enfermedad infradiagnosticada. A pesar de la recomendación actual de denominar a la enfermedad como AOS, esta incluye no solo eventos obstructivos, causados por el colapso de las vías respiratorias superiores [2], sino también centrales, en los que el cerebro no envía la señal necesaria a los músculos que controlan la respiración. Para confirmar la presencia de AOS, los eventos apnéicos deben ser principalmente obstructivos (>50%), mientras que cuando predomina el carácter central nos encontramos ante la denominada apnea central del sueño [1].

La técnica estándar utilizada para el diagnóstico de la apnea es la polisomnografía (PSG) nocturna, en la que se recogen hasta 32 señales fisiológicas diferentes, entre ellas: electroencefalograma (EEG), electrocardiograma (ECG), saturación de oxígeno, esfuerzo respiratorio torácico y abdominal [5]. A partir de estas señales, se contabiliza el número de eventos apnéicos por hora de sueño o índice de apnea-hipopnea (IAH). A partir del IAH se establecen los diferentes grados de severidad de la enfermedad: no apnea ( $IAH < 5$  e/h), leve ( $5 \text{ e/h} \leq IAH < 15$  e/h), moderada ( $15 \text{ e/h} \leq IAH < 30$  e/h) y grave ( $IAH \geq 30$  e/h) [2].

La PSG, a pesar de su efectividad, presenta claras desventajas: requiere que el paciente pase una noche completa en el hospital, tiene un coste muy elevado y requiere la presencia de un especialista durante todo el estudio [5]. Añadido a la alta prevalencia de la enfermedad, las unidades de sueño no son capaces de abordar todas las pruebas necesarias, derivando en largas listas de espera y retrasos en el diagnóstico y en el tratamiento. La necesidad de que las señales de las PSGs sean analizadas manualmente por médicos requiere una gran cantidad de tiempo, además de derivar en diagnósticos subjetivos. Por este motivo, la comunidad científica busca alternativas diagnósticas automáticas que simplifiquen la detección de la AOS. En esta búsqueda, surge la aplicación de algoritmos de *deep learning*, que permiten descubrir automáticamente patrones y características en los datos, evitando procedimientos costosos de preprocesado.

Trabajos anteriores han empleado las señales de esfuerzo respiratorio torácico y abdominal aisladamente [6] o junto con la señal de flujo de aire nasal [7] para la clasificación de tipos de eventos apnéicos. Nuestro grupo de investigación ha desarrollado modelos para la estimación del IAH a partir de otras señales, fundamentalmente la saturación de oxígeno en sangre [8, 9]. En el presente estudio se propone el uso de una arquitectura de redes convolucionales unidimensionales (1D) para la predicción del número de eventos apnéicos empleando fragmentos de 20 minutos de registro de sueño, de los que en conjunto se extrae la predicción del IAH de cada individuo. Para ello, se emplean tan solo dos señales que recogen los

movimientos respiratorios del paciente: esfuerzo respiratorio torácico y abdominal. El objetivo del estudio consiste en demostrar que estas señales son suficientes para la estimación del IAH en general y de la contribución de los eventos centrales en particular.

## 2. Sujetos y señales

Este estudio ha analizado las PSGs de la base de datos del *Sleep Heart Health Study* (SHHS) [10], accesibles a través del *National Sleep Research Resource* (NSRR). En dicho estudio participaron, en una primera etapa (SHHS1), 6441 sujetos (5722 estudios válidos), tanto hombres como mujeres de más de 40 años, desde noviembre de 1995 hasta enero de 1998. En una etapa posterior (SHHS2), 2651 de estos sujetos (2535 estudios válidos) se sometieron a una segunda PSG entre enero de 2001 y junio de 2003. Todas las PSGs fueron realizadas en el domicilio de los pacientes.

De todas las señales recogidas en las PSGs, se ha decidido estudiar únicamente dos, relacionadas con la respiración del sujeto: esfuerzo respiratorio torácico y abdominal, recogidas mediante bandas de pletismografía inductiva [11]. Ante la presencia de registros a diferentes frecuencias de muestreo, se decidió incluir en el estudio solamente aquellas PSGs en las que las señales de esfuerzo respiratorio se adquirieron a una tasa de 10 Hz. Todas las señales fueron filtradas para eliminar posibles artefactos. Se empleó un filtro paso bajo entre 0 y 1.5 Hz con ventana de Hamming y se segmentó cada señal en fragmentos de 20 minutos [9]. Cada fragmento de 20 minutos se etiquetó 4 veces: i) con el número de apneas obstructivas, ii) con el número de apneas centrales, iii) con el número de hipopneas, y iv) con el número total de eventos apnéicos, es decir, la suma de los 3 tipos anteriores.

Los sujetos fueron divididos en tres grupos: entrenamiento, validación y test, como se describe en la figura 1. El motivo de esta distribución de la población se encuentra en los individuos que participaron en ambas etapas del estudio: el uso de las señales de estos individuos en diferentes conjuntos de datos provocaría resultados sesgados. Por lo tanto, las señales de los pacientes que contribuyeron con PSGs en las dos fases del estudio se utilizaron como conjunto de test, mientras que las señales restantes se repartieron entre entrenamiento (80%) y validación (20%). Se aplicó *data augmentation* para aumentar el número de instancias en los grupos de entrenamiento y validación generando segmentos de señal con un solapamiento del 50%.

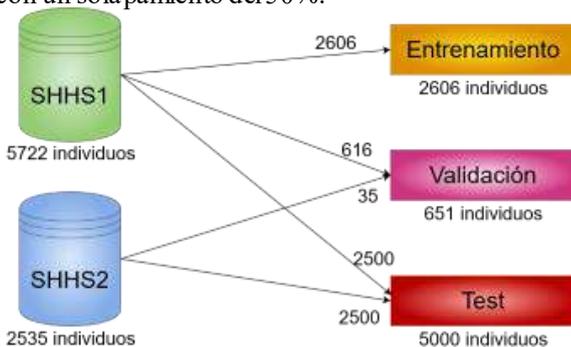


Figura 1. Distribución de los datos en entrenamiento, validación y test

## 3. Métodos

El objetivo de este estudio es la predicción mediante regresión del número de eventos apnéicos presentes en intervalos de 20 minutos, a partir de las dos señales recogidas por las bandas torácica y abdominal. Estas dos señales son complementarias ya que los pacientes pueden realizar el esfuerzo respiratorio de forma torácica o de forma abdominal, y por ello deben estudiarse conjuntamente. La estrategia seguida para ello ha sido tratar estas dos señales como una única dimensión, pero distribuidas en dos canales diferentes. En la figura 2 se muestra un esquema de esta distribución.

### 3.1. Arquitectura CNN

La arquitectura de redes convolucionales de este estudio está inspirada en el trabajo de Vaquerizo-Villar *et al.* [9] para la estimación del número de eventos apnéicos obstructivos en sujetos pediátricos mediante una señal de oximetría. Con esta arquitectura se trata de predecir el número de eventos apnéicos en segmentos de 20 minutos mediante un problema de regresión. En la figura 3 se muestra la arquitectura empleada, con ocho bloques secuenciales, cada uno de ellos formado por una capa convolucional, una normalización del *batch*, que normaliza los mapas de características [12], una capa ReLU, que aplica una función de activación [12], una capa de *pooling* máximo, que reduce la dimensión de los datos [12], y, por último, una etapa de *dropout* que elimina conexiones entre neuronas [12]. La entrada de esta red neuronal es el conjunto de fragmentos de 20 minutos con las señales de esfuerzo respiratorio torácico y abdominal. Una vez terminado el paso de los datos por los ocho bloques, se aplanan para terminar con un perceptrón simple. El resultado de este perceptrón es la predicción del número de eventos para cada fragmento de 20 minutos. A partir de estas predicciones, se calcula el IAH de cada individuo, medido en eventos por hora. En la búsqueda del mejor modelo, se ha utilizado como métrica el índice *kappa de Cohen*, discretizando el IAH según los grupos de severidad comúnmente empleados en la práctica clínica. El índice *kappa* permite evaluar el rendimiento de los modelos ajustados y busca el modelo con la mejor concordancia (mayor *kappa*) entre los valores estimados y los reales en el conjunto de validación.

La implementación de la arquitectura presentada se llevó a cabo con la biblioteca de aprendizaje automático PyTorch. El criterio de optimización utilizado fue la función de pérdida de Hubber, aplicando el optimizador Adam [13]. La tasa de aprendizaje utilizada fue de  $10^{-3}$ , el tamaño de *batch* elegido fue 32 y el número de épocas 200 [13].

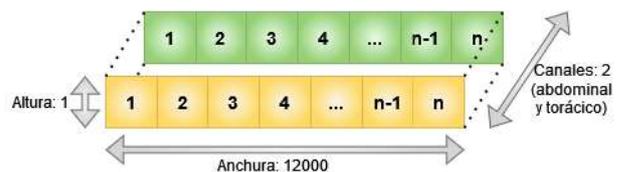


Figura 2. Distribución en canales de las señales de esfuerzo respiratorio torácico y abdominal

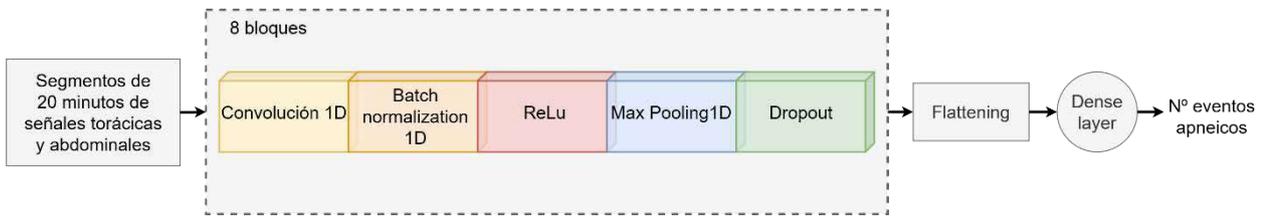


Figura 3. Arquitectura del modelo

Se decidió entrenar cuatro modelos. El primero ( $CNN-IAH_{GLOBAL}$ ) calculó el IAH para todos los tipos de eventos apneicos, siguiendo la discretización estándar en 4 clases. El segundo ( $CNN-IAH_{CENTRAL-4C}$ ), tan solo utilizó los eventos centrales para el cálculo del IAH y la misma discretización que el primer modelo. El tercero ( $CNN-IAH_{CENTRAL-3C}$ ), al no existir individuos con más de 30 e/h centrales, empleó únicamente las tres primeras categorías para el cálculo del IAH de los eventos centrales. En el cuarto modelo ( $CNN-IAH_{CENTRAL-4C_{NEW}}$ ), al existir un fuerte desbalanceo en las clases, se propuso una nueva categorización:  $IAH < 1$  e/h,  $1 \text{ e/h} \leq IAH < 5$  e/h,  $5 \text{ e/h} \leq IAH < 15$  e/h,  $IAH \geq 15$  e/h.

Para cada modelo se calculó la precisión global (Acc), índice kappa, sensibilidad (S), especificidad (E), valor predictivo positivo (VPP) y valor predictivo negativo (VPN), considerándose como falso positivo un sujeto cuyo IAH es superior al umbral de cada clase pero se clasifica como inferior y como falso negativo, el caso inverso.

## 4. Resultados

### 4.1. Optimización de hiperparámetros

En la búsqueda del modelo que proporcione unos mejores resultados, se ha estudiado un amplio espacio de búsqueda, utilizando arquitecturas con diferente número de capas, tamaño del *kernel*, número de canales de la última capa convolucional y diferente *dropout*. Tras un periodo completo de experimentación se pudo determinar la mejor combinación de hiperparámetros: 8 capas, tamaño de *kernel* de 32, 128 canales y un *dropout* de 0.3.

### 4.2. Evaluación del modelo óptimo

En la tabla 1 se muestran los rendimientos diagnósticos (Acc) de los modelos propuestos para la predicción del IAH de cada sujeto en el conjunto de test: para todos los tipos de eventos (obstructivos y centrales), eventos centrales, eventos centrales con 3 clases y eventos centrales con las 4 nuevas clases. El primero clasificó correctamente un 59.8% de casos, mientras que el segundo y el tercero un

	$CNN-IAH_{GLOBAL}$			$CNN-IAH_{CENTRAL-4C}$			$CNN-IAH_{CENTRAL-3C}$		$CNN-IAH_{CENTRAL-4C_{NEW}}$		
	5 e/h	15 e/h	30 e/h	5 e/h	15 e/h	30 e/h	5 e/h	15 e/h	1 e/h	5 e/h	10 e/h
S	0.212	0.675	0.882	0.997	0.999	-	0.996	0.999	0.971	0.993	0.999
E	0.986	0.871	0.663	0.5	0.318	-	0.531	0.455	0.632	0.594	0.5
VPP	0.445	0.750	0.877	0.990	0.997	-	0.991	0.998	0.962	0.992	0.998
VPN	0.960	0.824	0.675	0.75	0.7	-	0.708	0.667	0.694	0.633	0.688
Kappa	0.411			0.493			0.511		0.544		
Acc	59.8%			98.4%			98.4%		92.3%		

Tabla 1. Rendimiento sobre la población de test de los modelos propuestos para la predicción del IAH:  $CNN-IAH_{GLOBAL}$ ,  $CNN-IAH_{CENTRAL-4C}$ ,  $CNN-IAH_{CENTRAL-3C}$  y  $CNN-IAH_{CENTRAL-4C_{NEW}}$

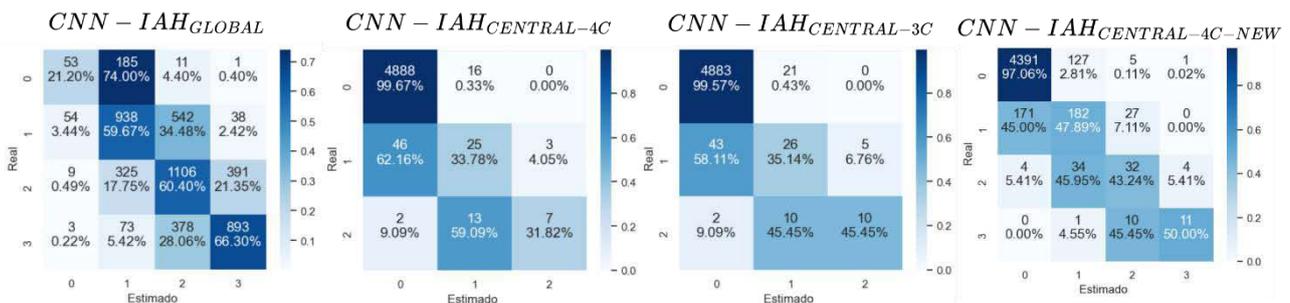


Figura 4. Matrices de confusión sobre test para la predicción del IAH para todos los tipos de eventos apneicos, eventos centrales con 4 clases, con 3 clases y con 4 nuevas clases. En los tres primeros: 0: no apnea ( $IAH < 5$  e/h), 1: leve ( $5 \text{ e/h} \leq IAH < 15$  e/h), 2: moderada ( $15 \text{ e/h} \leq IAH < 30$  e/h), 3: grave ( $IAH \geq 30$  e/h). En el último: 0:  $IAH < 1$  e/h, 1:  $1 \text{ e/h} \leq IAH < 5$  e/h, 2:  $5 \text{ e/h} \leq IAH < 15$  e/h, 3:  $IAH \geq 15$  e/h.

98.4% y el cuarto, un 92.3%.

La figura 4 muestra las matrices de confusión sobre el conjunto de test de los modelos entrenados para la predicción del IAH. En el segundo modelo, al no existir individuos con más de 30 eventos centrales, su matriz de confusión tiene una dimensión menos y no existen datos para el umbral de mayor gravedad, produciéndose una infra estimación del IAH que da lugar a valores más pequeños del índice  $kappa$ . El tercer modelo soluciona este problema. Finalmente, el modelo con la nueva discretización alcanza un índice  $kappa$  superior al resto.

## Discusión y conclusiones

En este estudio se ha evaluado la capacidad predictiva de una arquitectura *deep learning* de redes convolucionales 1D, aplicada a la predicción de la severidad del IAH en adultos, analizando de forma particular el desempeño de los modelos cuando los eventos apneicos son de tipo central, lo que se adecua al uso de las señales de esfuerzo respiratorio. El modelo diseñado para estimar la componente central del IAH usando la nueva discretización propuesta para las apneas centrales alcanzó un alto rendimiento diagnóstico, superior a la estimación tradicional del IAH de forma conjunta (obstructivos + centrales) y al empleo de la clasificación habitual de los grados de severidad del IAH.

Son escasos los estudios que analizan las dos señales de esfuerzo respiratorio exclusivamente [6] o junto con otras señales (flujo de aire nasal [7]). Estos proponen modelos SVM y redes convolucionales para la clasificación de segmentos en: respiración normal, hipopnea y obstructivos, con una precisión del 81.8% y 83.5%, respectivamente. Trabajos previos de nuestro grupo sí se han centrado en el desarrollo de modelos de *deep learning* para la predicción del IAH mediante las señales de flujo aéreo y oximetría, con una precisión del 72.55% ( $kappa$  del 0.6011) y 61.79% ( $kappa$  del 0.4469) en diferentes bases de datos pediátricas [8]. También mediante la señal de oximetría exclusivamente, con índices  $kappa$  de 0.515, 0.422 y 0.423 en diferentes poblaciones de estudio dentro del contexto de la AOS infantil [9].

Este estudio presenta algunas limitaciones. En primer lugar, todos los individuos que participaron en el SHHS eran mayores de 40 años, por lo que los resultados obtenidos deberían estudiarse en otros rangos de edad, principalmente en sujetos pediátricos, en los que la prevalencia de la apnea del sueño también es alta. En segundo lugar, los grupos de severidad de la apnea están desbalanceados, haciendo que las predicciones estén sesgadas hacia las clases mayoritarias. Esto se ve claramente reflejado en una infra estimación del IAH central que causa especificidades especialmente bajas en el modelo CNN-IAH<sub>CENTRAL-4C</sub>. Sería conveniente aumentar el número de instancias de las clases minoritarias, en busca de un mayor balanceo de los datos.

Para concluir, el rendimiento diagnóstico alcanzado en este estudio para la predicción de la severidad de la apnea, especialmente con eventos centrales, sugiere que el modelo de redes convolucionales 1D puede ser utilizado para

mejorar la capacidad diagnóstica de las señales de esfuerzo respiratorio, contribuyendo a simplificar la prueba diagnóstica y a facilitar la gestión de la enfermedad.

## Agradecimientos

Esta investigación se ha desarrollado al amparo de las ayudas PID2020-115468RB-I00 y PDC2021-120775-I00 financiadas por el Ministerio de Ciencia e Innovación/ Agencia Estatal de Investigación/ 10.13039/501100011033/ y ERDF A way of making Europe; y por el CIBER en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN) a través del Instituto de Salud Carlos III cofinanciado con fondos ERDF.

## Referencias

- [1] R. Berry *et al.*, “Rules for Scoring Respiratory Events in Sleep: Update of the 2007 AASM Manual for the Scoring of Sleep and Associated Events”, *Journal Of Clinical Sleep Medicine*, vol. 8, no. 5, pp. 597-619, 2012.
- [2] J. Muñoz-Lombo *et al.*, “Frequency and Factors Associated with Obstructive Sleep Apnea in Adults”, *RAMR*, vol. 20, no. 4, pp. 381-391, 2020.
- [3] A.N. Vgontzas, *et al.*, “Sleep Apnea and Daytime Sleepiness and Fatigue: Relation to Visceral Obesity, Insulin Resistance and Hypercytokinemia”, *The J. of Clinical Endocrinology & Metabolism*, vol. 85, no. 3, pp. 1151–1158, 2000.
- [4] C. Senaratna, *et al.*, “Prevalence of obstructive sleep apnea in the general population: A systematic review”, *Sleep Medicine Reviews*, vol. 34, pp. 70-81, 2017.
- [5] P. De Chazal, *et al.*, “Automated Detection of Obstructive Sleep Apnoea at Different Time Scales using the Electrocardiogram”, *Physiological Measurement*, vol. 25, no. 4, pp. 967–983, 2004.
- [6] Y.-Y. Lin, *et al.*, “Sleep apnea detection based on thoracic and abdominal movement signals of wearable piezoelectric bands,” *IEEE Journal of Biomedical and Health Informatics*, vol. 21, no. 6, pp. 1533–1545, 2017.
- [7] R. Haidar, *et al.*, “Convolutional Neural Networks on Multiple Respiratory Channels to Detect Hypopnea and Obstructive Apnea Events”, *IJCNN*, pp. 1-7, 2018.
- [8] J. Jiménez-García, *et al.*, “A 2D convolutional neural network to detect sleep apnea in children using airflow and oximetry”, *Computers in Biology and Medicine*, vol. 147, pp. 105784, 2022.
- [9] F. Vaquerizo-Villar, *et al.*, “A Convolutional Neural Network Architecture to Enhance Oximetry Ability to Diagnose Pediatric Obstructive Sleep Apnea”, *IEEE Journal Of Biomedical And Health Informatics*, vol. 25, no. 8, pp. 2906-2916, 2021.
- [10] G.-Q. Zhang, *et al.*, “The national sleep research resource: towards a sleep data commons”, *Journal of the American Medical Informatics Association*, vol. 25, no. 10, pp. 1351-1358, 2018.
- [11] S.F. Quan, *et al.*, “The Sleep Heart Health Study: design, rationale, and methods”, *Sleep*, vol. 20, no. 12, pp. 1077-85, 1997.
- [12] I. Goodfellow *et al.*, *Deep Learning*. MIT Press, 2016.
- [13] S. S. Mostafa *et al.*, “A systematic review of detecting sleep apnea using deep learning”, *Sensors*, vol. 19, no. 22, pp. 1–26, 2019.

# Clasificación explicada de las lesiones pigmentadas mediante la detección de los patrones dermatoscópicos de carcinoma

Carmen Serrano<sup>1</sup>, Manuel Lazo<sup>1</sup>, Amalia Serrano<sup>2</sup>, Tomás Toledo-Pastrana<sup>3</sup>, Rubén Barros-Tornay<sup>4</sup>, José A. Pérez-Carrasco<sup>1</sup> and Begoña Acha<sup>1</sup>

1 Dpto. Teoría de la Señal y Comunicaciones, Universidad de Sevilla, Camino de los Descubrimientos s/n, 41092 Seville, Spain; manu1235@gmail.com (M.L.); bacha@us.es (B.A.)

2 Hospital Universitario Virgen Macarena, Calle Dr. Fedriani, 3, 41009 Seville, Spain; amaliaserranog@gmail.com (A.S.); ttoledop@gmail.com (T.T.); rubbartor@gmail.com (R.B.)

3 Hospitales Quironsalud Infanta Luisa y Sagrado Corazón. Calle San Jacinto, 87. 41010 Seville Spain. ttoledop@gmail.com (T.T.); Hospital de Alta Resolución de Écija, Avda. Dr. Sánchez Malo 73, 41400 Écija, Seville, Spain

4 Hospital Universitario de Cabueñes. Los Prados, 395 33394 Gijón, Spain. rubbartor@gmail.com (R.B.)

## Resumen

*Antecedentes y objetivo.* El cáncer de piel es el más frecuente en todo el mundo. Uno de los tumores no melanoma más comunes es el carcinoma de células basales (BCC) que representa el 75% de todos los cánceres de piel. En este trabajo se presenta un nuevo método para identificar los diferentes patrones dermatoscópicos del BCC presentes en una lesión cutánea. Además, esta información se aplica para clasificar las lesiones cutáneas en BCC y no BCC.

*Métodos.* El método propuesto combina la información proporcionada por la imagen dermatoscópica original, introducida en una Red Neural Convolutiva (CNN), con características extraídas del análisis de color y textura de la imagen, combinando las diferentes entradas mediante una red de perceptrón multicapa (MLP).

*Resultados.* Los resultados muestran que el clasificador con características adicionales de color y textura supera a una CNN cuya entrada es sólo la imagen original. Otro hallazgo importante es que una nueva matriz de coocurrencia de color, propuesta en este trabajo, mejora los resultados obtenidos con otras medidas de textura. Finalmente, se consigue una sensibilidad de 0,99, una especificidad de 0,94 y una precisión de 0,97 cuando se clasifican las lesiones en BCC o no BCC.

*Conclusiones.* Hasta donde sabemos, es la primera vez que se propone una metodología para detectar todos los posibles patrones que pueden estar presentes en una lesión de BCC. Esta detección conduce a una clasificación clínicamente explicada. En este sentido, la clasificación de la herramienta propuesta se basa en la detección de las características dermatoscópicas que los dermatólogos emplean para su diagnóstico.

## 1. Motivación

El cáncer de piel es el más frecuente en todo el mundo [1]. Hay dos tipos principales de cáncer de piel: melanoma y no melanoma. Los tumores no melanoma más comunes son el carcinoma de células basales (BCC) y el carcinoma de células escamosas (SCC). El BCC representa el 75% de todos los cánceres de piel y es el tumor maligno más frecuente en la población blanca [2]. La detección de este cáncer se ejecuta mediante la inspección visual de un dermatólogo experto, pero hay muchas lesiones benignas

que pueden confundirse con este tipo de cánceres, lo que lleva a realizar biopsias innecesarias en una proporción de 5 biopsias frente a un caso real de cáncer [3].

La principal limitación de las herramientas de ayuda al diagnóstico por computador (CAD) basadas en el aprendizaje profundo (DL) es la falta de explicabilidad del resultado de la clasificación. Las redes neuronales profundas se consideran cajas negras, que dan una etiqueta para cada clase, pero sin revelar las decisiones internas que se han tomado para llegar a esa etiqueta. Últimamente, diferentes esfuerzos se han dirigido al desarrollo de herramientas de CAD para una clasificación clínicamente explicable del cáncer de piel [4-7]. En este sentido, en este trabajo se presenta una herramienta CAD que proporciona una detección explicable del BCC

Los dermatólogos diagnostican el BCC a partir de imágenes dermatoscópicas mediante la detección de diferentes características de alto nivel o criterios dermatoscópicos. El BCC tiene los criterios clínicos más claramente definidos [8, 2]. Los criterios dermatoscópicos para el BCC son vasos ramificados y lineales (arborización y telangiectasia superficial), glóbulos múltiples, ulceración, glóbulos gris azulados de tamaño variable (nidus ovoides), líneas radiales conectadas a una base común (como hoja de arce), líneas radiales que convergen a un punto central (áreas en forma de rueda de carro), estructuras concéntricas y ausencia de red pigmentada [9,2].

En los últimos años, algunos trabajos se han centrado en la detección de BCC en imágenes dermatoscópicas. Pero, hasta donde sabemos, hay un número muy limitado de trabajos dedicados a detectar algunas de las características dermatoscópicas que los dermatólogos emplean para diagnosticar el BCC y ninguno de estos trabajos se centra en la detección de todas las características dermatoscópicas del BCC.

El objetivo de este trabajo es, por una parte, la detección de patrones dermatoscópicos (patrones clínicos) en imágenes

dermatoscópicas; y por otra parte, según los patrones encontrados, la clasificación como BCC o no BCC.

## 2. Material y métodos

### 2.1. Base de datos

Como no existían bases de datos públicas que contuvieran imágenes con los criterios dermatoscópicos de BCC segmentados, para este trabajo ha habido que elaborar una base de datos.

Con el fin de recopilar una base de datos compuesta por imágenes que representen estructuras dermatoscópicas de BCC, hemos desarrollado un programa de fácil manejo que se ha instalado en ordenadores ubicados en el Hospital Universitario Virgen Macarena, Sevilla, España. Utilizando este software, los expertos han podido segmentar manualmente imágenes de 256×256 píxeles, donde se presentan uno o varios patrones. Dado que se trata de un trabajo tedioso, una herramienta informática de este tipo resulta útil y supone un gran ahorro de tiempo. Las imágenes se recogieron durante 9 meses aproximadamente. Todas las imágenes de BCC utilizadas en la evaluación han sido extirpadas y sometidas a biopsia. El intervalo de tiempo aproximado entre la adquisición de la imagen dermatoscópica y el diagnóstico del dermatólogo es inferior a 10 días. La biopsia se realiza en menos de 90 días desde la adquisición de la imagen. El criterio de elegibilidad para la selección de los casos de BCC fue el de pacientes con diagnóstico de BCC en consulta de teledermatología, y posteriormente confirmado por biopsia. Los pacientes pertenecen al Sistema Andaluz de Salud y fueron atendidos en centros médicos adscritos al área de influencia del Hospital Universitario Virgen Macarena de Sevilla (España). El criterio de inclusión fue el de pacientes mayores de 18 años (no pediátricos).

En total, la base de datos contiene 692 imágenes de BCCs y 671 no-BCC.

### 2.2. Análisis del color

Para mejorar y facilitar el proceso de aprendizaje del clasificador basado en redes neuronales, proponemos introducir, como información adicional, los principales colores presentes en la lesión. Debido a la pequeña base de datos, consideramos que el proceso de entrenamiento podría facilitarse mediante la introducción de los colores significativos de los patrones dermatoscópicos.

Experimentalmente, midiendo diferencias de color entre píxeles, se ha determinado que existían 20 colores significativos diferentes. Para ello se implementó un algoritmo de *clustering (K-means)*, adaptado a los distintos espacios de color escogidos, fusionando centroides cuyas distancias entre sí eran menor que un umbral.

Para realizar esta cuantificación del color se ha empleado el espacio de color uniforme derivado del Modelo de Apariencia del Color CIECAM16 [10], ya que respeta la apariencia visual de los colores y las distancias euclídeas en este espacio están bien correlacionadas con las diferencias percibidas.

### 2.3. Análisis de textura

Para el análisis de textura se emplean medidas derivadas de una nueva matriz de coocurrencia de color (CCM). Estas medidas se extraen a partir de las imágenes cuantificadas empleando CIECAM16. Como se cuantifica en 20 colores principales, se obtiene una matriz de coocurrencia 20×20. Para cada elemento de la matriz, para una determinada posición relativa entre pares de píxeles, se estima la probabilidad de coocurrencia del índice de color  $i$  y del índice de color  $j$ ,  $P_{ij}$ .

Para calcular los diferentes parámetros derivados de la CCM, se tiene en cuenta la información de color y las distancias entre colores. Así, en lugar de calcular la diferencias entre los índices de color de los píxeles como  $|i-j|$ , donde  $i$  y  $j$  son los índices, se calcula la diferencia de color,  $\Delta E$ , entre el color correspondiente al índice  $i$ ,  $C_i$ , y el color correspondiente al índice  $j$ ,  $C_j$ ,  $\Delta E$  se calcula de acuerdo con lo especificado en el espacio de color CAM16-UCS [10]. De acuerdo a esto, los parámetros calculados a partir de la CCM son:

Homogeneidad:

$$H = \sum_{i,j}^{20,20} \frac{P_{ij}}{1 + (\Delta E(C_i, C_j))^2}$$

Media:

$$\mu = \sum_{i,j}^{20,20} C_i P_{ij}$$

Varianza:

$$\sigma^2 = \sum_{i,j}^{20,20} (\Delta E(C_i, \mu_i))^2 P_{ij}$$

Correlación:

$$\rho = \sum_{i,j}^{20,20} \frac{\Delta E(C_i, \mu_i) \Delta E(C_j, \mu_j)}{\sqrt{\sigma_i^2 \sigma_j^2}} P_{ij}$$

Entropía:

$$S = \sum_{i,j}^{20,20} -P_{ij} \ln(P_{ij})$$

### 2.4. Clasificación

La arquitectura del clasificador se resume en la Figura 1. En dicho clasificador se combinan tres entradas diferentes: la imagen original, la imagen cuantificada a los 20 colores principales y las características de textura extraídas de la CCM. Las imágenes original y cuantificada son procesadas mediante las primeras capas de una VGG pre-entrenada con la base de datos ImageNet [11]. Las características de

textura son procesadas mediante un perceptrón multicapa (MLP). Las salidas de las tres redes neuronales se concatenan y alimentan otro MLP.

Como una misma imagen puede contener múltiples patrones, la última capa del clasificador contiene 7 neuronas con función de activación sigmoideal y la salida de cada una de estas 7 neuronas es umbralizada independientemente.

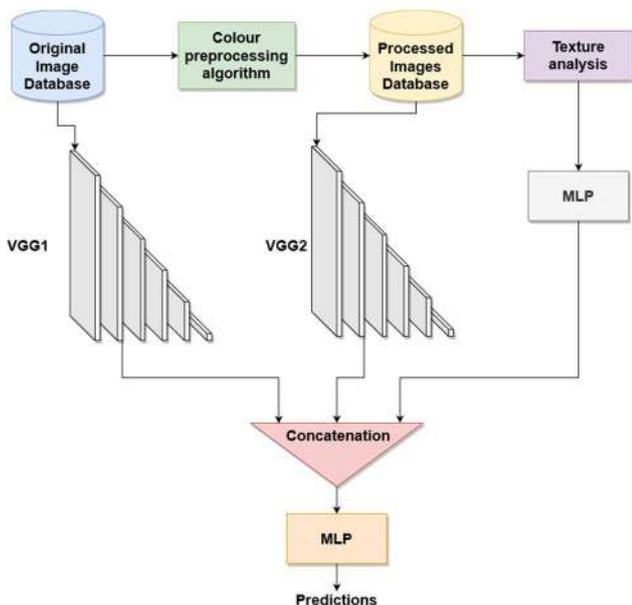


Figura 1. Esquema del clasificador empleado.

### 3. Resultados

En la Tabla 1 se resumen los resultados de detección en las imágenes de los 7 patrones dermatoscópicos que los dermatólogos emplean para diagnosticar el BCC. En dicha tabla aparece una comparativa que evidencia las ventajas de emplear el modelo de apariencia de color CIECAM16-UCS frente a otros espacios de color y de emplear los parámetros de textura propuestos frente a los tradicionales de la matriz de coocurrencia en escala de grises (GLCM).

Por último, en la Tabla 2, se muestran los resultados al clasificar todas las lesiones en BCC frente a no BCC, cuando las características dermatoscópicas del BCC se han estimado con la arquitectura de la Figura 1 y las características del CCM. Como se puede observar, los mejores resultados se obtienen cuando se estiman las diferencias de color en CIECAM16, alcanzando una sensibilidad de 0,9934. La Figura 2 muestra la curva ROC obtenida por este clasificador.

### 4. Conclusiones

En este trabajo, proponemos una arquitectura basado en redes neuronales profundas para detectar los patrones dermatoscópicos que los clínicos emplean para discriminar entre BCC y no BCC. En esta arquitectura, además de la imagen original, se introducen como entradas a la red una imagen de color cuantificada y características de color-textura, lo que mejora los resultados de la clasificación. La imagen de color cuantificada se ha cuantificado de acuerdo con las diferencias de color percibidas obtenidas de un espacio de color uniforme derivado de CIECAM16 [10].

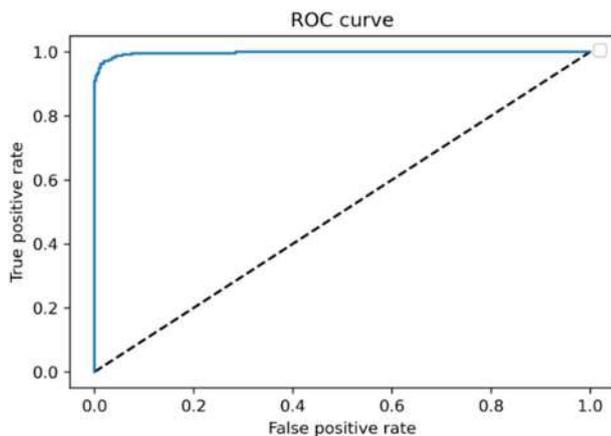
La mejora conseguida con esta metodología respecto a una red neuronal profunda VGG16 alimentada únicamente con imágenes RGB originales es la siguiente: el parámetro de especificidad ha aumentado de 0,75 a 0,82, el parámetro de sensibilidad ha aumentado de 0,77 a 0,90, y el AUC ha aumentado de 0,83 a 0,92.

Por último, cuando la lesión se clasifica en BCC o no BCC en función de las características dermatoscópicas encontradas, se alcanza una sensibilidad de 0,9934. Por lo tanto, esta clasificación puede proporcionar una clasificación precisa y explicable para el médico.

		SPEC	SENS	AUC
CIELAB+	R. Pigm.	0.92	0.99	0.99
GLCM	Ulcerac.	0.80	0.82	0.90
	Nido ov.	0.71	0.68	0.80
	Multigl.	0.78	0.77	0.86
	H. Arce	0.77	0.69	0.81
	R. Carro	0.72	0.86	0.89
	Telang.	0.72	0.85	0.87
	<b>Media</b>	<b>0.77</b>	<b>0.81</b>	<b>0.87</b>
CIECAM16	R. Pigm.	0.97	0.97	0.98
	+	Ulcerac.	0.76	0.81
GLCM	Nido ov.	0.60	0.88	0.79
	Multigl.	0.78	0.79	0.86
	H. Arce	0.80	0.77	0.85
	R. Carro	0.79	0.71	0.89
	Telang.	0.73	0.79	0.89
<b>Media</b>	<b>0.78</b>	<b>0.82</b>	<b>0.88</b>	
CIELAB+	R. Pigm.	0.98	0.97	0.99
CCM	Ulcerac.	0.82	0.75	0.89
	Nido ov.	0.74	0.84	0.86
	Multigl.	0.78	0.68	0.80
	H. Arce	0.78	0.68	0.85
	R. Carro	0.89	0.97	0.96
	Telang.	0.76	0.87	0.91
	<b>Media</b>	<b>0.82</b>	<b>0.82</b>	<b>0.89</b>
CIECAM16	R. Pigm.	0.98	0.97	0.99
+	Ulcerac.	0.86	0.92	0.94
CCM	Nido ov.	0.85	0.83	0.91
	Multigl.	0.79	0.87	0.89
	H. Arce	0.72	0.82	0.82
	R. Carro	0.87	0.93	0.96
	Telang.	0.68	0.99	0.91

	Media	0.82	0.90	0.92
<b>Tabla 1.</b> Resultados de la detección de los 7 patrones dermatoscópicos que los médicos emplean para diagnosticar el BCC				
	ACC	PPV	SPEC	SENS
CIELAB	0.9685	0.9789	0.9703	0.9673
CIECAM16	0.9699	0.9527	0.9423	0.9934

**Tabla 2.** Resultados de la de la clasificación de las lesiones en BCC o no BCC



**Figura 2.** Curva ROC de la clasificación entre BCC y no BCC

## Agradecimientos

Este trabajo ha sido financiado por Proyectos de Generación de Conocimiento 2021 del Ministerio de Ciencia e Innovación (PID2021-127871OB-I00) y proyecto FEDER-US Fondo Europeo de Desarrollo Regional (FEDER) y Consejería de Economía, Conocimiento, Empresas y Universidad de la Junta de Andalucía (US-1381640).

## Referencias

- [1] Skin Cancer Foundation, 2021. Skin cancer facts and statistics. <https://www.skincancer.org/skin-cancer-information/skin-cancer-facts/#:~:text=Skin%20cancer%20is%20the%20most,doubles%20your%20risk%20for%20melanoma> (accessed 30 March 2021)
- [2] Peris, K.; Fargnoli, M.C.; Garbe, C.; Kaufmann, R.; Bastholt, L.; Seguin, N.B.; Bataille, V.; Marmol, V.D.; Dummer, R.; Harwood, C.A.; et al. Diagnosis and treatment

of basal cell carcinoma: European consensus based interdisciplinary guidelines. *Eur. J. Cancer.* **2019**, *118*, 10–34. <https://doi.org/10.1016/j.ejca.2019.06.003>.

- [3] Breitbart, E.W.; Waldmann, A.; Nolte, S.; Capellaro, M.; Greinert, R.; Volkmer, B.; et al. Systematic skin cancer screening in northern Germany. *J. Am. Acad. Dermatol.* **2012**, *66*, 201–211. <https://doi.org/10.1016/j.jaad.2010.11.016>
- [4] Barata, C.; Celebi, E.; Marques, J.S. Explainable skin lesion diagnosis using taxonomies. *Pattern Recognit.* **2021**, *110*, 1071413. <https://doi.org/10.1016/j.patcog.2020.107413>
- [5] González-Díaz, I. DermaKNet: Incorporating the Knowledge of Dermatologists to Convolutional Neural Networks for Skin Lesion Diagnosis. *IEEE J. Biomed. Health Inform.* **2019**, *23*, 547–559. <https://doi.org/10.1109/JBHI.2018.2806962>.
- [6] Codella, N.C.F.; Lin, C.C.; Halpern, A.; Hind, M.; Feris, R.; Smith, J.R. Collaborative Human-AI (CHAD): Evidence-Based Interpretable Melanoma Classification in Dermoscopic Images. In *Lecture Notes in Computer Science 11038*; Stoyanov, D., Eds.; Springer (Cham, Switzerland) 2018. [https://doi.org/10.1007/978-3-030-02628-8\\_11](https://doi.org/10.1007/978-3-030-02628-8_11).
- [7] Barata, C.; Celebi, M.E.; Marques, J.S. A Survey of Feature Extraction in Dermoscopy Image Analysis of Skin Cancer. *IEEE J. Biomed. Health Inform.* **2019**, *23*, 1096–1109. <https://doi.org/10.1109/JBHI.2018.2845939>
- [8] Menzies, S.W.; Westerhoff, K.; Rabinovitz, H.; Kopf, A.W.; McCarthy, W.H.; Katz, B. Surface microscopy of pigmented basal cell carcinoma. *Arch. Dermatol.* **2000**, *136*, 1012–1016. <https://doi.org/10.1001/archderm.136.8.1012>.
- [9] Kittler, H.; Marghoob, A.A.; Argenziano, G.; Carrera, C.; Curiel-Lewandrowski, C.; Hofmann-Wellenhof, R.; Malvehy, J.; Menzies, S.; Puig, S.; Rabinovitz, H.; et al. Standardization of terminology in dermoscopy/dermatoscopy: Results of the third consensus conference of the International Society of Dermoscopy. *J. Am. Acad. Dermatol.* **2016**, *74*, 1093–1106. <https://doi.org/10.1016/j.jaad.2015.12.038>.
- [10] Li, C.; Li, Z.; Wang, Z.; Xu, Y.; Luo, M.R.; Cui, G.; Melgosa, M.; Brill, M.H.; Pointer, M. Comprehensive color solutions: CAM16, CAT16 and CAM16-UCS. *Color. Res. Appl.* **2017**, *42*, 703–718. <https://doi.org/10.1002/col.22131>.
- [11] Deng, J.; Dong, W.; Socher, R.; Li, L.-J.; Li, K.; Fei-Fei, L. ImageNet: A Large-Scale Hierarchical Image Database. In *Proceedings of the IEEE Computer Vision and Pattern Recognition (CVPR)*, Miami, FL, USA, 20–25 July 2009.

# Diseño de un clasificador interpretable de ruido clínico en ECG de monitorización prolongada mediante autoencoders

C. Plaza–Seco<sup>1\*</sup>, R. Holgado–Cuadrado<sup>1</sup>, J.L. Rojo–Álvarez<sup>2</sup> y M. Blanco–Velasco<sup>1</sup>

<sup>1</sup> Departamento de Teoría de la Señal y Comunicaciones, Universidad de Alcalá, {carmen.plazas, roberto.holgado, manuel.blanco}@uah.es

<sup>2</sup> Departamento de Teoría de la Señal y Comunicaciones y Sistemas Telemáticos y Computación, Universidad Rey Juan Carlos, joseluis.rojo@urjc.es

## Resumen

*La presencia de ruido en los registros de electrocardiograma (ECG) de monitorización prolongada es muy frecuente debido a su larga duración, lo que da lugar a que algunas partes sean inservibles para realizar un diagnóstico clínico. La definición de ruido clínico establece una métrica de calidad que, en lugar de evaluar el ruido desde un punto de vista cuantitativo, indica qué fragmentos de un ECG son válidos para realizar el diagnóstico clínico. Esta contribución surge en paralelo con un trabajo en el que se han clasificado tramas de ruido clínico con un FI-score de 0,74, utilizando modelos de máquinas de vectores soporte a partir de un conjunto de índices de calidad extraídos de registros anotados de monitorización prolongada. El enfoque de este trabajo se centra, no solo en la clasificación de ruido clínico, sino en la interpretabilidad del funcionamiento de los modelos, debido a su importancia en el ámbito clínico. Para ello, se propone el uso de autoencoders, que permiten extraer patrones relevantes de datos de alta dimensionalidad y estudiar su distribución en espacios latentes. Los resultados obtenidos sugieren que es posible igualar el rendimiento de clasificación desde una perspectiva interpretable y de menor dimensionalidad. El sistema podría ser una buena herramienta de apoyo para los profesionales sanitarios puesto que es capaz de discriminar tramas de ECG clínicamente interpretables y ofrecer explicabilidad e interpretabilidad de su funcionamiento.*

## 1. Introducción

Los registros de electrocardiograma (ECG) de monitorización prolongada (*Long-Term Monitoring*, LTM) se utilizan cuando existe la sospecha de la presencia de una patología cardíaca peligrosa e intermitente para el paciente. Estos registros se adquieren durante 7, 15 o incluso 21 días, de manera ambulatoria, habiéndose demostrado su utilidad para detectar numerosas patologías [1]. Sin embargo, su larga duración da lugar a que algunas partes del ECG estén muy contaminadas por distintos tipos de ruido, resultando ineficientes para el diagnóstico clínico. Por ello, la identificación de tramas válidas desde un punto de vista clínico es de gran importancia.

En [2], se demostró que la severidad clínica y el enfoque tradicional cuantitativo del ruido no están correlacionados, por lo que la evaluación de la calidad clínica de un registro LTM no puede limitarse a estimar cuantitativamente su nivel de corrupción de ruido. Así, se define el ruido clínico

que, mediante una clasificación en términos de calidad clínica, permite indicar qué fragmentos de un ECG son útiles para el diagnóstico. En un trabajo paralelo se ha demostrado que es posible distinguir segmentos clínicamente válidos mediante técnicas de aprendizaje automático o *Machine Learning* (ML). En concreto, se utilizaron máquinas de vectores soporte (*Support Vector Machine*, SVM) para abordar un problema binario en el que se clasificaron tramas en las categorías de *limpia* o *ruidosa* en términos de interpretabilidad clínica con un FI-score de 0,74.

Las técnicas de ML han demostrado ser una poderosa herramienta de apoyo en gran cantidad de procedimientos del ámbito clínico [3], ya que son capaces de aprender relaciones complejas existentes en datos de alta dimensionalidad. Sin embargo, en la mayoría de las ocasiones, es necesario comprender el funcionamiento del modelo para poder aplicarlo en tareas del mundo real, y más aún, en el entorno sanitario. De otro modo, no se generará la confianza suficiente entre el personal facultativo para que puedan hacer uso de estos instrumentos, siendo el principal obstáculo para su transferencia. La interpretabilidad permite generar conocimiento acerca de cómo se lleva a cabo la toma de decisiones que realiza un modelo de ML, lo que le dota de una mayor fiabilidad. Así, la interpretabilidad se convierte en una parte esencial del desarrollo de sistemas de ML.

Por tanto, este trabajo enfoca el problema de clasificación de ruido clínico desde una perspectiva interpretable. Para ello, se propone el uso de autoencoders (AE), modelos que permiten extraer patrones relevantes de los datos y estudiar cómo se distribuyen en espacios latentes, a partir de un conjunto de características que describen la forma de onda y el contenido del ECG.

## 2. Base de datos

Un trabajo introductorio de ruido clínico en el ECG se presentó inicialmente en [2] con el propósito de establecer una taxonomía en términos de interpretabilidad clínica, definiéndose 5 clases de severidad clínica del ruido. Siguiendo estos criterios, se etiquetó un conjunto de datos adquiri-

Tabla 1: Lista de las características y su descripción.

Característica	Descripción
$x_1$	Precisión en la detección de la onda R [4]
$x_2$	Periodicidad basada en la correlación
$x_3$	Comprobación de la línea de base [4]
$x_4$	Amplitud relativa del ruido de alta frecuencia [4]
$x_5$	Media
$x_6$	Desviación estándar
$x_7$	Skewness
$x_8$	Kurtosis
$x_9$	Desviación relativa del QRS [4]
$x_{10}$	Índice de Pureza Espectral
$x_{11}$	Entropía
$x_{12}$	Potencia relativa de la línea de base [4]
$x_{13}$	Potencia relativa del QRS [4]
$x_{14}$	Energía relativa del QRS [4]
$x_{15}$	Frecuencia central
$x_{16}$	Índice de calidad del ECG [5]

dos ( $f_s = 200\text{Hz}$ ) con un registrador externo de eventos (*External Event Recorded*, EER) de manera continuada a lo largo de toda su duración (aproximadamente 6,5 horas). Ese conjunto de datos de registros LTM ha sido utilizado en este trabajo y el problema de clasificación que se resuelve distingue entre dos clases: tramas limpias de tramas ruidosas. Los datos originales se segmentan en bloques de 5 segundos y se corrige el desbalanceo de clases descartando a los pacientes que presentan menos del 25 % de instancias en cualquier clase, lo que reduce el conjunto a 6 pacientes. Como resultado, se obtiene un conjunto de 2818 instancias (bloques de 5 segundos) correspondiente a 6 pacientes, donde un 47 % de ejemplos asociados a la clase *limpia* y el 53 % restante a la clase *ruidosa*. Se extraen múltiples características,  $x_j$ , de los segmentos de señal  $\mathbf{x} = [x_1, x_2, x_3, \dots, x_n]^T \in \mathbb{R}^n$  con  $n = 16$ , que informan acerca de (Tabla 1): la posición y distribución de los puntos críticos ( $j = 1, 2$ ); características del dominio del tiempo ( $j = 3, 4$ ); propiedades estadísticas ( $j = 5, \dots, 11$ ); y características espectrales ( $j = 12, \dots, 16$ ). Los datos a procesar se corresponden con la matriz  $\mathbf{X} \in \mathbb{R}^{m \times n}$ ,  $m = 2818$  y  $n = 16$ , cada ejemplo asociado al vector de etiquetas  $\mathbf{y} = [y^{(1)}, y^{(2)}, \dots, y^{(m)}]^T$ . Las características se normalizan en el rango  $[0, 1]$ .

### 3. Métodos

#### 3.1. SVM

Las SVM se han convertido en uno de los métodos de aprendizaje supervisado más destacados para resolver problemas de clasificación. Cabe destacar que este algoritmo es capaz de reducir los problemas de sobreajuste presentes en otros clasificadores, ya que, para problemas binarios, trata de separar óptimamente un conjunto de datos representados en el espacio mediante un hiperplano [6] eligiendo, entre las múltiples posibilidades, aquel que se encuentra a la mayor distancia entre los ejemplos de ambas clases.

#### 3.2. Autoencoders

Las técnicas de reducción de dimensionalidad permiten extraer información relevante de los datos y son ampliamente utilizadas para reducir la complejidad de los modelos

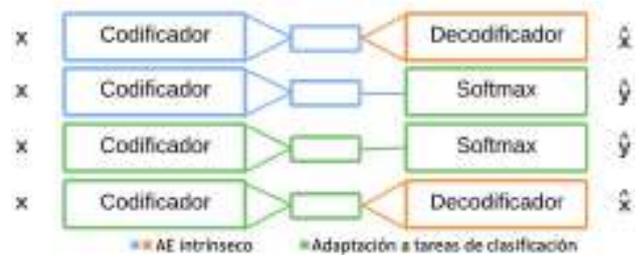


Figura 1: Fases de entrenamiento del autoencoder.

y optimizar su aprendizaje. Entre estos métodos, destacan los AE, debido a su capacidad para extraer patrones de datos de alta dimensionalidad y proyectar su distribución en espacios latentes de menor dimensión. Un AE es una red neuronal entrenada para replicar su entrada,  $\mathbf{x}$ , de forma no supervisada en la salida,  $\hat{\mathbf{x}}$ , como se puede ver en la primera fila de la Figura 1. El AE está formado por un primer bloque que comprime los datos, el codificador (azul), y un segundo bloque, el decodificador (naranja), que reconstruye los datos de entrada a partir de la información comprimida. El cuello de botella, denominado *espacio latente*, aprende características significativas del conjunto de datos, conocidas como *variables latentes*,  $l_i(\mathbf{x})$ . El proceso de aprendizaje consiste en minimizar una función de pérdida  $\mathcal{L}(\mathbf{x})$ , como el error cuadrático medio:

$$\mathcal{L}(\mathbf{x}) = \frac{1}{m} \sum_{i=1}^m (\mathbf{x}_i - \hat{\mathbf{x}}_i)^2. \quad (1)$$

La menor dimensión del espacio latente permite la representación visual de las variables latentes en función de las categorías de los datos de entrada, dando lugar a una localización espacial que se puede explotar en tareas de clasificación (verde), entrenando un *softmax* sobre las variables latentes, como se representa en la segunda fila de la Figura 1. El espacio latente, a su vez, se puede adaptar específicamente para resolver el problema de clasificación entrenando de forma supervisada el bloque codificador, indicado en la tercera fila de la Figura 1, donde se predice la clase  $\hat{\mathbf{y}}$  a partir de los datos de entrada  $\mathbf{x}$ . Así, se produce una compresión de los datos, pero al mismo tiempo se agrupan en función de la clase,  $\mathbf{y}$ , a la que pertenecen. Finalmente, el AE asociado a este espacio latente se obtendría mediante el entrenamiento del decodificador a partir del espacio latente definido en las etapas anteriores, es decir, manteniendo “congelados” los pesos del codificador, tal y como se representa en la cuarta fila de la Figura 1.

#### 3.3. Diseño experimental

Para evitar el sobreajuste al paciente, se propone el esquema de validación de la Figura 2, donde se asignan todos los datos de un paciente al conjunto de test. El 80 % de los datos de los cinco pacientes restantes se emplea para realizar el entrenamiento y ajuste de los hiperparámetros del modelo mediante una validación cruzada (VC) de 5 *folds*. El 20 % restante se utiliza como conjunto de validación, una vez optimizado el modelo en la etapa de entrenamiento. Se entrenan 6 modelos siguiendo este esquema para comprobar la consistencia de los resultados y compensar la escasa

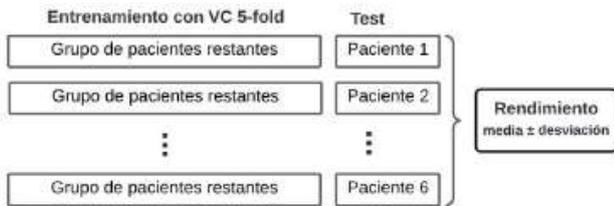


Figura 2: Esquema de validación.

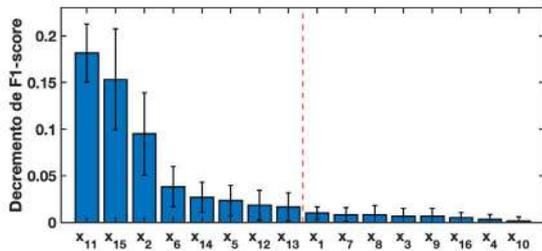


Figura 3: Importancia relativa de las características.

disponibilidad de pacientes. El rendimiento global se obtiene como la media y desviación estándar de las combinaciones descritas. La evaluación del rendimiento de los modelos se realiza en términos de detección en la clase *limpia* y se emplean cinco métricas estándar ampliamente conocidas: *accuracy* (*Acc*), *recall* (*Re*), *precision* (*Pr*) y *F1-score*.

## 4. Resultados

### 4.1. SVM e importancia de las características

El problema de clasificación planteado se ha resuelto mediante SVM [7], obteniéndose los resultados de la primera columna de la Tabla 2. La importancia de cada característica se ha estudiado aplicando el método *permutation feature importance* [6]. Consiste en analizar la disminución en el rendimiento cuando se anula la característica  $x_j$  mezclando aleatoriamente sus muestras. De esta manera se desvincula la relación entre dicha característica y las categorías a determinar, por lo que el decremento del rendimiento se convierte en un indicador de la dependencia del modelo con esa característica. La Figura 3 muestra los resultados de mayor a menor importancia de cada característica. Se concluye que las 3 primeras son las más significativas y que la disminución del rendimiento es poco apreciable a partir de la novena. Por ello, se diseñan y entrenan las SVM únicamente con las 8 características más relevantes, cuyos resultados se muestran en la segunda columna de la Tabla 2. De estos resultados, se desprende la importancia de seleccionar las características. Al entrenar los modelos únicamente con los 8 índices de calidad más relevantes, se observa una mejora en el grupo de test, lo que indica una mejor generalización de los modelos.

### 4.2. Autoencoder

En este experimento se aborda el problema de clasificación de ruido clínico utilizando la arquitectura mostrada en la Figura 1, donde se ha elegido un espacio latente de dimensión  $\mathbb{R}^3$  para poder representar la distribución de los datos, y al mismo tiempo, la frontera de clasificación obtenida. Tanto el codificador como el decodificador están formados

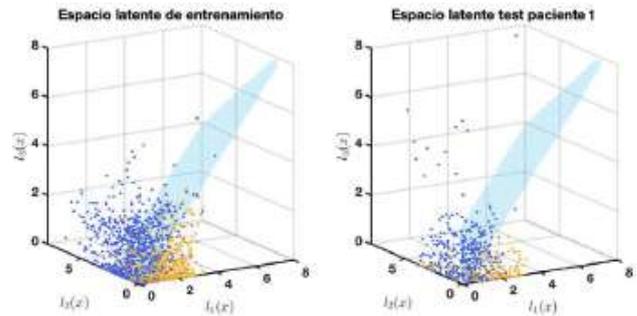


Figura 4: Espacio latente del AE y frontera de separación en la primera iteración del esquema de validación.

por una única capa intermedia *fully connected* de 50 neuronas. A la vista del comportamiento de los modelos SVM, se ha decidido entrenar los AE únicamente con los 8 índices de calidad más relevantes (ver Figura 3). Así, el uso de los AE, se traduce en el siguiente problema de reducción de dimensionalidad:  $\mathbb{R}^8 \rightarrow \mathbb{R}^3$ .

Para la evaluación de los 6 modelos con AE se han obtenido las métricas de evaluación en tres conjuntos diferentes: entrenamiento, validación y test, reflejadas en la tercera columna de la Tabla 2. El rendimiento de la clasificación en los conjuntos de entrenamiento y validación supera al del conjunto de test. Esto confirma que las características aprendidas de un grupo de pacientes durante el entrenamiento de los modelos dan lugar a una sobrestimación del rendimiento del sistema cuando se utilizan observaciones de los mismos pacientes en el conjunto de test. En la Tabla 3 se muestran el resto de métricas obtenidas al entrenar con estas 8 características, tanto para los modelos SVM como para los AE. Los resultados son similares con ambos esquemas, destacando una menor desviación estándar cuando la clasificación se efectúa con AE, lo que indica una mayor robustez frente a la variabilidad entre pacientes.

En la Figura 4 se muestra la proyección de los datos en el espacio latente del AE y la frontera de separación para el caso correspondiente a la primera iteración del esquema de validación de la Figura 2. Como se puede ver en la gráfica de la izquierda, la frontera se ajusta a los datos de entrenamiento dejando ambas clases claramente separadas. De esta manera, se ha conseguido proyectar las observaciones de cada categoría en distintas localizaciones del espacio; o lo que es lo mismo, diferentes regiones del espacio latente estarían asociadas a las clases que se están analizando. Por tanto, se espera que la regionalización sea similar en el caso del conjunto de test, de manera que la frontera aprendida durante el entrenamiento consiga separar las clases en el conjunto de test. Como se observa en la gráfica de la derecha, la generalización sucede de la manera que se ha explicado, aunque no es tan precisa en todos los casos analizados, debido a la variabilidad interpaciente, viéndose reflejada en la desviación estándar de las métricas del conjunto de test que se muestran en las Tablas 2 y 3.

Aunque los modelos SVM y AE presentan resultados de clasificación similares, el uso de estos últimos ofrece un matiz adicional: la proyección de los datos de entrada en regiones compactas y separadas de un espacio visualizable. Puesto que dicho espacio se emplea para la toma de deci-

Tabla 2: Resultados de *F1-score* con los modelos SVM y AE.

Conjunto	SVM entrada 16	SVM entrada 8	Autoencoder entrada 8
Entrenamiento	0.81 ± 0.03	0.81 ± 0.03	0.83 ± 0.01
Validación	0.80 ± 0.02	0.78 ± 0.04	0.82 ± 0.02
Test	0.74 ± 0.18	0.77 ± 0.15	0.76 ± 0.13

Tabla 3: Comparación entre SVM y AE con 8 características de entrada.

Test	Acc	Re	Pr
SVM	0.76 ± 0.16	0.78 ± 0.14	0.77 ± 0.19
Autoencoder	0.76 ± 0.14	0.76 ± 0.14	0.79 ± 0.18

sión, servirá, por tanto, como medio para aportar una explicación cualitativa acerca del funcionamiento del modelo. En términos de interpretabilidad, esto es una gran ventaja en el ámbito clínico, ya que permite entender de forma sencilla la manera en que el modelo realiza las predicciones sobre un nuevo paciente, simplemente identificando las zonas del espacio donde se proyectan sus datos.

## 5. Discusión y conclusiones

Junto con este trabajo, se ha presentado una contribución [7] en la que se aborda el diseño de un sistema de clasificación de ruido clínico mediante modelos SVM, cuyos resultados se emplean en este proyecto a título comparativo. La aportación adicional del presente trabajo, en la parte de SVM, se debe a la selección de características, donde, a tenor de los resultados, es posible mantener el rendimiento abordando el problema con una dimensionalidad menor.

La interpretabilidad se ha convertido en una parte fundamental para entender el funcionamiento de los sistemas de ML en el ámbito de la Ingeniería Biomédica. Es la clave para fomentar la confianza del personal facultativo en las soluciones que se ofrecen, siendo éste un aspecto necesario para la transferencia a tecnología. Para ello, la selección de características por medio del análisis de su importancia no parece suficiente. Es cierto que esto permite obtener soluciones más simples, sin embargo, en muchas ocasiones los médicos desconocen el significado de los predictores, ya que carecen de sentido clínico. Luego, interpretar el funcionamiento del sistema a través del conocimiento de las características predominantes no es una solución viable.

En este trabajo se muestra la importancia de emplear procedimientos que permitan representar espacios visualizables sobre los que se toman las decisiones. Aunque los espacios latentes obtenidos mediante AE carecen de interpretación clínica, la regionalización de los datos de entrada permite entender el funcionamiento de los sistemas. Así, si en dicho espacio existen regiones asociadas a las diferentes categorías, cuando llega un caso nuevo, se podrá interpretar la predicción ofrecida por el modelo al observar la zona del espacio en la que se proyecta.

De los resultados obtenidos en este trabajo, se infieren dos conclusiones. Por un lado, la selección de características permite eliminar variables que, aunque son descriptivas en cuanto a la morfología de la señal en su conjunto, aportan un menor valor predictivo, obteniéndose un incremento en

el rendimiento de clasificación tras su eliminación. Por otro lado, el mapeado de los datos de entrada a un espacio tridimensional permite resolver el problema de clasificación sin pérdida de rendimiento y con el valor añadido de la interpretabilidad en lo que a la toma de decisiones se refiere.

La mejora del presente trabajo está sujeta a la exploración de otras técnicas de reducción de dimensionalidad capaces de regionalizar los datos en espacios observables y adaptarlos a la tarea de clasificación, como sucede con los AE. Además, para aumentar el rendimiento de los modelos se proponen técnicas semi-supervisadas de *Active Learning* que permitan incorporar nuevos datos etiquetados.

## Agradecimientos

Queremos agradecer a los autores de [2] su contribución al desarrollo de la base de datos de ruido clínico. Además, este trabajo ha sido parcialmente financiado por la Comunidad de Madrid con el proyecto EPU-INV/2020/002.

## Referencias

- [1] F. J. Pastor-Pérez, S. Manzano-Fernández, R. Goya-Esteban, D. A. Pascual-Figal, *et al.*, “Comparison of detection of arrhythmias in patients with chronic heart failure secondary to non-ischemic versus ischemic cardiomyopathy by 1 versus 7-day holter monitoring,” *The American Journal of Cardiology*, vol. 106, no. 5, pp. 677–681, 2010.
- [2] E. Everss-Villalba, F. M. Melgarejo-Meseguer, M. Blanco-Velasco, F. J. Gimeno-Blanes, *et al.*, “Noise maps for quantitative and clinical severity towards long-term ECG monitoring,” *Sensors*, vol. 17, no. 11, p. 2448, 2017.
- [3] S. Yeasmin, “Benefits of artificial intelligence in medicine,” in *2019 2nd International Conference on Computer Applications Information Security*, 2019.
- [4] Q. Li, C. Rajagopalan, and G. D. Clifford, “A machine learning approach to multi-level ECG signal quality classification,” *Computer Methods and Programs in Biomedicine*, vol. 117, no. 3, pp. 435–447, 2014.
- [5] D. P. Tobón, T. H. Falk, and M. Maier, “MS-QI: A modulation spectrum-based ECG quality index for telehealth applications,” *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 8, pp. 1613–1622, 2014.
- [6] T. Hastie, R. Tibshirani, and J. Friedman, *The elements of statistical learning: data mining, inference and prediction*. Springer, 2 ed., 2009.
- [7] R. Holgado-Cuadrado, C. Plaza-Seco, L. Lovisolo, and M. Blanco-Velasco, “Clasificación de ruido clínico en ECG de monitorización prolongada mediante máquinas de vectores soporte,” CASEIB 2022.

# Análisis de factores de riesgo y predicción de enfermedades crónicas usando métodos de aprendizaje automático

A. Abad-González<sup>1</sup>, C. García-Vicente<sup>1</sup>, D. Chushig-Muzo<sup>1</sup>, C. Soguero-Ruiz<sup>1</sup>

<sup>1</sup> Departamento de Teoría de la Señal y las Comunicaciones y Sistemas Telemáticos, Universidad Rey Juan Carlos, Madrid, España, a.abad.2018@alumnos.urjc.es, {clara.garcia.vicente, david.chushig, cristina.soguero}@urjc.es

## Resumen

Las enfermedades crónicas (ECs) son la primera causa de muerte a nivel mundial. Dos de las patologías crónicas con mayor prevalencia son la cardiopatía coronaria (CAD) y la diabetes mellitus (DM). La detección precoz de ECs es clave para disminuir los efectos secundarios, el número de defunciones y los costes asociados. Los avances en Aprendizaje Automático (AA) han sido notables en los últimos años, especialmente en el ámbito de la investigación clínica, apoyando a profesionales sanitarios en la toma de decisiones y en la identificación del riesgo de padecer ECs. El objetivo del presente trabajo es doble. En primer lugar, validar la utilidad de métodos supervisados de AA para la predicción de CAD y DM, y, en segundo lugar, identificar los factores de riesgo más relevantes para el desarrollo de estas enfermedades. Los resultados experimentales demostraron que los métodos de AA tienen un buen desempeño para identificar pacientes con ECs. Para CAD, el regresor logístico alcanzó una tasa de acierto (accuracy) de 68,3%, mientras para DM, K-nearest neighbors obtuvo una accuracy de 96,5%. Los factores de riesgo identificados para CAD y DM fueron consistentes con la evidencia clínica y estudios previos. Este trabajo muestra las ventajas de aplicar métodos de AA en el ámbito clínico, específicamente para la predicción de ECs e identificación de factores de riesgo, apoyando así en la toma de decisiones a los clínicos, lo que permitiría a largo plazo una importante reducción de carga económica y sanitaria.

## 1. Introducción

Las enfermedades crónicas (ECs) son la principal causa de morbilidad y mortalidad a nivel mundial [1]. Éstas se caracterizan por ser enfermedades de larga duración y no tener cura, aunque si tienen tratamiento en la mayoría de los casos. En particular, la Cardiopatía Coronaria (CAD) y la Diabetes Mellitus (DM) son dos de las ECs con mayor prevalencia a nivel mundial [1]. La CAD es la enfermedad cardíaca más común en Estados Unidos, y en España fue la primera causa de muerte hasta la llegada del Covid-19 [2]. La DM es una patología causada por diversos factores biológicos y ambientales la cual representa la novena causa de muerte en el mundo [1]. Según estadísticas de la Federación Internacional de Diabetes, aproximadamente 463 millones de personas entre 20-79 años padecen DM, valor que se estima se duplicará en los próximos 25 años [3]. La prevalencia de estas dos enfermedades, su alta tasa de morbilidad, mortalidad y consumo de recursos sanitarios suponen un grave problema desde el punto de vista epidemiológico, sanitario y económico. La detección temprana de ECs es fundamental para aliviar la carga sanitaria, prevenir casos clínicos graves, y reducir el número de defunciones.

En los últimos años, se han producido importantes avances en las áreas de Inteligencia Artificial (IA) y Aprendizaje Automático (AA) [4]. El AA es una disciplina de la IA cuya finalidad es generar algoritmos que aprenden de manera autónoma relaciones/patrones a partir de datos sin necesidad de definirlos a priori [4]. En el ámbito clínico, múltiples trabajos han usado métodos de AA para identificar patrones, predecir la evolución de estados de salud, entre otros [4, 5]. La detección precoz de enfermedades crónicas es fundamental para disminuir los efectos secundarios, el número de defunciones y los costes asociados que suponen. El uso de la IA en este ámbito ha supuesto un gran avance, facilitando las labores de detección y obteniendo predicciones más fiables. La incorporación de estos métodos en la práctica clínica puede ser de gran ayuda en la toma de decisiones para el diagnóstico de enfermedades y permiten analizar grandes volúmenes de datos en un corto periodo de tiempo [6].

El objetivo de este trabajo es doble. En primer lugar, en base a una serie de variables clínicas, demográficas, y de estilo de vida, se predice si un paciente padece CAD y DM mediante métodos supervisados de AA. En segundo lugar, se identifican los principales factores de riesgo asociados con el desarrollo de ambas ECs a través de métodos de selección de características (FS, del inglés *Feature Selection*). De esta manera, se puede evaluar el riesgo de padecer la patología y analizar en qué medida ésta puede ser modificada mediante nuevos hábitos, disminuyendo así la probabilidad de sufrir la enfermedad.

## 2. Materiales

### 2.1. Bases de Datos

En este trabajo, se ha considerado una base de datos diferente para cada patología analizada, CAD y DM. Para DM, los datos fueron recopilados y compartidos por el hospital *Sylhet Diabetes* en Bangladesh [7]; mientras para CAD, los datos fueron obtenidos del *Framingham Heart Study* [8]. La primera base está formada por 520 observaciones, 16 variables binarias (género, poliuria, polidipsia, debilidad, polifagia, picor genital, visión borrosa, picores, irritabilidad, cicatrización, parestia parcial, rigidez, alopecia, obesidad) y 1 numérica (edad). La base de CAD incluye 4283 observaciones con 8 variables numéricas (edad, cigarrillos por día, colesterol, presión arterial (PA) diastólica, PA sistólica, índice de masa corporal (IMC), ritmo cardíaco, glucosa), y 7 variables categóricas (género, fumador, medicamentos PA, acci-

dente cerebrovascular (ACV), hipertensión, diabetes, educación). Los datos fueron anonimizados previamente, imposibilitando la identificación de pacientes. Estas bases de datos son públicamente accesibles en *UCI repository*. Se consideró la técnica *one-hot encoding*, para transformar las variables categóricas en variables binarias. En nuestro caso, la variable educación en la base de datos de CAD que tomaba 4 posibles valores resultó en 4 variables binarias. Además, se normalizaron los datos para ajustarlos al rango de valores en las variables y evitar dar mayor peso a unas variables que a otras. En este trabajo, la normalización estándar y min-max fueron consideradas, obteniendo mejores resultados en la predicción de las ECs con la segunda aproximación.

### 3. Métodos

A continuación, se describen los métodos de clasificación así como las técnicas de selección de características consideradas en este trabajo.

#### 3.1. Métodos supervisados para clasificación

Para la predicción de ECs se han entrenado los diferentes modelos de AA con las observaciones de cada una de las bases de datos de forma separada. Se consideraron métodos supervisados lineales, incluyendo *Linear Regression* (LR), LR con L1 como regularizador (LASSO) y Ridge. En cuanto a métodos no lineales, se consideraron: *Naive Bayes* (NB), *Decision Tree* (DT), *Support Vector Machine* (SVM), *Random Forest* (RF), *K-Nearest Neighbours* (KNN) y *XGBoost* (XGB). A continuación se describen brevemente:

\* *LR* estima la relación entre las variables de entrada y la variable a predecir (*target*) mediante una función logística, buscando minimizar la suma de los cuadrados de los residuos entre los valores reales y predichos [9]. En este trabajo se utilizó LR con regularización, incluyendo las variantes LASSO y Ridge.

\* *DT* es un método no paramétrico que está formado por un conjunto de nodos que se dividen recursivamente en ramas durante el entranamiento según varios criterios (ganancia, índice Gini) [9].

\* *RF* construye múltiples DTs a partir de subconjuntos aleatorios del conjunto de entrenamiento. Las estimaciones de cada DT se promedian para producir la estimación final [9]. Para realizar la predicción, el modelo toma todos los valores de salida de los diferentes DT y seleccionará aquella que cuente con más votos, es decir, la clase que más veces ha sido seleccionada.

\* *XGB* es una técnica basada en DTs que da más robustez mediante la creación de un conjunto de modelos secuenciales DT y usa un término de regularización para evitar el sobreajuste.

\* *SVM* es un método kernel que transforma el espacio de entrada en un espacio donde se crea un hiperplano

que separa las diferentes observaciones en dos clases [9].

\* *NB* es un método probabilístico que se basa en el Teorema de Bayes para predecir a qué clase pertenece una observación.

\* *KNN* calcula la distancia entre la observación objetivo que se desea clasificar en una clase y sus K vecinos más cercanos, los cuales se ordenan de manera ascendente y posteriormente, la observación se asignará a la clase a la que pertenezcan sus K vecinos más cercanos [9].

#### 3.2. Técnicas de selección de características

Las técnicas de FS buscan seleccionar un subconjunto de variables  $k$  del conjunto total de variables de entrada  $d$  ( $k < d$ ). Estos métodos buscan mejorar la eficiencia computacional y reducir el error de generalización de modelo predictivos eliminando variables irrelevantes/redundantes. Los métodos de FS generalmente se dividen en tres tipos: *filter*, *wrapper* y *embedded* [10].

Los métodos *filter* realizan la FS independiente del algoritmo de AA y se basan en un *scoring* para medir la relación entre las variables y la variable objetivo (*target*). Como *scoring* más usadas están la correlación, información mutua, varianza, o estadístico [10]. Los métodos *wrapper* crean modelos predictivos con diferentes subconjuntos de variables y utilizan el rendimiento de estos modelos como función objetivo para evaluar el subconjunto de variables. Se busca la combinación óptima de variables que maximiza la capacidad predictiva. En los métodos *embedded* se integra la FS en la propia construcción del modelo sin necesidad de una evaluación adicional de las variables. En este trabajo se consideraron métodos *filter*, incluyendo Información Mutua (IM), chi2, test basado en remuestreo bootstrap (TRB); *wrapper* (secuenciales *greedy* (SFS)); y *embedded* (LASSO, Ridge, DT, RF, XGB).

IM calcula la información mutua para cuantificar la relevancia de las variables con respecto a la variable *target* y realizar el FS. Chi2 mide la independencia entre dos variables. TRB calcula el valor medio de las observaciones de las clases A, B ( $\mu_A$ ,  $\mu_B$ ) y obtiene el intervalo de confianza ( $\Lambda_I$ ) de la diferencia de las medias/proporciones para cada variable [11]. La FS se realiza utilizando  $\Lambda_I$ , descartando las variables que incluyan el valor nulo ( $0 \in \Lambda_I$ ), pues no habría diferencia significativa entre clases para esa variable.

Los métodos SFS son algoritmos *greedy* que construye conjuntos de características de manera iterativa buscando mejorar la calidad de la predicción, pero no evalúa todos los conjuntos de características posibles [12]. Existen tres tipos de SFS, *Forward Feature Selection* (FFS) [12], *Backward Feature Selection* (BFS) [12] y *Stepwise Feature Selection* (SFFS) [12]. En cada iteración, el BFS elimina la característica que provoca la menor pérdida de rendimiento al eliminarla. FFS comienza con todas

las variables y se van eliminando una a una midiendo el accuracy en cada iteración. SFFS va añadiendo y eliminando variables al modelo evaluando el accuracy en cada iteración. En LASSO, la FS se logra identificando las variables con coeficientes distintos de cero. Para DT, RF y XGB se selecciona las variables con mayor valor de importancia (*feature importance*).

### 4. Resultados y discusión

En esta sección, se presentan los resultados de clasificación usando los modelos de AA, y además se presenta los variables más relevantes por las técnicas de FS.

#### 4.1. Diseño Experimental

Para cada base de datos, se crearon 5 particiones (*hold-out strategy*), y cada partición se se dividió en un conjunto de *training* y *test*, con un 70 % y 30 % de observaciones, respectivamente. El conjunto de *training* fue empleado para entrenar los diferentes modelos supervisados, mientras el conjunto de *test* fue usado para evaluar el rendimiento predictivo de estos modelos. Para la selección de hiperparámetros en los modelos, se usó *10-fold cross-validation (CV)* [13]. Los siguientes hiperparámetros fueron considerados:  $\lambda$  (LASSO),  $C$  (SVM), número de vecinos (KNN), mínimo de muestras para dividir un nodo y profundidad máxima del árbol (DT).

Las bases de datos consideradas presentaban desbalanceo, con una mayor proporción de observaciones de una de las clases que de otra, lo cual puede afectar en el entrenamiento de los métodos supervisados. Para abordar esto, se consideró la técnica de *random undersampling* [14] para el balanceo de los datos. Esta técnica reduce el número de observaciones de las clases mayoritarias tomando como referencia el número de observaciones de la clase minoritaria. Para cada base de datos los modelos se entrenaron utilizando dos enfoques: (i) incluyendo todas las variables; y (ii) considerando las variables seleccionadas por los métodos de FS.

#### 4.2. Interpretación de Factores de Riesgo

En esta subsección se presenta un análisis de los factores de riesgo asociados con CAD y DM mediante el uso de diferentes técnicas de FS. Las variables que se utilizarán en el entrenamiento de modelos supervisados son aquellas seleccionadas por un tercio o más de los métodos FS. En el caso de CAD, los factores de riesgo (variables más relevantes) seleccionados por cada métodos se representan en la Figura 1.

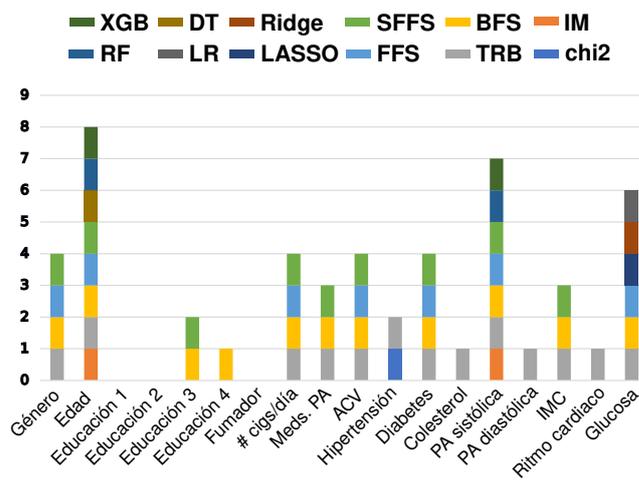


Figura 1: Variables seleccionadas por los métodos de FS para la base de datos de CAD.

Para CAD, las variable seleccionadas fueron género, edad, número de cigarrillos por día, ACV, diabetes, presión sistólica y glucosa.

En la Figura 2 se presentan las variables seleccionadas por los métodos de FS. Analizando el gráfico obtenido, se emplearán las características seleccionadas por un tercio o más de los modelos (4 modelos o más): poliuria, polidipsia, debilidad, polifagia, candidiasis genital, visión borrosa, irritabilidad, paresia parcial, género, retraso en la cicatrización y alopecia.

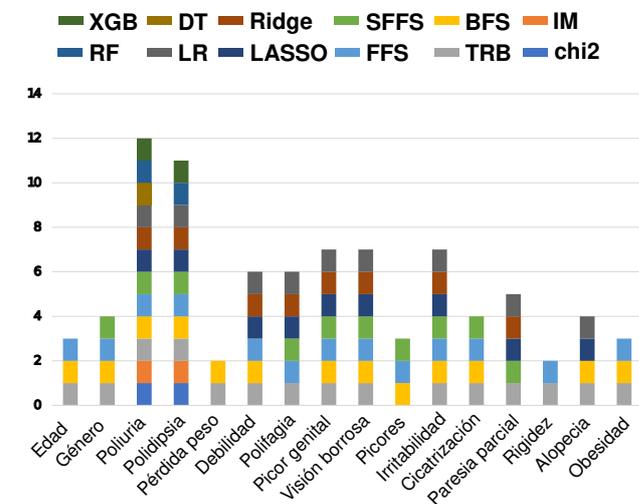


Figura 2: Variables seleccionadas por los métodos de FS para la base de datos de DM.

#### 4.3. Resultados de Predicción de CAD y DM

En esta subsección se presentan los resultados obtenidos con los métodos supervisados usando las bases de datos de CAD y DM. Como se mencionó, para entrenar los métodos se consideraron: (i) todas las variables; y (ii) las variables seleccionadas por en el análisis de FS. En la Tabla 1 se presentan los resultados para CAD. El modelo que mejores medidas de prestación obtiene fue LR con un *accuracy* de 68.3 % y usando las variables seleccionadas en FS.

Método	Accuracy	
	Sin FS	Con FS
LR	0.674 (0.013)	<b>0.683 (0.016)</b>
Ridge	0.665 (0.012)	0.682 (0.017)
LASSO	0.673 (0.014)	0.680 (0.018)
DT	0.620 (0.020)	0.625 (0.019)
RF	0.644 (0.013)	0.658 (0.018)
XGB	0.648 (0.013)	0.657 (0.014)
SVM	0.665 (0.020)	0.664 (0.014)
NB	0.602 (0.023)	0.572 (0.019)
KNN	0.606 (0.022)	0.650 (0.027)

**Tabla 1.** Valores medios (desviación estándar) de accuracy obtenidos por diferentes métodos supervisados para CAD. Considerando todas las variables (Sin FS) y las variables seleccionadas por los métodos FS (Con FS).

En la Tabla 2 se muestran los resultados obtenidos para la clasificación de DM. Se observa que RF tiene mejores resultados, con un accuracy de 96.5%. Actualmente, se ha visto que la mayoría de las personas que desarrollan una cardiopatía de este tipo es debido a un conjunto de factores que interactúan entre sí. De ahí, la importancia de que los facultativos sean capaces de predecir el riesgo de sufrir esa enfermedad, analizando un conjunto de factores, sin centrarse en un único factor de riesgo.

Método	Accuracy	
	Sin FS	Con FS
LR	0.903 (0.022)	0.913 (0.019)
Ridge	0.910 (0.021)	0.917 (0.018)
LASSO	0.917 (0.017)	0.913 (0.019)
DT	0.862 (0.015)	0.872 (0.019)
RF	<b>0.965 (0.010)</b>	0.942 (0.007)
XGB	0.952 (0.010)	0.927 (0.011)
SVM	0.962 (0.020)	0.953 (0.007)
NB	0.893 (0.023)	0.882 (0.023)
KNN	0.965 (0.014)	0.943 (0.006)

**Tabla 2.** Valores medios (desviación estándar) de accuracy obtenidos por diferentes métodos supervisados para DM. Considerando todas las variables (Sin FS) y las variables seleccionadas por los métodos FS (Con FS).

## 5. Conclusiones

En este trabajo se evaluaron diferentes métodos de FS para la identificación de factores de riesgo y técnicas de ML para predecir pacientes con CAD y DM. Para la CAD, los resultados muestran como la selección de variables mejora las predicciones. El modelo que mejores métricas obtiene es LR para los dos experimentos realizados con un accuracy de 68,3%. Para DM, el conjunto de datos con todas las variables logró mejores resultados con un RF alcanzando un accuracy de 96,5%.

## Agradecimientos

Este trabajo fue financiado por el Gobierno Español (AAVis-BMR PID2019-107768RA-I00); por el proyecto

Ref. 2020-661 y el proyecto I+D+i Jóvenes Investigadores Ref. F861 (AUTO-BA-GRAPH) financiado por la Comunidad de Madrid y Universidad Rey Juan Carlos.

## Referencias

- [1] Celestin Hategeka et al. Implementation research on noncommunicable disease prevention and control interventions in low-and middle-income countries: A systematic review. *PLoS Medicine*, 19(7):e1004055, 2022.
- [2] Joan B Soriano et al. La carga de enfermedad en España: resultados del estudio de la carga global de las enfermedades 2016. *Medicina Clinica*, 151(5):171–190, 2018.
- [3] Robert Wu et al. The next generation of clinical decision support: linking evidence to best practice. *Journal of Healthcare Information Management*, 16(4):50–55, 2002.
- [4] Riccardo Miotto et al. Deep learning for healthcare: review, opportunities and challenges. *Briefings in Bioinformatics*, 19(6):1236–1246, 2018.
- [5] David Chushig-Muzo et al. Learning and visualizing chronic latent representations using electronic health records. *BioData Mining*, 15(1):1–27, 2022.
- [6] Shahab Shamshirband et al. A review on deep learning approaches in healthcare systems: Taxonomies, challenges, and open issues. *Journal of Biomedical Informatics*, 113:103627, 2021.
- [7] MM Islam et al. Likelihood prediction of diabetes at early stage using data mining techniques. In *Computer Vision and Machine Intelligence in Medical Image Analysis*, pages 113–125. Springer, 2020.
- [8] Thomas R Dawber et al. Epidemiological approaches to heart disease: the framingham study. *American Journal of Public Health and the Nations Health*, 41(3):279–286, 1951.
- [9] Tammy Jiang et al. Supervised machine learning: a brief primer. *Behavior Therapy*, 51(5):675–687, 2020.
- [10] Beatriz Remeseiro and Veronica Bolon-Canedo. A review of feature selection methods in medical applications. *Computers in Biology and Medicine*, 112:103375, 2019.
- [11] Àlvar Hernández-Carnerero et al. Antimicrobial resistance prediction in intensive care unit for pseudomonas aeruginosa using temporal data-driven models. *International Journal of Interactive Multimedia and Artificial Intelligence*, 2021.
- [12] Joseph Bemister-Buffington et al. Machine learning to identify flexibility signatures of class a gpcr inhibition. *Biomolecules*, 10(3):454, 2020.
- [13] Payam Refaeilzadeh and et al. Cross-validation. *Encyclopedia of Database Systems*, 5:532–538, 2009.
- [14] Matloob Khushi et al. A comparative performance analysis of data resampling methods on imbalance medical data. *IEEE Access*, 9:109960–109975, 2021.

# Combinación de redes neuronales convolucionales y recurrentes para la detección de la apnea obstructiva del sueño en niños empleando las señales de flujo aéreo y oximetría

J. Jiménez García<sup>1,2</sup>, G.C. Gutiérrez Tobal<sup>1,2</sup>, M. García Gadañón<sup>1,2</sup>, F. Vaquerizo Villar<sup>1,2</sup>, D. Álvarez González<sup>1,2,3</sup>, F. del Campo<sup>3</sup>, L. Kheirandish-Gozal<sup>4</sup>, D. Gozal<sup>4</sup>, R. Hornero Sánchez<sup>1,2</sup>

<sup>1</sup> Grupo de Ingeniería Biomédica, Universidad de Valladolid, Valladolid, España, jorge.jimenez@gib.tel.uva.es

<sup>2</sup> Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), Instituto de Salud Carlos III, Valladolid, España

<sup>3</sup> Servicio de Neumología, Hospital Universitario Río Hortega, Valladolid, España, fsas@telefonica.net

<sup>4</sup> Dept. of Medical Pharmacology and Physiology, The University of Missouri School of Medicine, Columbia, Missouri, EEUU, gozald@health.missouri.edu

## Resumen

*El objetivo de este estudio fue evaluar la capacidad diagnóstica de un algoritmo de deep learning basado en la combinación de redes neuronales convolucionales y recurrentes para detectar la apnea obstructiva del sueño (AOS) en niños. Como fuentes de información se emplearon los registros nocturnos de flujo aéreo y oximetría, que se utilizaron para entrenar y validar un modelo de deep learning orientado a la estimación de la severidad de la AOS. La arquitectura del modelo está formada por sucesivas capas convolucionales, finalizado por una capa recurrente de tipo Gated Recurrent Unit (GRU). Se ha empleado una base de datos con 1.638 sujetos pediátricos repartidos aleatoriamente en conjuntos de entrenamiento, validación y test. El modelo de deep learning alcanzó una precisión del 73,20% en el conjunto de test al clasificar a los pacientes en los 4 niveles de severidad de la AOS. En los umbrales de severidad de 1, 5 y 10 eventos/hora se lograron precisiones del 86,27%, 93,46% y 93,14%, respectivamente. Estos resultados son superiores a los logrados en la literatura, mejorando también los alcanzados previamente empleando únicamente una arquitectura basada en una red neuronal convolucional. Como conclusión, la combinación de las arquitecturas convolucionales y recurrentes contribuyó a mejorar la detección de la AOS.*

## 1. Motivación

La apnea obstructiva del sueño (AOS) es un trastorno respiratorio que se caracteriza por interrupciones totales o parciales de la respiración (apneas/hipopneas) durante el sueño [1]. En el caso de la población infantil, la AOS afecta a nivel global a entre el 1% y el 5% de los niños, y puede provocar déficits neurocognitivos, diversos problemas conductuales y trastornos en el crecimiento y el desarrollo [1]. Para diagnosticarla, el método estándar de diagnóstico es la polisomnografía nocturna (PSG), que incluye el registro de señales neuromusculares y cardiorrespiratorias, entre otras. El índice de apnea-hipopnea (IAH) obtenido de la PSG es la tasa de eventos de apnea e hipopnea por hora de sueño (e/h), y se emplea comúnmente para determinar la severidad de la AOS [1]. Sin embargo, las limitaciones de la PSG (elevado coste y complejidad, incomodidad para los niños, baja disponibilidad de unidades del sueño especializadas) ha motivado la búsqueda de alternativas más sencillas [2]. Entre ellas, la reducción del número de

señales es la técnica más aceptada en la actualidad para simplificar el diagnóstico de la AOS. Las señales de flujo aéreo (FA) y oximetría (saturación de oxígeno en sangre, SpO<sub>2</sub>) forman parte de la PSG, y los eventos de apnea e hipopnea se definen principalmente atendiendo a la información que proporcionan [1]–[3]. Por ello, son dos de las señales más empleadas en las propuestas simplificadas, que lograrían mejorar el acceso al diagnóstico y reducir la complejidad y los costes asociados. Paralelamente, las técnicas de análisis automático han demostrado ser útiles para ayudar en el diagnóstico de la AOS tanto en adultos como en niños, simplificando la labor diagnóstica [4], [5].

Las técnicas de análisis automático de las señales procedentes de la PSG con el objetivo de detectar la AOS se han estudiado ampliamente en adultos. Los paradigmas más recientes se han basado en algoritmos de *deep learning* (DL), tanto redes neuronales convolucionales (CNN, *Convolutional Neural Network*) como recurrentes (RNN, *Recurrent Neural Network*) [6]. Algunas propuestas han abordado la combinación de CNN y RNN (CNN+RNN) para la detección de fases del sueño y eventos de apnea e hipopnea [7], [8]. Sin embargo, los estudios que han empleado técnicas de DL para la detección de la AOS infantil son aún muy limitados [5]. La mayoría de los estudios han analizado las señales de FA y SpO<sub>2</sub> combinando técnicas de extracción de características con algoritmos de aprendizaje automático tradicionales [5]. Recientemente se han aplicado algoritmos de DL basados en redes CNN aplicados sobre las señales de FA y SpO<sub>2</sub> para el diagnóstico de la AOS infantil [9], [10], pero no se ha abordado la combinación CNN+RNN en este contexto. La red RNN podría añadir a los patrones obtenidos por la CNN información complementaria sobre la recurrencia de los eventos apneicos.

La hipótesis de este estudio es que un algoritmo de DL basado en la combinación CNN+RNN puede ser útil en la detección automática de la AOS en niños empleando las señales de FA y SpO<sub>2</sub>. Por tanto, el objetivo de este estudio es entrenar y validar una red CNN+RNN orientada a establecer la presencia y severidad de la AOS mediante la estimación del IAH a partir de las señales de FA y SpO<sub>2</sub>.

## 2. Materiales y métodos

### 2.1. Sujetos y señales

En este estudio se empleó la base de datos pública del *Childhood Adenotonsillectomy Trial* (CHAT), proporcionada por el *National Sleep Research Resource* (<https://sleepdata.org/datasets/chat>) [11]. Esta base de datos está formada por 1.638 registros de PSG realizados a 1.232 niños entre 5 y 10 años, procedentes de un ensayo aleatorizado con el objetivo de evaluar la eficacia de un tratamiento quirúrgico para la AOS frente a un tratamiento conservador. En este estudio, los sujetos de la base de datos se dividieron de forma aleatoria en conjuntos de entrenamiento (60%), validación (20%) y test (20%) independientes entre sí. Los dos primeros conjuntos se emplearon para entrenar y optimizar el algoritmo de DL, y su rendimiento diagnóstico se evaluó sobre el conjunto de test. En la Tabla 1 se muestran los datos sociodemográficos y clínicos de los sujetos.

Las señales de FA adquiridas mediante un termistor y de SpO<sub>2</sub> procedentes de un sensor de pulsioximetría fueron extraídas de cada estudio de PSG. Después, se remuestrearon a una tasa común de 4 Hz con los objetivos de reducir la carga computacional del algoritmo y de lograr que ambas señales tuvieran la misma longitud de cara a ser procesadas por la red CNN+RNN de manera conjunta. Las señales de FA se filtraron con un filtro paso-bajo de ventana de Kaiser, con frecuencia de corte 1,5 Hz [10]. Las amplitudes de ambas señales se normalizaron restando su valor medio y dividiendo entre la desviación estándar.

Finalmente, las señales se segmentaron formando secuencias de 30 minutos, que resultaron las más adecuadas para el procesamiento de la red CNN+RNN. Estas secuencias de larga duración son aptas para poder analizar la distribución temporal de los eventos de apnea/hipopnea, que típicamente se producen en agrupaciones (clústeres) [9], [12]. Las secuencias están divididas en 6 épocas de 5 minutos cada una (300s, 1.200 muestras), una duración que se ha mostrado útil en el pasado para la estimación del IAH en niños empleando una CNN [10]. Por tanto, cada secuencia es una matriz con un formato de 6 épocas  $\times$  1200 muestras/época  $\times$  2 canales, etiquetada con la cantidad de eventos de apnea y/o hipopnea presentes en la misma ( $y_i$ )

	Entrenam.	Validación	Test
Sujetos (n)	1006	326	306
Edad (años)	7 [6; 8]	7 [6; 8]	6,9 [6; 8]
Varones (%)	46,8%	47,9%	43,8%
IMC (kg/m <sup>2</sup> )	17,4 [15,6; 21,7]	17,1 [15,4; 21,8]	17,6 [15,7; 21,7]
IAH (eventos/h)	2,6 [1,1; 5,9]	2,4 [1,2; 5,8]	2,3 [1,1; 6,2]
IAH<1 (n)	219	69	67
1≤IAH<5 (n)	496	168	148
5≤IAH<10 (n)	160	44	49
IAH≥10 (n)	131	45	42

Tabla 1. Datos sociodemográficos y clínicos de los sujetos.

[10], [13]. De esta forma, el entrenamiento de la arquitectura desarrollada se orientó a realizar una tarea de regresión del número de eventos apneicos.

### 2.2. Modelo de deep learning

El modelo CNN+RNN propuesto en este estudio está formado por  $NC$  sucesiones de capas convolucionales (*Conv2D*), de normalización (*Batch Normalization*), activación de tipo *Rectified Linear Unit* (ReLU), reducción de dimensionalidad (*Max Pooling*) y regularización (*Dropout*). Estas 5 capas son bloques que forman parte de una arquitectura CNN propuesta en trabajos anteriores [9], [10], y que en este estudio están orientadas al aprendizaje de características abstractas que se emplean en las capas recurrentes [8]. En la Figura 1 se muestra un esquema general del modelo de CNN+RNN propuesto. En las capas convolucionales se implementó un número de filtros 2D ( $NF$ ) en cada capa, cada uno de ellos de un tamaño  $9 \times 2$ . De acuerdo con estudios previos, se fijaron un *max pooling* 2D de  $2 \times 1$  y una tasa de *dropout*  $p=0,1$  [10]. Después de los bloques CNN se implementó una capa de tipo *Global Average Pooling* (GAP), que reduce la dimensionalidad calculando el promedio de cada uno de los  $NF$  canales de salida [7]. Todas las capas descritas anteriormente están albergadas dentro de capas de tipo *Time Distributed* (TD) que procesan los datos de cada época de la secuencia individualmente (Figura 1). De esta forma, una arquitectura de sucesivas capas TD conserva el formato de secuencia de los datos a medida que se procesan en las capas de la CNN. Una capa de *Dropout Gaussiano* ( $p=0,3$ ) aplica ruido gaussiano a las características aprendidas durante el entrenamiento para reducir el sobreajuste [7]. La última capa del modelo propuesto es una RNN formada por unidades de tipo *Gated Recurrent Unit* (GRU) bidireccional con una dimensión  $NG$ , que procesa las secuencias de características aprendidas en ambos sentidos para predecir el número de eventos de apnea/hipopnea

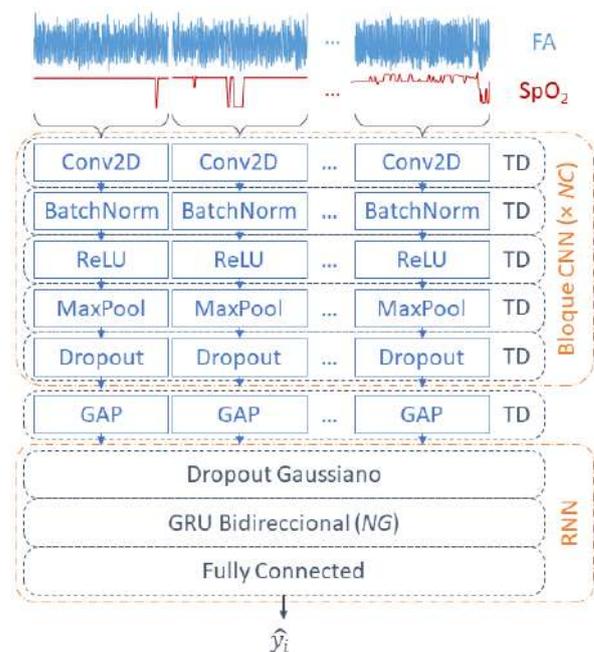


Figura 1. Esquema de la arquitectura CNN+RNN empleada.

presentes mediante regresión ( $\hat{y}_i$ ) [13]. Se evaluaron distintos valores de los hiperparámetros  $NC$ ,  $NF$  y  $NG$  del modelo CNN+RNN con el objetivo de maximizar el rendimiento diagnóstico en el conjunto de validación.

La arquitectura CNN+RNN se implementó empleando las librerías *Keras* y *Tensorflow*. El algoritmo de optimización Adam empleado durante entrenamiento de la arquitectura CNN+RNN se orientó a minimizar el error empleando la función de pérdidas de Huber, utilizada con frecuencia en tareas de regresión robusta. Se fijó una tasa de aprendizaje inicial de 0,0001 y un tamaño de bloque de 128 ejemplos.

### 2.3. Estimación del IAH

Una vez estimado el número de eventos de apnea/hipopnea en cada segmento de 30 minutos, se calculó el IAH del registro nocturno. Para ello se obtuvo el número de eventos estimado en todos los segmentos del registro dividido por su duración en horas ( $IAH_{CNN+RNN}$ ). Esta duración es mayor que el tiempo total de sueño, por lo que este cálculo infraestima el IAH. Para compensar este sesgo, la estimación del IAH final ( $IAH_{EST}$ ) se obtuvo mediante una regresión lineal [9], [10], fijando en el conjunto de entrenamiento los coeficientes de sesgo ( $a_0$ ) y pendiente ( $a_1$ ) de la ecuación:

$$IAH_{EST} = a_0 + a_1 \cdot IAH_{CNN+RNN} \quad (1)$$

### 2.4. Análisis estadístico

La concordancia entre el IAH estimado y el IAH real de cada registro se evaluó mediante el coeficiente de correlación intraclass ( $ICC$ , *Intraclass Correlation Coefficient*). El rendimiento diagnóstico del modelo CNN+RNN se evaluó en los umbrales del IAH habituales en niños (1, 5 y 10 e/h para AOS leve, moderado y severo, respectivamente), empleando la matriz de confusión y la precisión de 4 clases ( $Acc_4$ ), así como la precisión ( $Acc$ ), la sensibilidad ( $Se$ ), la especificidad ( $Sp$ ) y los valores predictivos positivo y negativo ( $VPP$  y  $VPN$ , respectivamente) para la clasificación binaria.

## 3. Resultados

Los valores óptimos de los hiperparámetros fueron  $NC=7$ ,  $NF=64$  y  $NG=16$ . Valores más altos de estos hiperparámetros no mejoraron el rendimiento diagnóstico en el conjunto de validación, a la vez que incrementaron el sobreajuste del modelo a los datos de entrenamiento. En la Figura 2 se muestra el gráfico de Bland-Altman correspondiente a la estimación del IAH mediante el modelo CNN+RNN frente al IAH real. La arquitectura CNN+RNN tiene una ligera tendencia a infraestimar el IAH de los sujetos (-0,2271 e/h de media), mientras que el  $ICC=0,9354$  muestra una elevada concordancia. La Figura 3 muestra la matriz de confusión obtenida en la clasificación en los 4 niveles de severidad. La Tabla 2 muestra los resultados de clasificación binaria en los tres umbrales habituales del IAH alcanzados en el conjunto de test. Se logró una  $Acc_4=73,20\%$ , con 42 casos (13,73%) en los que se infraestimó el nivel de severidad frente a 40 (13,07%) en los que se asignó una severidad mayor.

## 4. Discusión y conclusiones

En este estudio se ha desarrollado y evaluado un modelo de DL basado en una arquitectura combinada CNN+RNN para estimar la severidad de la AOS en niños empleando las señales de FA y SpO<sub>2</sub>. Para ello, se ha diseñado, entrenado y validado un modelo que combina capas convolucionales y recurrentes dirigido hacia la estimación del IAH. El algoritmo propuesto logró una elevada concordancia en la predicción del IAH ( $ICC=0,9354$ ), que se trasladó también a una elevada precisión en la estimación del nivel de severidad ( $Acc_4=73,20\%$ ). Además, la capacidad diagnóstica del modelo CNN+RNN alcanzó niveles de  $Acc$  superiores al 85% en cada uno de los umbrales, mientras que la  $Se$  fue superior al 75% en todos

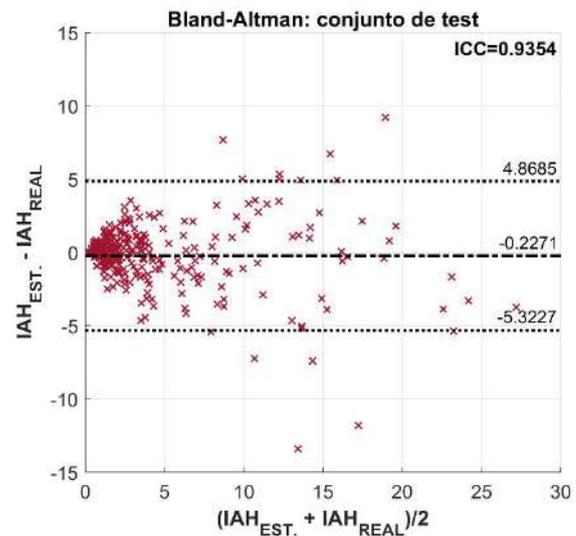


Figura 2. Gráfico de Bland-Altman de la estimación del IAH.

		CNN+RNN			
Severidad AOS real	0	42 0.63	25 0.37	0 0.00	0 0.00
	1	17 0.11	126 0.85	4 0.03	1 0.01
	2	0 0.00	15 0.31	24 0.49	10 0.20
	3	0 0.00	0 0.00	10 0.24	32 0.76
		0	1	2	3
		Severidad AOS est			

Figura 3. Matriz de confusión del modelo CNN+RNN para la clasificación en 4 clases en el conjunto de test (0: No AOS; 1: AOS leve; 2: AOS moderado; 3: AOS severo).

	$Se$	$Sp$	$Acc$	$VPP$	$VPN$
1 e/h	92,89	62,69	86,27	89,88	71,19
5 e/h	83,52	97,67	93,46	93,83	93,33
10 e/h	76,19	95,83	93,14	74,42	96,20

Tabla 2. Resultados de clasificación binaria del modelo CNN+RNN. Valores expresados en porcentaje.

los casos. Por tanto, la arquitectura CNN+RNN logró modelar la información de las señales de FA y SpO<sub>2</sub> y analizar su dependencia temporal a lo largo de secuencias de larga duración. Solo dos estudios previos basados en DL aplicaron una arquitectura basada únicamente en CNN para estimar el IAH: a partir de la señal de SpO<sub>2</sub> [9], así como de la combinación FA y SpO<sub>2</sub> [10]. En este sentido, la aplicación de una arquitectura CNN+RNN es una contribución novedosa en el contexto de la ayuda al diagnóstico de la AOS infantil. En comparación con el estudio basado en FA y SpO<sub>2</sub> que utilizó la misma base de datos e idéntica partición [10], este trabajo alcanzó una *Acc*<sub>4</sub> ligeramente superior (73,20% vs. 72,55%) y mejoró la *Acc* en 1 e/h (86,27% vs. 84,64%) a la vez que logró además pares *Se-Sp* más equilibrados, con *Acc* similares en 5 y 10 e/h al comparar con el modelo CNN [10]. Por lo tanto, la capacidad diagnóstica del modelo CNN+RNN supera ligeramente a la del modelo CNN.

Otros estudios han analizado las señales de FA y SpO<sub>2</sub> con el objetivo de detectar la AOS empleando métodos clásicos de aprendizaje automático. El análisis *wavelet* de la señal de FA combinado con el índice de desaturación de SpO<sub>2</sub> alcanzó valores de *Acc* entre 77,97% y 90,99% [14]. De manera similar, una metodología basada en la aplicación del *bispectrum* sobre la señal de FA logró valores de *Acc* entre 82,16% y 90,15% [15]. Esto indicaría que la capacidad diagnóstica lograda en este estudio supera la mostrada en estudios previos.

Este estudio tiene ciertas limitaciones. En primer lugar, el algoritmo se ha desarrollado y validado en una única base de datos. Sería deseable evaluar la capacidad de generalización de una manera más exhaustiva empleando otras bases de datos de AOS pediátrico. Además, la interpretabilidad del modelo CNN+RNN es muy baja, por lo que sería adecuado aplicar técnicas de *Explainable Artificial Intelligence* para identificar la información de las señales que el modelo emplea específicamente para la detección de la AOS.

En conclusión, una arquitectura de DL basada en CNN+RNN ha demostrado ser útil en la detección de la AOS infantil empleando señales de FA y SpO<sub>2</sub>. El rendimiento diagnóstico superó ligeramente al de estudios previos centrados en CNN. Estos resultados sugieren que el modelo CNN+RNN es capaz de analizar información útil sobre la distribución temporal de las apneas/hipopneas presentes en las señales de FA y SpO<sub>2</sub> para detectar la AOS en niños.

## Agradecimientos

Este estudio ha sido financiado por el MCIN/AEI 10.13039/501100011033, FEDER y por NextGeneration EU/PRTR (PID2020-115468RB-I00 PDC2021-120775-I00), por el CIBER-BBN ISCIII-MCIN (CB19/01/00012), Sociedad Española de Neumología y Cirugía Torácica (SEPAR 649/2018) y por la Sociedad Española de Sueño (Beca de Investigación SES 2019). D. Álvarez cuenta con una ayuda Ramón y Cajal (RYC2019-028566-I) del MCIN/AEI y el FSE. J. Jiménez cuenta con un contrato predoctoral de la Universidad de Valladolid.

## Referencias

- [1] E. Dehlink *et al.*, “Update on paediatric obstructive sleep apnoea,” *J Thorac Dis*, vol. 8, no. 2, pp. 224–235, 2016, doi: 10.3978/j.issn.2072-1439.2015.12.04.
- [2] H.-L. Tan *et al.*, “Overnight Polysomnography versus Respiratory Polygraphy in the Diagnosis of Pediatric Obstructive Sleep Apnea,” *Sleep*, vol. 37, no. 2, pp. 255–260, 2014, doi: 10.5665/sleep.3392.
- [3] R. B. Berry *et al.*, “Rules for Scoring Respiratory Events in Sleep: Update of the 2007 AASM Manual for the Scoring of Sleep and Associated Events,” *J. Clin. Sleep Med.*, vol. 8, no. 5, pp. 597–619, Oct. 2012, doi: 10.5664/jcsm.2172.
- [4] F. Mendonca *et al.*, “A Review of Obstructive Sleep Apnea Detection Approaches,” *IEEE J. Biomed. Heal. Informatics*, vol. 23, no. 2, pp. 825–837, Mar. 2019, doi: 10.1109/JBHI.2018.2823265.
- [5] G. C. Gutiérrez-Tobal *et al.*, “Reliability of machine learning to diagnose pediatric obstructive sleep apnea: Systematic review and meta-analysis,” *Pediatr. Pulmonol.*, vol. 57, no. 8, pp. 1931–1943, Aug. 2022, doi: 10.1002/ppul.25423.
- [6] S. S. Mostafa *et al.*, “A systematic review of detecting sleep apnea using deep learning,” *Sensors (Switzerland)*, vol. 19, no. 22, pp. 1–26, 2019, doi: 10.3390/s19224934.
- [7] H. Korkalainen *et al.*, “Accurate Deep Learning-Based Sleep Staging in a Clinical Population with Suspected Obstructive Sleep Apnea,” *IEEE J. Biomed. Heal. Informatics*, vol. 24, no. 7, pp. 2073–2081, 2020, doi: 10.1109/JBHI.2019.2951346.
- [8] F. Vaquerizo-Villar *et al.*, “Modelo de deep learning basado en la combinación de redes neuronales convolucionales y recurrentes para clasificar eventos de apnea e hipopnea mediante la señal de oximetría,” in *XXXIX Congreso Anual de la Sociedad Española de Ingeniería Biomédica*, 2021, pp. 63–66.
- [9] F. Vaquerizo-Villar *et al.*, “A Convolutional Neural Network Architecture to Enhance Oximetry Ability to Diagnose Pediatric Obstructive Sleep Apnea,” *IEEE J. Biomed. Heal. Informatics*, vol. 25, no. 8, pp. 2906–2916, Aug. 2021, doi: 10.1109/JBHI.2020.3048901.
- [10] J. Jiménez-García *et al.*, “A 2D convolutional neural network to detect sleep apnea in children using airflow and oximetry,” *Comput. Biol. Med.*, vol. 147, no. August, p. 105784, Aug. 2022, doi: 10.1016/j.combiomed.2022.105784.
- [11] C. L. Marcus *et al.*, “A Randomized Trial of Adenotonsillectomy for Childhood Sleep Apnea,” *N. Engl. J. Med.*, vol. 368, no. 25, pp. 2366–2376, 2013, doi: 10.1056/nejmoa1215881.
- [12] R. T. Brouillette *et al.*, “Nocturnal Pulse Oximetry as an Abbreviated Testing Modality for Pediatric Obstructive Sleep Apnea,” *Pediatrics*, vol. 105, no. 2, pp. 405–412, Feb. 2000, doi: 10.1542/peds.105.2.405.
- [13] S. Nikkonen *et al.*, “Artificial neural network analysis of the oxygen saturation signal enables accurate diagnostics of sleep apnea,” *Sci. Rep.*, vol. 9, no. 1, pp. 1–9, 2019, doi: 10.1038/s41598-019-49330-7.
- [14] V. Barroso-García *et al.*, “Wavelet analysis of overnight airflow to detect obstructive sleep apnea in children,” *Sensors*, vol. 21, no. 4, pp. 1–19, 2021, doi: 10.3390/s21041491.
- [15] V. Barroso-García *et al.*, “Bispectral analysis of overnight airflow to improve the pediatric sleep apnea diagnosis,” *Comput. Biol. Med.*, vol. 129, no. August 2020, 2021, doi: 10.1016/j.combiomed.2020.104167.

# Clasificación de ruido clínico en ECG de monitorización prolongada mediante máquinas de vectores soporte

R. Holgado–Cuadrado<sup>1\*</sup>, C. Plaza–Seco<sup>1</sup>, L. Lovisololo<sup>1</sup> y M. Blanco–Velasco<sup>1</sup>

<sup>1</sup> Departamento de Teoría de la Señal y Comunicaciones, Universidad de Alcalá, {roberto.holgado, carmen.plazas, manuel.blanco, lisandro.lovisolo}@uah.es

## Resumen

*El ruido y sus artefactos afectan en gran medida a la calidad del electrocardiograma (ECG) en registros de monitorización prolongada, haciendo que algunas partes sean inservibles para el diagnóstico clínico. La definición de ruido clínico establece una métrica de calidad que, en lugar de evaluar el ruido desde un punto de vista cuantitativo, indica qué fragmentos de un ECG son válidos para realizar un diagnóstico clínico. Este trabajo propone el uso de modelos máquinas de vectores soporte para clasificar ruido clínico en distintas categorías. Se utiliza una base de datos anotada según una taxonomía definida específicamente para ruido clínico. El principal objetivo es resolver un problema de clasificación binaria basado en identificar segmentos de ECG válidos para el diagnóstico clínico. Para ello, se desarrolla una metodología que tiene en cuenta el equilibrio de las clases, la separación de los pacientes y la permutación de los mismos en el conjunto de test para evitar sobreajuste. Con el sistema de aprendizaje propuesto se obtiene un buen rendimiento de clasificación, alcanzando un F1-score de 0,74 en el conjunto de test. El método propuesto, que proporciona una solución de clasificación para evaluar la calidad clínica del ECG, puede ayudar a los profesionales a discriminar tramas de ECG clínicamente interpretables en los registros de monitorización prolongada.*

## 1. Introducción

El electrocardiograma (ECG) ha demostrado ser una herramienta eficaz y fiable para estudiar la actividad eléctrica del corazón [1]. Aunque existen distintos tipos de registros de ECG, ante la sospecha de la presencia de una patología intermitente, es necesario monitorizar la actividad cardíaca durante periodos prolongados, debido a las lesiones irreversibles que esta podría causar. Dichos registros se denominan de monitorización prolongada (*Long-Term Monitoring*, LTM) y se adquieren durante 7, 15 o incluso 21 días, de manera ambulatoria. Estos registros LTM son útiles para detectar numerosas patologías cardíacas [2]. Sin embargo, su larga duración da lugar a que algunas partes del ECG estén muy contaminadas por distintos tipos de ruido, como el ruido de la red eléctrica, los movimientos del paciente o de los electrodos (entre otros), por lo que parte de estos registros resultan ineficaces para el diagnóstico. Por ello, la identificación de segmentos útiles desde un punto de vista clínico sería de gran beneficio.

La severidad clínica del ruido se define como el impacto que el ruido causa en la validez clínica de los segmentos del ECG para establecer un diagnóstico. En este contex-

to, se dice que un fragmento de ECG posee calidad clínica cuando contiene información útil para el diagnóstico. En [3], se demostró que la severidad clínica y el enfoque tradicional cuantitativo del ruido no están correlacionados, por lo que la evaluación de la calidad clínica de un registro LTM no puede limitarse a estimar cuantitativamente su nivel de corrupción de ruido.

En este trabajo se aborda un enfoque novedoso de análisis de ruido en ECG en términos de interpretabilidad clínica para averiguar si es posible la distinción entre segmentos clínicamente válidos mediante técnicas de aprendizaje automático o *Machine Learning* (ML). Nuestro principal objetivo es determinar si estos sistemas pueden utilizarse para reconocer patrones de ruido clínico, abordando un problema de clasificación binaria. Se lleva a cabo una metodología basada en máquinas de vectores soporte (*Support Vector Machine*, SVM) [4], en la que se utilizan como predictores un conjunto de características que describen la forma de onda y el contenido informativo del ECG.

## 2. Base de datos y predictores

### 2.1. Base de datos

Un trabajo introductorio de ruido clínico en el ECG se presentó inicialmente en [3] con el propósito de establecer una taxonomía en términos de interpretabilidad clínica. Así, se definieron 5 clases de severidad clínica del ruido [3]: *Noise-free* o tipo 0 (T0) para segmentos sin ruido; *Low noise* o tipo 1 (T1) si existe algún ruido presente en el segmento, pero las ondas P, T y los complejos QRS son legibles y se puede identificar su morfología; *Moderate noise* o tipo 2 (T2) si sólo los complejos QRS se identifican de forma fiable, en al menos tres latidos consecutivos; *Hard noise* o tipo 3 (T3) si los complejos QRS son difícilmente reconocibles o irreconocibles; y *Other noise* o tipo 4 (T4) si hay ausencia total de la señal.

Según estos criterios, se ha etiquetado un conjunto de datos adquiridos ( $f_s = 200\text{Hz}$ ) con un registrador externo de eventos (*External Event Recorded*, EER) de manera continuada a lo largo de toda su duración, resultando en más de 6,5 horas de registro de 10 pacientes con 2 derivaciones. La Figura 1 muestra un fragmento de una de estas grabaciones

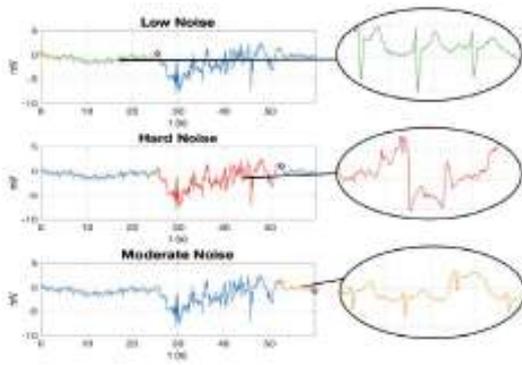


Figura 1: Fragmentos etiquetado de los datos originales procedentes del EER.

con una transición entre las clases T1-T3 y T3-T2.

El enfoque propuesto en este trabajo se basa en distinguir únicamente dos clases: tramas limpias de tramas ruidosas, a pesar de que el ruido clínico se divide en varias categorías. Se considera este criterio basándonos en que el interés clínico es determinar si las tramas de ECG son clínicamente válidas o no para hacer un diagnóstico. Así, las cinco categorías iniciales se reducen a dos, conformando la clase *limpia* con las categorías T0 y T1 y la clase *ruidosa* con T2 y T3. La quinta T4, no se tiene en cuenta porque se corresponde con segmentos sin señal fácilmente identificables con técnicas sencillas de procesamiento de señales.

Para diseñar los modelos de aprendizaje, la base de datos del EER se segmenta mediante un procedimiento que consta de dos etapas. En primer lugar, los registros se separan por categorías a partir de las etiquetas, de forma que se obtiene un conjunto de segmentos de diferente longitud, cada uno referente a una sola clase. En el ejemplo de la Figura 1, el primer paso de la segmentación generaría 3 segmentos de diferente longitud y clase. Los sistemas se entrenan y evalúan con información de segmentos de la misma longitud, por lo que la segunda etapa extrae bloques de la misma longitud a partir de los segmentos anteriores,  $s(t)$ ,  $0 \leq t \leq t_b$ . Cada segmento puede considerarse como un conjunto de  $L$  muestras empaquetadas en el vector

$$\mathbf{s}^{(i)} = [s_0^{(i)}, s_1^{(i)} \cdots, s_{L-1}^{(i)}]^T, \quad i = 1, \dots, m \quad (1)$$

donde  $m$  es el número de instancias totales de la base de datos. En este trabajo, hemos tomado una longitud de  $t_b = 5$  segundos (s) que garantiza la presencia de, al menos, 3 latidos del ECG, en concordancia con la definición de ruido dada para la clase T2 [3].

Analizando los datos por pacientes se observa un fuerte desequilibrio entre las clases, como puede verse en la Figura 2 para los pacientes 1, 3, 5 y 8. Estos individuos pueden ser inadecuados para el proceso de entrenamiento y validación, debido a la dificultad que pueden tener los modelos para aprender de ellos. Por esta razón, hemos decidido descartar a los pacientes que presentan menos del 25% de instancias en cualquier clase, lo que reduce el conjunto a los 6 pacientes restantes. Finalmente, la base de datos está formada por  $m = 2818$  instancias, donde un 47% se corresponde con ejemplos de la clase *limpia* y el 53%

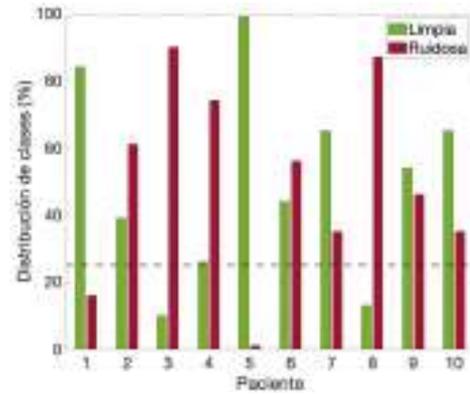


Figura 2: Distribución de las clases *limpia* y *ruidosa* por paciente.

restante con ejemplos de la clase *ruidosa*. El equilibrio entre ambas clases es razonable, aunque no lo es tanto si se analiza en las categorías originales (T0-T1 y T2-T3). Este desequilibrio podría tener un impacto significativo en el proceso de entrenamiento, ya que puede dar lugar a un mayor error de clasificación para las clases menos frecuentes.

## 2.2. Extracción de características

Como resultado del proceso de segmentación de la base de datos, se ha obtenido un conjunto formado por 2818 bloques de ECG de 5 s de duración. A todos ellos se les aplica un preprocesamiento básico donde se elimina la línea de base, la interferencia de la línea eléctrica y se aplica un filtro paso banda de  $0,5 - 40Hz$  para preservar la banda espectral del ECG. Cada bloque  $\mathbf{s}^{(i)}$  se normaliza:

$$\mathbf{s}^{(i)} := \frac{\mathbf{s}^{(i)} - \mu_s^{(i)}}{\sigma_s^{(i)}} \quad (2)$$

donde  $\mu_s^{(i)}$  y  $\sigma_s^{(i)}$  representan la media y la desviación estándar respectivamente, del  $i$ -ésimo bloque de señal.

Este trabajo se afronta como un problema de clasificación a partir de múltiples índices de calidad de la señal  $\mathbf{x} = [x_1, x_2, x_3, \dots, x_n]^T \in \mathbb{R}^n$ . Se seleccionaron dieciséis características ( $n = 16$ ) para representar la estructura de cada bloque. Las variables  $x_j$  se enumeran en la Tabla 1 y se corresponden a características que informan sobre: la posición y distribución de los puntos críticos ( $j = 1, 2$ ); características en el dominio del tiempo ( $j = 3, 4$ ); propiedades estadísticas ( $j = 5, \dots, 11$ ); y características espectrales ( $j = 12, \dots, 16$ ). Algunas de ellas se inspiran en [5]. Otras, como  $x_5$  hasta  $x_8$ , y  $x_{11}$ , son estadísticos bien conocidos.

Las características  $\mathbf{x}^{(i)} = [x_1^{(i)}, x_2^{(i)}, \dots, x_n^{(i)}]^T$  extraídas de cada bloque de ECG, conforman un conjunto de datos de dimensión  $\mathbf{X} \in \mathbb{R}^{m \times n}$ , donde  $m = 2818$  y  $n = 16$ . Cada fila de la matriz corresponde a un ejemplo asociado a una de las dos categorías (*limpia* o *ruidosa*) y, por tanto, las categorías del conjunto de datos configuran el vector de etiquetas  $\mathbf{y} = [y^{(1)}, y^{(2)}, \dots, y^{(m)}]^T$ . Así, la  $i$ -ésima instancia viene dada por el par:  $(\mathbf{x}^{(i)}, y^{(i)})$ . Las características se normalizan a partir de  $z$ -score con la media y la desviación estándar de cada  $x_j$  del conjunto del entrenamiento.

Tabla 1: Lista de las características y su descripción.

Característica	Descripción
$x_1$	Precisión en la detección de la onda R [5]
$x_2$	Periodicidad basada en la correlación
$x_3$	Comprobación de la línea de base [5]
$x_4$	Amplitud relativa del ruido de alta frecuencia [5]
$x_5$	Media
$x_6$	Desviación estándar
$x_7$	<i>Skewness</i>
$x_8$	<i>Kurtosis</i>
$x_9$	Desviación relativa del QRS [5]
$x_{10}$	Índice de Pureza Espectral [6]
$x_{11}$	Entropía
$x_{12}$	Potencia relativa de la línea de base [5]
$x_{13}$	Potencia relativa del QRS [5]
$x_{14}$	Energía relativa del QRS [5]
$x_{15}$	Frecuencia central
$x_{16}$	Índice de calidad del ECG [7]

### 3. Métodos

#### 3.1. SVM

Las SVM se han convertido en uno de los métodos de aprendizaje supervisado más destacados para resolver problemas de clasificación. Cabe destacar que este algoritmo es capaz de reducir los problemas de sobreajuste presentes en otros clasificadores, ya que, para problemas binarios, trata de separar óptimamente un conjunto de datos representados en el espacio mediante un hiperplano. Para separar dos clases de datos puede haber muchos hiperplanos posibles, pero SVM trata de encontrar el que está a la mayor distancia entre los ejemplos de ambas clases para maximizar la capacidad de generalización del modelo. Por lo tanto, el modelo que se obtiene resulta del problema de optimización:

$$\begin{aligned} \underset{\mathbf{w}, b, \zeta}{\text{minimizar}} \quad & \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^n \zeta_i, \\ \text{sujeto a} \quad & y_i(\mathbf{w}^T + b) \geq 1 - \zeta_i, \\ & \zeta_i \geq 0, \end{aligned}$$

donde  $C$  es el parámetro de regularización a seleccionar para aumentar o disminuir de manera controlada la generalización del clasificador.

#### 3.2. Diseño experimental

La variabilidad interpaciente (entre distintos pacientes) en el ECG es bien conocida en cardiología, por lo que cualquier modelo aprenderá las características propias del ECG asociadas a los pacientes que estén en el conjunto de entrenamiento. Como resultado, se puede producir el riesgo de sobrestimar el rendimiento del sistema si estos pacientes también se han incluido en el conjunto de test, causando sobreajuste al paciente. En nuestro caso, este fenómeno puede tener un impacto significativo, ya que sólo disponemos de 6 pacientes. Para evitarlo, no se van a asignar pacientes del conjunto de entrenamiento al conjunto de test.

La Figura 3 representa el esquema de validación del sistema de aprendizaje automático propuesto, donde un paciente se asigna al grupo de test, mientras que los cinco restantes se utilizan para entrenar y validar el modelo. Ca-

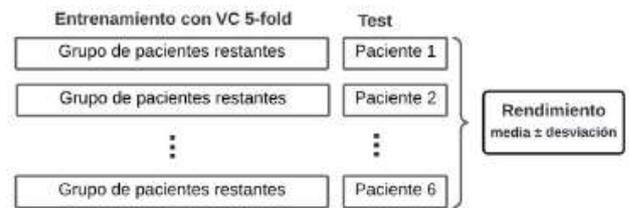


Figura 3: Esquema de validación.

da combinación se utiliza para diseñar un modelo, cuyos hiperparámetros se ajustan utilizando una validación cruzada (VC) de 5 *folds* sobre el 80 % aleatorio del grupo de pacientes de entrenamiento. El otro 20 % restante se utiliza para proporcionar una métrica en un conjunto de validación. Por otro lado, el conjunto de test lo conforman datos del único paciente que no se ha utilizado para el entrenamiento. Así, se evalúan un total de 6 modelos diferentes. El rendimiento global estimado se calcula como la media de todas las combinaciones, junto con la desviación estándar, que se utiliza como indicador de consistencia. Siguiendo esta configuración, la escasa disponibilidad de pacientes se compensa con el diseño de varios modelos que siguen una metodología común para minimizar el sobreajuste.

La evaluación del rendimiento de los modelos se realiza en términos de detección en la clase *limpia* y se emplean cinco métricas estándar ampliamente conocidas: *accuracy* (*Acc*), *recall* (*Re*), *precision* (*Pr*) y *F1-score*.

### 4. Resultados

Para la evaluación de los 6 modelos de SVM se han obtenido las métricas de evaluación en tres conjuntos diferentes: entrenamiento, validación y test, reflejadas en la Tabla 2. El rendimiento de la clasificación en los conjuntos de entrenamiento y validación supera al del conjunto de test. Esto confirma que las características aprendidas de un grupo de pacientes durante el entrenamiento del modelo sobrestiman el rendimiento del sistema cuando se utilizan los mismos pacientes en el conjunto de test. Esto puede interpretarse como un ejemplo de la importancia que tiene evitar el sobreajuste al paciente. En cuanto a la variabilidad de las métricas, se puede observar que la desviación estándar es significativamente mayor en el conjunto de test que en los conjuntos de entrenamiento y validación. Esto, se puede observar de manera más significativa en la Tabla 2, donde se detalla el rendimiento de los distintos modelos SVM cuando cada paciente se encuentra en el conjunto de test. Los valores mínimos y máximos de *F1-score* se obtienen para los modelos evaluados con el paciente 6 (0,42) y el paciente 4 (0,93) respectivamente. El rendimiento menor del paciente 6 podría ser indicativo de que tiene características diferentes que no se aprenden de los otros cinco, lo que hace que el sistema tenga mayor dificultad para clasificar sus instancias correctamente. Por el contrario, casi todos los bloques del paciente 4 se clasifican correctamente, lo que indica, o bien que los otros 5 proporcionan información suficientemente representativa, o bien que el paciente tiene características fáciles de aprender.

Tabla 2: Resultados de clasificación con SVM. Se indica el valor medio y la desviación estándar de los 6 modelos diseñados.

Conjunto	Acc	Re	Pr	F1-score
Entrenamiento	0.82 ± 0.03	0.80 ± 0.04	0.81 ± 0.03	0.81 ± 0.03
Validación	0.82 ± 0.03	0.80 ± 0.02	0.81 ± 0.03	0.80 ± 0.02
Test	0.73 ± 0.20	0.75 ± 0.17	0.75 ± 0.22	0.74 ± 0.18

Tabla 3: Resultados de clasificación detallados por paciente cuando se considera como conjunto de test.

Paciente Test	Acc	Re	Pr	F1-score
2	0.75	0.67	0.68	0.67
4	0.97	0.94	0.93	0.93
6	0.37	0.53	0.35	0.42
7	0.75	0.98	0.73	0.83
9	0.80	0.73	0.88	0.80
10	0.75	0.66	0.94	0.78

## 5. Discusión y conclusiones

Existen varios trabajos que abordan el problema del ruido en el ECG utilizando técnicas de ML para evaluar la calidad del ECG. Sin embargo, en la mayoría de los casos se realiza una caracterización del ruido en términos de severidad cuantitativa, donde, o bien no definen una escala de calidad o bien la validación se realiza sobre una base de datos con ruido sintético. En este trabajo se propone la caracterización del ruido según una escala de severidad clínica [3], utilizando para ello una base de datos anotada manualmente desde una perspectiva clínica. La principal diferencia de este trabajo con respecto a los anteriores es que tratamos con datos de LTM adquiridos en condiciones reales y etiquetados de forma continua.

El tamaño de la base de datos disponible es una limitación, porque comprende una duración de aproximadamente 6,5 horas con sólo 10 pacientes. Esta restricción se ve afectada además por el proceso de segmentación que hemos efectuado con el objeto de obtener un conjunto de datos adecuado para diseñar los modelos de aprendizaje, implicando una reducción tanto del número de datos como de pacientes. Sin embargo, el mayor valor a tener en cuenta reside en sus anotaciones, obtenidas a partir del etiquetado continuo de las señales. Así, aunque la información disponible es limitada, la aplicación de una metodología sólida apoya la hipótesis de que las clases de ruido clínico pueden distinguirse mediante técnicas de ML. Las limitaciones comentadas también tienen impacto en el rendimiento de los modelos dando lugar a resultados moderados a altos correspondientes con un 0,75 tanto en *recall* como en *precision*. A pesar de ello, estos valores muestran que el uso de sistemas basados en aprendizaje automático en, concretamente, modelos SVM, para resolver un problema de clasificación de ruido clínico en ECG. Esto ha sido posible gracias al uso una metodología robusta que tiene en cuenta el equilibrio de clases y la variabilidad interpaciente.

La aplicación de estos sistemas en el ámbito clínico está ligada a entender mejor su funcionamiento y mejorar la confianza de los facultativos en las soluciones que ofrecen. Para ello, está cobrando gran interés la explotación de arquitecturas que proporcionen una mejor interpretabilidad de los resultados. Una alternativa puede orientarse mediante el uso de autoencoders [8] al permitir proyectar espacios

de alta dimensionalidad en un conjunto de características observables en mapas de dos o tres dimensiones. Por otro lado, la mejora del presente trabajo también está sujeta a la exploración con otras técnicas de ML e incluso metodologías más avanzadas. Métodos como *Active Learning* pueden desempeñar un papel importante para la incorporación de nuevos datos, que, a su vez, permita superar la tasa de detección de los sistemas de aprendizaje.

## Agradecimientos

Queremos agradecer a los autores de [3] su contribución al desarrollo de la base de datos de ruido clínico. Además, este trabajo ha sido parcialmente financiado por la Comunidad de Madrid con el proyecto EPU-INV/2020/002.

## Referencias

- [1] L. M. Buja and J. Butany, *Cardiovascular Pathology*. San Diego: Academic Press, Fourth ed., 2016.
- [2] F. J. Pastor-Pérez, S. Manzano-Fernández, R. Goya-Esteban, D. A. Pascual-Figal, *et al.*, “Comparison of detection of arrhythmias in patients with chronic heart failure secondary to non-ischemic versus ischemic cardiomyopathy by 1 versus 7-day holter monitoring,” *The American Journal of Cardiology*, vol. 106, no. 5, pp. 677–681, 2010.
- [3] E. Everss-Villalba, F. M. Melgarejo-Meseguer, M. Blanco-Velasco, F. Gimeno-Blanes, *et al.*, “Noise maps for quantitative and clinical severity towards long-term ECG monitoring,” *Sensors*, vol. 17, no. 11, p. 2448, 2017.
- [4] T. Hastie, R. Tibshirani, and J. Friedman, *The elements of statistical learning: data mining, inference and prediction*. Springer, 2 ed., 2009.
- [5] Q. Li, C. Rajagopalan, and G. D. Clifford, “A machine learning approach to multi-level ECG signal quality classification,” *Computer Methods and Programs in Biomedicine*, vol. 117, no. 3, pp. 435–447, 2014.
- [6] M. Blanco-Velasco, F. Cruz-Roldán, J. I. Godino-Llorente, and K. E. Barner, “Nonlinear Trend Estimation of the Ventricular Repolarization Segment for T-Wave Alternans Detection,” *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 10, pp. 2402–2412, 2010.
- [7] D. P. Tobón, T. H. Falk, and M. Maier, “MS-QI: A modulation spectrum-based ECG quality index for telehealth applications,” *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 8, pp. 1613–1622, 2014.
- [8] C. Plaza-Seco, M. Blanco-Velasco, J. Rojo-Álvarez, and S. Muñoz-Romero, “P-wave detection using latent spaces from autoencoders and domain description classifiers,” pp. 91–94, CASEIB 2021, Nov 2021.

# Señales Biomédicas III

---

Viernes, 25 de noviembre de 2022

# Simulación de señales ECG incluyendo dinámica del intervalo PQ con el ritmo cardiaco y ruido muscular variante en el tiempo

C. Pérez<sup>1</sup>, E. Pueyo<sup>1,2</sup>, JP. Martínez<sup>1,2</sup>, L. Sörnmo<sup>3</sup>, P. Laguna<sup>1,2</sup>

<sup>1</sup>Grupo BSICoS, I3A, IIS Aragón, Universidad de Zaragoza, Zaragoza, España {cperez, epueyo, jpmart, laguna}@unizar.es

<sup>2</sup>CIBER-Bioingeniería, Biomateriales y Nanomedicina (CIBER BBN), España

<sup>3</sup>Department of Biomedical Engineering, Lund University, Sweden, leif.sornmo@bme.lth.se

## Resumen

Contar con simuladores de electrocardiogramas (ECG) capaces de generales señales fieles a las dinámicas reales tiene un gran valor en la actualidad. Gracias a ellas se puede desde evaluar algoritmos de procesamiento de señal en un entorno controlado hasta crear señales que ayuden a entrenar modelos de Machine Learning. Una dinámica realista del intervalo PQ y el modelado de ruido muscular variante en el tiempo son las dos contribuciones novedosas recogidas en la versión de un simulador presentada en este trabajo. Se ha observado en datos reales que existe una relación lineal entre el intervalo PQ y el ritmo cardiaco cuando este supera cierto valor, mientras que por debajo de este valor se mantiene constante. Se incluye también un modelo paramétrico para simular ruido muscular que incluye dos partes: una es la caracterización frecuencial del ruido y otra la inclusión de variación de la potencia a lo largo del tiempo.

## 1. Introducción

El procesamiento del electrocardiograma (ECG) y la extracción de características del mismo son campos con gran número de investigaciones. A lo largo de estos años se han desarrollado algunos simuladores de ECG como los presentados en [1] y [2] para evaluar el rendimiento de algoritmos de procesamiento de ECG en un entorno controlado. Sin embargo, su campo de aplicación es muy específico: Behar et al. [1] se centraron en la simulación de ECGs maternos y fetales, mientras que Petreñas et al. [2] en el estudio de episodios de fibrilación auricular.

Actualmente tiene un gran auge el uso de herramientas de Machine Learning, para las que se entrena con un procedimiento de caja negra, con el objetivo de que ayude a clasificar, por ejemplo, en aplicaciones donde se desea agrupar sujetos en función de su señal ECG o realizar un diagnóstico usando esta señal. Al contrario de lo que sucede con herramientas de simulación de señales de ECG, este proceso no está basado en los fundamentos fisiológicos o el modelado matemático de las señales cardíacas observadas [3]. Por otro lado, las señales de ECG creadas por los simuladores pueden utilizarse como señales para entrenar redes neuronales de una forma controlada, aumentando así las señales disponibles y ayudando a vencer el fenómeno del overfitting. Pero esto será tanto más cierto en tanto las señales simuladas sean creadas lo más realísticamente posible.

Este trabajo presenta una nueva versión mejorada del simulador multiderivacional presentado por Petreñas et al. [2], que incluye las dos mejoras que se exponen en este

trabajo: una dinámica realista del intervalo PQ con el ritmo cardiaco y el modelado de ruido muscular de una forma variante en el tiempo.

Existe una laguna en la literatura en el estudio de la relación entre el intervalo PQ y el ritmo cardiaco. Una de las principales dificultades es que, conforme aumenta el ritmo cardiaco, es más difícil delinear la onda P debido al solapamiento que se produce entre esta onda y la onda T. Sin embargo, algunos estudios han demostrado que la longitud del intervalo PQ no es constante y que tiene cierta dependencia del ritmo cardiaco [4,5].

Por otro lado, el modelado de la variación en el tiempo del ruido muscular presente en las señales de ECG también es un campo poco investigado. Behar et al. [1] definieron un modelo que modificaba ligeramente el espectro frecuencial a lo largo del tiempo, pero no añadía una variación de la potencia como ocurre frecuentemente en registros reales.

Tanto la consideración de la dinámica del intervalo PQ, como la inclusión de ruido en señales simuladas son características esenciales para probar los algoritmos de forma controlada, para cualquier valor del ritmo cardiaco, a diferentes niveles de ruido y en situaciones altamente variantes en el tiempo, como pruebas de esfuerzo.

## 2. Dinámica del intervalo PQ

### 2.1. Hipótesis

El intervalo PQ y la variabilidad del ritmo cardiaco (representada en el intervalo RR) no están correlacionados para valores bajos del ritmo cardiaco. Sin embargo, se observa una relación lineal cuando el ritmo cardiaco supera un cierto valor. De modo que esta relación entre el intervalo PQ y el intervalo RR predecesor se puede modelar como:

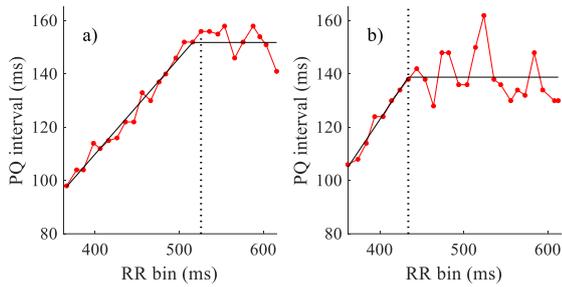
$$PQ(d_{RR}) = \begin{cases} A & \text{if } d_{RR} \geq \lambda \\ A + B(d_{RR} - \lambda) & \text{if } d_{RR} < \lambda \end{cases} \quad (1)$$

donde el intervalo PQ depende del intervalo RR cuando está por debajo de cierto valor  $\lambda$ .

### 2.2. Datos

Para ilustrar la variación temporal del intervalo PQ y determinar los parámetros de (1), la dinámica del intervalo PQ respecto a RR se estudió en un conjunto de 42 señales ECG registradas en el Hospital Universitario de Tampere (Finlandia) durante una prueba de esfuerzo y muestreadas a 500 Hz [6]. Los registros corresponden a pacientes con

bajo riesgo de sufrir enfermedad coronaria, según los síntomas desarrollados en la prueba de esfuerzo y su historial clínico.



**Figura 1.** Modelado del intervalo PQ en función del intervalo RR en a) la zona de aceleración y b) la zona de desaceleración.

Las líneas negras representan el mejor ajuste de regresión lineal, tras calcular el punto de cambio  $\lambda$  (línea discontinua negra), de los valores medianos (puntos rojos) de PQ de todos los sujetos para valores del intervalo RR tomados cada 10 ms.

### 2.3. Proceso

Para delinear el ECG se calculó el punto inicial del complejo QRS aplicando un detector de picos, inicios y finales de onda [7] a las 8 derivaciones estándar disponibles. El reto fue calcular el inicio de la onda P, ya que conforme aumenta el ritmo cardíaco, la distancia entre las ondas T y P se va reduciendo, llegando a solaparse. Por ello, haciendo uso de las 8 derivaciones disponibles, se obtuvo una derivación transformada por medio de la técnica de reducción espacial de análisis por componentes periódicas aplicada a la onda P, de manera análoga como se ha hecho en otros trabajos sobre las ondas T [8]. El inicio de la onda P se obtuvo al delinear solamente la derivación transformada (en este caso, entrenada para enfatizar la onda P) con el delineador mencionado anteriormente.

Tras haber calculado el intervalo PQ para todas las señales, se agruparon estas medidas por rangos de intervalos de RR de 10 ms y posteriormente se calculó el valor mediano para cada rango. De modo que se tiene un valor de PQ para cada rango de intervalos RR, estando equiespaciados los valores de RR cada 10 ms. El punto de cambio  $\lambda$ , donde PQ deja de tomar un valor constante y decrece linealmente con el intervalo RR, se obtuvo mediante la búsqueda del mínimo error cuadrático medio (MSE), calculado al ir ajustando los resultados de la dependencia de las series de PQ con RR y el modelo descrito en la ecuación (1).

El proceso descrito anteriormente se realiza de forma independiente para los datos de PQ y RR registrados durante la fase de ejercicio (desde un estado basal hasta el pico de ejercicio), y para aquellos registrados durante la fase de recuperación (desde el pico de ejercicio hasta un estado de reposo) de la prueba de esfuerzo, dado que el comportamiento del intervalo PQ con el RR puede presentar el fenómeno de histéresis que hace que la respuesta del intervalo PQ ante incrementos y decrementos del intervalo RR sea distinta. El resultado de este ajuste se puede ver en la Fig.1a para la zona de ejercicio (fase de aceleración) y Fig.1b para la zona de recuperación (fase de desaceleración) donde cada punto rojo es el valor mediano de PQ a lo largo del rango de RR, la línea vertical punteada

indica el punto de cambio  $\lambda$ , obtenido de resolver la minimización, y las líneas negras son el ajuste según la ecuación (1).

## 3. Modelado ruido muscular

### 3.1. Hipótesis

El modelado del ruido muscular (*MN por sus siglas en inglés*), se ha desglosado en dos bloques: un bloque que modela la parte cuasi-estacionaria, la cual refleja la distribución espectral del ruido, y otra que modela la varianza, centrándose este bloque en la potencia del ruido. El esquema de funcionamiento de ambos bloques y su interacción se recoge en el diagrama de la Fig. 2.

### 3.2. Datos

Se cuenta con 25 señales de diferentes sujetos que recogen el ruido registrado en un ECG grabado durante una prueba de esfuerzo [9]. Cada señal cuenta con 9 derivaciones: V1, RV4, V3-V6, I, II y III muestreadas a 1000 Hz.

### 3.3. Procesado

Como solo se desea modelar el ruido muscular, en primer lugar, se aplica un filtro Butterworth de tercer orden pasa-banda con fase cero, siendo las frecuencias de corte 10 Hz y 80 Hz respectivamente. Además, se elimina la interferencia de la línea eléctrica de 50 Hz por medio de un filtro no lineal [10].

Posteriormente, se normaliza la varianza de la señal de ruido muscular real resultante, es decir, se calcula una señal con potencia homogénea. Para ello, se divide la señal filtrada entre su envolvente suavizada. Para finalizar con el modelado de la parte cuasi-estacionaria del MN, la señal normalizada se remuestra a 200 Hz para calcular los coeficientes de un modelo AR de orden 4 mediante el método de Burg. De este modo, se cuenta con una librería de 25 modelos AR (uno por cada señal real disponible). En cada simulación, se seleccionan los coeficientes de un modelo y, para aumentar la variabilidad, se modifica ligeramente el valor los polos asociados cada 10 s empleando el algoritmo de camino aleatorio (o *random walk*), garantizando que permanezcan siempre dentro del círculo unidad.

En cuanto a la potencia del ruido muscular, se modela su desviación estándar,  $\sigma_w(n)$ , según el bloque inferior izquierdo de la Fig.2. Básicamente  $\sigma_w(n)$  está compuesta por dos componentes:

$$\sigma_w(n) = m_{\sigma_w}(n) + x_{\sigma_w}(n) \quad (2)$$

donde  $m_{\sigma_w}(n)$  es una función exógena que define el nivel de ruido y  $x_{\sigma_w}(n)$  es la salida del filtro lineal  $H(z)$ :

$$H(z) = \frac{1}{1 - \kappa z^{-1}} \quad (3)$$

siendo la entrada de este filtro un ruido blanco  $v_{\sigma_w}(n)$ , Gaussiano con varianza  $\sigma_v^2$ . Este filtro define la variación entre los valores de  $\sigma_w(n)$  en función de su coeficiente  $\kappa$ , el cual varía entre  $0 < \kappa < 1$ . La señal  $\sigma_w(n)$  está muestreada a 200 Hz.

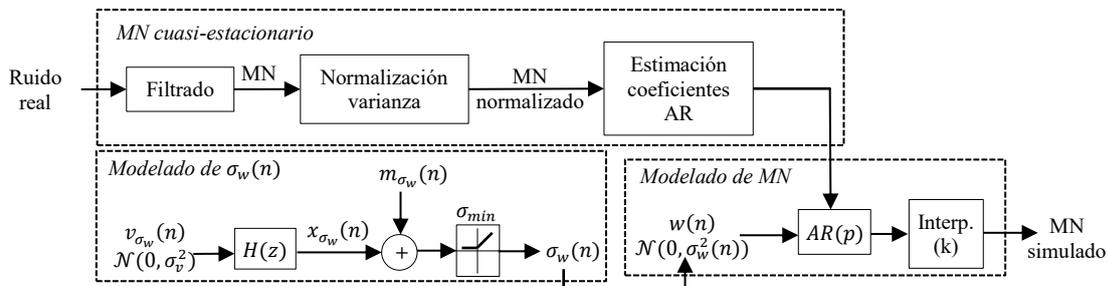


Figura 2. Diagrama del modelado del ruido muscular a partir de una señal de ruido real.

Para evitar valores negativos en  $\sigma_w(n)$  se ha añadido un rectificador de media onda imponiendo que el valor mínimo sea  $\sigma_{min}$ .

Finalmente, el ruido muscular simulado se obtiene al computar el proceso AR de orden 4 (cuyos coeficientes se describieron en el primer bloque) con una entrada definida por un ruido blanco  $w(n)$ , Gaussiano y con una varianza  $\sigma_w^2(n)$  que cambia a lo largo del tiempo. La señal de ruido muscular simulado resultante se remuestra a 1000 Hz, que es la frecuencia de trabajo del simulador al cual se incorpora este proceso.

#### 4. Resultados

En relación a la dinámica del intervalo PQ, se aplicó el proceso descrito de forma independiente tanto en la fase de ejercicio como en la de recuperación de la prueba de esfuerzo. De este modo, pudimos analizar los resultados tanto en valores de aceleración (ejercicio) como de deceleración (recuperación) del ritmo cardiaco. En la Tabla 1 se recogen los valores de los parámetros del modelado descrito en la ecuación (1) para cada caso, así como el coeficiente de determinación para el ajuste en la zona con pendiente ( $R_s^2$ ).

Cabe señalar que la duración de la onda P se estableció en 120 ms, ya que se puede considerar invariante frente a variaciones en el ritmo cardiaco.

	A (ms)	B	$\lambda$ (ms)	$R_s^2$
<b>Ejercicio</b>	152	0.362	520	0.98
<b>Recuperación</b>	139	0.470	430	0.97

Tabla 1. Valores de los parámetros de la ecuación (1) calculados en la zona de ejercicio y de recuperación.

En la Fig.3a se muestra una señal ECG simulada que incluye la dinámica del intervalo PQ propuesta. Para ello, los intervalos de los primeros latidos tienen una frecuencia cardiaca por debajo del punto de cambio  $\lambda$ , y el ritmo cardiaco a partir de los 500s supera el valor  $\lambda$ . En la Fig.3b se ha añadido otra señal simulada sin tener en cuenta la dinámica del intervalo PQ según el intervalo RR, apreciando que cuando el ritmo cardiaco aumenta la onda P queda solapada con la onda T. Además, se ha añadido un ruido muscular simulado, donde nos hemos centrados en el caso de modelar este tipo de ruido para ECGs simulados durante una prueba de esfuerzo. Por ello, se ha definido la función exógena  $m_{\sigma_w}(n)$  mediante dos funciones exponenciales de base 2: una de ellas creciente desde el

inicio de la prueba hasta el pico de esfuerzo; y la otra, decreciente desde el pico de esfuerzo hasta el final de la prueba. El valor mínimo de estas funciones corresponde con un tercio del nivel de ruido, cuyo valor es definido por el usuario, y el valor máximo es igual al nivel de ruido definido. En este ejemplo el nivel de ruido se ha establecido en 20  $\mu V$ .

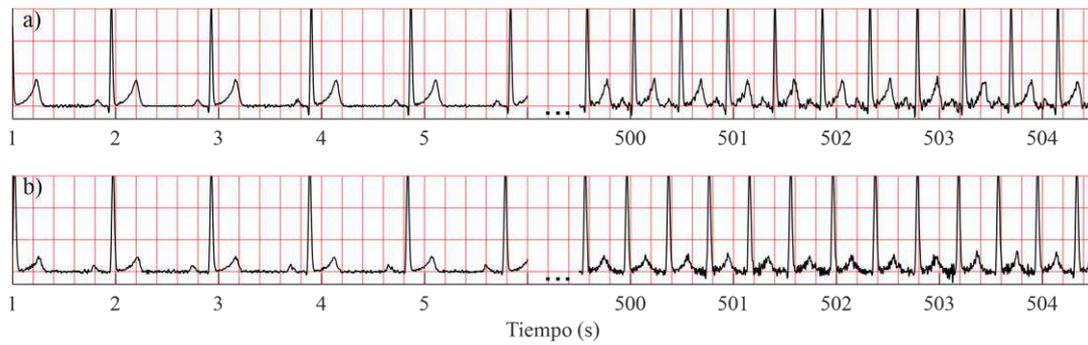
#### 5. Discusión

Los métodos propuestos en este trabajo surgen de la necesidad de analizar, en un entorno controlado, algoritmos para delinear el final de la onda T [8] en señales ECG grabadas durante prueba de esfuerzo. Las aportaciones descritas han sido añadidas al simulador de arritmias de Petrénas et al. [2].

La hipótesis sobre la relación de los intervalos PQ y RR, modelada de acuerdo con la ecuación (1), también fue observada por Atterhög et al. [4], pero utilizando el intervalo PR. Pilhall et al. [11] estudiaron los cambios del segmento QRS durante el ejercicio, concluyendo que la duración QR no varía significativamente para altos valores de ritmo cardiaco. Por tanto, modelar la dinámica del intervalo PQ en lugar del intervalo PR queda justificada.

Cabasson et al. [5] estudiaron la dinámica del intervalo PR en señales de pruebas de esfuerzo, observando que existe histéresis en la relación PR-RR estudiada en ejercicio y en recuperación de forma independiente. Sin embargo, la duración del intervalo PR de sus señales simuladas permaneció constante, mientras que la distancia entre las ondas T y P disminuía linealmente conforme aumentaba el ritmo cardiaco.

En cuanto al modelado del ruido muscular presente en una señal ECG, Behar et al. [1] propusieron un modelo basado en una función AR, cuyos coeficientes fueron calculados de una señal de ruido muscular de un solo sujeto y utilizaron el algoritmo de random walk para calcular nuevos coeficientes. En este trabajo, se ha presentado un modelado de ruido muscular más detallado, el cual se ha dividido en dos partes. La primera modela la distribución espectral de este ruido. En el espectrograma de las señales reales disponibles se observó un espectro bastante plano en el rango 10-80 Hz. Tampoco se observó ninguna correlación entre las distintas derivaciones. La elección de un modelo AR de bajo orden para modelar la parte cuasi-estacionaria del ruido muscular se debe a que se desea que exista cierta variabilidad en sus características entre distintas simulaciones. Además, se potencia la variabilidad del modelo entre diferentes señales simuladas al disponer



**Figura 3.** Señal cardíaca simulada con ruido muscular incorporado cuando a) se incluye la dinámica del intervalo PQ según el intervalo RR, b) la longitud del intervalo PQ es constante para cualquier valor de RR. Los primeros latidos corresponden a un valor del intervalo RR mayor que  $\lambda$ , mientras que a partir de los 500s el intervalo RR es menor que  $\lambda$ .

de 25 señales reales, de las cuales se calculan los coeficientes del modelo AR, incorporando la idea de Behar et al. [1] de alterar ligeramente el valor de estos coeficientes en diferentes ventanas de tiempo.

Se deja a voluntad del usuario definir tanto la media de la desviación estándar como la entrada exógena. De esta forma, se puede simular ruido con diferente nivel, así como definir un ruido que evolucione a lo largo del tiempo, como es el caso de la prueba de esfuerzo. También debe notarse el efecto del coeficiente  $\kappa$ , obteniéndose una señal más oscilatoria conforme más cerca está el coeficiente del círculo unidad. Centrándonos en nuestro trabajo de simular señales ECG durante una prueba de esfuerzo, en la Fig.3 se observa que el valor de ruido muscular ha aumentado al aumentar el ritmo cardíaco. Sin embargo, en la simulación mostrada se ha seleccionado un nivel de ruido bajo para apreciar el comportamiento de las ondas P y T.

Por último, analizando las dinámicas incluidas en el simulador, podemos observar en la Fig.3a cómo la distancia entre las ondas P y T se reduce al depender del intervalo RR cuando este es menor que el umbral  $\lambda$ .

## 6. Conclusiones

En este trabajo se han introducido dos aspectos relevantes para dar realismo a los simuladores de ECG, de particular utilidad cuando se simulan señales de prueba de esfuerzo. Se ha identificado y cuantificado una relación lineal entre el intervalo PQ y la variación del ritmo cardíaco cuando este supera cierto umbral. Se ha modelado el ruido variante en el tiempo para simular ruido muscular, el cual se compone de dos bloques: uno de ellos modela la distribución espectral y el otro la variación de la potencia a lo largo del tiempo. Ambas aportaciones se han incluido en un simulador de propósito general disponible para la comunidad científica.

## Agradecimientos

Este trabajo se pudo llevar a cabo gracias al Gobierno de Aragón por la concesión de un contrato Predoctoral (orden IJU/796/2019) y a la Universidad de Zaragoza, la Fundación Bancaria Ibercaja y Fundación CA de Estancias de Investigación (referencia IT 3/22). Está respaldado por los proyectos ERC-StG 638284 (Consejo Europeo de Investigación), PID2019-105674RB-I00 y PID2019-104881RB-I00 (Ministerio de Ciencia e Innovación) y el

proyecto LMP94 21 y grupo de referencia T39-20R (Gobierno de Aragón co-financiado por FEDER 2014-2020 “Construyendo Europa desde Aragón”). Los cálculos se realizaron mediante ICTS NANBIOSIS (unidad HPC de la University of Zaragoza).

## Referencias

- [1] Behar, J., et al. An ECG simulator for generating maternal-foetal activity mixtures on abdominal ECG recordings. *Physiological Measurement*, 2014, vol. 35, no 8, p. 1537.
- [2] Petrenas, A., et al. Electrocardiogram modeling during paroxysmal atrial fibrillation: application to the detection of brief episodes. *Physiological Measurement*, 2017, vol. 38, no 11, p. 2058.
- [3] Mincholé, A., et al. Machine learning in the electrocardiogram. *Journal of Electrocardiology*, 2019, vol. 57, p. S61-S64.
- [4] Atterhög, JH., et al. PR interval in relation to heart rate during exercise and the influence of posture and autonomic tone. *Journal of Electrocardiology*, 1977, vol. 10, no 4, p. 331-336.
- [5] Cabasson, A., et al. Quantifying the PR interval pattern during dynamic exercise and recovery. *IEEE Transactions on Biomedical Engineering*, 2009, vol. 56, no 11, p. 2675-2683.
- [6] Virtanen, M., et al. Heart rate variability derived from exercise ECG in the detection of coronary artery disease. *Physiological Measurement*, 2007, vol. 28, no 10, p. 1189.
- [7] Martínez, J.P., et al. A wavelet-based ECG delineator: evaluation on standard databases. *IEEE Transactions on Biomedical Engineering*, 2004, vol. 51, no 4, p. 570-581.
- [8] Pérez, C., et al. Eigenvector-based spatial ECG filtering improves QT delineation in stress test recordings. *55th Asilomar Conference on Signals, Systems, and Computers. IEEE*, 2021. p. 261-264.
- [9] Bailón, R., et al. Coronary artery disease diagnosis based on exercise electrocardiogram indexes from repolarisation, depolarisation and heart rate variability. *Medical and Biological Engineering and Computing*, 2003, vol. 41, no 5, p. 561-571.
- [10] Mortara, R.W. Digital filters for ECG signals. *Computers in Cardiology*, 1977, p. 511-514.
- [11] Pilhall, M., et al. Changes in the QRS segment during exercise: effects of acute  $\beta$ -blockade with propranolol. *Clinical Physiology*, 1993, vol. 13, no 2, p. 113-131.

# Simulación de los efectos de fármacos antiarrítmicos en la electrofisiología auricular

V. Puche-García<sup>1</sup>, L. Romero<sup>1</sup>, J. Saiz<sup>1</sup>

<sup>1</sup> Centro de Investigación e Innovación en Bioingeniería (Ci2B), Universitat Politècnica de València, España

[jsaiz@ci2b.upv.es](mailto:jsaiz@ci2b.upv.es)

## Resumen

La terapia farmacológica es la primera línea contra la fibrilación auricular; aunque los mecanismos no se comprenden por completo. Nuestro objetivo es comparar datos experimentales a nivel unicelular de varios fármacos con la respuesta de tres modelos computacionales diferentes: Courtemanche (CRN), Grandi y Skibbye (SKI) empleando un modelo de bloqueo de poro simple, así como una variante de modelo dinámico para los fármacos flecainida y vernakalant en CRN (+I<sub>Na,Marvok</sub>). La mayoría de las simulaciones consiguen reproducir la dependencia del aumento del APD<sub>90</sub> y la disminución de la  $dV/dt_{max}$  con la frecuencia y la concentración. Además, las respuestas de CRN, CRN+I<sub>Na,Marvok</sub> y SKI reprodujeron de manera más adecuada las tendencias encontradas en los resultados experimentales. El análisis realizado contribuye a la mejora de la simulación de fármacos antiarrítmicos.

## 1. Introducción

La fibrilación auricular (FA) es la arritmia sostenida más común en humanos, afectando a casi el 2% de la población. Aunque no se considera una arritmia mortal, conlleva riesgos potencialmente mortales para el paciente [1].

A pesar de las implicaciones relacionadas, los mecanismos subyacentes a esta patología no se comprenden completamente y los tratamientos existentes no son completamente efectivos. La terapia de ablación y la farmacoterapia son las dos estrategias principales contra la FA; sin embargo, aproximadamente en un tercio de todos los pacientes los resultados no son óptimos.

La administración de agentes antiarrítmicos es la primera línea de tratamiento contra la FA. Los fármacos actúan sobre canales iónicos específicos de manera que se modifica la electrofisiología y la propagación del impulso en el tejido cardíaco, provocando la terminación de las múltiples ondas de actividad eléctrica que mantienen la FA y eventualmente restaurando el ritmo sinusal. Sin embargo, los fármacos actuales tienen una eficacia limitada y en algunos casos se han identificado efectos secundarios graves como favorecer arritmias potencialmente mortales en los ventrículos tipo *torsades de pointes*.

Debido a la alta incidencia y a la baja efectividad de las estrategias actuales, existe un gran interés en estudiar en profundidad los mecanismos de acción de los agentes antiarrítmicos para poder entender mejor su efecto y conseguir desarrollar biomarcadores que ayuden al clínico a elegir la terapia óptima.

En este trabajo se lleva a cabo un estudio computacional para evaluar y entender mejor el efecto de cuatro fármacos

comúnmente utilizados en el tratamiento de la FA (amiodarona, dofetilide, flecainida, vernakalant) sobre el potencial de acción auricular (PA).

## 2. Métodos

### 2.1. Modelo de miocito humano auricular

Para reproducir la electrofisiología presente en las aurículas se implementaron tres modelos computacionales de miocito humano auricular: Courtemanche-Ramirez-Nattel [2], Grandi [3] y Skibbye [4].

### 2.2. Modelado del bloqueo por fármaco

Los fármacos actúan interaccionando con ciertas familias de canales bloqueando parte de su corriente. Este efecto se puede modelar a través de un modelo de poro simple o mediante un modelo dinámico mediante una formulación de Markov. El modelo de poro simple se implementó para todos los fármacos en los modelos CRN, GRND y SKI. Adicionalmente, también se implementó el modelo dinámico para los fármacos vernakalant y flecainida en el modelo CRN (CRN+I<sub>NaMarvok</sub>).

Para simular el efecto del fármaco mediante el modelo de poro simple, la conductancia máxima de las corrientes iónicas se redujo de acuerdo con la ecuación de Hill:

$$g_{i,drug} = g_i \cdot \left[ 1 + \left( \frac{D}{IC_{50,i}} \right)^{nH} \right]^{-1}$$

donde  $g_{i,drug}$  es la conductancia máxima de la corriente  $i$  en presencia del fármaco,  $IC_{50,i}$  es la concentración inhibitoria media máxima del fármaco,  $D$  la concentración de fármaco libre y  $nH$  el coeficiente de Hill indicando el número de moléculas del fármaco que se supone que son suficientes para bloquear el canal. Para cada fármaco, los valores de  $IC_{50}$  y  $nH$  para las diferentes corrientes se muestran en la Tabla 1, publicados por Llopis-Lorente et al. [5].

**Tabla 1.** Valores experimentales de  $IC_{50}$  ( $\mu M$ ) y coeficiente de Hill para los diferentes fármacos antiarrítmicos.

Fármaco	IC <sub>50</sub> de la corriente iónica ( $\mu M$ ) (Coeficiente de Hill si es distinto de 1)					
	I <sub>CaL</sub>	I <sub>Kr</sub>	I <sub>Ks</sub>	I <sub>Kur</sub>	I <sub>to</sub>	I <sub>Na</sub>
<i>Amiodarona</i>	1.59 (0.65)	0.56 (1.32)	1.74 (0.72)	-	3.76 (0.4)	0.48 (0.7)
<i>Dofetilide</i>	2.3 (5.4)	0.01	100	-	-	1460 (5.1)
<i>Flecainida</i>	26.35 (1.19)	1.5 (0.88)	20	-	9.26 (0.7)	5.8
<i>Vernakalant</i>	84	20	-	15	15	90

El modelo dinámico permite simular la afinidad y dependencia del fármaco al estado del canal. Para simular la inhibición de la corriente rápida de sodio ( $I_{Na}$ ) producida por el vernakalant y la flecainida en el modelo CRN, se incluyó el modelo de Markov publicado por Moreno et al. [6] y modificado por Sutanto et al. [7].

Las respuestas de los modelos en presencia de los diferentes fármacos se compararon y validaron frente a datos experimentales publicados en la literatura [8, 9, 10, 11, 12, 13, 14].

### 2.3. Protocolo de simulación

Las ecuaciones diferenciales ordinarias de los modelos unicelulares CRN, GRND y SKI se resolvieron utilizando el *solver* de paso de tiempo variable *ode15s* en MATLAB.

Los modelos de PA auricular fueron estimulados aplicando una corriente 1.5 veces la umbral y una duración de 2 ms. Los modelos se estabilizaron estimulándolos a 1 y 2.5 Hz, durante 20 s en el caso de CRN y 500 s para GRND y SKI. A continuación, se aplicaron 150 pulsos más añadiendo el efecto del fármaco, manteniendo la misma frecuencia.

En las simulaciones se reprodujeron los efectos de diferentes concentraciones de fármacos y diferentes periodos de estimulación (BCL, *Basic Cycle Length*) sobre los PAs, la duración del PA al 90% de repolarización ( $APD_{90}$ ) y sobre la velocidad máxima de despolarización ( $dV/dt_{max}$ ). Estos biomarcadores se calcularon en el pulso número 150 tras la aplicación del fármaco.

## 3. Resultados y discusión

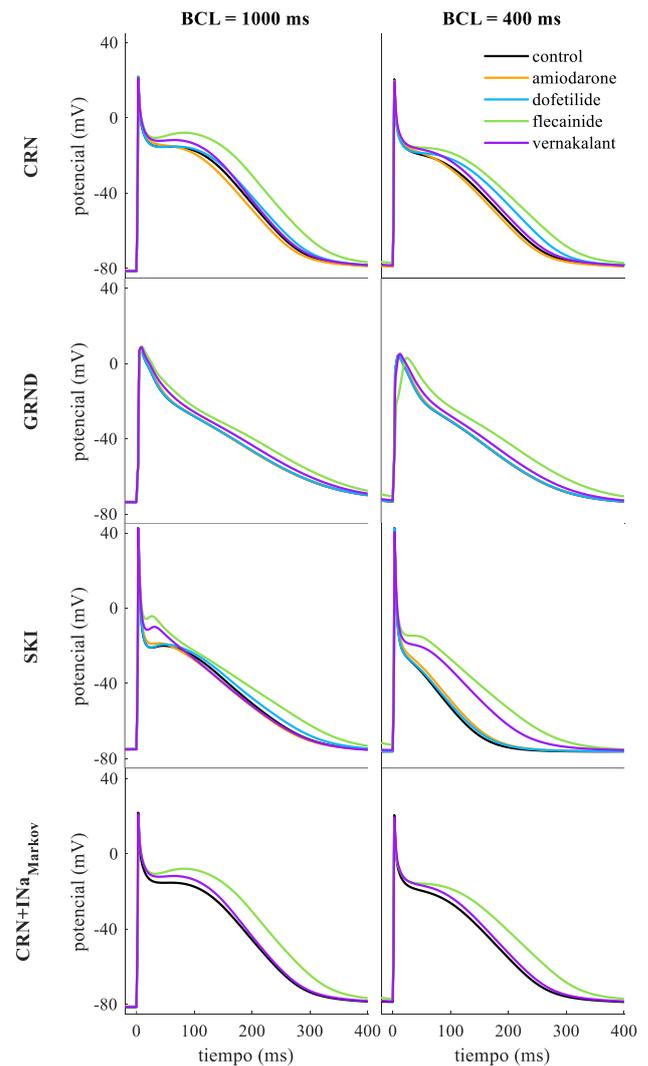
### 3.1. Simulaciones unicelulares

La Figura 1 muestra el PA para los 4 modelos electrofisiológicos (CRN, GRND, SKI,  $CRN+I_{Na,Markov}$ ) tras la aplicación de los 4 fármacos (amiodarona, dofetilide, flecainida, vernakalant) para 1 y 2.5 Hz. Se han utilizado sus concentraciones terapéuticas (EFTPC, *Effective Therapeutic Concentration*) [5, 7]: amiodarona = 0.063  $\mu$ M, dofetilide = 0.002  $\mu$ M, flecainida = 1.448  $\mu$ M, vernakalant = 3  $\mu$ M. De manera general, los fármacos consiguen aumentar el  $APD_{90}$  y reducir la  $dV/dt_{max}$ .

La flecainida consigue el mayor aumento en la prolongación del  $APD_{90}$ ; por ejemplo, en el caso de CRN (BCL = 1000 ms) consigue incrementar el  $APD_{90}$  en un 17.03%, mientras que dofetilide lo aumenta un 4.40% y vernakalant, un 1.61%. Esto se debe a que el bloqueo de la corriente de potasio rectificadora de activación rápida ( $I_{Kr}$ ), la cual contribuye principalmente en la repolarización del PA, es mayor en el caso de la flecainida (bloqueo de un 49%) mientras que para el resto de los fármacos el bloqueo no supera el 20%, teniendo en cuenta para cada fármaco el  $IC_{50}$  y la EFTPC. El bloqueo de corrientes de potasio salientes ( $I_K$ ) actúa como principio antiarrítmico ayudando a reducir la vulnerabilidad a la FA ya que provoca el aumento en la duración del potencial y, por lo tanto, favorece la refractariedad, generando un aumento en el período refractario efectivo.

El modelo GRND no consigue reproducir adecuadamente el alargamiento en el  $APD_{90}$  debido a que la sensibilidad

de este modelo a la corriente  $I_{Kr}$  es muy baja (datos no mostrados), por lo que el bloqueo de esta corriente no produce un cambio significativo en la prolongación del PA.



**Figura 1.** PAs para los 4 modelos y los dos periodos de estimulación bajo condiciones control y bajo el efecto de los fármacos a su concentración terapéutica.

Para el modelo CRN, el aumento en el  $APD_{90}$  es mayor para la frecuencia rápida obteniendo un aumento de 4.40% en frecuencia normal frente un 4.51% en frecuencia rápida en el caso de dofetilide, un 17.03% frente un 21.06% en el caso de flecainida y un 1.61% frente a un 3.27% en el caso de vernakalant por lo que la prolongación del  $APD_{90}$  es dependiente de la frecuencia.

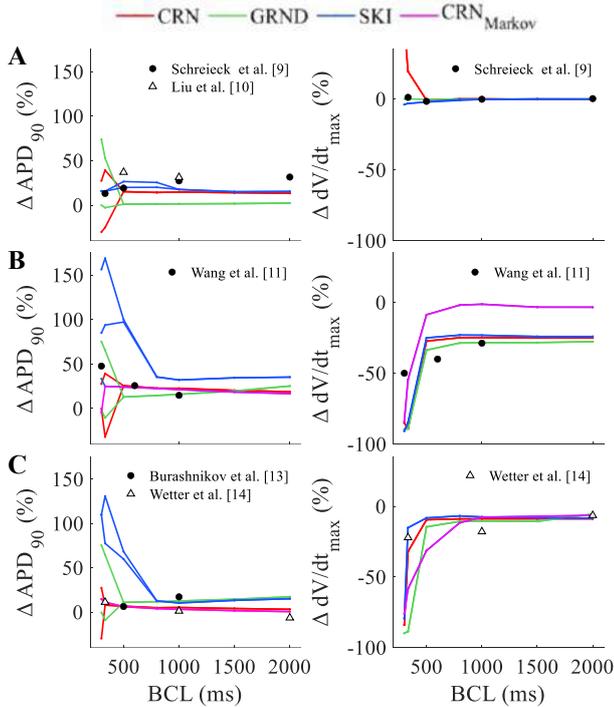
La amiodarona casi no varía el  $APD_{90}$ , pero sí logra reducir la  $dV/dt_{max}$  produciendo una disminución del 3.97% para frecuencia normal y 4.46% para la frecuencia rápida para el modelo SKI. De la misma forma, el resto de fármacos reducen la  $dV/dt_{max}$  de manera dependiente a la frecuencia consiguiendo un decremento en frecuencia normal de 0.08% frente a un 0.66% en frecuencia rápida en el caso de dofetilide, un 16.53% frente un 36.28% en el caso de flecainida y un 1.75% frente a un 6.91% en el caso de vernakalant para el modelo SKI.

En líneas generales, la dependencia del bloqueo y sus efectos sobre el  $APD_{90}$  y la  $dV/dt_{max}$  con la frecuencia es

coherente con estudios experimentales y se observa que los fármacos producen una acción selectiva en ritmos fibrilatorios.

### 3.2. Modulación del efecto de fármacos en función del periodo de estimulación

La Figura 2 muestra las variaciones de  $APD_{90}$  y  $dV/dt_{max}$  en función del BCL comparado con resultados experimentales. La concentración utilizada en la simulación para cada fármaco coincide con aquella usada en los experimentos referidos.



**Figura 2.** Simulación de la variación porcentual del  $APD_{90}$  y  $dV/dt_{max}$  en función del BCL para (A) dofetilide ( $0.01 \mu M$ ), (B) flecainida ( $2.25 \mu M$ ) y (C) vernakalant ( $10 \mu M$ ) para los diferentes modelos comparados con datos experimentales (●, Δ) publicados en la literatura.

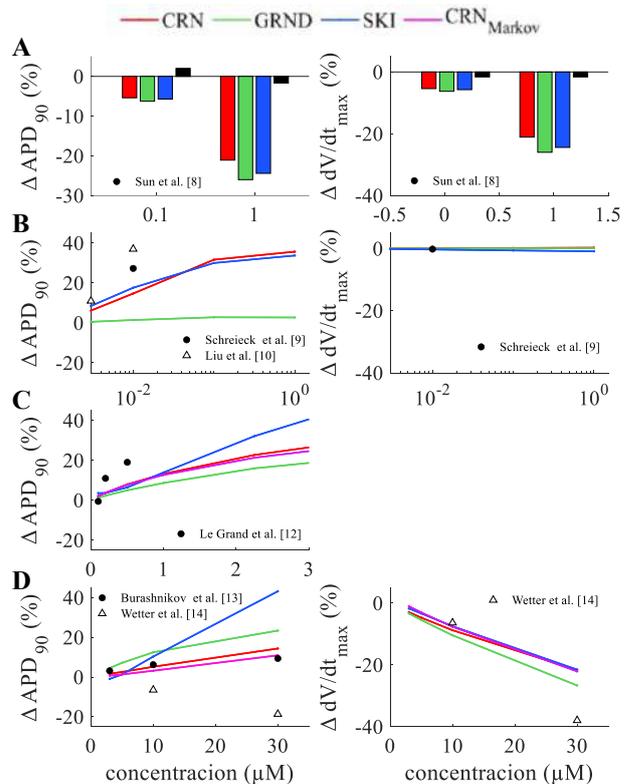
El dofetilide (Figura 2-A) produce un incremento en el  $APD_{90}$  que se mantiene estable para diferentes BCLs de forma coherente con los registros experimentales, excepto en el caso de GRND donde no hay apenas variación debido a la baja sensibilidad del modelo a la  $I_{Kr}$ . De la misma manera, este fármaco no genera una variación significativa de la  $dV/dt_{max}$  para diferentes frecuencias debido a que el fármaco actúa principalmente sobre la corriente  $I_{Kr}$  la cual actúa fuera de la despolarización. Este efecto se consigue reproducir con GRND y SKI, mientras que para CRN aumenta bastante a ritmos rápidos. Aun así, en los registros experimentales publicados por Schreieck et al. [9] se muestra un ligero incremento de la  $dV/dt_{max}$  de  $0.28\%$  a una frecuencia de  $3\text{Hz}$  comparado con el incremento de  $8.24\%$  para CRN.

Tanto flecainida (Figura 2-B) como vernakalant (Figura 2-C) aumentan su efecto con la frecuencia: mayor incremento del  $APD_{90}$  y mayor reducción de la  $dV/dt_{max}$ . Ambos fármacos bloquean la  $I_{Na}$  reduciendo con ello la  $dV/dt_{max}$  y consecuentemente aumentando el umbral de excitación de manera que se consigue reducir la excitabilidad y

automaticidad ectópica auricular. La mayor potencia de estos fármacos a ritmos rápidos conduce a una mayor inhibición de la  $I_{Na}$  en estas condiciones, como es el caso de la FA, lo que ayuda a limitar la actividad de reentrada y restaurar el ritmo sinusal en esta patología. Para estos fármacos, los modelos CRN y  $CRN+I_{Na,Markov}$  reproducen las tendencias de los fármacos de manera óptima tanto para la variación en la  $dV/dt_{max}$  como en el  $APD_{90}$ . Para el modelo SKI, el  $APD_{90}$  se acerca bastante a la respuesta experimental, aunque muestra incrementos demasiado elevados a la hora de reproducir el efecto de flecainida y vernakalant en ritmos rápidos.

### 3.3. Modulación del efecto de fármacos en función de su concentración

Estudios experimentales muestran un efecto de los fármacos dependiente de la concentración aplicada. Los resultados simulados se representan en la Figura 3. El BCL implementado en la simulación para cada fármaco coincide con aquel empleado en el protocolo experimental. En este caso, todas las simulaciones se realizaron a  $1\text{Hz}$ .



**Figura 3.** Simulación de la variación porcentual del  $APD_{90}$  y  $dV/dt_{max}$  en función de la concentración ( $\mu M$ ) para (A) amiodarona (B) dofetilide, (C) flecainida y (D) vernakalant para los diferentes modelos a  $1\text{Hz}$  comparados con datos experimentales (●, Δ) publicados en la literatura.

Excepto por el efecto de dofetilide en la  $dV/dt_{max}$ , los fármacos muestran un efecto dependiente a la concentración, de manera que, a mayor concentración, mayor es el grado de bloqueo sobre los canales y por lo tanto mayores son la prolongación del  $APD_{90}$  y la reducción de la  $dV/dt_{max}$ .

Como se adelantaba en las simulaciones bajo EFTPC, la amiodarona (Figura 3-A) produce poca variación sobre el

APD<sub>90</sub>, teniendo un mayor efecto sobre la  $dV/dt_{\max}$ . En el caso del dofetilide (Figura 3-B), el modelo GRND vuelve a fallar en representar el aumento del APD<sub>90</sub>. Además, la  $dV/dt_{\max}$  muestra apenas variación debido a que este fármaco no actúa sobre  $I_{Na}$  a diferencia del resto de fármacos simulados.

Nuevamente, el modelo CRN, CRN+ $I_{Na,Markov}$  y SKI reproducen las tendencias de forma similar a los resultados experimentales, aunque el modelo SKI genera un incremento elevado del APD<sub>90</sub> en el caso de vernakalant a concentraciones elevadas (Figura 3-D).

### 3.4. Limitaciones y líneas futuras

Para comprender mejor el efecto de los fármacos sobre la FA, las simulaciones unicelulares se pueden ampliar a modelos multidimensionales que permitan incluir las características anatómicas y electrofisiológicas. Debido a la disponibilidad de datos en la literatura, la acción de los fármacos se simuló mediante modelos de poro simple en algunos casos como en otros estudios similares [15]. Extender el modelo de Markov al resto de fármacos y modelos computacionales permitiría reproducir la dependencia en frecuencia y la inhibición de la  $I_{Na}$  de manera más realista.

## 4. Conclusiones

En este estudio *in-silico* se ha demostrado que los modelos computacionales resultan útiles a la hora de predecir el efecto de fármacos sobre la electrofisiología auricular. Se puede observar que los modelos que reproducen mejor los resultados experimentales son CRN y CRN+ $I_{Na,Markov}$  seguidos del modelo SKI, ya que este se aleja de los mismos para flecainida y vernakalant, especialmente para altas frecuencias. Respecto a los modelos dinámicos basados en CRN+ $I_{Na,Markov}$ , es necesario tener en cuenta el coste computacional que conlleva simular con un esquema de Markov en el modelo para decidir cuál sería más indicado para simulaciones multicelulares, como aquellas que utilizarían modelos 3D realistas de la aurícula. Este trabajo constituye un paso previo al uso de la simulación computacional como herramienta para diseñar nuevos fármacos en el tratamiento de la FA.

## Agradecimientos

Este trabajo ha sido financiado por el Programa Prometeo de la Generalitat Valenciana (2020/043) y por el programa de investigación e innovación Horizon 2020 de la Unión Europea bajo subvención No 101016496 (SimCardioTest).

## Referencias

- [1] Camm AJ, Kirchhof P, Lip GY, Schotten U, Savelieva I, Ernst S, Zupan I. Guidelines for the management of atrial fibrillation: the Task Force for the Management of Atrial Fibrillation of the European Society of Cardiology (ESC). *European heart journal*, vol 31, sup 19, 2010, pp 2369-2429
- [2] Courtemanche M, Ramirez RJ, Nattel S. Ionic mechanisms underlying human atrial action potential properties: insights from a mathematical model. *American Journal of Physiology-Heart and Circulatory Physiology*, vol 275, sup 1, 1998, pp H301-H321
- [3] Grandi E, Pandit SV, Voigt N, Workman AJ, Dobrev D, Jalife J, Bers DM. Human atrial action potential and Ca<sup>2+</sup> model: sinus rhythm and chronic atrial fibrillation. *Circulation research*, vol 109, sup 9, 2011, pp 1055-1066
- [4] Skibsbye L, Jespersen T, Christ T, Maleckar MM, van den Brink J, Tavi P, Koivumäki JT. Refractoriness in human atria: Time and voltage dependence of sodium channel availability. *Journal of molecular and cellular cardiology*, vol 101, 2016, pp 26-34
- [5] Llopis-Lorente J, Gomis-Tena J, Cano J, Romero L, Saiz J, Trenor B. In silico classifiers for the assessment of drug proarrhythmicity. *Journal of Chemical Information and Modeling*, vol 60, sup 10, 2020, pp 5172-5187
- [6] Moreno J, Zhu Z, Yang P, Bankston J, Jeng M, Kang C, Wang L, Bayer J, Christini D, Trayanova N, Ripplinger C, Kass R, Clancy C. A computational model to predict the effects of class I anti-arrhythmic drugs on ventricular rhythms. *Science translational medicine*, vol 3, sup 98, 2011, pp 98ra83
- [7] Sutanto H, Laudy L, Clerx M, Dobrev D, Crijns HJ, Heijman J. Maastricht antiarrhythmic drug evaluator (MANTA): A computational tool for better understanding of antiarrhythmic drugs. *Pharmacological research*, vol 148, 2019, pp 104444
- [8] Sun W, Sarma JS, Singh BN. Chronic and acute effects of dronedarone on the action potential of rabbit atrial muscle preparations: comparison with amiodarone. *Journal of cardiovascular pharmacology*, vol 39, sup 5, 2002, pp 677-684
- [9] Schreieck J, Wang Y, Gjini V, Korth M, Zrenner B, Schömig A, Schmitt C. Differential Effect of  $\beta$ -Adrenergic Stimulation on the Frequency-Dependent Electrophysiologic Actions of the New Class III Antiarrhythmics Dofetilide, Ambasilide, and Chromanol 293. *Journal of cardiovascular electrophysiology*, vol 8, sup 12, 1997, pp 1420-1430
- [10] Liu G, Xue X, Gao C, Huang J, Qi D, Zhang Y, Dong JZ, Ma C, Yan GX. Synergistic effect of dofetilide and mexiletine on prevention of atrial fibrillation. *Journal of the American Heart Association*, vol 6, sup 5, 2017, pp e005482
- [11] Wang ZG, Pelletier LC, Talajic M, Nattel S. Effects of flecainide and quinidine on human atrial action potentials. Role of rate-dependence and comparison with guinea pig, rabbit, and dog tissues. *Circulation*, vol 82, sup 1, 1990, pp 274-283
- [12] Le Grand B, Le Heuzey JY, Perier P, Peronneau P, Lavergne T, Hatem S, Guize L. Cellular electrophysiological effects of flecainide on human atrial fibres. *Cardiovascular research*, vol 24, sup 3, 1990, pp 232-238
- [13] Burashnikov A, Pourrier M, Gibson JK, Lynch JJ, Antzelevitch C. Rate-dependent effects of vernakalant in the isolated non-remodeled canine left atria are primarily due to block of the sodium channel: comparison with ranolazine and dl-sotalol. *Circulation: Arrhythmia and Electrophysiology*, vol 2, sup 400-408, 2012, pp 5
- [14] Wettwer E, Christ T, Endig S, Rozmaritsa N, Matschke K, Lynch JJ, Pourrier M, Gibson JK, Fedida D, Knaut M, Ravens U. The new antiarrhythmic drug vernakalant: ex vivo study of human atrial tissue from sinus rhythm and chronic atrial fibrillation. *Cardiovascular research*, vol 98, sup 1, 2013, pp 145-154
- [15] Dasí A, Roy A, Sachetto R, Camps J, Bueno-Orovio A, Rodriguez B. In-silico drug trials for precision medicine in atrial fibrillation: From ionic mechanisms to electrocardiogram-based predictions in structurally-healthy human atria. *Frontiers in physiology*, 2022, pp 1927

# DetECCIÓN automática de complejos QRS en pacientes con actividad eléctrica sin pulso durante la parada cardiorrespiratoria

J. Urteaga<sup>1</sup>, A. Elola<sup>2</sup>, E. Aramendi<sup>1,3</sup>, D. Herráez<sup>4</sup>, A. Norvik<sup>5</sup>, E. Unneland<sup>5</sup>, E. Skogvoll<sup>5</sup>

<sup>1</sup> Departamento de Ingeniería de Comunicaciones, Universidad del País Vasco, 48013 Bilbao, España

<sup>2</sup> Departamento de Tecnología Electrónica, Universidad del País Vasco, 48013 Bilbao, España

<sup>3</sup> Instituto de Investigación Sanitaria Biocruces Bizkaia, Hospital Universitario Cruces, 48903 Barakaldo, España

<sup>4</sup> Hospital Universitario Cruces, 48903 Barakaldo, España

<sup>5</sup> Norwegian University of Science and Technology (NTNU), 7491 Trondheim, Noruega

## Resumen

*La actividad eléctrica sin pulso (AESP) es uno de los principales ritmos cardíacos presentes durante la parada cardiorrespiratoria (PCR), y consiste en un ritmo cardíaco regular, pero sin contracciones efectivas. El principal tratamiento de la PCR es la reanimación cardiopulmonar (RCP) mediante compresiones torácicas y ventilaciones, acompañada de desfibrilaciones o fármacos cuando son necesarios. La terapia ante una AESP está condicionada por las características del ECG, principalmente de la morfología del complejo QRS, cuya detección es crucial. En la actualidad la mayoría de los métodos de detección de complejos QRS están diseñados para pacientes hemodinámicamente estables y con ritmo con pulso (RP). El objetivo de este trabajo es desarrollar un método automático de detección de complejos QRS en pacientes con AESP durante la PCR. La base de datos de este estudio cuenta con 5128 segmentos de AESP y 34704 segmentos de RP de 336 pacientes con una duración total de 2097 minutos. La señal ECG se descompone utilizando la transformada wavelet estacionaria, se aplica un detector de picos en una componente de detalle y se realiza una valoración multicomponente para la identificación de los complejos QRS. Finalmente se realiza un ajuste en tiempo de los complejos detectados usando la señal ECG original. El método propuesto presenta unos valores media (desviación estándar) de Se/PPV/F1 de 94.4(13.6)/85.0(17.6)/88.0(16.0) para segmentos de AESP, y de 95.8 (9.8)/89.3(12.5)/91.4(11.4) para todos los ritmos organizados en general.*

## 1. Introducción

La parada cardiorrespiratoria (PCR) es una de las principales causas de muerte en los países desarrollados, con una incidencia de 55 personas por 100.000 personas cada año y una tasa de supervivencia por debajo del 8.4% [1,2]. Una rápida detección de la parada y un tratamiento adecuado son esenciales para aumentar las probabilidades de supervivencia del paciente. El principal tratamiento ante una PCR es la reanimación cardiopulmonar (RCP), que consiste en compresiones torácicas y ventilaciones, junto a desfibrilaciones y tratamiento farmacológico cuando sea necesario [3-5].

Uno de los ritmos cardíacos más frecuentes durante la PCR es la actividad eléctrica sin pulso (AESP), con una incidencia de entre 40 y 60% en las PCR en hospitales [6]. La AESP se caracteriza por la presencia de una actividad eléctrica organizada en el corazón, pero sin contracciones efectivas, es decir incapaz de generar pulso [7]. Su

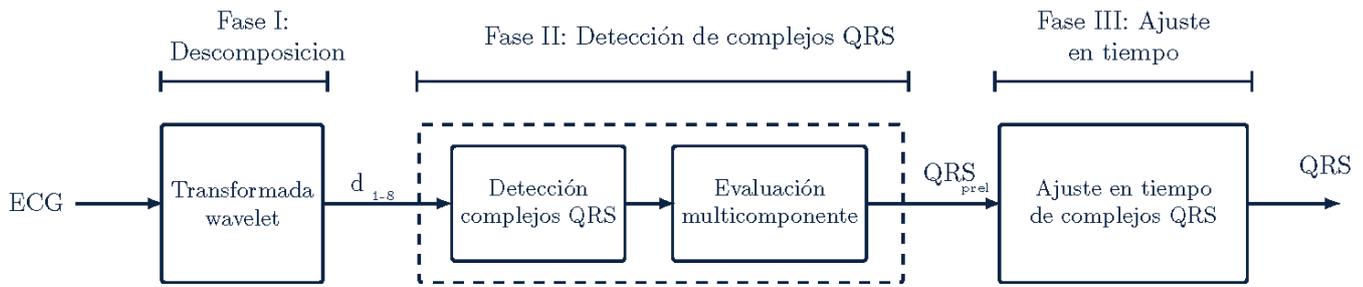
tratamiento durante la PCR está ligado al tipo concreto de AESP, que se refleja en la forma de onda del ECG, y principalmente en las características del QRS. Así, hay evidencias de que los denominados pseudo-AESP con QRS de estrecha duración y elevada pendiente tienen una evolución más positiva que las AESP estrictas, en las que se aconseja un tratamiento farmacológico inmediato. [8-10].

En este contexto, es esencial la correcta detección y delineación de los complejos QRS como primer paso de la caracterización automática del AESP. Una delineación correcta permitiría calcular características como la frecuencia cardíaca, duración de los complejos QRS, pendiente del complejo QRS etc. que ayudarían a conformar métodos de ayuda y guiado del rescatador.

En la literatura pueden encontrarse diferentes métodos automáticos desarrollados para la delineación de los complejos QRS en el contexto de pacientes hemodinámicamente estables, es decir, con ritmo organizado y con pulso (RP). Todos parten de una primera detección del complejo QRS, identificando el punto R de mayor amplitud [11-12]. Estos métodos no han sido evaluados para pacientes en PCR con AESP. El objetivo de este estudio es desarrollar un método automático para la detección de complejos QRS de pacientes con AESP durante la PCR.

## 2. Materiales

La base de datos utilizada en este estudio ha sido recolectada en los hospitales Sant Olav University Hospital (Noruega), Hospital of the University of Pennsylvania (EEUU) y Penn Peresbyterian Mecial Center (EEUU). Todos los episodios han sido grabados durante PCR utilizando los desfibriladores LifePak 20 y HeartSart MRx, con una frecuencia de muestreo de 250 Hz, e incluyen las señales de ECG e impedancia torácica. El suministro de compresiones torácicas durante la PCR, introduce artefactos en la señal ECG, lo que inhabilita su uso para el análisis automático. Es por ello que en este estudio se han incluido intervalos entre compresiones con ritmo cardíaco organizado y una duración mínima de 3 s. Los complejos QRS han sido revisados y anotados por los clínicos expertos en PCR.



**Ilustración 1.** Esquema general del método propuesto para la detección de los complejos QRS

Los intervalos se han dividido en segmentos de señal de 3s: 5128 segmentos de AESP con una duración total de 335 minutos y 19085 latidos; y 34704 segmentos de RP con una duración total de 1762 minutos y 154807 latidos.

Adicionalmente, se han utilizado tres bases de datos públicas de pacientes hemodinámicamente estables y con RP para evaluar el método propuesto. Se trata de las bases de datos son MIT-BIH Arrhythmia Database [13], European ST-T database [14] y QT database [15]. Se han identificado un total de 220 episodios de un total de 13474 minutos y 1003474 latidos.

### 3. Métodos

En la Ilustración 1 se observan las fases del método propuesto. En la primera, la señal se descompone en subbandas. En la segunda fase se detectan los complejos QRS usando las diferentes subbandas, y en la última se ajustan en tiempo.

#### 3.1. Preprocesado

La señal ECG se descompone en ocho subbandas utilizando la transformada wavelet estacionaria (SWT) con wavelet madre Daubechies-3 y umbrales suaves. Descomponer en ocho subbandas permite el análisis por separado de las diferentes bandas y se ha demostrado que esto es útil en la detección de complejos QRS [2,16].

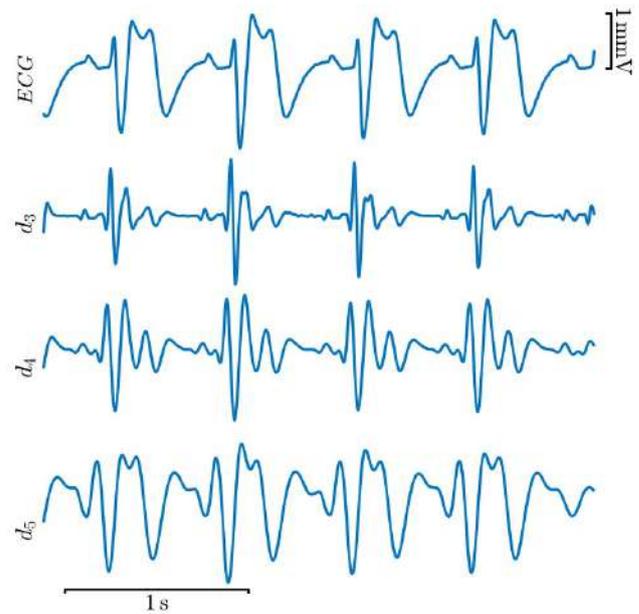
En la Ilustración 2 se pueden observar la señal de ECG y tres de los detalles resultado de la descomposición SWT.

#### 3.2. Detección de picos

Los potenciales complejos QRS son detectados primeramente en el tercer nivel de detalle ( $d_3$ ), correspondiente a la banda de frecuencias 15-30 Hz. Se aplica un detector de picos de alta sensibilidad con un umbral estático  $Th_3$  calculado sobre todo el segmento de 3 s, calculado a partir del percentil 99 según:

$$Th_3 = \text{percentil}_{99}(-d_3)$$

Y estableciendo una distancia mínima de 250ms entre dos picos consecutivos. Ese periodo entre latidos correspondería a una frecuencia cardiaca de 240 latidos por minuto, un valor suficientemente alto para que no se presenten ritmos de RP y AESP con frecuencias más altas.



**Ilustración 2.** Ejemplo de la señal de ECG y sus componentes de detalle. De arriba a abajo: ECG,  $d_3$ ,  $d_4$  y  $d_5$ .

Seguidamente, se aplica una evaluación multicomponente en los niveles de detalle  $d_4$  y  $d_5$ , correspondientes a las bandas 7.5-15 Hz y 3.75-7.5 Hz respectivamente. Se limitan los picos detectados en el detalle 3, a los que sobrepasen el umbral  $Th_4$  y  $Th_5$  calculados según las siguientes formulas:

$$Th_4 = 0.4 * \text{percentil}_{95}(-d_4)$$

$$Th_5 = 0.2 * \text{percentil}_{95}(-d_5)$$

#### 3.3. Ajuste en tiempo de los complejos QRS

Los instantes de los complejos QRS detectados en los pasos previos son ajustados en tiempo usando la señal ECG original. Se identifica el pico de máxima amplitud en la señal ECG en una banda de tiempo de 250ms (125ms antes y después del instante R en  $d_3$ ).

### 3.4. Evaluación estadística

Los instantes R manualmente anotados por los clínicos se han utilizado como referencia para evaluar el algoritmo propuesto. Un pico detectado se considera correcto si está a menos de 100ms de la referencia. Los tres algoritmos se evalúan en términos de sensibilidad (Se): porcentaje de picos detectados correctamente; valor predictivo positivo (PPV): porcentaje de picos detectados que realmente son instantes R; y valor-F (F1): la media harmónica de Se y PPV. Las métricas de desempeño se calculan por paciente y los resultados se dan en términos de media (desviación estándar, SD) para todos los pacientes.

Los algoritmos de Martínez et al. [17] y Elola et al. [18] han sido implementados en la base de datos de PCR y las bases públicas de pacientes hemodinámicamente estables y los resultados han sido comparados con el algoritmo propuesto.

## 4. Resultados

La Tabla 1 muestra las métricas de desempeño del algoritmo para los segmentos de AESP. Se observa que la propuesta supera al algoritmo de la literatura con mejores resultados en 5.3 puntos de F1.

	Se (%)	PPV (%)	F1 (%)
Este estudio	94.4 (13.6)	85.0 (17.6)	88.0 (16.0)
Martínez et al. [16]	75.7 (16.4)	89.1 (18.9)	80.0 (16.6)
Elola et al. [17]	82.7 (28.5)	84.5 (28.7)	82.7 (28.3)

**Tabla 1.** Métricas de desempeño en segmentos de AESP, representados como media (SD)

En la Tabla 2 se muestran las métricas para los segmentos de RP. En este caso, la mejora respecto al mejor método del estado del arte es de 11.4 puntos de F1.

	Se (%)	PPV (%)	F1 (%)
Este estudio	97.0 (6.7)	93.0 (7.3)	94.3 (6.9)
Martínez et al. [16]	74.9 (10.2)	96.1 (11.4)	82.9 (10.3)
Elola et al. [17]	79.0 (33.6)	82.5 (34.4)	79.9 (33.6)

**Tabla 2.** Métricas de desempeño en segmentos de RP, representados como media (SD)

En la tabla 3 se muestran las métricas conjuntas para AESP y RP.

	Se (%)	PPV (%)	F1 (%)
Este estudio	95.8 (9.8)	89.3 (12.5)	91.4 (11.4)
Martínez et al. [16]	74.9 (11.5)	93.1 (13.8)	81.5 (11.9)
Elola et al. [17]	79.6 (31.7)	82.2 (32.6)	80.0 (31.8)

**Tabla 3.** Métricas de desempeño segmentos AESP y RP, representados como media (SD)

Finalmente, la Tabla 4 muestra las métricas de los algoritmos para pacientes hemodinámicamente estables con RP. En el caso del algoritmo de Elola et al., no funciona en los segmentos de la base de datos pública porque excede la longitud máxima admitida. Se observa, que el algoritmo propuesto proporciona resultados parecidos a los del método de Martínez et al.

	Se (%)	PPV (%)	F1 (%)
Este estudio	97.0 (5.9)	99.6 (1.2)	98.2 (4.0)
Martínez et al. [16]	97.7 (3.5)	99.5 (2.0)	98.6 (2.5)
Elola et al. [17]	-	-	-

**Tabla 4.** Métricas de desempeño de los algoritmos evaluados con las bases de datos públicas, representados como media (SD)

## 5. Discusión

La detección de complejos QRS es ampliamente utilizada para la caracterización de los distintos ritmos cardíacos en general, y de los ritmos AESP en particular. La mayoría de los métodos de caracterización publicados se han diseñado y validado con pacientes hemodinámicamente estables.

La propuesta consiste en el primer algoritmo diseñado para la detección de complejos QRS de pacientes en PCR con ritmo organizado, AESP y RP. Este estudio viene a ampliar y mejorar los métodos diseñados para la caracterización de AESP propuestos por Urteaga et al. [19].

Al comparar nuestro algoritmo con métodos diseñados para ritmos estables como el método de Martínez et al., se observa una mejora de más de 10 puntos de F1. Dos son las posibles razones para este comportamiento. Por una parte, la técnica propuesta se adapta mejor a los complejos QRS caóticos y variables, tanto en forma como en amplitud, característicos de un paciente en PCR. Esto se refuerza observando los resultados del análisis en bases de datos públicas con RP en pacientes hemodinámicamente estables, donde los complejos QRS son más regulares y ambas técnicas detectan muy bien todos los complejos. Por otra, el tratamiento de RCP implica que los intervalos de análisis del ECG están limitados a las pausas entre compresiones, con una duración de 3-10 s. A diferencia de otros métodos publicados, el algoritmo propuesto ha sido optimizado para segmentos de corta duración,

proporcionando mejores resultados tanto en ritmos AESP como en RP.

Cualquiera de los algoritmos analizados, muestra un mejor comportamiento con ritmos RP, lo que demuestra la complejidad de identificar los complejos QRS asociados a ritmos AESP, más caóticos y menos estacionarios que los del RP, y por tanto más difíciles de identificar.

Este trabajo está sujeto a una serie de limitaciones. Por una parte, la base de datos cuenta con un número limitado de pacientes, y la proporción de segmentos de AESP/RP está bastante desbalanceada, lo que hace más difícil desarrollar un método optimizado para pacientes con AESP. Por otra parte, los umbrales  $Th_3$ ,  $Th_4$  y  $Th_5$ , así como las tolerancias del 0.2 y 0.4 en han sido calculadas de forma heurística. En un futuro, con una base de datos más amplia se podría utilizar un método de optimización para su cálculo. Por último, el algoritmo asume que todos los ritmos son organizados, AESP o PR, y requeriría ser adaptado si se desea incluir los ritmos sin complejos QRS también presentes en episodios de PCR.

Este estudio viene a apoyar el desarrollo de métodos automáticos de caracterización de complejos QRS durante la PCR, que ayuden y guíen al clínico en la definición del tratamiento de resucitación. Decisiones sobre el tratamiento de la AESP, tradicionalmente basados en la experiencia previa del clínico, podrían así verse sustentadas por herramientas de ayuda al diagnóstico de la AESP, lo que redundaría en un tratamiento más adecuado y en una mayor probabilidad de supervivencia del paciente en parada.

## Agradecimientos

Este trabajo ha sido parcialmente financiado por el Ministerio de Ciencia, Innovación y Universidades a través del proyecto RTI2018-101475-BI00, en conjunto con el Fondo Europeo de Desarrollo Regional (FEDER), por el Gobierno Vasco por medio del proyecto IT1717-22 y PRE\_2021\_2\_0173 por la Universidad del País Vasco (UPV/EHU) a través del proyecto COLAB20/01.

## Referencias

- [1] Gräsner, J. T., & Bossaert, L. (2013). Epidemiology and management of cardiac arrest: what registries are revealing. *Best practice & research Clinical anaesthesiology*, 27(3), 293-306.
- [2] Atwood, C., Eisenberg, M. S., Herlitz, J., & Rea, T. D. (2005). Incidence of EMS-treated out-of-hospital cardiac arrest in Europe. *Resuscitation*, 67(1), 75-80.
- [3] Monsieurs, K. G., Nolan, J. P., Bossaert, L. L., et al. (2015). European resuscitation council guidelines for resuscitation 2015: section 1. Executive summary. *Resuscitation*, 95, 1-80.
- [4] Neumar, R. W., Shuster, M., Callaway, C. W., et al. (2015). Part 1: executive summary: 2015 American Heart Association guidelines update for cardiopulmonary resuscitation and emergency cardiovascular care. *Circulation*, 132(18\_suppl\_2), S315-S367.
- [5] Rad, A. B., Eftestøl, T., Engan, K., et al. (2017). ECG-based classification of resuscitation cardiac rhythms for retrospective data analysis. *IEEE Transactions on Biomedical Engineering*, 64(10), 2411-2418.
- [6] Urteaga, J., Aramendi, E., Elola, A., et al. (2021). A Machine Learning Model for the Prognosis of Pulseless Electrical Activity during Out-of-Hospital Cardiac Arrest. *Entropy*, 23(7), 847.
- [7] Urteaga, J., Aramendi, E., Elola, A., et al. (2021). A Machine Learning Model for the Prognosis of Pulseless Electrical Activity during Out-of-Hospital Cardiac Arrest. *Entropy*, 23(7), 847.
- [8] Bakke, M., Borgen, A., Norvik, A., et al. (2019). Physiological Effects of Adrenaline in Human Cardiac Arrest With Pulseless Electrical Activity: A Pilot Study. *Circulation*, 140(Suppl\_2), A166-A166.
- [9] Hauck, M., Studnek, J., Heffner, A. C., et al. (2015). Cardiac arrest with initial arrest rhythm of pulseless electrical activity: do rhythm characteristics correlate with outcome?. *The American Journal of Emergency Medicine*, 33(7), 891-894.
- [10] Norvik, A., Unneland, E., Bergum, D., et al. (2022). Pulseless electrical activity in in-hospital cardiac arrest—A crossroad for decisions. *Resuscitation*.
- [11] Li, C., Zheng, C., & Tai, C. (1995). Detection of ECG characteristic points using wavelet transforms. *IEEE Transactions on biomedical Engineering*, 42(1), 21-28.
- [12] Zong, W., Moody, G. B., & Jiang, D. (2003, September). A robust open-source algorithm to detect onset and duration of QRS complexes. In *Computers in Cardiology, 2003* (pp. 737-740). IEEE.
- [13] Moody, G. B., & Mark, R. G. (2001). The impact of the MIT-BIH arrhythmia database. *IEEE Engineering in Medicine and Biology Magazine*, 20(3), 45-50.
- [14] Taddei, A., Distanti, G., Emdin, M., et al. (1992). The European ST-T database: standard for evaluating systems for the analysis of ST-T changes in ambulatory electrocardiography. *European heart journal*, 13(9), 1164-1172.
- [15] Laguna, P., Mark, R. G., Goldberg, A., et al. (1997, September). A database for evaluation of algorithms for measurement of QT and other waveform intervals in the ECG. In *Computers in cardiology 1997* (pp. 673-676). IEEE.
- [16] Kalidas, V., & Tamil, L. (2017). Real-time QRS detector using stationary wavelet transform for automated ECG analysis. In *2017 IEEE 17th International Conference on Bioinformatics and Bioengineering (BIBE)* (pp. 457-461). IEEE.
- [17] Martínez, J. P., Almeida, R., Olmos, S., et al. (2004). A wavelet-based ECG delineator: evaluation on standard databases. *IEEE Transactions on biomedical engineering*, 51(4), 570-581.
- [18] Elola, A., Aramendi, E., Irusta, U., et al. (2019). ECG-based pulse detection during cardiac arrest using random forest classifier. *Medical & biological engineering & computing*, 57(2), 453-462.
- [19] Urteaga, Jon, et al. "Automated algorithm for QRS detection in cardiac arrest patients with PEA" *2022 Computing in Cardiology Conference (CinC)*. IEEE, 2022.

# Caracterización de los Cambios en las Medidas de la HRV durante Episodios de Apnea del Sueño

P. Armañac-Julián<sup>1,2,\*</sup>, A. Martín-Montero<sup>1,3</sup>, J. Lázaro<sup>1,2</sup>, S. Kontaxis<sup>1,2</sup>, D. Álvarez<sup>1,3</sup>, D. Gozal<sup>4</sup>, R. Hornero<sup>1,3</sup>, P. Laguna<sup>1,2</sup>, G. Gutiérrez-Tobal<sup>1,2</sup>, R. Bailón<sup>1,2</sup>, E Gil<sup>1,2</sup>

<sup>1</sup> CIBER-BBN, Instituto de Salud Carlos III, Madrid, España, (\*) parmanac@unizar.es

<sup>2</sup> Grupo BSICoS, Universidad de Zaragoza, I3A-IIS Aragón, España

<sup>3</sup> Grupo de Ingeniería Biomédica, Universidad de Valladolid, Valladolid, España

<sup>4</sup> Child Health Research Institute, Universidad de Missouri School of Medicine, Columbia MO, Estados Unidos

## Resumen

La literatura científica ha establecido que la apnea obstructiva del sueño (OSA) genera una dominancia del sistema nervioso simpático durante el sueño de los pacientes pediátricos afectados. Estos estudios se basan en la medida de la variabilidad de la frecuencia cardíaca (HRV). Sin embargo, no se ha realizado un análisis profundo que compare las medidas de la HRV en los propios episodios de apnea durante el sueño y en los periodos de respiración normal por separado, lo que permitiría caracterizar si el aumento de la dominancia simpática se debe solo a una respuesta autonómica instantánea causada por cada apnea o si hay una activación simpática sostenida a lo largo de toda la noche. En este trabajo, se evalúan tanto las métricas de HRV temporales como frecuenciales, junto con nuevas bandas espectrales, recientemente identificadas como específicas en 311 pacientes pediátricos con OSA. El análisis estadístico pareado no paramétrico muestra diferencias claramente significativas en la HRV de cada paciente pediátrico, comparando entre eventos de apnea y respiración normal durante el sueño. Tales hallazgos son aplicables para todas las métricas de HRV excepto RMSSD y la potencia de banda HF, cuya interpretación durante los eventos de apnea pasa a ser cuestionable. Si bien todos los estudios anteriores analizaron la HRV ya sea para los registros completos o después de eliminar los eventos de apnea, los resultados obtenidos en este trabajo resaltan la necesidad de realizar el análisis de la HRV teniendo en cuenta tanto los periodos de apnea como los de respiración normal.

## 1. Introducción

Cuando la apnea obstructiva del sueño (*Obstructive Sleep Apnea*, OSA) pediátrica es persistente, las alteraciones fisiológicas inducidas por la OSA derivan en importantes comorbilidades cardiovasculares tales como somnolencia excesiva, hipertensión o hipercolesterolemia. Estudios previos, usando polisomnografía (*Polysomnography*, PSG), revelaron una alteración del sistema nervioso autónomo (*Autonomic Nervous System*, ANS) durante el sueño en pacientes que padecían OSA moderada o severa en comparación con sujetos controles sanos [1]. En este contexto, se viene utilizando la monitorización no invasiva del ANS, a través del análisis de la variabilidad del ritmo cardíaco (*Heart Rate Variability*, HRV), para cuantificar este aumento de la dominancia simpática.

La respuesta fisiológica a la apnea tiene un patrón característico de bradicardia-taquicardia que termina con un aumento agudo de la actividad simpática, la cual puede

desencadenar una alta excitación [1]. Tal y como se ha mencionado, estas variaciones debidas a la OSA se relacionan con un mayor riesgo de morbilidad cardiovascular en los niños [2]. Sin embargo, todos los estudios que se han realizado hasta el momento han abordado las diferencias en las etapas del sueño pero han obviado por completo los efectos que tienen la presencia o no de los periodos de apnea en los valores de HRV que se obtienen [3,4], o incluso han excluido directamente de su análisis aquellos segmentos de HRV que contenían episodios apneicos [5–7]. En otras palabras, el aumento de la dominancia simpática observado en los pacientes con OSA puede estar sesgado por la activación simpática tan característica y fuerte que tiene lugar solo en respuesta a los episodios de apnea [8], y estará más sesgada todavía en los pacientes con mayor cantidad de eventos apneicos. Esto enmascara el comportamiento de cualquier índice ANS y es posible que los resultados de HRV no reflejen el estado basal del organismo [1].

El objetivo de este trabajo es caracterizar y comparar las métricas de HRV, considerando la presencia o no de eventos apneicos. Se analiza en un conjunto de pacientes pediátricos con síntomas de OSA, con la HRV obtenido en tres partes por separado: durante los propios eventos de apnea, con actividad simpática aguda; durante la respiración normal en el sueño y para los registros nocturnos completos. Esto tiene como finalidad ayudar a interpretar los índices fisiológicos de HRV, y así evaluar si el aumento de la dominancia simpática se debe solo a una respuesta autonómica instantánea a cada apnea o si hay una activación simpática sostenida a lo largo del sueño.

De manera adicional, se estudian también las diferencias en tres nuevas bandas espectrales de HRV específicas de la OSA pediátrica. Estas bandas han sido recientemente definidas y caracterizadas, y se ha demostrado que establecen una relación causal entre los cambios en la actividad en esas bandas y los cambios tanto en la gravedad como en la resolución de la OSA [9].

## 2. Materiales

En torno a 2012, se realizó un ensayo prospectivo aleatorizado para la evaluación de la prueba de adenoamigdalectomía infantil (*Childhood Adenotonsillectomy*, CHAT). Fue un ensayo diseñado para evaluar la

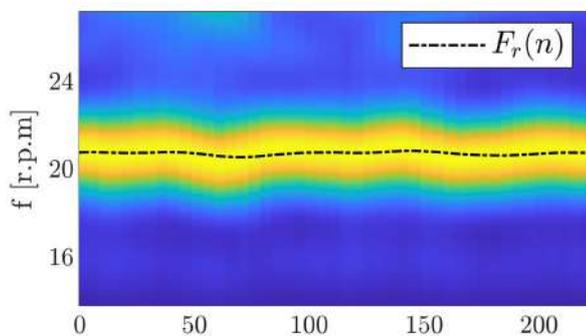
eficacia de dicho tratamiento para la OSA en pacientes pediátricos [10]. Nuestro estudio incluye a 311 de esos niños con síntomas de OSA, de entre 5-10 años. Los pacientes pediátricos fueron reclutados para una PSG nocturna en el laboratorio clínico antes del tratamiento de la OSA. Se utilizó el índice de apnea-hipopnea (*Apnea-Hypopnea Index*, AHI) para establecer la gravedad de la OSA [10], distinguiendo entre OSA leve ( $1 \leq AHI < 5$ ), moderada ( $5 \leq AHI < 10$ ) y severa ( $10 \leq AHI$ ). Como base del estudio, no había ningún sujeto sin OSA ( $AHI < 1$ ). Del total de pacientes analizados, hay 130 con OSA leve, 105 con OSA moderada y 76 con OSA severa.

De los registros de la PSG, utilizamos el electrocardiograma, y las señales respiratorias torácicas y abdominales. Además, se extraen los segmentos de apnea-hipopnea del sueño etiquetados por los clínicos especialistas de OSA, además del AHI correspondiente de cada paciente.

### 3. Métodos

#### 3.1. Análisis de la HRV guiado por respiración

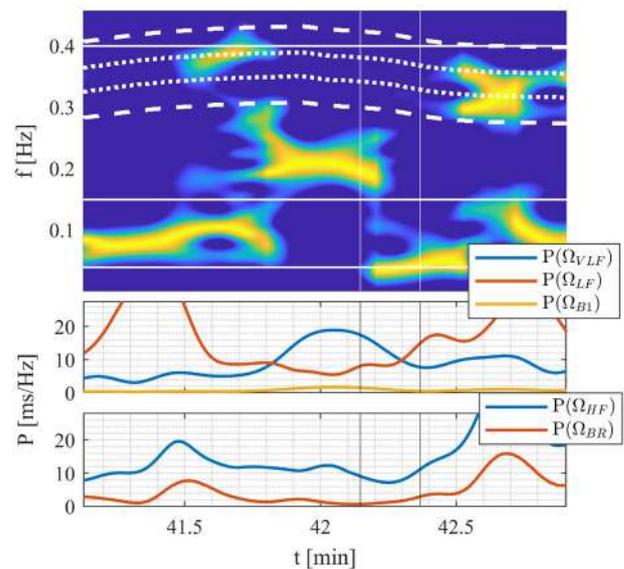
Debido a la mayor tasa respiratoria observada en niños, en ocasiones por encima de las 24 r.p.m (0.4 Hz), es imprescindible realizar un análisis de la HRV guiada por la respiración. Para ello, es necesario estimar primero la frecuencia respiratoria instantánea,  $F_r(n)$ . Esto se consigue a partir de las señales respiratorias torácica y abdominal, utilizando un método de promediado espectral condicionado por la picudez de la señal [11]. Se muestra un ejemplo ilustrativo de este algoritmo en la Figura 1.



**Figura 1.** Estimación de la frecuencia respiratoria instantánea,  $F_r(n)$ , mediante el promediado espectral condicionado por la picudez de la señal [11], en respiraciones por minuto [r.p.m.].

A continuación, los latidos del corazón se detectan en el electrocardiograma con un algoritmo basado en *Wavelets* [12]. A partir de la serie de latidos R-R, el modelo de modulación de frecuencia de pulso integral (*Integral Pulse Frequency Modulation*, IPFM) se utiliza para estimar la señal HRV [13], muestreada a 4 Hz. Después de la corrección de los latidos ectópicos [13], el intervalo de latidos normal a normal (NN) se usa para calcular las métricas temporales de la HRV: la frecuencia cardíaca media (mHR), la desviación estándar de los intervalos NN (*Standard Deviation of NN intervals*, SDNN) y la raíz cuadrada media de las diferencias sucesivas de NN (*Root Mean Square of Successive Differences*, RMSSD).

Las métricas frecuenciales de HRV son la potencia en la banda de baja frecuencia (*Low Frequency*, LF):  $\Omega_{LF} = [0,04, 0,15]$  Hz; y la potencia en la banda de muy baja frecuencia (*Very Low Frequency*, VLF):  $\Omega_{VLF} = [0,0033, 0,04]$  Hz. La banda de alta frecuencia (*High Frequency*, HF) se redefine, y es variante en el tiempo y centrada en la frecuencia respiratoria:  $\Omega_{HF}(n) = F_r(n) \pm 0,125$  Hz [8]. Además, estudios previos identificaron nuevas bandas espectrales de HRV específicas para la estratificación de la OSA pediátrica [9], siendo:  $\Omega_{B1} = [0,001, 0,005]$  Hz,  $\Omega_{B2} = [0,028, 0,074]$  Hz y  $\Omega_{BR}(n) = F_r(n) \pm 0,04/2$  Hz. Todos los parámetros de frecuencia se calculan utilizando una distribución de tiempo-frecuencia perteneciente a la clase de Cohen [14]. De esta forma, la potencia en cada banda frecuencial se monitoriza de manera continua para los registros nocturnos completos, a la frecuencia de remuestreo de 4 Hz, tal y como se puede observar en la Figura 2.



**Figura 2.** Mapa Tiempo-Frecuencia de la señal de HRV y señales continuas de potencia obtenidas. Existe un tramo de apnea marcado: 42':14''–42':36''. Las líneas blancas discontinuas y de puntos en el panel superior se corresponden con las bandas centradas en la respiración,  $\Omega_{HF}(n)$  y  $\Omega_{BR}(n)$ , respectivamente.

#### 3.2. Efectos de la Apnea sobre la HRV

Para analizar el efecto de las apneas sobre la HRV, se identifica cada uno de los segmentos de apnea-hipopnea anotados. Como se ha mencionado, la potencia frecuencial de cada banda HRV se obtiene continuamente a 4 Hz, durante toda la noche. Por tanto, a partir de las señales de potencia continuas para todo el registro, se puede realizar fácilmente un análisis por separado para los tramos de respiración normal durante el sueño y para los episodios de apnea, allá donde los episodios de apnea estén etiquetados, tal y como se puede observar en la Figura 2. Cabe destacar que, para cada evento apneico, se incluyen también los 5 segundos previos al inicio de la apnea y los 15 segundos después de su finalización, incluyendo así el patrón de compensación/recuperación de taquicardia posterior al episodio de apnea [8].

<i>(N=311)</i>	<i>En Respiración Normal</i>			<i>En Apnea</i>			<i>Registro Completo</i>			<i>p-valor</i> <i>Test Friedman</i>
	<i>leve</i>	<i>moderada</i>	<i>severa</i>	<i>leve</i>	<i>moderada</i>	<i>severa</i>	<i>leve</i>	<i>moderada</i>	<i>severa</i>	
<i>% tiempo</i>	99%	98%	95%	1%	2%	5%	100%	100%	100%	
<i>mHR</i> <sup>1</sup>	87	90	92	89	90	92	87	90	92	**
<i>SDNN</i> <sup>2</sup>	104	108	89	102	96	97	105	109	91	**
<i>RMSSD</i> <sup>2</sup>	77	78	58	60	56	60	76	77	58	.
<i>P(Ω<sub>VLF</sub>)</i> <sup>3</sup>	1,81	1,77	1,82	11,92	11,10	11,36	1,89	1,93	2,25	**
<i>P(Ω<sub>LF</sub>)</i> <sup>3</sup>	7,70	6,86	6,93	33,63	34,30	36,16	7,90	7,27	8,07	**
<i>P(Ω<sub>HF</sub>)</i> <sup>3</sup>	24,78	23,38	25,34	10,64	10,31	13,02	24,56	22,96	24,47	.
<i>P(Ω<sub>B1</sub>)</i> <sup>3</sup>	0,17	0,17	0,17	1,11	1,02	1,03	0,18	0,18	0,21	**
<i>P(Ω<sub>B2</sub>)</i> <sup>3</sup>	2,00	1,94	2,00	13,31	12,65	13,25	2,08	2,10	2,45	**
<i>P(Ω<sub>BR</sub>)</i> <sup>3</sup>	13,49	12,87	13,83	1,32	1,17	1,84	13,17	12,28	12,25	**

**Tabla 1.** Valores promedios de la mediana para las métricas de HRV de los 311 pacientes, calculados por separado en la respiración normal, en los episodios de apnea y en los registros completos. A la derecha del todo, se muestra el resultado del análisis estadístico Friedman pareado, con una marca en el caso de que exista significancia ( $p$ -valor  $< 0.01$ ). Unidades: <sup>1</sup>[b.p.m]; <sup>2</sup>[ms]; <sup>3</sup>[ms/Hz].

### 3.3. Análisis Estadístico

Para cada paciente se obtiene el valor mediano de cada una de las métricas de la HRV, de los registros completos, en los eventos de apnea y después de excluirlas. Por último, para cada paciente, se aplica una prueba Friedman de rango con signo, pareada, para evaluar las diferencias observables ( $p$ -valor  $< 0.01$ ), para cada paciente, entre los valores obtenidos en los registros completos, frente a los valores obtenidos tanto en los eventos de apnea como en los registros nocturnos después de excluir los eventos de apnea.

## 4. Resultados y Discusión

Para cada paciente y cada métrica se obtiene el valor de la mediana. Los valores promedios se muestran en la Tabla 1, por separado: en los episodios de apnea, en la respiración normal durante el sueño, después de eliminar los eventos de apnea, y para los registros completos. Además, con fines ilustrativos, los valores promedio de todos los pacientes se representan por separado con respecto a los grupos de OSA leve, moderada y severa. También se muestra la significancia estadística del análisis pareado, comparando, para cada paciente, los valores en episodios de apnea vs. valores en respiración normal durante el sueño y los valores en el registro completo.

Se observan claras diferencias significativas ( $p$ -valor  $< 0.01$ ) para todas las métricas, excepto para los parámetros RMSSD y  $P(\Omega_{HF})$ . Estas diferencias tan claras eran algo esperable, aunque no se habían reportado en estudios previos. Con estos resultados podemos afirmar la importancia de derivar índices por separado en los eventos de apnea y en la respiración normal durante el sueño. Al contrario de lo que se ha hecho hasta ahora, se debe hacer una interpretación fisiológica diferente y particular en estas dos situaciones, para evitar sesgos en la interpretación. Tanto los eventos de apnea como la respiración normal durante el sueño contienen información del SNA, pero las situaciones son totalmente diferentes: en el primero, los

pacientes están bajo un estrés severo y en el otro, están totalmente relajados.

Algunas diferencias claras son visibles en los valores promediados que comparan los resultados de apnea frente a los de no apnea. La banda  $\Omega_{B2}$  está fuertemente relacionada con la presencia de apnea, y los cambios en esta banda se atribuyeron específicamente a los mediadores de resolución de OSA [9]. El aumento repentino y significativo de las bandas  $\Omega_{B2}$ ,  $\Omega_{VLF}$  y  $\Omega_{LF}$  durante la apnea puede explicarse por el patrón característico de taquicardia-bradicardia durante la apnea, que coincide con la frecuencia de estas bandas. Notar que en promedio es posible que no se observen diferencias entre los parámetros comparando, por ejemplo, valores en respiración normal con respecto a valores en el registro completo. Sin embargo, es necesario remarcar el hecho de que se ha realizado un análisis estadístico pareado, por lo que las diferencias que se obtienen son intrínsecas a cada paciente, aunque las alteraciones en el ANS a causa de los tramos de apnea parecen enmascararse con la mayor presencia de tramos de respiración normal al analizar el registro completo.

Vale la pena mencionar el hecho de que la  $P(\Omega_{HF})$ , y el RMSSD relacionado con él, no dieron como resultado diferencias significativas. En primer lugar, como ya se ha mencionado, el uso de la banda de HF modificada, en lugar de la clásica, está específicamente incentivado por el aumento de la tasa respiratoria observado en niños, superior a 24 rpm, es decir, 0,4 Hz. Sin embargo, esta banda no mostró diferencias entre los eventos de apnea y la respiración normal durante los períodos de sueño. Por tanto, la interpretación de estos índices durante las apneas resulta cuestionable. En el caso de  $P(\Omega_{HF})$ , hay dos problemas. El primero, que es metodológico, es qué estimar y cómo estimar la frecuencia respiratoria durante los episodios de apnea. El segundo es fisiológico, como ocurre con la RMSSD, ya que estos parámetros están relacionados con la respiración y la actividad parasimpática, y es justo lo contrario de lo que ocurre durante un evento de apnea.

Si bien todos los estudios anteriores analizaron la HRV para los registros nocturnos o después de eliminar los eventos de apnea, los resultados obtenidos en este trabajo resaltan la necesidad de realizar el análisis por separado para una interpretación completa de los índices de HRV, tanto en eventos de apnea como en respiración normal. Además, es muy posible y recomendable que también sea necesario separar por etapas del sueño, con el objetivo de evaluar sus contribuciones y efectos sobre la desregulación autonómica en la OSA pediátrica [5].

De hecho, para los estudios que eliminan los segmentos de HRV con eventos apneicos [5-7], sus resultados estarán sesgados por el nivel de severidad de los pacientes, ya que han eliminado una cantidad considerablemente mayor de segmentos en los pacientes del grupo OSA severo o moderado, con respecto a los pacientes con una OSA leve.

## 5. Conclusión

Este trabajo exhibe las importantes diferencias que existen en las métricas de HRV, considerando el análisis completo, separando en episodios de apnea, con fuerte actividad simpática y los valores obtenidos en respiración normal durante el sueño. De hecho, este trabajo evidenció que la interpretación de la potencia en la banda de HF y la RMSSD pueden no ser medidas adecuadas durante los eventos de apnea.

## Agradecimientos

Este trabajo fue apoyado por CIBER-BBN (ISCIII) bajo el proyecto SleepyHeart, cofinanciado con el “Fondo Europeo de Desarrollo Regional” (FEDER), por el “Ministerio de Ciencia e Innovación” bajo los proyectos PID2020-115468RBI00, PDC2021-120775 y PID2021-126734OB-C21, y por Gobierno de Aragón (Grupo de Referencia BSICoS T39-20R).

## Referencias

[1] E. Tobaldini, L. Nobili, S. Strada, K. R. Casali, A. Braghiroli, and N. Montano, “Heart rate variability in normal and pathological sleep,” *Front. Physiol.*, vol. 4, no. October, pp. 1–11, 2013, doi: 10.3389/fphys.2013.00294.

[2] C. L. Marcus *et al.*, “Diagnosis and Management of Childhood Obstructive Sleep Apnea Syndrome,” *Pediatrics*, vol. 130, no. 3, pp. e714–e755, Sep. 2012, doi: 10.1542/peds.2012-1672.

[3] R. S. C. Horne *et al.*, “The impact of sleep disordered breathing on cardiovascular health in overweight children,” *Sleep Med.*, vol. 41, pp. 58–68, Jan. 2018, doi: 10.1016/j.sleep.2017.09.012.

[4] Y. Wu *et al.*, “Autonomic nervous function and low-grade inflammation in children with sleep-disordered

breathing,” *Pediatr. Res.*, vol. 91, no. 7, pp. 1834–1840, Jun. 2022, doi: 10.1038/s41390-021-01691-4.

[5] L. C. Nisbet *et al.*, “Nocturnal autonomic function in preschool children with sleep-disordered breathing,” *Sleep Med.*, vol. 14, no. 12, pp. 1310–1316, 2013, doi: 10.1016/j.sleep.2013.07.010.

[6] A. Vlahandonis *et al.*, “Long-term changes in heart rate variability in elementary school-aged children with sleep-disordered breathing,” *Sleep Med.*, vol. 15, no. 1, pp. 76–82, 2014, doi: 10.1016/j.sleep.2013.06.023.

[7] D. Liao *et al.*, “Sleep-disordered breathing in children is associated with impairment of sleep stage-specific shift of cardiac autonomic modulation,” *J. Sleep Res.*, vol. 19, no. 2, pp. 358–365, Mar. 2010, doi: 10.1111/j.1365-2869.2009.00807.x.

[8] J. Milagro *et al.*, “Autonomic Dysfunction Increases Cardiovascular Risk in the Presence of Sleep Apnea,” *Front. Physiol.*, vol. 10, no. 620, May 2019, doi: 10.3389/fphys.2019.00620.

[9] A. Martín-Montero *et al.*, “Heart rate variability as a potential biomarker of pediatric obstructive sleep apnea resolution,” *Sleep*, vol. 45, no. 2, pp. 1–9, Feb. 2022, doi: 10.1093/sleep/zsab214.

[10] S. Redline *et al.*, “The Childhood Adenotonsillectomy Trial (CHAT): Rationale, Design, and Challenges of a Randomized Controlled Trial Evaluating a Standard Surgical Procedure in a Pediatric Population,” *Sleep*, vol. 34, no. 11, pp. 1509–1517, Nov. 2011, doi: 10.5665/sleep.1388.

[11] S. Kontaxis *et al.*, “ECG-Derived Respiratory Rate in Atrial Fibrillation,” *IEEE Trans. Biomed. Eng.*, vol. 67, no. 3, pp. 905–914, Mar. 2020, doi: 10.1109/TBME.2019.2923587.

[12] J. P. Martínez, R. Almeida, S. Olmos, A. P. Rocha, and P. Laguna, “A Wavelet-Based ECG Delineator: Evaluation on Standard Databases,” *IEEE Trans. Biomed. Eng.*, vol. 51, no. 4, pp. 570–581, Apr. 2004, doi: 10.1109/TBME.2003.821031.

[13] R. Bailón, G. Laouini, C. Grao, M. Orini, P. Laguna, and O. Meste, “The Integral Pulse Frequency Modulation Model With Time-Varying Threshold: Application to Heart Rate Variability Analysis During Exercise Stress Testing,” *IEEE Trans. Biomed. Eng.*, vol. 58, no. 3, pp. 642–652, Mar. 2011, doi: 10.1109/TBME.2010.2095011.

[14] M. Orini, R. Bailon, L. T. Mainardi, P. Laguna, and P. Flandrin, “Characterization of Dynamic Interactions Between Cardiovascular Signals by Time-Frequency Coherence,” *IEEE Trans. Biomed. Eng.*, vol. 59, no. 3, pp. 663–673, Mar. 2012, doi: 10.1109/TBME.2011.2171959.

Contacto: Pablo Armañac-Julián, Universidad de Zaragoza, I3A, Ed. Ada Byron, Zaragoza (España), parmanac@unizar.es

# Métodos de filtrado adaptativo para la restauración de la forma de onda del ECG durante la reanimación cardiopulmonar

Alvaro Iza<sup>1</sup>, Andoni Elola<sup>2</sup>, Iraia Isasi<sup>3</sup>, Erik Alonso<sup>3</sup>, Elisabete Aramendi<sup>1</sup>

<sup>1</sup>Departamento de Ingeniería de Comunicaciones, Universidad del País Vasco (UPV/EHU), Bilbao, España, {alvaro.iza, elisabete.aramendi}@ehu.eus.

<sup>2</sup>Departamento de Tecnología Electrónica, Universidad del País Vasco (UPV/EHU), Eibar, España, {andoni.elola}@ehu.eus.

<sup>3</sup>Departamento de Matemática Aplicada, Universidad del País Vasco (UPV/EHU), Bilbao, España, {iraia.isasi, erik.alonso}@ehu.eus.

## Resumen

*Las compresiones torácicas inducen artefactos en el electrocardiograma (ECG) que dificultan la toma de decisiones clínicas fiables durante la reanimación cardiopulmonar (RCP). Este estudio analiza la eficacia de diferentes filtros de supresión del artefacto RCP. Para su desarrollo, se mezclaron de forma artificial ECGs libres de artefacto y artefactos RCP adquiridos durante asistolia a diferentes relaciones de señal a ruido. Se testaron varias configuraciones de 4 filtros del estado del arte: recursive least squares, least mean squares, el filtro de Kalman y una versión modificada de este último. El rendimiento de los filtros se evaluó comparando los ECGs libres de artefacto y los ECGs filtrados en términos de la relación señal a ruido restaurada ( $SNR_{res}$ ) y el coeficiente de correlación de Pearson (PCC). El Kalman Smoother resultó ser la mejor estrategia para la eliminación del artefacto RCP, obteniendo una mejora media de 1.4 dBs en SNR y un PCC medio de 0.77, respectivamente, en el punto de trabajo óptimo. En conclusión, este trabajo propone métodos de filtrado del artefacto RCP adecuados para la restauración de la forma de onda del ECG, lo que permitiría mejorar la fiabilidad del análisis del ECG durante las compresiones torácicas.*

## 1. Motivación

La reanimación cardiopulmonar (RCP) de alta calidad y la desfibrilación temprana son fundamentales para la supervivencia de los pacientes en parada cardiorespiratoria extrahospitalaria (PCREH). En particular, la administración de compresiones torácicas (CTs) ininterrumpidas es de vital importancia durante la RCP [2]. Mientras que los profesionales de soporte vital básico se basan en el algoritmo de decisión de shock implementado en los desfibriladores comerciales para decidir si el paciente necesita una descarga eléctrica o no, los médicos de soporte vital avanzado evalúan visualmente el ECG para determinar las intervenciones terapéuticas adecuadas. Desafortunadamente, en ambos escenarios, las CTs deben ser interrumpidas ya que introducen interferencias en el ECG que impiden su correcta interpretación. Estas interrupciones, comprometen la presión de perfusión coronaria, y disminuyen la probabilidad de supervivencia del paciente en PCREH [2].

En la literatura se han propuesto varios filtros adaptativos para eliminar los artefactos RCP del ECG, desde filtros que utilizan señales de referencia adicionales correlacionadas con el artefacto, hasta filtros más simples, pero menos eficaces que analizan únicamente el ECG [2]. Aprovechando la naturaleza cuasiperiódica del artefacto RCP, se han explorado también filtros adaptativos basados en el modelo multiarmónico del artefacto [4]. Estos últimos son los que ofrecen, hasta la fecha, el mejor compromiso entre simplicidad y rendimiento. Durante las CTs, la forma de onda del ECG subyacente se desconoce, y, por lo tanto, el ECG filtrado no puede ser directamente comparado con el ECG real del paciente, lo que complica la medición del rendimiento del filtro. Es por esto que, la mayoría de los filtros adaptativos desarrollados hasta la fecha se han evaluado de forma indirecta en términos de sensibilidad y especificidad de un algoritmo de decisión de shock comercial aplicado al ECG filtrado.

El objetivo de este trabajo es abordar el problema que acabamos de mencionar mediante el uso de mezclas artificiales de ECGs libres de artefacto y artefactos RCP obtenidos en ausencia de actividad eléctrica del corazón (asistolia) para evaluar el rendimiento de varios filtros adaptativos del estado del arte. El modelo de mezcla permite conocer el ritmo subyacente del paciente, y, por lo tanto, el rendimiento del filtro puede evaluarse en términos de restauración de la forma de onda del ECG filtrado. De esta forma, se podrían determinar configuraciones de filtro adecuadas para la restauración de la forma de onda del ECG durante la RCP permitiendo la toma de decisiones clínicas fiables sin interrumpir las CTs.

## 2. Materiales

Los artefactos RCP de este estudio se obtuvieron de un ensayo clínico prospectivo diseñado para medir la calidad de la RCP durante la PCREH. Los detalles del estudio pueden encontrarse en [5]. Se exportaron a Matlab y se remuestrearon a 250 Hz las señales de ECG y profundidad de compresión (PC) adquiridas mediante los desfibriladores Heartstart 4000 (Philips Medical Systems, Andover, USA). De este estudio, se extrajo un subconjunto de 1192 segmentos de 10 s de duración correspondientes a 177 pacientes con señales simultáneas de ECG y PC. Estos

segmentos se adquirieron en presencia de asistolia y durante CTs. La asistolia es la ausencia de actividad eléctrica en el corazón. Por lo tanto, en estos intervalos la única actividad en el ECG es la inducida por las CTs. Estos segmentos permiten simular los artefactos RCP en las mezclas artificiales.

Los ECGs libres de artefacto utilizados en el modelo de mezcla lineal, se extrajeron de 3 bases de datos públicas: CUDB, VFDB y AHADB [6]. La base de datos final consiste en 5724 segmentos de ECG de 10 s de duración correspondientes a 67 pacientes. Estos segmentos incluyen tanto ritmos desfibrilables como no-desfibrilables.

### 3. Métodos

#### 3.1. Modelo de mezcla

El ECG corrupto artificial,  $x(n)$ , se obtiene mediante la mezcla de un segmento ECG libre de artefacto,  $s_{\text{ecg}}(n)$ , y un segmento de artefacto RCP,  $s_{\text{rcp}}(n)$ :

$$x(n) = s_{\text{ecg}}(n) + \alpha s_{\text{rcp}}(n) \quad (1)$$

La relación señal a ruido ( $SNR$ ) de  $x(n)$  está controlado por el peso del valor  $\alpha$ :

$$SNR_{\text{in}} = 10 \log_{10} \left( \frac{P_{\text{ecg}}}{\alpha^2 P_{\text{rcp}}} \right) \quad (2)$$

donde  $P_{\text{ecg}}$  y  $P_{\text{rcp}}$  son la potencia de  $s_{\text{ecg}}(n)$  y  $s_{\text{rcp}}(n)$ , respectivamente. Por cada paciente, primero, se seleccionaron 4 segmentos de ECG de forma aleatoria (2 correspondientes a ritmos desfibrilables y 2 a ritmos no-desfibrilables), obteniendo un total de  $67 \cdot 4 = 268$  segmentos. Éstos se mezclaron con 15 artefactos RCP,  $s_{\text{rcp}}(n)$ , seleccionados al azar para 7 valores diferentes de  $SNR_{\text{in}}$ , que van desde -15 dB hasta 15 dB en pasos de 5 dB. Por lo tanto, la base de datos artificial final está compuesta por  $268 \cdot 15 \cdot 7 = 28140$  ECGs corruptos.

#### 3.2. Filtrado adaptativo

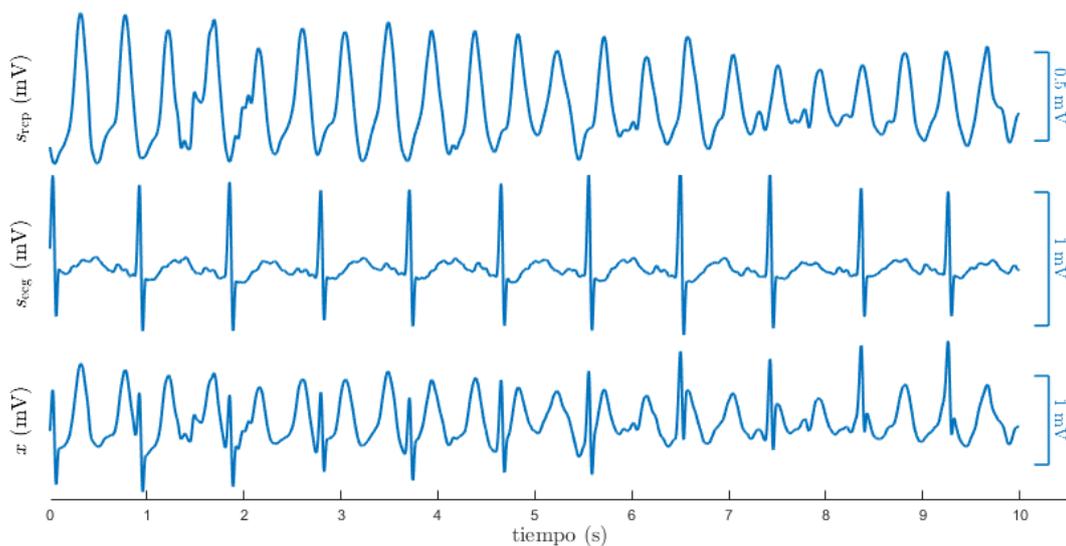
Durante las CTs, el artefacto RCP se modela como una interferencia cuasiperiódica compuesta por  $N$  armónicos y una frecuencia fundamental variable,  $f_0(n)$ :

$$\hat{s}_{\text{rcp}}(n) = \sum_{k=1}^N a_k(n) \cos(k\omega_0(n)n) + b_k(n) \sin(k\omega_0(n)n) \quad (4)$$

$$f_0(n) = \frac{1}{t_k - t_{k-1}} \quad t_{k-1} < nT_s \leq t_k \quad (5)$$

donde la frecuencia angular discreta es  $\omega_0(n) = 2\pi f_0(n)T_s$ ,  $T_s$  es el periodo de muestreo y  $t_k$  son los instantes de compresión (en segundos) que se marcaron automáticamente en la señal de PC utilizando un detector de picos negativos con un umbral de -1.5 cm. Los coeficientes de Fourier,  $a_k(n)$  y  $b_k(n)$ , definen el filtro adaptativo que se ajusta a las características, variables en el tiempo, del artefacto. El ECG estimado,  $\hat{s}_{\text{ecg}}(n)$ , se obtuvo sustrayendo la estimación del artefacto,  $\hat{s}_{\text{rcp}}(n)$ , del ECG corrupto artificial,  $x(n)$ .

En este estudio se utilizan los siguientes filtros para la estimación de los coeficientes  $a_k(n)$  y  $b_k(n)$ : recursive least squares (RLS) [7], least mean squares (LMS) [4], el filtro de Kalman y una versión modificada de este último, Kalman Smoother [8]. Todos los filtros emplean criterios para minimizar el error entre  $x(n)$  y  $\hat{s}_{\text{rcp}}(n)$  en los armónicos de  $f_0(n)$ . En resumen, el filtro LMS actualiza sus coeficientes en cada instante  $n$  utilizando incrementos proporcionales al error cuadrático y al tamaño de paso  $\mu$  [4]. El filtro RLS amplía la ventana de observación del error cuadrático mediante un factor de olvido exponencial,  $\lambda$  [7]. Por último, los filtros Kalman y Kalman Smoother se basan en un modelo de variables de estado donde la varianza del ruido de observación,  $q$ , y el factor de olvido exponencial,  $\lambda$ , controlan la tasa de ajuste de los coeficientes [8].



**Figura 1.** Modelo de mezcla. De arriba abajo: un segmento de artefacto RCP registrado durante CTs y en presencia de ritmo asistólico, un segmento de ECG libre de artefacto registrado durante la ausencia de CTs y un segmento de ECG artificialmente corrupto derivado de la mezcla de las señales anteriores para un  $SNR_{\text{in}}$  de -5 dB.

Se probaron diferentes número de armónicos y coeficientes de adaptabilidad para cada filtro dentro de los siguientes rangos de trabajo:  $1 \leq N \leq 10$ ,  $10^{-4} \leq \mu \leq 10^{-2}$  para el LMS,  $0.985 \leq \lambda \leq 0.9999$  para el RLS y  $10^{-3} \leq \lambda \leq 0.6/3 \cdot 10^{-3} \leq q \leq 10^{-1}$  para los filtros Kalman y Kalman Smoother. La mayor diferencia entre estos dos últimos filtros, radica en que el Kalman Smoother trabaja de forma bidireccional. Con el fin de agilizar la carga computacional, el punto de trabajo óptimo del Kalman Smoother se calculó en base al Kalman original.

### 3.3. Métricas de rendimiento

Se calcularon dos métricas para evaluar el rendimiento de los filtros en términos de restauración de la forma de onda del ECG: el  $SNR$  del ECG filtrado/restaurado ( $SNR_{res}$ ) y el coeficiente de correlación de Pearson ( $PCC$ ) que mide la similitud entre el ECG filtrado,  $\hat{s}_{ecg}(n)$ , y el ECG libre de artefacto,  $s_{ecg}(n)$ :

$$SNR_{res} = 10 \log_{10} \left( \frac{P_{ecg}}{P_e} \right) \quad (6)$$

$$PCC = \frac{\sum_{n=1}^L s_{ecg}(n) \cdot \hat{s}_{ecg}(n)}{\sqrt{\sum_{n=1}^L s_{ecg}^2(n) \cdot \sum_{n=1}^L \hat{s}_{ecg}^2(n)}} \quad (7)$$

donde  $P_{ecg}$  y  $P_e$  son las potencias de  $s_{ecg}(n)$  y del error de estimación dado por  $e(n) = s_{ecg}(n) - \hat{s}_{ecg}(n)$ , respectivamente.

La métricas de rendimiento se calcularon sobre el intervalo 2-10 s de la señal de ECG estimada,  $\hat{s}_{ecg}(n)$ . Se descartaron los primeros 2 s para evitar los transitorios de los filtros adaptativos.

## 4. Resultados

La Tabla 1 muestra la mejor configuración de cada filtro en términos del  $SNR_{res}$  medio obtenido sobre todo el rango de valores de  $SNR_{in}$  testeados. Además, se muestran los valores medios de  $SNR_{res}$  y  $PCC$  obtenidos mediante dichas configuraciones. Como se puede apreciar, el filtro Kalman Smoother fue significativamente la mejor opción (valor-p < 0.05), obteniendo valores de  $SNR_{res}$  y  $PCC$  medios de 4.8 dB y 0.77, respectivamente.

**Tabla 1.** Configuraciones óptimas de los filtros adaptativos junto con los valores de  $SNR_{res}$  y  $PCC$  medios obtenidos mediante dichas configuraciones.

	Configuración	$SNR_{res}$ (dB)	PCC
RLS	$N = 2$ $\lambda = 0.9994$	3.3	0.74
LMS	$N = 4$ $\mu = 0.0013$	3.2	0.74
Kalman	$N = 5$ $\lambda = 0.068$ $q = 0.033$	3.4	0.74
Kalman Smoother	$N = 5$ $\lambda = 0.068$ $q = 0.033$	4.8	0.77

El panel izquierdo de la Figura 2 se corresponde con un gráfico tridimensional de superficie donde se muestran por

colores los valores de  $SNR_{res}$  medio en función de los parámetros de adaptabilidad del filtro Kalman (filtro utilizado también para la optimización del Kalman Smoother) para  $N = 5$ . Aunque la configuración que mayor  $SNR_{res}$  obtiene sea la indicada en la Tabla I, se consiguen valores de  $SNR_{res} > 3$  dB para una amplia gama de configuraciones de  $\lambda$  y  $q$ .

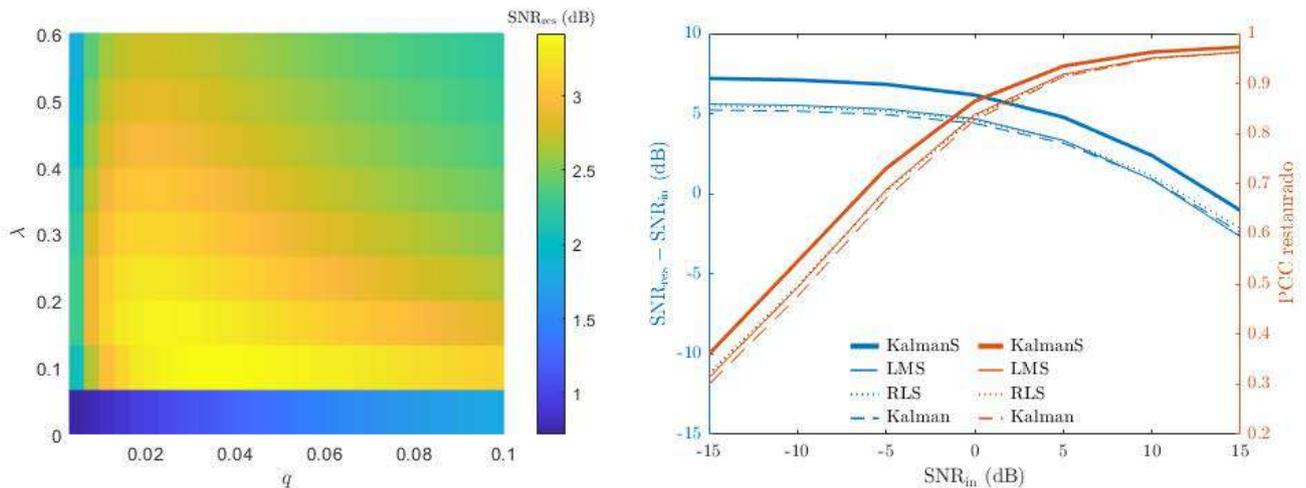
El panel derecho de la Figura 2 muestra el incremento medio obtenido en el  $SNR$  ( $SNR_{res} - SNR_{in}$ ) y los valores medios de  $PCC$  alcanzados para cada  $SNR_{in}$  cuando se utilizan las configuraciones óptimas de cada filtro. De nuevo se puede apreciar que el rendimiento del Kalman Smoother tanto en términos de la mejora del  $SNR$  como del  $PCC$  es superior al resto de filtros independientemente del  $SNR_{in}$  aplicado. A niveles de corrupción elevados (hasta -5 dB), el aumento de la  $SNR$  fue superior a 6.5 dB para el Kalman Smoother, mientras que el resto de filtros apenas superaban los 5 dB. A medida que aumenta el  $SNR_{in}$  la mejora en el  $SNR$  es menor para todos los filtros. Aun así, para un  $SNR_{in} = 10$  dB, donde la potencia del ECG limpio es 10 veces mayor a la del artefacto RCP, el ruido se reduce en 2.4 dB en la salida del Kalman Smoother y aproximadamente en 1.1 dB en la salida del resto de filtros. Respecto al  $PCC$ , la similitud entre el ECG restaurado y el ECG limpio aumenta a medida que el valor de  $SNR_{in}$  incrementa, obteniendo valores de  $PCC$  superiores a 0.7 (Kalman Smoother) y 0.6 (resto de filtros) para valores de  $SNR_{in} > -5$  dB.

## 5. Conclusiones

El presente estudio proporciona una evaluación de la restauración de la forma de onda del ECG tras la cancelación adaptativa de los artefactos RCP. El modelo de mezcla nos ha permitido proponer diferentes índices de calidad de la señal que proporcionan una buena estimación de la precisión con la que el ritmo subyacente del paciente se puede restablecer cuando se aplican las CTs.

Se compararon 4 filtros adaptativos del estado del arte utilizando el  $SNR_{res}$  y el  $PCC$  como índices de calidad de la señal filtrada: RLS, LMS, Kalman y Kalman Smoother. El Kalman Smoother resultó ser con un amplio margen la mejor estrategia para la eliminación del artefacto RCP tanto en términos de  $SNR_{res}$  como de  $PCC$  y para cualquier nivel de  $SNR_{in}$  de entrada.

En conclusión, se han determinado configuraciones de filtro adecuadas para restaurar el ECG durante la RCP, lo que permitiría realizar una interpretación fiable del ECG sin interrumpir la terapia de las CTs.



**Figura 2.** El gráfico de la izquierda muestra por colores los valores de  $SNR_{res}$  medios obtenidos por el filtro Kalman para todo el rango de valores de  $SNR_{in}$  en función de los parámetros de adaptabilidad del filtro,  $\lambda$  y  $q$  (el número de armónicos se fija a la configuración óptima,  $N=5$ ). El panel derecho muestra la mejora media obtenida en el SNR ( $SNR_{res} - SNR_{in}$ ) y los valores medios de PCC alcanzados tras el filtrado en términos de  $SNR_{in}$ .

## Agradecimientos

Este trabajo fue financiado por el Ministerio Español de Ciencia Innovación y universidades (RTI2018-101475-BI00) conjuntamente con el Fondo Europeo de Desarrollo Regional (FEDER), por el Gobierno Vasco bajo la convocatoria IT1717-22 y por la Universidad del País Vasco (UPV/EHU) con la ayuda COLAB20/01.

## Referencias

- [1] Perkins GD et al. European Resuscitation council guidelines for resuscitation 2015: Section 2. Adult basic life support and automated external defibrillation 2015. *Resuscitation* 2015, vol 85: pp 81-89.
- [2] Vaillancourt C et al. The impact of increased chest compression fraction on return of spontaneous circulation for out-of-hospital cardiac arrest patients not in ventricular fibrillation. *Resuscitation* 2011, vol 82(12): pp. 1501-1507.
- [3] De Gauna S.R et al. Rhythm analysis during cardiopulmonary resuscitation: past, present, and future. *BioMed research international* 2014, vol 2014: e386010.
- [4] Irusta U et al. A least mean-square filter for the estimation of the cardiopulmonary resuscitation artifact based on the frequency of the compressions. *IEEE Transactions on Biomedical Engineering* 2009, vol 56(4): pp. 1052-1062.
- [5] Wik L et al. Quality of cardiopulmonary resuscitation during out-of-hospital cardiac arrest. *Jama* 2005, vol 293(3): pp. 299-304.
- [6] Picon A et al. Mixed convolutional and long short-term memory network for the detection of lethal ventricular arrhythmia. *PloS one* 2019, vol 14(5): e0216756.
- [7] Isasi I et al. A multistage algorithm for ECG rhythm analysis during piston-driven mechanical chest compressions. *IEEE Transactions on Biomedical Engineering* 2018, vol 66(1): pp. 263-272.
- [8] Elola A et al. Multimodal algorithms for the classification of circulation states during out-of-hospital cardiac arrest. *IEEE Transactions on Biomedical Engineering* 2020, vol 68(6): 1913-1922.

# Análisis de la variabilidad del ritmo cardíaco durante la realización de apneas controladas en sujetos sanos

B.F. Giraldo Giraldo<sup>1,2,3</sup>, D. Ferré López<sup>1</sup>, J. Solà Soler<sup>1,2,3</sup>

<sup>1</sup> Universitat Politècnica de Catalunya (UPC), Dept. ESAII – EEBE, Barcelona, España, {Beatriz.Giraldo, Jordi.Sola}@upc.edu, david.ferre.lopez@estudiantat.upc.edu

<sup>2</sup> Institute of Bioengineering of Catalonia (IBEC), Barcelona, España, {bgiraldo, jsola}@ibecbarcelona.eu

<sup>3</sup> CIBER de Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), España

## Resumen

*La interacción cardiorrespiratoria está relacionada con la variabilidad de la frecuencia cardíaca (HRV) y la respiración. La regulación de estos sistemas, cardíaco y respiratorio es gestionada por el sistema nervioso autónomo (SNA). Según el método Wim Hof (WHM), basado en respiraciones profundas controladas, seguidas de apneas, se puede controlar la respuesta del SNA en el sistema cardiovascular. Este trabajo propone un análisis temporal y frecuencial de las etapas del método WHM, a partir de las señales ECG (RR) y de volumen respiratorio, de 36 sujetos sanos (20 hombres, 16 mujeres). Se han definido 11 intervalos: inicio de respiración basal (SRI=1), respiraciones profundas controladas (CDB=3), apnea espiratoria larga (LEA=3), apnea inspiratoria corta (SIA=3), final de respiración basal (FRI=1). Al analizar todos los sujetos, se observan diferencias estadísticamente significativas en los valores medios de la frecuencia cardíaca entre SRI y FRI ( $p<0.001$ ), entre las medias cuadrática de RR en los intervalos CDB ( $p<0.01$ ), y entre la variabilidad de RR en CDB1 y CDB2 ( $p<0.001$ ). La amplitud de la frecuencia de pico máxima de la banda frecuencial alta (HF) presenta diferencias estadísticamente significativas entre SRI y FRI ( $p=0.016$ ), y entre los intervalos CDB ( $p=0.009$ ). Entre sexos, la máxima frecuencia de pico de HF presenta diferencias en CDB1 ( $p<0.05$ ), y en CDB2 y CDB3 ( $p<0.01$ ). En conclusión, el método WHM ha presentado disminución en la frecuencia cardíaca de los sujetos al final del protocolo. Ejercicios de respiración especiales podrían ayudar a mantener el cuerpo en condiciones óptimas, y su control en diferentes condiciones.*

## 1. Introducción

El sistema nervioso autónomo (SNA) contribuye a la regulación de procesos fisiológicos involuntarios como la frecuencia cardíaca, la presión arterial y la respiración, entre otros. La variabilidad de la frecuencia cardíaca (HRV) es una variable fisiológica ampliamente utilizada que evalúa de forma no invasiva el sistema nervioso autónomo cardíaco midiendo los cambios en el ritmo cardíaco a lo largo del tiempo [1, 2]. La HRV se considera como un reflejo de los cambios de las ramas simpática y parasimpática cardíaca. Está asociada a la frecuencia cardíaca y relacionada con la frecuencia respiratoria de las personas y, por tanto, alteraciones en estos parámetros pueden producir cambios en la HRV [3].

La variabilidad de los sistemas cardíaco y respiratorio, así como su interacción, pueden analizarse mediante técnicas de respiración controlada. El método Wim Hof (WHM) propone variaciones controladas en la respiración que podrían influir en la respuesta del sistema endocrino e

inmune de las personas. La cantidad de oxígeno que se inhala en una respiración influye en la cantidad de energía que se libera en las células del cuerpo. Según la descripción de este método, la forma en que se respira afecta fuertemente las actividades químicas y fisiológicas del cuerpo [4].

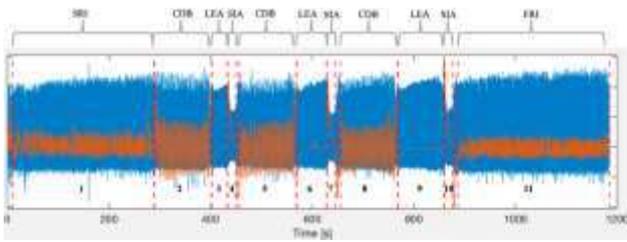
En este estudio se propone analizar la variabilidad del ritmo cardíaco y de la frecuencia cardíaca durante la realización de respiraciones controladas, incluyendo apneas con diferentes períodos de duración. A partir de señales cardíaca (ECG) y de volumen respiratorio, registradas aplicando el método WHM, se obtuvieron parámetros temporales y frecuenciales para analizar la respuesta del sistema cardíaco en función de los cambios en la respiración, así como según el sexo.

## 2. Método Wim Hof

La técnica de respiración del método Wim Hof (WHM) consiste en realizar respiraciones profundas controladas seguidas de una apnea espiratoria larga y finalmente una apnea inspiratoria corta [4]. El protocolo implementado consiste en una respiración basal durante 4.5 minutos al inicio y al final, y tres etapas de respiraciones controladas. Las etapas de respiraciones controladas comienzan con un primer período de 30 respiraciones profundas durante 50 segundos, seguidos de una apnea espiratoria larga de 30 segundos, y finalmente una apnea inspiratoria corta de 15 segundos. En la segunda etapa la apnea espiratoria larga es de 60 segundos, y en la tercera es de 90 segundos. En total se registraron 11 intervalos.

En este estudio se ha trabajado con una señal ECG (derivación II), y las señales de volumen respiratorio (torácico y abdominal), obtenidas mediante bandas pletismográficas por impedancia. Las señales se registraron con un sistema Biopac Systems Inc MP160, a una frecuencia de muestreo de 500 Hz, en la Escuela de Ingeniería de Barcelona Est, Universitat Politècnica de Catalunya. Todos los sujetos firmaron su correspondiente consentimiento de participación en el estudio.

La figura 1 es un ejemplo del protocolo aplicado, donde se presentan una señal de volumen respiratorio y una señal ECG con los correspondientes intervalos: intervalo inicial de relajación (SRI), respiraciones profundas controladas (CDB), apnea espiratoria larga (LEA), apnea inspiratoria corta (SIA), e intervalo final de relajación (FRI).



**Figura 1.** Señal de volumen respiratorio (naranja) superpuesta a la señal ECG (azul), de un sujeto sano. SRI: (1) intervalo inicial de relajación, CDB: (2, 5, 8) respiraciones profundas controladas, LEA: (3, 6, 9) apnea espiratoria larga, SIA: (4, 7, 10) apnea inspiratoria corta, y FRI: (11) intervalo final de relajación

### 3. Material y métodos

#### 3.1. Base de datos

La tabla 1 presenta la información demográfica de los 36 sujetos sanos analizados en este estudio.

	Nº	Edad (Años)	Altura (m)	Peso (kg)
<b>Hombres</b>	20	22.2 ± 1.4	1.8 ± 0.1	74.8 ± 6.3
<b>Mujeres</b>	16	21.1 ± 1.1	1.7 ± 0.1	57.1 ± 3.5
<b>Total</b>	36	21.7 ± 1.4	1.7 ± 0.1	66.7 ± 1.3

**Tabla 1.** Información demográfica sujetos sanos (media ± desviación estándar)

Adicionalmente, se anotó si los sujetos eran fumadores, y se registró información relacionada con la actividad deportiva, considerando un hábito regular quien practicaba deporte 5 horas semanales o más (Tabla 2).

A vs B	Hombres vs Mujeres	Fumadores vs no-Fumadores	Deporte vs no-Deporte
<b>A</b>	20	9	15
<b>B</b>	16	26	21
<b>Total</b>	36	35	36

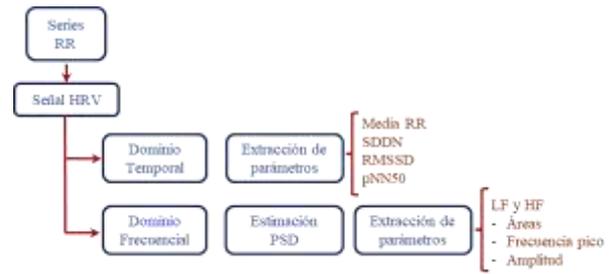
**Tabla 2.** Grupos analizados: Hombres (A) vs Mujeres (B), Fumadores (A) vs no fumadores (B) y Deporte (A) vs no deporte (B)

#### 3.2. Procesado de las señales

En primer lugar, las señales fueron pre-procesadas, eliminando los artefactos y valores atípicos. Luego, estas fueron procesadas para obtener las series temporales de los intervalos de duración de cada ciclo, utilizando algoritmos propios basados en umbrales y cruces por cero. A partir de la señal ECG se obtuvieron los intervalos RR, y se calculó la HRV.

Del dominio temporal se determinaron los siguientes parámetros: frecuencia cardíaca (HR), valor medio del intervalo RR, su desviación estándar (SDNN), raíz cuadrada de las diferencias medias cuadráticas entre intervalos RR sucesivos (RMSSD), porcentaje de pares de intervalos RR sucesivos que difieren en más de 50 ms (pNN50). Para el análisis espectral, se interpoló la señal RR a 4 Hz., y luego se obtuvo la densidad espectral de potencia (PSD) considerando las bandas de baja frecuencia LF [0.04 a 0.15 Hz] y alta frecuencia HF [0.15 a 0.4 Hz] [5]. Para cada banda frecuencial se determinó el área del

espectro ( $A$ ), la frecuencia del pico máximo ( $f_k$ ) y su amplitud ( $A_k$ ) (Figura 2).



**Figura 2.** Parámetros obtenidos a partir de las series temporales RR

#### 3.3. Análisis estadístico

Para determinar diferencias estadísticamente significativas ( $p \leq 0.05$ ) al comparar los diferentes intervalos y grupos de sujetos (por sexo, fumadores o no, con actividad deportiva regular o no), se utilizó la prueba estadística no-paramétrica de Wilcoxon.

### 4. Resultados y discusión

La tabla 3 y 4 presentan los resultados obtenidos con los parámetros temporales y frecuenciales extraídos de las series RR, para todos los sujetos.

Intervalo	Time-domain values for all subjects				
	HR [BPM]	RR [ms]	SDNN [ms]	RMSSD [ms]	pNN50 [%]
1 SRI	74.5 ± 7.3	821 ± 81	71.9 ± 26.5	20.1 ± 7.3	2.33 ± 2.98
2 CDB	88.7 ± 10.9	692 ± 82	57.2 ± 23.3	17.5 ± 10.9	2.04 ± 4.56
3 LEA	80.1 ± 10.3	768 ± 94	66.0 ± 28.0	10.2 ± 10.3	0.96 ± 4.03
4 SIA	79.4 ± 12.5	787 ± 130	85.0 ± 51.0	19.1 ± 12.5	5.00 ± 10.90
5 CDB	86.6 ± 9.1	710 ± 70	82.0 ± 33.9	22.0 ± 9.1	3.65 ± 6.33
6 LEA	78.1 ± 9.2	789 ± 94	79.2 ± 30.7	12.7 ± 9.2	1.35 ± 3.54
7 SIA	82.7 ± 14.9	760 ± 132	76.7 ± 54.3	18.3 ± 14.9	3.76 ± 10.30
8 CDB	84.9 ± 7.7	722 ± 62	75.9 ± 30.3	20.6 ± 7.7	2.83 ± 5.49
9 LEA	78.0 ± 9.7	791 ± 103	78.7 ± 23.5	11.2 ± 9.9	0.80 ± 1.41
10 SIA	81.3 ± 12.2	770 ± 123	94.4 ± 42.0	24.2 ± 12.2	5.49 ± 11.70
11 FRI	69.7 ± 6.3	877 ± 79	78.6 ± 27.4	20.5 ± 6.3	2.55 ± 3.58

**Tabla 3.** Parámetros temporales, en términos de media ± desviación estándar, de los 11 intervalos de señal. HR: frecuencia cardíaca; RR: Intervalo de tiempo entre dos picos consecutivos de la onda R; SDNN: desviación estándar de RR; RMSSD: raíz cuadrada de las diferencias medias cuadráticas entre intervalos RR; pNN50: porcentaje de pares de intervalos RR sucesivos que difieren en más de 50 ms

Intervalo	Area LF		Area HF	
	$\bar{x}$ [ $s^2$ ]	$\sigma$ [ $s^2$ ]	$\bar{x}$ [ $s^2$ ]	$\sigma$ [ $s^2$ ]
1 SRI	2.0E-03 ± 1.7E-03	2.6E-03 ± 3.2E-03	1.5E-03 ± 1.3E-03	2.6E-03 ± 2.6E-03
2 CDB	8.0E-04 ± 6.0E-04	1.1E-03 ± 7.0E-04	1.2E-03 ± 1.5E-03	1.0E-03 ± 2.5E-03
3 LEA	2.2E-03 ± 1.9E-03	2.8E-03 ± 3.5E-03	8.0E-04 ± 9.0E-04	1.0E-03 ± 1.5E-03
4 SIA	4.0E-03 ± 4.7E-03	3.4E-03 ± 3.8E-03	2.1E-03 ± 5.4E-03	1.6E-03 ± 2.9E-03
5 CDB	1.2E-03 ± 1.1E-03	1.6E-03 ± 1.3E-03	1.7E-03 ± 2.2E-03	1.6E-03 ± 1.7E-03
6 LEA	1.5E-03 ± 1.1E-03	1.9E-03 ± 1.5E-03	8.0E-04 ± 1.4E-03	1.1E-03 ± 1.9E-03
7 SIA	4.4E-03 ± 4.3E-03	3.6E-03 ± 2.9E-03	1.8E-03 ± 4.3E-03	1.4E-03 ± 2.6E-03
8 CDB	1.0E-03 ± 8.0E-04	1.6E-03 ± 1.6E-03	1.6E-03 ± 2.0E-03	1.3E-03 ± 1.6E-03
9 LEA	1.4E-03 ± 1.0E-03	2.0E-03 ± 2.0E-03	6.0E-04 ± 6.0E-04	1.2E-03 ± 1.5E-03
10 SIA	3.9E-03 ± 3.5E-03	3.4E-03 ± 3.3E-03	2.2E-03 ± 4.2E-03	1.7E-03 ± 2.3E-03
11 FRI	2.2E-03 ± 1.9E-03	2.7E-03 ± 2.3E-03	1.6E-03 ± 2.2E-03	1.8E-03 ± 2.5E-03

**Tabla 4.** Parámetros frecuenciales, en términos de media ± desviación estándar, para los 11 intervalos de señal. Área y desviación estándar, en segundos<sup>2</sup>, de las bandas de baja (LF) y alta (HF) frecuencia

\* SRI(1): intervalo de inicio de respiración, CDB(2,5,8): respiraciones profundas controladas, LEA(3,6,9): apnea espiratoria larga, SIA(4,7,10): apnea inspiratoria corta y FRI(11): intervalo final de relajación

De acuerdo con estos resultados, se observa que los valores de los intervalos SRI y FRI son similares en ambas bandas frecuenciales; el valor medio del área en HF es ligeramente superior que en LF. En los tres intervalos CDB, el área presenta valores mayores en HF que en LF. Por el contrario, los intervalos LEA y SIA presentaron valores de área significativamente más elevados en LF que en HF. Los mayores valores de la desviación estándar se encuentran en los intervalos SRI y FRI.

La Figura 3 presenta un ejemplo del comportamiento del RR en los diferentes intervalos de tiempo, diferenciado por sexo a) mujeres, b) hombres. Estos registros presentan una mayor oscilación en los intervalos de respiraciones profundas controladas, las cuales podrían asociarse a una interacción cardiorrespiratoria de los sistemas.

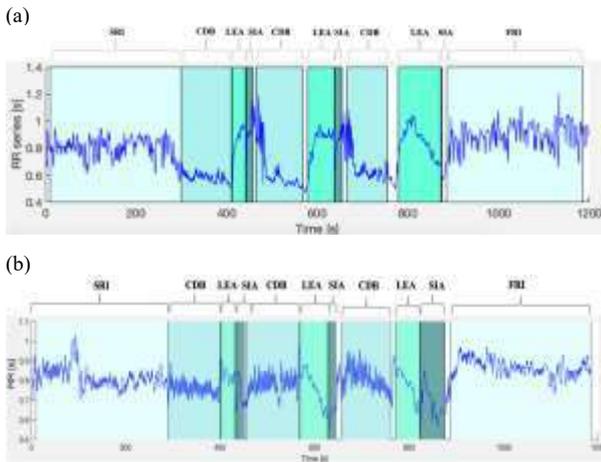


Figura 3. Señal RR de un registro de a) mujer, b) hombre

Analizada la frecuencia cardíaca en función del sexo, se observa, que en promedio, las mujeres presentan un valor más elevado en los intervalos de respiración profunda controlada que los hombres (Figura 4).

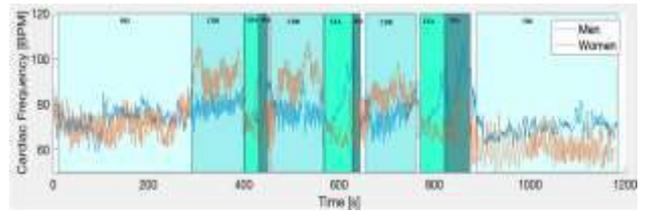


Figura 4. Ejemplo de comparación de la frecuencia cardíaca en un registro de mujer (naranja) y hombre (azul). SRI: intervalo de inicio de respiración, CDB: respiraciones profundas controladas, LEA: apnea espiratoria larga, SIA: apnea inspiratoria corta y FRI: intervalo final de relajación

La figura 5 presenta el comportamiento de los diferentes parámetros temporales para cada intervalo de señal. Se han encontrado diferencias estadísticamente significativas al comparar la frecuencia cardíaca media entre SRI y FRI ( $p < 0.001$ ), SDNN entre SRI y FRI ( $p = 0.038$ ), SDNN entre intervalos LEA ( $p < 0.01$ ), RMSSD entre intervalos CDB ( $p < 0.01$ ), y pNN50 entre CDB1 y CDB2 ( $p = 0.009$ ). También se observan diferencias significativas al comparar los intervalos de respiración y de apneas ( $p < 0.05$ ).

Al analizar estos parámetros por sexo, se observan diferencias estadísticamente significativas en la frecuencia cardíaca en los últimos intervalos de apnea LEA y SIA ( $p < 0.01$ ) (Figura 6.a), en RMSSD en el último intervalo CDB ( $p < 0.05$ ) (Figura 6.b), y en pNN50 en el intervalo medio de LEA ( $p < 0.05$ ) (Figura 6.c).

En el dominio frecuencial, se han obtenido diferencias estadísticamente significativas con los siguientes parámetros:  $fk$ -LF en los intervalos LEA1 y LEA2 ( $p < 0.044$ ),  $fk$ -HF entre los intervalos LEA ( $p < 0.01$ ),  $Ak$ -LF entre los intervalos LEA ( $p < 0.05$ ),  $Ak$ -HF entre los intervalos SRI y FRI ( $p = 0.016$ ), y  $Ak$ -HF entre los intervalos CDB ( $p = 0.009$ ) (Figura 7).

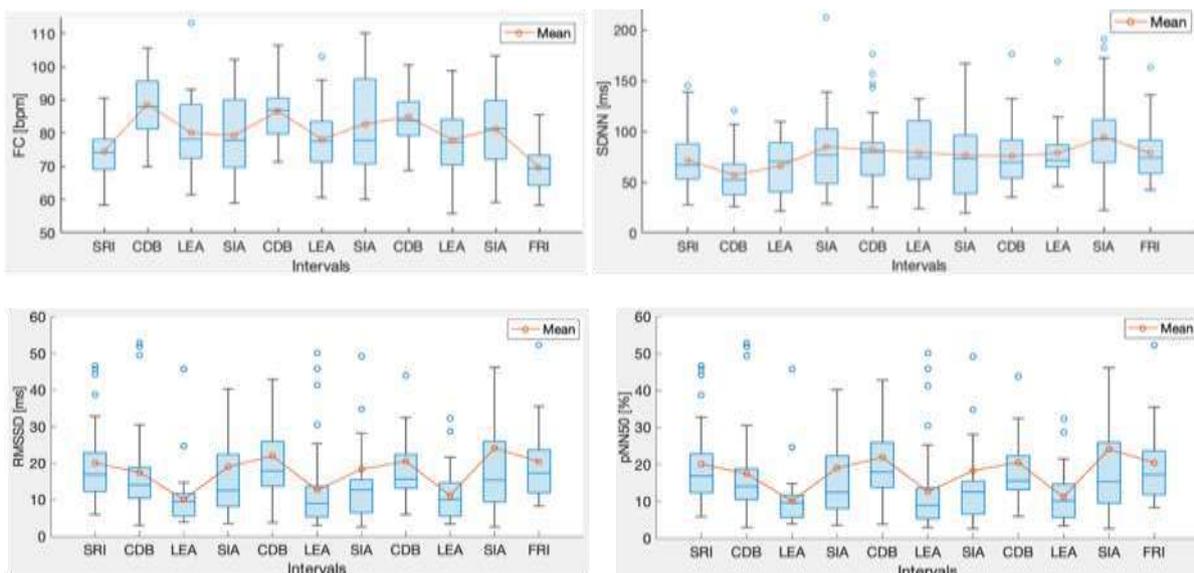
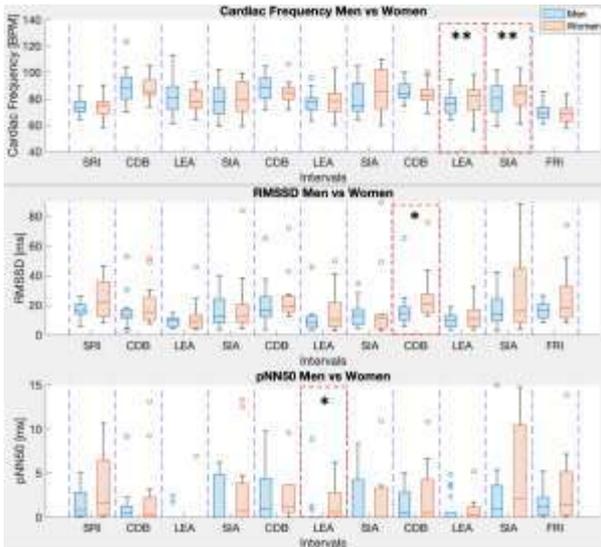
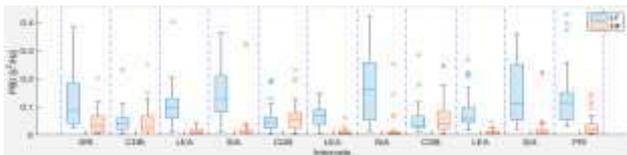


Figura 5. Comportamiento promedio de los parámetros temporales obtenidos a partir de las series temporales RR para cada intervalo de señal: frecuencia cardíaca (FC), SDNN, RMSSD, pNN50. SRI: intervalo de inicio de respiración, CDB: respiraciones profundas controladas, LEA: apnea espiratoria larga, SIA: apnea inspiratoria corta y FRI: intervalo final de relajación

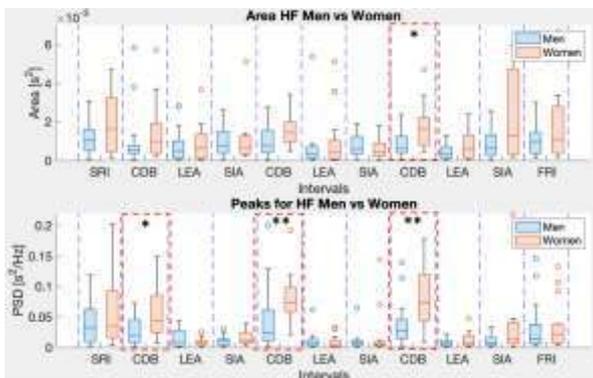


**Figura 6.** Parámetros temporales que presentaron diferencias estadísticamente significativas entre sexos: a) frecuencia cardíaca – intervalos LEA y SIA ( $p < 0.01$ ); b) RMSSD en el último intervalo CDB ( $p < 0.05$ ); c) pNN50 en el intervalo medio de LEA ( $p < 0.05$ )



**Figura 7.** Amplitud de los picos de frecuencia considerando las bandas frecuenciales baja y alta (LF y HF) para cada intervalo de señal, de todos los sujetos

Al comparar la respuesta de los parámetros espectrales por sexo, se observan diferencias estadísticamente significativas en: áreas en la banda frecuencial HF del último intervalo CDB2 ( $p < 0.05$ ) (Figura 8.a),  $F_k$ -HF en CDB1 ( $p < 0.05$ ), y en CDB2 y CDB3 ( $p < 0.01$ ) (Figura 8.b).



**Figura 8.** Parámetros espectrales con diferencias estadísticamente significativas entre sexos: A-HF último intervalo CDB ( $p < 0.05$ )

Al comparar los grupos por fumadores, se observan diferencias estadísticamente significativas en SIA-LF del intervalo 10 ( $p < 0.05$ ), así como una mayor dispersión en los valores de  $A_k$  para los fumadores en todos los

intervalos. Al analizar los sujetos según la actividad deportiva, no se observan diferencias estadísticamente significativas entre ellos, pero si una tendencia de valores más elevados en  $f_k$  – HF, en los intervalos SRI y FRI.

## 5. Conclusión

De acuerdo con el análisis global entre los sujetos se observan valores elevados en la frecuencia cardíaca durante las respiraciones profundas controladas, y bajos en las etapas de relajación, presentando una elevada dispersión entre los intervalos. La variabilidad del ritmo cardíaco presenta diferencias significativas al comparar intervalos de respiración y de apnea. Los grupos con mayor variabilidad del ritmo cardíaco son las mujeres y los fumadores. Al comparar los grupos por el deporte, no se observan diferencias significativas.

Estos resultados indican una relación directa entre la respuesta del sistema cardíaco a respiraciones profundas y a apneas según el método de Wim Hof. Sin embargo, se necesitan más estudios para consolidar estas conclusiones.

## Agradecimientos

Este trabajo ha sido financiado parcialmente por el programa CERCA de la Generalitat de Catalunya, por la Secretaría de Universidades e Investigación del Departamento de Economía y Conocimiento de la Generalitat de Catalunya (GRC 2017 SGR 1770), y por el Ministerio de Economía y Competitividad de España (PID2021-126455OB-I00 (MCIN/AEI/FEDER, UE)).

Los autores agradecen especialmente a los Ingenieros Biomédicos Sergi Coma Roca y Natalia Caballero Diaz por su participación en la creación de este proyecto, así como a todos los estudiantes que participaron como voluntarios para el registros y aplicación del protocolo.

## Referencias

- [1] Mejía-Mejía E, Budidha K, Abay TY, May JM and Kyriacou PA (2020) Heart Rate Variability (HRV) and Pulse Rate Variability (PRV) for the Assessment of Autonomic Responses. *Front. Physiol.* 11:779. doi: 10.3389/fphys.2020.00779.
- [2] Shaffer, F., and Ginsberg, J. (2017). An overview of heart rate variability metrics and norms. *Front. Public Heal.* 5:258. doi: 10.3389/fpubh.2017.00258.
- [3] Gasior JS, Sacha J, Jelén PJ, Zielinski J and Przybylski J (2016). Heart Rate and Respiratory Rate Influence on Heart Rate Variability Repeatability: Effects of the Correction for the Prevailing Heart Rate. *Front. Physiol.* 7:356. doi: 10.3389/fphys.2016.00356.
- [4] What is the Wim Hof Method? | Practice The Method. *Welcome to the Official Wim Hof Method Website*[online]. [accessed February 2022]. Available in: <http://www.wimhofmethod.com/practice-the-method>
- [5] Task-Force. Heart rate variability: standards of measurement, physiological interpretation and clinical use. Task force of the European society of cardiology and the north American society of pacing and electrophysiology. *Circulation* 1996; 93:1043–1065.

# Simulación y Planificación Quirúrgica

---

Viernes, 25 de noviembre de 2022

# Estudio *in silico* de la Inducibilidad de Arritmias Reentrantes en Modelos Personalizados de Ventrículos Infartados

Javier Villar<sup>1</sup>, Juan F. Gomez<sup>2</sup>, David Soto-Iglesias<sup>3</sup>, Diego Penela<sup>3</sup>, Antonio Berruezo<sup>3</sup>, Beatriz Trenor<sup>2</sup>

<sup>1</sup> Centro de Investigación e Innovación en Bioingeniería, Universitat Politècnica de València, Valencia, Spain

<sup>2</sup> Valencian International University, Valencia, Spain

<sup>3</sup> Arrhythmia Department, Heart Institute, Teknon Medical Center, Barcelona, Spain

\* Correspondencia: btrenor@eln.upv.es

## Resumen

*En las etapas posteriores a un infarto de miocardio, la evolución en la morfología de la cicatriz del infarto y la zona de borde juega un papel fundamental en la aparición de arritmias potencialmente mortales. Los modelos computacionales 3D del corazón pueden ser de gran ayuda para comprender en profundidad los mecanismos que subyacen a las arritmias reentrantes en unas condiciones tan particulares.*

*En este trabajo, se han construido modelos personalizados, a partir de imágenes de resonancia magnética con realce tardío de gadolinio, de corazones de pacientes que han sufrido un infarto de miocardio. Se ha simulado la actividad eléctrica de estos corazones, y se han diseñado distintas configuraciones para reproducir de forma realista el remodelado que tiene lugar en la zona de borde de la cicatriz tras el infarto. Tras replicar en los corazones virtuales el mismo protocolo de estimulación que emplean los cardiólogos en su práctica clínica en pacientes con y sin taquicardia ventricular (TV), se consigue simular con éxito el mismo comportamiento arritmogénico de los corazones en condiciones de remodelado estructural de la zona de borde.*

*Las simulaciones llevadas a cabo reproducen patrones de reentrada realistas y consistentes con los resultados obtenidos en la clínica, lo que aporta información útil sobre los mecanismos de aparición de las TVs.*

En los últimos años, la resonancia magnética cardiaca con realce tardío de gadolinio se ha convertido en una punta de lanza en lo referente a la caracterización de las lesiones isquémicas *in vivo*; supone también una gran ayuda para la planificación de posibles intervenciones tras el infarto como la ablación por radiofrecuencia, cuyo objetivo es eliminar, precisamente, los circuitos reentrantes por los que puede haber una propagación anómala, reduciendo así el riesgo de arritmias, fibrilación y, en última instancia, la muerte del paciente [3]. A lo largo de la última década, han ido surgiendo nuevos métodos de diagnóstico y tratamiento para este tipo de patologías; metodologías basadas en simulaciones computacionales *in vitro* de modelos personalizados de los pacientes, que aportan un enfoque distinto al que aborda la práctica clínica, y que ayudan a comprender los mecanismos que favorecen las arritmias, así como sus causas subyacentes.

Los cambios electrofisiológicos en las células de un ventrículo infartado, en la zona de borde particularmente, se pueden resumir como la prolongación de la duración del potencial de acción, y alteraciones en las dinámicas de calcio [4]. Estas alteraciones en la actividad eléctrica de los cardiomiocitos se deben al remodelado funcional de los canales iónicos de las células, de las bombas electrogénicas y de los cotransportadores iónicos. En cuanto al remodelado estructural de la zona de borde, uno de los factores más importantes es la aparición de fibrosis, que aumenta considerablemente el riesgo de padecer arritmias. La presencia de fibroblastos intersticiales afecta a la conducción del impulso eléctrico tanto o más que el comportamiento eléctrico de los miocitos tras el remodelado iónico [5, 6].

En este trabajo se desarrollan modelos computacionales de ventrículos anatómicamente realistas y personalizados para pacientes que han sufrido un infarto de miocardio, y se trata de inducir actividad reentrante en ellos para analizar las características de la cicatriz y de la zona de borde, que son responsables de los mecanismos de la reentrada. Toda la metodología descrita en este artículo ofrece una herramienta robusta para la ayuda al diagnóstico y a las técnicas de terapia de ablación y de tratamientos preventivos contra futuras arritmias.

## 1. Introducción

De acuerdo con la Organización Mundial de la Salud (OMS), las enfermedades cardiovasculares son la principal causa de muerte en todo el mundo, matando a 17.9 millones de personas en 2019, lo que supone el 32% de todas las muertes registradas en el mundo. De estas muertes, el 85% se deben a infartos de miocardio [1]. La OMS señala que la detección temprana y el tratamiento de los pacientes con enfermedades cardiovasculares es esencial dada su alta mortalidad.

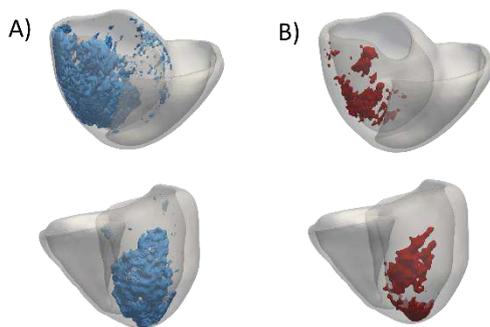
Además, de entre los pacientes que sufren un infarto de miocardio, el 10% tienen un alto riesgo de muerte en los próximos años, siendo el 50% de estas muertes causadas por arritmias ventriculares [2]. Esto se debe al remodelado tisular en el tejido circundante a la cicatriz del infarto tras el episodio isquémico una vez que el proceso de curación ha alcanzado el estadio crónico [3]. El tejido remodelado se conoce como zona de borde del infarto, y conforma un sustrato que favorece la aparición de circuitos de conducción reentrantes y propagaciones anómalas.

## 2. Materiales y métodos

### 2.1. Datos de los pacientes

La construcción de modelos personalizados requiere de un conjunto de imágenes de alta resolución, y de información completa sobre la anatomía y disposición de las regiones infartadas del corazón. El centro médico Teknon (Barcelona) ha provisto un conjunto de imágenes de dos adquisiciones distintas de resonancia magnética con realce tardío de gadolinio de corazones infartados; dichas adquisiciones corresponden a sendos pacientes en etapas posteriores al infarto, uno de los cuales ha presentado taquicardia ventricular en la clínica, mientras que el otro no ha referido ningún episodio arrítmico. Los datos de los pacientes han sido anonimizados previamente por el centro médico.

En las adquisiciones de los ventrículos tras el infarto se consideran las regiones de la cicatriz del infarto y de la zona de borde aplicando unos umbrales a la intensidad de píxel de las imágenes de resonancia magnética. Se clasifica la región del ventrículo como núcleo de la cicatriz si la intensidad de píxel es superior al 60% de la intensidad máxima de píxel, como zona de borde si se encuentra entre el 40% y el 60% de la intensidad máxima de píxel y como tejido sano si se encuentra por debajo del 40% [7].



**Figura 1.** Modelo cardíaco anatómico personalizado de dos pacientes distintos. Fila superior: paciente que presenta taquicardia ventricular clínica (paciente 1). Fila inferior: paciente que no presenta taquicardia ventricular clínica (paciente 2). (A) Etiquetado de la zona de borde de la cicatriz del infarto en azul (B) Etiquetado del núcleo de la cicatriz del infarto en rojo.

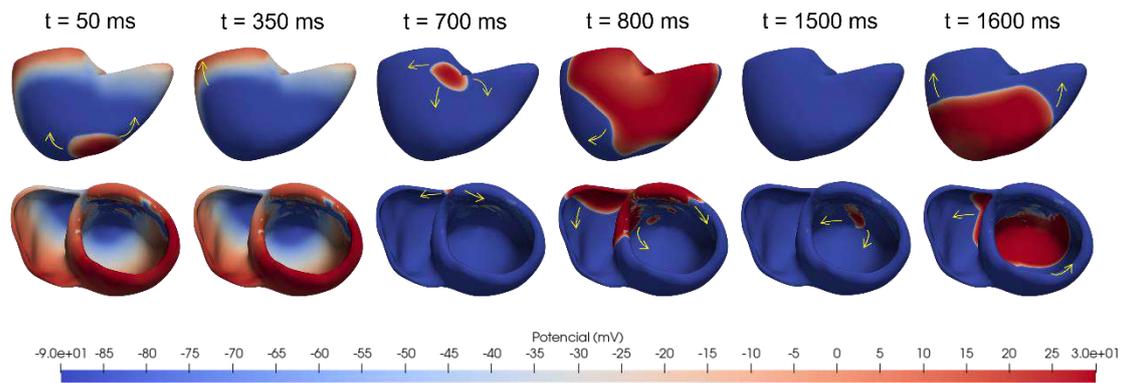
### 2.2. Reconstrucción del modelo específico del paciente

Para la reconstrucción de los modelos personalizados a partir de las imágenes adquiridas de resonancia magnética se ha realizado una segmentación de las estructuras de interés (anatomía del ventrículo, zona de borde y núcleo de la escara, mostrados en la Figura 1) y, una vez están segmentadas las distintas regiones de los ventrículos, se genera un volumen tridimensional con el software Seg3D. Además, se aplica un suavizado exhaustivo al volumen generado para que el resultado sea anatómicamente realista, y se genera una malla tridimensional con el programa MeshGems de elementos hexaédricos con una resolución espacial de 0.4 mm en la que reproducir simulaciones computacionales. Una vez el volumen está correctamente mallado se etiquetan las distintas estructuras

del volumen: Regiones del ventrículo derecho, ventrículo izquierdo y septo ventricular, regiones de zona de borde y núcleo de la cicatriz. Las regiones de endocardio, midmiocardio y epicardio se etiquetan en función del grosor de la pared cardíaca, siendo el 17% de la pared más interna endocardio, el siguiente 41% midmiocardio y, por último, el 42% restante hasta la superficie externa se etiqueta como epicardio [3]. Una vez los volúmenes están mallados y etiquetados se generan archivos con las características de las mallas, que sirven de entrada para el software ELVIRA, un código de elementos finitos que utiliza el modelo monodominio anisotrópico para modelar el tejido cardíaco que describe la propagación del potencial de membrana [8]. El programa permite realizar simulaciones del comportamiento eléctrico de los volúmenes ventriculares reconstruidos; a partir de los nodos y de los elementos hexaédricos de la malla y las características definidas para cada zona, el programa representa las distintas células como ecuaciones diferenciales en cada elemento y las resuelve para devolver los resultados de una simulación a lo largo de un tiempo definido. Así pues, los volúmenes se resolvieron en ELVIRA utilizando una versión modificada del modelo de miocitos ventriculares humano de O'Hara et al (2011) [9, 10] en los distintos nodos de la malla. El núcleo de la cicatriz del infarto se consideró tejido no conductivo, y la zona de borde se remodeló reduciendo la conductividad tanto en la dirección longitudinal como en la transversal de la fibra en un 90% respecto de la conductividad del tejido sano, esto implica una reducción en la velocidad de conducción que depende de la dirección de la fibra y de la región del ventrículo, pero que oscila entre una reducción 55% y un 70% de la velocidad de conducción respecto de los miocitos sanos.

Además, para simular el comportamiento electrofisiológico de los fibroblastos presentes en la zona de borde en etapas posteriores al infarto de miocardio, se ha utilizado el modelo celular de fibroblasto de MacCannell et al. (2007) [6] combinado con el modelo de O'Hara et al. (2011) modificado de miocito ventricular [10]. Se introducen distintos porcentajes de fibroblastos en la zona de borde según una función probabilística para asignar de forma aleatoria la disposición de los fibroblastos en la zona de borde y los nodos de la malla a los que se asigna el modelo de fibroblasto de MacCannell et al. (2007) en lugar del de miocito. El acoplamiento entre miocito y fibroblasto se simula reduciendo el coeficiente de difusión un 50% [5], lo que supone mayor reducción de la velocidad de conducción en las regiones con mayor concentración de fibroblastos y un mayor fraccionamiento del frente de onda de propagación eléctrico, lo que puede propiciar la aparición de arritmias reentrantes.

Por último, y basándonos en estudios previos, se ha diseñado un remodelado iónico para los miocitos ventriculares de la zona de borde. Se han adaptado los cambios en las conductancias de los canales iónicos que se han validado en modelos celulares distintos para obtener



**Figura 2.** Evolución temporal del potencial de acción cardíaco del paciente 1 con 40% de fibrosis en zona de borde (\* ver Tabla 1). Los instantes temporales corresponden al tiempo transcurrido desde la aplicación del último estímulo. Fila superior: vista coronal del volumen biventricular. Fila inferior: vista axial del volumen biventricular.

las mismas alteraciones en el potencial de acción documentados en estudios experimentales, pero en el modelo de miocito de O'Hara et al (2011) modificado [10]. Estos cambios se pueden resumir de la siguiente manera: Reducción de la conductancia de la corriente  $I_{Na}$  a un 38% de su valor normal,  $I_{CaL}$  a un 23%, y las corrientes  $I_{Kr}$  e  $I_{Ks}$  a un 80% y un 70% respectivamente [11]. Estos cambios conducen a un aumento de la duración del potencial de acción en un 17%, a una disminución de la velocidad de despolarización del 44'9% y una disminución del valor de potencial máximo de 42'35%, lo que es consistente con los resultados experimentales en cardiomiocitos de la zona de borde consultados en la bibliografía [11].

### 2.3. Protocolo de estimulación

Una de las condiciones necesarias para el desarrollo y mantenimiento de una arritmia es la aparición de un estímulo prematuro. Los modelos anatómicamente realistas con zona de borde y núcleo de la cicatriz son el sustrato perfecto para propiciar la aparición de una reentrada si el estímulo prematuro se aplica durante el periodo de tiempo que comprende la ventana vulnerable.

El protocolo que se describe a continuación ha sido utilizado por los cardiólogos en el centro médico para tratar de inducir taquicardia ventricular *in vivo* para la planificación de procedimientos de ablación, y ha sido reproducido en este estudio *in silico* aplicando una serie de estímulos, siempre en el ápex del ventrículo derecho. En primer lugar, se aplica un tren de 6 estímulos basales (S1-S6) espaciados 430 ms para asegurar la estabilización de las corrientes iónicas del modelo. Después de esto, se aplican hasta 3 extraestímulos prematuros (E1-E3); cada nuevo estímulo se aplica 300 ms después del último estímulo (basal o extra) y el intervalo temporal se va reduciendo en pasos de 10 ms hasta que, bien se alcanza el periodo refractario de la zona estimulada para la que el estímulo no se propaga, bien se induce una arritmia reentrante mantenida. Si el estímulo prematuro se reduce hasta el periodo refractario y no hay reentrada, se fija este estímulo 20 ms después del periodo refractario y se aplica un nuevo extraestímulo. Este protocolo se repite hasta añadir 3 extraestímulos, si después de esto no se induce una arritmia, se considera que el modelo no tiene actividad reentrante.

## 3. Resultados

Paciente 1	Sin remodelado iónico	Con remodelado iónico
0% fibrosis	No TV	No TV
30% fibrosis	No TV	No TV
40% fibrosis	No TV	TV*
50% fibrosis	TV	TV
Paciente 2	Sin remodelado iónico	Con remodelado iónico
0% fibrosis	No TV	No TV
30% fibrosis	No TV	No TV
40% fibrosis	No TV	No TV
50% fibrosis	No TV	No TV

**Tabla 1.** Resultado del protocolo de simulación para distintas configuraciones de la zona de borde. TV indica que se ha observado taquicardia ventricular sostenida; No TV indica que el protocolo de estimulación ha finalizado sin arritmias. La configuración destacada (\*) es la empleada en la Figura 2.

El protocolo de estimulación descrito anteriormente se ha repetido en los volúmenes reconstruidos de ambos pacientes para distintos modelos de la zona de borde y configuraciones de fibrosis. Para cada paciente se repite el protocolo de inducción de arritmias para cuatro configuraciones distintas en función del volumen de fibroblastos en la zona de borde: sin fibrosis, con un 30% de fibrosis en la zona de borde, con un 40% de fibrosis y con un 50% de fibrosis; además, todas ellas se repitieron con la adición del remodelado iónico desarrollado y sin él, resultando en un total de 8 protocolos por paciente.

Como se muestra en la tabla 1, en el caso del paciente 1, que presenta taquicardia ventricular en la clínica, es posible inducir la taquicardia mediante el protocolo de estimulación en la simulación computacional solo cuando la fibrosis de la zona de borde supera un valor determinado (a partir del 40% de fibroblastos) en los casos en que se

aplica remodelado; y, en las simulaciones sin el remodelado iónico propuesto, el desarrollo de la reentrada requiere de un mayor volumen de fibroblastos.

Por otra parte, en el paciente que no presentaba taquicardia en la clínica, tampoco se induce taquicardia en ninguna de las simulaciones realizadas, independientemente de la cantidad de fibrosis o remodelado iónico.

La figura 2 muestra la evolución temporal de la propagación del potencial de acción después del último estímulo prematuro del protocolo de estimulación de la simulación destacada en la Tabla 1 (\*). Se puede observar cómo el impulso se propaga normalmente y se extingue, pero algunas pequeñas regiones de zona de borde quedan activas. El retardo en la activación de zona de borde provoca que se genere un nuevo impulso anómalo en el epicardio del ventrículo izquierdo; este impulso se propaga por todo el volumen ventricular y, antes de extinguirse por completo, la propagación vuelve a reentrar por un sitio distinto de salida al primer ciclo de la arritmia, el endocardio del ventrículo izquierdo, lo que resulta en una taquicardia ventricular persistente.

#### 4. Discusión y conclusiones

Las características de la cicatriz cardiaca después de un infarto constituyen un factor fundamental a estudiar, hasta tal punto que pueden ser determinantes en la ocurrencia, o no, de arritmias potencialmente mortales. En este estudio, se han desarrollado con éxito distintos modelos ventriculares para dos pacientes que habían sufrido infarto de miocardio, con el fin de analizar las características de la cicatriz que conducen a una taquicardia ventricular.

La novedad de esta investigación radica en el uso de herramientas personalizadas con un alto grado de detalle desde la construcción del modelo hasta la caracterización del remodelado que experimenta la zona de borde del infarto. Estos modelos son capaces de predecir con precisión la ocurrencia de arritmias y, de acuerdo con los resultados extraídos de las simulaciones, para el paciente con taquicardia clínica, tanto el remodelado iónico como la densidad de fibroblastos en la zona de borde juegan un papel fundamental en el desarrollo de la taquicardia ventricular; mientras que, en el paciente que no muestra TV clínica, ni el remodelado ni la fibrosis tienen ningún efecto en la ocurrencia de TV por no disponer de las estructuras anatómicas que favorecen la arritmia reentrante.

Por otra parte, son necesarias más investigaciones aumentando el número de casos de estudio para poder caracterizar y entender en mayor profundidad el papel del remodelado estructural y electrofisiológico en la generación de reentradas después de un infarto.

#### Agradecimientos

Este trabajo fue apoyado por Grant PRE2020-091849 financiado por MCIN/AEI/10.13039/501100011033 y “ESF Investing in your future”, y por Grant PID2019-104356RB-C41 financiado por MCIN/AEI/10.13039/501100011033. Los autores también agradecen los proyectos de Teknon y los recursos informáticos de

MareNostrum así como el soporte técnico proporcionado por Barcelona Supercomputing Center (IM-2021-1-0001 e IM-2021-3-0001).

#### Referencias

- [1] Who.int. 2022. Cardiovascular diseases (CVDs). [online] Available at: <[https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-\(cvds\)](https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds))> (Consultado marzo 2022).
- [2] Bhar-Amato, J., Davies, W. and Agarwal, S., Ventricular Arrhythmia after Acute Myocardial Infarction: ‘The Perfect Storm’. *Arrhythmia & Electrophysiology Review*, 6(3), 2017, pp.134-139. (ISSN: 1944-7981).
- [3] Lopez-Perez, A., Sebastian, R., Izquierdo, M., Ruiz, R., Bishop, M. and Ferrero, J., Personalized Cardiac Computational Models: From Clinical Data to Simulation of Infarct-Related Ventricular Tachycardia. *Frontiers in Physiology*, vol 10, 2019 (ISSN=1664-042X).
- [4] Decker, K. and Rudy, Y. Ionic mechanisms of electrophysiological heterogeneity and conduction block in the infarct border zone. *American Journal of Physiology-Heart and Circulatory Physiology*, 299(5), 2010, pp H1588-H1597 (ISSN: 0002-9513).
- [5] Gomez, J., Cardona, K., Romero, L., Ferrero, J. and Trenor, B. Electrophysiological and Structural Remodeling in Heart Failure Modulate Arrhythmogenesis. *1D Simulation Study*. *PLoS ONE*, 9(9), 2014, p.e106602 (ISSN: 1932-6203).
- [6] MacCannell, K., Bazzazi, H., Chilton, L., Shibukawa, Y., Clark, R. and Giles, W. A Mathematical Model of Electrotonic Interactions between Ventricular Myocytes and Fibroblasts. *Biophysical Journal*, 92(11), 2007, pp.4121-4132 (ISSN: 0006-3495).
- [7] Soto-Iglesias, D., Penela, D., Jáuregui, B., Acosta, J., Fernández-Armenta, J., Linhart, M., Zucchelli, G., Syrovnev, V., Zaraket, F., Terés, C., Perea, R., Prat-González, S., Doltra, A., Ortiz-Pérez, J., Bosch, X., Camara, O. and Berruezo, A. Cardiac Magnetic Resonance-Guided Ventricular Tachycardia Substrate Ablation. *JACC: Clinical Electrophysiology*, 6(4), 2020, pp.436-447 (ISSN: 2405-500X).
- [8] Heidenreich E, Eisler H, Lengheimer T, Dorninger P, Steinboeck F. A mutation-promotive role of nucleotide excision repair in cell cycle-arrested cell populations following UV irradiation. *DNA Repair (Amst)*. 9(1), 2010, pp 96-100 (ISSN: 1568-7856)
- [9] O'Hara, T., Virág, L., Varró, A. and Rudy, Y. Simulation of the Undiseased Human Cardiac Ventricular Action Potential: Model Formulation and Experimental Validation. *PLoS Computational Biology*, 7(5), 2011, p.e1002061 (ISSN: 1553-734X).
- [10] Carpio EF, Gomez JF, Sebastian R, et al. Optimization of Lead Placement in the Right Ventricle During Cardiac Resynchronization Therapy. A Simulation Study. *Front Physiol*. 2019, 10:74. (ISSN: 1873-1732).
- [11] Mendonca Costa C, Plank G, Rinaldi CA, Niederer SA, Bishop MJ. Modeling the Electrophysiological Properties of the Infarct Border Zone. *Front Physiol*. 9, 2018, 9:356 (ISSN: 1664042X).

# DetECCIÓN DEL GRADO DE LIMPIEZA DE LA MUCOSA BASADO EN INTELIGENCIA ARTIFICIAL COMO SISTEMA DE APOYO PARA LA EVALUACIÓN DE CALIDAD EN COLONOSCOPIAS.

P. García-García<sup>1</sup>, C. Guzmán-García<sup>1</sup>, Álvaro Martínez- Alcalá García<sup>2</sup>, I. Oropesa<sup>1</sup>, E. J. Gómez<sup>1,3</sup>, P. Sánchez-González<sup>1,3</sup>

<sup>1</sup> Grupo de Bioingeniería y Telemedicina, ETSI Telecomunicación, Centro de Tecnología Biomédica, Universidad Politécnica de Madrid, Madrid, Spain; paloma.garcia@alumnos.upm.es, carmen.guzmang, i.oropesa, enriquejavier.gomez, p.sanchez{@upm.es}

<sup>2</sup> Servicio de Aparato Digestivo, Hospital Infanta Leonor, Madrid, Spain. alvaromalcala@gmail.com

<sup>3</sup> Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina, Madrid, Spain

## Resumen

*La detección automática del grado de limpieza durante la colonoscopia desempeña un papel importante en la evaluación de calidad del procedimiento. Esta aporta información relevante sobre la tasa de preparación intestinal del procedimiento, siendo esta uno de los principales indicadores de calidad para la colonoscopia. Los sistemas actuales de detección con inteligencia artificial en este ámbito están poco extendidos y desarrollados, siendo el principal objetivo de estos la detección de pólipos. Como método de evaluación del indicador se emplea a lo largo de este estudio la escala validada de Boston (BBPS). Para el desarrollo e implementación del sistema propuesto se emplea un conjunto de videos anotados y clasificados en cuatro grupos según las cuatro puntuaciones de la escala BBPS. La evaluación del sistema resultó en un valor de F1 de 0.939. Este sistema tiene potencial para seguir siendo desarrollado e implementado en aplicaciones para evaluación de calidad de colonoscopias.*

## 1. Introducción

El cáncer colorrectal (CCR) es uno de los cánceres más frecuentemente diagnosticados en 2022 con 43.370 casos nuevos respecto del año anterior, de los cuales aproximadamente un 60% son varones [1]. Se conoce que, en 2020 hubo una prevalencia en España de 191.884 casos en total, siendo el cáncer más común y se estima que al año fallecen más de 14.000 solo en nuestro país [1].

La principal estrategia para reducir esta incidencia es la detección precoz mediante programas de cribado en la población, ya que el 70% de los casos de CCR son esporádicos y en personas sin factores de riesgo conocidos. La técnica más habitualmente empleada para el diagnóstico es la colonoscopia, en la que se utiliza un instrumento llamado colonoscopio (formado por una sonda flexible y una cámara fijada en su extremo) para visualizar la parte interna del colon, el recto y el final del intestino delgado (íleon terminal) [2]. De esta forma se pueden detectar y tratar lesiones y pólipos colorrectales, precursores de este cáncer. Este procedimiento se realiza de manera ambulatoria en una sala de intervenciones del centro médico u hospital [2].

Durante los últimos 20 años, se ha extendido el interés entre las distintas sociedades científicas en evaluar y cuantificar la calidad del procedimiento, aplicando una serie de criterios y parámetros que se resumen en forma de indicadores. Aplicado a la endoscopia, una aproximación del concepto de calidad sería el “conjunto de propiedades que hace que la endoscopia cumpla los requisitos para los que se realiza y se aprecie como igual, mejor o peor que el resto de los servicios de su misma categoría” [3]. Es en este punto donde se introduce la calidad como un término relativo, con una condición tácita de comparación. Con ello surge la necesidad de poder valorar y contrastar los procesos con una norma o estándar [4], que resumen datos e información sistemática obtenida al aplicar una serie de criterios y parámetros, resumidos en forma de indicadores. Estos están definidos en las guías clínicas de la Sociedad Europea de Endoscopia Gastrointestinal (ESGE) [5], [6] y son una medición que describe el proceso, evaluando el comportamiento en el tiempo y desempeño, constituyendo una fuente de información para la toma de decisiones [3].

La adquisición de estos criterios de calidad no siempre está exenta de barreras o dificultades, entre las que destaca la resistencia al cambio o el tiempo necesario para medir, cuantificar y auditar la adhesión a los mismos ya que en muchos casos son criterios dependientes del observador o con cierto grado de subjetividad y que requieren análisis manual o poco automatizado. En los últimos años las diferentes sociedades científicas han publicado artículos de posicionamiento [7] y promocionado el uso de aplicaciones específicas [8].

El primer paso para conseguir un examen clínico exitoso, es que los intestinos estén limpios y vacíos [2]. Esto es lo que se conoce como preparación del intestino, limpieza colónica o mucosa [2]. Se busca una mejor y más eficiente visualización de la luz del intestino. La limpieza de la mucosa intestinal va a influir tanto en la peligrosidad como en la adecuación del examen (es decir, una limpieza deficiente conlleva un incremento de lesiones no detectadas, una incompleta intubación cecal y una alteración del seguimiento, lo que suele implicar repeticiones de pruebas en un intervalo más corto por falta de resultados [9]). El indicador de calidad en el que se

centra este estudio es la tasa de preparación intestinal [3], [9], [10]. Para evaluar la preparación se han descrito varias escalas validadas, recomendadas por la ESGE, como la Boston Bowel Preparation Score (BBPS), Ottawa (OBPS), Aronchick (ABPS) o Harefield [11]. Queda establecido que la limpieza, para ser considerada buena o excelente, debe llegar al 90% como mínimo [3], [10].

La escala empleada durante el estudio es la BBPS [12], debido a su fácil aplicación tras las maniobras de limpieza y por ser la escala utilizada en el etiquetado de datos empleados en el trabajo. Con ella se evalúa a través de una puntuación de 0 a 3 para cada segmento intestinal evaluado por separado (colon izquierdo, transversal y descendente), siendo el 0 una preparación muy mala y el 3 excelente. El resultado final debe ser entre 2 y 3 para cada segmento para ser considerado adecuado y permitirle al endoscopista identificar pólipos de más de 5 mm de tamaño. Por el contrario, los pacientes con puntuaciones menores tendrán una mayor tasa de adenomas no detectados y, por tanto, les será fechada otra prueba o evaluación antes de los 6 meses [13].

Zhou et al. [14] emplean ENDOANGEL, un sistema basado en inteligencia artificial (IA), para la evaluación de la preparación de la mucosa, con el objetivo de mostrar cada 30 segundos la puntuación BBPS, de forma que pueda establecerse un valor para cada segmento, obteniendo una precisión del 93,3% para la aplicación en imágenes y un 89,4% en videos. En el artículo de Hyung Nam et al. [15] los autores desarrollaron un sistema capaz de evaluar la calidad de la preparación de intestino delgado, logrando una precisión del 93%. Young Lee et al. [16] buscan también desarrollar un algoritmo capaz de evaluar objetivamente la preparación intestinal mediante el uso de redes neuronales, empleando la BBPS, consiguiendo distinguir la adecuación de la preparación en un 85% de los videos. Pogorelov et al. [17] crean la base de datos Nerthus, un conjunto de datos público con 21 videos de colonoscopias anotados y clasificados en cuatro grupos según las cuatro puntuaciones de la escala BBPS, obteniendo además en su ensayo de clasificación una puntuación F1 de 0,899.

En este estudio se plantea mejorar la detección automática de la tasa de preparación intestinal mediante la implementación de un modelo basado en redes neuronales convolucionales, capaz de obtener una cuantificación y calificación objetiva.

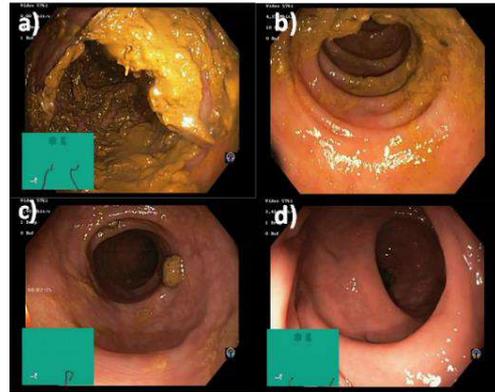
## 2. Materiales y métodos

### 2.1. Materiales

Para el desarrollo e implementación del sistema se emplean dos bases de datos: Nerthus [17] y HyperKvasir [18]. Los videos de ambas bases de datos fueron recopilados en el hospital de Bærum, Vestre Viken Hospital Trust en Noruega y fueron anotados por expertos de diversos orígenes.

La base de datos Nerthus contiene 5.525 fotogramas de 21 videos agrupadas según los cuatro niveles de limpieza de la escala BBPS (Figura 1). Esta base de datos se utiliza como conjunto de entrenamiento de nuestro modelo: el

80% del conjunto para el entrenamiento y el 20% restante como validación del modelo.



**Figura 1.** Imágenes de Nerthus correspondientes a los cuatro niveles de limpieza BBPS [13]. a) Nivel 0; b) Nivel 1; c) Nivel 2; d) Nivel 3.

Por otra parte, la base de datos HyperKvasir contiene, entre otros datos, 1.794 imágenes relacionadas con los niveles de limpieza BBPS. En este caso las imágenes se agrupan únicamente en dos grupos. En el primero de ellos se incluyen las pertenecientes a las puntuaciones BBPS de 0 y 1, mientras que el segundo de ellos contiene las de 2 y 3. Las imágenes de esta base de datos se utilizan como conjunto de prueba.

### 2.2. Entrenamiento y evaluación

Primero se realiza un preprocesamiento, que consiste en un aumento de datos igual para todas las clases volteando los fotogramas horizontalmente y rotándolos aleatoriamente. Se aplica también un cambio de tamaño de imagen para normalizarlo, siendo el alto 160 píxeles y el ancho 160 píxeles, y una normalización de los datos para adaptarlos a la entrada de los distintos modelos seleccionados.

Concretamente, se han escogido varias arquitecturas de redes neuronales para probar su desempeño con los datos: MobileNetV2, ResNet50V2, InceptionV3, VGG-16, Xception y EfficientNetB0. Todas ellas entrenadas previamente con ImageNet, una base de datos con más de 14 millones de imágenes, ampliamente utilizada para el entrenamiento de modelos para reconocimiento de objetos.

Para entrenar de nuevo el modelo con nuestros datos, se comienza congelando todas las capas del cuerpo, a excepción de la última, que será realmente la entrenada en este estudio. De esta forma evitamos que se actualicen los pesos de las capas anteriores, correspondientes al entrenamiento previo de las redes, y la posibilidad de perder información en futuras épocas del entrenamiento. Se agrega una capa como encabezado de clasificación que realiza una agrupación promedio los mapas de características de las distintas ubicaciones espaciales, de forma que queden resumidas en un único vector por imagen. Por último, se añade una capa densa de dimensionalidad 4, correspondiente a los grupos de clasificación. De esta forma, estas últimas capas, no congeladas, aprenden a convertir características de las capas anteriores en predicciones sobre nuestro conjunto de datos nuevo. La función de activación para esta última capa

fue *softmax*, una generalización de la regresión logística ampliamente utilizada en la salida de problemas de clasificación multiclase.

Se comienza entrenando el modelo durante 10 épocas. Como función de pérdida se emplea la entropía cruzada categórica (al ser un problema de clasificación multiclase). Ésta evalúa la diferencia entre el valor predicho por la red neuronal y el real. La idea es minimizar ese valor ajustando los pesos, buscando una mayor precisión del modelo. A lo largo del ensayo se cambian algunos parámetros, como la tasa de aprendizaje y el optimizador, para lograr un resultado más eficiente, tanto en el ámbito temporal como en la consecución de una mayor precisión y sensibilidad. En lo referido a los optimizadores, se prueba la eficiencia de 3 de los más comúnmente extendidos en el entrenamiento de redes para clasificación como la propuesta en este estudio: Adam, AdaDelta y SGD. La función de los optimizadores consiste en generar parámetros cada vez mejores durante el entrenamiento, de forma que disminuya la pérdida y el modelo converja más rápidamente. Su funcionamiento se basa en el cálculo de gradientes de la función de pérdida con derivadas parciales por cada peso de la red. Del mismo modo, se realiza una adaptación de la tasa de aprendizaje, probando 0,1, 0,001 y 0,0001.

Para la evaluación del modelo, se hace uso de las métricas precisión, exhaustividad y F1.

$$Precisión = \frac{TP}{TP + FP}$$

$$Exhaustividad = \frac{TP}{TP + FN}$$

$$F1 = \frac{2 * (Precisión * Exhaustividad)}{(Precisión + Exhaustividad)}$$

Con TP: verdaderos positivos; TN: verdaderos negativos; FP: falsos positivos y FN: falsos negativos.

### 3. Resultados y discusión

La Tabla 1 recoge el rendimiento de las distintas redes empleadas a lo largo del ensayo, en términos de las métricas mencionadas previamente, junto con el coste computacional. Estos resultados se corresponden al entrenamiento realizado durante 10 épocas, con un tamaño de lote de 32, optimizador Adam y tasa de aprendizaje 0,0001. El comportamiento más notable es el perteneciente al modelo EfficientNetB0 con un F1 de 0.939.

Red	F1	Tiempo entrenamiento (min)	Tiempo predicción (s)
Xception	0,903	50	136
EfficientNetB0	<b>0,939</b>	26	87
InceptionV3	0,868	210	65
MobileNetV2	0,917	15	44
ResNet50V2	0,889	40	101
VGG16	0,913	120	286

**Tabla 1.** Comportamiento de los distintos modelos durante la prueba

Se empleó esta misma arquitectura para la comparación del funcionamiento entre optimizadores. El modelo entrenado con Adam mostró un comportamiento superior al entrenado con SGD y AdaDelta con una precisión del 89,47% para el grupo 0-1 y un 95,47% para el 2-3, además de un coste computacional menor (Tabla 2). Esto puede deberse a la fijación de la tasa de aprendizaje incorporada en SGD y AdaDelta, frente a la adaptabilidad proporcionada por Adam. De hecho, Adam demuestra aprender más rápido y ser más estable sin perder precisión en el resultado de la predicción.

Optimizador	Precisión media entre grupos	Pérdida entrenamiento
Adam	92,471	0,251
SGD	56,041	1,359
AdaDelta	64,466	0,930

**Tabla 2.** Evaluación de los optimizadores con red EfficientNetB0.

Para el último ajuste del modelo se mantiene la arquitectura EfficientNetB0 y el calificado como mejor optimizador, Adam, y se varía la tasa de aprendizaje. Los resultados quedan descritos en la Tabla 3, demostrando un mejor funcionamiento el modelo con el valor de tasa de aprendizaje definida desde el primer momento: 0,0001. Esto indica que aumentando la velocidad con que el algoritmo converge no logramos la estabilidad que el modelo requiere. Además, observamos que la predicción no se demora en demasía, por lo que se concluye que mediante esta configuración se logra un buen equilibrio entre coste computacional y capacidad de convergencia.

Learning rate	PRE	Exhaustividad	F1
0.1	0,755	0,787	0,795
0.001	0,755	0,773	0,734
0.0001	<b>0,954</b>	<b>0,930</b>	<b>0,940</b>

**Tabla 3.** Evaluación del desempeño de la red EfficientNetB0 con optimizador Adam y diferentes tasas de aprendizaje.

Como la clasificación de las imágenes del conjunto de prueba estaba dividida en 2 grupos únicamente, fueron aceptadas como verdaderos positivos en el primero grupo, por ejemplo, tanto los valores detectados de clase 1, como los de 0, sin hacer distinción entre ellos; aunque a la salida de la prueba el modelo sí detectara diferencias. Por una parte, esto pierde importancia si entendemos que 0 y 1 son considerados como no adecuados, mientras que 2 y 3 sí lo son. Entonces, la clasificación dentro de un subgrupo no sería tan relevante al realizar la suma de las puntuaciones en los distintos segmentos. Además, esto contrarrestaría el posible sesgo introducido por los profesionales que etiquetaran las imágenes. Sin embargo, al realizarse de esta forma se está produciendo una pérdida de información que realmente se puede extraer de las imágenes. La solución a este inconveniente consistiría en la creación de otro

conjunto de datos que sí clasifique realmente en las mismas 4 clases que el conjunto de entrenamiento, o bien aumentar considerablemente el número de datos y vídeos de la primera base de datos utilizada, de forma que el número de datos final sea similar al empleado en los conjuntos privados de [14]–[16].

En cualquier caso, en este estudio se logra una mayor precisión que en otros estudios similares. Concretamente, Pogorelov et al. obtuvieron una F1 de 0.899 [17], Young Lee et al. una exactitud de 0.85 [16], y Zhou et al. una precisión de 0.933 [14].

#### 4. Conclusiones

En este estudio se han investigado e implementado estrategias con las que desarrollar un sistema de apoyo para la evaluación de calidad en colonoscopias basadas en técnicas de inteligencia artificial, alcanzando con el algoritmo final una puntuación F1 de 0,939. Esto demuestra que es posible reconocer el grado de limpieza de la mucosa intestinal de manera objetiva y, posteriormente, cuantificar el indicador de tasa de preparación intestinal a partir de imágenes de una colonoscopia.

Los sistemas basados en IA están ampliamente extendidos en el ámbito de la detección y clasificación de pólipos, pero es menor la dedicación a otros indicadores de calidad como el propuesto en este trabajo. Aunque existan implementaciones capaces de detectar y evaluar la limpieza y preparación del colon, estas tienen aún margen de mejora para sacarles el máximo partido. Realizar una detección de calidad sobre vídeo permitiría la obtención de información que podría crear un informe objetivo con el porcentaje de áreas de adecuada limpieza inicial, así como el número de veces que se necesita limpiar el área para favorecer la visión.

El objetivo final es reunir en una aplicación algoritmos similares al propuesto que cubran, además, otros indicadores de calidad, de forma que puedan ser cuantificados y calificados objetivamente. Esto conllevaría una mayor adherencia y aceptación hacia los criterios y recomendaciones por parte de los profesionales, de forma que se vería reflejado en mejoras en el rendimiento.

#### Referencias

[1] Sociedad Española de Oncología Médica, “Las cifras del cáncer en España,” 2022.

[2] A. M. D. Wolf et al., “Colorectal cancer screening for average-risk adults,” *A Cancer J. Clin.*, vol. 68, pp. 250–281, 2018.

[3] A. Sánchez del Río, P.-R. Shirley, J. López- Picazo, F. Alberca de las Parras, J. Júdez, and J. León-Molina, “Indicadores de calidad en colonoscopia. Procedimiento de la colonoscopia,” *Rev. Española Enfermedades Dig.*, vol. 110, no. 5, pp. 316–326, 2018.

[4] A. Arcelay et al., “Adaptación de un Modelo de Gestión de Calidad Total al sector sanitario,” *Rev. calid. asist.*, vol. 15, no. 3, pp. 184–191, 2000.

[5] B. Rembacken et al., “Quality in screening colonoscopy: Position statement of the European Society of

Gastrointestinal Endoscopy (ESGE),” *Endoscopy*, vol. 44, pp. 957–968, 2012.

[6] M. F. Kaminski et al., “Performance measures for lower gastrointestinal endoscopy: a European Society of Gastrointestinal Endoscopy (ESGE) Quality Improvement Initiative,” *Endoscopy*, vol. 49, pp. 1–20, 2017.

[7] R. Bisschops et al., “Overcoming the barriers to dissemination and implementation of quality measures for gastrointestinal endoscopy: European Society of Gastrointestinal Endoscopy (ESGE) and United European Gastroenterology (UEG) position statement,” *Endoscopy*, vol. 53, no. 2, pp. 196–202, 2021.

[8] ESGE, “ESGE Quality Check App.” [Online]. Available: <https://www.esge.com/about-us/esge-committees/esge-quality-improvement-committee/quality-check-app/>. [Accessed: 31-Aug-2022].

[9] F. González-Huix Lladó, M. Figa Francesch, and C. Huertas Nadal, “Criterios de calidad que deben exigirse en la indicación y en la realización de la colonoscopia,” *Gastroenterol. Hepatol.*, vol. 33, no. 1, pp. 33–42, 2010.

[10] S. Morán Sánchez et al., “Valoración de la calidad en la práctica de la colonoscopia,” *Rev. Esp. Enfermedades Dig.*, vol. 101, no. 2, pp. 107–116, 2009.

[11] L. Rivero-Sánchez and M. Pellisé, “Preparación para colonoscopia. ¿Algún avance significativo en el horizonte?,” *Gastroenterol. Hepatol.*, vol. 38, no. 4, pp. 287–300, 2015.

[12] E. J. Lai, A. H. Calderwood, G. Doros, O. K. Fix, and B. C. Jacobson, “The Boston bowel preparation scale: a valid and reliable instrument for colonoscopy-oriented research,” *Gastrointest. Endosc.*, vol. 150, no. 2, pp. 396–405, 2009.

[13] B. T. Clark et al., “Quantification of Adequate Bowel Preparation for Screening or Surveillance Colonoscopy in Men,” *Gastroenterology*, vol. 69, no. 3, pp. 620–625, 2016.

[14] J. Zhou et al., “A novel artificial intelligence system for the assessment of bowel preparation (with video),” *Gastrointest. Endosc.*, vol. 91, no. 2, pp. 428–435.e2, 2020.

[15] J. H. Nam, D. J. Oh, S. Lee, H. J. Song, and Y. J. Lim, “Development and verification of a deep learning algorithm to evaluate small-bowel preparation quality,” *Diagnostics*, vol. 11, no. 6, p. 1127, 2021.

[16] J. Y. Lee, A. H. Calderwood, W. Karnes, J. Requa, B. C. Jacobson, and M. B. Wallace, “Artificial intelligence for the assessment of bowel preparation,” *Gastrointest. Endosc.*, vol. 95, no. 5, pp. 512–518.e1, 2022.

[17] K. Pogorelov et al., “Nerthus: A bowel preparation quality video dataset,” in *Proceedings of the 8th ACM Multimedia Systems Conference, MMSys 2017*, 2017, pp. 170–174.

[18] H. Borgli et al., “HyperKvasir, a comprehensive multi-class image and video dataset for gastrointestinal endoscopy,” *Sci. Data*, vol. 7, no. 1, pp. 1–14, 2020.

# Evaluación de la ablación quirúrgica mediante un láser Nd:YAG en aplicaciones dermatológicas sobre tejido porcino

L. Santamaría-Bustamante, Á. Isla-Calvo, J. L. Arce-Diego, F. Fanjul-Vélez

Grupo de Ingeniería Biomédica, Departamento TEISA, Universidad de Cantabria, Av. de los Castros 46, 39005 Santander (Cantabria), España

[arcedj@unican.es](mailto:arcedj@unican.es), [fanjulf@unican.es](mailto:fanjulf@unican.es)

## Resumen

*Uno de los efectos de la radiación óptica sobre tejidos biológicos con utilidad clínica es la ablación. La capacidad de producir la remoción de una parte del tejido, de forma controlada y sin contacto, constituye un elemento del máximo interés quirúrgico, tanto en procedimiento no invasivos o mínimamente invasivos. Para que este proceso ablativo sea controlado, es preciso conocer la influencia de los parámetros de la fuente, habitualmente láser, particularmente la irradiación, el tiempo de exposición, la tasa de repetición, el tiempo o la energía del pulso, así como la distribución transversal del haz, entre otros. El láser Nd:YAG es uno de los más empleados. En este trabajo se estudia la interacción ablativa de un láser Nd:YAG sobre tejidos biológicos epiteliales porcinos. Para ello se varían los parámetros principales del haz, se aplica sobre diferentes muestras y se evalúan los efectos tanto ablativos como térmicos. Además, se trata de identificar automáticamente el efecto producido mediante algorítmica de clasificación sobre propiedades ópticas medidas. Tanto el estudio de los efectos ablativos y térmicos como la identificación del estado del tejido permiten contribuir a la precisión de las intervenciones de cirugía láser guiada, como por ejemplo la resección tumoral.*

## 1. Introducción

Los láseres permiten aplicaciones biomédicas en las que pueden destruir y manipular tejidos biológicos con una precisión sin precedentes. En particular, aunque los láseres aparecieron en la década de los sesenta, el proceso de ablación no se informó hasta 1970, lo que puede explicarse por una comprensión limitada de los mecanismos que gobiernan la interacción entre la luz y el tejido biológico. Pero, desde entonces, el uso del láser en los campos de la medicina ha experimentado un crecimiento importante gracias a sus aplicaciones tanto de cirugía como de diagnóstico no invasivas.

La ablación de los tejidos biológicos mediante el uso del láser es una técnica muy utilizada en la actualidad debido a que este tipo de luz puede trabajar a prácticamente una sola longitud de onda y, gracias a que tiene una emisión muy concentrada y una gran potencia, puede llegar a atravesar las células y profundizar en el tejido ofreciendo múltiples ventajas, como la precisión, la minimización de cicatrices y un menor tiempo de recuperación [1]. Existen varios tipos de láseres para la aplicación de esta técnica ablativa, pero uno de los más utilizados en cirugía es el Nd:YAG. El uso de estas fuentes de luz en aplicaciones médicas ofrece ventajas respecto a otro tipo de radiación,

o incluso frente a procedimientos de corte mecánico, como por ejemplo la posibilidad de integrar estas fuentes en sistemas no invasivos y realizar cortes en ausencia de contacto.

Pero un aspecto crítico en el uso de estas aplicaciones médicas es el control del volumen ablacionado, debido a que hay que tener en cuenta factores como el tipo de láser utilizado, sus parámetros y por supuesto el tipo de tejido biológico que se está tratando [2]. Este aspecto es de gran importancia en intervenciones que se realizan en zonas donde existe la posibilidad de provocar daño térmico a tejido sano adyacente e incluso existir el riesgo de su perforación. Por lo que, para desarrollar cualquier aplicación médica en la que utilicemos fuentes ópticas se deben conocer las propiedades ópticas de los tejidos biológicos a tratar y los posibles efectos que esta interacción luz-tejido pueda causar para tener así una adecuada planificación quirúrgica y poder estimar el volumen ablacionado. Además, resulta del máximo interés tratar de identificar de manera automática el estado de un tejido biológico, especialmente en el ámbito de las intervenciones de cirugía guiada, en las que han podido producirse efectos de coagulación o térmicos no deseados.

En este trabajo se explicarán, en primer lugar, los mecanismos de interacción luz-tejido biológico, particularmente la ablación inducida por plasma. En segundo lugar, se establece el procedimiento experimental de aplicación de haces ópticos sobre tejidos biológicos epiteliales porcinos. Posteriormente se llevan a cabo procedimientos de ablación experimentales en diferentes muestras y con diferentes parámetros, con el objetivo de determinar y evaluar el volumen ablacionado, así como el estado del tejido automáticamente.

## 2. Mecanismos de interacción luz-tejido

El tejido biológico es un medio ópticamente no homogéneo y absorbente, y la propagación de la luz dentro de un tejido depende de las propiedades de esparcimiento y absorción de sus componentes, es decir, de las células, las estructuras de células y fibras y los orgánulos celulares. Además, en particular, los parámetros como el tamaño, la forma y densidad de las partículas, las propiedades de la sustancia fundamental alrededor de las partículas que se esparcen y los estados de polarización de la luz incidente juegan un papel muy importante en la propagación de la luz en medios turbios.

Existen múltiples mecanismos que pueden ocurrir al aplicar luz láser en un tejido biológico y se van a diferenciar dependiendo de las características específicas del tejido y de los parámetros del láser [2]. Son de gran importancia en estos tipos de mecanismos, las propiedades ópticas del tejido que determinan la transmisión total del tejido a una determinada longitud de onda, concretamente los coeficientes de reflexión, absorción y esparcimiento. También son determinantes las propiedades térmicas, es decir, la capacidad y conductividad térmicas en el tejido, y los parámetros de la radiación láser, como la longitud de onda, tiempo de exposición, energía aplicada, tamaño del punto focal, densidad de energía y potencia.

Se establecen cinco categorías de tipos de interacción luz-tejido biológico. Las interacciones fotoquímicas, interacciones térmicas, fotoablación, ablación inducida por plasma y fotorruptura. Estos procesos se distinguen y son controlados por la irradiancia y la duración de la exposición láser, pero pueden compartir un valor común en forma de densidad de energía, que varía de 1 J/cm<sup>2</sup> a 100 J/cm<sup>2</sup>.

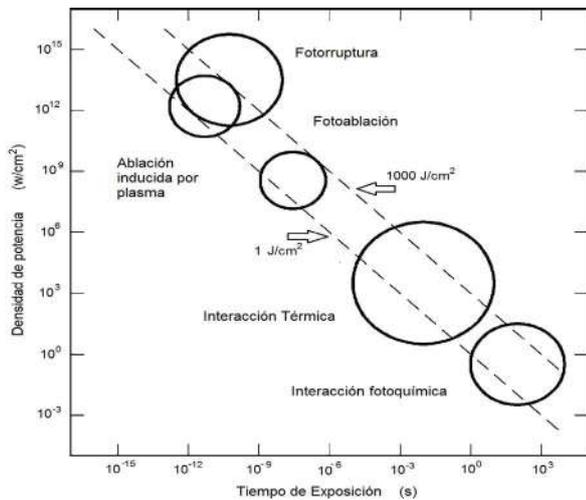


Figura 1. Representación esquemática de los efectos posibles de la interacción luz-tejido biológico en función de la irradiancia y del tiempo de exposición [2].

En la Figura 1 se presenta el esquema de la interacción luz-tejido biológico donde se encuentran representados los cinco mecanismos de interacción. Las dos diagonales representan los flujos de energía constante a 1 J/cm<sup>2</sup> y a 100 J/cm<sup>2</sup>. En los ejes de abscisas y ordenadas se encuentran expresados el tiempo de exposición (s) y la densidad de potencia aplicada o irradiancia (W/cm<sup>2</sup>), respectivamente. Se puede dividir en cinco partes dependiendo de la escala temporal, que va a ser el parámetro encargado de diferenciar esos mecanismos, junto con la irradiancia. Las ondas continuas o tiempos de exposición >1s para interacciones fotoquímicas, 1 ms hasta 1 μs para interacciones térmicas, 1 μs hasta 1 ns para fotoablación y <1ns para ablación inducida por plasma y fotorruptura [3], [4].

En este trabajo es de particular interés el mecanismo de la ablación inducida por plasma, así como los efectos térmicos. Mediante el proceso de ablación inducida por

plasma se puede lograr una extracción muy limpia y bien definida del tejido sin evidencia de daño térmico o mecánico, eligiendo los parámetros láser apropiados. Este tipo de ablación es causada principalmente por la propia ionización del plasma, y es el principal efecto buscado en aplicaciones quirúrgica mediante fuentes láser.

La intensidad de campo eléctrico local determina cuándo se logra la ruptura óptica por lo que es el parámetro más importante de este mecanismo. Si este campo eléctrico aplicado fuerza la ionización de moléculas es cuando se produce la ruptura óptica, es decir, cuando el campo eléctrico ha superado un umbral. Se puede relacionar la propia fuerza del campo eléctrico con la densidad de potencia local, I, a través de la ecuación electrodinámica básica, donde para pulsos de picosegundos, la intensidad de umbral típica es de 10<sup>11</sup> W/cm<sup>2</sup>, por lo tanto, el campo eléctrico correspondiente es de 10<sup>7</sup> V/cm, que proporciona la condición necesaria para la ionización del plasma. En unos pocos cientos de picosegundos, se crea una densidad de electrones libres muy grandes con valores típicos de 10<sup>18</sup> cm<sup>-3</sup> en volumen focal del haz láser. Y esta generación de plasma debido a un intenso campo eléctrico se denomina ruptura dieléctrica.

Este tipo de interacción de la ablación inducida por plasma puede utilizarse tanto en aplicaciones quirúrgicas como para el diagnóstico. Mediante un análisis espectroscópico de la chispa del plasma inducido, la densidad de electrones libres y la temperatura pueden ser evaluados. Además, la información sobre la consistencia química del objetivo puede obtenerse, lo que permite sacar ciertas conclusiones sobre el estado de salud del volumen del tejido investigado [5], [6]. Asimismo, pueden aparecer efectos térmicos no deseados en el tejido biológico ablacionado, así como en tejido sano adyacente como efecto secundario.

### 3. Procedimiento experimental

Para obtener los diferentes tipos de tejidos se utiliza un modelo porcino. Las muestras de tejido in vitro se obtienen de un proveedor local. La preparación de las muestras se realiza mediante su corte con el uso de un bisturí, obteniendo muestras de tejido epitelial de unos 3 mm de espesor. Estas muestras se colocan en placas y así se delimita el área de acción donde se aplicará la radiación láser en un total de nueve puntos por muestra.

La fuente óptica utilizada en este procedimiento experimental es un láser de tipo Nd:YAG de clase IV, que trabaja a una longitud de onda de 532 nm (radiación en el verde, armónico doblado de la emisión fundamental del Nd:YAG), con duraciones de pulso en el rango de los nanosegundos, en modo Q-conmutado y a una frecuencia de repetición de pulso de 15 Hz. Se utilizan energías de pulso de 50 mJ, 85 mJ y 120 mJ, con tiempos de exposición de 180 seg, 320 seg y 720 seg.

Las muestras se colocan separadas 15 cm de la fuente láser (Figura 2) y se irradian en cada punto delimitado, aumentando tanto la energía como el tiempo de exposición a medida que avanzamos por filas y columnas,

respectivamente, partiendo desde el punto uno hasta el nueve de la Figura 3.



Figura 2. Montaje de irradiación de la muestra mediante láser Nd:YAG.

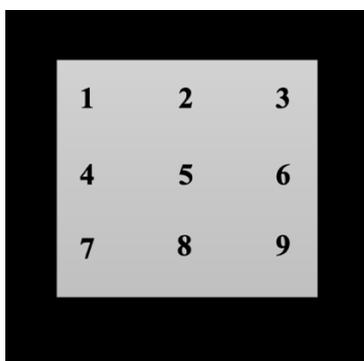


Figura 3. Esquema de irradiación de la muestra.

Además de identificar las profundidades de ablación en función de la energía de los pulsos y tiempo de exposición, se lleva a cabo un análisis de los efectos térmicos no deseados en las muestras tratadas. Así, junto a la zona ablacionada, en caso de que se produzca, existen partes del tejido sano adyacente que sufren o bien una desnaturalización de tipo coagulación, o bien un efecto de carbonización (visible claramente en algunas muestras de la Figura 4). En un procedimiento quirúrgico de ablación láser constituyen efectos en principio no deseados, por cuanto se pretende que el único efecto sea una ablación limpia sin afcción al tejido sano adyacente. Dado que en un entorno de intervención quirúrgica la aparición de estos efectos secundarios puede no ser advertida, es de gran interés tratar de implementar un procedimiento automatizado de identificación de efectos de desnaturalización y carbonización basado en la medida del coeficiente de absorción.

#### 4. Resultados y discusión

Una vez aplicada la radiación láser en los nueve puntos de acción de cada tejido, se observan varios efectos, tanto térmicos como la completa ablación del tejido. La aparición de unos efectos u otros va a depender de la combinación de tiempo de exposición y energía de pulso que se ha utilizado.

Los efectos térmicos más observados en cada uno de los puntos son la coagulación del tejido y su carbonización,

pero en los puntos con los parámetros más altos se puede ver en la Figura 4 cómo se logra la ablación del tejido.



Figura 4. Muestra de tejido epitelial tratado siguiendo el esquema de la Figura 3.

Una vez se han estudiado los efectos causados debido a esta interacción luz-tejido biológico, se procede a medir tanto la profundidad de ablación y efectos térmicos como los diámetros de cada uno de los puntos tratados. En la Figura 5 se puede ver cómo existe una evolución de esta profundidad a medida que aumentamos los parámetros del láser, teniendo en el punto de la derecha solo efectos de coagulación del tejido con una profundidad de 0.7 mm, pasando por el del centro, en el que ya se presenta una ablación junto con la carbonización del tejido con una profundidad de 1.4 mm y 3 mm junto con la coagulación del tejido. Y, por último en el punto de la izquierda se presenta ahora mayor volumen ablacionado que el punto anterior. En este caso, la imagen corresponde a la segunda columna del tejido tratado de la Figura 4, por lo que se ha utilizado un tiempo de exposición de 320 seg y lo que se ha variado es la energía de pulso, partiendo de 50 mJ, 85mJ y 120mJ de derecha a izquierda.

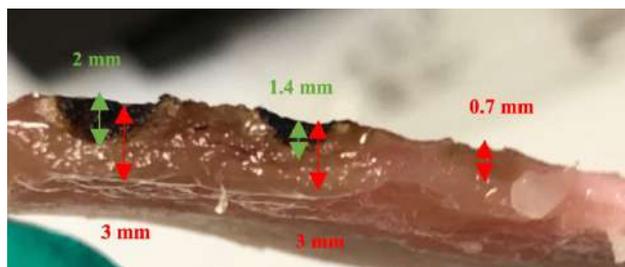


Figura 5. Profundidades de ablación de la segunda columna del tejido epitelial tratado con radiación láser.

Si se representan gráficamente estos valores de profundidad de ablación, conjuntamente para los valores obtenidos de otras muestras de tejido epitelial tratado también con los mismos parámetros láser, se puede observar su evolución y la escasa desviación que presentan cada uno de los resultados en cada punto para valores elevados de energía y tiempo de exposición (Figura 6).

Con estos resultados es posible establecer una primera correlación entre los valores de energía de pulso, tiempo de exposición y profundidad de ablación. Estos resultados son de interés potencial en la planificación quirúrgica.

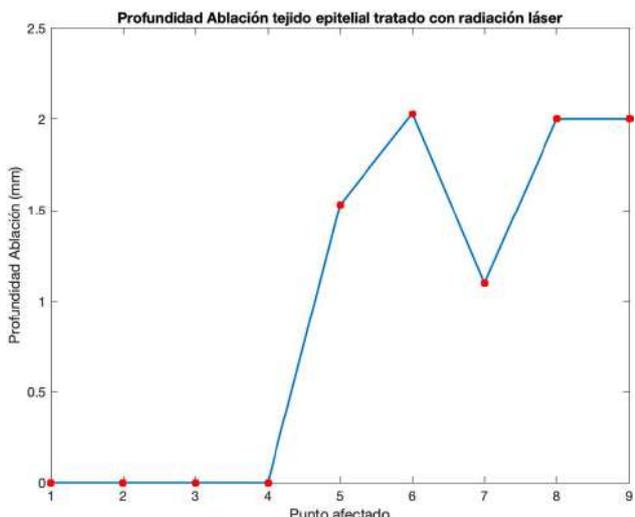


Figura 6. Profundidad de ablación provocada al aplicar radiación láser en el tejido epitelial.

A continuación se realizan medidas del coeficiente de absorción de cada uno de los puntos tratados. Se seleccionan los puntos en los que se aprecia desnaturalización o carbonización, y se establecen grupos de clasificación. Se realiza un análisis de componentes principales (PCA), y se emplean los coeficientes en un sistema de clasificación de tipo Máquina de Soporte Vectorial (SVM). Los resultados aparecen gráficamente en la Figura 7 con solo dos componentes principales para poder representarlo. Los resultados del error de correlación cruzada arrojan un valor de 0.1429, lo que resulta esperanzador para un sistema automático de identificación de estos efectos en tiempo real.

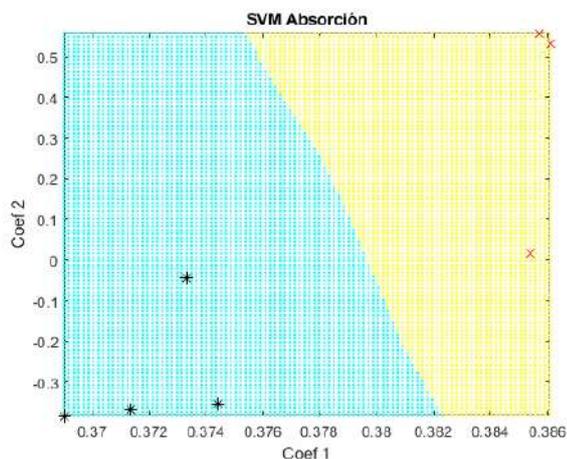


Figura 7. Representación gráfica del resultado del clasificador de tipo SVM aplicado sobre los coeficientes de absorción. Las cruces rojas se corresponden con las muestras desnaturalizadas, mientras que los asteriscos negros indican muestras carbonizadas.

## 5. Conclusiones

Tras los experimentos realizados, se puede observar que a medida que se incrementan los parámetros de la fuente óptica, en este caso el láser de Nd:YAG, se obtienen volúmenes de ablación mayores y la aparición de efectos térmicos como la coagulación o carbonización del tejido

son cada vez más significativos, como era de esperar. En todo caso, las medidas obtenidas permiten hacer una estimación a priori del efecto resultante en cuanto a profundidad de ablación, clave en las intervenciones quirúrgicas, especialmente epiteliales.

Además, con objeto de identificar automáticamente daños térmicos en los tejidos sanos adyacentes, se han realizado medidas del coeficiente de absorción de los puntos irradiados. Los resultados del algoritmo de clasificación indican una tasa significativa de acierto. Este hecho permitiría realizar valoraciones de la cirugía en tiempo real, de una forma automatizada e incluso estar integrado en el dispositivo quirúrgico que se esté utilizando para avisar al cirujano. Un campo de aplicación clínica sería el de la resección tumoral controlada.

## Agradecimientos

Este Trabajo ha sido parcialmente financiado por el proyecto del Plan Nacional de I+D+i “Cribado diagnóstico de microorganismos mediante microscopia avanzada e inteligencia artificial en patologías humanas” (PID2021-127691OB-I00), del Ministerio de Ciencia e Innovación, cofinanciado con fondos FEDER.

## Referencias

- [1] Félix Fanjul-Vélez, Irene Salas-García and José Luis Arce-Diego. “Analysis of laser surgery in non-melanoma skin cancer for optimal tissue removal”, *Laser Physics* 25, 025606 (8pp) (2015).
- [2] Markolf H. Niemz. *Laser-Tissue Interactions*. Third, Enlarged Edition. Springer-Verlag Berlin Heidelberg 2004, 2007 (ISBN: 3-540-40553-4).
- [3] F. H. Loesel, M. H. Niemz, J. F. Bille, T. Juhasz, “Laser-Induced Optical Breakdown on Hard and Soft Tissues and Its Dependence on the Pulse Duration: Experiment and Model,” *IEEE Journal of Quantum Electronics* 32, 1717-1722 (1996).
- [4] A. Vogel, V. Venugopalan, “Mechanisms of Pulsed Laser Ablation of Biological Tissues,” *Chem. Rev.* 103, 577-644 (2003).
- [5] American Society for Gastrointestinal Endoscopy, “Mucosal ablation devices,” *Gastrointestinal Endoscopy* 68, 1031-1042 (2008).
- [6] N. Muguruma, K. Okamoto, T. Kimura, K. Kishi, T. Okahisa, S. Okamura, T. Takayama, “Endoscopic ablation therapy for gastrointestinal superficial neoplasia,” *Digestive Endoscopy* 24, 139–149 (2012).

# Metodología de diseño y fabricación de guías personalizadas mediante impresión 3D para cirugía de columna. Patología: Escoliosis idiopática en adolescentes

P. Candel Cano<sup>1</sup>, G. Martínez Bordes<sup>2</sup>

1 Escuela Superior de Ingeniería, Ciencia y Tecnología, Universidad Internacional de Valencia (VIU), 46002 Valencia, España, {pcandelc@gmail.com }

2 Universidad Austral de Chile, Valdivia, Chile {gabriela.martinez@uach.cl}

## Resumen

Se describe y aplica una metodología para el diseño y fabricación de guías quirúrgicas, particularizadas para el instrumental utilizado en la instalación de tornillos intrapediculares durante la intervención de corrección de deformaciones de columna por escoliosis. Inicialmente, se realiza un estudio sobre la patología, las técnicas quirúrgicas y las tecnologías aplicables, y posteriormente se propone una metodología que engloba desde los métodos y parámetros de adquisición de imágenes diagnósticas, hasta el diseño y fabricación del instrumental quirúrgico. Para el diseño se utilizaron como referencia las imágenes diagnósticas resultantes de una tomografía computarizada (TAC) del paciente, que una vez tratadas y transformadas en modelos tridimensionales, sirven de base para la planificación de la cirugía y diseño de guías de trayectoria para el taladrado previo a la instalación de los tornillos. La fabricación, es realizada mediante técnicas de impresión 3D (en este caso, FDM o modelado por deposición fundida). Finalmente, las guías fueron evaluadas *In Vitro* en un fantoma de columna y validadas por un facultativo experto en cirugías de columna.

## 1. Introducción

La escoliosis es la deformidad del esqueleto axial en el plano anteroposterior. Incluye deformidad tridimensional con rotación vertebral y puede acompañarse de alteraciones en el plano sagital. Para que se considere escoliosis debe tener más de 10° de angulación [1].

La proporción de casos de escoliosis idiopática se aproxima al 2% de la población, siendo cinco veces más frecuente en niñas que en niños [2].

Para el diagnóstico de esta patología se realizan diversas pruebas y mediciones entre los que se encuentran el test de Adam que consiste en una exploración del paciente flexionado hacia delante, con el objetivo de observar el desnivel de hombros, asimetría del talle o la percepción de deformidad o giba [1]. Una vez obtenido un resultado positivo del test de Adams, se procederá a realización de un estudio radiológico del paciente para evaluar la curva del ángulo de Cobb, que es el ángulo que forman las vértebras más inclinadas en la parte superior e inferior de la curva que describe la columna vertebral. Esta medición se compara con unas tablas de clasificación que ayudan al facultativo a decidir el tratamiento más apropiado.

Una vez diagnosticado al paciente con escoliosis, se procederá a prescribir diversos tratamientos en función de la gravedad del estado de este, siempre intentando evitar/posponer la intervención quirúrgica.

Como tratamiento alternativo a la cirugía se recomiendan el uso de un corsé ortopédico personalizados para cada individuo, hasta alcanzar la madurez esquelética, además de la realización de una actividad física moderada. [1].

En la siguiente Figura 1 se muestra un diagrama resumen de los posibles tratamientos aplicables, según diversos parámetros medidos durante el diagnóstico.

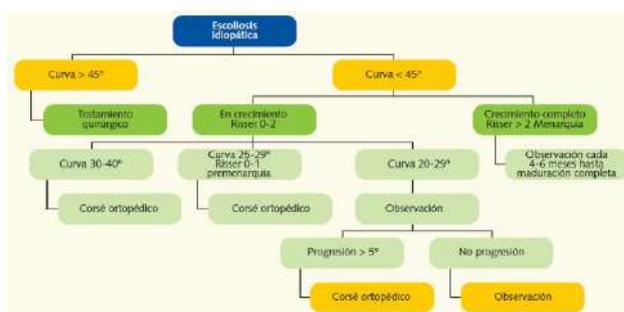


Figura 1. Algoritmo de tratamiento de la escoliosis[1]

Una de las técnicas quirúrgicas más utilizadas actualmente es la colocación de tornillos pediculares. La técnica, que consiste en la instalación de tornillos monoaxiales o poliaxiales (Figura 2) en los pedículos de las vértebras dorsales y/o lumbares, que sirven de anclaje para la instalación de barras metálicas que aportan la rigidez necesaria para la rectificación y desrotación de la columna del paciente [3].



Figura 2. Ejemplo de tornillos intrapediculares; monoaxial (Izq.) y poliaxial (Dcha.)

Dada la complejidad y riesgo de una cirugía de columna, es de gran ayuda poner todas las tecnologías al alcance de nuestras manos, para reducir al mínimo la incertidumbre, los riesgos o posibles errores humanos durante una intervención.

En la siguiente imagen (Figura 3) se puede observar una radiografía de la paciente preoperatoria (izq.) y postoperatoria (drcha.).



**Figura 3.** Radiografías anteroposteriores de raquis completo de paciente intervenida quirúrgicamente por escoliosis dorsolumbar. Situación preoperatoria y postoperatoria [4]

Tras los avances en tecnología de los últimos años, es posible desarrollar nuevas metodologías de trabajo que beneficien en gran medida a facultativos, los cuales encuentran habitualmente pequeñas dificultades, o incomodidades, a la hora de realizar ciertos procedimientos, que son fácilmente subsanables con la ayuda de tecnologías de uso cotidiano en otros sectores, pero que no son aplicadas habitualmente en el ámbito hospitalario. De ahí la importancia de la sinergia entre distintas disciplinas profesionales como la ingeniería y la medicina.

Es necesario tener en cuenta que las radiaciones ionizantes son potencialmente cancerígenas, y sus dosis son acumulativas. Durante el proceso de diagnóstico de escoliosis, se realizan diversas exposiciones, superando fácilmente la veintena [5]-[6]. Si el caso requiere intervención quirúrgica, el número de exposiciones se incrementa notablemente.

En un estudio realizado ya en el año 2000 sobre la relación entre el cáncer y la exposición a radiación durante los procesos de diagnóstico, cirugía y control postoperatorio, se demostró, que la mortalidad por cáncer de mama se había multiplicado por tres en las mujeres que padecieron escoliosis durante su infancia [7].

Con el uso de guías para la instalación de los tornillos, podría reducirse muy considerablemente el número de irradiaciones a las que se somete al paciente en la cirugía, reduciendo notablemente la probabilidad de padecer cáncer a lo largo de la vida del paciente.

La justificación de este proyecto viene dada por la necesidad de crear un marco metodológico para el diseño

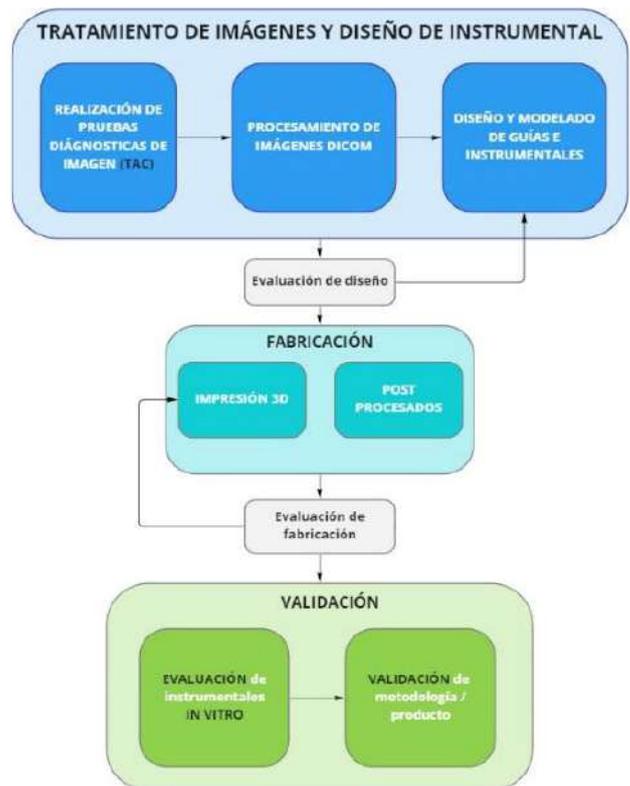
de nuevos instrumentales personalizados para cada paciente y tipo de cirugía. En él se describen los pasos a seguir durante el proceso de diseño y fabricación de guías quirúrgicas, y diversos criterios a tener en cuenta, particularmente para una cirugía de columna, en la que sea necesaria la instalación de tornillos intrapediculares.

De esta manera se pretende mejorar de la técnica quirúrgica actual para avanzar en aspectos como: planificación quirúrgica, reducción la irradiación intraoperatoria, éxito quirúrgico, reducción de tiempos en el quirófano y posibilidad de mejoras a coste reducido.

Valga entender, que esta metodología, desde una visión generalizada, y sin ahondar en las particularidades de la patología de columna, es aplicable para casi cualquier caso en el que sea necesaria el diseño y fabricación de instrumental personalizado.

## 2. Metodología de diseño y fabricación de guías quirúrgicas mediante impresión 3D

En el siguiente diagrama (Figura 4), se describe el flujo de trabajo propuesto a realizar durante el estudio de cada caso/paciente.



**Figura 4.** Flujo de trabajo

### 2.1. Realización de pruebas diagnósticas de imagen

Se realizará una tomografía computarizada del paciente con los parámetros adecuados para la posterior reconstrucción del modelo 3D, pudiendo ser realizable mediante otro tipo de imágenes de diagnóstico como, por ejemplo, la resonancia magnética.

## 2.2. Procesamiento de imágenes DICOM (Segmentación)

Procesamiento de imágenes y la generación de un modelo 3D a partir de las imágenes 2D procedentes de la prueba diagnóstica.

Se utilizará un software especializado en procesamiento de imágenes del que se obtendrán los modelos 3D necesarios de la anatomía del paciente para la planificación de la cirugía, como se muestra en la Figura 5.



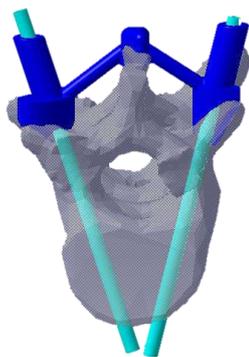
**Figura 5.** Reconstrucción ósea del paciente en 3D

## 2.3. Diseño y modelado de guías e instrumentales

Importación de los modelos anatómicos a un software de diseño CAD/CAM, para el diseño y modelado del instrumental quirúrgico.

En este paso es fundamental la participación de facultativos, que aporten los conocimientos y experiencia durante el proceso de diseño y su posterior validación.

A continuación, se muestra, a modo referencial, un diseño de guía y trayectorias, sobre vertebras torácicas para la instalación de tornillos intrapediculares (Figura 6).

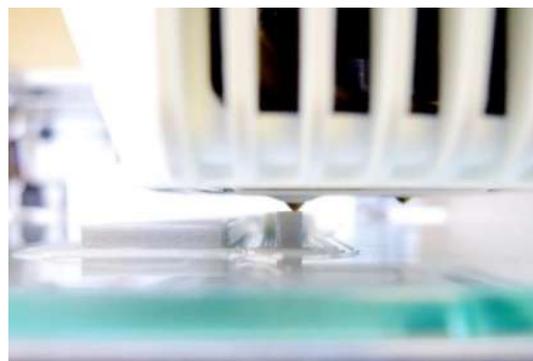


**Figura 6.** Ejemplo de diseño de guía (azul) y simulación de trayectorias de los tornillos pediculares (azul celeste)

## 2.4. Fabricación

La fabricación se realizará mediante técnicas de impresión 3D (FDM, SLA, SLS...) siendo necesaria la utilización de materiales biocompatibles (ver Figura 7).

Tras la fabricación es posible que haya que realizar ligeros trabajos de postprocesado, como eliminación de soportes de fabricación, eliminación de rebabas, repaso de taladros.



**Figura 7.** Fabricación mediante impresión 3D

Es importante remarcar que la fabricación de este tipo de instrumentales mediante tecnologías convencionales sería prácticamente inviable, debido, a los tiempos y los costes de fabricación, entre otros factores.

## 2.5. Validación

Al finalizar el proceso de fabricación, el dispositivo médico será sometido a diversos controles de calidad tanto por parte del ingeniero responsable de diseño y fabricación, como por el facultativo/responsable clínico, además, de a una correcta limpieza y esterilización, apropiada al tipo de material de fabricación, siendo en el caso del PLA, esterilización de baja temperatura con plasma de peróxido de hidrógeno. La documentación generada deberá ser recogida en informes trazables, que certifiquen la seguridad del paciente.

## 2.6. Certificación

Cualquier dispositivo médico debe cumplir unos requerimientos de calidad y seguridad exigidos por las normativas ISO 13485 y 93/42/ECC, incluidos los productos a medida que sirvan para la planificación quirúrgica, instrumentales, implantes y software especializados.

La actual regulación sanitaria clasifica los dispositivos médicos según la finalidad y el tiempo de exposición al paciente, siendo las guías quirúrgicas consideradas de **clase IIa**. Es por ello que requieren de diferentes procesos de certificación entre los que se encuentran; validación, control dimensional, biocompatibilidad de materiales, trazabilidad y controles del proceso de diseño y fabricación, y esterilización, entre otros. Este proceso no debería implicar un incremento relevante de tiempo de entrega del dispositivo.

## 3. Resultados

Durante el proceso de diseño se realizaron distintas modificaciones de la geometría de las guías de taladrado, de mano del facultativo especialista en cirugía de columna, con la finalidad de alcanzar una correcta adaptación a la técnica quirúrgica, y al instrumental a utilizar simultáneamente. Es muy importante tener en cuenta algunos parámetros de obligado cumplimiento, cómo son; zonas de apoyo, tolerancias de fabricación, ajustes, trayectorias, siendo posible un diseño orgánico y ajustable del resto modelo de la guía.

La figura 8 muestra el posicionamiento de una guía impresa sobre fantoma óseo del paciente. Se utilizó como material de fabricación ácido poliláctico (PLA) e impresión por deposición fundida (FDM) como técnica de fabricación. Se observa el ajuste del instrumental diseñado.



**Figura 8.** Posicionamiento de guía sobre fantoma óseo del paciente

Tras la evaluación de la guía fabricada y probada sobre fantoma óseo de la vértebra acompañado del instrumental a utilizar conjuntamente, se obtuvo una valoración positiva del facultativo.

#### 4. Conclusiones y recomendaciones

Tras el desarrollo del proyecto, es posible obtener las siguientes conclusiones:

- Es un proceso capaz de aportar un gran valor añadido al procedimiento estándar de realización de la cirugía, dotando al cirujano de herramientas personalizadas que le ayuden en la planificación y desarrollo de la cirugía, aportándole tranquilidad y **seguridad**, reduciendo los **tiempos** de quirófano, y por consiguiente, sus respectivos **costes**.
- Podría suponer una solución para **reducir** considerablemente el número de **irradiaciones** intraoperatorias en este tipo de cirugías.
- Es una tecnología de **bajo coste** que permite la fabricación de **dispositivos personalizados**, que sería inviable mediante otras técnicas de fabricación más convencionales. Según estudios recientes, los costes de quirófano se reducen considerablemente (entre 30' y 120' dependiendo del tipo de cirugía) y el tiempo de recuperación postoperatoria, suponiendo una reducción importante de costes que permite implementar este tipo de metodologías sin que suponga un coste añadido para el hospital.[8]
- Es una metodología **extrapolable** a multitud de patologías.

Tras el estudio y aplicación de esta metodología, se proponen distintas recomendaciones desprendidas a lo largo de su desarrollo que facilitarían y agilizarían la aplicación de la misma.

- Establecer, de mano de radiólogos especializados, **protocolos** estándar para la realización de pruebas diagnósticas de imagen orientados a la optimización de tiempo y resultados de este tipo de metodologías.

- **Automatización de procesos** repetitivos mejorando los algoritmos de segmentación actuales, mediante tecnologías como la inteligencia artificial, para el reconocimiento de patologías o deformaciones, que reduzca el tiempo de tratamiento de imágenes y amplíe el marco de especialización del personal a cargo.
- Se recomienda, que durante el proceso de modelado se utilice un **diseño paramétrico** que permita la adaptación fácil y rápida, de guías ya diseñadas a distintos pacientes con patologías similares o variaciones de instrumentales estándar, mediante la modificación de dichos parámetros.
- Realización de estudios y **protocolos de esterilización**, dependiendo del material utilizado en la fabricación, para evitar deformaciones o el propio deterioro y consecuente invalidación del instrumental diseñado.
- Sería interesante, crear un **marco estandarizado de validación** según el tipo de dispositivo, aun siendo complicado debido a la personalización y singularidad del mismo, pudiendo alcanzar unos parámetros generales adaptables a cada caso.

#### Referencias

- [1] Álvarez García de Quesada LI, Núñez Giralda A. Escoliosis idiopática. Rev Pediatr Aten Primaria. 2011;13:135-46.
- [2] “Tratamiento Quirúrgico De Escoliosis En Personas Menores De 25 Años.” Orientación En Salud. Superintendencia De Salud, Gobierno De Chile., [www.supersalud.gob.cl/difusion/665/w3-article-584.html](http://www.supersalud.gob.cl/difusion/665/w3-article-584.html)
- [3] Langlais, T., R. Vialle, and J. Sales de Gauzy. 2021. “Tratamiento Quirúrgico de Las Escoliosis Idiopáticas.” EMC - Técnicas Quirúrgicas - Ortopedia y Traumatología 13(1):1–16. doi: 10.1016/s2211-033x(20)44764-4.
- [4] Ardura Aragón, F., Noriega González, D. C., & Hernández Ramajo, R. (2014). Deformidades de la columna vertebral. *Pediatría Integral*, 18(7), 468–477.
- [5] Luan, F.J., Wan, Y., Mak, K.C. et al. Cancer and mortality risks of patients with scoliosis from radiation exposure: a systematic review and meta-analysis. *Eur Spine J* 29, 3123–3134 (2020). <https://doi.org/10.1007/s00586-020-06573-7>
- [6] Gérvas, J. (2016, 7 mayo). La escoliosis y la muerte por cáncer de mama. *Acta Sanitaria*.
- [7] Doody, M. M., Lonstein, J. E., Stovall, M., Hacker, D. G., Luckyanov, N., & Land, C. E. (2000). Breast cancer mortality after diagnostic radiography: findings from the U.S. Scoliosis Cohort Study. *Spine*, 25(16), 2052–2063. <https://doi.org/10.1097/00007632-200008150-00009>
- [8] Rodríguez-Arias, J. P., Tapia, B., Pampín, M. M., Morán, M. J., Gonzalez, J., Barajas, M., Del Castillo, J. L., et al. (2022). Clinical Outcomes and Cost Analysis of Fibula Free Flaps: A Retrospective Comparison of CAD/CAM versus Conventional Technique. *Journal of Personalized Medicine*, 12(6), 930. MDPI AG. Retrieved from <http://dx.doi.org/10.3390/jpm12060930>

# Evaluación mediante simulación de nuevos enfoques en la terapia de resincronización cardiaca

Cristóbal R Ruiz<sup>1</sup>, Juan F Gómez<sup>1</sup>, Eduardo Castellanos<sup>2</sup>, Jesús Almendral<sup>2</sup>, Beatriz Trenor<sup>1</sup>

<sup>1</sup>Centro de Investigación e Innovación en Bioingeniería (Ci2B), Valencia, España

<sup>2</sup>Fundación Grupo Hospitales Madrid (FGHM), Madrid, España

## Resumen

*La terapia de resincronización cardiaca (CRT) es una técnica ampliamente utilizada en la práctica clínica para solucionar problemas de desincronización en la contracción ventricular. Entre las diferentes variantes de la CRT, la más extendida es la estimulación biventricular (BiVP), no obstante, nuevos enfoques están empezando a ser evaluados. Es aquí donde el empleo de modelos cardiacos computacionales específicos de cada paciente podría constituir una poderosa herramienta para el electrofisiólogo, pues permiten evaluar la efectividad de diferentes enfoques de CRT sin riesgo para el paciente. El objetivo del presente estudio fue por tanto evaluar, mediante simulación, el desempeño de la CRT convencional (BiVP) frente a enfoques más novedosos como la estimulación del área de rama izquierda (LBBAP) o un enfoque híbrido: la CRT optimizada mediante estimulación del área de la rama izquierda (LOT-CRT). Para ello, se utilizó un modelo biventricular en el que se modelaron diferentes patologías: bloqueo de rama izquierda (LBBB), insuficiencia cardiaca (HF) y retardos en el sistema de conducción periférico del ventrículo izquierdo (LVd). Posteriormente se analizaron las señales de ECG y los mapas de activación temporal simulados para determinar el desempeño de cada terapia. Los resultados mostraron que en ausencia de LVd, los nuevos enfoques (LBBAP y LOT-CRT) resultaron ser efectivos y superiores a la BiVP. En presencia de LVd, el beneficio de la terapia fue menor, sin un claro vencedor entre la BiVP y la LOT-CRT.*

## 1. Introducción

La insuficiencia cardiaca (HF) es una enfermedad cardiovascular caracterizada por la incapacidad del corazón de bombear sangre suficiente para satisfacer las necesidades del organismo. A menudo, la insuficiencia cardiaca va acompañada de alteraciones en el sistema de conducción cardiaco (CCS) como el bloqueo de rama izquierda (LBBB) u otros retardos más periféricos del sistema de conducción del ventrículo izquierdo (LVd) [1,2]. Dichas alteraciones bloquean o retrasan la propagación normal del impulso eléctrico por el sistema de conducción del ventrículo izquierdo (VI), provocando una contracción ventricular desincronizada.

La terapia de resincronización cardiaca (CRT) es una técnica ampliamente extendida en el ámbito clínico [3]. La CRT intenta corregir la desincronización de la contracción ventricular mediante estímulos eléctricos aplicados en determinadas regiones del corazón que reemplazan o se coordinan con el estímulo fisiológico proveniente del nodo sinusal para lograr una activación simultánea de ambos ventrículos y mejorar la función cardiaca.

Por lo general, la CRT se implementa en forma de estimulación biventricular (BiVP). Este enfoque consiste en aplicar dos estímulos externos en el miocardio ventricular, sin embargo, gracias al desarrollo de la tecnología médica se están comenzando a utilizar nuevos enfoques [4]. Este es el caso de la estimulación del área de la rama izquierda (LBBAP) que emplea un único estímulo en esa ubicación específica o la CRT optimizada mediante estimulación del área de la rama izquierda (LOT-CRT), que estimula tanto en la región de la rama izquierda como en el epicardio del VI [5,6].

Sin embargo, en la práctica clínica no es posible evaluar múltiples enfoques de CRT en un mismo paciente, pues cada uno de ellos conlleva una serie de riesgos. En este sentido, los modelos cardiacos *in silico* podrían ser de gran utilidad, pues permitirían predecir y comprender los resultados de cada terapia sin riesgo para el paciente. De esta forma, en Strocchi et al. [7] evalúan mediante simulación la eficacia de la CRT al estimular en el haz de His (HBP) o en la rama izquierda (LBBAP) frente al uso de la CRT convencional (BiVP) en casos de LBBB. Sin embargo, en sus simulaciones no tienen en cuenta la presencia de HF o LVd. Otros estudios de simulación tratan de optimizar la CRT estándar variando la posición de los electrodos o el retardo programado [8], pero no consideran otros enfoques de CRT. Por otro lado, hasta dónde saben los autores, no hay estudios de simulación que analicen el desempeño de la LOT-CRT.

Así pues, el presente estudio tuvo como objetivo evaluar, mediante simulación, la efectividad de la LBBAP y la LOT-CRT frente a la BiVP en un mismo paciente. Dónde además de considerar diferentes condiciones patológicas de partida, se tuvieron en cuenta distintos retardos auriculoventriculares (AVD), es decir, las terapias se aplicaron en diferentes instantes temporales con el fin de optimizar los resultados.

## 2. Métodos

### 2.1. Modelo 3D cardiaco

El modelo cardiaco *in silico* utilizado en este estudio fue desarrollado previamente por nuestro equipo [9]. Brevemente, consiste en un modelo biventricular 3D obtenido a partir de imágenes de resonancia magnética con realce tardío (DE-MRI) de un paciente específico.

El modelo también incluye el CCS, formado a partir de elementos lineales y generado en base a un método de

crecimiento estocástico [10]. El CCS en el modelo comienza a nivel del haz de His y se divide en las ramas derecha e izquierda, que posteriormente continúan ramificándose conformando el sistema de Purkinje (PS).

## 2.2. Modelos patológicos

Aunque se ha demostrado que la CRT puede aportar beneficios en pacientes con diversos cuadros clínicos, la mayoría de pacientes candidatos a la implantación de un dispositivo CRT son casos de pacientes con HF y LBBB. Algunos de estos casos, pueden presentar además LVd.

Es por ello que en el presente trabajo se han considerado dos casos de estudio. El primero es un paciente con bloqueo de rama izquierda e insuficiencia cardiaca (LBBB+HF), y el segundo, cursa además con LVd (LBBB+HF+LVd). Para modelar el LBBB, la conductividad en dos elementos lineales consecutivos del CCS modelado se hizo nula. Por otro lado, las condiciones de HF y LVd se reprodujeron reduciendo la conductividad en todo el miocardio y únicamente la porción izquierda del CCS respectivamente. En particular, para modelar la HF, la conductividad global se fijó al 50% de su valor basal, de acuerdo con la reducción en la proteína connexin43 encontrada en dichos pacientes [11], mientras que para modelar el LVd, la conductividad se redujo aún más hasta que el ECG simulado mostró un patrón análogo al encontrado en la práctica clínica [12]. De esta forma, la condición de "LBBB+HF" y "LBBB+HF+LVd" se obtiene al aplicar simultáneamente cada una de las patologías modeladas (LBBB, HF y LVd).

## 2.3. Modelo electrofisiológico

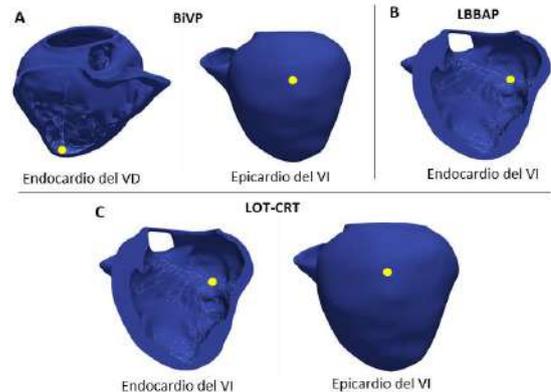
Para describir el comportamiento eléctrico a nivel celular, se ha utilizado una versión modificada del modelo de O'Hara et al. [9,13] para las células ventriculares, mientras que para las células del CCS se utilizó el modelo de Stewart et al. [14].

A nivel tisular, la propagación eléctrica es descrita por el modelo monodominio. Aunque dicho modelo es una simplificación del modelo bidominio, es adecuado para el propósito del estudio y requiere de menor carga computacional [15].

## 2.4. Protocolos de estimulación

Para reproducir los diferentes enfoques de CRT en nuestro modelo, se han aplicado varios protocolos de estimulación en el miocardio y el CCS de acuerdo al procedimiento seguido en la práctica clínica (Fig. 1). Primero, para emular la BiVP, se han aplicado simultáneamente dos estímulos externos, uno en el ápex del ventrículo derecho (VD) y otro en el epicardio de la pared libre del VI. En el caso de la LBBAP, solo se aplicó un estímulo en la región de la rama izquierda, cerca de la bifurcación en sus correspondientes fascículos y capturando simultáneamente tanto el CCS como el miocardio circundante. Finalmente, para reproducir la LOT-CRT se estimuló tanto en el epicardio de la pared libre del VI como en la región septal donde se ubica la rama izquierda. Los estímulos aplicados fueron de 900  $\mu\text{A}/\mu\text{F}$  de intensidad y 2 ms de duración.

Además, con el fin de optimizar los resultados, se han probado dos AVD diferentes. Con ello se persigue lograr la mejor coordinación de los estímulos externos con el estímulo fisiológico, lo que se ha demostrado que mejora los resultados [16]. Estos AVD fueron 195 y 210 ms.



**Figura 1.** Puntos de estimulación (en amarillo) utilizados para simular BiVP (A), LBBAP (B) y LOT-CRT (C). VI = ventrículo izquierdo; VD = ventrículo derecho.

## 2.5. Simulaciones 3D

Para ejecutar las simulaciones se ha utilizado el software ELVIRA [17]. ELVIRA es un solucionador FEM desarrollado específicamente para resolver la ecuación de reacción-difusión anisotrópica del modelo monodominio en problemas de electrofisiología cardiaca. Para las soluciones numéricas se utilizó el método del gradiente conjugado con un paso de tiempo de integración de 0,02 ms.

## 2.6. Simulación del ECG

Para reproducir la señal del ECG de superficie, se utilizó un modelo de torso previamente desarrollado [18] en el que se acopló el modelo biventricular. De esta forma, a partir de los potenciales eléctricos calculados previamente a nivel de órgano usando ELVIRA, se calcularon los potenciales extracelulares a través del modelo de torso utilizando una aproximación del modelo bidominio descrita en Keller et al. [19]. Posteriormente, se obtuvieron las derivaciones precordiales del ECG extrayendo el valor del potencial eléctrico en aquellas ubicaciones del modelo de torso donde habitualmente se colocan los electrodos para adquirir la señal de ECG en la práctica clínica.

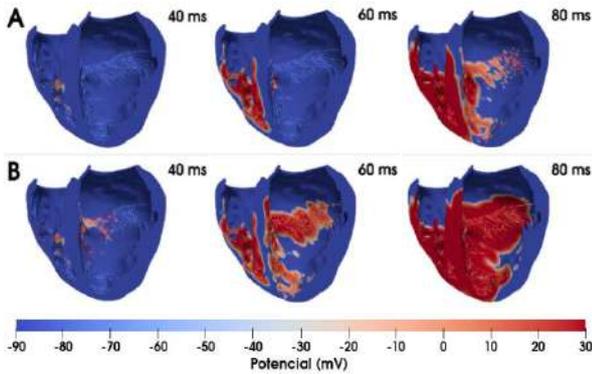
## 2.7. Parámetros evaluados

Para evaluar la efectividad de cada terapia, se calcularon diversas medidas a partir de la señal de ECG simulada y del mapa de isócronas obtenido en las simulaciones (mapa 3D que muestra la activación ventricular en función del tiempo). En particular, a partir de la señal de ECG simulada, se midió la duración del complejo QRS (QRSd), valor ampliamente utilizado en el entorno clínico para evaluar la eficacia de la CRT. No obstante, en estudios previos [9], se ha visto que otros parámetros parcialmente correlacionados con QRSd son también buenos predictores de la eficacia de la CRT. Estos son, el tiempo en el que se alcanza el 90% del área encerrada por el complejo QRS ( $a_{90}$ ) y el intervalo de tiempo hasta alcanzar el 90% de la activación ventricular ( $t_{90}$ ). Aunque son medidas

análogas, la diferencia radica en que una se ha obtenido a partir de la señal de ECG simulada (a90) y la otra a partir del mapa de isócronas (t90), por lo que ambas se complementan.

### 3. Resultados

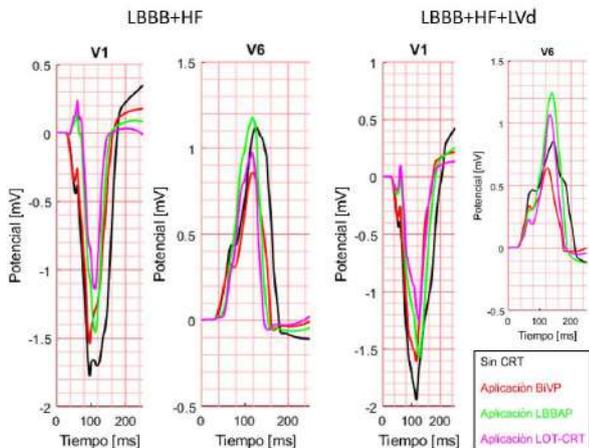
Al aplicar cada una de las terapias, el patrón de activación del tejido cardíaco varía. La siguiente figura (Fig. 2) muestra lo que sucede al aplicar la LBBAP en el caso de LBBB+HF.



**Figura 2.** Propagación eléctrica en condiciones de LBBB+HF (A) y después de aplicar la LBBAP (B).

Como se puede observar, la LBBAP produce una activación más temprana del septo del VI, lo que mejorará la sincronía en la contracción ventricular. De manera análoga, la aplicación de la BiVP y la LOT-CRT genera frentes de activación provenientes del ápex del VD y el septo del VI respectivamente junto a un frente de activación común en ambos enfoques proveniente de la pared lateral del VI. Esto, de nuevo, provocará una mayor sincronía en la activación ventricular mejorando la contracción.

En cuanto a los cambios en la morfología de la señal de ECG al aplicar los distintos enfoques de CRT, se puede observar una disminución no solo en la amplitud de la señal (potencial eléctrico) sino también en la duración del complejo QRS, aproximándose a los valores fisiológicos normales (Fig. 3).



**Figura 3.** Complejo QRS de las precordiales V1 y V6 en ambos casos de estudio (LBBB+HF y LBBB+HF+LVd) sin aplicar ninguna terapia (negro) y tras aplicar los distintos enfoques de CRT planteados (BiVP: rojo, LBBAP: verde; LOT-CRT: rosa).

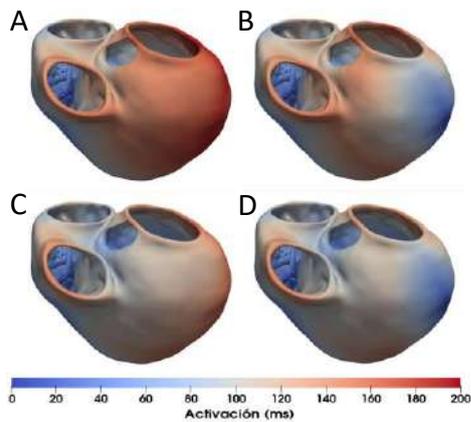
Nótese que en dicha figura se muestran solo los complejos QRS de las derivaciones precordiales V1 y V6, pues es aquí donde resulta más fácil evaluar la patología y el efecto de la terapia. Además, en condiciones de LBBB y HF, al usar enfoques que estimulan el septo del VI (LBBAP y LOT-CRT) se consigue restaurar el patrón rS encontrado habitualmente en la derivación V1 en el caso sano. Los valores de QRSd y a90 medidos para cada caso y terapia se pueden encontrar en la Tabla 1.

**Tabla 1.** Valores de QRSd, a90 y t90 obtenidos para cada caso de estudio y terapia simulada. Los recuadros rojos indican el valor patológico de partida en cada caso de estudio y los recuadros verdes muestran el enfoque de CRT y AVD en el que se logra la mayor reducción.

		QRSd (ms)	a90 (ms)	t90 (ms)			
Caso de estudio	Sano	-	-	95	72	77	
	Valor inicial	-	-	166	123	134	
	LBBB+HF	Aplicación BiVP	AVD = 195 ms	156	114	115	
		Aplicación LBBAP	AVD = 210 ms	151	113	120	
	LBBB+HF+LVd	Aplicación LBBAP	AVD = 195 ms	142	104	108	
		Aplicación LOT-CRT	AVD = 210 ms	146	113	121	
	LBBB+HF	Valor inicial	-	-	198	143	161
		Aplicación BiVP	AVD = 195 ms	161	117	124	
		Aplicación LBBAP	AVD = 210 ms	165	124	131	
			AVD = 195 ms	162	124	135	
LBBB+HF+LVd	Aplicación LBBAP	AVD = 210 ms	175	130	147		
	Aplicación LOT-CRT	AVD = 195 ms	162	122	114		
LBBB+HF+LVd	Valor inicial	-	-	167	124	125	
	Aplicación LOT-CRT	AVD = 210 ms	167	124	125		

Respecto a los mapas de isócronas, en la Figura 4 se muestra lo que sucede en el caso de LBBB+HF+LVd al aplicar cada una de las terapias.

Como se aprecia, en el caso patológico (A) la última región en ser activada es la zona de la pared lateral del VI, dando lugar a un tiempo total de activación (TAT) que ronda los 200 ms. Sin embargo, al aplicar los distintos enfoques de CRT (B, C y D) los tiempos de activación de las distintas regiones cardíacas varían, obteniéndose en algunos casos un patrón más similar al fisiológico (C), en cualquier caso, tanto el TAT como el correspondiente t90 se reducen significativamente. En la Tabla 1 pueden encontrarse los valores de t90 medidos para cada caso y terapia evaluada.



**Figura 4.** Mapas de isócronas obtenidos en condiciones de LBBB+HF+LVd (A), y tras aplicar los distintos enfoques de CRT (B: BiVP; C: LBBAP; D: LOT-CRT).

#### 4. Discusión

En ambos casos de estudio, se han conseguido reducir significativamente los valores patológicos de partida al aplicar cada una de las distintas terapias, mostrando una eficacia de leve ( $\leq 20$  ms) a moderada (20-40 ms) tanto para un AVD de 195 como de 210 ms.

En particular, en el caso de LBBB+HF la mejor terapia resultó ser la LOT-CRT reduciendo el QRSd de 166 a 141 ms, el a90 de 123 a 99 ms y el t90 de 134 a 98 ms; lo cual concuerda con los resultados obtenidos en la práctica clínica [6].

En el caso de LBBB+HF+LVd no hay un claro vencedor, siendo la BiVP ligeramente superior a la LOT-CRT en cuanto a QRSd y a90 se refiere (161 vs 162 ms y 117 vs 122 ms respectivamente), pero obteniéndose un menor valor del t90 al aplicar la LOT-CRT (124 vs 114 ms). Dada la ausencia de estudios de LOT-CRT y LBBAP que hayan podido evidenciar la presencia de LVd no se ha podido contrastar los resultados con datos experimentales.

Por último, en cuanto a la influencia del AVD, se puede observar que en general se obtienen mejores resultados al utilizar el AVD de 195 ms. Esto puede explicarse debido a la mayor coordinación de los estímulos externos aplicados con el estímulo fisiológico que viaja por la rama sana (la derecha) que se consigue al utilizar dicho retardo.

#### 5. Conclusión

A través de este estudio, se ha demostrado claramente la capacidad de los modelos *in silico* para comprender, evaluar y optimizar diferentes enfoques de CRT en un mismo paciente, lo cual podría ayudar al electrofisiólogo a planificar el procedimiento y tomar la mejor decisión para cada paciente, reduciendo el tiempo de la intervención y mejorando los resultados.

#### Agradecimientos

Esta publicación es parte del proyecto I+D+i PID2019-104356RB-C41, financiado por MCIN/AEI/10.13039/501100011033 y por el programa de investigación e innovación Horizon 2020 de la Unión Europea bajo el acuerdo de subvención No 101016496 (SimCardioTest).

#### Referencias

- [1] A. L. Clark et al., “The prevalence and incidence of left bundle branch block in ambulant patients with chronic heart failure,” *Eur. J. Heart Fail.*, vol. 10, no. 7, pp. 696–702, Jul. 2008.
- [2] D. G. Strauss et al., “Defining Left Bundle Branch Block in the Era of Cardiac Resynchronization Therapy,” *Am. J. Cardiol.*, vol. 107, no. 6, pp. 927–934, Mar. 2011.
- [3] F. Leyva, S. Nisam, and A. Auricchio, “20 years of cardiac resynchronization therapy,” *J. Am. Coll. Cardiol.*, vol. 64, no. 10, pp. 1047–1058, 2014.
- [4] Z. Wang et al., “Cardiac resynchronization therapy in heart failure patients: tough road but clear future,” *Hear. Fail. Rev.* 2020 263, vol. 26, no. 3, pp. 735–745, Oct. 2020.
- [5] P. Vijayaraman et al., “Left Bundle Branch Area Pacing for Cardiac Resynchronization Therapy: Results From the International LBBAP Collaborative Study Group,” *JACC Clin. Electrophysiol.*, vol. 7, no. 2, pp. 135–147, Feb. 2021.
- [6] M. Jastrzębski et al., “Left bundle branch-optimized cardiac resynchronization therapy (LOT-CRT): Results from an international LBBAP collaborative study group,” *Hear. Rhythm*, vol. 19, no. 1, pp. 13–21, Jan. 2022.
- [7] M. Strocchi et al., “His-bundle and left bundle pacing with optimized atrioventricular delay achieve superior electrical synchrony over endocardial and epicardial pacing in left bundle branch block patients,” *Hear. Rhythm*, vol. 17, no. 11, pp. 1922–1929, Nov. 2020.
- [8] R. Miri, I. M. Graf, and O. Dössel, “Efficiency of timing delays and electrode positions in optimization of biventricular pacing: A simulation study,” *IEEE Trans. Biomed. Eng.*, vol. 56, no. 11, pp. 2573–2582, Nov. 2009.
- [9] E. F. Carpio et al., “Optimization of lead placement in the right ventricle during cardiac resynchronization therapy. A simulation study,” *Front. Physiol.*, vol. 10, no. 74, Feb. 2019.
- [10] R. Sebastian et al., “Characterization and modeling of the peripheral cardiac conduction system,” *IEEE Trans. Med. Imaging*, vol. 32, no. 1, pp. 45–55, Jan. 2013.
- [11] R. Coronel, R. Wilders et al., “Electrophysiological changes in heart failure and their implications for arrhythmogenesis,” *Biochim. Biophys. Acta - Mol. Basis Dis.*, vol. 1832, no. 12, pp. 2432–2441, Dec. 2013.
- [12] G. Pastore et al., “Patients with right bundle branch block and concomitant delayed left ventricular activation respond to cardiac resynchronization therapy,” *Europace*, vol. 20, no. 11, pp. E171–E178, Nov. 2018.
- [13] T. O’Hara, L. Virág, A. Varró, and Y. Rudy, “Simulation of the undiseased human cardiac ventricular action potential: Model formulation and experimental validation,” *PLoS Comput. Biol.*, vol. 7, no. 5, May. 2011.
- [14] P. Stewart et al., “Mathematical models of the electrical action potential of Purkinje fibre cells,” *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.*, vol. 367, no. 1896, pp. 2225–2255, Jun. 2009.
- [15] M. Potse et al., “A comparison of monodomain and bidomain reaction-diffusion models for action potential propagation in the human heart,” *IEEE Trans. Biomed. Eng.*, vol. 53, no. 12, pp. 2425–2435, Dec. 2006.
- [16] A. Alturki et al., “Fusion pacing in patients with right bundle branch block who undergo cardiac resynchronization therapy,” *J. Electrocardiol.*, vol. 64, pp. 66–71, Jan. 2021.
- [17] E. Heidenreich, J. Ferrero, M. Doblaré, and J. Rodríguez, “Adaptive macro finite elements for the numerical solution of monodomain equations in cardiac electrophysiology,” *Ann. Biomed. Eng.*, vol. 38, no. 7, pp. 2331–2345, Jul. 2010.
- [18] A. Ferrer et al., “Detailed Anatomical and Electrophysiological Models of Human Atria and Torso for the Simulation of Atrial Activation,” *PLoS One*, vol. 10, no. 11, Nov. 2015, doi: 10.1371/JOURNAL.PONE.0141573.
- [19] D. U. J. Keller et al., “Ranking the influence of tissue conductivities on forward-calculated eegs,” *IEEE Trans. Biomed. Eng.*, vol. 57, no. 7, pp. 1568–1576, Jul. 2010.

#### Contacto:

Beatriz Trenor.  
 Centro de Investigación e Innovación en Bioingeniería (Ci2B),  
 Universitat Politècnica de València, Camino de Vera s/n, Edificio  
 8B, 46022 Valencia, Valencia, Spain.  
[btrenor@eln.upv.es](mailto:btrenor@eln.upv.es)

# Utilización de dispositivos de realidad mixta en cirugía maxilofacial

Carlos Plaza de Miguel <sup>1</sup>, Carlos Lobato Gómez <sup>2</sup>, José C. Moreno Vázquez <sup>3</sup>,  
Juan A. Sánchez-Margallo <sup>3</sup>, Francisco M. Sánchez Margallo <sup>4</sup>

<sup>1</sup> TREMIRS, Centro de Cirugía de Mínima Invasión Jesús Usón, Cáceres, España,  
[cplaza@ccmijesususon.com](mailto:cplaza@ccmijesususon.com)

<sup>2</sup> Unidad de Bioingeniería, Centro de Cirugía de Mínima Invasión Jesús Usón, Cáceres, España,  
[clobato@ccmijesususon.com](mailto:clobato@ccmijesususon.com), [jasanchez@ccmijesususon.com](mailto:jasanchez@ccmijesususon.com)

<sup>3</sup> Neofacial, Badajoz, España, [josecarlosmoreno@neofacial.com](mailto:josecarlosmoreno@neofacial.com)

<sup>4</sup> Dirección Científica, Centro de Cirugía de Mínima Invasión Jesús Usón, Cáceres, España,  
[msanchez@ccmijesususon.com](mailto:msanchez@ccmijesususon.com)

## 1. Resumen

El proyecto presentado consiste en el desarrollo y utilización de un sistema de Realidad Mixta en implantología, que permita realizar la colocación de implantes dentales de forma precisa, en tiempo real y según una planificación computarizada previa. Para este estudio, se ha realizado un caso de uso en un entorno controlado con un cirujano experto utilizando el dispositivo HoloLens 2. Los implantes se han llevado a cabo sobre un simulador para implantes maxilofaciales en cuatro moldes diferentes: una mandíbula con dientes, una sin dientes, un maxilar con dientes y un maxilar sin dientes.

El sistema ha permitido cargar los diferentes estudios preoperatorios mediante una interfaz holográfica controlada con gestos manuales o comandos de voz. Del mismo modo, el cirujano puede visualizar los implantes de forma aislada o en conjunto, así como modificar aspectos gráficos para que los hologramas interfieran lo mínimo posible con el campo de visión quirúrgico. El sistema ha permitido un guiado holográfico sin apartar la vista de la zona operatoria, mejorando la percepción espacial del cirujano.

## 2. Introducción

Han pasado más de 15 años desde las primeras publicaciones de Cirugía Asistida por Ordenador en Implantología [1-2]. Actualmente existen dos tipos, la Cirugía Guiada Estática y la Cirugía Guiada Dinámica o de Navegación, sobre las que existe amplia evidencia científica sobre su precisión, indicaciones y limitaciones [3-6].

La aparición de exploraciones radiológicas 3D y programas informáticos que permiten virtualizar la posición 3D real del implante ha supuesto una segunda revolución en la implantología, al permitir visualizar la posición ideal del implante antes de la cirugía y planificar

la intervención de una forma más precisa [7]. Esto es lo que conocemos como planificación 3D.

El uso de las gafas de Realidad Mixta (RM) está actualmente en proceso de desarrollo y validación en el ámbito quirúrgico, con un número muy limitado de publicaciones en la literatura científica que investigan la aplicación de esta tecnología en odontología. Pellegrino et al. [8] presentan un informe clínico piloto de dos casos de rehabilitación con implantes en el área del premolar superior mediante el uso de un sistema de planificación basado en RM. Sin embargo, este sistema requiere de un método adicional de registrado basado en un tracker óptico externo, así como de un procedimiento de calibración previo a la cirugía.

La Realidad Aumentada, y concretamente la RM, en cirugía es todavía un método experimental. Su uso como alternativa a las técnicas tradicionales existentes en implantología presenta desafíos importantes que requieren soluciones técnicas innovadoras.

## 3. Materiales y Métodos

Para el correcto funcionamiento del sistema, han sido necesarios software de diferentes temática, por lo que se ha separado el contenido en las siguientes subsecciones:

### 3.1. Preprocesado del estudio preoperatorio

El estudio preoperatorio se ha exportado en diferentes archivos tridimensionales. El primer paso ha sido la conversión de los distintos archivos .STL extraídos de la tomografía axial computarizada (TAC) con la información de los implantes, la anatomía y los marcadores fiduciales en un único archivo .OBJ mediante el software libre Blender [9]. Se ha separado y jerarquizado el contenido de una forma adecuada para que el sistema sea capaz de interpretar los diferentes elementos tridimensionales.

Esto permite que posteriormente el cirujano pueda elegir visualizar cada implante de forma aislada, visualizar la

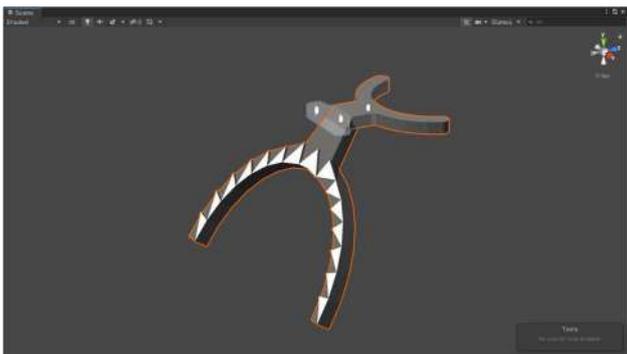
anatomía superpuesta o modificar los aspectos visuales de los elementos holográficos. Los tres marcadores fiduciales obtenidos del TAC serán los puntos de referencia para el correcto alineamiento.



**Figura 1.** Preprocesado (separación y jerarquización) de los elementos tridimensionales del estudio preoperatorio en Blender.

### 3.2. Impresión 3D del frame

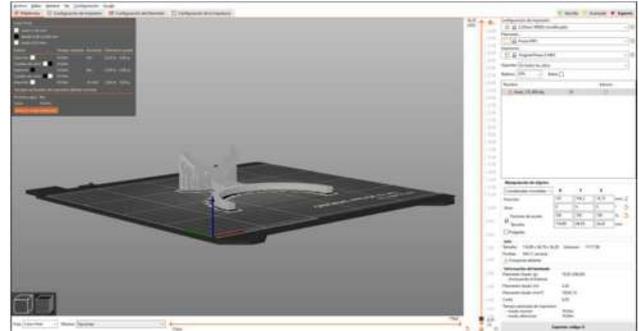
Para que el sistema desarrollado pueda detectar y alinear los implantes utilizando las librerías de visión artificial Vuforia, se ha diseñado en Blender [9] e impreso un modelo tridimensional de un arco de navegación con tres orificios en los cuales se insertaron los marcadores fiduciales de referencia (Figura 1). Estos marcadores de referencia están fabricados con un material radiopaco, visible en el estudio de TAC. Este arco impreso o frame, permitirá obtener los puntos de referencia para alinear el modelo tridimensional de la anatomía del paciente, obtenido a partir del estudio de TAC, con el frame anexo a la mandíbula o al maxilar en los cuales se desee realizar los implantes.



**Figura 2.** Diseño 3D del frame junto con los marcadores fiduciales que servirán para una correcta alineación.

Para proceder con la impresión, se exportó el modelo del frame en formato .STL (Figura 2) y se cargó en el software de impresión 3D de Prusa Slicer [10] para establecer los parámetros de la impresora i3 MK3S (Figura 3). Para llevar a cabo las adecuadas

comprobaciones iniciales, se imprimieron varios prototipos a diferentes niveles de resolución. Los modelos finales fueron impresos en mayor calidad permitida, lo cual necesitó de aproximadamente 5 horas de impresión.



**Figura 3.** Preparación de la impresión del modelo en diferentes materiales con PrusaSlicer.

Se ha optado por utilizar como material de impresión dos filamentos de PLA (ácido poliláctico) de color blanco y negro. Después de hacer varias pruebas con las imágenes obtenidas por el equipo de TAC, no se encontraron problemas en la segmentación de los marcadores fiduciales del frame, diferenciándose fácilmente el material PLA de los marcadores fiduciales de oro empleados (Figura 4).



**Figura 4.** Diferentes impresiones realizadas en PLA con la impresora Prusa i3 MK3S.

### 3.3. Marcadores fiduciales

Como marcadores fiduciales se han empleado los marcadores de oro de 0,8x3mm de la empresa RT-IDEa Fiducial Gold Markers [11]. Estos marcadores fiduciales han sido introducidos con agujas de implatación en el frame impreso y posteriormente sellados, para que, una vez se realice el TAC al paciente, se pueda extraer la geometría de dichos marcadores radiopacos (Figura 5). Esto permite utilizar esas posiciones para llevar a cabo el registrado de los contenidos holográficos de planificación quirúrgica con la anatomía del paciente y la localización de los implantes.



Figura 5. Agujas de implantación de los 3 marcadores fiduciales Gold Markers.

### 3.4. Aplicación de Realidad Mixta

La identificación del frame impreso en 3D se llevó a cabo mediante la cámara RGB de las gafas HoloLens 2. Se trata de un dispositivo comercial puntero en tecnología de realidad mixta. La segunda iteración introduce importantes mejoras respecto a la primera versión como son la detección de las manos, la inclusión de eye tracking y la extensión al doble del campo de visión en el que podrán visualizarse hologramas. Para desarrollar aplicaciones para este dispositivo, se recomienda hacer uso del motor de videojuegos Unity3D y el lenguaje de programación C# junto con el framework Mixed Reality Toolkit [12].

Se ha desarrollado un método para llevar a cabo el registrado de los modelos 3D preoperatorios de planificación quirúrgica basado en la posición de los marcadores fiduciales incluidos en el frame y las marcas visuales presentes en el arco de navegación identificadas por el dispositivo HoloLens 2. En la Figura 6 se puede ver el esquema de la arquitectura desarrollada:

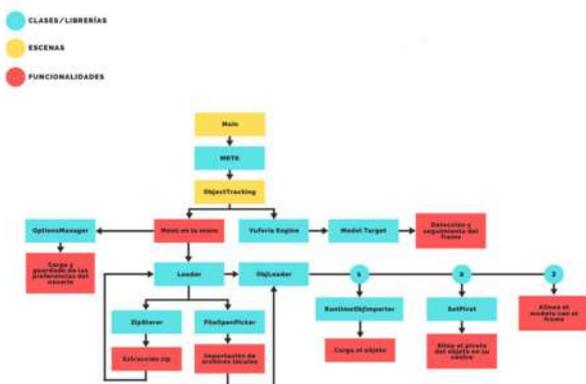


Figura 6. Esquema del sistema desarrollado en lenguaje de programación C# con Unity3D.

### 3.5. Librería Vuforia

Según la documentación aportada por la librería de visión por computador Vuforia [13], para facilitar la detección y estimación de la pose de los marcadores visuales, se recomienda introducir variaciones de color/textura dentro del modelo tridimensional a detectar. Por ello, se han utilizado dos filamentos de PLA blanco y negro, lo cual mejora la detección y estimación de la posición con Vuforia, permitiendo solapar el modelo tridimensional con la anatomía real mediante la cámara RGB de las gafas HoloLens 2.

La herramienta Model Target Generator [14], permite entrenar su detección con los algoritmos de deep learning de Vuforia y posteriormente exportar un training set que se introducirá en la aplicación de Unity3D, haciendo que el frame impreso sea detectable (Figura 7). Los algoritmos utilizados por Vuforia para la detección del modelo impreso están encapsulados, por lo que el usuario de esta herramienta simplemente podrá configurar el ángulo de visión para reconocer el modelo introducido, no pudiendo acceder al código fuente ni modificar ningún parámetro de los algoritmos de deep learning utilizados.



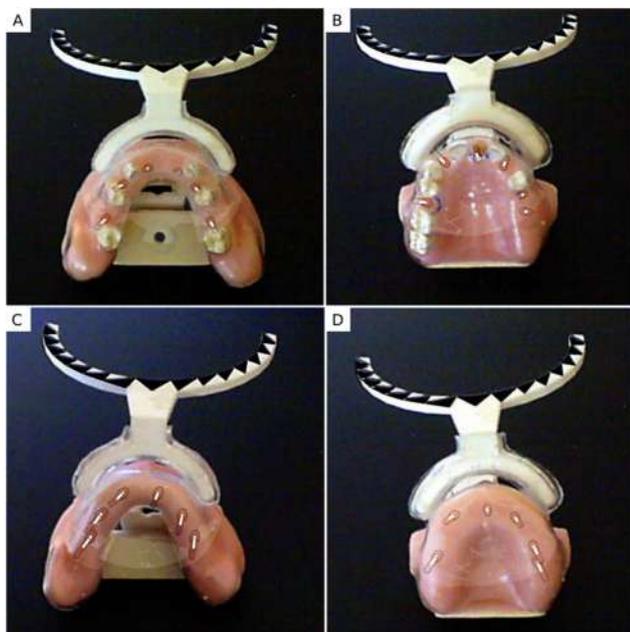
Figura 7. Herramienta Model Target Generator para entrenar la detección y estimación de la pose del frame impreso.

Para mejorar el proceso de detección del arco de navegación mediante el software de visión por computador, se determinó la combinación de colores de los elementos del arco de navegación que mejorasen su identificación. Del mismo modo, se determinó la inclinación más adecuada del arco de navegación con respecto al frame para facilitar al cirujano el abordaje quirúrgico. Como resultado, se decidió usar un diseño en blanco y negro con un ángulo de inclinación de 30 grados.

## 4. Resultados y discusión

Para este trabajo, se han cargado de forma satisfactoria cuatro estudios preoperatorios, permitiendo al cirujano visualizar los implantes sin interferir con el campo de visión. En total, se han realizado una serie de 6 implantes en 4 modelos (dos mandíbulas y dos maxilares).

Para confirmar el correcto alineamiento de los implantes, el feedback para que el cirujano pueda proceder a realizar los implantes fueron los triángulos negros del frame. Cuando los triángulos de estos hologramas están correctamente alineados y solapan posiciones con los reales, el cirujano sabrá que puede realizar los fresados (Figura 8).



**Figura 8.** Visualización de los hologramas con las guías de los implantes a realizar en dos maxilares (B-D) y dos mandíbulas (A-C).

Se ha prestado especial atención a la usabilidad del sistema, en dos aspectos: por una parte, facilitar la interacción con la interfaz y los hologramas mediante gestos manuales y comandos de voz. Por otra, minimizar el impacto visual de la información mostrada para que interfiera lo menos posible con el campo de visión del cirujano. Para esto último, se han elaborado materiales que solo muestren el contorno de los implantes y de la anatomía del estudio preoperatorio en una especie de suave parpadeo constante de blanco a transparente.

En siguientes estudios, se comparará la precisión de los implantes realizados con los planificados en los estudios preoperatorios para poder evaluar el sistema propuesto con datos objetivos. El alineamiento de los elementos ha requerido un preprocesamiento manual de los 4 TACs realizados, algo que en un futuro será conveniente automatizar en la medida de lo posible mediante Blender. También se ampliará el número de cirujanos, empleando una encuesta subjetiva para poder extraer conclusiones que hagan referencia a la posible utilización de esta herramienta en cuestiones como la ergonomía, la rapidez del sistema o la mejora de la percepción espacial.

## Agradecimientos

Este estudio ha sido parcialmente financiado por la Junta de Extremadura (España), Fondo Social Europeo, Fondo Europeo de Desarrollo Regional (FEDER) "Una forma de hacer Europa", y el Ministerio de Ciencia e Innovación de España (TA18023, IG122, CPI-2019-2033-1-TRE-14).

## Referencias

- [1] Kraut RA, ImplantDent. Utilization of 3D/dental software for precise implant site selection: clinical reports. 1:134-139.1992.
- [2] Rothman SLG, Chafetz N, Rhodes ML, Schwartz MS: CT in the preoperative assessment of the mandible and maxilla. For endosseous implant surgery. Radiology 168: 171.1988.
- [3] Sarment DP y col: Accuracy of implant placement with a stereolithographic surgical guide. Int J Oral Maxillofac Implants 18(4):571-7. 2003.
- [4] Tardieu PB y col. Computer-assisted implant placement. A case report: treatment of the mandible. Int J Oral Maxillofac Implants18(4): 599-604. 2003.
- [5] Cushen SE, Turkyilmaz I: Impact of operator experience on the accuracy of implant placement with stereolithographic surgical templates: An in vitro study. J Prosthet Dent 109:248–54, 2013.
- [6] Sarment DP, Sukovic P, Clinthorne N. Accuracy of implant placement with a stereolithographic surgical guide. Int J Oral Maxillofac Implants [Internet]. [cited2014Oct18]; 18(4):571–7. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/12939011>
- [7] Tahmaseb A, Wismeijer D, Coucke W, Derksen W. Computer technology applications in surgical implant dentistry: a systematic review. Int J Oral Maxillofac 29 Suppl (25-42)2014.
- [8] Pellegrino G, Mangano C, Mangano R, Ferri A, Taraschi V, Marchetti C. Augmented reality for dental implantology: a pilot clinical report of two cases. BMC Oral Health 2019, 19(1):158. doi:10.1186/s12903-019-0853-y
- [9] Blender. <https://www.blender.org/about>
- [10] Prusa Slicer. [https://www.prusa3d.com/es/pagina/prusaslicer\\_424/](https://www.prusa3d.com/es/pagina/prusaslicer_424/)
- [11] Gold Anchor. <https://www.goldanchormarker.com/>
- [12] Microsoft. Introducción a MRTK para Unity. <https://docs.microsoft.com/es-es/windows/mixed-reality/develop/unity/mrtk-getting-started>
- [13] Vuforia Developer Library. Best Practices for Designing and Developing Image-Based Targets. <https://library.vuforia.com/features/images/image-targets/best-practices-for-designing-and-developing-image-based-targets.html>
- [14] Vuforia Developer Library and Model Target Generator. <https://developer.vuforia.com/downloads/tool>

## Pósteres III

---

Viernes, 25 de noviembre de 2022

# Simulador clínico para el entrenamiento de la técnica de transferencia embrionaria

C. Aguilar Sabido<sup>1</sup>, D. Moreno Blanco<sup>1</sup>, O. Armijo Suárez<sup>2</sup>, F. J. Rubio<sup>3</sup>, E. J. Gómez<sup>1,4</sup>, M. Quintana<sup>3</sup>, P. Sánchez Gonzalez<sup>1,4</sup>

<sup>1</sup> Grupo de Bioingeniería y Telemedicina, ETSI Telecomunicación, Centro de Tecnología Biomédica, Universidad Politécnica de Madrid, España, {claudia.aguilar.sabido}@alumnos.upm.es {diego.morenob, p.sanchez, enriquejavier.gomez}@upm.es

<sup>2</sup> Servicio de Ginecología y Obstetricia, Hospital Universitario La Paz de Madrid, 28046 Madrid, España

<sup>3</sup> Fundación para la Investigación Biomédica del Hospital Universitario La Paz de Madrid, 28046 Madrid, España

<sup>4</sup> Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina, Madrid, España

## Resumen

*Las técnicas de reproducción asistida permiten a la población en edad fértil que padece algún tipo de anomalía congénita u adquirida que le ha provocado esterilidad lograr un embarazo. La técnica más utilizada a nivel mundial es la Fecundación in Vitro (FIV-ICSI). Una fase de gran importancia dentro de esta técnica es la transferencia embrionaria. Consiste en la deposición del embrión dentro del útero de la mujer. En el Centro Avanzado de Simulación y Entrenamiento Clínico del Instituto de Investigación del Hospital Universitario de La Paz en Madrid, se detectó, junto con el servicio de Ginecología y Obstetricia del citado hospital, la necesidad de desarrollar un simulador que permitiese el entrenamiento de la técnica de transferencia embrionaria. Para el desarrollo del simulador se ha integrado en un útero impreso en 3D los componentes electrónicos necesarios para poder sensorizarlo y recoger métricas que permiten valorar objetivamente la calidad del procedimiento realizado. Paralelamente, el simulador consta de una interfaz web donde se pueden resolver los casos propuestos y visualizar los resultados obtenidos. Finalmente se ha llevado a cabo una validación clínica que permite concluir que el simulador cumple con los objetivos planteados y permite a los alumnos adquirir habilidades que antes no podían. Por supuesto, también ha servido para detectar los puntos débiles del simulador y sentar las bases de las futuras líneas de investigación que habría que seguir para en un futuro próximo, desarrollar un simulador mucho más completo e integrado que permita la práctica transversal de la transferencia embrionaria.*

## 1. Introducción

La fertilidad es la capacidad de concebir un hijo. Los datos obtenidos por el Banco Mundial [1], muestran que la tasa de fertilidad en mujeres ha ido decayendo a lo largo de los años. Observando más en profundidad, se puede comprobar como España ha estado siempre muy por debajo de la media mundial en cuanto a Tasa de Fertilidad.

Los problemas relacionados con la fertilidad pueden clasificarse en dos categorías: infertilidad o esterilidad [2]. La infertilidad es la incapacidad de concluir un embarazo a término después de que se haya producido la fecundación. Por el contrario, la esterilidad es la incapacidad de uno de los miembros de la pareja o incluso de ambos, para establecer un embarazo. Se estima que actualmente, entre un 8 y un 12% de las parejas en edad reproductiva padecen problemas de infertilidad [3].

La reproducción asistida es el conjunto de técnicas y tratamientos que se llevan a cabo para facilitar un embarazo, sustituyendo así, el proceso natural de la reproducción [4]. Estas técnicas se utilizan cuando el paciente o los pacientes padecen de esterilidad o infertilidad, cuando el paciente padece enfermedades como la endometriosis grave o trastornos de la ovulación, cuando existen alteraciones inmunológicas con transcendencia reproductiva o cuando se ha experimentado fracasos en la inseminación.

Dentro de las TRA se encuentra la Fecundación In Vitro (FIV). La FIV requiere la extracción de óvulos y espermatozoides, y la unión de estos en el laboratorio para crear un embrión. Una vez se seleccionan los mejores embriones, éstos son transferidos al útero de la mujer en el proceso que se conoce como Transferencia Embrionaria (ET) [5]. Es una técnica compleja que requiere de una intervención quirúrgica y, por tanto, tiene un coste económico relevante. Además, tiene una mayor tasa de éxito que otras TRA, llegando a obtener resultados satisfactorios incluso en mujeres que padecen de endometriosis.

La transferencia embrionaria es uno de los pasos clave de la FIV. El proceso consiste en introducir directamente los embriones en el útero de la mujer y se divide en cuatro fases: (1) Preparación del Paciente, (2) Preparación del Catéter, (3) Transferencia de Embriones y (4) Cuidados Posteriores. Hay muchos factores relacionados con el desarrollo de la técnica de la ET que pueden disminuir la posibilidad de embarazo siendo uno de los más frecuentes el contacto con las paredes uterinas o bien con el fondo uterino. El contacto con el fondo uterino es la principal causa que provoca las contracciones uterinas. Estas contracciones facilitan la expulsión del embrión, disminuyendo así la posibilidad de embarazo.

El objetivo de este trabajo es realizar un simulador de transferencia embrionaria que permita entrenar esta técnica a los residentes y estudiantes de medicina, así como a médicos ya formados que deseen practicar esta técnica. Para ello se utilizará un maniquí físico sensorizado que permita evaluar objetivamente la realización de la técnica, proporcionando una retroalimentación a los estudiantes.

## 2. Diseño del sistema

El simulador desarrollado en este trabajo ha sido codiseñado mediante entrevistas entre los profesionales del Grupo de Bioingeniería y Telemedicina de la Universidad Politécnica de Madrid, los profesionales del Centro Avanzado de Simulación y Entrenamiento Clínico (CEASEC) del Instituto de Investigación del Hospital Universitario de La Paz en Madrid y el servicio de Ginecología y Obstetricia del citado hospital.

El sistema diseñado tiene la arquitectura que se puede observar en la Figura 1. Como se observa en el diagrama, el sistema está formado por un maniquí físico sensorizado, que mediante un microcontrolador establece una conexión UDP con un servidor. Dicho servidor se encarga de almacenar las métricas y de asignarlas a los distintos usuarios. El sistema también consta de una aplicación web que permite la visualización y análisis de las métricas, así como la gestión de usuarios.

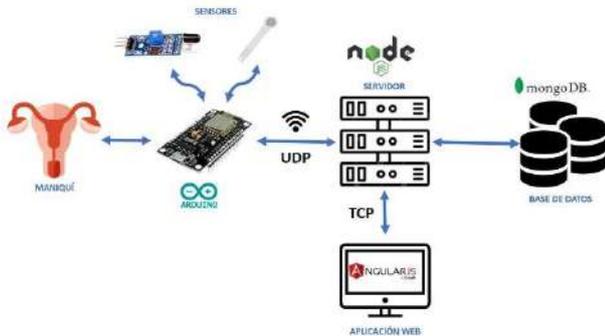


Figura 1. Diagrama de arquitectura.

Dentro del sistema se han definido dos roles de usuario:

- El rol de docente, que se encargará de supervisar, tutorizar y personalizar los casos de los alumnos.
- El rol de alumno, que resolverá los casos y al que se evaluará en su desempeño permitiendo observar y supervisar su aprendizaje.

Cada uno de los diferentes roles tendrá asignados los casos de uso que se observan en la Figura 2.

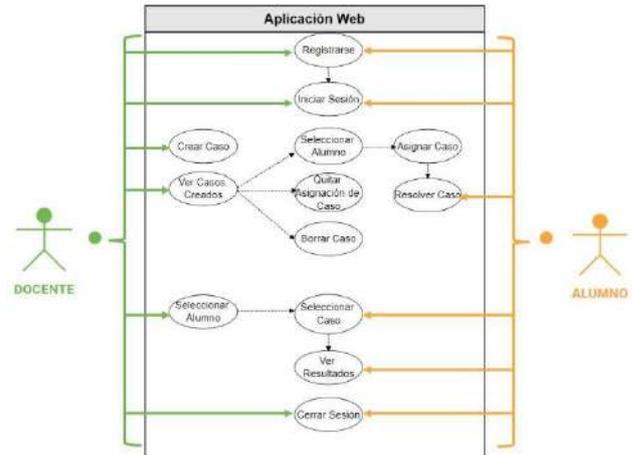


Figura 2. Diagrama de casos de uso.

Además, se han recogido los diferentes requisitos funcionales y no funcionales resultantes del proceso de codiseño. El flujo resultante para el sistema se puede observar en el diagrama representado en la Figura 3.

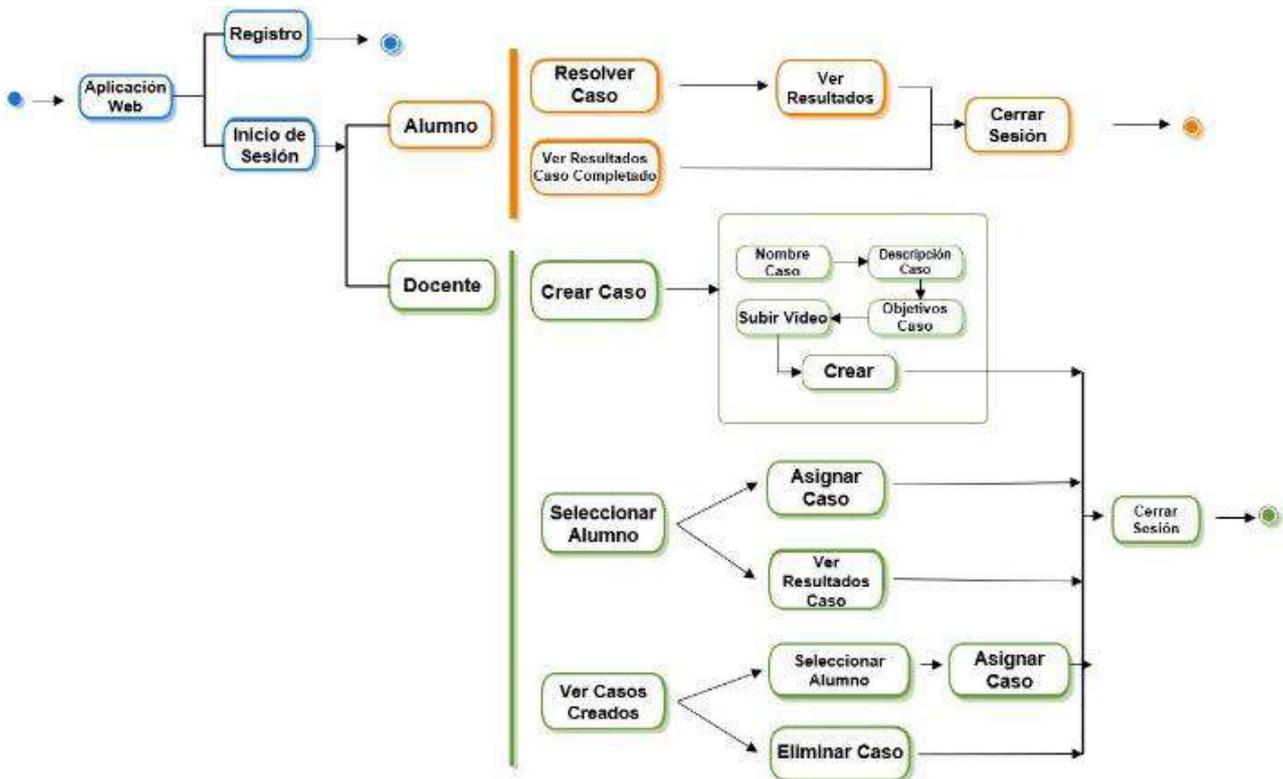


Figura 3. Diagrama de flujo.

### 2.1. Definición de las métricas de evaluación

Las métricas de evaluación del proceso de transferencia embrionaria para determinar la calidad de la realización de éste han sido definidas por el servicio de Ginecología y Obstetricia del Hospital Universitario La Paz de Madrid. Así pues, los profesionales de dicho servicio han determinado que los hechos más relevantes que pueden ocurrir durante este procedimiento son: (1) Contacto con el fondo del útero. Se considera un fallo eliminatorio. (2) Contacto con el canal cervicouterino. Se considera un fallo leve. (3) Profundidad a la que se deposita el embrión. En caso de depositarse antes de la mitad del útero se considera un fallo grave. (4) Cambio de presión abrupto en el émbolo del catéter. Se considera un fallo grave. (5) Velocidad de entrada y salida del catéter. En caso de detectar una velocidad excesiva se considera un fallo grave. (6) Duración total del procedimiento. El procedimiento no debe tener una duración excesiva para evitar el estrés a la paciente. En caso de detectarse una duración excesiva se considera un fallo leve.

Además de estos aspectos, también es importante que el estudiante interiorice la rutina clínica de preparación de la paciente y del material a utilizar. Por ello se incluye una lista de acciones que el usuario tendrá que realizar y marcar.

### 2.2. Red de sensores y circuito electrónico

Para monitorizar el procedimiento de ET al completo se han incluido diferentes sensores en el sistema. Toda esta red de sensores está conectada a un microcontrolador que envía vía Wifi mediante protocolo UDP las diferentes métricas al resto de componentes del sistema. Para la red de sensores se han utilizado 7 sensores de infrarrojos para detectar la posición del catéter en cada momento, un sensor de presión para medir la presión ejercida en el embolo del catéter y pintura conductora para detectar los contactos entre la punta del catéter (que se ha modificado mediante un hilo conductor) con el canal cervicouterino y el fondo del útero de forma diferenciada. El circuito electrónico para la conexión de toda la red al microcontrolador puede observarse en la Figura 4.

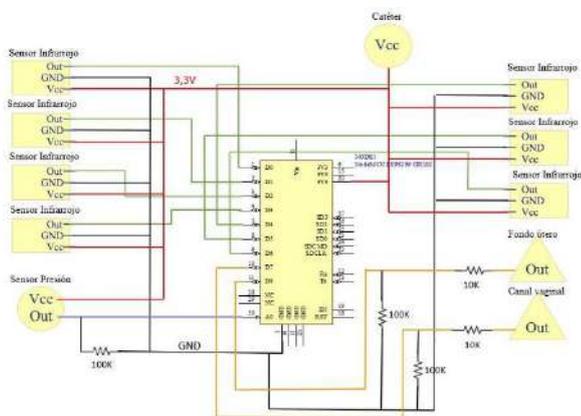


Figura 4. Circuito diseñado para la red de sensores.

### 2.3. Maniquí físico

Para la realización del presente simulador se ha contado con un maniquí que simula las extremidades inferiores de

una mujer proporcionado por el CEASEC. Para el correcto desempeño de la simulación es necesario colocar el maniquí en posición de litotomía, tal y como se hace en un caso real. También se ha utilizado un modelo 3D del útero donde se integra la red de sensores.

### 2.4. Aplicación web

Además de todo lo anterior se ha realizado un mockup con el diseño inicial de la aplicación web. Este mockup se ha validado con el personal clínico del servicio de Obstetricia y Ginecología del Hospital Universitario La Paz de Madrid. Esta aplicación web será la herramienta mediante la cual los profesionales consulten los resultados de los estudiantes y la cual utilizarán los estudiantes para resolver los casos, incluyendo el visionado de videos ecográficos en tiempo real que simulen el caso que están resolviendo.

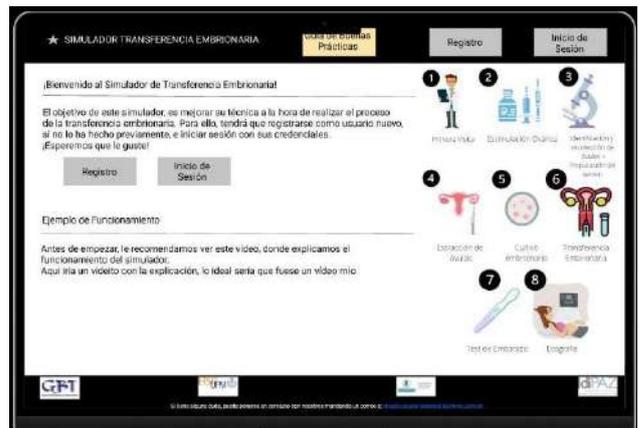


Figura 5. Interfaz de inicio incluida en el mockup.

### 2.5. Validación

Con el objetivo de poder realizar una validación clínica con el personal médico del HULP, se ha propuesto realizar una validación del simulador con el cuestionario *System Usability Scale* (SUS) para evaluar la usabilidad del prototipo desarrollado. En la validación se ha tenido en cuenta tanto la validación del útero como la validación de la aplicación web. Además, se han añadido otras preguntas donde se evaluar por separado la usabilidad del maniquí y de la aplicación web. Finalmente, también se han añadido dos preguntas relacionadas con la experiencia del usuario en la realización de la técnica de la ET y en el uso de simuladores previamente.

## 3. Resultados

En la Figura 6 se puede observar el modelo 3D de útero impreso. Es un modelo que consta de dos partes, lo que facilita la integración de la red de sensores y su manipulación.

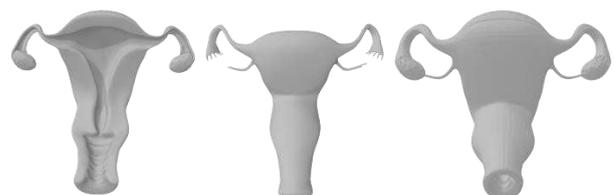


Figura 6. Modelo 3D de útero impreso en 3D.

Una vez integrada la red de sensores, se ha introducido el sistema en el maniquí, como puede observarse en la Figura 7. Esto dota al sistema de un mayor realismo y acerca el simulador a la práctica clínica, ayudando a los estudiantes a acercar esta simulación a su próxima realidad clínica. El catéter está unido a este sistema ya que debe estar conectado al microcontrolador utilizado.

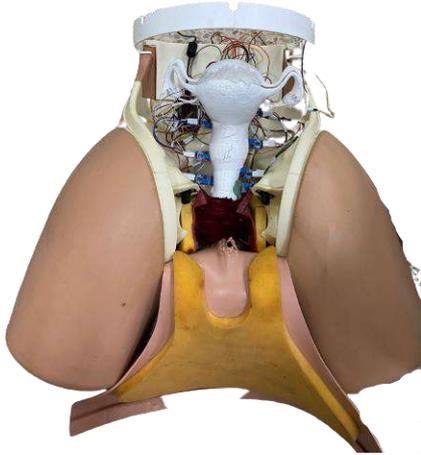


Figura 7. Vista final del sistema.

En cuanto a la aplicación web, se muestra una de las pantallas más relevantes del sistema. La pantalla de resolución de caso (Figura 8), donde el estudiante puede observar un video de ultrasonidos de un caso real que se actualiza en función de la posición del catéter en el maniquí, y una lista de tareas que debe completar durante el proceso de transferencia embrionaria.

#### 4. Discusión y conclusiones

Esta primera versión del simulador está siendo validada con el personal del servicio de Obstetricia y Ginecología del Hospital Universitario de La Paz. En los resultados preliminares obtenidos se observa que los profesionales tienen un alto grado de satisfacción con el prototipo desarrollado.

Sin embargo, también hemos encontrado limitaciones que serán tenidas en cuenta en las futuras versiones y mejoras

del presente simulador. En primer lugar, el tamaño anatómico es algo más grande que el real. Esto nos ha permitido la sensorización mediante sensores comerciales de bajo coste, pero hace que el procedimiento pierda algo de realismo. Además, el uso de sensores infrarrojos para detectar el posicionamiento del catéter en el útero limita las posiciones del vídeo y hace que este pierda fluidez y realismo.

En futuros trabajos de esta línea de investigación se va a trabajar en mejorar el simulador para eliminar las limitaciones anteriores. Además, se van a explorar alternativas a los sensores infrarrojos como el seguimiento mediante técnicas de análisis de imagen o el uso de otro tipo de sensores.

#### Referencias

- [1] Banco Mundial, “Tasa de fertilidad, total (nacimientos por cada mujer),” *Tasa de fertilidad*, 2016. <http://datos.bancomundial.org/indicador/SP.DYN.TFRT.IN/countries/1W-Z4?display=default> (accessed Aug. 28, 2022).
- [2] M. Vander Borgh and C. Wyns, “Fertility and infertility: Definition and epidemiology,” *Clinical Biochemistry*, vol. 62, pp. 2–10, Dec. 2018, doi: 10.1016/J.CLINBIOCHEM.2018.03.012.
- [3] S. Brugo-olmedo, C. Chillik, and S. Kopelman, “Infertility: Causes and Definitions,” *Revista Colombiana de Obstetricia y Ginecología*, vol. 54, no. 1055, pp. 228–248, 2003.
- [4] J. Y. J. Huang and Z. Rosenwaks, “Assisted reproductive techniques,” *Methods in Molecular Biology*, vol. 1154, pp. 171–231, 2014, doi: 10.1007/978-1-4939-0659-8\_8/COVER.
- [5] A. A. Brauer and G. Schattman, “Embryo transfer,” *Methods in Molecular Biology*, vol. 1154, pp. 541–548, 2014, doi: 10.1007/978-1-4939-0659-8\_25.

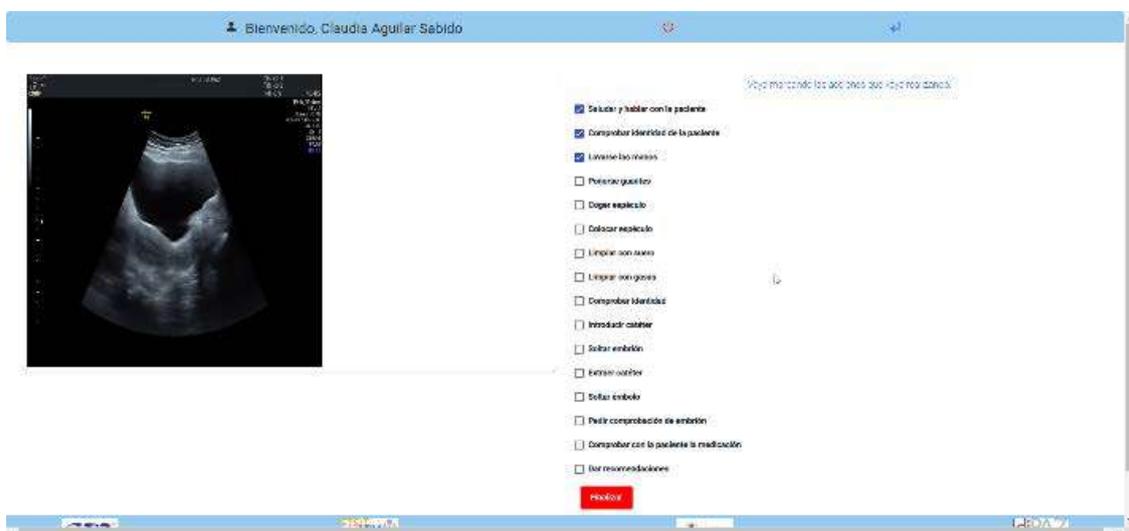


Figura 8. Pantalla de realización de caso de la aplicación web.

# Algoritmo y herramienta para la segmentación y volumetría esplénica sobre imágenes tomográficas

J. De Benavides<sup>1</sup>, E. Milara<sup>1</sup>, R. Alonso<sup>2,3</sup>, A.P. Seiffert<sup>1</sup>, A. Gómez-Grande<sup>3,4</sup>, E.J. Gómez<sup>1,5</sup>, P. Sánchez-González<sup>1,5</sup>

<sup>1</sup> Grupo de Bioingeniería y Telemedicina, ETSI Telecomunicación, Centro de Tecnología Biomédica, Universidad Politécnica de Madrid, Madrid, España; j.debenavides@alumnos.upm.es, {eva.milara.hernando, ap.seiffert, enriquejavier.gomez, p.sanchez}@upm.es

<sup>2</sup> Servicio de Hematología y Hemoterapia e Instituto de Investigación Sanitaria (i+12), Hospital Universitario 12 de Octubre, CNIO, CIBERONIC, Madrid, España; rafaelalberto.alonso@salud.madrid.org

<sup>3</sup> Facultad de Medicina, Universidad Complutense de Madrid, Madrid, España

<sup>4</sup> Servicio de Medicina Nuclear, Hospital Universitario 12 de Octubre, Madrid, España; adolfo.gomez@salud.madrid.org

<sup>5</sup> Centro de Investigación Biomédica en Red de Bioingeniería, Biomateriales y Nanomedicina, Instituto de Salud Carlos III, Madrid, España

## Resumen

*El volumen del bazo se encuentra entre 120 y 480 cm<sup>3</sup>. Sin embargo, existen numerosas patologías caracterizadas por el agrandamiento del bazo, también conocido como esplenomegalia. En la práctica clínica actual, se utiliza el índice esplénico como gold standard para la medida de esta volumetría, suponiendo una métrica imprecisa por su cálculo a partir de tres medidas unidimensionales tomadas manualmente. En este trabajo, se propone un algoritmo semiautomático de segmentación y cálculo de la volumetría esplénica para imágenes de tomografía computarizada abdominal. Por un lado, se estudia la bondad de las segmentaciones realizadas, que además son validadas por una radióloga experta, y se discute sobre la segmentación en presencia de lesiones esplénicas. Por otro lado, se estudia la correspondencia de los resultados con el diagnóstico recibido por el especialista, alcanzando un 70% de buenas segmentaciones con sensibilidad y especificidad de 87% y 80% respectivamente. De esta manera se trata de una herramienta útil para los clínicos.*

## 1. Introducción

El bazo es una de las regiones más grandes del sistema linfático. El peso medio de un bazo en estado normal es de unos 150 gramos y su volumen oscila entre 120 – 480 cm<sup>3</sup>, con una media de 250 cm<sup>3</sup>, independientemente del sexo o del peso [1]. Dentro de su participación en el sistema linfático, es un órgano linfoide secundario que juega un papel destacado en el desarrollo de la inmunidad humoral y celular. Además, tiene un rol esencial en el sistema circulatorio con funciones como filtrado sanguíneo de partículas extrañas, fagocitosis de microorganismos circulantes y eliminación de células sanguíneas envejecidas o anómalas [2]. Además, el bazo tiene la capacidad de satisfacer ciertas necesidades fisiológicas de las que no es responsable como, por ejemplo, la hematopoyesis extramedular cuando es deficiente en la médula ósea. Sin embargo, como esta función no es fisiológica del bazo, el metabolismo esplénico aumenta agrandando su tamaño, condición conocida como esplenomegalia [3]. Esta alteración del bazo aparece como síntoma de múltiples enfermedades hematológicas,

hepáticas, autoinmunes e infecciosas, así como en neoplasias malignas [4]. La esplenomegalia se evalúa de dos formas distintas: mediante palpación y mediante imágenes médicas. Sin embargo, es la imagen médica la técnica preferida para esta valoración, ya que permite tomar medidas volumétricas más precisas. Las técnicas más utilizadas son la imagen por ultrasonidos, la resonancia magnética y, especialmente, la tomografía computarizada (TC) [5].

Desde 1983 [6], el índice esplénico se sitúa como el *gold standard* de la medida volumétrica del bazo, alcanzando un 99,1% de correlación con el volumen real. En las últimas décadas, otros autores [7]–[10] han definido diversos marcadores correlacionados con el volumen esplénico basados en medidas unidimensionales del bazo e, incluso, modelos de segmentación basados en Aprendizaje Profundo [11]. Sin embargo, los resultados obtenidos no superan a los del índice esplénico. La definición de esta métrica viene dada como el producto de las medidas esplénicas en las tres dimensiones del espacio: el largo, distancia entre el primer y último corte en el que aparece el bazo; el ancho máximo del órgano entre esos cortes, y el grosor a media altura del corte de ancho máximo. Al ser estas medidas tomadas manualmente y, por lo tanto, dependientes del experto clínico que lo mida, junto a la forma irregular del bazo, el índice esplénico resulta ser una métrica imprecisa.

El principal objetivo de este trabajo es desarrollar un algoritmo de segmentación y cálculo semiautomáticos del volumen esplénico con el objetivo de detectar esplenomegalia en imágenes de TC abdominal. Además, se implementa el algoritmo desarrollado en una herramienta de ayuda a la decisión clínica para ser usada en la práctica clínica.

## 2. Metodología

Para el desarrollo de la metodología, se realiza un estudio observacional analítico a partir de dos cohortes: control (10 sujetos) y pacientes diagnosticados con esplenomegalia

causada por mielofibrosis (18 sujetos), con un total de 20 y 38 imágenes de TC abdominal respectivamente. La diferenciación entre ambas cohortes se realiza por la revisión de los informes clínicos. Con el objetivo de desarrollar una herramienta aplicable a cualquier imagen de TC abdominal, la base de datos cuenta con imágenes de diferentes protocolos de adquisición (sin contraste de fase, con contraste en fase arterial y en fase venosa) y diferentes escáneres de TC. Todas las imágenes han sido recogidas en el Servicio de Hematología y Hemoterapia del Hospital Universitario 12 de Octubre de Madrid, España.

Tanto para la implementación del algoritmo de segmentación como para la implementación de la herramienta, se utiliza el entorno de lenguaje de desarrollo *Python 3.9* (Python Software Foundation, EEUU).

### 2.1. Algoritmo de segmentación

En primer lugar, las imágenes TC se importan en formato DICOM y son convertidas a Unidades Hounsfield (UH) siguiendo la Ecuación 1:

$$UH = I_{\text{pixel}} \times \text{RescaleSlope} + \text{RescaleIntercept}$$

**Ecuación 1.** Transformación de niveles de gris a UH.

Siendo  $I_{\text{pixel}}$  la intensidad de píxel en nivel de gris, y *RescaleSlope* y *RescaleIntercept* las variables de pendiente y coordenada en el origen de la transformación establecidas por el fabricante del escáner TC.

Posteriormente, se selecciona un corte central del bazo y, en ese corte, una semilla localizada en el interior del bazo. Tomando la semilla como centro, se crea un entorno con forma de cubo con 11 vóxeles de arista, y se calcula el valor medio de UH del mismo como umbral. Todos los vóxeles con UH de valores 40 UH por encima o por debajo de este umbral son convertidos a -1000 UH para diferenciar el bazo del resto de estructuras mostradas en el corte.

Partiendo de la imagen umbralizada y la semilla seleccionada, se realiza un proceso de crecimiento de regiones. En primer lugar, se generan dos volúmenes vacíos del mismo tamaño que el volumen original. Uno de ellos, el 'Volumen Segmentado', se usa para ir almacenando el resultado del crecimiento de regiones y, cuando finalice el algoritmo, será la segmentación final. El otro, el 'Volumen Comprobado', se usa para asegurar que el algoritmo de crecimiento de regiones ha recorrido, al menos, todo el volumen correspondiente al bazo. A su vez, se crea una lista, la 'Lista de Vóxeles a Comprobar', que establece el orden en el que se van comprobando los vóxeles del volumen original en busca de vóxeles pertenecientes al bazo. Cuando esta lista queda vacía, termina el algoritmo.

Para comenzar el algoritmo de crecimiento de regiones se incluye la semilla como parte del 'Volumen Segmentado' y como parte del 'Volumen Comprobado'. Además, se incluye su entorno, los seis vóxeles colindantes en conectividad 6 3D, a la 'Lista de Vóxeles a Comprobar' siguiendo el orden: izquierda, derecha, delante, detrás, corte inferior y corte superior. Cuando el último vóxel añadido a la lista no está incluido en el 'Volumen Comprobado' se procede a incluirlo en el mismo y

comprobar si cumple el criterio de comparación para ser incluido en el 'Volumen Segmentado'. El criterio de comparación consiste en el cálculo del valor medio de las UH de todos los vecinos inmediatamente colindantes en conectividad 26 3D. Si el valor medio de este entorno es mayor o igual que el valor tomado para umbralizar menos 20 UH, el vóxel es también añadido al 'Volumen Segmentado'. Consecuencia inmediata de añadir un vóxel al 'Volumen Segmentado' es que se procede a añadir su entorno a la 'Lista de Vóxeles a Comprobar'. El algoritmo se repite hasta que no queden vóxeles en dicha lista.

Para optimizar el proceso de crecimiento de regiones, se implementa la detección de la frontera del órgano cuando un vóxel se añade al 'Volumen Segmentado' y su vóxel colindante en conectividad 6 3D no es incluido. Para estos casos se establece una condición de direccionalidad que establece que el resto de la columna o fila de vóxeles en esa dirección es incluida en 'Volumen Comprobado' directamente y desestimada de 'Volumen Segmentado'. Teniendo en cuenta la curvatura convexa del bazo, algunos píxeles se podrían perder con esta condición de direccionalidad. Por ello, se realiza el algoritmo de crecimiento de regiones tres veces, cambiando el orden en el que se incluyen los píxeles del entorno en la 'Lista de Vóxeles a Comprobar'. De esta forma, se obtienen tres máscaras de 'Volumen Segmentado', uno por cada dirección, considerando que los vóxeles que estén incluidos en cualquiera de los tres pertenecen al 'Volumen Segmentado' final. Para recuperar los bordes del bazo subestimados durante el proceso de umbralización, se realiza una doble dilatación de la máscara final del 'Volumen Segmentado'.

### 2.2. Cálculo del volumen esplénico

Una vez se obtiene la máscara del bazo, se calcula el volumen del mismo mediante la Ecuación 2:

$$V (\text{cm}^3) = \frac{PS[0] \times PS[1] \times ST/2 \times N}{1000}$$

**Ecuación 2.** Cálculo del volumen esplénico  $V$  en  $\text{cm}^3$ .

Siendo  $V$  el volumen segmentado medido en  $\text{cm}^3$ ;  $PS$  (*PixelSpacing*) el tamaño de los píxeles en las direcciones 0 y 1 del espacio, es decir, ancho y largo;  $ST$  (*SliceThickness*) el grosor del corte; y  $N$  el número de vóxeles incluidos en el 'Volumen Segmentado' final.

### 2.3. Diseño de la herramienta

Todo el algoritmo explicado con anterioridad está implementado en Python y se ejecuta a través del uso de una herramienta llamada Splenic. Esta aplicación incluye un paso inicial de validación de credenciales del usuario para aumentar la seguridad del tratamiento de los datos. Una vez el usuario entra, le permite seleccionar la carpeta con las imágenes que quiere segmentar. Estas imágenes son convertidas a UH y se muestra un corte axial, permitiendo al usuario que cambie el corte visualizado. Una vez seleccionado el corte, el usuario marca la semilla dentro del bazo y comienza el proceso de segmentación. Cuando el proceso de segmentación finaliza, la interfaz muestra el volumen original en UH con la máscara

segmentada solapada en color verde, permitiendo al usuario moverse por los diferentes cortes. Junto al volumen mostrado, se indica el volumen en cm<sup>3</sup> calculado a partir de la máscara y la etiqueta asignada a esa segmentación, siendo ‘BAZO NORMAL’ si el volumen es menor de 480 cm<sup>3</sup> o ‘BAZO ESPLENOMEGÁLICO’ en caso contrario.

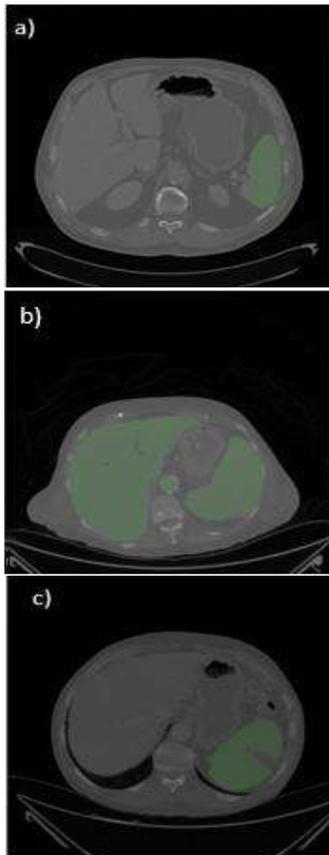
### 2.4. Validación de los resultados

Se evalúan los resultados diferenciando las segmentaciones obtenidas entre ‘Buena segmentación’ cuando visualmente el volumen segmentado es adecuado, ‘Mala segmentación’ cuando el volumen segmentado incluye otras estructuras externas al bazo, y ‘Lesiones’ cuando el bazo no es segmentado por completo debido a la presencia de lesiones esplénicas hipodensas. Además, se validan mediante la valoración visual de una radióloga experta, utilizando una escala del 0 al 5 para indicar como de buena es la segmentación del bazo, siendo 5 el valor seleccionado para una mejor segmentación. Finalmente, partiendo de la diferenciación entre cohorte control y pacientes diagnosticados con esplenomegalia, se calcula la sensibilidad y especificidad del algoritmo para diferenciar entre ambos grupos.

## 3. Resultados y discusión

### 3.1. Algoritmo de segmentación

En la Figura 1, se muestra el resultado de la segmentación de tres casos distintos.



**Figura 1.** Resultados de la segmentación del volumen esplénico sobre el volumen original. a) Buena segmentación, b) mala segmentación y c) bazo con lesiones esplénicas.

El primero de ellos (a) corresponde a una imagen de un bazo sano cuya segmentación final es ‘Buena segmentación’. El segundo (b) corresponde a una imagen en la que el bazo está considerablemente aumentado y, en algunos cortes, se encuentra en contacto con el hígado. Para estos casos la segmentación final es considerada ‘Mala segmentación’ debido a que se incluyen otras regiones no pertenecientes al bazo. Finalmente, la última imagen (c) se corresponde con una segmentación etiquetada como ‘Lesiones’. En la Tabla 1, se muestran los resultados de las segmentaciones obtenidas medidas en porcentaje de imágenes clasificadas en cada resultado de segmentación entre el total de imágenes de control, de pacientes diagnosticados con esplenomegalia y el total de ambas categorías. Se puede observar como para los pacientes de control se alcanza un 85% de ‘Buena segmentación’, mientras que para esplenomegálicos es más complicado alcanzar buenos resultados.

	Buena	Lesiones	Mala
Control	85%	0%	15%
Esplenomegálicos	60,5%	7,9%	31,6%
Total	69%	5%	26%

**Tabla 1.** Resultados del algoritmo de segmentación.

Teniendo en cuenta que, para los casos de la cohorte en los que hígado y bazo están en contacto, la presencia de esplenomegalia es evidente mediante valoración visual y que los casos clasificados como ‘Lesiones’ obtienen una segmentación incompleta del bazo, se calculan los resultados eliminando las imágenes de hígado y bazo en contacto y se reclasifican como ‘Mala segmentación’ todas aquellas imágenes incluidas en ‘Lesiones’ (Tabla 2). El resultado final obtiene un 81,3% de buenas segmentaciones, similar al 82.1% obtenido con un modelo basado en aprendizaje profundo [11]. A pesar de que el porcentaje de clasificación es inferior, el algoritmo desarrollado en el presente trabajo no requiere partir de una amplia base de datos ni las segmentaciones manuales para su prueba. De esta manera, se llega a una solución similar a la del modelo de aprendizaje profundo consumiendo menos recursos de cómputo y tiempo de segmentación manual.

	Buena	Mala
Control	85%	15%
Esplenomegálicos	79,3%	20,6%
Total	81,3%	18,7%

**Tabla 2.** Resultados del algoritmo de segmentación con reclasificación de imágenes.

En cuanto a la validación realizada por la radióloga experta, se obtiene una valoración de 3,6 para los pacientes control y 2,47 para los esplenomegálicos. Eliminando las imágenes en las que bazo e hígado están en contacto, la valoración de los esplenomegálicos sube a 3,24.

### 3.2. Cálculo del volumen esplénico

En cuanto a la evaluación de la medida volumétrica y su diferenciación entre casos esplenomegálicos (EM+) o

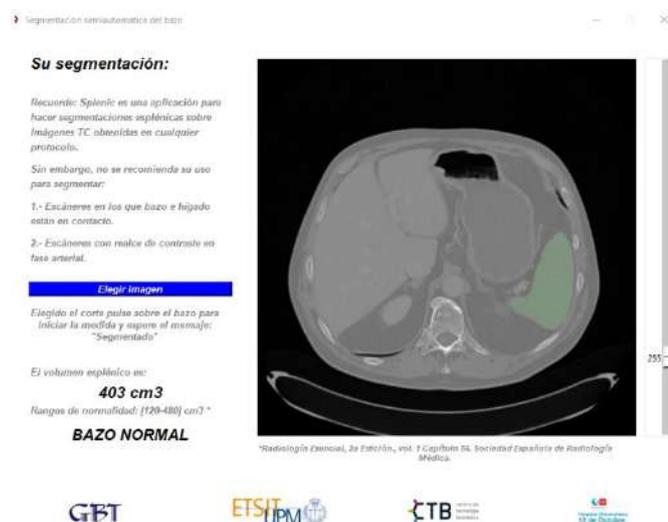
control (EM-), se obtienen los resultados mostrados en la Tabla 3, alcanzando unos valores de Sensibilidad y Especificidad del 87% y 80% respectivamente, inferiores a los del índice esplénico [6]. Sin embargo, considerando que el cálculo de la volumetría incluye el bazo al completo, y no solo tres índices unidimensionales, se espera que en los casos de ‘Buena segmentación’, la medida volumétrica sea más precisa que el índice esplénico.

		Valor real	
		EM+	EM-
Valor predicho	EM+	33	4
	EM-	5	16

**Tabla 3.** Matriz de confusión de la medida volumétrica diferenciando entre control (-) y esplenomegálicos (+).

### 3.3. Diseño de la herramienta

Todos los resultados son adquiridos mediante la segmentación de la base de datos de imágenes de TC en la herramienta diseñada. En la Figura 2, se muestra la interfaz gráfica de la herramienta diseñada con el ejemplo de una imagen de la cual se obtiene una ‘Buena segmentación’ del bazo. En la izquierda de la pantalla se muestran las instrucciones, el botón para abrir el explorador de archivos y seleccionar la carpeta donde se encuentra el volumen, y los resultados de la medida de la volumetría y su clasificación entre normal o patológico. En la derecha se muestra la segmentación del bazo sobre la imagen original.



**Figura 2.** Interfaz gráfica de Splenic.

La metodología desarrollada en el presente trabajo permite realizar una medida de la volumetría esplénica semiautomática ya que solo hay que seleccionar la semilla, reduciendo la arbitrariedad y eliminando el trabajo manual de obtener las medidas para calcular el índice esplénico. Además, es evaluada positivamente por la experta en radiología.

## 4. Conclusiones

En este trabajo, se propone un algoritmo de segmentación y cálculo de la volumetría esplénica a partir de imágenes de TC abdominal. La metodología desarrollada se implementa en una herramienta para la ayuda a la decisión

clínica en la detección de esplenomegalia como síntoma de numerosas patologías, con una medida más precisa al obtener la volumetría completa del bazo.

## Agradecimientos

Este trabajo está parcialmente financiado por una beca predoctoral (ayuda del Programa Propio de I+D+i 2020) de la Universidad Politécnica de Madrid.

## Referencias

- [1] J. L. del Cura, S. Pedraza, À. Gayete, and À. Roira, *Radiología Esencial*, 2ª. Sociedad Española de Radiología Médica, 2018.
- [2] A. P. N. A. De Porto, A. J. J. Lammers, R. J. Bennink, I. J. M. Ten Berge, P. Speelman, and J. B. L. Hoekstra, “Assessment of splenic function,” *Eur. J. Clin. Microbiol. Infect. Dis.*, vol. 29, pp. 1465–1473, 2010, doi: 10.1007/s10096-010-1049-1.
- [3] S. Slot *et al.*, “The value of bone marrow, liver, and spleen imaging in diagnosis, prognostication, and follow-up monitoring of myeloproliferative neoplasms: a systematic review,” *Cancer Imaging*, vol. 21, no. 36, 2021, doi: 10.1186/s40644-021-00405-7.
- [4] Y. Inoue, A. Nakajima, S. Mizukami, and H. Hata, “Effect of Breath Holding on Spleen Volume Measured by Magnetic Resonance Imaging,” *PLoS One*, vol. 8, no. 6, p. 68670, 2013, doi: 10.1371/journal.pone.0068670.
- [5] G. Treece, R. Prager, A. Gee, and L. Berman, “3D ultrasound measurement of large organ volume,” *Med. Image Anal.*, vol. 5, no. 1, pp. 41–54, 2001, doi: 10.1016/S1361-8415(00)00034-7.
- [6] L. Cools, M. Osteaux, L. Divano, and L. Jeanmart, “Prediction of splenic volume by a simple CT measurement: A statistical study,” *Journal of Computer Assisted Tomography*, vol. 7, no. 3, pp. 426–430, 1983, doi: 10.1097/00004728-198306000-00007.
- [7] A. S. Bezerra, G. D’ippolito, S. Faintuch, J. Szejnfeld, and M. Ahmed, “Determination of Splenomegaly by CT: Is There a Place for a Single Measurement?,” *Am. J. Roentgenol.*, vol. 184, no. 5, pp. 1510–1513, 2005, doi: 10.2214/ajr.184.5.01841510.
- [8] I. Kucybała, S. Ciuk, and J. Teńczar, “Spleen enlargement assessment using computed tomography: which coefficient correlates the strongest with the real volume of the spleen?,” *Abdom. Radiol.*, vol. 43, pp. 2455–2461, 2018, doi: 10.1007/s00261-018-1500-9.
- [9] V. Indiran, V. Singh, T. Ramachandra Prasad, and P. Maduraimuthu, “Does coronal oblique length of spleen on CT reflect splenic index?,” *Abdom. Radiol.*, vol. 42, pp. 1444–1448, 2017, doi: 10.1007/s00261-017-1055-1.
- [10] S. O. Gunes and Y. Akturk, “Determination of splenomegaly by coronal oblique length on CT,” *Jpn. J. Radiol.*, vol. 36, pp. 142–150, 2018, doi: 10.1007/s11604-017-0704-1.
- [11] L. C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, “Encoder-decoder with atrous separable convolution for semantic image segmentation,” in *European Conference on Computer Vision*, 2018, pp. 801–818. doi: 10.1007/978-3-030-01234-2\_49.

# Discriminación Histológica mediante Propiedades Ópticas de Fase de Muestras Tisulares de Colon

J. L. Ganoza-Quintana, F. Fanjul-Vélez, J. L. Arce-Diego

Grupo de Ingeniería Biomédica, Departamento TEISA, Universidad de Cantabria, Av. de los Castros 46, 39005 Santander (Cantabria), España

[fanjul@unican.es](mailto:fanjul@unican.es), [arcedj@unican.es](mailto:arcedj@unican.es)

## Resumen

*La biopsia convencional es el principal estándar del diagnóstico histológico. La histología digital incluye automatización y diagnóstico mejorado, digitaliza imágenes microscópicas de muestras histológicas y analiza parámetros similares. El presente enfoque propone el uso novedoso del contraste de fase en la histología digital clínica para mejorar el diagnóstico. El uso de cortes de tejido fresco sin etiquetas evita los dispositivos de procesamiento y reduce el tiempo de procesamiento. Los parámetros de contraste de fase que se implementan y calculan son: la escala externa, la dimensión fractal, el factor de anisotropía de esparcimiento, el coeficiente de esparcimiento y la varianza del índice de refracción. Se emplean imágenes de muestras sanas y tumorales de colon. Se han utilizado un total de 84 imágenes con aumentos de 10x, 20x y 40x. La importancia de la discriminación entre tejidos sanos y tumorales se evalúa estadísticamente con ANOVA (valor  $p < 0,005$ ). El análisis se realiza para diferentes aumentos. Se muestra una dependencia del tipo de parámetro y del aumento de la imagen. El p-valor de los parámetros más significativos está por debajo de  $10^{-5}$ . Las muestras de los tejidos de colon presentan un conjunto de parámetros de contraste de fase significativos. En concreto la dimensión fractal a 10x es significativa para este tipo de tejido bajo análisis. Estos resultados son prometedores para el uso del contraste de fase avanzado en la praxis clínica de histología digital.*

## 1. Introducción

La histopatología es el principal estándar para el diagnóstico [1]. Los métodos actuales se basan en la observación microscópica directa de cortes de tejido teñidos por parte del patólogo. La biopsia convencional está especialmente indicada para la discriminación de tejidos tumorales, incluso en estadios precoces. El procedimiento presenta varias limitaciones, que incluyen dispositivos para el procesamiento de tejidos, retrasos en el diagnóstico causados por el cribado manual o errores provocados por el sesgo del patólogo. La histología digital contribuye a la automatización del procedimiento y a la mejora del diagnóstico mediante parámetros más objetivos. La histología digital se implementa mediante la digitalización de imágenes microscópicas de intensidad de muestras de biopsias convencionales. Las imágenes sin etiquetas podrían mejorar aún más el análisis histológico digital. Eliminaría los dispositivos de procesamiento y reduciría el tiempo de procesamiento. Uno de los problemas es que los tejidos biológicos no proporcionan suficiente contraste intrínseco para distinguir las

estructuras de interés y se necesita la tinción. Los tejidos biológicos son intrínsecamente turbios y son medios de bajo contraste [2]. Este hecho complica la observación de estructuras internas sin cortes físicos.

En este trabajo, se propone un enfoque novedoso que consiste en explotar la variación del índice de refracción de los tejidos biológicos para el diagnóstico manteniendo una resolución muy alta. Se propone emplear cortes de tejido fresco sin etiquetas e imágenes de contraste de fase para el análisis de muestras biológicas. Este enfoque permite, en primer lugar, eliminar los artefactos que pueden aparecer en las etapas de fijación y tinción de los procedimientos de biopsia convencionales. Esos procedimientos se pueden evitar debido a la variación intrínseca del índice de refracción que explota el contraste de fase. También reduce el tiempo de procesamiento de la muestra. Las imágenes de contraste de fase proporcionan características tisulares novedosas que son particularmente relevantes en los cambios morfoestructurales, que son típicos en los tejidos tumorales, [3].

Este trabajo comienza con la descripción del modelo teórico que da lugar a la definición de los parámetros de contraste de fase bajo estudio. Posteriormente se describe el procedimiento experimental de toma de muestras e imágenes. A continuación, se obtienen los parámetros identificados y se analizan estadísticamente para comprobar su potencialidad discriminadora. Por último se discuten los resultados y se extraen conclusiones.

## 2. Modelo teórico

La microscopía de contraste de fase (PCM) presenta ventajas sobre la microscopía convencional. La microscopía de contraste de fase fue desarrollada por Zernike en la década de 1930. Permite obtener imágenes no invasivas y sin etiquetas de células vivas.

El estudio de los modelos de la distribución del índice de refracción comenzó con los modelos de los medios continuos atmosféricos y oceánicos y luego se extrapoló al modelo de medio continuo de los tejidos biológicos. La distribución del índice de refracción en un medio turbulento general requiere una descripción estadística del haz que se propaga. En este caso estamos ante muestras biológicas turbias, por lo que no se espera variación temporal. En consecuencia, el análisis de los parámetros de

contraste de fase debe basarse en la teoría de los procesos aleatorios [4].

De la distribución espacial del índice de refracción, las características estadísticas más relevantes son los dos primeros momentos: el valor medio

$$n_0(\vec{r}) = \langle n(\vec{r}) \rangle_M \quad (1)$$

Y su función de covarianza,

$$B_n(\vec{r}_1, \vec{r}_2) = \langle [n(\vec{r}_1) - n_0(\vec{r}_1)][n(\vec{r}_2) - n_0(\vec{r}_2)] \rangle_M \quad (2)$$

En un campo que se propaga a través de un medio de esparcimiento débil, la relación de dispersión asociada se puede expresar como,

$$\langle k^2 \rangle = n_0^2 \beta_0^2 \left( 1 + \frac{\sigma_n^2}{n_0^2} \right) \quad (3)$$

Donde  $\langle k^2 \rangle$  es el momento de segundo orden del vector de onda,  $\langle k^2 \rangle = \langle k_x^2 \rangle + \langle k_y^2 \rangle + \langle k_z^2 \rangle$ , con paréntesis angulares que indican el promedio del conjunto,  $n_0$  es el promedio del índice de refracción,  $\beta_0$  es el número de onda en el vacío,  $\beta_0 = \omega/c$ , y  $\sigma_n^2$  es la variación espacial del índice de refracción. La variación del índice de refracción de una muestra transparente se puede calcular a partir de la imagen de contraste de fase medida como [5],

$$\sigma_n^2 = \frac{1}{\beta_0^2} \langle |\nabla\phi|^2 \rangle \quad (4)$$

La ecuación (4) implica que la variación del índice de refracción se puede extraer de la intensidad media del gradiente de las imágenes de contraste de fase.

Otro parámetro de contraste de fase que se puede estimar es el coeficiente de esparcimiento ( $\mu_s$ ). El coeficiente de esparcimiento se puede calcular utilizando el teorema de fase de esparcimiento [6],

$$\mu_s = \frac{\langle \Delta\phi^2(r) \rangle_r}{L} \quad (5)$$

Donde  $\langle \Delta\phi^2(r) \rangle$  indica la variación espacial de la fase y  $L$  es el grosor de la muestra de tejido.

Además, considerando el carácter intrínsecamente esparcador de los tejidos biológicos, también se puede estimar el factor de anisotropía de esparcimiento. La estimación también hace uso del teorema de fase de esparcimiento y puede expresarse como,

$$g = 1 - \frac{1}{2k_0^2} \frac{\langle |\nabla[\phi(r)]|^2 \rangle_r}{\langle \Delta\phi^2(r) \rangle_r^2} \quad (6)$$

donde  $k_0$  es el vector de onda incidente y  $\nabla[\phi(r)]$  es el gradiente de fase.

La turbulencia óptica suele explicarse por la presencia de irregularidades en el índice de refracción, los llamados “remolinos turbulentos”. Aparecen debido a fluctuaciones en varias propiedades físicas de la materia, como la temperatura, la presión y la concentración no homogénea de especies, y es de aplicación habitual en modelos atmosféricos o marinos. Las posibilidades de mezcla de

diferentes partes de la muestra por factores externos o internos, como el crecimiento celular o la transferencia de fluidos en tejidos biológicos, provocan transferencia de energía entre remolinos de diferentes tamaños. El mayor tamaño posible de un remolino se define como la escala exterior  $L_0$  de turbulencia. Los remolinos más grandes se descomponen en otros más pequeños hasta que su tamaño alcanza el límite inferior y la energía se disipa. Para estudiar estas características en muestras histológicas, el espectro de inhomogeneidades del índice se ajusta a la ecuación,

$$\Phi(\kappa) = \frac{4\pi\sigma_n^2 L_0^2 (m-1)}{(1+\kappa^2 L_0^2)^m} \quad (7)$$

que tiene la forma del espectro de Von Karman, con excepción del exponente  $m$ , que puede asumir valores distintos a  $4/3$ . En este caso,  $m$  es aproximadamente igual a la mitad de la pendiente medida en el rango de escala de ley de potencia [7]. Su magnitud está relacionada con la dimensión fractal de una superficie bidimensional,  $d_f = 4 - m$ .

Se ha demostrado que la distribución espacial de los principales componentes de los tejidos biológicos generales presenta autosimilitud. Esto hace posible el uso potencial del análisis fractal también en estas estructuras. El objetivo es tratar de identificar regularidades estadísticas espaciales que puedan ser descritas por leyes de potencia. En particular, la dimensión fractal define las características de estabilidad de una imagen, con respecto a las transformaciones de escala. La dimensión fractal de una imagen se puede estimar mediante varias técnicas: conteo de cajas; correlación; sandbox; o incluso el espectro de Fourier. El método de conteo de cajas consiste en una cuadrícula con cajas de tamaño  $\varepsilon$  superpuestas a una imagen. El número de casillas que contienen cualquier parte de la figura se registra como  $N(\varepsilon)$ . El procedimiento se repite para diferentes tamaños de  $\varepsilon$ , y la dimensión fractal  $d_f$  se calcula como,

$$d_f = -\frac{\log[N(\varepsilon)]}{\log(\varepsilon)} \quad (8)$$

La dimensión fractal  $d_f$  generalmente se calcula a partir del valor negativo de la pendiente de la regresión lineal de  $\log[N(\varepsilon)]$  en  $\log(\varepsilon)$ ,

$$\log[N(\varepsilon)] = -d_f \log(\varepsilon) \quad (9)$$

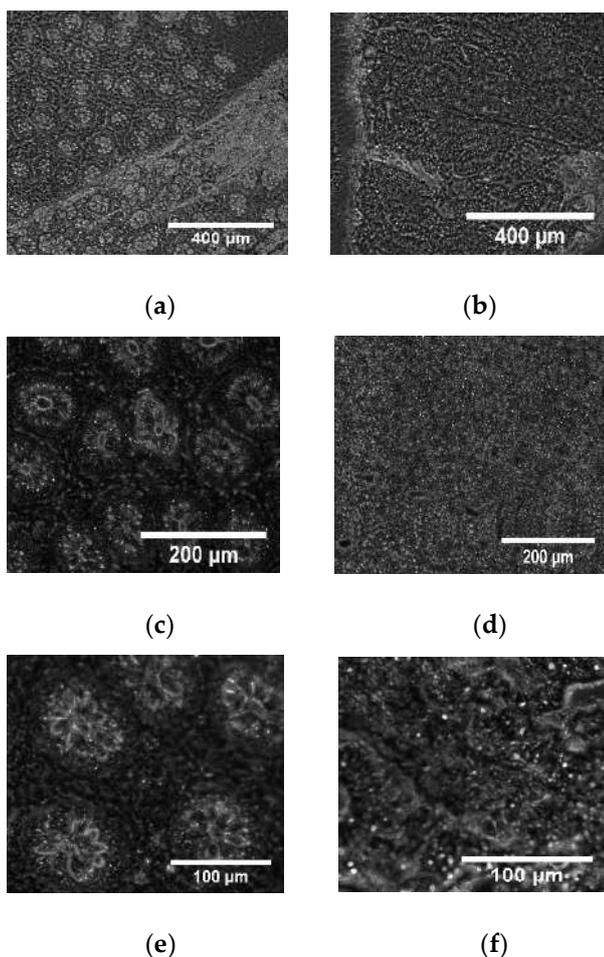
### 3. Materiales y métodos

Se empleó una configuración de microscopía de contraste de fase. Las muestras se colocaron en un escenario motorizado para el barrido espacial automático. El contraste de fase se implementó mediante un anillo de fase externo, combinado con varios objetivos de contraste de fase con aumentos de 10x, 20x y 40x (Nikon 10x Plan Fluor 0.30NA Ph1 DLL, 20x SPlan Fluor ELWD 0.45 NA ph1 ADM y 40x SPlan Fluor ELWD 0.60 NA ph2 ADM). Las imágenes se grabaron con una cámara CCD (Orca-R2,

Hamamatsu) y se almacenaron en formato de escala de grises de 12 bits [8].

Cada muestra histológica se visualiza mediante la configuración de microscopía de contraste de fase descrita anteriormente. Las imágenes se obtienen con aumentos de 10x (866x660  $\mu\text{m}$ ; tamaño de píxel 0,64  $\mu\text{m}$ ), 20x (433x330  $\mu\text{m}$ ; tamaño de píxel 0,3225  $\mu\text{m}$ ) y 40x (219x167  $\mu\text{m}$ ; tamaño de píxel 0,1632  $\mu\text{m}$ ). Se presta especial atención a la saturación de píxeles y al campo de visión. Se adquirieron un total de 84 imágenes.

Se emplearon muestras frescas de tejidos sanos y tumorales del biobanco de tejidos biológicos del Hospital Marqués de Valdecilla. Particularmente las muestras biológicas procedían de colon. Las muestras fueron congeladas y cortadas por microtomo, con un espesor menor a 6  $\mu\text{m}$ . No se aplicó ningún procedimiento de fijación o tinción a las muestras, ya que estos procesos podrían alterar las propiedades del tejido, en particular las propiedades del índice de refracción. Las muestras se recogieron sobre portaobjetos de vidrio de microscopía tratados para su medición. Las muestras se conservaron a 4°C durante dos horas antes de la medición por microscopía de contraste de fase. En la Figura 1 se presentan ejemplos de imágenes de tejido de colon sano y tumoral a 10x, 20x y 40x.

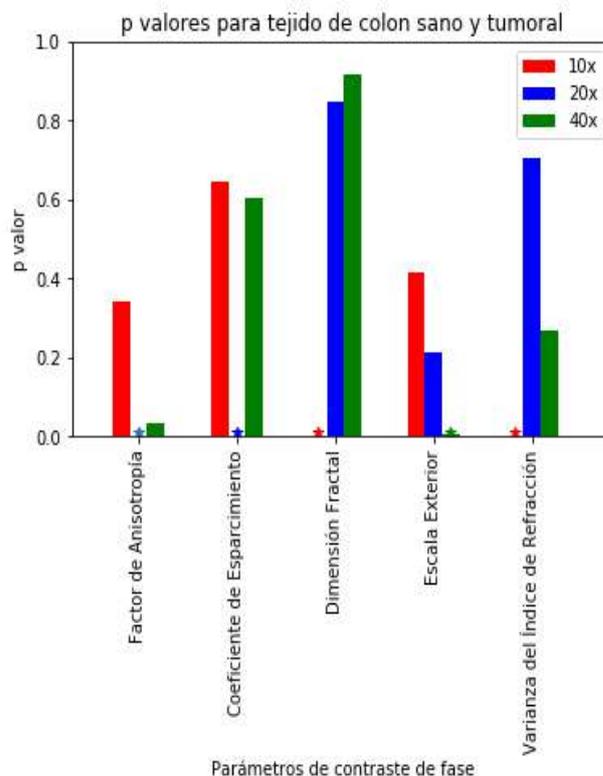


**Figura 1.** Imagen histológica de tejido de colon sano (a) y tumoral (b) con aumento de 10x; imagen histológica de tejido de colon sano (c) y tumoral (d) con aumento de 20x; imagen histológica de tejido de colon sano (e) y tumoral (f) con aumento de 40x.

#### 4. Obtención y análisis estadístico de parámetros ópticos de fase

A partir de las imágenes de contraste de fase obtenidas, se lleva a cabo una implementación del cálculo de los parámetros de contraste de fase identificados anteriormente, ya que estos parámetros influyen en el modelado matemático de la propagación óptica en los tejidos biológicos, [4]. Se han usado 42 imágenes de tejido de colon sano y 42 imágenes de tejido de colon tumoral. Particularmente, se lleva a cabo el cálculo del factor de anisotropía de esparcimiento, el coeficiente de esparcimiento, la dimensión fractal, la escala exterior y la varianza del índice de refracción. Se realiza el cálculo mediante una implementación propia para cada una de las imágenes obtenidas, y se separan los resultados en dos grupos, correspondientes a las muestras sanas y patológicas.

Se necesita un análisis más profundo para estudiar la relevancia estadística de los parámetros y su consiguiente potencialidad diagnóstica. Como estamos tratando de analizar la importancia de los posibles valores promedio de los parámetros, se realiza un análisis estadístico ANOVA. Estos análisis se realizan para tejidos sanos y tumorales en cada aumento, para los diferentes parámetros de contraste de fase obtenidos. La Figura 2 y la Tabla 1 muestran los resultados de los p-valores del análisis ANOVA para muestras de tejido de colon a diferentes aumentos. Como la prueba estadística verifica la hipótesis de medias iguales de los parámetros, habrá diferencias significativas para p-valores pequeños.



**Figura 2.** Resultados de p-valores del análisis ANOVA para parámetros de contraste de fase al comparar muestras de colon sano y tumoral, para diferentes aumentos. Los valores mostrados con un \* son valores inferiores a 0,009.

Parámetro de contraste	10x	20x	40x
FA	0.3401	0.0027	0.0349
CE	0.6446	0.0022	0.6036
DF	5.4414e-04	0.8466	0.9181
EE	0.4176	0.2155	0.0086
VIR	6.6435e-04	0.7047	0.2696

**Tabla 1.** Resultados de p-valores del análisis ANOVA para parámetros de los siguientes parámetros de contraste: Factor de Anisotropía (FA), Coeficiente de Esparcimiento (CE), Dimensión Fractal (DF), Escala Exterior (EE) y Varianza del Índice de Refracción (VIR), en muestras de colon sano y tumoral a diferentes magnificaciones.

## 5. Resultados y discusión

Los resultados mostrados en las secciones anteriores permiten estimar la potencialidad de los nuevos parámetros de contraste de fase propuestos para el diagnóstico de tejidos. La principal diferencia entre el análisis histológico convencional, incluida la histología digital, se basa en el hecho de que se recopila y emplea la información de fase de la señal óptica. Este hecho permite el uso de cortes de tejido fresco, que se pueden obtener más rápidamente, ya que no requieren procesamiento adicional, al contrario de las muestras fijas convencionales. Cada una de las imágenes, correspondientes a colon, ya sea sano o tumoral, y con diferentes aumentos a 10x, 20x y 40x, se procesó para calcular los parámetros de contraste de fase propuestos: factor de anisotropía de esparcimiento, coeficiente de esparcimiento, dimensión fractal, escala exterior y varianza del índice de refracción. Se observan unas tendencias que indican que la validez de los parámetros parece depender del factor de magnificación de las imágenes. Se ha realizado un análisis estadístico utilizando ANOVA de los parámetros de contraste en los diferentes tejidos a diferentes magnificaciones. Se debe prestar atención a los p-valores bajos. Estos valores indican que los grupos sanos y tumorales son estadísticamente significativos, en la medida expresada por el p-valor. Se establece un criterio de un p valor inferior a 0,005 para ser considerado estadísticamente significativo. Por ejemplo, en el caso de las muestras de colon, los parámetros de contraste de fase significativos, según la Figura 2, serían el factor de anisotropía de esparcimiento a 20x, el coeficiente de esparcimiento a 20x, la dimensión fractal a 10x, la escala exterior a 40x y la varianza del índice de refracción a 10x. Los resultados más significativos son para la dimensión fractal a 10x y la variación del índice de refracción a 10x. Las variaciones con respecto al aumento tienen que ver con la diferencia estructural del tejido en relación con el factor de magnificación. Estos resultados muestran una aplicabilidad prometedora de los parámetros de contraste de fase en histología digital en muestras de tejido fresco. En particular, para las muestras de tejido bajo análisis, existe una coincidencia significativa en varios parámetros con un aumento de 10x. Aun así, la aplicación

clínica de estos parámetros para un conjunto general de muestras biológicas probablemente requeriría la consideración de aumentos y parámetros combinados en un algoritmo de clasificación sofisticado. El hecho de que las muestras sean frescas podría ser, por un lado, una ventaja, ya que se elimina el procesamiento de las muestras y el procedimiento resulta más rápido, pero por otro lado dificultaría la conservación, que solo podría hacerse por congelación. También, se tendría como futura investigación, estudiar como los cambios morfoestructurales de los tejidos tumorales de colon en las diferentes magnificaciones, afectan a los parámetros de contraste. Los resultados de este trabajo muestran la potencialidad de los parámetros de contraste de fase para la histología digital en tejido de colon.

## Agradecimientos

Este Trabajo ha sido parcialmente financiado por el proyecto del Plan Nacional de I+D+i “Cribado diagnóstico de microorganismos mediante microscopía avanzada e inteligencia artificial en patologías humana” (PID2021-127691OB-I00), del Ministerio de Ciencia e Innovación, cofinanciado con fondos FEDER.

## Referencias

- [1] Tortora G. J., Derrickson B. H. *Principles of Anatomy and Physiology*; 15th ed.; Wiley, 2018.
- [2] Tuchin V.V. *Tissue Optics: Light Scattering Methods and Instruments for Medical Diagnosis*; Society of Photo-Optical Instrumentation Engineers (SPIE), 2015; ISBN 978-1-62841-516-2.
- [3] Wang Z, Tangella K, Balla A, Popescu G. Tissue refractive index as marker of disease. *J Biomed Opt.* 2011 Nov;16(11):116017. doi: 10.1117/1.3656732. PMID: 22112122; PMCID: PMC3223513.
- [4] Chen X, Korotkova O. Optical beam propagation in soft anisotropic biological tissues. *OSA Continuum Vol. 1, Issue 3*, pp. 1055-1067 (2018)
- [5] Shan M., Kandel M. E., Popescu G., Refractive index variance of cells and tissues measured by quantitative phase imaging. *Optics Express* 25, 1573-1581 (2017).
- [6] Wang, Z., Ding, H., Popescu, G. Scattering-Phase Theorem. *Opt. Lett.* 2011, 36, 1215–1217, doi:10.1364/OL.36.001215.
- [7] Schmitt J. M., Kumar, G., Turbulent nature of refractive-index variations in biological tissue, OSA Publishing, *Optics Letters*, Vol. 21, Issue 16, pp. 1310-1312 (1996).
- [8] Ganoza-Quintana, J.L., Fanjul-Vélez, F., Arce-Diego, J.L. Digital Histology by Phase Imaging Specific Biomarkers for Human Tumoral Tissues Discrimination. *Appl. Sci.* 2021, 11, 6142. <https://doi.org/10.3390/app11136142>.

# Efecto del plomo en la generación de arritmias. Estudio de simulación en un modelo 3D de aurícula humana

L.C. Palacio<sup>1</sup>, C. Tobón<sup>1</sup>, J. Saiz<sup>2</sup>

<sup>1</sup> MATBIOM, Universidad de Medellín, Medellín, Colombia, lcpt91@gmail.com, ctobon@udemedellin.edu.co

<sup>2</sup> CI<sup>2</sup>B, Universitat Politècnica de València, Valencia, España, jsaiz@ci2b.upv.es

## Resumen

La exposición a contaminantes atmosféricos como el plomo (Pb) aumenta el riesgo de padecer arritmias cardíacas y contribuye a la generación de fibrosis miocárdica. En este trabajo, evaluamos los efectos del Pb en condiciones de ritmo sinusal a concentraciones baja, media y alta sobre las características electrofisiológicas y la generación de patrones de propagación arrítmicos en la aurícula. Para ello, se implementaron modelos celulares de miocitos y fibroblastos auriculares en un modelo 3D de aurícula humana. Los resultados sugieren un efecto proarrítmico del Pb a medida que incrementa su concentración, donde el remodelado iónico y estructural (fibrosis), conducen a un escenario propicio para la generación de rotorez, favoreciendo la generación de arritmias auriculares.

## 1. Introducción

La contaminación atmosférica se ha indicado como una de las principales causas de morbilidad y mortalidad, siendo responsable de más de 6,5 millones de muertes al año a nivel mundial [1]. Los daños a la salud asociados a la exposición a la contaminación atmosférica representa un coste económico anual a nivel mundial de \$8,1 trillones de dólares [2], convirtiéndose en un importante problema de salud pública [3]. De acuerdo con la Agencia de Protección Ambiental de Estados Unidos (EPA), los contaminantes más peligrosos para la salud son: el dióxido de nitrógeno, el material particulado, el dióxido de azufre, el monóxido de carbono y el plomo (Pb).

Estudios recientes sugieren una fuerte correlación entre la exposición a contaminantes atmosféricos y efectos adversos para la salud, asociados a enfermedades respiratorias y cardíacas, principalmente [4]–[6]. A nivel cardíaco, la exposición a contaminantes contribuye a una remodelación adversa, a lesiones degenerativas y fibróticas [7], [8], y al empeoramiento de la fibrosis miocárdica [9], incrementando el riesgo a padecer arritmias cardíacas [6], [10]. El Pb es considerado uno de los contaminantes metálicos más peligrosos para la salud [11], incluso a bajo nivel de exposición [12], causando estrés oxidativo y propiciando a su vez, la generación de inflamación, fibrosis y apoptosis [13]; siendo la fibrosis es un sello distintivo de la remodelación estructural arritmogénica y del agravamiento de arritmias auriculares [14].

Sin embargo, el efecto del Pb sobre las características electrofisiológicas y patrones de propagación eléctrica auricular en humanos, ha sido poco estudiado. La implementación de modelos auriculares realistas permite evaluar el efecto de diferentes concentraciones de Pb teniendo en cuenta los efectos observados a nivel iónico (remodelado iónico) y estructural (fibrosis) sin las

limitaciones éticas, de complejidad y coste que implica el desarrollo de estudios experimentales y clínicos. En este sentido, el objetivo de este trabajo es evaluar el efecto del Pb a diferentes concentraciones, sobre las características electrofisiológicas y patrones de propagación en modelos auriculares humanos.

## 2. Métodos

### 2.1. Modelos celulares de auricular humana

Para simular el potencial de acción de miocitos cardíacos, se implementaron los modelos matemáticos de célula de aurícula humana desarrollados por Koivumäki *et al.* [81] y Courtemanche–Ramirez–Nattel–Kneller [15], [16]. El voltaje transmembrana ( $V_m$ ) es dado por la siguiente ecuación:

$$C_m \frac{dV_m}{dt} + I_{ion} + I_{est} = 0 \quad (1)$$

donde  $C_m$  es la capacidad de membrana,  $I_{ion}$  es el total de corrientes transmembrana involucradas en cada modelo en específico, y  $I_{est}$  es la corriente de simulación externa.

### 2.2. Modelo del efecto del Pb

Para modelar la relación concentración-respuesta por bloqueo de los canales de calcio tipo L ( $I_{CaL}$ ) [17], generada por el Pb, se implementó la ecuación de Hill a través de la siguiente ecuación de fracción de bloqueo ( $b_{Pb}$ ):

$$b_{Pb} = \frac{1}{1 + \left(\frac{169}{D_{Pb}}\right)^{0.5}} \quad (2)$$

Donde la concentración máxima a la cual se bloquea al 50% el canal de calcio ( $IC_{50}$ ) es de 169 nM [17]; el valor del 0,5 del coeficiente de Hill indica una unión cooperativa. Se simuló concentraciones de Pb ( $D_{Pb}$ ) de 0 nM (control), 34 nM (baja), 85 nM (media) y 170 nM (alta). Estas concentraciones corresponden al ~ 80%, 200%, y 400%, respectivamente, del nivel máximo de material particulado permisible, de acuerdo con las guías de calidad del aire de la Organización Mundial de la salud (75  $\mu\text{g}/\text{m}^3$ ). La ecuación de la corriente  $I_{CaL}$  modificada para modelar el efecto del Pb es la siguiente:

$$I_{CaL} = (1 - b_{Pb})g_{CaL}df_{Ca}(V_m - 65) \quad (3)$$

Adicionalmente, para simular las condiciones de remodelación estructural generadas por el Pb, se implementó el modelo matemático de fibroblastos propuesto por MacCannell *et al.* [18]. Este modelo fue asignado aleatoriamente al 6% de los nodos que conforman la aurícula izquierda (excluyendo venas pulmonares),

distribuidos de manera que se intercalan uniformemente con los cardiomiocitos (fibrosis difusa).

### 2.3. Modelo 3D virtual de aurícula humana

Se implementó un modelo tridimensional (3D) de aurícula humana [19]. Este modelo está compuesto por 515.010 elementos hexaédricos, con una resolución espacial uniforme de 300  $\mu\text{m}$ . El modelo incluye orientación realista de las fibras, heterogeneidad electrofisiológica y anisotropía.

### 2.4. Modelo de propagación

La propagación cardiaca en el tejido es descrita por el siguiente modelo monodominio de propagación eléctrica:

$$\frac{1}{S_v} \nabla \cdot (D \nabla V_m) = C_m \frac{\partial V_m}{\partial t} + I_{ion} + I_{st} \quad (4)$$

Donde  $S_v$  es el radio de superficie/volumen y  $D$  es el tensor de conductividad. Esta ecuación fue resuelta numéricamente usando el método de elementos finitos en el software EMOS [20].

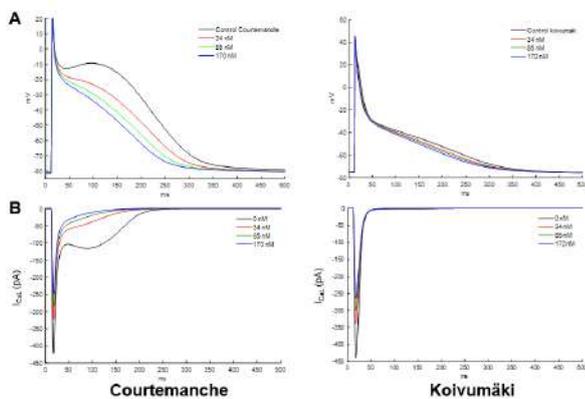
### 2.5. Protocolos de estimulación

A nivel celular, se aplicó un tren de estímulos a una longitud de ciclo básico (BCL) de 1000 ms. Se realizaron mediciones de la duración del potencial de acción al 90% de la repolarización ( $APD_{90}$ ), del potencial de reposo de la membrana (RMP) y de la corriente iónica  $I_{CaL}$ , para cada uno de los modelos de miocito auricular.

En el modelo 3D, se aplicó un protocolo de estimulación estándar S1-S2, que consiste en pulsos rectangulares de 2 ms de duración y 28 pA/pF de amplitud, donde S1 corresponde al ritmo sinusal a un BCL de 500 ms; y el estímulo S2 corresponde a dos focos ectópicos, localizados en el centro de la base de las venas pulmonares derechas y el otro en el istmo cavotricuspidéo cerca del seno coronario. Se realizaron mediciones del  $APD_{90}$  y del RMP. Las simulaciones tuvieron una duración de 5 segundos.

## 3. Resultados

En la Figura 1 se observa el efecto generado a diferentes concentraciones de Pb sobre el potencial de acción y la corriente de calcio  $I_{CaL}$ , para cada modelo.



**Figura 1.** Efecto del Pb a diferentes concentraciones en el modelo de miocito auricular de Courtemanche y Koivumäki sobre el A) potencial de acción y B) corriente  $I_{CaL}$ .

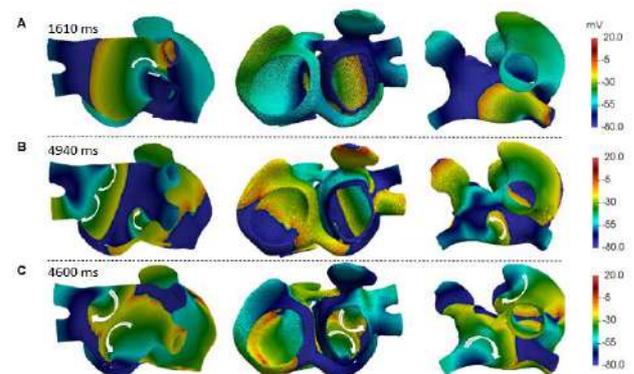
Sin concentración de Pb (control), el pico máximo de la corriente de  $I_{CaL}$  fue de  $-423$  y  $-439$  pA, para los modelos de Courtemanche y Koivumäki, respectivamente, el cual se redujo progresivamente de manera concentración-dependiente, lo cual se evidencia en la pérdida del domo de la fase de meseta y en la reducción del APD, principalmente en el modelo auricular de Courtemanche.

En la Tabla 1 se muestran los valores del efecto del Pb a las diferentes concentraciones, sobre el pico máximo de la corriente  $I_{CaL}$  y  $APD_{90}$ , para cada modelo de miocito. A alta concentración de Pb (170 nM), el pico de la corriente  $I_{CaL}$  se redujo, respecto a control, un 42% y un 40%, en los modelos de Courtemanche y Koivumäki, respectivamente. Ocasionando una reducción del  $APD_{90}$  de un 24% y un 14%, en los modelos de Courtemanche y Koivumäki, respectivamente. El RMP no presentó cambios significativos en ambos modelos.

[Pb]	Courtemanche		Koivumäki	
	$APD_{90}$	$I_{CaL}$	$APD_{90}$	$I_{CaL}$
Control	309 ms	$-423$ pA	289 ms	$-439$ pA
Baja	278 ms	$-324$ pA	271 ms	$-340$ pA
Media	254 ms	$-284$ pA	262 ms	$-300$ pA
Alta	236 ms	$-249$ pA	250 ms	$-264$ pA

**Tabla 1.** Valores de  $APD_{90}$  y pico máximo de la corriente  $I_{CaL}$ , a diferentes concentraciones de Pb.

Con base en los resultados unicelulares, se seleccionó el modelo de Courtemanche para el estudio 3D. Al aplicar el protocolo S1-S2 al modelo 3D de aurícula humana, en la simulación de control (0 nM de Pb) se propagan los estímulos de forma regular sin generar actividad arritmica; a baja concentración (34 nM), el frente de onda reentrante se encuentra con tejido en periodo refractario por lo que no se favorece un escenario para la generación de arritmia (Figura 2A). Sin embargo, al incrementar la concentración de Pb, se observa un efecto proarritmico, que empeora con el incremento de la concentración, debido a la reducción del APD.



**Figura 2.** Efecto del Pb sobre la propagación de la actividad eléctrica auricular a concentración A) baja, B) media y C) alta. Las flechas blancas señalan la dirección de los frentes de onda reentrantes.

A concentración media de Pb (85 nM), la actividad arritmica se caracterizó por la presencia de ondas reentrantes en la pared posterior y superior de la aurícula

izquierda. Se observaron dos rotores relativamente estables localizados en las bases de las venas pulmonares izquierdas, conformando una reentrada en figura de ocho (Figura 2B), mientras que en la aurícula derecha se observaron colisiones de frentes de onda. A concentración alta de Pb (170 nM), la actividad arritmica se caracterizó por la presencia de ondas reentrantes en ambas aurículas. En la aurícula derecha, se generó un rotor en la unión entre la vena cava superior y el haz de Bachmann que migra al apéndice derecho; mientras que en la aurícula izquierda, se observaron dos rotores estables en la pared posterior, conformando una reentrada en figura de ocho, y un rotor en la pared superior (Figura 2C).

#### 4. Discusión

El Pb es considerado como uno de los metales más peligrosos para la salud humana [11], siendo un elemento característico de la contaminación atmosférica. La inhalación de partículas de Pb generadas por la combustión de materiales que contienen este metal puede causar anemia, hipertensión, disfunción renal, inmunotoxicidad, el desarrollo de aterosclerosis, enfermedades cardiovasculares, entre otras [13]. Asimismo, en los seres humanos, las altas concentraciones del Pb en la sangre se han asociado con anomalías electrocardiográficas, enfermedad arterial periférica, hipertrofia del ventrículo izquierdo y mortalidad por enfermedades cardiovasculares [21], [22]. Sin embargo, el efecto del Pb sobre las características electrofisiológicas y arritmias auriculares en humanos, ha sido poco estudiado.

En este estudio se evaluó el efecto del Pb en las características electrofisiológicas y propagación de la actividad eléctrica en modelos auriculares humanos. Nuestros resultados a nivel unicelular muestran los efectos electrofisiológicos del Pb al bloquear el canal de calcio tipo L de forma concentración-dependiente, observándose reducción del domo en la fase de meseta del potencial de acción, generando una morfología más triangular, y un acortamiento en el APD, a medida que incrementa la concentración del Pb. El Pb ha sido evaluado de forma individual en células aisladas animales, Bernal *et al.* [23] evaluaron en miocitos ventriculares de rata el efecto de diferentes concentraciones de Pb (1 y 10 mM), demostrando su efecto sobre el bloqueo de los canales de calcio tipo L [23]. Análogamente, Ferreira de Mattos *et al.* [17] evaluaron el efecto del Pb a diferentes concentraciones (1 - 200  $\mu$ M) en células HEK29, reportando el bloqueo de los canales Cav1.2; observaron además una reducción del influjo de calcio sarcolémico y de la actividad ATPasa de la miosina en la fase inicial de la despolarización y una reducción de la meseta del potencial de acción [17]. Este hallazgo coincide con un estudio desarrollado en el miocardio ventricular de ratas [24], el cual indica que la administración aguda de Pb disminuye la contractilidad miocárdica al reducir el influjo de calcio sarcolémico y la actividad de la miosina. Estos efectos han conducido a una disminución de la contractilidad y la excitabilidad miocárdica [17], [24], y a una reducción de la velocidad de conducción; y en general, a un escenario propenso para la generación de arritmias cardíacas.

Por otra parte, los efectos de los contaminantes atmosféricos son tanto a nivel celular como estructural (de tejido). La presencia de fibrosis debida a la exposición a contaminantes atmosféricos, es un sello distintivo en el desencadenamiento y sostenimiento de los patrones de propagación arritmica cardíaca [6]–[10]. El estrés oxidativo ha sido implicado en la patogénesis, como parte de los mecanismos reportados en la literatura para tratar de explicar los efectos dañinos de la exposición del Pb en el cuerpo humano [25]. En un estudio realizado en niños uruguayos, se reportó una alta presencia de marcadores de estrés oxidativo en la orina, por exposición a bajos niveles de Pb [26]. El estrés oxidativo a su vez, propicia la generación de inflamación, fibrosis y apoptosis [13]. En nuestro estudio, implementando un modelo realista 3D de aurícula humana, modelamos no solo el efecto del Pb a nivel iónico (bloqueo sobre la  $I_{CaL}$ ), sino también a nivel estructural (fibrosis). Nuestras simulaciones mostraron un efecto proarritmico del Pb, donde a altas concentraciones se observó actividad arritmica caracterizada por la presencia principalmente de rotores. Se ha informado ampliamente que los rotores son mecanismos importantes que mantienen las arritmias auriculares [27], [28].

#### 5. Conclusión

Los resultados de simulación señalan el efecto proarritmico del Pb, donde el remodelado iónico, el cual genera cambios morfológicos en el potencial de acción reduciendo el APD, junto con un remodelado estructural (fibrosis), conducen a un escenario propicio para la generación de rotores de manera concentración dependiente, favoreciendo la generación de arritmias auriculares.

#### Agradecimientos

Este trabajo fue financiado por el Ministerio de Ciencia, Tecnología e Innovación (MINCIENCIAS) de Colombia, a través del proyecto No. 120677757994, el “Programa de excelencia doctoral del Bicentenario 2019” de MINCIENCIAS y por la Dirección General de Política Científica de la Generalitat Valenciana (PROMETEO/2020/043).

#### Referencias

- [1] R. Fuller *et al.*, “Pollution and health: a progress update,” *Lancet Planet. Heal.*, vol. 0, no. 0, May 2022, doi: 10.1016/S2542-5196(22)00090-0.
- [2] World Bank, “The Global Health Cost of PM2.5 Air Pollution: A Case for Action Beyond 2021,” The World Bank, Apr. 2022. doi: 10.1596/978-1-4648-1816-5.
- [3] Organización Mundial de la salud, “Salud y medio ambiente,” May 2016.
- [4] I. Cattani-Cavaliere *et al.*, “Dimethyl Fumarate Attenuates Lung Inflammation and Oxidative Stress Induced by Chronic Exposure to Diesel Exhaust Particles in Mice,” *Int. J. Mol. Sci.*, vol. 21, no. 24, p. 9658, Dec. 2020, doi: 10.3390/ijms21249658.
- [5] A. Karoui *et al.*, “Nitrogen Dioxide Inhalation Exposures Induce Cardiac Mitochondrial Reactive Oxygen Species Production, Impair Mitochondrial Function and Promote Coronary Endothelial Dysfunction,” *Int. J. Environ. Res. Public Health*, vol. 17, no. 15, pp. 1–14, Aug. 2020, doi:

- 10.3390/IJERPH17155526.
- [6] A. M. de Oliveira-Fonoff *et al.*, “The role of air pollution in myocardial remodeling,” *PLoS One*, vol. 12, no. 4, p. e0176084, 2017, doi: 10.1371/journal.pone.0176084.
- [7] D. E. Newby *et al.*, “Expert position paper on air pollution and cardiovascular disease,” *Eur. Heart J.*, vol. 36, no. 2, pp. 83–93b, Jan. 2015, doi: 10.1093/eurheartj/ehu458.
- [8] Y. Liu, J. M. Goodson, B. Zhang, and M. T. Chin, “Air pollution and adverse cardiac remodeling: clinical effects and basic mechanisms,” *Front. Physiol.*, vol. 6, p. 162, 2015, doi: 10.3389/fphys.2015.00162.
- [9] L. E. Wold *et al.*, “Cardiovascular Remodeling in Response to Long-Term Exposure to Fine Particulate Matter Air Pollution,” *Circ. Hear. Fail.*, vol. 5, no. 4, pp. 452–461, Jul. 2012, doi: 10.1161/CIRCHEARTFAILURE.112.966580.
- [10] R. D. Brook *et al.*, “Particulate Matter Air Pollution and Cardiovascular Disease,” *Circulation*, vol. 121, no. 21, pp. 2331–2378, Jun. 2010, doi: 10.1161/CIR.0b013e3181d8bec1.
- [11] World Health Organization, “Lead poisoning and health,” 2018. .
- [12] A. L. Wani, A. Ara, and J. A. Usmani, “Lead toxicity: a review,” *Interdiscip. Toxicol.*, vol. 8, no. 2, p. 55, Jun. 2015, doi: 10.1515/INTOX-2015-0009.
- [13] N. D. Vaziri, “Mechanisms of lead-induced hypertension and cardiovascular disease,” *Am. J. Physiol. - Hear. Circ. Physiol.*, vol. 295, no. 2, pp. 454–465, Aug. 2008, doi: 10.1152/AJPHEART.00158.2008.
- [14] N. Akoum *et al.*, “Atrial fibrosis helps select the appropriate patient and strategy in catheter ablation of atrial fibrillation: a DE-MRI guided approach,” *J. Cardiovasc. Electrophysiol.*, vol. 22, no. 1, pp. 16–22, Jan. 2011, doi: 10.1111/j.1540-8167.2010.01876.x.
- [15] M. Courtemanche, R. J. Ramirez, and S. Nattel, “Ionic mechanisms underlying human atrial action potential properties: insights from a mathematical model,” *Am. J. Physiol.*, vol. 275, no. 1, pp. H301–H321, Jul. 1998, doi: 10.1152/ajpheart.1998.275.1.H301.
- [16] J. Kneller, R. Zou, E. J. Vigmond, Z. Wang, L. J. Leon, and S. Nattel, “Cholinergic Atrial Fibrillation in a Computer Model of a Two-Dimensional Sheet of Canine Atrial Cells With Realistic Ionic Properties,” *Circ. Res.*, vol. 90, no. 9, pp. E73–E87, May 2002, doi: 10.1161/01.RES.0000019783.88094.BA.
- [17] G. Ferreira de Mattos, C. Costa, F. Savio, M. Alonso, and G. L. Nicolson, “Lead poisoning: acute exposure of the heart to lead ions promotes changes in cardiac function and Cav1.2 ion channels,” *Biophys. Rev.*, vol. 9, no. 5, pp. 807–825, Oct. 2017, doi: 10.1007/s12551-017-0303-5.
- [18] K. Andrew MacCannell, H. Bazzazi, L. Chilton, Y. Shibukawa, R. B. Clark, and W. R. Giles, “A Mathematical Model of Electrotonic Interactions between Ventricular Myocytes and Fibroblasts,” *Biophys. J.*, vol. 92, no. 11, pp. 4121–4132, Jun. 2007, doi: 10.1529/biophysj.106.101410.
- [19] C. Tobón, C. Ruiz-Villa, E. Heidenreich, L. Romero, F. Hornero, and J. Saiz, “A Three-Dimensional Human Atrial Model with Fiber Orientation. Electrograms and Arrhythmic Activation Patterns Relationship,” *PLoS One*, vol. 8, no. 2, p. e50883, Feb. 2013, doi: 10.1371/journal.pone.0050883.
- [20] E. A. Heidenreich, J. M. Ferrero, M. Doblaré, and J. F. Rodriguez, “Adaptive macro finite elements for the numerical solution of monodomain equations in cardiac electrophysiology,” *Ann. Biomed. Eng.*, vol. 38, no. 7, pp. 2331–2345, 2010, doi: 10.1007/s10439-010-9997-2.
- [21] S. E. Schober, L. B. Mirel, B. I. Graubard, D. J. Brody, and K. M. Flegal, “Blood Lead Levels and Death from All Causes, Cardiovascular Disease, and Cancer: Results from the NHANES III Mortality Study,” *Environ. Health Perspect.*, vol. 114, no. 10, pp. 1538–1541, Oct. 2006, doi: 10.1289/ehp.9123.
- [22] Y. Aoki, D. J. Brody, K. M. Flegal, T. H. I. Fakhouri, J. D. Parker, and D. A. Axelrad, “Blood Lead and Other Metal Biomarkers as Risk Factors for Cardiovascular Disease Mortality,” *Medicine (Baltimore)*, vol. 95, no. 1, 2016, doi: 10.1097/MD.0000000000002223.
- [23] J. Bernal, J. Lee, L. L. Cribbs, and E. Perez-reyes, “Full Reversal of Pb<sup>++</sup> Block of L-Type Ca<sup>++</sup> Channels Requires Treatment with Heavy Metal Antidotes,” *J Pharmacol Exp Ther*, vol. 282, no. 1, pp. 172–180, 1997.
- [24] D. V. Vassallo, E. C. Lebarçh, C. M. Moreira, G. A. Wiggers, and I. Stefanon, “Lead reduces tension development and the myosin ATPase activity of the rat right ventricular myocardium,” *Brazilian J. Med. Biol. Res.*, vol. 41, no. 9, pp. 789–795, 2008, doi: 10.1590/S0100-879X2008000900008.
- [25] R. C. Patra, A. K. Rautray, and D. Swarup, “Oxidative Stress in Lead and Cadmium Toxicity and Its Amelioration,” *Vet. Med. Int.*, vol. 2011, 2011, doi: 10.4061/2011/457327.
- [26] A. Roy, E. Queirolo, F. Peregalli, N. Mañay, G. Martínez, and K. Kordas, “Association of blood lead levels with urinary F2-8α Isoprostane and 8-hydroxy-2-deoxy-Guanosine concentrations in first-grade Uruguayan children,” *Environ. Res.*, vol. 140, p. 127, Jul. 2015, doi: 10.1016/J.ENVRES.2015.03.001.
- [27] S. Nattel and D. Dobrev, “Controversies About Atrial Fibrillation Mechanisms,” *Circ. Res.*, vol. 120, no. 9, pp. 1396–1398, Apr. 2017, doi: 10.1161/CIRCRESAHA.116.310489.
- [28] M. S. Guillem, A. M. Climent, M. Rodrigo, F. Fernández-Avilé, F. Atienza, and O. Berenfeld, “Presence and stability of rotors in atrial fibrillation: evidence and therapeutic implications,” *Cardiovasc. Res.*, vol. 109, pp. 480–492, 2016, doi: 10.1093/cvr/cvw011.

# Realidad Virtual para la rehabilitación motora de miembro inferior

A. Rojo Agustí<sup>1,2</sup>, R. Raya López<sup>1</sup>, Juan C. Moreno Sastoque<sup>2</sup>

<sup>1</sup> Universidad San Pablo-CEU, CEU Universities, Madrid, España. { }

<sup>2</sup> Grupo Neuro-Rehabilitación, Instituto Cajal, CSIC. Madrid, España.

## Resumen

La literatura reciente demostró que el ejercicio del pedaleo tiene el potencial de proporcionar un alto número de repeticiones de flexión y extensión de la extremidad inferior en períodos de tiempo razonables para mejorar la actividad muscular, la fuerza y el equilibrio en los ancianos. El empleo de tecnologías interactivas durante el entrenamiento puede promover un aumento de la adherencia y la continuidad de la actividad física. Lo cual puede llegar a tener un efecto positivo en el funcionamiento motor de los adultos mayores y los pacientes con lesión motora en miembro inferior. Aunque la importancia de la motivación en todos los ámbitos físicos es innegable, el uso específico de la tecnología de realidad virtual para la generación de un sistema de promoción de la actividad de pedaleo que sea universal para todas las estaciones de pedaleo y no requiera cables, no ha sido estudiada hasta la fecha. El presente trabajo aporta un desarrollo que permite seguir con precisión el movimiento de pedaleo del individuo en tiempo real, sin limitaciones de oclusión y baja latencia. Esta solución de realidad virtual se ha evaluado técnicamente para conocer la validez y fiabilidad de este sistema de estimación de la cadencia de pedaleo. Posteriormente, se ha realizado un ensayo experimental con población anciana y con pacientes neurológicos para evaluar las características principales de realidad virtual del sistema. A la luz de los estudios realizados, la herramienta de realidad virtual ha sido validada en términos de funcionalidad y de usabilidad.

## 1. Motivación

Los tratamientos convencionales para las personas con deterioro motor (DM) comprenden una combinación de fisioterapia, ortopedia y terapia ocupacional. Sin embargo, en términos de motivación, las terapias tradicionales presentan dos inconvenientes principales: la monotonía del ejercicio y la falta de estímulos interesantes para el paciente que aumenten su motivación. Para mitigar este problema, se han propuesto soluciones gamificadas que requieren de nuevos enfoques de diseño y uso de dispositivos de juego comerciales. El principal objetivo de este estudio es la creación de una herramienta de software interactiva que permita la realización de actividades físicas de rehabilitación de miembro inferior. Esta plataforma tendrá como fin la mejora de la movilidad y funcionalidad de las personas con patologías neurológicas o discapacidades motoras como beneficio de

la continuidad y adherencia al entrenamiento físico promovido por el sistema gamificado. Esta plataforma se ha desarrollado utilizando Oculus Quest 2 de acuerdo con los requisitos de diseño de la actividad física a realizar, el modo de uso, el sistema de análisis de movimiento y la arquitectura del sistema

## 2. Realidad virtual con ejercicio de pedaleo

Existe una considerable evidencia científica que indica que la práctica de ejercicio de pedaleo con retroalimentación visual puede mejorar el control neuromuscular, el fortalecimiento de la extremidad inferior, la recuperación funcional de la marcha y el rendimiento general. Además, varios estudios han puesto de manifiesto que las soluciones de rehabilitación basadas en entornos inmersivos son más eficaces que las tradicionales debido a tres cualidades esenciales. En primer lugar, mejoran el compromiso del paciente en la realización de la actividad [1]. Los *exergames* aumentan el gasto energético del usuario e implican tareas tanto cognitivas como fisiológicamente gratificantes. Estas actividades altamente motivadoras probablemente promuevan la adherencia al juego [2]. En segundo lugar, permiten la fidelidad física a un movimiento real. El paciente realiza movimientos similares a los que haría en una situación análoga durante su vida cotidiana. En último lugar, aportan fidelidad cognitiva a una situación real, pues el paciente debe realizar las actividades en un entorno diseñado para ser similar al mundo real.

Por ello, se propone construir un sistema que permita controlar el entorno virtual adaptando el estímulo visual a la cadencia de pedaleo estimada a partir del análisis biomecánico del movimiento de pedaleo. Para ello, se ha implementado un algoritmo de estimación de cadencia de en tiempo real integrando el sistema de captura de movimiento ENLAZA (Werium Assistive Solutions), con el dispositivo de realidad virtual (RV) Oculus Quest 2 (Facebook Technologies) [3]. Esta comunicación continua de datos cinemáticos permite realizar el análisis de la cadencia en el dispositivo de RV a la par que se renderiza el entorno digital. Como resultado, el sistema propuesto puede generar entradas sensibles para influir en la cadencia de pedaleo del usuario con el fin de mejorar el rendimiento en la actividad y fomentar los beneficios del ejercicio funcional. La principal contribución científico-técnica de este desarrollo consiste en la generación de un sistema compacto que integra un dispositivo de RV comercial de altas prestaciones con un dispositivo vestible

que permite la estimación de la cadencia de pedaleo basándose únicamente en el rango de movimiento (*range of motion*, ROM) de la cadera y es independiente del equipo de pedales.

### 3. Metodología

#### 3.1. Estudio de validación funcional del sistema RV para pedaleo

Con el desarrollo de esta herramienta denominada, PedaleoVR, se ha llevado a cabo su caracterización técnica y evaluación funcional con el fin de estudiar su fiabilidad y precisión [4]. Se realizó un estudio comparativo entre sistemas en el que participan 13 sujetos sanos que realizan 2 sesiones independientes de pedaleo que consisten en 3 minutos de pedaleo a 3 velocidades distintas. Se tomó como sistema de referencia el dispositivo *gold-standard* en la práctica clínica, MOTomed. Así, en primer lugar, se evaluó la fiabilidad del sistema a partir del análisis de la repetibilidad de medidas basado en el Índice de Correlación Intraclass (ICC). Este índice mide el grado de acuerdo o consistencia entre medidas cuyas varianzas puede ser atribuida a variaciones entre los evaluadores que realizan las medidas o entre sistemas de medida. En este caso se aplicó el test estadístico ICC (3,1), cuyo resultado representa la fiabilidad de las mediciones realizadas por un instrumento en el mismo sujeto y en las mismas condiciones y el efecto de los calificadores es despreciable.

Por otra parte, se evaluó la precisión en la estimación de cadencia a distintas velocidades: baja (30rpm), moderada (60rpm) y alta (90rpm), mediante la correlación de medidas entre ambos sistemas aplicando la correlación de Spearman.

#### 3.2. Estudio de usabilidad del sistema de RV para pedaleo en población anciana y pacientes con ataxia y hemiparesia

El siguiente estudio consistió en la evaluación de la aceptación, satisfacción, credibilidad, motivación intrínseca, usabilidad y seguridad con el fin de contrastar las características de diseño de PedaleoVR. La validación de estos aspectos se llevó a cabo en dos poblaciones con deterioro motor o lesión motora que afecta al miembro inferior. Se reclutaron 20 pacientes neurológicos del Centro Lescer, con ataxia o hemiparesia, y 18 personas mayores de 65 años residentes en el Grupo Alberta Servicios Sociosanitarios. Todos ellos realizaron una sesión experimental con PedaleoVR (ver Fig. 1.) y posteriormente completaron los siguientes cuestionarios estandarizados.

A los adultos con ataxia y hemiparesia se les consultó por la usabilidad, la credibilidad y la motivación intrínseca de la plataforma de RV a través de los cuestionarios traducidos al español: el cuestionario de Motivación Intrínseca (*Intrinsic Motivation Questionnaire*, IMI) [5], cuestionario de Credibilidad y Expectativa de un Sistema (*Credibility and Expectancy Questionnaire*, CEQ) [6] y la escala de Usabilidad de Sistemas (*System Usability Scale*,

SUS) [7]. A los adultos mayores de 65 años se les consultó por la usabilidad, la satisfacción de uso, la seguridad y/o efectos adversos y la percepción de sensación de presencia en RV. Para ello se emplearon los siguientes cuestionarios traducidos al español: la escala de efectos adversos asociados al uso de tecnología RV o cybersickness (*Simulator Sickness Questionnaire*, SSQ) [8], el cuestionario de Presencia (*Presence Questionnaire*, PQ) [9], la escala de Usabilidad de Sistemas (*System Usability Scale*, SUS) y la escala de Satisfacción de Experiencia de Usuario en juegos reducida de 18 ítems (*Game User Experience Satisfaction Scale*, GUESS -18) [10].

### 4. Resultados

#### 4.1. Resultados del estudio de validación funcional de PedaleoVR

A continuación, se presentan los resultados reportados en el estudio [4]. Las tablas 1 y 2 muestran los valores de ICC para el sistema de medida de referencia, MOTomed, y PedaleoVR. De la comparación de resultados entre ambas tablas se observa que los ICC de PedaleoVR son muy similares a los de MOTomed para cada una de las tres velocidades. A partir del análisis de estos datos se concluyó que el uso de esta plataforma de RV es adecuada para el ciclismo a velocidad baja y moderada, no recomendando su uso en velocidad altas debido a la reducción en la precisión de la estimación de la cadencia por el sistema. La extensión de estas conclusiones a la práctica de actividad física implicaría que la plataforma puede emplearse como herramienta interactiva complementaria en rehabilitación de miembro inferior.



Fig. 1. La imagen superior muestra la visualización del entorno virtual de PedaleoVR. La imagen inferior muestra el uso del sistema por un sujeto sano.

**Tabla 1. ICC (con intervalo de confianza de 95%) entre las sesiones tomadas por el sistema MOTOMed organizadas por la velocidad del objetivo.**

Velocidad	ICC	IC 95%	Significancia
30 rpm	0.743	[0.083, 0.855]	0.015
60 rpm	0.445	[-0.649, 0.825]	0.151
90 rpm	0.913	[0.719, 0.973]	≤0.001

**Tabla 2. ICC (con intervalo de confianza de 95%) entre las sesiones tomadas por PedaleoVR organizadas por la velocidad del objetivo**

Velocidad	ICC	IC 95%	Significancia
30 rpm	0.746	[0.173, 0.922]	0.014
60 rpm	0.416	[-0.871, 0.821]	0.183
90 rpm	0.851	[0.504, 0.955]	≤0.001

#### 4.2. PedaleoVR en adultos con ataxia y hemiparesia

Las puntuaciones obtenidas en las tres subescalas del cuestionario de motivación intrínseca fueron superiores a 4.5 sobre 7.0. La subescala de credibilidad del CEQ, obtuvo 18.3/27.0 puntos mientras que la subescala de expectativa obtuvo 15.0/27.0. En cuanto a la usabilidad, los participantes con ataxia y hemiparesia valoraron PedaleoVR con un 80.4/100.0 en el cuestionario de usabilidad. A partir de estas valoraciones positivas, se caracteriza PedaleoVR como una herramienta creíble y motivadora para que los adultos con trastornos neuromotores realicen ejercicio en bicicleta.

#### 4.3. Resultados del estudio de usabilidad en adultos mayores de 65

Para la evaluación de la seguridad de PedaleoVR, en términos de generación de efectos adversos o cybersickness, se tomaron los valores del SSQ previos a la exposición de la persona de edad avanzada a la RV y los SSQ posteriores. Las puntuaciones promedias de todas las subescalas del SSQ (Nausea: 3.710 ± 9.613, Oculomotor: 3.79 ± 8.089 y Desorientación 3.866 ± 9.051) se sitúan por debajo de 20 puntos en la diferencia de las evaluaciones pre y post exposición.

A partir del análisis de los resultados de las métricas PQ, GUESS-18 y SUS se observa que se ha valorado positivamente la sensación de presencia y el grado de satisfacción que genera la plataforma virtual en la población mayor de 65 años. Ambas pruebas coincidieron en destacar los aspectos de calidad de la interfaz (0,8/1,0) y realismo (0,7/1,0) del PQ con la estética visual (6,781/7,000) del GUESS-18. Asimismo, en cuanto a la aceptación y satisfacción de su uso, observando los resultados de las subescalas del cuestionario GUESS-18 se extrae que la subescala de gratificación personal (6,9/7,0), la de disfrute del juego (6,2/7,0), la de narrativa (6,4/7,0) y la de usabilidad (6,6/7,0), fueron efectivamente las mejor valoradas por las personas mayores. Con ello, y tomando la puntuación SUS de

68,5/100,0, se podría considerar que la plataforma de RV para el pedaleo obtuvo una respuesta muy favorable de aceptación de la herramienta en términos de validez y usabilidad.

## 5. Discusión

Este trabajo perseguía conseguir un desarrollo que cumpliera con las siguientes características. En primer lugar, para facilitar el manejo a estos pacientes, cuyos daños neurológicos podrían afectar también al miembro superior, se descartó el uso de mandos de RV. En consonancia con las mejoras en el software de reconocimiento de manos, que se utiliza cada vez más en el campo de las aplicaciones de rehabilitación [11]. Por ello, se implementaron las interacciones con el VE usando el seguimiento de manos y reconocimiento de gestos de Oculus (*hand-tracking*). En segundo lugar, el uso de sensores inerciales para el control de exergames se ha convertido en el más utilizado en las herramientas de entrenamiento virtual [12, 13] ya que son cada vez más baratos, su precisión es cada vez mayor. Además de estas razones, se incluye el sensor ENLAZA™ con la intención principal de incorporar un sistema de captura del movimiento de pedaleo no obstructivo durante la actividad que sea adaptable para todos los pacientes y todas las estaciones de pedaleo estacionarias.

Como resultado del primer estudio, destacamos de nuestro prototipo que no depende de ergómetros de ciclismo o hardware específicos [14] ya que estima la cadencia basada en la cinemática de la cadera del usuario en lugar del ángulo de la manivela del ergómetro igualmente [15, 16]. La facilidad y accesibilidad de la configuración establece la universalidad del sistema, siendo útil como solución de rehabilitación virtual en centros de día, clínicas de rehabilitación o incluso para terapias en casa. Además, en un futuro próximo, tenemos previsto mejorar el diseño de los escenarios virtuales para permitir diferentes usos del sensor de seguimiento, por ejemplo, para capturar los movimientos circulares de la extremidad superior. Otro aspecto positivo es que nuestro enfoque es perfectamente portátil y sencillo de montar, a diferencia de los sistemas propuestos [15, 16] que se basan en la adaptación mecánica de manivelas/pedales de ergómetros comerciales o en desarrollos de hardware a medida.

En lo que respecta al segundo estudio de usabilidad y aceptación de la tecnología, sólo se han podido realizar estadísticas descriptivas con dos poblaciones. Por este motivo, nuestros resultados deben manejarse con la máxima cautela. Los resultados de este estudio proporcionan un contexto para el uso de PedaleoVR en dos poblaciones diferentes, y describen cómo se percibe y acepta el uso de este novedoso sistema. Sin embargo, el análisis de los efectos de PedaleoVR como herramienta para mejorar el ejercicio de pedaleo y sus potenciales beneficios en pacientes con discapacidades motoras requiere más estudios, así como su implicación en el fortalecimiento de las extremidades inferiores en personas mayores. No obstante, los resultados de motivación y satisfacción concuerdan con los estudios revisados que

mostraron que los grupos de intervención con exergames estaban más motivados para hacer ejercicio [17] y encontraban el entrenamiento más atractivo que los ejercicios tradicionales [18].

Entre las limitaciones encontradas en los estudios funcionales y de usabilidad con ambas poblaciones, se destaca el reducido número de participantes con el que se ha contado en estos experimentos preliminares. También se identifica la necesidad de llevar a cabo un futuro estudio integral de PedaleoVR para evaluar el impacto de su uso en una intervención física a largo plazo y analizar sus posibles efectos en la mejora física de los usuarios.

Para concluir, resaltar que este estudio busca contribuir al ámbito de la rehabilitación apoyada en tecnologías inmersivas e interactivas, en el que todavía se están explorando estrategias de rehabilitación de miembro inferior efectivas apoyadas en estas tecnologías.

## Agradecimientos

Agradecer a los participantes y facultativos del Centro Lescer y Grupo Albertia Servicios Sociosanitarios por su amable colaboración en el proyecto de investigación, financiado por la Comunidad Autónoma de Madrid con referencia: IND2019/TIC17090 y en la colaboración al proyecto de investigación con referencia PID2021-127096OB-I00 financiado por MCIN/ AEI /10.13039/501100011033/ FEDER, UE.

## Referencias

- [1] Lohse, K. R., Shirzad, N., Verster, A., Hodges, N., & Van der Loos, H. F. (2013). Video games and rehabilitation: using design principles to enhance engagement in physical therapy. *J. neurol. Phys. Ther.* 37 (4), 166-175.
- [2] Subramanian, S., Dahl, Y., Skjæret Maroni, N., Vereijken, B., & Svanæs, D. (2020). Assessing Motivational Differences Between Young and Older Adults When Playing an Exergame. *Games Health J.* 9, 24-30.
- [3] Rojo, A., Cortina, J., Sánchez, C., Urendes, E., García-Carmona, R., & Raya, R. (2022). Accuracy study of the Oculus Touch v2 versus inertial sensor for a single-axis rotation simulating the elbow's range of motion. *Virtual Reality*, 1-12
- [4] Rojo, A., Raya, R., Moreno, J. C. (2022) Virtual reality application for real-time pedalling cadence estimation based on hip ROM tracking with inertial sensors: a pilot study. *Virtual Reality*, 1-15
- [5] Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *Am. Psychol.* 55(1), 68-78.
- [6] Deviilly, Grant, Borkovec, & Thomas. (2000). Psychometric properties of the Credibility/Expectancy Questionnaire. *Journal of Behavior Therapy and Experimental Psychiatry*, 31, 73-86.
- [7] Brooke, J. e. (1996). SUS: A quick and dirty usability scale. *Usabil. Eval. Ind.* 189 (194), 4-7.
- [8] Kennedy, R. S., Lane, N. E., Berbaum, K. S., & Lilienthal, M. G. (1993). Simulator sickness questionnaire: an enhanced method for quantifying simulator sickness. *Int J Aviat Psychol*, 3 (3), 203-220.
- [9] Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: a presence questionnaire. *Presence: teleope. Virtual Environ.* 7 (3), 225-240.
- [10] Keebler, J.R.; Shelstad, W.J.; Smith, D.C.; Chaparro, B. S.; Phan, M.H. (2020) Validation of the GUESS-18: a short version of the Game User Experience Satisfaction Scale (GUESS). *Journal of Usability Studies.* 16 (1):49.
- [11] Lange, B.; Koenig, S.; Chang, C-Y.; McConnell, E.; Suma, E.; Bolas, M.; Rizzo, A. (2012) Designing informed game-based rehabilitation tasks leveraging advances in virtual reality. *Disabil Rehabil.* 34. 1863-1870.
- [12] Saposnik, G.; Teasell, R.; Mamdani, M.; Hall, J.; McIlroy, W.; Cheung, D.; Thorpe, K. E.; Cohen, L. G.; Bayley, M. (2010). Effectiveness of Virtual Reality Using Wii Gaming Technology in Stroke Rehabilitation. A Pilot Randomized Clinical Trial and Proof of Principle. *Stroke* 41, 1477-1484.
- [13] Kosse, N.; Caljouw, S.; Vuijk, P.; CJC, L. (2011) Exergaming: interactive balance training in healthy community-dwelling elderly. *Journal of Cyber Therapy & Rehabilitation.* 4. 399-407.
- [14] Chen, S.; Fu, F.-F.; Meng, Q.-L.; Yu, H.-L. (2017) Development of a Lower Limb Rehabilitation Wheelchair System Based on Tele-Doctor--Patient Interaction. En: *Wearable Sensors and Robots.* Springer.211-223.
- [15] Ferreira, J., Coimbra, A. P., Crisóstomo, M., & Liu, T. (2020). Virtual interface for an active motorized pedal exerciser for human leg rehabilitation. *XV Mediterranean Conference on Medical and Biological Engineering and Computing – MEDICON 2019.*
- [16] Grani, F., & Bruun-Pedersen, J. R. (2017). Giro: Better Biking in Virtual Reality. *IEEE 3rd Workshop on Everyday Virtual Reality (WEVR).*
- [17] Fitzgerald, D.; Trakarnatanakul, N.; Smyth, B.; Caulfield, B. (2010) Effects of a wobble board-based therapeutic exergaming system for balance training on dynamic postural stability and intrinsic motivation levels. *J Orthop Sports Phys Ther.* 40. 1-19.
- [18] Betker, A.; Szturm, T.; Moussavi, Z. (2005) Development of an interactive motivating tool for rehabilitation movements. En: *Conf Proc IEEE Eng Med Biol Soc 2005*, 3. 2341-2344.

# Desarrollo de un encaje transfemoral para prótesis de miembro inferior asequible mediante el uso de tecnologías disruptivas

C. Bernal Benítez<sup>1</sup>, A. Pino Ramirez<sup>2</sup>, L. Gavidia Ceballos<sup>1</sup>, R. Diaz Rivas<sup>3</sup>

<sup>1</sup> Universidad Internacional de Valencia, Valencia, España, carmenbernal@outlook.com, liliana.gavidia@campusviu.com

<sup>2</sup> Universidad Simón Bolívar, Caracas, Venezuela, 16-90461@ubs.ve

<sup>3</sup> Universidad de los Andes, Mérida, Venezuela, rossiori@gmail.com

## Resumen

*La alta incidencia de reportes sobre amputados de extremidad inferior que no tiene acceso a una prótesis y el uso de tecnologías disruptivas como alternativa a los problemas que acarrea la técnica tradicional de construcción de estos dispositivos fueron las principales motivaciones de la presente investigación. Su objetivo es desarrollar un encaje o cuenca anatómica que será utilizada en una prótesis de miembro inferior para un paciente amputado transfemoral, haciendo uso de fabricación digital (escáner, software CAD-CAM e impresora FDM), y estudiar su viabilidad como método de bajo coste. A efectos de lo anterior, se realizó una revisión bibliográfica y se determinó una metodología de trabajo tecnológica experimental, enfocada en un caso clínico de paciente con amputación transfemoral. De esta forma, se modeló y fabricó un encaje específico para el amputado y se simuló tres casos de carga crítica (choque de talón, apoyo medio y despegue), así como pruebas superficiales en el paciente. Por último, se llevó a cabo una evaluación de costes de construcción de una prótesis completa y se comparó el precio final con los precios de prótesis similares del mercado actual, determinándose que es posible obtener cuencas transfemorales anatómicas para prótesis de miembro inferior a bajo coste empleando tecnologías disruptivas.*

## 1. Introducción

Acorde con un artículo reciente [1], se estima que en el mundo hay 58 millones de personas que carecen de alguna extremidad debido a la amputación de uno o más de sus miembros. De acuerdo con la Organización Mundial de la Salud (OMS) [2] el 0,5% de la población del mundo, es decir, de 35 a 40 millones de personas, requiere alguno de estos dispositivos.

Asimismo, la OMS estima que, actualmente, solo una de cada diez personas que necesitan ayudas técnicas prioritarias, incluidas las prótesis y ortesis, tienen acceso a ellos [2]. La situación es aún más desfavorable para los países de bajos y medios ingresos económicos (LMIC, por sus siglas en inglés), donde se estima que el 64% de las personas viven con amputaciones [3].

La mayoría de las amputaciones realizadas cada año son de miembros inferiores, representando más del 60% de los 1,5 millones anuales [3], por consiguiente, el reporte se centra en las prótesis de miembro inferior.

Las principales barreras que se presentan a los amputados de los países LMIC son financieras y logísticas, puesto que hay pocas unidades de servicios protésicos, y la mayoría de

la población proviene del medio rural, por lo que deben incluir altos costos indirectos para acceder al servicio [3]. Además, la fabricación tradicional de esta clase de dispositivos requiere diferentes tipos de equipos y maquinaria, tales como horno, succión al vacío y taladros, por lo tanto, es difícil establecer nuevas unidades [3].

Bajo esta perspectiva, entra en gran valor el uso de las tecnologías disruptivas como alternativa para el desarrollo de prototipos que imiten los movimientos humanos y que puedan acoplarse a los mismos, mejorando así la calidad de vida de las personas. En el campo de la protésica, estas tecnologías se centran en sustituir la maquinaria pesada por herramientas digitales, como un escáner, un ordenador portátil o una impresora 3D, con el fin de reducir costes y utilizar un archivo digital remoto que subsane las dificultades logísticas, especialmente en la fabricación de encajes adaptados al miembro residual del paciente.

En 1991, [4] se comenzó a investigar con la impresión por estereolitografía (SLS, por sus siglas en inglés), y poco tiempo después, investigadores como Lee *et al.* [5] exploraron el modelado por deposición fundida del material (FDM, por sus siglas en inglés), y en particular, más adelante, el uso de impresión 3D de baja gama [6].

En base a esta información, autores como Hervert *et al.* [7] propusieron metodologías muy similares que aglutinasen todas las técnicas digitales, es decir, reconstrucción anatómica por escáner, modelado y simulación CAD/CAM y fabricación por FDM o SL, sin embargo, no dejaron constancia del coste que implicó su realización, y, por tanto, surge la incertidumbre de si realmente el empleo de las tecnologías disruptivas para fabricar prótesis consigue abaratar los costes, desmarcándose del mercado actual.

Por ende, con el fin de estudiar la viabilidad de las tecnologías disruptivas como solución a los problemas mencionados anteriormente, en el presente trabajo se desarrolla un encaje, cuenca o socket para una prótesis de miembro inferior haciendo uso de fabricación digital (escáner, software CAD-CAM e impresora FDM), y se analiza su coste y las ventajas que presenta con respecto a la construcción tradicional.

## 2. Materiales

Los materiales empleados en la investigación son los siguientes:

- **Materias primas.** Poliuretano termoplástico (TPU) y ácido poliláctico (PLA).
- **Softwares.** Meshmixer , Ultimaker Cura y SolidWorks Simulation.
- **Maquinaria.** Escáner 3D portátil, modelo 3D Scanner Pro, de XYZ printing; y una impresora FDM de uso doméstico, modelo Creality CR10S.

### 2.1. Caso de estudio

Paciente masculino de 58 años con amputación transfemoral y muñón blando de longitud media. El paciente convive con la amputación desde 2018, utiliza muletas y su estilo de vida se clasificó dentro del nivel K2 descrito por Medicare [8]. Acorde a sus características se seleccionó un encaje de contención isquiática.

## 3. Métodos

El tipo de investigación llevada a cabo fue tecnológica experimental, basada en un caso clínico de paciente con amputación transfemoral realizado en Venezuela en el año 2021. La metodología propuesta en el proyecto combina la sucesión secuencial de acciones y la iteración de determinadas fases que requieren retroalimentación.

Para la elaboración del encaje se realizó, en primer lugar, la caracterización del paciente y establecimiento de requisitos; seguidamente, el modelado del encaje y un análisis del comportamiento mecánico, y, por último, se fabricó un prototipo y se realizaron pruebas con el paciente. Tras ello, se suceden la selección del resto de los componentes, el enlazado de todas las piezas y su presupuesto.

### 3.1. Reconstrucción tridimensional

En primer lugar, se obtuvo una imagen tridimensional del miembro residual del paciente mediante un escáner 3D portátil, colocando al sujeto un calcetín protésico suave (calcetín protésico Knit Rite A-Plus) con el fin de homogeneizar la superficie física (Figura 1).



Figura 1. Escaneo digital del miembro residual del paciente

### 3.2. Modelado

En la Figura 2 se observa el proceso de diseño definitivo, en el cual, primero, se realizó una depuración del modelo tridimensional eliminando las interferencias e irregularidades del escaneo (A), se suavizaron aquellas zonas que no se correspondían fielmente con la realidad y se creó la malla virtual de la geometría del muñón que serviría como modelo de referencia (B).

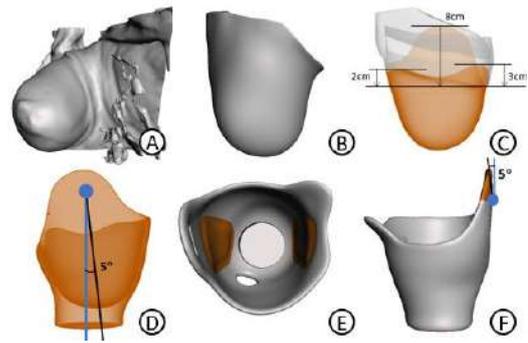


Figura 2. Modelado del encaje de contención isquiática

A continuación, siguiendo dicha referencia se modeló el encaje (C). Éste se dividió en dos partes, un marco flexible capaz de proporcionar al paciente comodidad y amortiguamiento, y un marco rígido para la transmisión del peso, similar al sistema de encaje flexible islandés sueco de Nueva York (ISNY). En el diseño se tuvo en cuenta la alineación con el resto de la prótesis considerando 5° en el plano sagital (D), y, además, se realizó un incremento de material en las caras internas medial y lateral (E) con el propósito de generar una pretensión en los músculos y encauzar el fémur en el apoyo medio [9]. Su contorno (C) se definió acorde a los parámetros recogidos en [10], y la zona comprendida entre la cabeza del fémur y la espina iliaca anterosuperior (EIAS) se rotó 5° en el plano frontal (F) para cerrar la superficie ajustándola así a la anatomía del amputado, y evitar que se separase con su peso y con ello se generase inestabilidad durante la marcha.

### 3.3. Simulación

El ciclo de marcha es un proceso dinámico, sin embargo, debido a la complejidad que conllevaba realizar este tipo de estudio, se optó por tres estudios estáticos que aglutinasen las situaciones de carga más críticas producidas durante la marcha, las fases de choque de talón, apoyo medio y despegue [11]. Se empleó un modelo simplificado del encaje rígido en el que se eliminaron los orificios y ranuras cuya influencia no era relevante en los resultados.

#### Material

Se definió como material PLA de impresión 3D, con las propiedades obtenidas de [12], y se asumió como material isotrópico elástico lineal, dado que su comportamiento plástico no era de interés.

#### Condiciones de contorno

Se estableció una sujeción de *Geometría fija* en la cara distal inferior (Figura 3, vectores verdes) representando el adaptador que se colocaría en dicha cara.

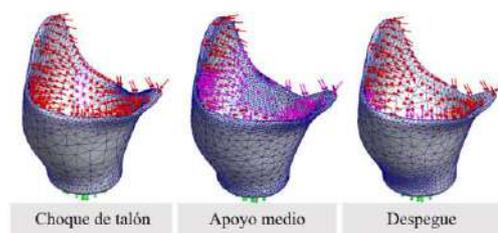


Figura 3. Casos de estudio realizados para la simulación del encaje

**Cargas**

Se incluyó una presión de contacto de 3,9 kPa en la zona distal, medial y lateral y una presión media de 0,9 kPa en el resto de la superficie (Figura 3, vectores rojos), tomándose como guía el estudio de Surapureddy [13].

Por otra parte, se aplicaron las cargas relativas al ciclo de marcha (Figura 3, vectores rosas). En la fase de apoyo medio, se asumió que todo el peso era soportado por la pierna amputada, y, por tanto, la fuerza establecida fue de 1000N (se consideró un peso de 100 kg para ir del lado de la seguridad). Además, acorde con la teoría, la carga en esta fase debía ser asumida por la parte proximal del encaje, por lo que la fuerza se distribuyó en el primer tercio de éste. En la fase de choque de talón y despegue, se consideró la mitad del peso del paciente, y, por ende, una fuerza de 500N. En tales situaciones, la diferencia se produjo en la zona de aplicación, mientras que en la fase de choque de talón el peso se concentró en la cara posterior, en la fase de despegue fue en la cara anterior.

**Mallado**

Se realizó un análisis de convergencia con el cual se definieron como tamaños de malla máximos 15 mm en la fase de apoyo medio, y 30 mm en las fases de choque de talón y despegue. Asimismo, en todas las fases se aplicó un control de mallado de 4,33 mm en las zonas que requerían mayor precisión.

**3.4. Ensamblaje**

Una vez realizada la cuenca, se obtuvieron el resto de los componentes protésicos, como eran la rodilla, los adaptadores, pilones y el pie, y se ensamblaron para formar la prótesis completa. Su selección se basó en la bibliografía anterior [14].

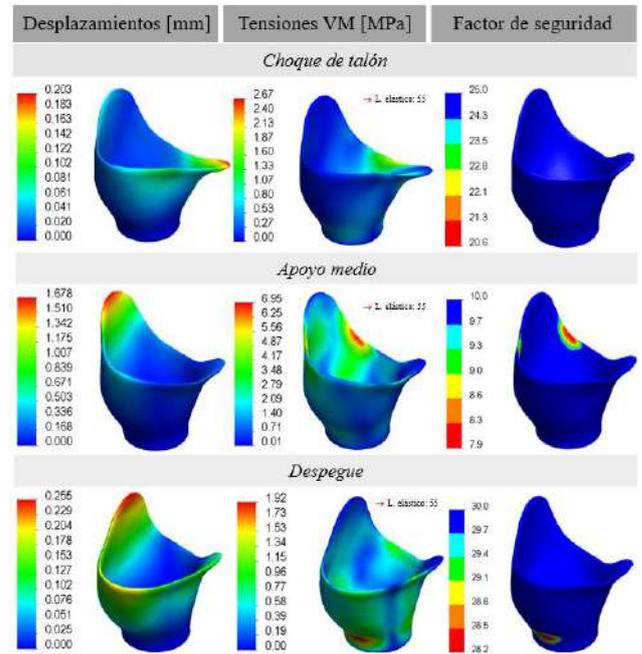
**3.5. Valoración económica**

Se tuvieron en cuenta los distintos costes asumidos en la fabricación de la prótesis desarrollada, contemplando tanto los costes de diseño como los de fabricación, y también, los precios entre los que oscilaban las prótesis con características similares en el mercado.

**4. Resultados**

Los resultados de desplazamientos, tensiones de Von Mises y distribución del factor de seguridad en cada fase analizada son apreciables en la Figura 4. En ella se puede apreciar como en la fase de choque de talón los desplazamientos son máximos en la región isquiática ( $U_{RES} = 0,203mm$ ), mientras que los esfuerzos máximos ( $\sigma_{VMmax} = 2,67MPa$ ) se originan en la parte posterior. En lo relativo a la fase de apoyo medio los desplazamientos más relevantes son en la cara lateral con un valor de 1,678mm. Asimismo, se percibe que la base del anillo y el área de mando son las partes más afectadas por las tensiones, con una máxima de  $\sigma_{VM} = 6,95MPa$ . Por otra parte, en lo que concierne a la fase de despegue, los desplazamientos más elevados se generaron en el extremo lateral del encaje ( $U_{RES} = 0,255mm$ ), y se distinguen también desplazamientos relevantes en la cara anterior y medial del componente protésico. Además, las tensiones más significativas resultaron en la cara anterior del encaje,

concretamente en la zona distal, con un valor de 1,92MPa. Por último, el factor de seguridad en todos los casos de estudios es superior a 1.



**Figura 4.** Resultados de Desplazamientos [mm], Tensiones de VM [MPa] y FDS en la fase de choque talón, apoyo medio y despegue

La Figura 5 muestra la prótesis completa y las pruebas superficiales realizadas en el paciente, quien comentó sentirse cómodo con la adaptación anatómica de la cuenca rígida y la interfaz suave, manifestando un alto nivel de satisfacción con el producto desarrollado.



**Figura 5.** Prótesis completa ensamblada (izquierda) y ajustada al paciente (derecha)

El coste final de la prótesis fue de 265 euros, mientras que el coste medio de una prótesis de características similares en el mercado (encaje básico adaptado al paciente, rodilla monocéntrica y pie tipo Sach) fue de 3200 euros, por ende, el prototipo representó un coste aproximado del 92% por debajo del precio de un producto similar en el mercado.

**5. Discusión**

En el presente proyecto se estableció como objetivo evaluar la viabilidad de desarrollo de un encaje protésico para un paciente con una amputación transfemoral utilizando tecnologías disruptivas, siendo el precio final de

la prótesis más asequible que el resto de las prótesis disponibles en el mercado.

En lo que concierne al desarrollo del encaje protésico, los resultados mostraron que pudo satisfacer de forma correcta su función. Por una parte, la caracterización tridimensional del miembro residual del paciente resultó satisfactoria para su posterior modelado pese a que requiriese una depuración inicial de interferencias. Por otro lado, la deformada que se apreció en los casos de estudio supuestos se correspondió con su predicción en base a la ubicación de la concentración de cargas, y, por consiguiente, los modelos de simulación planteados se consideraron consistentes con la realidad.

Como podía esperarse al ser la hipótesis con mayor peso, la fase de apoyo medio fue en la que se originaron mayores esfuerzos y desplazamientos, y como consecuencia de soportar la fuerza en el primer tercio del encaje, la parte lateral con menos sujeción es la que tuvo el desplazamiento más elevado. En las fases de choque de talón y despegue, los mayores esfuerzos se generaron acordes a las zonas geométricas del diseño más susceptibles para la combinación de cargas establecida, la región isquiática y la región anterior distal respectivamente. No obstante, el mayor desplazamiento en la fase de despegue, pese a no recibir fuerza, se ocasionó en la parte lateral a causa de la influencia del movimiento de la cara anterior. Por último, el factor de seguridad en todos los casos estudios analizados garantiza la estabilidad de la pieza.

El comportamiento del material concordó con las suposiciones de linealidad que fundamentaron el análisis, por lo que el error debido a las simplificaciones se trató como despreciable. Las tensiones alcanzadas resultaron muy por debajo de su límite elástico, lo que se atribuyó a la distribución de fuerza por toda la superficie seleccionada. El PLA demostró otorgar la rigidez suficiente para soportar el peso del paciente, y, por tanto, podría ser empleado en la fabricación definitiva del encaje.

No existen muchos trabajos que evalúen la capacidad de resistencia del encaje protésico, por lo que es difícil su comparativa, si bien, los órdenes de magnitud (0,003 a 6,95 MPa) son similares al encaje transtibial (0,44 a 5,068 MPa) analizado en [15].

Finalmente, los resultados obtenidos de la valoración económica y la comparativa con el precio de mercado dejaron constancia de que la prótesis obtenida cumplió el requisito de bajo coste comparada con los precios de otros fabricantes. Asimismo, los aspectos relativos al confort y adaptación del prototipo al miembro residual consultados al paciente fueron satisfactorios. No obstante, son necesarios más estudios que evalúen aspectos como la durabilidad o las cargas dinámicas.

## 6. Conclusión

Atendiendo a los resultados expuestos es posible afirmar la factibilidad de desarrollar una cuenca transfemoral anatómica para una prótesis de miembro inferior empleando tecnologías innovadoras, y, en consecuencia, conseguir reducir su precio en relación con las prótesis actuales disponibles.

## Referencias

- [1] C. L. McDonald, S. Westcott-McCoy, M. R. Weaver, J. Haagsma, y D. Kartin, «Global prevalence of traumatic non-fatal limb amputation», *Prosthet. Orthot. Int.*, 2020, doi: 10.1177/0309364620972258. Doarn CR, Adilova F, Lam D. Review of telemedicine in Uzbekistan. *Journal of Telemedicine and Telecare*, vol 11, sup 3, 2005, pp 135-9 (ISSN: 1357-633X).
- [2] Organización Mundial de la Salud, Normas de ortoprotésica de la OMS. Ginebra PP - Ginebra: Organización Mundial de la Salud.
- [3] Iniciativa Clinton de Acceso a la Salud, «Un panorama del mercado y un enfoque estratégico para incrementar el acceso a las prótesis y servicios relacionados en los países de bajos y medianos ingresos», Reino Unido, 2020. [En línea]. Disponible en: <https://forms.gle/kQdJTR9uXRj8g5aYA>.
- [4] W. E. Rogers, R. H. Crawford, J. J. Beaman, y N. E. Walsh, «Fabrication of prosthetic socket by selective laser sintering», 1991.
- [5] D. Freeman y L. Wontorcik, «Stereolithography and prosthetic test socket manufacture: A cost/benefit analysis», *JPO J. Prosthetics Orthot.*, vol. 10, n.o 1, pp. 17-20, 1998.
- [6] P. Lee, K. C. Tan, K. F. Tam, y S. L. Lye, «Biomechanical evaluation of prosthetic sockets fabricated using fused depository method», 1998.
- [7] N. Herbert, D. Simpson, W. D. Spence, y W. Ion, «A preliminary investigation into the development of 3-D printing of prosthetic sockets.», *J. Rehabil. Res. Dev.*, vol. 42, n.o 2, 2005.
- [8] J. Vargas Duque, «Aportaciones metodológicas para el diseño y fabricación de dispositivos protésicos: Aplicación en prótesis de miembro inferior», 2018.
- [9] A. P. Murillo, J. M. Martínez, y C. I. López, «PLM Strategy Definition in Product Development: Case Study on Lower Limb Sockets».
- [10] C. I. López-Gualdrón, L. E. Bautista-Rojas, y J. A. Machuca-Gelvez, «Reconstrucción 3D para el desarrollo de prótesis de miembro inferior», *Rev. UIS Ing.*, vol. 19, n.o 1, pp. 73-85, 2020, doi: 10.18273/revuin.v19n1-2020007.
- [11] Gallardo Riquelme SF, «Diseño y Fabricación de Socket Transtibial para Prótesis de Extremidad Inferior Diseño y Fabricación de Socket Transtibial para Prótesis de Extremidad Inferior», *Univ. Concepción*, pp. 1-85, 2018.
- [12] J. P. Villar, «Estudio y caracterización de materiales utilizados para la construcción de prótesis impresas mediante tecnología FDM», 2019.
- [13] R. Surapureddy, A. Schönning, y S. Stagon, «Predicting pressure distribution between transfemoral prosthetic socket and residual limb using finite element analysis», *Int. J. Exp. Comput. Biomech.*, vol. 4, n.o 1, pp. 32-48, 2016.
- [14] G. Colombo, G. Facoetti, D. Regazzoni, y C. Rizzi, «A full virtual approach to design and test lower limb prosthesis», *Virtual Phys. Prototyp.*, vol. 8, n.o 2, pp. 97-111, 2013.
- [15] C. A. Alvarado Silva, J. C. Vives Garnique, «Modelado, optimización y simulación de estructuras impresas en 3D con PLA y PET para fabricación de prótesis para pacientes de bajos recursos con amputación transtibial», 2020.

# Design and implementation of cases for serious games to train soft skills for surgical teams

LF Sánchez-Peralta<sup>1</sup>, J León<sup>2</sup>, I. Tacheci<sup>3</sup>, K. Berner-Juhos<sup>4</sup>, A. Arco<sup>5</sup>, JF Ortega<sup>1</sup>,  
E Mediero<sup>2</sup>, T. Otčenášková<sup>3</sup>, A. Ferencz<sup>4</sup>, H. Arco<sup>5</sup>, Á. Calvo<sup>2</sup>,  
FM Sánchez-Margallo<sup>6</sup>, JB Pagador<sup>1</sup>

<sup>1</sup> Bioengineering and Health Technologies Unit, Jesús Usón Minimally Invasive Surgery Centre, Cáceres, Spain. {lfsanchez, jfortega, jbpagador}@ccmijesususon.com

<sup>2</sup> Viral Studios, Badajoz, Spain, {joakilm2, enmediero, alvaro.calvo11}@gmail.com

<sup>3</sup> Charles University, Faculty of Medicine in Hradec Králové, Hradec Králové, Czech Republic, ilja.tacheci@fnhk.cz, tereza.otcenaskova@uhk.cz

<sup>4</sup> Semmelweis University, Budapest, Hungary, {juhoskriszta89, andrea.ferencz}@gmail.com

<sup>5</sup> Superior School of Health, Polytechnic of Portalegre, Portalegre, Portugal, {a.arco, helenarco}@ippportalegre.pt

<sup>6</sup> Scientific Direction, Jesús Usón Minimally Invasive Surgery Centre, Cáceres, Spain. msanchez@ccmijesususon.com

## Abstract

*Soft skills are usually neglected in surgical training, but these have been identified to be as relevant as technical ones for a safe and efficient practice. In this regard, serious games and virtual reality provide a safe environment for training, as it is possible to fully recreate a 3D environment in which actual situations in an operating room where soft skills play a role might be replayed as many times as necessary. Therefore, this paper presents the design and technical implementation of a serious game for soft skills training for surgical teams. It also includes the development of an ad hoc tool for dialogue creation, which facilitates the implementation of new cases. In all, five cases with 98 situations and 330 possible answers were developed. These cases present a story flow that varies depending on the player choices (modelled as decision trees), as it was designed as a graphical adventure game, which was implemented using the dialogue creation tool. This tool provides four different nodes: situation, question, answer, and dialogue, with a set of associated parameters, to implement the decision trees. Over 1.110 nodes (mainly dialogue and answer nodes) and 1.250 links were necessary to create the story flow of the 5 proposed cases. The presented serious game can be used as training tool for soft skills, but further validation studies are still required to prove its pedagogical value.*

## 1. Introduction

Soft skills are interpersonal, cognitive and personal resource skills that are complementary to technical skills but equally important for safe and efficient performance in surgical practice [1]. Currently, soft skills are trained following traditional approaches such as best practices, situation-based training or case studies [2].

Serious games (SG) have arisen with great success in several domains for last years. Health professionals and surgical training have also benefited from their potential, as they allow repeated training in safe environment in an amusing yet pedagogical way [3][4].

Similarly, virtual reality (VR) has become more accessible, with the potential to be used for medical education [5].

Regarding to surgical training, it allows for training in an immersive environment that do not compromise patient security nor require an actual operating room (OR) [6].

A recent review [5] identified that surgery outperforms other fields, such as anatomy, dentistry or emergency medicine in the number of VR applications for technical skills training, but none was found for communication or other soft skills. Therefore, there is currently no SG based on VR for soft skills training for surgical teams. So, this paper presents the design and implementation of a VR-based SG for soft skills training for surgical teams together with a dialogue creation tool to facilitate the implementation of new cases.

## 2. Design and implementation

In this section, the technical aspects for the development of the proposed SG are described.

### 2.1. General aspects of the serious game

The main aim of the SG is to provide surgical team members with a training tool to acquire and develop soft skills to foster their professional development. Its main learning objective is the acquisition of soft skills required for working in an OR, such as communication, teamwork or leadership. The roles involved in the OR (surgeon, nurse and anaesthesiologist) are present in the SG as 3D avatars, recreating the typical interaction during the surgical procedure.

The SG is conceived as a graphical adventure game, in which the player should make decisions based on which the story flow changes correspondingly. Decisions should be made within a predefined time limit. In this way, a twofold objective is achieved: the player is prevented from losing interest and it is approached more effectively a real case of an OR, where all decisions must be made in within an adequate response time that is not unlimited.

The 3D environment recreates a hospital hall as well as an actual OR, in which details, such as surgical instruments, laparoscopic towers or anesthetic machines were recreated to improve the feeling of immersion. In this regard, ambient audio has also been included both in the hospital hall and OR.

The perspective of the camera within the SG will depend on the case and it is either first- or third-person view. In first-person view, the camera is the same as the point of view of the player, who can look around the 3D environment. On the other hand, in the third-person view, the scene is seen globally, as an external spectator, and from the back of the main character.

Besides the character played by the player, in the case of first-person view, the rest of characters can be either interactive or non-playable. While the formers are involved in dialogues or the player’s actions have an impact on their actions, the latter appear in the SG have no specific purpose other than to improve the immersion of the player.

**2.2. Implementation**

The SG was implemented in Unity (Unity Software Inc) to be played with the Oculus Quest VR headset (Facebook Technologies LLC). The controllers will be the input devices to select options within the SG. Alternatively, a desktop version is also provided, which can be handled with the mouse.

Scenes hierarchy starts with a tutorial scene, which besides being used for control learning is also implemented as menu for cases and language selection.

Several scripts were created. The most relevant ones follow:

- **UI\_Manager:** it manages the interactive panel for case and language selections and the game loop once one case is selected (single instance), including all related aspects, such as audios, scores or context setup. It also contains the methods for translating dialogues into different languages, being English the default one.
- **Language manager:** together with the UI\_Manager, it handles language settings and translation as well as reading the information from the dialogue creation tool nodes, which will be explained in section 2.4.
- **QuestionUI:** it manages the logic of the questions, their text and which answers should be presented, as well as the countdown timer. It also supports UI\_Manager to display the selection panel.
- **AnswerUI:** similarly, it manages the logic of answers and their interface, as well as their scores.

**2.3. Content creation**

Selection of cases was done based on the premise of recreating actual situations in OR where soft skills play a significant role, for example how to handle if a team member is repeatedly making mistakes or if some surgical instrument is missing. For each case, a storyline was created, including dialogues and mood of each involved character, in form of decision tree, allowing for different endings depending on the player’s choices (Figure 1). This

way, the case can be played more than once with the possibility of having new gaming experiences each time. In the decision tree, sub-situations might be included, and the next situation after each answer is indicated.

**2.4. Dialogue creation tool**

In order to facilitate the case creation and inclusion of the dialogues in the SG, a dialogue creation tool was developed ad hoc, so it can be used in Unity for graph creation. Any case is automatically included in the interactive panel for selection.

Different types of nodes were defined:

- **Situation node (SN)** corresponds to scenes where a situation happens. Interaction with the user is not required and it is used to contextualize the dialogue or add intermediate scenes.
- **Question node (QN)** represents the case when one of several alternative questions are presented to the player.
- **Answer node (AN)** corresponds to one of the possible answers for a question node. It can contain dialogues.
- **Dialogue node (DN)** represents any type of dialogue that occurs during the SG.

Therefore, each question, situation or answer can include one or more dialogues.

Each node type has paraments to store the necessary information, that might be mandatory or optional (Table 1).

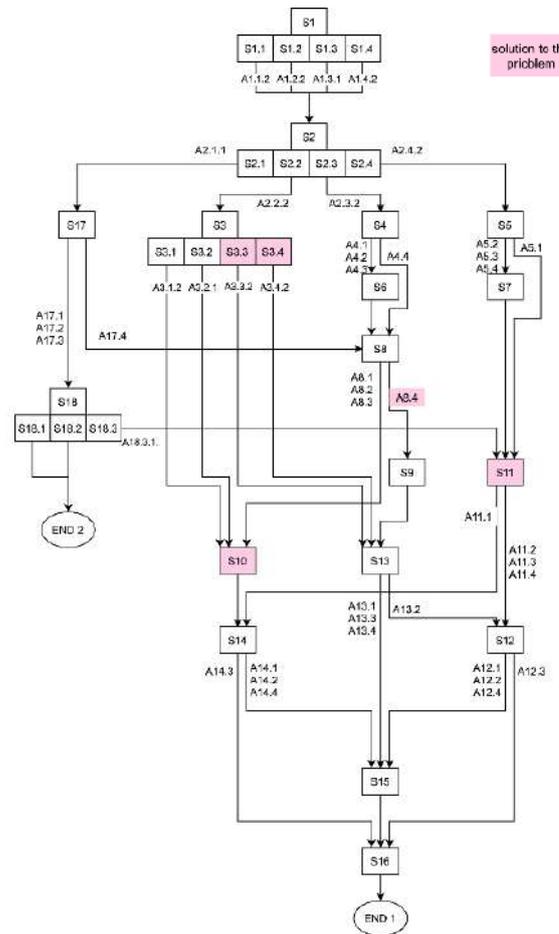


Figure 1. Example of decision tree

Node	Parameter	Description	Mandatory
Situation (SN)	SituationID	Unique identifier of situations, starting always with S1.	Yes
	Context	Description of the scenario in which the action takes place.	No
	Input port	Connection to the previous node. For S1, it will be start node.	Yes
	Dialogue port	Connection to the dialogue node that occurs in the situation.	Yes
	Question	Connection to the question node for the situation. This output can be created with the “New question” button.	No
Dialogue (DN)	DialogueID	Unique identifier of dialogues.	Yes
	Text	Field for the text.	Yes
	Speaker	Actor in scene that will say the line.	Yes
	Mood	Way in which the text is said, this will affect the movement animation that the actor will perform, the possible moods follow: Ironic, Regretful, Aggressive, Nervous and Calm. By default, the status will be Calm.	Yes
	AudioID	Name of the audio file. It should be loaded in the corresponding folder of the audio folder within resources folder in assets.	No
	Input port	Connection to previous situation, answer or dialogue node.	Yes
	Dialogue port	Connection to other dialogue node if necessary for a conversation.	No
Question (QN)	QuestionID	Unique identifier of questions.	Yes
	Description	Text of the question.	Yes
	Input port	Connection to previous situation.	Yes
	Answer port	Connection to possible answer. As many outputs as necessary can be created with the “New answer” button.	Yes
Answer (AN)	AnswerID	Unique identifier of answers.	Yes
	Feedback	Text to be shown to the user if the answer terminates the game.	No
	isEnd	Checkbox that indicates if the answer terminates the game.	No
	Score	Points obtained (positive or negative) if the answer is selected.	Yes
	Input port	Connection to previous question node.	Yes
	Dialogue port	Connection to dialogue node if necessary	No
	Situation port	Connection to next situation to be displayed	No

*Table 1. Parameters per node type*

### 2.5. Minimum requirements

The SG is optimized to be used in the Oculus Quest 2 device, but it is also compatible with Oculus Quest 1. For proper functioning of the desktop version, the minimum system requirements are an Intel Core i5 processor or similar (Intel Core i7 processor or similar recommended), at least 5GB of disk memory, 4GB of RAM (8GB of RAM recommended) and Nvidia GTX 960 or similar.

### 3. Results

Five cases have been implemented in the SG, which in all include 98 situations (including alternative sub-situations) with 330 possible answers. All cases are available in English, Spanish, Portuguese, Hungarian and Czech, to

increase acceptance of the SG as soft skills training is usually preferred in mother tongue.

The implementation of the cases using the dialogue creation tool (Figure 2) required over 1,100 nodes, being mainly situation dialogue nodes (57.87%) and answer nodes (28.80%) (Table 2).

Time to implement each case with the dialogue creation tool is directly related to the complexity of the case, being the requested time for the proposed cases around 10 minutes each. The more nodes the case requires, the longer the loading time. Maximum number of nodes for proper functioning depends on the computer being used, so a limit cannot be set universally. In this regard, designed cases with a number of nodes ranging from 110 to 331 did not present any problem in the tested devices.

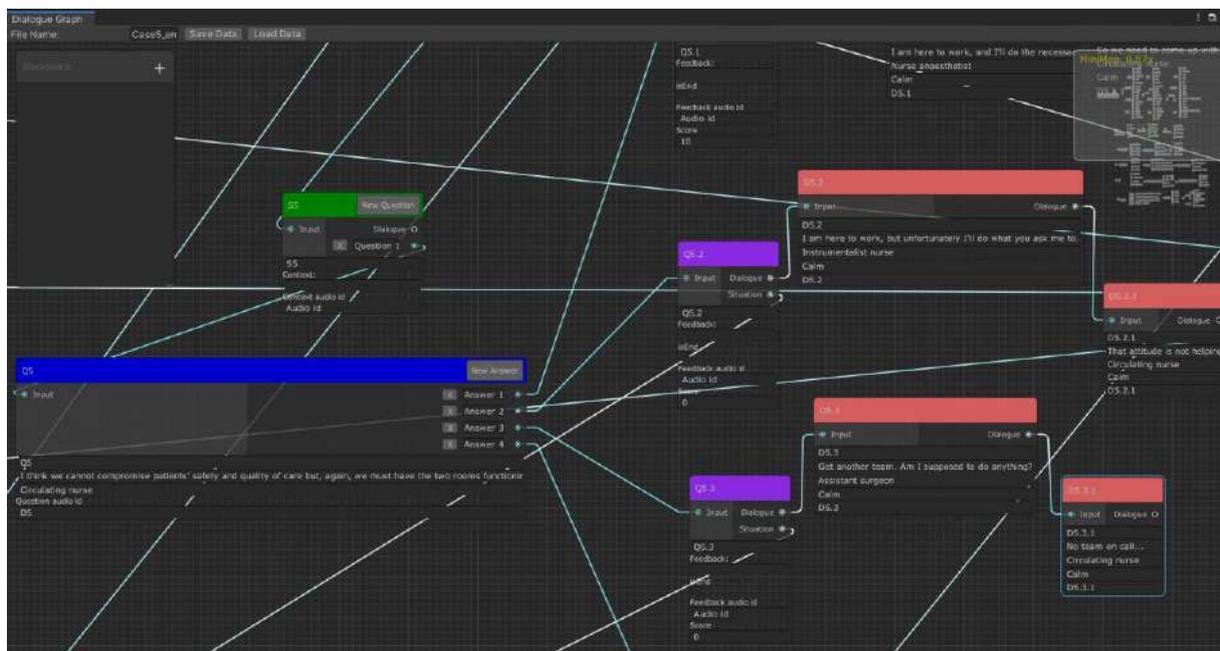


Figure 2. Example of situation from a larger graph in the dialogue tool

Case	Situations/ Answers	SN	DN	QN	AN	Links
1	11/34	11	75	8	32	147
2	33/110	18	174	29	110	416
3	26/82	7	89	24	82	275
4	19/76	19	152	19	76	305
5	9/28	9	71	6	24	127
Total	98/330	64	651	86	324	1.270

Table 2. Number of situations and nodes per case

#### 4. Conclusions

This paper presents the design and implementation of five cases for soft skills training for surgical teams by using SG and VR. The developed cases include several endings depending on the player actions, allowing for several plays without repeating the same story flow.

As future improvements, the multilanguage option could be included as node parameter for the different node types, which would facilitate even more the update of the SG to include other languages, besides the five already provided.

Nevertheless, this works has a limitation that should be acknowledged: a full validation process as well as the evaluation of the SG as a training tool are still pending. In this regard, it would be necessary to test the SG with the end users. It would be advisable to use a validated test, such as the Serious Games Evaluation Scale proposed by Fokides *et al.* [7]. It includes 53 items clustered in 12 factors that evaluate the player perception about narrative, realism, or feedback among others.

#### Acknowledgments

Project reference number: 2018-1-ES01-KA202-050943. This project has been funded with support from the

European Commission. This communication reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

This work has been also funded by the grant GR21201 from Catálogo de Grupos from Junta de Extremadura.

#### References

- [1] Agha RA, Fowler AJ, Sevdalis N. The role of non-technical skills in surgery. *Annals of Medicine and Surgery*, vol. 4, no. 4, 2015, pp 422-427 (ISSN: 2049-0801).
- [2] S4Game consortium. Handbook of soft skills training using virtual reality and serious games for surgical teams in the operating room. 2021, (ISBN:978-84-09-31179-8).
- [3] Wang R, DeMaria S, Goldberg A, Katz D. A systematic review of serious games in training health care professionals. *Simulation in Healthcare*, vol. 11, no. 1, 2016, pp 41-51 (ISSN: 1559-2332).
- [4] Graafland M, Schraagen JM, Schijven MP. Systematic review of serious games for medical education and surgical skills training. *British Journal of Surgery*, vol. 99, no. 10, 2012, pp 322-330 (ISSN: 1365-2168).
- [5] Barteit S, Lanfermann L, Bärnighausen T, Neuhann F, Beiersmann C. Augmented, Mixed, and Virtual Reality-Based Head-Mounted Devices for Medical Education: Systematic Review. *JMIR Serious Games*, vol. 9, no. 3, 2021, pp e29080 (ISSN: 2291-9279).
- [6] Pottle J. Virtual reality and the transformation of medical education. *Future Healthcare Journal*, vol. 6, no. 3, 2019, pp 181-185 (ISSN: 2514-6645).
- [7] Fokides E, Atsikpasi P, Kaimara P, Deliyannis I. Let players evaluate serious games. Design and validation of the Serious Games Evaluation Scale. *ICGA Journal*, vol 41, no. 6, 2019, pp 116-137 (ISSN: 2468-2438).

# Uso de la inteligencia artificial para la formación en cirugía mínimamente invasiva: Revisión sistemática

Francisco M. González Nuño<sup>1</sup>, Carlos Plaza de Miguel<sup>1</sup>, Juan A. Sánchez-Margallo<sup>1</sup>, Francisco M. Sánchez-Margallo<sup>1</sup>

<sup>1</sup>Centro de Cirugía de Mínima Invasión Jesús Usón, Cáceres, Spain, {fgonzalez, cplaza, jasanchez, msacnhez}@ccmijesususon.com

## Resumen

La cirugía mínimamente invasiva (CMI) está en constante evolución con la inclusión de nuevas técnicas y tecnologías, las cuales requieren de un entrenamiento más intenso hacia los cirujanos noveles. Aprovechar los avances tecnológicos para potenciar y mejorar la calidad de la enseñanza en el ámbito de la CMI debe de ser el paso más evidente. Con este fin, la inteligencia artificial (IA) presenta un gran potencial gracias a su versatilidad y diversidad. En este punto, se nos plantean varias cuestiones: ¿qué técnicas se emplean con este fin? ¿Qué rol tiene la IA en estos estudios?, las cuales pretendemos responder en la presente revisión.

## 1. Introducción y motivación

La cirugía de mínima invasión (CMI) ha revolucionado el ámbito médico y el panorama quirúrgico alrededor del mundo durante las últimas tres décadas. Esto ha sido posible gracias a distintos avances técnicos y tecnológicos como son la imagen médica o robótica quirúrgica.

Entre las distintas tecnologías interdisciplinarias emergentes, destaca la inteligencia artificial (IA), cada vez más empleada en CMI. Entre sus aplicaciones en el campo, cabe remarcar su uso en la ayuda de toma de decisiones. Sin embargo, no es su única aplicación disponible en este campo, pues otro uso posible es la mejora de la calidad de la formación de los profesionales sanitarios, y de CMI en particular.

Este tipo de cirugías, como la cirugía laparoscópica, son técnicamente muy demandantes y con una amplia curva de aprendizaje, por lo que requiere de una formación adicional a la cirugía convencional. Aparte de los métodos de formación tradicionales, como los simuladores físicos, recientemente se están aplicando nuevas tecnologías, como la mencionada IA, las cuales presentan un amplio potencial para su aplicación en el entrenamiento en CMI.

Por consiguiente, el objetivo de este estudio es analizar qué rol desempeña la IA como herramienta para la ayuda en el entrenamiento en CMI.

## 2. Metodología

Se ha realizado una búsqueda metodológica en dos bases de datos electrónicas (PubMed y *Web of Science* (WoS)) desde el 1 de junio de 2017 al 1 de junio de 2022 usando

las palabras clave “artificial intelligence” o “ai”, “training” y “surgery”, así como sus derivados.

Se seleccionaron los artículos que incluían algún tipo de inteligencia artificial y cuyo propósito sea entrenar cirujanos o alguna habilidad quirúrgica en CMI. De la búsqueda, se excluyeron artículos que no estuviesen escritos en inglés, “reviews”, “cartas”, “case reports” o “conference proceedings”.

Algunos de los aspectos analizados en esta búsqueda han sido qué rol se le otorga a la inteligencia artificial en la formación en CMI, qué técnica de IA se utiliza y cómo está diseñada la comunicación entre la IA y su entorno de trabajo/estudio.

## 3. Resultados

De los 2313 artículos resultantes de la búsqueda, fueron preseleccionados 72, de los cuales fueron finalmente incluidos 12. Este proceso se detalla siguiendo el esquema de la Figura 1.

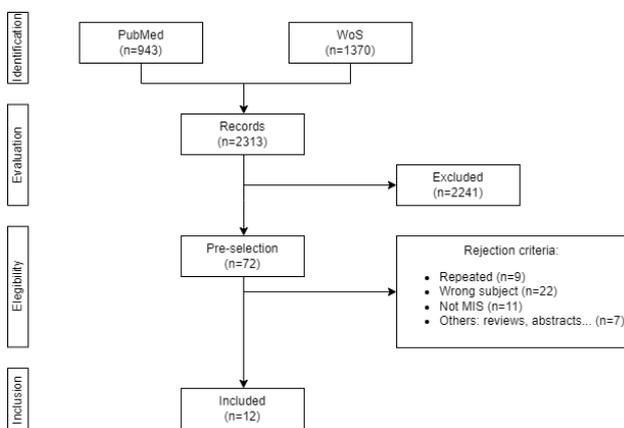


Figura 1. Flujograma para la inclusión de artículos en la revisión pertinente.

La Tabla 1 muestra un resumen de los resultados generales obtenidos en esta revisión.

Como se puede apreciar en la Tabla 1, un total de ocho estudios de los doce totales (un 67%) utilizan la IA como clasificador, principalmente de experiencia [1,3,7,9,11], también de mejora [4], de actividad quirúrgica [6] y de carga de trabajo [8].

Las aplicaciones propuestas más innovadoras pueden ser el uso de IA para el incremento de realismo en simuladores virtuales [5], y el uso combinado de distintas

técnicas de IA con distintos propósitos con el objetivo de generar trayectorias óptimas en la cirugía laparoscópica de manera demostrativa para su formación.

**Tabla 1.** Resumen de resultados de la revisión sistemática. *dVS*: sistema da Vinci; *VR*: realidad virtual; *OR*: quirófano; *KNN*: *k*-vecinos cercanos; *LR*: regresión logística; *SVM*: support vector machine; *(D/A/C)NN*: (Deep/Artificial/Convolutional) red neuronal; *GAN*: generative adversarial networks; *LSTM*: long short-term memory; *PPO*: proximal policy optimization; *GAIL*: generative adversarial imitation learning; *Mask R-CNN*: Mask Region-based convolucional neural network.

Estudio	Aplicación	Técnica de IA	Configuración	Tipos de datos	Tareas	Precisión
[1]	Clasificador de experiencia	KNN (k=2), LR, SVM	Simulador	Kinematics data/ time/ forces	Arthroscopy and basic laparoscopic tasks	84%
[2]	Predictor de alcance de nivel	DNN	Simulador dVS	Score per attempt	Basic tasks	90.4%
[3]	Clasificador de experiencia	ANN	VR Simulador	Forces and number of contacts	ACDF	83.3%
[4]	Clasificador de mejora	LR, Naïve Bayes and SVM	Simulador dVS	Engagement information	Basic Tasks	72.2%
[5]	Generación de texturas realistas	NN (Tensor flow, GANs)	VR simulador	Images from OR procedures	-	-
[6]	Clasificador de actividad quirúrgica	CNN-LSTM	OR	Endoscopic video and dVS data	Prostatectomy	85%
[7]	Clasificador de experiencia + Feedback	Decision trees; evaluación basada en reglas difusas	Simulador	Kinematics data	Basic task	-
[8]	Clasificador de carga de trabajo	Naïve Bayes	Simulador dVS	Performance's score, pupil's diameter, gaze entropy, fixation duration and PERCLOS [8]	Basic tasks	84.7%
[9]	Clasificador de experiencia	CNN-LSTM	dVS	Kinematics data	Basic tasks	97.17%
[10]	Generación de trayectoria	PPO	Simulador personalizado	-	Basic tasks	-
	Demostración	GAIL		Paths from PPO and experts		
	Seguimiento de herramientas quirúrgicas	Mask R-CNN		Endoscopic images		
[11]	Clasificador de experiencia	ANN	Simulador	Binary images with instrument's tip's path	Basic tasks	90.98%
[12]	Feedback	-	Simulador y OR	-	Ventriculostomy	-

La técnica de IA más empleada ha sido el desarrollo de redes neuronales, seguido de algoritmos habituales en inteligencia artificial tales como Naïve-Bayes o regresión logística, con la aparición de técnicas de mayor complejidad como CNN-LSTM.

La mayoría de los estudios se realizan en simulador, ya sea convencional, robótico (usando el sistema de simulación da Vinci) o virtual (83%).

De la misma manera, la mayoría de los estudios revisados realiza el estudio sobre tareas básicas (67%) en lugar de procedimientos quirúrgicos más complejos (33%).

El tipo de conjunto de datos empleado es muy diverso: los tipos más comunes son datos cinemáticos (tales como posición y movimiento), así como imágenes o puntuación obtenida de la tarea o procedimiento.

#### 4. Discusión y conclusión

La mayoría de las aplicaciones propuestas de las técnicas de IA para el apoyo a la formación en CMI, se basan en la clasificación de experiencia o la actividad quirúrgica, si bien está relacionado de manera indirecta con la educación y entrenamiento en cirugía. La mayoría de estos estudios [1,3,7,9,11] se limitan a clasificar la experiencia en un mínimo de dos categorías (novel y experto); otras aplicaciones, como el clasificador de mejora del progreso [4], el clasificador de actividad quirúrgica [6] o de carga de trabajo [8], pueden resultar de mayor utilidad a un experto que tutorice y haga seguimiento del entrenamiento correspondiente.

En cuanto a los resultados obtenidos de los clasificadores, el estudio que muestra una mayor precisión es el presentado por *Holden et al.* [7] con un 97.17% con 8

sujetos. Sin embargo, teniendo en cuenta el número de sujetos incluidos en el estudio, el mejor resultado es el obtenido por *Moglia et al.* [2], con un 90.4% y 176 sujetos.

Algunos de los estudios revisados [7,12] proponen un sistema de feedback al usuario. El primero propone un feedback posterior al procedimiento de manera que, en base a una serie de métricas simples (tiempo, ángulo entre el plano de la aguja y del ultrasonido, etc.) proporciona comentarios preestablecidos (tales como “sigue practicando con la técnica apropiada para mejorar tu eficiencia temporal”); en ausencia de errores en las métricas medidas, da un mensaje de “bien hecho”. Por otro lado, en [12] se propone un sistema de feedback basado en un módulo de inteligencia artificial, el cual ha sido definido con poco detalle. Este módulo entrena varios algoritmos con el fin de proporcionar asistencia para la mejora del procedimiento quirúrgico.

En contraste, aparecen los estudios [5] y [10], que hacen uso de estas técnicas con aplicaciones más novedosas. El propósito de *Tan et al.* [10] es quizás el más directamente relacionado con la enseñanza en CMI de los artículos estudiados. Presenta un sistema cuya finalidad es la demostración de trayectorias óptimas durante el desarrollo de un procedimiento de cirugía laparoscópica mediante la conjunción de tres técnicas distintas e independientes de IA con propósitos individuales: diseño de una trayectoria óptima, posicionamiento de las herramientas quirúrgicas en el entorno de trabajo y la demostración de la trayectoria óptima. Por otro lado, el enfoque de *Engelhardt et al.* [5] consiste en incrementar el realismo de texturas para simulaciones en entornos virtuales, las cuales son y serán cada vez más frecuentes debido al avance de esta tecnología.

Existe una gran variedad en los datos utilizados. Si bien los más comunes son imágenes, datos cinemáticos o puntuación, esta información es muy variada según los autores: *Reich et al.* [3] utiliza también el número de contactos entre instrumental y tejido como métrica; *Wu et al.* [8] hace uso del diámetro de la pupila y de la entropía de la mirada; otros autores como *Tan et al.* [10] o *Alonso-Silverio et al.* [11] requieren del recorrido del instrumento, ya sea real o generado.

Una de las mejores posibilidades que abre el uso de la IA en la enseñanza es su uso en formación personalizada a cada individuo, adaptándose a su progreso, sus errores más comunes, sus aptitudes, fortalezas y debilidades. Sería muy interesante una aplicación en este sentido.

Como trabajos futuros se plantean algunas alternativas: hacer una búsqueda más rigurosa de las distintas aplicaciones de la IA para la formación en CMI, aumentar la amplitud de la búsqueda o aumentar el rango temporal de la búsqueda.

En conclusión, la aplicación de la IA para la formación en CMI es un campo pendiente de explorar, el cual presenta mucho potencial y requiere de la sinergia de los ámbitos clínicos y técnicos. El desarrollo de ellos por separado

sería insuficiente para hacer progresar esta técnica y que llegue a ser productiva en su aplicación clínica final.

## Agradecimientos

Este estudio ha sido parcialmente financiado por la Junta de Extremadura (España), el Fondo Social Europeo, el Fondo Europeo de Desarrollo Regional (FEDER) y el Ministerio de Ciencia e Innovación de España (IG122, TA18023, CPI-2019-2033-1-TRE –14).

## Referencias

- [1] S Shedage et al., Development of Virtual Skill Trainers and Their Validation Study Analysis Using Machine Learning, *5th international conference on information system and data mining (ICISDM 2021)*, DOI: 10.1145/3471287.3471296.
- [2] A Moglia et al., Ensemble deep learning for the prediction of proficiency at a virtual simulator for robot-assisted surgery, *Surg Endosc 2022 Jan 12*, DOI: 10.1007/s00464-021-08999-6.
- [3] A Reich et al., Artificial Neural Network Approach to Competency-Based Training Using a Virtual Reality Neurosurgical Simulation, *Oper Neurosurg (Hagerstown) 2022 Jul 1*, DOI: 10.1227/ons.000000000000173.
- [4] C Wu et al., Sensor-based indicators of performance changes between sessions during robotic surgery training, *Appl Ergon 2021 Jan*, DOI: 10.1016/j.apergo.2020.103251.
- [5] S Engelhardt et al., Cross-Domain Conditional Generative Adversarial Networks for Stereoscopic Hyperrealism in Surgical Training, *MICCAI 2019*, DOI: 10.1007/978-3-030-32254-0\_18.
- [6] A Zia et al., Novel evaluation of surgical activity recognition models using task-based efficiency metrics, *Int J CARS 14 (2019)*, DOI: 10.1007/s11548-019-02025-w.
- [7] MS Holden et al., Machine learning methods for automated technical skills assessment with instructional feedback in ultrasound-guided interventions, *Int J Comput Assist Radiol Surg 2019 Nov*, DOI: 10.1007/s11548-019-01977-3.
- [8] C Wu et al., Eye-Tracking Metrics Predict Perceived Workload in Robotic Surgical Skills Training, *Hum Factors 2020 Dec*, DOI: 10.1177/0018720819874544.
- [9] XA Nguyen et al., Surgical skill levels: Classification and analysis using deep neural network model and motion signals, *Computer Methods and Programs in Biomedicine (2019)*, DOI 10.1016/j.cmpb.2019.05.008.
- [10] X Tan et al., Robot-Assisted Training in Laparoscopy Using Deep Reinforcement Learning, *IEEE Robotics and Automation Letters (2019)*, DOI: 10.1109/LRA.2019.2891311.
- [11] GA Alonso-Silverio et al., Development of a Laparoscopic Box Trainer Based on Open Source Hardware and Artificial Intelligence for Objective Assessment of Surgical Psychomotor Skills, *Surgical Innovation. 2018*, DOI: 10.1177/1553350618777045.
- [12] E Azimi et al., Interactive Training and Operation Ecosystem for Surgical Tasks in Mixed Reality, *ARE CLIP OR 2.0 ISIC 2018. Lecture Notes in Computer Science, vol 11041. Springer, Cham.*, DOI: 10.1007/978-3-030-01201-4\_3.

# Knowledge elicitation of pedagogical needs for medical learning based on 3D models

C. Guzmán García<sup>1</sup>, I. Oropesa<sup>1</sup>, M.K. Chmarra<sup>2</sup>, J.A. Sánchez Margallo<sup>3</sup>, Thomas Langø<sup>4</sup>, Calin Tiu<sup>5</sup>, Jan-Magne Gjerde<sup>6</sup>, F.M. Sánchez-Margallo<sup>3</sup>, J. Dankelman<sup>2</sup>, E.J. Gómez<sup>1,7</sup>, P. Sánchez-González<sup>1,7</sup>

<sup>1</sup> Grupo de Bioingeniería y Telemedicina, ETSI Telecomunicación, Centro de Tecnología Biomédica, Universidad Politécnica de Madrid, Madrid, España; {i.oropesa, p.sanchez, carmen.guzmang, enriquejavier.gomez}@upm.es

<sup>2</sup> Department of Biomechanical Engineering, Faculty of Mechanical, Maritime and Materials Engineering (3mE), Delft University of Technology, Delft, Países Bajos; mchmarra@gmail.com, j.dankelman@tudelft.nl

<sup>3</sup> Centro de Cirugía de Mínima Invasión Jesús Usón, Cáceres, España; {jasanchez, msanchez}@ccmijesususon.com

<sup>4</sup> SINTEF Health Research, Medical Technology, Trondheim, Noruega, thomas.lango@sintef.no

<sup>5</sup> Medis Foundation, Campina, Rumanía, tiucalin@gmail.com

<sup>6</sup> St. Olavs hospital, Trondheim, Noruega, jan-Magne.Gjerde@stolav.no

<sup>7</sup> Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina, Zaragoza, España

## Summary

*The European Alliance MIREIA aims to provide medical students and residents with a set of innovative tools and guidelines to access immersive, interactive and lifelike learning contents based on 3D models to support student-centered learning of basic technical and non-technical skills. This study presents the knowledge elicitation process of identifying and defining a list of pedagogical needs of the solutions to be developed in the MIREIA project. These pedagogical needs will be used to set the guidelines for 3D models' design, ensuring that they offer adequate affordances to implement sound learning resources. They will also serve as a reference for the functional requirements for the MIREIA Alliance.*

## 1. Motivation

Medical education is a long and demanding process, requiring learning of extensive theoretical knowledge as well as mastering a set of technical and non-technical skills [1]. Traditionally, during the early stages of education, learning methods are often based on static didactic content and, sometimes, far removed from actual clinical practice [2,3]. Currently, these methods are being replaced by new approaches based on the use of information and communication technologies [4]. New technologies, such as extended reality (XR, which includes virtual (VR), augmented (AR) and mixed reality (MR)) and three-dimensional (3D) printing, are being applied in different aspects of the medical field, including education [4].

Despite the recent popularization of these new technologies, several challenges remain to be addressed in order to extract the full didactic potential of 3D models. The first one concerns model availability. There are no widely available technologies for quick and automatic generation of intracorporeal 3D anatomical models, making the task arduous and non-accessible to anyone without a background in medical image processing. This means that models must be obtained from third parties, and that personalization to specific training needs sought for the model is lost. The second challenge concerns didactic usage of the novel immersive technologies with pedagogical guidelines in medical training, as there are no accepted EU standards to exploit them, seeking the

enhancement of knowledge transfer to actual clinical practice. Finally, the third challenge involves the scarcity of scientific evidence to support the validity of personalized 3D models as learning and training tools [5].

MIREIA is a unique Knowledge Alliance (Erasmus+, 621668-EPP-1-2020-1-ES-EPPKA2-KA) involving Higher Education institutions and companies that will combine the use of cutting-edge technology in immersive virtual technology and 3D printing with personalized learning content to promote student-centered learning for medical students and residents. This Alliance proposes the development of an innovative methodology and tools to provide interactive pedagogical contents for customized training based on 3D models, such as anatomical models (with and without pathologies) built from real-patient cases (e.g., medical imaging studies) or virtual scenarios for basic training in minimally invasive surgery.

Contents will be accessible anytime and anywhere using portable devices, XR visualization technologies, or as printed 3D models. This will allow students to train through immersive virtual environments or in physical simulators that use personalized 3D printed models. Mentors will also be able to create and share their clinical experience in the form of learning content for students including medical imaging studies, 3D anatomical models based on preoperative studies, or video sequences of surgical procedures, following specific methodological guidelines. In addition, innovative tools will be implemented for the semi-automatic creation of customized 3D models for educational purposes.

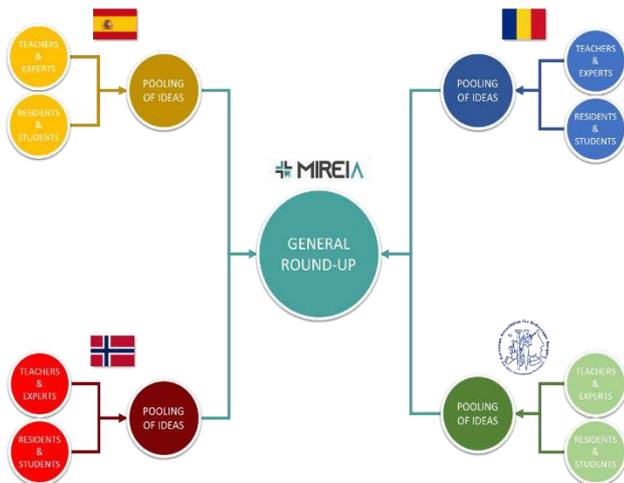
This study focuses on identifying the pedagogical needs for 3D model creation and printing, to be used by the MIREIA Alliance. To do that, the consortium designed a knowledge elicitation (KE) process consisting of workshops and questionnaires with medical stakeholders, based on the consortium know-how and previous experience [6]. The gathered knowledge will be used to set the guidelines for instructional design to create 3D models and print them guided by pedagogical sustainment, as well as serve as a reference for the functional requirements for the MIREIA Alliance.

## 2. KE Methodology

KE is composed of two steps: 1) focused discussion groups and 2) a general dissemination questionnaire, bridging both the analytic- and expert-leading KE techniques [7].

The discussion groups were carried out as 3 local workshops at clinical partner sites (Spain, Romania, Norway) as well as with the European Association for Endoscopic Surgery (EAES; <https://eaes.eu/>) (Figure 1). When possible, the workshops were carried out in native languages (except the EAES workshop, which was carried out in English), and the key findings were translated into English.

Before the workshops, we sent participants a list of questions regarding 3D models in medical training (similar to the questions posed in the workshop) to reflect upon and to prepare for the workshop. To avoid domination of the discussion on behalf of more experienced subjects, each workshop included two parallel sessions: one for experts and teachers, the other one for residents and students. To provide a safe space and encourage participation, sessions were limited to 7-8 participants maximum. At the end of each session, all participants came together to pool the ideas given by both groups and draw the final conclusions for that participation group. The conclusions extracted from each workshop were rounded up in a final workshop with representatives from the clinical sites.



**Figure 1.** Diagram of the workshop functioning. (Each session is indicated by a circle.)

Each workshop was attended by at least two members of the Alliance. A local staff member acted as director of the workshop, presenting different questions for debate to the participants and ensuring that everybody gets to participate. A secretary from the Alliance was present to take minutes.

Based on the conclusions extracted from the workshops a brief, focused questionnaire was prepared and distributed online across the surgical community. The questionnaire was divided into 5 main domains: the skills to be trained using 3D models, the means for integrating the models within curricula, the methods for 3D model creation, the barriers of 3D models with respect to their creation and integration, and the methods for interacting with the

models. This allowed us to ratify or discard some of the ideas discussed during the workshops.

## 3. Results and discussion

### 3.1. Participants

A total of 26 experts and 15 residents participated in the workshops. Participants from Romania and EAES had previous experience with 3D models. However, participants in Spain claimed to have rather little experience with both virtual and 3D printed models and experts at Norway mentioned to have experience with 3D models for clinical applications but not so much for training systems. All participants expressed their interest in prediction and planning actual procedures, rather than for training.

The questionnaire resulted in 85 responses (62.8% male respondents, 37.2% female respondents). 58.1% of respondents were from Spain, 24.4% from Norway, 2.3% from Romania, and the remaining 15.2% from different Latin American countries (i.e., Argentina, Colombia, Mexico, Ecuador, Peru and Brazil). The age groups of the respondents were 22-30 (5%), 31-45 (42%), 46-60 (29%) and over 61 (9%). The specialties of respondents varied; most of respondents specialized in *general surgery* (30.2%), followed by *urology* (17.4%) and *obstetrics and gynecology* (14.1%).

### 3.2. Skills

From the workshops, we found that the skills for which 3D models can be most beneficial are (ordered by the number of times they were mentioned during the workshops at different sites): cognitive, technical, decision making, stress management, and teamwork.

Basic technical skills were considered the most beneficial for training using virtual and printed 3D models by participants in the questionnaire, followed by cognitive skills (both related to anatomy and procedure) in the case of virtual models, and advanced technical skills and anatomical cognitive skills for printed models. This differentiation between virtual and printed models was also noted during the workshops, where 3D printed models were considered more useful for technical skills whereas virtual models could be useful for non-technical ones as well.

Soft skills (stress management, leadership, interpersonal skills) were considered to benefit less from training with virtual and printed models.

The reasoning that technical skills can benefit more from 3D models than soft skills is consistent with the literature; there are more articles including 3D models for technical skills than for soft skills [8–13].

### 3.3. Integration

Regarding the integration of 3D models within learning programs, participants of the workshops reinforced the potential of 3D models to replace animal models and cadavers, but they stated this replacement should be gradual (i.e., not forcing an immediate replacement) and it should always be done before hands-on practice. However,

participants in the questionnaire did not agree with this statement, and highlighted the usefulness of 3D models “as an aid to learn anatomical and procedural competencies”, as well as “as a complement to current learning/hands-on programs”. This shows that further research should be done as to how to best incorporate 3D models to medical training programs.

In the final workshop, a new idea surfaced with respect to the optimal model (either virtual or printed) for each learning target. Specifically, experts identified three learning targets: undergraduate, graduate and postgraduate students. Hence, virtual models would be preferred for undergraduate students (since the main focus is to engage them into medical practice and virtual simulations are indeed eye-catching), graduate students would make use of both virtual and printed models depending on the task complexity (i.e., they would start with simple models and require more complex ones as they advance in their residency), and so would postgraduate students (specifically, the study of pathological models obtained from actual patients could be beneficial for them).

Both experts and non-experts in the final workshop suggested that 3D models should be used for surgical planning, and specified the usefulness of the combination of both printed and virtual 3D models in the training for surgical actions (including sutures and resections).

### 3.4. Creation

In general, participants in the workshops suggested that the creation of 3D models should be realistic, reproducible, adaptable to the needs of trainees and trainers, and shareable. Moreover, experts stated that models should respect normal anatomy but allow to add modifications (e.g., anatomical variants). Specifically for printed models, participants in the workshops suggested that the materials used to print the models should be flexible (similar to actual tissues).

When creating 3D models, the respondents of the questionnaire suggested that the most important features to consider according to them are: the possibility to include different levels of complexity and layers of information, and the ability to interact with those models in real time.

In the final workshop, the importance of anatomical modifications was stressed, implying that pathological models could be relevant for overall training. They also suggested that CT images (or any other imaging technique, e.g., MRI) could be used to create patient-specific models. In addition, when asked about the possibility to use creation tools, surgical and medical experts were not eager to learn (and train students) to use them, while those with technical background saw great advantages in these creation tools, which would allow them to design models from 2D images from actual patients without extra expenses.

### 3.5. Barriers

Based on the results of the workshops, we have identified three main barriers, which may deter participants from using 3D models: 1) logistic (e.g., lack of printers

available, lack of areas for 3D printing labs, lack of space to place, and keep, the 3D printed models, etc.), 2) financial (e.g., lack of financial resources to buy equipment, lack of personnel specialized in 3D model creation and running a 3D print lab, etc.), and 3) technical (e.g., lack of personnel trained in 3D model creation and printing, long time required to create and print the models, etc.).

Some other barriers mentioned in the workshops were regulatory aspects (e.g., issues with privacy from patients, difficulties to adapt the learning program to include innovative models...), lack of pedagogy associated to the models, lack of secure/approved and available repositories where models are stored.

Financial and logistic issues were selected as the most limiting barriers for the use of 3D models (both virtual and printed) by respondents of the questionnaire. The next issue in the case of virtual models is technical as well as for printed models (i.e., the long time required to create the model).

In the final workshop, technical and financial barriers were agreed as the most limiting ones, although experts recognize virtual models to have less barriers than printed ones. In addition, experts remarked that virtual models account for multiple possible virtual scenarios (varying in complexity and skills trained), while printed models' applications are more limited from their design and 3D printing technology available.

### 3.6. Interaction

Participants in the workshops suggested that the interaction with printed models should be done directly with the hands. As for the interaction with virtual models, it was proposed to use hand tracking, head-set visualization with haptic feedback, or even interacting the same way as in a gaming console.

Respondents of the questionnaire highlighted that the interaction with virtual models is more suitable using haptic devices and hand tracking. One of the respondents suggested to leave the interaction method to the wishes of the trainer/trainee. A different respondent suggested to interact with the virtual model directly from the computer or the smartphone to make it more accessible.

During the final workshop, experts agreed that immersive environments are interesting due to their ability to train different skills and variability, while non-immersive ones are usually less expensive and quicker to implement. In conclusion: none of the environments is preferred over the other and they should be applied depending on the trainee's needs.

## 4. Pedagogical needs

Taking all the information extracted from the KE process, 16 pedagogical needs were detected (9 common to virtual and printed models, 5 exclusive for virtual models and 2 exclusive for printed models), as summarized in Table 1.

Common to virtual and printed models
Ability to personalize contents
Ability to access realistic contents

Ability to incorporate models within MR applications  
 Ability to include different levels of complexity and layers of information  
 Ability to interact with models in real time  
 Ability to access models at reduced prizes  
 Ability to incorporate models within AR applications  
 Ability to access to creation tools  
 Ability to adapt models to training needs

---

**Exclusive to virtual models**

---

Ability to model organs in motion  
 Ability to share and easily download models  
 Ability to interact with different tools (e.g., hand tracking, VR headset, haptic devices)  
 Ability to incorporate models within VR applications  
 Ability to freely move, rotate and scale models

---

**Exclusive to printed models**

---

Ability to reduce printing times  
 Wide range of materials for printing according to tissues' needs

---

**Table 1.** List of pedagogical needs.

Specifically, we took into account the most important features extracted from the KE process (i.e., realism, personalization and ability for real-time interaction), and the main ideas surfaced in workshops to ensure model creation and integration were easily provided to teachers and students, including the ability to share, download, and incorporate models into MR, AR and VR applications. Simultaneously, we recognized the main barriers (technical and financial) and added needs to avoid those barriers (printing time reduction, reduced prizes). It is worth to mention that there were no significant differences between people of different medical specialties.

The list of pedagogical needs will serve as a roadmap to define the guidelines to create learning contents from 3D models that will be presented in upcoming stages of the MIREIA Alliance. Moreover, they will become a key part for defining functional requirements for the MIREIA solution, ensuring its proper pedagogical foundations.

Pedagogical needs can be considered as a subgroup of user needs. However, every user need in a learning environment is going to impact the learning experience. So, in a sense, all user needs may also be considered as pedagogical needs. For the sake of clarity, we include only user needs that are related to actions that have emerged from the knowledge elicitation process. The complete set of user needs, including those with weaker pedagogical base and, as such, not incorporated here, will be considered in further stages of the project. This complete and exhaustive definition of user needs is essential for developers, who must translate them into technical (functional and non-functional) requirements to be implemented.

## 5. Conclusions

The KE process presented in this article has brought together the appropriate expertise from the MIREIA Alliance in medicine, higher education, technology and educational psychology. As a result of the KE, 16 pedagogical needs were defined falling under 3 classes: common to virtual and printed models, exclusive to virtual models, and exclusive to printed models. These results will

provide evidence-informed guidelines for the construction of the MIREIA solution.

## Acknowledgements

This work has been partially funded by the Erasmus+ Program of the European Union (621668-EPP-1-2020-1-ES-EPPKA2-KA).

## References

- [1] Yule S, Flin R, Paterson-Brown S, et al. Non-technical skills for surgeons in the operating room: A review of the literature. *Surgery*, 139(2), 2006, 140-9.
- [2] Oropesa I, Gutiérrez D, Chmarra M, et al. Can effective pedagogy be ensured in minimally invasive surgery e-learning? *Minim. Invasive Ther. Allied Technol.* **2020**, 1–11.
- [3] Sánchez-Peralta LF, Sánchez-Margallo FM, Moyano-Cuevas JL, et al. Construct and face validity of SINERGIA laparoscopic virtual reality simulator. *Int. J. Comput. Assist. Radiol. Surg.* **2010**, 5, 307–315.
- [4] Billet S. Learning through practice: models, traditions, orientation and approaches. Springer Dordrecht, 2011 (ISBN: 9789248139392)
- [5] Nickel F, Bintintan VV, Gehrig T, et al. Virtual reality does not meet expectations in a pilot study on multimodal laparoscopic surgery training. *World J. Surg.* **2013**, 37, 967–973.
- [6] Oropesa, I.; Chmarra, M.K.; Gutiérrez, D.; Sánchez-González, P.; Guzmán-García, C.; Sánchez-Peralta, L.F.; Juhos, K.; Negoita, A.; Weber, G.; Tiu, C.; et al. Knowledge elicitation of pedagogical needs for TEL-based minimally invasive surgery learning of multiple skills. In Proceedings of the XXXVI Congreso Anual de la Sociedad Española de Ingeniería Biomédica; 2018.
- [7] Gavrilova, T.; Andreeva, T. Knowledge elicitation techniques in a knowledge management context. *J. Knowl. Manag.* **2012**, 16, 523–537, doi:10.1108/13673271211246112.
- [8] Gillis CJ, Bishop N, Walsh G, et al. Evaluation of a Novel 3D-Printed Urinary Catheterization Simulation Model in Undergraduate Medical Teaching. *Cureus* **2020**.
- [9] Hyde ER, Berger LU, Ramachandran N, et al. Interactive virtual 3D models of renal cancer patient anatomies alter partial nephrectomy surgical planning decisions and increase surgeon confidence compared to volume-rendered images. *Int. J. Comput. Assist. Radiol. Surg.* **2019**, 14, 723–732.
- [10] Jacobo OM, Giachero VE, Hartwig DK, et al. 3D modeling: application in maxillofacial and hand fractures and resident training. *Eur. J. Plast. Surg.* **2018**, 41, 137-146.
- [11] Cherkasskiy L, Caffrey JP, Szweczyk AF, et al.. Patient-specific 3D models aid planning for triplane proximal femoral osteotomy in slipped capital femoral epiphysis. *J. Child. Orthop.* **2017**, 11, 147–153.
- [12] Loke YHH, Harahsheh AS, Krieger A, et al. Usage of 3D models of tetralogy of Fallot for medical education: impact on learning congenital heart disease. *BMC Med. Educ.* **2017**, 17, 54.
- [13] Mashiko T, Kaneko N, Konno T, et al. Training in Cerebral Aneurysm Clipping Using Self-Made 3-Dimensional Models. *J. Surg. Educ.* **2017**, 74, 681–689.

# Modelos de predicción de la evaluación neuropsicológica a partir de cuestionarios sobre factores percibidos de salud y estilo de vida

P. Chausa<sup>1</sup>, J. Moreno<sup>1</sup>, J. Solana<sup>2,3</sup>, G. Cattaneo<sup>2,3</sup>, J.M. Tormos<sup>2,3</sup>, P. Sánchez-González<sup>1,4</sup>, D. Bartrés-Faz<sup>2,5,6</sup>, A. Pascual-Leone<sup>2,7,8</sup>, E.J. Gómez<sup>1,4</sup>

<sup>1</sup> Grupo de Bioingeniería y Telemedicina, ETSI Telecomunicación, Centro de Tecnología Biomédica, Universidad Politécnica de Madrid, Madrid, España, [paloma.chausa/p.sanchez/enriquejavier.gomez@upm.es](mailto:paloma.chausa/p.sanchez/enriquejavier.gomez@upm.es)

<sup>2</sup> Institut Guttmann, Institut Universitari de Neurorehabilitació adscrit a la UAB, Badalona, Spain, [jsolana/gcattaneo/jmtormos@guttmann.com](mailto:jsolana/gcattaneo/jmtormos@guttmann.com); [dbartres@ub.edu](mailto:dbartres@ub.edu); [apleone@hsl.harvard.edu](mailto:apleone@hsl.harvard.edu)

<sup>3</sup> Fundació Institut d'Investigació en Ciències de la Salut Germans Trias i Pujol, Badalona, Spain

<sup>4</sup> CIBER de Bioingeniería, Biomateriales y Nanomedicina, Instituto de Salud Carlos III

<sup>5</sup> Institut d'Investigacions Biomèdiques August Pi i Sunyer, Barcelona, Spain

<sup>6</sup> Departament de Medicina, Facultat de Medicina i Ciències de la Salut, Universitat de Barcelona, Barcelona, Spain

<sup>7</sup> Hinda and Arthur Marcus Institute for Aging Research and Deanna and Sidney Wolk Center for Memory Health, Hebrew SeniorLife, Harvard Medical School, Boston, MA, USA

<sup>8</sup> Department of Neurology, Harvard Medical School, Boston, MA, USA

## Resumen

*Una mayor comprensión de los factores que influyen en la salud cerebral podría ayudar a su preservación mediante cambios en el estilo de vida de las personas, disminuyendo así la aparición y evolución de trastornos neurológicos y reduciendo el impacto social y económico de los mismos. Para crear evidencia clínica en el ámbito de la salud cerebral, la Barcelona Brain Health Initiative (BBHI) está recogiendo información demográfica, socioeconómica, medidas de salud y estilos de vida, así como imágenes médicas y evaluaciones clínicas en un estudio longitudinal y prospectivo de cohortes en el que participan más de 5.500 voluntarios. Este trabajo de investigación describe la aplicación de técnicas de machine learning a los datos recogidos en el BBHI para crear modelos de predicción que permitan inferir los resultados de las evaluaciones neuropsicológicas presenciales a partir de las respuestas a cuestionarios online autoadministrados. Estos modelos podrían ayudar a conocer los determinantes de la función cognitiva, así como proporcionar un método alternativo de valoración, más sencillo y eficaz. A partir de la evaluación de los modelos podemos concluir que su capacidad de predicción es buena (exactitud=0,75).*

## 1. Introducción

La salud cerebral, un concepto emergente y en expansión, es un estado en el que cada individuo puede darse cuenta de sus propias capacidades y optimizar su funcionamiento cognitivo, emocional, psicológico y conductual para hacer frente a las situaciones de la vida. Numerosos determinantes sociales y biológicos interconectados (incluida la genética) desempeñan un papel en el desarrollo y la salud del cerebro desde antes de la concepción hasta el final de la vida. Estos determinantes influyen en la forma en que nuestro cerebro se desarrolla, se adapta y responde al estrés y la adversidad, dando paso a estrategias tanto de promoción como de prevención.

A lo largo de la vida pueden surgir diferentes trastornos neurológicos que se caracterizan por interrupciones en el

crecimiento normal y/o el funcionamiento del cerebro. La prevalencia de la mayoría de los trastornos neurológicos incapacitantes aumenta considerablemente con la edad y su etiología y determinantes siguen sin comprenderse bien ya que en su gran mayoría involucran múltiples factores de riesgo genéticos y ambientales [1]. A medida que las poblaciones envejecen, nuestras sociedades enfrentarán desafíos cada vez mayores relacionados con los tratamientos, rehabilitaciones y consecuencias directas e indirectas de estos trastornos neurológicos. Ya en 2016 los trastornos neurológicos fueron la principal causa de discapacidad (276 millones de años de vida ajustados por discapacidad) y la segunda causa principal de muerte (9 millones) a nivel mundial [2].

Con el objetivo de avanzar en el conocimiento sobre qué factores ejercen una mayor influencia en la preservación de la salud cerebral, la iniciativa BBHI [3] iniciada en 2017, está llevando a cabo un estudio longitudinal y prospectivo de cohortes de base poblacional en el que participan más de 5.500 voluntarios entre 40 y 70 años, libres de diagnóstico facultativo de cualquier enfermedad neurológica o psiquiátrica en el momento de ingresar al estudio. Esta iniciativa recoge información demográfica, socioeconómica, medidas de salud y estilos de vida, así como imágenes médicas y evaluaciones clínicas, con el objetivo de comprender y caracterizar marcadores biológicos, conductuales y ambientales y sus interacciones, relacionados con el mantenimiento de la salud cerebral en la mediana edad. Durante la implementación del estudio, se recoge información de los voluntarios de forma periódica a través de cuestionarios autoadministrados online que evalúan los factores de estilos de vida y salud percibidos en siete dominios diferentes: salud general, ejercicio físico, actividad cognitiva, sueño, nutrición, interacciones sociales y plan vital. Un subgrupo de 1.000 personas ha participado en una serie de evaluaciones

clínicas presenciales para determinar un fenotipado profundo de la función cerebral, incluidos exámenes médicos, neurológicos y psiquiátricos, evaluación de la aptitud física, evaluaciones neuropsicológicas, imágenes de resonancia magnética cerebral estructural y funcional, electroencefalografía y evaluaciones de la actividad cerebral con estimulación cerebral no invasiva basada en perturbaciones, así como la recogida de muestras biológicas (sangre y saliva).

Este trabajo de investigación describe la aplicación de técnicas de machine learning para la definición, implementación y evaluación de modelos de predicción para inferir el resultado de la valoración neuropsicológica a partir de la información recogida en los cuestionarios online autoadministrados sobre factores percibidos de salud y estilos de vida.

## 2. Materiales y métodos

### 2.1. Fuentes de datos

Los datos de entrada de los modelos de predicción propuestos son las respuestas a los cuestionarios online autoadministrados, los cuales están accesibles de forma segura y a demanda en un data warehouse creado por la iniciativa BBHL. Para este trabajo de investigación se han utilizado todos los cuestionarios del primer año del estudio, en los que se recoge información demográfica, socioeconómica, antropométrica, presencia de diagnóstico médico y otros factores de riesgo para la salud (p. ej., tabaquismo, consumo de alcohol), así como información específica sobre los hábitos de vida en los 7 dominios mencionados anteriormente. De forma específica, las variables de entrada de los modelos predictivos son los ítems individuales de los siguientes cuestionarios: PROMIS® (Patient-Reported Outcomes Measurement Information System; 4), PROMIS® Cognitive Abilities and Cognitive Concerns [5], DASS (Escala de Depresión, Ansiedad y Estrés; 6), MeDAS (Escala de Adherencia a la Dieta Mediterránea; 7), Cuestionario de actividad física en el tiempo libre de Godin-Shepard [8], Cuestionario de evaluación del sueño de Jenkins [9], Escala Ryff de bienestar psicológico [10], Cuestionario de reserva cognitiva [11], ELS (Escala de Vida Comprometida; 12), SOC (Escala de Sentido de Coherencia; 13), SHH-Q (Self Health Horizon Questionnaire; 14), PSQI (Índice de Calidad del Sueño de Pittsburg; 15) y LSNS (Escala de la Red Social de Lubben; [16]).

Los datos de salida de los modelos de predicción se corresponden con los resultados de la evaluación neuropsicológica realizada al subgrupo de 1.000 personas participantes de la fase de evaluaciones presenciales. Las evaluaciones neuropsicológicas evalúan los siguientes dominios cognitivos: inteligencia general/fluida (WAIS-IV Matrices; 17), inteligencia premórbida (National Adult Reading Test [18]), vocabulario (fluidez verbal semántica), atención selectiva (Test de cancelación WAIS-IV; 17), funciones ejecutivas (Trial Making Test B, fluidez verbal fonémica; 19), memoria de trabajo (Dígito hacia atrás, Secuenciación de números de letras; 17), memoria episódica (RAVLT; 20), velocidad de procesamiento (sustitución de símbolos de dígitos; 17), velocidad motora

y visomotora (toque de dedos, TMT-A; 21) y habilidades visoespaciales (Block design; 17).

### 2.2. Metodología de análisis de datos

#### 2.2.1. Fase de selección y preprocesamiento de datos

Esta primera fase consiste en la comprensión y preparación de los datos para la posterior aplicación de los modelos de machine learning. En primer lugar, ha sido necesario identificar los subconjuntos de datos que forman las entradas y salidas del modelo y aplicar diferentes técnicas de preprocesamiento para eliminar posibles errores, normalizar variables y reducir la dimensionalidad.

En el preprocesamiento de las variables de entrada, se ha aplicado la técnica del umbral de varianza para identificar aquellas variables que tienen un valor alto del índice de correlación (estableciendo un umbral de 0,4). Gracias a la aplicación de este método se ha reducido el set de entrada un 20%, pasando de 441 variables a 352.

En el caso del set de salida, los resultados de la evaluación neuropsicológica se han procesado para obtener una puntuación global que describa de forma conjunta el estado de las funciones cognitivas de los participantes. Este valor se ha obtenido mediante la suma de los valores normalizados de cada test respecto a la media. Posteriormente, las puntuaciones globales se han categorizado en 4 clases. Para ello se ha normalizado el valor de puntuación global con respecto a la media y calculado la desviación estándar. La clase etiquetada como “-2” comprende los valores desde el mínimo hasta la media menos dos veces la desviación estándar. La clase “-1” va desde la frontera de la clase “-2” hasta la media. La clase “1” va desde la media hasta la media más dos veces la desviación estándar. Por último, la clase “2” va desde la frontera de la clase “1” hasta el valor máximo del conjunto. Las cuatro clases, por tanto, clasifican a las personas según el estado de su función cognitiva respecto a la media de la muestra, desde situaciones de mayor afectación (clase -2) a menor (clase 2).

Una vez los conjuntos de datos de entrada y salida han sido preprocesados, la base de datos ha sido balanceada para evitar el sobreajuste de los modelos.

#### 2.2.2. Fase de análisis y evaluación de los modelos

Esta fase consiste en la aplicación de técnicas de machine learning a los datos previamente preprocesados. Durante este trabajo de investigación se ha aplicado regresión logística (RL), K-vecinos más cercanos (KNN) y árboles de decisión. En todos los casos se ha utilizado validación cruzada y, en el caso de los árboles de decisión, se ha seleccionado el índice de Gini como medida de pureza del nodo y pre-pruning como método de optimización. El desarrollo y optimización de los modelos se ha realizado haciendo uso la librería Scikit-learn en Python, mediante la plataforma Anaconda.

La evaluación de los modelos se ha realizado en términos de precisión, sensibilidad, exactitud y F-score.

### 3. Resultados

A continuación, se muestra la matriz de confusión y los resultados de las métricas obtenidos al aplicar cada uno de los tres algoritmos propuestos. En concreto se muestra la precisión, sensibilidad y valor F obtenido para cada una de las clases de salida. Por último, se refiere el valor de exactitud global del modelo.

#### 3.1. Regresión Logística

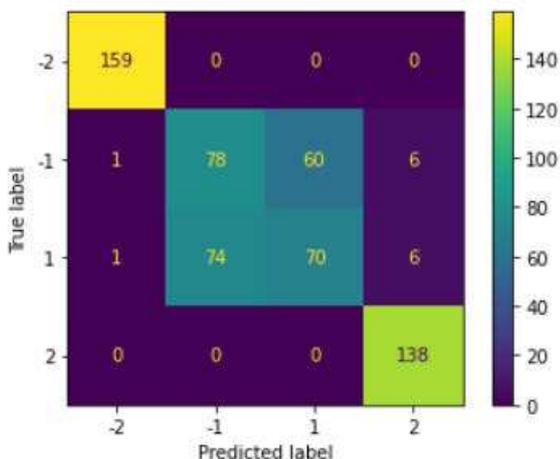


Figura 1. Matriz de confusión obtenida al aplicar RL

	P	Se	Valor F
Clase “-2”	0,99	1,00	0,99
Clase “-1”	0,51	0,54	0,53
Clase “1”	0,54	0,46	0,50
Clase “2”	0,92	1,00	0,96

Tabla 1. Resultados de las métricas de evaluación al aplicar RL

Obteniéndose una exactitud final de 0,75.

#### 3.2. KNN

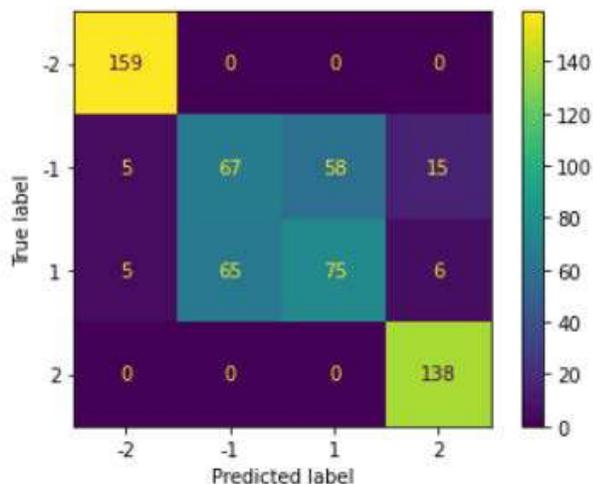


Figura 2. Matriz de confusión obtenida al aplicar KNN

	P	Se	Valor F
Clase “-2”	0,94	1,00	0,97
Clase “-1”	0,51	0,46	0,48
Clase “1”	0,56	0,50	0,53
Clase “2”	0,87	1,00	0,93

Tabla 2. Resultados de las métricas de evaluación al aplicar KNN

Obteniéndose una exactitud final de 0,74.

#### 3.3. Árboles de decisión

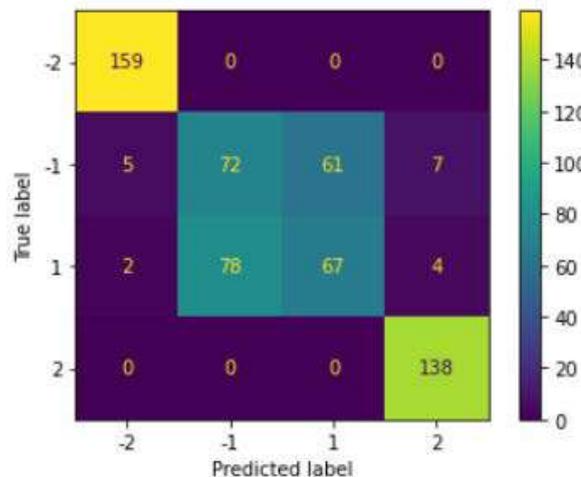


Figura 3. Matriz de confusión obtenida al aplicar Árboles de Decisión

	P	Se	Valor F
Clase “-2”	0,96	1,00	0,98
Clase “-1”	0,48	0,50	0,49
Clase “1”	0,52	0,44	0,48
Clase “2”	0,93	1,00	0,96

Tabla 3. Resultados de las métricas de evaluación al aplicar Árboles de Decisión

Obteniéndose una exactitud final de 0,74.

### 4. Conclusiones

Este trabajo de investigación describe la definición, implementación y evaluación de modelos de predicción para inferir el resultado de la valoración neuropsicológica a partir de la información recogida en los cuestionarios sobre factores percibidos de salud y estilo de vida. El objetivo final sería disponer de un método más sencillo y con menos coste de obtener una primera valoración del estado de las funciones cognitivas de una persona. Esto sería posible siempre y cuando el modelo tuviese la suficiente capacidad de predicción. Los resultados obtenidos demuestran que los modelos implementados poseen una buena capacidad de predicción, que podría ser mejorada mediante la exploración de nuevos procedimientos de extracción y reducción de

características, la aplicación de otras técnicas de balanceo y análisis y/o el estudio de diferentes categorizaciones de la variable de salida. De forma concreta, se propone la optimización de los modelos basados en árboles de decisión mediante la aplicación de técnicas de bagging y boosting.

Durante la implementación de los modelos de predicción se han aplicado una serie de procedimientos para preprocesar los datos disponibles. Se ha realizado una reducción de dimensionalidad de los datos de entrada mediante la técnica de umbral de varianza. También se han elaborado modelos que reducen la dimensionalidad mediante la selección manual de las variables e incluso sin reducción alguna. Los resultados de estos modelos alternativos han sido muy similares a los mostrados en este artículo lo que demuestra que la reducción de dimensionalidad realizada nos permite simplificar el modelo sin perder capacidad de predicción, sobre todo para el caso de uso de los algoritmos de KNN y árboles de decisión.

Una de las limitaciones de los modelos de predicción implementados es su falta de interpretabilidad, algo de especial importancia en el área de la salud. Para mejorar la interpretabilidad de los resultados y conocer qué variables influyen en mayor medida en el mantenimiento de la salud cerebral y más en concreto en la capacidad cognitiva de las personas, se podrían aplicar técnicas de selección automática de variables y métodos que permitan conocer qué variables tienen más peso, como la regresión de Lasso. Por otro lado, se ha planteado la posibilidad de realizar un análisis de componentes principales en los resultados de las evaluaciones neuropsicológicas para poder implementar modelos específicos para cada uno de los dominios cognitivos resultantes, lo que podría ayudar a tener una valoración más precisa.

## Agradecimientos

Esta investigación ha recibido financiación de “La Caixa” Foundation (grant number LCF/PR/PR16/11110004), del Institut Guttmann y de la Fundació Abertis.

## Referencias

- [1] GBD 2016 Neurology Collaborators. Global, regional, and national burden of neurological disorders, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Neurol*, 18 (2019), pp. 459–480
- [2] Dumurgier, J., & Tzourio, C. (2020). Epidemiology of neurological diseases in older adults. *Revue Neurologique*, 176(9), 642–648. <https://doi.org/10.1016/j.neurol.2020.01.356>
- [3] Cattaneo G., Bartrés-faz D., Morris T. P., Solana Sanchez J., Maciá D., Tarrero C, et al. (2018). The Barcelona Brain Health Initiative: A Cohort Study to Define and Promote Determinants of Brain Health. *Frontiers in Aging Neuroscience*, 10(October), 321.
- [4] Ader, D. (2007). Developing the patient-reported outcomes measurement information system (PROMIS). *Med. Care*
- [5] Fieo, et al. (2016). Measurement Equivalence of the Patient Reported Outcomes Measurement Information System (PROMIS) Applied Cognition – General Concerns, Short Forms in Ethnically Diverse Groups. *Psychological test and assessment modeling*. 58. 255–307. 45(5 Suppl. 1), S3–S11. doi: 10.1097/01.mlr.0000260537.45076.74
- [6] Brown, T. A., et al. (1997). Psychometric properties of the depression anxiety stress scale (DASS) in clinical samples. *Behav. Res. Ther.* 35, 79–89. doi: 10.1016/S0005-7967(96)00068-X
- [7] Schröder, H., et al. (2010). A short screener is valid for assessing Mediterranean diet adherence among older. *J. Nutr.* 141, 1140–1145. doi: 10.3945/jn.110.135566
- [8] Godin, G., and Shephard, R. J. (1997). Godin leisure-time exercise questionnaire. *Med. Sci. Sports Exerc.* 29, 14–15
- [9] Jenkins, C. D., et al. (1988). A scale for the estimation of sleep problems in clinical research. *J. Clin. Epidemiol.* 41, 313–321. doi: 10.1016/0895-4356(88)90138-2
- [10] Díaz, D., et al. (2006). Adaptación española de las escalas de bienestar psicológico de Ryff. *Psicothema*, 18(Número 3), 572–577
- [11] Rami, L., et al. (2011). Cognitive reserve questionnaire. Scores obtained in a healthy elderly population and in one with Alzheimer’s disease. *Rev. Neurol.* 52, 195–201
- [12] Trompetter, H. R., et al. (2013). Measuring values and committed action with the engaged living scale (ELS): psychometric evaluation in a nonclinical sample and a chronic pain sample. *Psychol. Assess.* 25, 1235–1246. doi: 10.1037/a0033813
- [13] Eriksson M, Lindström B. Antonovsky's sense of coherence scale and the relation with health: a systematic review. *J Epidemiol Community Health.* 2006 May;60(5):376–81. doi: 10.1136/jech.2005.041616. PMID: 16614325; PMCID: PMC2563977.
- [14] Düzel, S., et al. (2015). The subjective health horizon questionnaire (SHH-Q): assessing future time perspectives for facets of an active lifestyle. *Gerontology* 62, 345–353. doi: 10.1159/000441493
- [15] Buysse, D. J., et al. (1989). The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry Res.* 28, 193–213. doi: 10.1016/0165-1781(89)90047-4
- [16] Lubben, J. E. (1998). Assessing social networks among elderly populations. *Fam. Commun. Health* 11, 42–52. doi: 10.1097/00003727-198811000-00008
- [17] Weschler, D. (2008). Wechsler Adult Intelligence Scale, 4th Edn. San Antonio, TX: NCS Pearson.
- [18] Bright, P., Jaldow, E., & Kopelman, M. D. (2002). The National Adult Reading Test as a measure of premorbid intelligence: A comparison with estimates derived from demographic variables. *Journal of the International Neuropsychological Society*, 8(6), 847–854. <https://doi.org/10.1017/S1355617702860131>
- [19] Reitan, R. M., and Wolfson, D. (1985). The Halstead–Reitan Neuropsychological Test Battery: Therapy and Clinical Interpretation. Tucson, AZ: Neuropsychological Press.
- [20] Rey, A. (1958). L’examen Clinique en Psychologie [The Psychological Examination]. Paris: Presses Universitaires de France.
- [21] Shimoyama, I., Ninchoji, T., and Uemura, K. (1990). Fingertapping test. *Arch. Neurol.* 47, 681–684. doi: 10.1001/archneur.1990.00530060095025

# Modelos de predicción del pronóstico neurológico de pacientes con hemorragia subaracnoidea (HSA) con ventilación asistida

H. Araujo Rodríguez<sup>1</sup>, L.E. Mujica Delgado<sup>1</sup>, R. Mellado-Artigas<sup>2</sup>, L. Zattera<sup>2</sup>, L. Llull<sup>3</sup>, S. Amaro<sup>3</sup>, M. Ruiz Ordoñez<sup>1</sup>

<sup>1</sup> Departament de Matemàtiques, CoDALab, Escola d'Enginyeria de Barcelona Est (EEBE), Universitat Politècnica de Catalunya, Barcelona, España, {helena.maria.araujo, luis.eduardo.mujica, magda.ruiz}@upc.edu

<sup>2</sup> Departamento de Anestesiología, Hospital Clínic, Institut D'investigació August Pi i Sunyer, Barcelona, España, {rmellado, lzattera}@clinic.cat

<sup>3</sup> Servicio de Neurología, Hospital Clínic, Barcelona, España, {bllull, samaro}@clinic.cat

## Resumen

*La hemorragia subaracnoidea (HSA) es un tipo de ictus poco común cuya aparición puede causar lesiones permanentes o la muerte del paciente. Este trabajo tiene como finalidad la creación de una herramienta de apoyo a la toma de decisiones para el personal médico de la UCI sobre el tratamiento de pacientes con HSA según su pronóstico neurológico. La base de datos empleada fue proporcionada por el Hospital Clínic de Barcelona y consta de un total de 130 pacientes con un nivel de gravedad en la World Federation of Neurosurgical Societies Scale (WFNS) de 4 o 5, donde cada paciente tiene 133 variables. En un principio se realiza un estudio univariable para determinar aquellas variables que podrían permitir la predicción del pronóstico neurológico. Finalmente, se realiza un análisis multivariable en el que se empleará un modelo de Machine Learning (ML) compuesto por los algoritmos de Random Forest, Support Vector Machine, Regresión Logística y K-Nearest-Neighbors, obteniendo una exactitud de 0.8564 a día cero del paciente en el hospital y de 0.9259 considerando el efecto de la intubación en el paciente. De esta manera, se ha empleado el conocimiento médico y el ML para obtener nuevas opciones de ayuda al personal sanitario que sigan siendo de utilidad a medida que pasan los días en los que el paciente se mantiene en ventilación mecánica.*

## 1. Introducción

La hemorragia subaracnoidea (HSA) es una afección neurológica que supone un 5% de los accidentes cerebrovasculares (ACV), afectando en torno a 10 de cada 100,000 individuos por año [1]. Esta patología suele tener un pronóstico neurológico negativo sobre el conjunto de pacientes que la sufre. Debido a ello, incluir en el diagnóstico los datos del paciente y sus constantes médicas a su ingreso en UCI podría ser de ayuda a los médicos para la toma de decisiones a la hora de aplicar un tratamiento. Por otra parte, dentro del ámbito de Ingeniería Biomédica, el mundo del Machine Learning (ML) como herramienta para el análisis de datos y creación de modelos analíticos y predictivos está

aumentando considerablemente su importancia en los últimos años. Actualmente no existen trabajos que enfoquen el empleo de ML en esta patología. De esta forma este trabajo combian varias ramas de la ingeniería que fusionen el análisis de estos datos de forma más profunda comprendiendo la parte fisiológica de la enfermedad que además incluya el efecto de la ventilación mecánica. Al final se obtienen algoritmos automáticos (en Rstudio) que podrían ayudar a la toma de decisiones para el pronóstico médico.

## 2. Metodología

### 2.1. Preprocesamiento de datos

La base de datos proporcionada por el Hospital Clínic de Barcelona consta de un total de 130 pacientes con HSA grave, medida por una puntuación de 4 o 5 en la escala del World Federation of Neurosurgical Societies Scale (WFNS) [2].

Los datos registrados de todos estos pacientes incluyen tanto datos demográficos como información sobre las constantes vitales de estos a día 0, cronologías o índices de gravedad. Por último, contiene el estado de los pacientes una vez han salido de la UCI y del hospital y su ranking neurológico a 3 meses vista (**rankin\_3\_meses**). Este conjunto de datos iniciales precisa de un preprocesamiento en el que se eliminan manualmente en base a criterios clínicos aquellas variables que carezcan de interés para el estudio o posean gran cantidad de datos faltantes. Tras la limpieza, la base de datos queda con 38 variables predictoras. La variable de respuesta (*outcome*) de este estudio es una dicotomización de la variable **rankin\_3\_meses** cuyo rango va de los valores 0 (vivo sin secuelas) hasta 6 (muerte). Se ha clasificado de manera que de 0 a 3 el estado del paciente sea bueno (“good”) y que de 4 a 6 el paciente haya sufrido un gran número de secuelas e incluso la muerte, clasificándose como “poor”, siendo la variable resultante **rankin.cat**.

## 2.2. Análisis univariable

De forma previa al empleo de ML para la creación de la herramienta de apoyo a la toma de decisiones es necesario conocer la población estudiada. Con el *dataframe* ya preprocesado, se procede a realizar un análisis univariable basado en estadística descriptiva que permitirá conocer al grupo de pacientes estudiado y su distribución.

## 2.3. Tabla de demográficos

Un total de 40 pacientes tuvieron un pronóstico neurológico positivo y 90 de ellos uno negativo, siendo este más del doble de pacientes. Esta diferencia entre ambos grupos puede traer problemas a la hora de entrenar la herramienta por lo que será necesario el remuestreo del grupo menos extenso.

Variable	Niveles	rankin.cat		p-valor		
		good	poor			
death.on.mv	0	40 (100.0%)	33 (36.7%)	<0.001		
	1	0 (0.0%)	57 (63.3%)			
Días_ingreso_hospita- lario	Median	32.0	6.0	<0.001		
	(IQR)	(25.0 to 47.0)	(1.0 to 31.2)			
Días_ingreso_en_UCI	Median	14.0	5.0	<0.001		
	(IQR)	(7.8 to 19.0)	(1.0 to 14.2)			
LET_de_entrada	0	40 (100.0%)	62 (68.9%)	<0.001		
	1	0 (0.0%)	28 (31.1%)			
Edad	Median	53.0	60.5	0.001		
	(IQR)	(47.8 to 62.0)	(53.0 to 77.2)			
CTPY_N <sup>4</sup>	0	20 (50.0%)	70 (77.8%)	0.002		
	1	20 (50.0%)	20 (22.2%)			
Ingreso_UCI	0	0 (0.0%)	14 (15.6%)	0.005		
	1	40 (100.0%)	76 (84.4%)			
Creatinina	Median	0.7	0.8	0.007		
	(IQR)	(0.6 to 0.8)	(0.7 to 1.0)			
WFNS_ingreso <sup>5</sup>	4	16 (40.0%)	17 (18.9%)	0.016		
	5	24 (60.0%)	73 (81.1%)			
GCS_ingreso <sup>6</sup>	3	13 (32.5%)	51 (56.7%)	0.019		
	4	2 (5.0%)	9 (10.0%)			
	5	2 (5.0%)	4 (4.4%)			
	6	3 (7.5%)	6 (6.7%)			
	7	1 (2.5%)	7 (7.8%)			
	8	6 (15.0%)	4 (4.4%)			
	9	3 (7.5%)	3 (3.3%)			
	10	3 (7.5%)	2 (2.2%)			
	11	3 (7.5%)	2 (2.2%)			
	12	2 (5.0%)	2 (2.2%)			
	13	0 (0.0%)	0 (0.0%)			
	14	2 (5.0%)	0 (0.0%)			
	15	0 (0.0%)	0 (0.0%)			
	Glicemia	Median	151.0		177.0	0.022
		(IQR)	(132.0 to 187.0)		(140.5 to 212.5)	
Tratamiento_ endovascular	0	12 (30.0%)	47 (52.2%)	0.022		
	1	28 (70.0%)	43 (47.8%)			
Monocitos	Median	0.5	0.6	0.038		
	(IQR)	(0.4 to 0.7)	(0.4 to 1.4)			
DVE <sup>7</sup>	0	10 (25.0%)	38 (42.2%)	0.077		
	1	30 (75.0%)	52 (57.8%)			
Pérdida_conciencia_i- nicio	0	11 (27.5%)	12 (13.3%)	0.079		
	1	29 (72.5%)	78 (86.7%)			
Plaquetas	Median	235.0	217.0	0.079		
	(IQR)	(198.2 to 299.0)	(182.0 to 248.8)			
Tratamiento_quirúr- gi-co	0	29 (72.5%)	78 (86.7%)	0.079		
	1	11 (27.5%)	12 (13.3%)			
mv.days <sup>8</sup>	Median	8.0	4.0	0.109		
	(IQR)	(2.5 to 13.0)	(1.0 to 10.0)			

Para poder conocer la demografía del conjunto de pacientes estudiados se realiza un estudio demográfico de las diferentes variables separadas por la variable de salida (**rankin.cat**), añadiendo el criterio del test de *Fisher Least discriminant* para el cálculo del *p-valor* con nivel de significación 0.05 para conocer si se relacionan con el *outcome*.

En la Tabla 1 se puede observar que a mayor número de días de hospitalización, el pronóstico es mejor. Los niveles de las variables **Monocitos**, **Creatinina** y **Glicemia** (en las que unos mayores niveles se relacionan con un pronóstico “*poor*”) de los pacientes que tienen un pronóstico “*good*” son ligeramente inferiores.

Variable	Niveles	rankin.cat		p-valor
		good	poor	
Aneurisma_dic	0	5 (12.5%)	8 (8.9%)	0.119
	1	35 (87.5%)	74 (82.2%)	
	2	0 (0.0%)	8 (8.9%)	
Traqueostomía	0	25 (62.5%)	69 (76.7%)	0.136
	1	15 (37.5%)	21 (23.3%)	
Sexo	Female	28 (70.0%)	52 (57.8%)	0.242
	Male	12 (30.0%)	38 (42.2%)	
Troponina_I_24h	Median	0.1	0.3	0.301
	(IQR)	(0.0 to 2.3)	(0.0 to 4.8)	
Tabaquismo	0	24 (60.0%)	62 (68.9%)	0.325
	1	16 (40.0%)	28 (31.1%)	
Troponina_I_ingreso	Median	0.1	0.1	0.373
	(IQR)	(0.0 to 2.0)	(0.0 to 2.9)	
Hidrocefalia_ingreso	0	10 (25.0%)	30 (33.3%)	0.413
	1	30 (75.0%)	60 (66.7%)	
Leucocitos	Median	13.4	13.3	0.472
	(IQR)	(10.1 to 16.6)	(10.7 to 18.5)	
Hemoglobina	Median	13.7	13.4	0.558
	(IQR)	(12.4 to 14.8)	(12.2 to 14.8)	
Potasio	Median	3.8	3.7	0.605
	(IQR)	(3.4 to 4.2)	(3.4 to 4.2)	
Cefalea_centinela	0	31 (77.5%)	74 (82.2%)	0.630
	1	9 (22.5%)	16 (17.8%)	
Sodio	Median	140.0	140.0	0.641
	(IQR)	(139.0 to 142.2)	(138.0 to 143.0)	
aneurysm.origin	Anterior	14 (35.0%)	22 (24.4%)	0.645
	Carotid	3 (7.5%)	13 (14.4%)	
	Middle	6 (15.0%)	13 (14.4%)	
	Posterior	12 (30.0%)	26 (28.9%)	
	Unknown	5 (12.5%)	16 (17.8%)	
Hematocrito	Median	41.0 (37.8 to 45.0)	41.0 (38.0 to 44.0)	0.649
	(IQR)			
IOT_VM <sup>9</sup>	0	1 (2.5%)	5 (5.6%)	0.666
	1	39 (97.5%)	85 (94.4%)	
Rankin_previo	0	36 (90.0%)	70 (77.8%)	0.674
	1	3 (7.5%)	11 (12.2%)	
	2	1 (2.5%)	7 (7.8%)	
	3	0 (0.0%)	1 (1.1%)	
	4	0 (0.0%)	0 (0.0%)	
	5	0 (0.0%)	1 (1.1%)	
Crisis_epilépticas_ini- cio	0	33 (82.5%)	77 (85.6%)	0.793
	1	7 (17.5%)	13 (14.4%)	
	2	0 (0.0%)	2 (2.2%)	
mFisher <sup>10</sup>	3	2 (5.0%)	3 (3.3%)	0.832
	4	38 (95.0%)	85 (94.4%)	
	0	37 (92.5%)	83 (92.2%)	
Alcohol	0	37 (92.5%)	83 (92.2%)	1.000
	1	3 (7.5%)	7 (7.8%)	

Tabla 1. Demografía de la población estudiada

<sup>4</sup> TAC de perfusión, <sup>5</sup> World Federation of Neurosurgical Societies Scale, <sup>6</sup> Glasgow Coma Scale, <sup>7</sup> Drenaje ventricular, <sup>8</sup> Días de ventilación mecánica, <sup>9</sup> Intubación del paciente, <sup>10</sup> Fisher modificado.

Adicionalmente, se demuestra que un mayor número de días de ventilación mecánica (**mv.days**) parece aumentar las posibilidades de tener un buen pronóstico. Esta variable a pesar de poseer un p-valor bajo (0.109) se incluye para estudiar el efecto del tiempo de intubación sobre los modelos.

Para complementar el estudio de demográficos, se calcula la matriz de correlación (no mostrada), la cual ratifica la información señalada por la Tabla 1. En conclusión se observa que que las variables relativas a las escalas **WFNS** y el *Glasgow Coma Scale* (**GCS**) se encuentran altamente correlacionadas, así como los niveles de las variables **Hemoglobina** y **Hematocrito** o **Hidrocefalia\_ingreso** y **DVE**. Teniendo el criterio médico, las variables Hematocrito y DVE se incluyen en la creación de modelos de *Machine Learning*.

### 2.2.2. Gráfico de distribución

Para poder comprender de forma más visual las distribuciones de variables en el análisis univariable se realizan gráficos de densidad de las variables numéricas (Figura 1) e histogramas de las categóricas (Figura 2) para conocer su distribución dependiendo de la variable de salida, **rankin.cat**.

Con este análisis se puede observar (confirmando así lo aportado por el p-valor) que en ninguna variable existe una división clara que permita conocer por sí sola aquellos pacientes que podrán tener un pronóstico neurológico bueno o grave y por lo tanto es necesario un análisis multivariable de los datos basado en técnicas de *Machine Learning* creando una herramienta de apoyo a la toma de decisiones en la UCI.

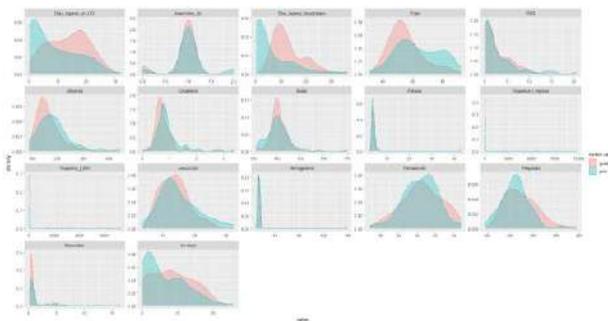


Figura 1. Distribución variables numéricas

### 2.3. Imputación de datos

En la base de datos hay un total hay 100 datos faltantes suponiendo un 2% de los datos (100 de 5070 datos totales). La mayoría de los pacientes (69 de 130) tienen todos la información completa y el número máximo de datos faltantes por paciente es de cuatro por lo que no es necesario eliminar ningún paciente del estudio. Se utiliza el método de imputación multivariable por medio de ecuaciones encadenadas (multivariate imputation by chained equations -MICE) [3]. Se estiman 10 valores por cada valor faltante, las nuevas bases de datos se complementan utilizando las medidas de la media, moda y mediana. Quedando finalmente esta última ya que es más insensible ante valores extremos.

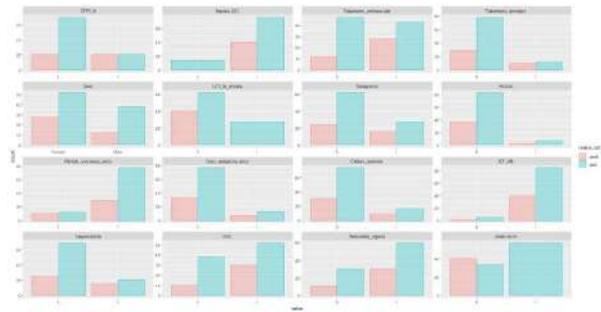


Figura 2. Distribución variables categóricas

## 2.4. Análisis multivariable

El empleo de *Machine Learning* permite la construcción de un modelo que pueda ser de ayuda para la toma de decisiones del personal médico en la UCI. Para ello, se entrenarán múltiples modelos con el fin de obtener una herramienta de apoyo replicable, estableciendo una semilla fija que permita obtener resultados reproducibles.

### 2.4.1. Selección de variables

De las 38 variables preseleccionadas que aportan información al pronóstico neurológico del paciente, se escogen las que a día cero son útiles mediante la ayuda de un experto médico (22): **Aneurisma\_dic**, **Sexo**, **Edad**, **Pérdida\_conciencia\_inicio**, **Crisis\_epilépticas\_inicio**, **Cefalea\_centinela**, **GCS\_ingreso**, **PCR**, **Glicemia**, **Creatinina**, **Sodio**, **Troponina\_I\_ingreso**, **Troponina\_I\_24h**, **Leucocitos**, **Hemoglobina**, **Plaquetas**, **Monocitos**, **DVE**, **IOT\_VM**, **mFisher** y **aneurysm.origin**.

### 2.4.2. Entrenamiento de los modelos

Con este grupo de 22 variables seleccionadas y el 75% del grupo de pacientes (datos de entrenamiento) se entrenan diversos modelos combinando diferentes criterios de selección de variables con diferentes técnicas de clasificación. Para evaluar el desempeño de los modelos y a su vez, evitar el overfitting, se utiliza validación cruzada (CV) repetida de 10 particiones y 50 repeticiones. De esta forma, se generalizan los resultados separando iterativamente los grupos de entrenamiento y validación.

Como criterio de selección de variables se han usado: todas las variables (22), selección de Akaike -AIC- (5), criterio médico reducido (4), *Random Forest* -RF- (16) y Máquinas de soporte vectorial -SVM- (6). Como técnicas de clasificación de ML se han usado: Regresión logística (RL), Análisis discriminante lineal (LDA), K-vecinos cercanos (KNN), Máquinas de soporte vectorial (SVM) y *Random Forest* (RF). Se han entrenado modelos con diferentes conjuntos de hiperparámetros y escogiendo los conjuntos con mayor exactitud (proporción de aciertos sobre el total de pacientes) [4].

Se obtiene una exactitud máxima de 0.7857 pero con sensibilidades de 0.4747. Esto debido al desbalance de las clases. Se aumenta el número de pacientes con buen pronóstico por medio del remuestreo (Bootstrapping).

La base de datos resultante, tiene un 47% de pacientes “good” y un 53% de pacientes con “poor” con los que se vuelve a entrenar los diferentes modelos. Esto permite una mejora de la sensibilidad, pero sacrificando un poco la especificidad y la exactitud.

Para mejorar la sensibilidad sin perjudicar a ningún parámetro se vuelven a entrenar los modelos mediante aprendizaje conjunto (*ensemble training*) por el reentrenamiento de los modelos de RF (mtry =1, min.node.size=10, splitrule = ‘gini’, num.trees=500), SVM (C=100, sigma = 0.01), RL y KNN (k=20) y el empleo de voto por mayoría. Con este entrenamiento, se obtiene una exactitud de 0.8564, una sensibilidad de 0.8405 y especificidad de 0.8705, con las variables seleccionadas mediante criterio AIC (**Edad, GCS\_ingreso, Troponina\_I\_ingreso, DVE y Monocitos**).

Finalmente, este modelo se prueba con el restante 25% de los datos originales (datos de prueba) obteniéndose una exactitud de 0.8333, una sensibilidad de 0.7727 y especificidad de 0.9. La curva ROC representando sensibilidad vs especificidad se observa en la Figura 4.

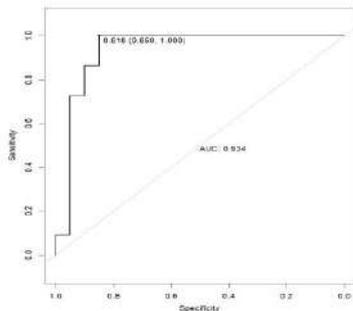


Figura 4. Curva ROC del modelo seleccionado

### 2.5. Efecto de la ventilación mecánica

Es infrecuente que los médicos tomen decisiones en cuanto al pronóstico de un paciente con HSA grave al día 0 de ingreso. Sin embargo, a medida que pasan los días de ingreso, la pregunta de cómo será la recuperación, cobra especial sentido, sobre todo en aquellos pacientes que siguen requiriendo intubación. Por ello, en este trabajo se ha considerado analizar la utilidad de los modelos analizados en un subgrupo de pacientes que continuaban intubados al tercer día de ingreso. Por ello, se vuelve a entrenar el modelo con solamente aquellos pacientes que han estado tres o más días intubados (85 pacientes). Se obtiene una exactitud de 0.9259, sensibilidad de 0.8462 y especificidad de 1. La curva ROC (Figura 5) permite observar que el modelo sigue siendo de utilidad a partir del tercer día de ventilación mecánica.

### 3. Discusión y conclusiones

Este análisis ha servido de base para el estudio del efecto de diversos predictores sobre el pronóstico neurológico de pacientes que han padecido hemorragia subaracnoidea con el fin de poder encontrar una herramienta que sirva de apoyo en la toma de decisiones por parte del personal médico.

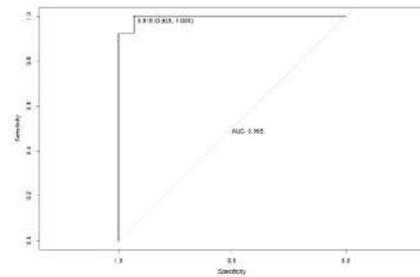


Figura 5. Curva ROC modelo seleccionado.

Mediante el uso de algoritmos de clasificación, seleccionando sus hiperparámetros óptimos y las variables por diversos criterios, se ha podido crear una herramienta de apoyo útil tanto a día cero como al tener en cuenta la intubación, con exactitud de 0.8564 y 0.9259.

Los modelos realizados con las variables seleccionadas mediante criterio AIC (**Edad, GCS\_ingreso, Troponina\_I\_ingreso, DVE y Monocitos**), que son variables con sentido clínico, y remuestreados con *bootstrapping* y entrenados con algoritmos de aprendizaje conjunto permiten obtener herramientas que presentan muy buenas predicciones tanto a día 0 como a día 3 y son unos candidatos muy buenos a la hora de poder predecir el pronóstico neurológico de pacientes con HSA y constituir una herramienta de apoyo al personal médico.

Más detalles de los modelos entrenados, configuración e hiperparámetros utilizados, se pueden encontrar en el trabajo final de grado [5].

### Agradecimientos

UCI Quirúrgica y Unidad de Ictus del Hospital Clínic de Barcelona y todo su equipo por su continuo soporte médico.

Grupo de investigación CoDALab que financia la asistencia al congreso a través del proyecto AGRUPS 2022-UPC.

### Referencias

- [1] C. Gauba and P. N. Renjen, “Subarachnoid hemorrhage,” *ICU Protoc. A Stepwise Approach*, pp. 223–228, 2012, doi: 10.1007/978-81-322-0535-7\_27.
- [2] D. S. Rosen and R. L. MacDonald, “Subarachnoid hemorrhage grading scales,” *Neurocrit. Care*, vol. 2, no. 2, pp. 110–118, 2005, doi: 10.1385/NCC:2:2:110.
- [3] S. Van Buuren and K. Grothuis-oudshoorn. MICE: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software*, [S.l.], v. 45, Issue 3, p. 1 - 67, dec. 2011.
- [4] B. Nithya and V. Ilango, “Predictive analytics in health care using machine learning tools and techniques,” *Proc. 2017 Int. Conf. Intell. Comput. Control Syst. ICICCS 2017*, vol. 2018-Janua, pp. 492–499, 2017, doi: 10.1109/ICCONS.2017.8250771.
- [5] H. Araujo. “Modelos de predicción del pronóstico neurológico de pacientes con hemorragia subaracnoidea con ventilación mecánica”. Universitat Politècnica de Catalunya, Barcelona. Junio, 2022.

# Towards early detection of melanoma using deep learning and multimodal data

L. Marugán Romero<sup>1</sup>, M. Castro Fernández<sup>2</sup>, C. Soguero-Ruiz<sup>1</sup>

- 1 Department of Signal Theory and Communications and Telematic Systems and Computation, Rey Juan Carlos University, Camino del Molino s/n, 28943 Fuenlabrada, Madrid, Spain
- 2 Research Institute for Applied Microelectronics (IUMA), University of Las Palmas de Gran Canaria, Spain

## Abstract

Melanoma is one of the least frequent types of skin cancer and at the same time one of the most lethal. Therefore, early detection of this pathology is key to increase patient survival and reduce the costs associated with its treatment. With this we aim to improve melanoma detection in the development of melanoma detection and diagnosis systems using deep learning algorithms and databases that incorporate not only images of the lesion, but also clinical data of each patient (SIIM-ISIC 2019, SIIM-ISIC 2020, and PAD-UFES-20). Statistical models are applied to evaluate the relevance of each clinical variable to detect melanoma. This allows to increase the information about the data, and to reduce the dimensionality of the data. Since multimodality data is analysed, artificial neural networks are considered to analyse the clinical data, whereas pre-trained convolutional neural networks are evaluated for the images of the lesion. The results show how the use of clinical data combined with images of the lesion improves the algorithm's performance to predict whether a patient has melanoma.

## 1. Introduction

Melanoma, one of the most aggressive types of skin cancer, arises through the uncontrolled growth of melanocytes [1]. The likelihood of developing metastases is quite high, even when patients have the lesion localised at the time of diagnosis and are treated by surgical excision [2]. Therefore, melanoma is considered a socioeconomic problem worldwide. Although the incidence of many types of malignant neoplasms is currently decreasing considerably, the incidence of melanoma continues to increase [3].

Among the tests used in the detection of melanoma, laser scanning confocal microscopy and epiluminescence microscopy or dermoscopy are noteworthy [4]. The problem with these methods is the variability of lesion evaluation depending on the subject and the subject's experience. For this reason, numerous efforts have been made in recent years to automate the detection process. Thanks to the development of artificial intelligence, important advances have been made in the field of medicine. Among these improvements are pattern recognition in images and computer-aided diagnosis. Most of these algorithms use only images as a source of data. However, in practice, physicians base part of their diagnosis on patient data (skin tone, history of cancer, origin, whether the subject is a smoker, etc.).

Computer-aided diagnosis can be a tool to help save some of the time and money spent in the diagnosis of melanoma skin cancer. By speeding up the workflow with these methods, a conclusion about the patient's health status could be obtained in a shorter period of time. This would allow earlier initiation of the corresponding treatment and thus

facilitate patient adherence to treatment. This would increase the probability of obtaining the results and, therefore, save costs associated with treatment complications. Moreover, if such tools could be implemented on mobile devices, the mentioned benefits of applying them may have a higher impact. However, it would imply having lower computational capabilities to run the algorithms too. Therefore, the motivation of this work is to provide an approach in the field of computer-aided detection, using as a source of information not only images, but also light-weighted algorithms that associate the images with the clinical context of the patient. Although previous work had already been done with databases that include metadata [5], image segmentation [6] or transfer learning [7], the novelty of this project lies in the combination of images and their metadata to improve the classification process and the usage of feature selection methods to ensure that only relevant information is provided to the algorithms.

## 2. Materials and methods

Figure 1 shows the working methodology applied in this study. Two databases were used: one with dermoscopic images (a combination of SIIM-ISIC-2019 [8]–[10] and SIIM-ISIC-2020 [11]) and one with clinical images (PAD-UFES-20 [12]). Both are further described in the next section. Once suitable databases were found, the images and metadata were pre-processed. The images were centralised and resized. The numerical variables were normalised, and the categorical variables were converted into binary ones. In addition, duplicate or null instances were removed, and finally the datasets were balanced. Then, we perform a descriptive analysis of the metadata and apply feature selection techniques to them. Artificial neural networks and deep learning (DL) algorithms were used to detect melanoma based on metadata and images. Both binary and multiclass classification is performed in different experiments that will be described in detail below.

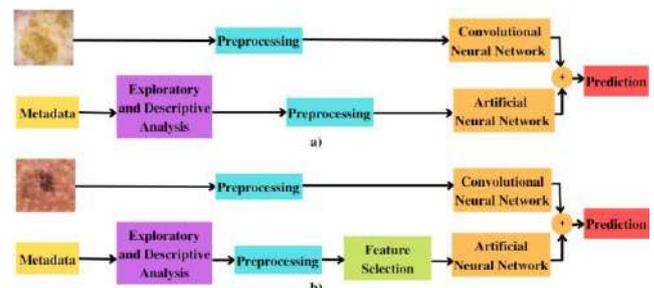


Figure 1. Working scheme with a) SIIM-ISIC and b) PAD-UFES-20 datasets.

## 2.1. Databases

The databases used in this work contain images and patient information providing the clinical context of the individual.

The first dataset analyzed in this study was obtained by concatenating entries from the SIIM-ISIC-2019 and SIIM-ISIC-2020 databases. It is composed of a total of 58457 samples (5106 of them are melanoma), that is, dermoscopic images associated with up to 8 different variables: patient ID, lesion ID, age and sex of the patient, anatomical region where the lesion is located, the diagnosis of the lesion, the malignancy of the lesion and the label (*target* = 1 for melanoma and *target* = 0 for non-melanoma).

The second dataset considered, PAD-UFES-20, is composed of a total of 2298 observations (62 of them are melanoma cases), associated with 22 variables, including the information mentioned for the previous dataset, and other important data such as whether the patient consumes alcohol or drugs, the origin of his parents, his history of cancer, the type of skin, whether he is usually exposed to pesticides or whether the patient has access to drinking water and a sewage system at home. They also provide information on the size of the lesion and whether the patient has witnessed a change in the appearance of the lesion or that the lesion has become raised, enlarged, itchy, painful or bleeding. The images in PAD-UFES-29 were captured with smartphones.

There are as many classes as there are unique values for the variable "diagnosis" in each dataset. Thus, the SIIM-ISIC database is composed of 8 classes (melanoma, nevus, basal cell carcinoma, benign keratosis, seborrheic keratosis, squamous cell carcinoma, vascular lesion and dermatofibroma). However, PAD-UFES is composed of 6 classes (melanoma, melanocytic nevus, basal cell carcinoma, seborrheic keratosis, actinic keratosis and squamous cell carcinoma).

## 2.2. Data pre-processing

The datasets used in this study contain quantitative and categorical variables. The One Hot Encoding Technique [13] was applied to categorical variables and the numerical variables were normalized.

Both datasets are imbalanced. Therefore, data balancing techniques were performed to reduce the difference between the number of samples in each class. Specifically, an undersampling technique was applied, which consists of reducing the number of samples in the majority class. Within the data pre-processing stage, it is convenient to identify the most relevant features of the dataset. For this purpose, the Mutual Information [14], Bootstrap resampling [15] and Decision Tree [16] methods were used as feature selection strategies.

## 2.3. Deep Learning model selection

Artificial Neural Networks (ANN) were selected as learning models to address the classification problem. To work with images in the database Convolutional Neural Networks (CNN) were used, which are commonly used in the field of pattern recognition in computer vision and there are plenty of different architectures. In this work, one main objective is to focus on light-weighted algorithms, so they can be used even in mobile devices. For such reason two architectures

were selected: MobileNetV2 [17] and EfficientNetB3 [18], which are known for their low number of parameters and their high performance (for their size). To study the metadata, they were passed through a "metadata-block" and then both the CNN and the metadata-block outputs were concatenated, before a fully connected layer serving to make the classification. This metadata-block contains two fully connected layers, with batch normalization and sigmoid activation function and 512 and 128 neurons respectively. Additionally, a dropout layer in between (dropout rate equal to 0.3) for regularization purposes.

Training a network is time consuming. The transfer learning technique is often used to address this problem [19]. This is based on taking advantage of the training of a pre-existing CNN and adjusting the model parameters to adapt it to the data to be worked with [20]. Thanks to this technique, it is not necessary to train the network from scratch. In this study, ImageNet dataset [21] was used to pre-train the MobileNetV2 and EfficientNet-B3 models.

## 2.4. Training, validation and evaluation stages

Each dataset was divided into three parts: training, validation and test sets, where the distribution is 60%, 20% and 20%, respectively for the SIIM-ISIC dataset and 70%, 15% and 15% for the PAD-UFES-20 dataset. A larger percentage was considered to the training of the PAD-UFES-20 database due to the small number of melanoma samples contained in the original dataset. Finally, the test or evaluation set is used to examine the performance of the algorithm. For this, the predictions obtained were evaluated using various metrics (accuracy, specificity, sensitivity, precision, F1 parameter and Area Under Curve). In addition, the hyperparameters of the model are: number of outputs (two for binary problems and more than two for multi-class problems), learning rate ( $3 \cdot 10^{-5}$ ), image size (256x256 pixels), number of epochs (15), batch size (64 samples for SIIM-ISIC dataset and 5 samples for the PAD-UFES-20 dataset, due to the reduced number of instances it contains).

## 3. Experiments

In this study, the following experiments were proposed for each of the databases mentioned above:

- **Experiment I.** Comparison between approaching the target problem using a binary model versus approaching it using a multiclass model. With this approach, the algorithm classifies a lesion according to the value of the "diagnostic" variable, instead of the "target".
- **Experiment II.** Comparison between using only images as input data to a convolutional neural network versus using these images combined with clinical data.
- **Experiment III.** Comparison between using a pre-trained convolutional neural network of MobileNetV2 type versus a pre-trained convolutional neural network of EfficientNet-B3 type.
- **Experiment IV.** Comparison between using feature selection techniques on patient clinical data versus not implementing such techniques. This experiment has only been performed on the PAD-UFES-20 database.

### 3.1. Software used

This work has been developed using the Python 3.7.12 programming language and the following open-source libraries: Pandas (1.3.5), numpy (1.21.6), scipy (1.4.1), matplotlib (3.2.2), seaborn (0.11.2) and scikit-learn (1.0.2) libraries have been used during the pre-processing, feature selection and data visualization and analysis stage. Pytorch (1.11.0) was used for the learning stage.

### 3.1. Experimental setup

The dataset was divided into five equal parts in the SIIM-ISIC database and seven PAD-UFES. In each dataset, one of the splits is reserved for validation and one for evaluation. During training, the validation process will be carried out in order to avoid overfitting phenomena [22]. For this purpose, the k-fold cross validation technique will be used [23].

During validation, the parameters of the models that have obtained the best metrics are saved. These models will then be tested on the evaluation data set. The overall performance of the algorithm will be calculated as the average of the metrics obtained with each of the models used during the evaluation stage.

## 4. Results

In this section results obtained for SIIM-ISIC and PAD-UFES-20 datasets are presented and discussed.

### 4.1 Results with the SIIM-ISIC dataset

Table 1 shows the results obtained when performing experiments, I, II, III and IV, using the SIIM-ISIC dataset.

Model	AUC	Accu- racy	Speci- ficity	Preci- sion	Sensi- tivity	Parame- ter F1	Mean
MobileNetV2 + MD	89,2	80,57	81,62	80,36	79,45	79,88	81,85
MobileNetV2	88,24	79,11	80,2	78,86	77,95	78,36	80,45
EfficientNet-B3 + MD	89,8	<b>81,19</b>	<b>82,06</b>	<b>80,9</b>	80,27	80,56	<b>82,46</b>
EfficientNet-B3	88,91	80,43	79,6	79,01	81,3	80,13	81,56

a) Binary model (2 classes)

Model	AUC	Accu- racy	Speci- ficity	Preci- sion	Sensi- tivity	Parame- ter F1	Mean
MobileNetV2 + MD	89,93	75,06	73,07	75,37	<b>87,01</b>	<b>80,75</b>	80,2
MobileNetV2	88,73	72,69	73,58	74,77	82,73	78,52	78,5
EfficientNet-B3 + MD	<b>90,37</b>	74,65	77,81	78,07	83,42	80,62	80,82
EfficientNet-B3	88,91	72,49	73,3	74,75	83,64	78,94	78,67

b) Multiclass model (8 classes)

**Table 1.** Metrics obtained when experimenting with EfficientNet-B3 and MobileNetV2 architectures, including or not patient metadata (MD) as data source. The experiment has been performed by posing the problem with both a binary and a multiclass approach. “MD” denotes the use of metadata as input to the algorithm.

All the models presented in Table 1 improve when applying a binary approach versus a multi-class approach. The metrics increase when combining clinical data from the patient with the images associated with the lesion. Regarding the architecture used, no relevant improvements are observed depending on the model used. In summary, it can be said that of all the experiments performed with the SIIM-ISIC database, the one with the best metrics is the model that uses an EfficientNet-B3 architecture, which differentiates only between two groups of data and introduces images of the lesion and metadata associated with the clinical episode as

input to the model. Using this model, the mean score obtained with the metrics is 82.46%.

### 4.2. Results with the PAD-UFES-20 dataset

Table 2 shows the results obtained by performing experiments I, II, III and IV, using the PAD-UFES-20 dataset.

Model	AUC	Accu- racy	Speci- ficity	Preci- sion	Sensi- tivity	Parame- ter F1	Mean
MobileNetV2 + MD	93,20	84,52	85,71	85,62	83,33	84,24	86,10
MobileNetV2 + MD + FS	94,56	<b>88,10</b>	<b>92,86</b>	<b>92,36</b>	83,33	87,14	<b>89,73</b>
MobileNetV2	91,16	84,52	88,10	87,60	80,95	83,66	86,00
EfficientNet-B3 + MD	95,24	80,95	69,05	76,31	92,86	83,18	82,93
EfficientNet-B3 + MD + FS	<b>95,92</b>	86,9	83,33	86,53	90,48	<b>87,65</b>	88,47
EfficientNet-B3	84,35	75	64,29	71,82	85,71	77,65	76,47

a) binary model (2 classes)

Model	AUC	Accu- racy	Speci- ficity	Preci- sion	Sensi- tivity	Parame- ter F1	Mean
MobileNetV2 + MD	94,90	60,71	57,14	71,03	<b>100</b>	82,75	77,76
MobileNetV2 + MD + FS	92,52	53,57	40,48	62,83	92,86	74,18	69,41
MobileNetV2	89,80	55,95	52,38	68,50	92,86	77,69	72,86
EfficientNet-B3 + MD	83,67	52,38	28,57	57,04	92,86	70,46	64,16
EfficientNet-B3 + MD + FS	86,73	54,76	47,62	63,89	90,48	74,63	69,69
EfficientNet-B3	75,85	46,43	11,90	50,75	90,48	64,97	56,73

b) multiclass model (6 classes)

**Table 2.** Metrics obtained when experimenting with EfficientNet-B3 and MobileNetV2 architectures (Experiment III), including or not patient metadata as a data source (Experiment II). Feature selection techniques have been applied in some of the experiments with metadata (Experiment IV); The experiment has been performed by approaching the problem with both a binary and a multiclass approach (Experiment I). “MD” denotes the use of metadata as input to the algorithm, “FS” denotes the use of feature selection techniques on metadata.

The proposed framework cannot handle the multiclass problem. Looking at specificity, it can be seen that the model is more likely to generate an erroneous prediction than an accurate one, by predicting that a patient does not have melanoma when in fact he does not. Furthermore, analysing the sensitivity values (all above 90%) together with the specificity values (none above 60%), it can be seen that the model is not able to discern between classes, as it considers that almost all patients have melanoma. Therefore, addressing the problem using a binary model improves the predictive ability of the algorithm.

Focusing on the analysis of the binary models, it can be observed how the mean of their metrics improves when images combined with metadata are applied. Within these experiments, those in which feature selection techniques have also been used show better results. In summary, it can be said that of all the experiments performed with the PAD-UFES-20 database, the one with the best metrics is the model using a MobileNetV2 architecture, which differentiates only between patients with and without melanoma and introduces as input values to the model images of the lesion and metadata associated with the clinical episode. This metadata was selected using feature selection techniques. Using this model, the mean score obtained is 89.73%.

## 5. Conclusions

In this report, an innovative strategy has been presented to address the problem of automatic detection of melanoma

lesions. For this purpose, pretrained convolutional networks have been used as classifiers.

Public databases presenting both images and clinical patient data have been used as information sources. Conventionally, diagnostic computer systems for melanoma prediction have used only images to establish predictions. Therefore, the search for a public database that combines images and clinical data associated with the patient has involved an exhaustive period of research.

After training, validating, and evaluating the learning models, the same conclusions have been obtained with both databases: the best performing model is the one that uses a binary approach and introduces both images and metadata as input data to the model. In the case of the PAD-UFES-20 database, variables have been selected using feature selection techniques.

## Acknowledgements

This work was partly funded by the Spanish Government by the Grant AAVis-BMR PID2019-107768RA-I00); and by the Project Ref. 2020-661, financed by Rey Juan Carlos University (URJC) and Community of Madrid.

## References

- [1] M. Berwick *et al.*, ‘Melanoma epidemiology and prevention’, in *Melanoma*, vol. 167, Kluwer Academic Publishers, 2016, pp. 17–49. doi: 10.1007/978-3-319-22539-5\_2.
- [2] L. M. D. Duncan, ‘The classification of cutaneous melanoma’, *Hematol. Oncol. Clin. North Am.*, vol. 23, no. 3, pp. 501–513, Jun. 2009, doi: 10.1016/J.HOC.2009.03.013.
- [3] R. M. Mackie, A. Hauschild, and & A. M. M. Eggermont, ‘Epidemiology of invasive cutaneous melanoma’, *Ann. Oncol.*, vol. 20, pp. 1–7, 2009, doi: 10.1093/annonc/mdp252.
- [4] S. A. Leachman *et al.*, ‘Methods of melanoma detection’, in *Melanoma*, vol. 167, Kluwer Academic Publishers, 2016, pp. 51–105. doi: 10.1007/978-3-319-22539-5\_3.
- [5] D. N. Anggraini Ningrum *et al.*, ‘Deep Learning Classifier with Patient’s Metadata of Dermoscopic Images in Malignant Melanoma Detection’, *J. Multidiscip. Healthc.*, vol. 14, pp. 877–885, Apr. 2021, doi: 10.2147/JMDH.S306284.
- [6] D. N. H. Thanh, V. B. S. Prasath, L. M. Hieu, and N. N. Hien, ‘Melanoma Skin Cancer Detection Method Based on Adaptive Principal Curvature, Colour Normalisation and Feature Extraction with the ABCD Rule’, *J. Digit. Imaging*, vol. 33, no. 3, pp. 574–585, Dec. 2019, doi: 10.1007/S10278-019-00316-X.
- [7] Q. Ha, B. Liu, and F. Liu, ‘Identifying Melanoma Images using EfficientNet Ensemble: Winning Solution to the SIIM-ISIC Melanoma Classification Challenge’, Oct. 2020, doi: 10.48550/arxiv.2010.05351.
- [8] M. Combalia *et al.*, ‘BCN20000: Dermoscopic Lesions in the Wild’, Aug. 2019, doi: 10.48550/arxiv.1908.02288.
- [9] P. Tschandl, C. Rosendahl, and H. Kittler, ‘The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions’, *Sci. Data 2018 51*, vol. 5, no. 1, pp. 1–9, Aug. 2018, doi: 10.1038/sdata.2018.161.
- [10] N. C. F. Codella *et al.*, ‘Skin Lesion Analysis Toward Melanoma Detection: A Challenge at the 2017 International Symposium on Biomedical Imaging (ISBI), Hosted by the International Skin Imaging Collaboration (ISIC)’, *Proc. - Int. Symp. Biomed. Imaging*, vol. 2018-April, pp. 168–172, Oct. 2017, doi: 10.48550/arxiv.1710.05006.
- [11] V. Rotemberg *et al.*, ‘A patient-centric dataset of images and metadata for identifying melanomas using clinical context’, *Sci. Data 2021 81*, vol. 34, no. 1, pp. 1–8, Jan. 2021, doi: 10.1038/s41597-021-00815-z.
- [12] A. G. C. Pacheco and R. A. Krohling, ‘The impact of patient clinical information on automated skin cancer detection’, *Comput. Biol. Med.*, vol. 116, no. November 2019, p. 103545, 2020, doi: 10.1016/j.combiomed.2019.103545.
- [13] F. Chollet, ‘Deep Learning with Python (2nd ed.) MEAP’, in *Books*, 2020, pp. 1–9.
- [14] J. R. Vergara and P. A. Estévez, ‘A review of feature selection methods based on mutual information’, *Neural Comput. Appl. 2013 241*, vol. 24, no. 1, pp. 175–186, Mar. 2013, doi: 10.1007/S00521-013-1368-0.
- [15] T. J. DiCiccio and B. Efron, ‘Bootstrap confidence intervals’, <https://doi.org/10.1214/ss/1032280214>, vol. 11, no. 3, pp. 189–228, Sep. 1996, doi: 10.1214/SS/1032280214.
- [16] B. Taha Jijo and A. Mohsin Abdulazeez, ‘Classification Based on Decision Tree Algorithm for Machine Learning’, *J. Appl. Sci. Technol. Trends*, vol. 2, no. 01, pp. 20–28, Mar. 2021, doi: 10.38094/jastt20165.
- [17] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L. C. Chen, ‘MobileNetV2: Inverted Residuals and Linear Bottlenecks’, *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 4510–4520, Jan. 2018, doi: 10.48550/arxiv.1801.04381.
- [18] M. Tan and Q. V. Le, ‘EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks’, *36th Int. Conf. Mach. Learn. ICML 2019*, vol. 2019-June, pp. 10691–10700, May 2019, doi: 10.48550/arxiv.1905.11946.
- [19] Q. Yang, Y. Zhang, W. Dai, and S. J. Pan, *Transfer Learning*. Cambridge University Press, 2020. doi: 10.1017/9781139061773.
- [20] K. O’Shea and R. Nash, ‘An Introduction to Convolutional Neural Networks’, Nov. 2015, doi: 10.48550/arxiv.1511.08458.
- [21] A. C. B. and L. F.-F. Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, ‘ImageNet Large Scale Visual Recognition Challenge’, *Int. J. Comput. Vis.*, 2015.
- [22] N. Srivastava, G. Hinton, A. Krizhevsky, and R. Salakhutdinov, ‘Dropout: A Simple Way to Prevent Neural Networks from Overfitting’, *J. Mach. Learn. Res.*, vol. 15, pp. 1929–1958, 2014.
- [23] Jason Brownlee, ‘Evaluate the Performance of Machine Learning Algorithms with Resampling’, in *Machine Learning Mastery with Python*, V1.20., 2021, p. 59.

## ÍNDICE DE AUTORES

Abad-González, A.	423	Bailón, R.	448
Abella, M.	10, 213	Balastegui García, G.	72
Abellán Rubio, E.	266	Barquero-Pérez, O.	245
Abete, I.	52	Barrenechea Carrasco, M.	345
Acha, B.	18, 415	Barros-Tornay, R.	415
Aginagalde Lopez, A.	217	Barroso García, V.	39
Aguilar-Aguilar, E.	52	Bartrés-Faz, D.	521
Aguilar Paredes, F.	237	Belda Calabuig, A.	142
Aguilar Sabido, C.	486	Bernabéu-Wittel, J.	18
Albaladejo-Belmonte, M.	146, 398	Bernal Benítez, C.	506
Alberola-López, C.	85, 167	Berner-Juhos, K.	510
Alberola-Rubio, J.	146, 398	Berrocoso, E.	130
Albusac, J.	97	Berruezo, A.	461
Alcaraz Martínez, R.	365, 394	Blanco-Velasco, M.	419, 431
Alcon, C.	113	Bocos-Corredor, E.	262
Aldazabal, J.	249	Bonse, J.	217
Alfano, F.	274	Bragulat Arévalo, M.	369
Algaba-Vidoy, M.	101	Buendia Aviles, S.	357
Almendral, J.	477	Bueno, G.	2
Alonso, E.	452	Bueno Zamora, O.	274
Alonso, J. F.	283	Burgaleta Ledesma, A.	117
Alonso-Linaje, G.	180	Burgos Llamo, A.	121
Alonso, R.	22, 490	Buruaga Lamarain, L.	217
Álvarez, D.	39, 411, 427, 448	Cabal-Rodríguez, R.	77
Álvarez Rodríguez, M.	184, 233	Cabezaolías, C.	258
Alves-Ferreira, M.	287, 402	Cabrera-Umpiérrez, M. F.	56, 117
Amaro, S.	525	Cadena Bonfanti, A.	89
André, J. M.	192	Calvillo Arbizu, J.	93
Antón-Rodríguez, M.	270	Calvo, A.	510
Aramendi, E.	155, 175, 444, 452	Calvo Manuel, F.	171
Araujo Rodríguez, H.	525	Cámara-Vázquez, M. A.	245
Arbelo Lainez, E.	369	Campos Añón, F.	381
Arce-Diego, J. L.	469, 494	Cànaves Llabrés, A. I.	201
Arco, A.	510	Candel Cano, P.	473
Arco, H.	510	Candel, M. J.	81
Arias Palomares, M. A.	365	Candela Muries, T.	291
Armañac-Julián, P.	448	Cano del Pozo, M.	109
Armijo Suárez, O.	486	Cano Mayo, P.	105
Arnal, M. L.	81	Cappella, B.	217
Arredondo Waldmeyer, M. T.	117	Carneiro, P.	353
Ayala Fernández, U.	345	Carretero-Calvo, P.	109
Aznar, S. D.	81	Casado, E.	68
Azorín, J. M.	188, 291, 308, 320	Castellanos, E.	477
Bachiller, A.	283, 299	Castells, F.	31, 134, 138, 150
		Castillo-Passi, C.	85, 167
		Castillo Rabazo, J.	266

Castro de la Nuez, P.	93	Ferreira-Santos, D.	163
Castro Fernández, M.	529	Ferrer-Montiel, A.	361
Cattaneo, G.	521	Ferreres García, K.	274
Cenarro, R.	22	Ferrero, J. M.	361, 386
Cenis, J. L.	81	Ferrero, L.	308
Ceruelo Abajo, S.	97, 233	Ferrero, V.	320
Chausa, P.	52, 521	Fico, G.	117
Chmarra, M. K.	517	Fraile Marinero, J. C.	180
Chorro, J.	150	Franco Terriza, L.	60
Chushig-Muzo, D.	423	Fuster Palà, A.	221
Cisnal, A.	180	Gabilondo Cuellar, I.	345
Climent, A. M.	245	Galarregui, C.	52
Costa, V.	258, 278, 349	Gallardo, C.	406
Crespo Aguirre, M.	138	Ganoza-Quintana, J. L.	494
Cunill Rodríguez, M.	357	García-Barrio, N.	121
Dankelman, J.	517	García-Blas, J.	10
de Benavides, J.	490	García-Carmona, R.	262
de Ciurana Gay, J.	201	García-Casado, F. J.	142, 146, 398
de los Reyes Guzmán, A.	97, 184, 233	García-Cazorla, A.	299
del-Ama, A. J.	328	García Fernandez, I.	159
del Campo, F.	39, 427	García-García, P.	465
del Cerro, C. F.	10, 213	García-Lizarribar, A.	113
del Riego Gómez, S.	349	García, M.	14, 427
Delgado Atencio, J. A.	357	García-Molina, A.	52
Descó, M.	10, 213	García-Real, M. I.	10
Díaz-Pernas, F. J.	270	García-Rudolph, A.	52
Díaz Rivas, R.	506	García-Sáez, G.	341, 406
Eimil-Suarez, E.	77	García Teruel, M.	365
Elena, M.	60	García-Vicente, C.	423
Elola, A.	155, 444, 452	Garnés Camarena, O.	43
Elorza Fernández, D.	377	Garrido Rubiales, B.	68
Escribano Cano, P.	365	Gavidia Ceballos, L.	506
Espinosa-Salinas, I.	68	Gil-Agudo, A.	184, 233, 328
Esteban Gata, M. J.	60	Gil-Castillo, J.	328
Fabregat-Sanjuan, A.	241	Gil, E.	448
Falagán, S.	68	Gil Simoes, R.	105, 333
Fanjul-Vélez, F.	357, 469, 494	Gilete Preciado, V.	64
Ferencz, A.	510	Giraldo Giraldo, B. F.	456
Fernández-Arias, I.	85, 167	Gjerdje, J.-M.	517
Fernández-Linsenbarth, I.	283	Goicolea, J. M.	274
Fernández Moreno, G.	105	Gómez, A.	213
Fernández-Poyatos, M.	411	Gómez, C.	47, 109, 287, 303, 390, 402
Fernández-Rodríguez, A.	192	Gómez, E. J.	6, 22, 52, 68, 225, 353, 377, 465, 486, 490, 517, 521
Ferré López, D.	456	Gómez-García, J. A.	101
		Gómez-Grande, A.	6, 22, 490

Gómez, J. F.	477	Irrarrazaval, P.	85, 167
Gómez-Pilar, J.	163, 287, 402	Isasi, I.	452
Gomis Jimeno, F.	72	Isla-Calvo, A.	469
Gomis Roman, P.	369	Iza, A.	452
Gomis-Tena, J.	35	Jané, R.	237
González-Alonso, J.	270	Jara Muriel, M.	266
González, L.	2	Jarrín Vera, I.	121
González Morcillo, C.	97	Jaureguibeitia, X.	155, 175
González-Navarro, C. J.	52	Jiménez Carrizosa, M.	105, 333
González Nuño, F. M.	266, 514	Jiménez García, J.	427
González-Ortega, D.	270	Jordà Burgos, P.	369
Gorrin, N.	274	Juan, J. V.	188
Gozal, D.	39, 163, 411, 427, 448	Junquera Godoy, I.	35
Gracia Laso, D. I.	291	Jutinico Alarcón, A.	89
Guerrero, A. L.	287, 402	Kamarova, K.	390
Guerrero Pérez, M.	64	Kheirandish-Gozal, L.	39, 427
Guijarro, E.	146, 398	Kontaxakis, G.	337
Guill, A.	138, 150	Kontaxis, S.	448
Guillem, M. S.	245	Kotzki, S.	353
Gutiérrez, A.	406	Lafuente Alcázar, B.	253
Gutiérrez-de Pablo, L.	303	Laguna, P.	436, 448
Gutiérrez-de Pablo, V.	47, 287, 303, 402	Lamata, P.	274
Gutiérrez Delgado, M.	142	Langø, T.	517
Gutiérrez Fernández-Calvillo, M.	245	Lara-Doña, A.	130
Gutiérrez-Tobal, G. C.	39, 163, 411, 427, 448	Lázaro, J.	448
Gutiez-Bravo, B.	225	Lazo, M.	415
Guzmán-García, C.	465, 517	Ledesma Carbayo, M. J.	274
Hernández-Romero, I.	245	Leñero-Bardallo, J. A.	18
Hernando, M. E.	52, 68, 341, 406	León, J.	510
Herráez, D.	444	Liberos, A.	159
Herrera Tirado, V.	97	Lillo Crespo, M.	72
Herrera-Valenzuela, D.	328	Liu, V.	209
Herrero, M.	14	Lizarraga, S.	274
Hingorani-Jai Prakash, S.	361	Llull, L.	525
Holgado-Cuadrado, R.	419, 431	Lobato Gómez, C.	373, 481
Hornero, R.	14, 27, 39, 47, 125, 163, 180, 196, 287, 316, 324, 390, 402, 411, 427, 448	Loidi, A.	386
Hoshi, H.	390	López Gálvez, M. I.	14
Huerta Herraiz, A.	394	López López, S.	278
Iáñez, E.	188, 291, 308, 320	Lopez Martinez, M. J.	113
Ibáñez-Molina, A. J.	295	López-Sánchez, N.	361
Idris, A. H.	175	Lopour, B.	209
Iglesias-Parro, S.	295	Lora-Millán, J. S.	312
Imaz Higera, J.	142	Loria-Kohen, V.	52
Iniesta-Chamorro, J. M.	52, 68	Lovisoló, L.	431

Lozano Berrio, V.	184	Moreno-Rubio, J.	68
Lozano-García, M.	237	Moreno Vázquez, J. C.	481
Lozano, M.	159	Muñoz Carrero, A.	121
Malanda Trigueros, A.	43	Muñoz Pérez, M.	357
Mañanas, M. A.	283, 299	Mujica Delgado, L. E.	525
Marcos Jorquera, D.	205	Navallas Irujo, J.	43
Marcos-Martínez, D.	27, 125, 196, 316, 324	Navas-Carretero, S.	52
Marcos-Pasero, H.	52	Navas, P.	274
Martín-Fernández, A.	27, 125	Niñerola Baizán, A.	381
Martín Martínez, I.	312	Nohales-Alfonso, F. J.	146, 398
Martín-Montero, A.	448	Norvik, A.	444
Martín-Pintado-Zugasti, A.	262	Núñez, P.	109
Martínez-Alcalá García, A.	465	Oliveira-Barroso, F.	101
Martínez Bordes, G.	473	Oropesa, I.	353, 465, 517
Martínez-Cagigal, V.	27, 125, 180, 196, 316, 324	Ortega, J. F.	64, 510
Martínez-de-Juan, J. L.	35, 142	Ortiz-Álvarez, J.	18
Martínez Donázar, M.	249	Ortiz, M.	188, 291, 308, 320
Martínez, J. A.	52	Ortuño, J. E.	274, 337
Martínez, J. P.	436	Otero, A.	262
Martínez Rodrigo, A.	394	Otero Arteseros, M.	377
Martínez Ruiz, B.	381	Otčenášková, T.	510
Martínez-Zarzuela, M.	270	Pagador, J. B.	64, 510
Marugán Romero, L.	529	Pagan, A.	81
Más Sesé, G.	142	Palacio, L. C.	498
Maturana-Candelas, A.	47	Palmieri, F.	369
Mediero, E.	510	Pàmies-Vilà, R.	241
Medina López, R.	93	Pancorbo, L.	150
Medina-Rivera, I.	299	Pantoja-Gómez, Y.	77
Medrano Gil, A. M.	56	Paredes Puente, J.	249
Megía-García, A.	328	Pareja Galeano, H.	278
Meiriño Seoane, R. M.	171	Pascau, J.	171, 274
Mellado-Artigas, R.	525	Pascual-Carrasco, M.	121
Merino, B.	117	Pascual-Leone, A.	521
Mijancos-Martínez, G.	283	Pascual-Rubio, V.	241
Milara, E.	6, 22, 490	Patiño Mejido, M. J.	262
Millet, J.	31, 134, 138, 150	Patrocinio, D.	373
Mira-Tomás, J. M.	142, 146, 398	Peñate Arrieta, M.	333
Moggioli, F.	262	Pedregosa, E.	373
Molina, V.	283	Pedreira-Jiménez, M.	121
Moreno-Arribas, M. V.	68	Peinado, H.	377
Moreno Blanco, D.	377, 486	Penela, D.	461
Moreno-Calderón, S.	27, 180, 196, 316, 324	Perales Gómez, R.	97, 184
Moreno, J.	521	Perea, L.	278
Moreno, J. C.	101, 328, 502	Pereira-Rodrigues, P.	163
Moreno-Manzano, V.	361		

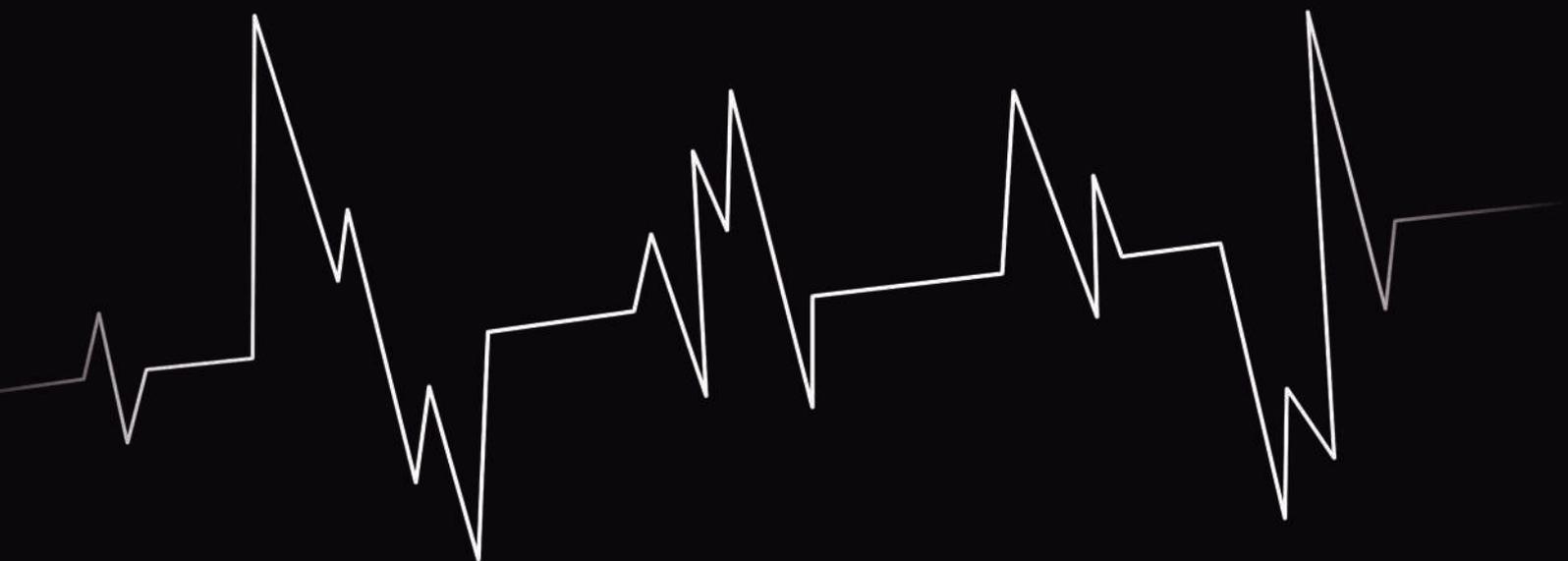
Pérez-Blanco, J. G.	77	Rollano Corroto, M.	229
Pérez, C.	436	Romero-Bascones, D.	345
Pérez-Carrasco, J. A.	18, 415	Romero García, A.	64
Pérez-Gandía, C.	406	Romero, L.	440
Pérez-Rigueiro, J.	81	Romero Mila, B.	209
Pérez Turiel, J.	180	Romero-Oraá, R.	14
Pérez Valdivia, M. A.	93	Romero, P.	159
Pérez-Velasco, S.	27, 125, 196, 303, 316, 324	Romero, S.	283, 299
Pina-de-Paz, J.	328	Romero Zayas, I.	381
Pino Ramirez, A.	506	Romitti, G. S.	159
Pinto, N.	287, 402	Ron-Angevin, R.	192
Pires Louça, L. C.	266	Rubio, F. J.	225, 377, 486
Plaza de Miguel, C.	481, 514	Rufo Rafael de la Cruz, N.	337
Plaza-Seco, C.	419, 431	Ruipérez-Campillo, S.	31, 134, 138, 150
Pose Díez de la Lastra, A.	171	Ruiz, C. R.	477
Poza, J.	47, 109, 303, 390, 402	Ruiz-Díez, S.	101
Prats-Boluda, G.	35, 142	Ruiz Fernández, D.	72
Puche-García, V.	361, 440	Ruiz-Fernández, M. J.	10
Pueyo, E.	436	Ruiz Ordoñez, M.	525
Puig Miquel, T.	201		
		Sahelices, B.	411
Quiles, V.	308, 320	Saiz, J.	35, 440, 498
Quintana, M.	225, 377, 486	Sala Llonch, R.	381
		Samitier, J.	113
Ramírez, A.	52, 68	Sanches, J.	353
Ramirez-Hernández, B.	77	Sánchez-de Madariaga, R.	121
Ramos Hortal, R.	105, 333	Sánchez del Valle, C.	184, 233
Raya, R.	258, 278, 349, 502	Sánchez-González, P.	6, 22, 225, 377, 465, 486, 490, 517, 521
Reglero, G.	52	Sánchez López, A.	171
Regueiro-Gómez, A.	77	Sánchez López de Pablo, C.	278, 349
Reina Tosina, J.	93	Sánchez-Margallo, F. M.	64, 266, 373, 481, 510, 514, 517
Revilla-Vallejo, M.	390	Sánchez-Margallo, J. A.	266, 373, 481, 514, 517
Rieta Ibáñez, J. J.	365, 394	Sánchez-Morillo, D.	130
Rigla, M.	406	Sánchez-Peralta, L. F.	64, 510
Rigo-Vidal, A.	241	Sanderson, D.	10
Roa Romero, L. M.	93	Santamaría-Bustamante, L.	469
Rocon, E.	312	Santamaría-Vázquez, E.	27, 196, 316, 324
Ródenas García, J.	365	Santos, A.	274
Rodrigo, M.	159	Santos-Ceballos, J. C.	77
Rodríguez-Cayetano, M.	85, 167	Santos-Paz, J. A.	262
Rodríguez Falces, J.	43	Santos Virosta, M.	381
Rodríguez-González, V.	109, 287, 303, 402	Sarandeses, P.	6
Rodríguez-Matesanz, M.	225	Schwibbert, K.	217
Rodríguez, R.	353	Sebastian, R.	159
Rojas García, A.	377		
Rojo Agustí, A.	349, 502		
Rojo-Álvarez, J. L.	419		

Segarra, I.	134, 138	Torres-Fortuny, A.	77
Seiffert, A. P.	6, 22, 490	Torres-Pardo, A.	101
Sereno Moyano, M.	68	Torres-Sánchez, S.	130
Serra, D.	159	Torricelli, D.	101
Serrano, A.	415	Tost, A.	299
Serrano, C.	18, 415	Trenor, B.	461, 477
Shigihara, Y.	390	Tudela Fernández, R.	381
Shrey, D.	209	Unneland, E.	444
Silva, M. J.	353	Urendes Jiménez, E.	349
Simmross-Wattenberg, F.	85, 167	Urteaga, J.	444
Simón-Martínez, C.	270	Valiente Jaén, S.	245
Skogvoll, E.	444	Valle Morgado, C.	43
Soguero-Ruiz, C.	423, 529	Vállez, N.	2
Solà Soler, J.	456	Vaquerizo Villar, F.	39, 427
Solana, J.	521	Varona Torralvo, M. I.	56
Solar Risueño, E. M.	205	Velarde-Reyes, E.	77
Somolinos-Simón, F. J.	341	Velasco-Álvarez, F.	192
Soriano, J.	221	Velasco, J. I.	6
Sörnmo, L.	436	Velásquez, A.	278
Soto-Iglesias, D.	461	Vélez Díaz, J.	89
Sou, G.	353	Vidal, N.	353
Stashuk, D.	43	Villa-Muñoz, P.	146, 398
Tachecí, I.	510	Villacorta-Aylagas, P.	85, 167
Talaminos Barroso, A.	93	Villalba-Riquelme, E.	361
Talayero, C.	233	Villar, J.	461
Tapia-Galisteo, J.	341	Villarino Sanz, M.	68
Tello Portillo, J. P.	89	Villasante, A.	113
Tiu, C.	517	Wang, H. E.	175
Tobón, C.	498	Ye Lin, Y.	142
Tola-Arribas, M. A.	109	Zabala Eguren, A.	217
Toledo-Pastrana, T.	415	Zattera, L.	525
Tormos, A.	138, 150	Zubia Aranburu, J.	217
Tormos, J. M.	521	Zulet, M. A.	52
Tornero, D.	221		
Tornero, J.	188		



**CASEIB**  
**2022**

XL Congreso Anual  
de la Sociedad  
Española de  
Ingeniería  
Biomédica



ISBN 978-84-09-45972-8



9 788409 459728