| VOLUMEN 46 | NUMBER 1 | **JANUARY - APRIL 2025**

**Una publicación de:**
**SOMIB**
Sociedad Mexican
de Ingeniería Biomé

# ib
# Revista Mexicana de Ingeniería Biomédica

## Modalidad de publicación

***Publicación continua*:**
**una vez que acepta y prepara un**
**manuscrito, se publicará en línea**

## Publication Modalitu:

***Continuous publication*:**
**once a manuscript is accepted and**
**prepared, it will be released online**

**http://www.rmib.mx**

**ISSN 2395-9126**

**SOMIB**
Sociedad Mexicana
de Ingeniería Biomédica

# Sociedad Mexicana de Ingeniería Biomédica

La Mesa Directiva de la Sociedad Mexicana de Ingeniería Biomédica hace una extensa invitación a las personas interesadas en participar, colaborar y pertenecer como Socio Activo de la SOMIB. La SOMIB reúne a profesionistas que se desarrollan en áreas de Ingeniería Biomédica, principalmente ingenieros biomédicos, así como otros profesionistas afines con el desarrollo de tecnología para la salud.

### Membresía Estudiante

***$1,400.00* PESOS MXN**

15% de descuento para grupos de 5 o más personas.

### Membresía Profesionista

***$2,400.00* PESOS MXN**

15% de descuento para grupos de 5 o más personas.

### Membresía Institucional

***$11,600.00* PESOS MXN**

No aplica descuento.

### Membresía Empresarial

***$20,000.00* PESOS MXN**

No aplica descuento.

EL PAGO CUBRE UN AÑO DE CUOTA. EN CASO DE REQUERIR FACTURA
FAVOR DE SOLICITARLA, ADJUNTANDO COMPROBANTE DE PAGO Y ESPECIFICANDO CONCEPTO,
AL CORREO ELECTRÓNICO: ***gerencia@somib.org.mx***

**ib** Revista Mexicana de Ingeniería Biomédica

### AUTORES

Los trabajos a publicar en la RMIB, deben ser originales, inéditos y de excelencia. Los costos de publicación para autores son los siguientes:

| | |
|---|---|
| **NO SOCIOS:** | ***$5,000* PESOS MEXICANOS** |
| **SOCIOS:** | ***$4,000* PESOS MEXICANOS** |

### PUBLICIDAD

A las empresas e instituciones interesadas en publicitar su marca o productos en la RMIB, los costos por número son los siguientes:

| | |
|---|---|
| **MEDIA PLANA:** | ***$4,999.00* PESOS MXN** (INCLUYE I.V.A.) |
| **UNA PLANA:** | ***$6,799.00* PESOS MXN** (INCLUYE I.V.A.) |
| **CONTRAPORTADA:** | ***$7,799.00* PESOS MXN** *(INCLUYE I.V.A.)* |
| **FORROS INTERIORES:** | ***$7,799.00* PESOS MXN** *(INCLUYE I.V.A.)* |

DESCUENTO DEL **20%** AL CONTRATAR PUBLICIDAD EN DOS O MÁS NÚMEROS.

## Para ser socio

- Realiza el pago de derechos, de acuerdo a la categoría que te corresponde.
- Ingresa a **www.somib.org.mx/membresias** y elige el tipo de membresía por el cual reali- zaste el pago de derechos.
- Completa el formulario correspondiente y envíalo.
- Se emitirá carta de aceptación y número de socio por parte de la mesa directiva (aprobada la solicitud).
- Para mayor información sobre beneficios, ingresa a **www.somib.org.mx**; o escribe a **gerencia@somib.org.mx.**

## Datos bancarios

- **Beneficiario:** Sociedad Mexicana de Ingeniería Biomédica A. C.
- **Banco:** *Scotiabank*
- **Referencia:** *1000000333*
- **Cuenta:** *11006665861*
- **CLABE Interbancaria:** *044770110066658614*

## INFORMES

Juan Vázquez de Mella #481,
Polanco I Sección,
Alc. Miguel Hidalgo, C. P. 11510, Ciudad de México, México,
(555) 574-4505
rib.somib@gmail.com

# Revista Mexicana de Ingeniería Biomédica

**Fundador**
**Dr. Carlos García Moreira**

## COMITÉ EDITORIAL

**Editora en Jefe**
**Dra. Dora-Luz Flores**
UNIVERSIDAD AUTÓNOMA DE BAJA CALIFORNIA

### Editores Asociados Nacionales

**Dr. Christian Chapa González**
UNIVERSIDAD AUTÓNOMA DE CIUDAD JUÁREZ

**Dra. en C. Citlalli Jessica Trujillo Romero**
DIVISIÓN DE INVESTIGACIÓN EN INGENIERÍA MÉDICA
INSTITUTO NACIONAL DE REHABILITACIÓN "LUIS GUILLERMO IBARRA IBARRA"

**Dr. Rafael Eliecer González Landaeta**
UUNIVERSIDAD AUTÓNOMA DE CIUDAD JUÁREZ

**Dra. Rebeca Romo Vázquez**
UNIVERSIDAD DE GUADALAJARA

**Dra. Isela Bonilla Gutiérrez**
UNIVERSIDAD AUTÓNOMA DE SAN LUIS POTOSÍ

### Comité Editorial Internacional

**Dr. Leonel Sebastián Malacrida Rodríguez**
UNIVERSIDAD DE LA REPÚBLICA, URUGUAY

**Dra. Elisa Scalco**
INSTITUTE OF BIOMEDICAL TECHNOLOGY
ITALIAN NATIONAL RESEARCH COUNCIL, MILAN, ITALY

**Dra. Natali Olaya Mira**
INSTITUTO TECNOLÓGICO METROPOLITANO
ITM, MEDELLÍN, COLOMBIA

**Índices**
La Revista Mexicana de Ingeniería Biomédica aparece en los siguientes índices científicos:
**Sistema de Clasificación de Revistas Científicas y Tecnologías del CONACYT - Q4, SCOPUS, SciELO, EBSCO, LATINDEX, Medigraphic Literatura Biomedica, Sociedad Iberoamericana de Información Científica - SIIC.**

*www.rmib.mx*
**ISSN 2395-9126**

**Asistente Editorial**
**Carla Ivonne Guerrero Robles**

**Editor Técnico y en Internet**
**Sandra Sánchez Jáuregui**

Se autoriza la reproducción parcial o total de cualquier artículo a condición de hacer referencia bibliográfica a la Revista Mexicana de Ingeniería Biomédica y enviar una copia a la redacción de la misma.

SOMIB
Sociedad Mexicana de Ingeniería Biomédica

**Sociedad Mexicana de Ingeniería Biomédica**
Juan Vázquez de Mella #481, Polanco I Sección, Alc. Miguel Hidalgo, C. P. 11510, Ciudad de México, México, (555) 574-4505

SOMIB
Sociedad Mexicana
de Ingeniería Biomédica

**Afiliada a:**
International Federation of Medical and Biological Engineering (IFMB-IUPSM-ICSU)
Federación de Sociedades Científicas de México, A.C. (FESOCIME)
Consejo Regional de Ingeniería Biomédica para América Latina (CORAL)

**SOMIB**
Juan Vázquez de Mella #481, Polanco I Sección, Alc. Miguel Hidalgo, C. P. 11510, Ciudad de México, México (555) 574-4505
*www.somib.org.mx*

**REVISTA MEXICANA DE INGENIERÍA BIOMÉDICA,** Vol. 46, No. 1, January-April 2025, es una publicación cuatrimestral editada por la Sociedad Mexicana de Ingeniería Biomédica A.C., Juan Vázquez de Mella #481, Polanco I Sección, Alc. Miguel Hidalgo, C. P. 11510, Ciudad de México, México, (555) 574-4505, www.somib.org.mx, rib.somib@gmail.com. Editora responsable: Dra. Dora-Luz Flores. Reserva de Derechos al Uso Exclusivo No. 04-2015-041310063800-203, ISSN (impreso) 0188-9532; ISSN (electrónico) 2395-9126, ambos otorgados por el Instituto Nacional del Derecho de Autor. Responsable de la última actualización de este número: Lic. Sandra Sánchez Jáuregui, Juan Vázquez de Mella #481, Polanco I Sección, Alc. Miguel Hidalgo, C. P. 11510, Ciudad de México, México, (555) 574-4505, fecha de última modificación, 31 de agosto del 2023.

**Disponible en línea:**
*www.rmib.mx*

# CONTENTS
## *CONTENIDO*

RESEARCH ARTICLE

VOL. 46 | NO. 1 | **JANUARY-APRIL 2025** | PP 1 - 15

Revista Mexicana de Ingeniería Biomédica

https://dx.doi.org/10.17488/RMIB.46.1.1

E-LOCATION ID: e1456

# Fast Computational Modeling Based on the Boundary Element Method Towards the Design of an Ultrasonic Biomedical Applicator

## Modelado Computacional Rápido Basado en el Método del Elemento de Frontera Hacia el Diseño de un Aplicador Biomédico Ultrasónico

*Raquel Martínez-Valdez*[1], *Ivonne Bazán*[1]

[1]Universidad Autónoma de Aguascalientes, Aguascalientes - México

## ABSTRACT

The aim of this work is to analyze the usage of the boundary element method (BEM) as a fast computational tool for solving large ultrasonic field problems, *i.e.* 3D models. A proposed tridimensional radiating surface $S_R$ was modeled by means of BEM and the finite element method (FEM). Four time-harmonics models were developed: two containing the entire $S_R$ and two considering a symmetrical plane at half-length of the radiator. BEM solutions were validated with FEM models by contours at -3 dB and -6 dB pressure decays, areas within the contours, elliptical shape ratio $E_r$ and ellipsoidal focal volume approximations. The average differences in pressure and distance at the focus were 39.875 Pa and 0.4515 mm, respectively; the areas within the contours show differences between 0.6 mm$^2$ and 2.3 mm$^2$. The $E_r$ of the focal zone was over 92 %, while the ellipsoidal volume approximation showed differences between 0.0817 mm$^3$ to 1.4632 mm$^3$ at -3 dB, and 1.2354 mm$^3$ to 4.1144 mm$^3$ at -6 dB. Analyzed data suggest the use of BEM to model the ultrasonic beam pattern in a lossless medium during ultrasonic biomedical applicators design, reducing the solution time from 22 h with FEM to 2 min with BEM.

**KEYWORDS:** **acoustic field modeling, boundary element method, finite element method, focused ultrasound, ultrasonic biomedical applicators**

## RESUMEN

El objetivo de este trabajo es analizar el uso del método del elemento de frontera (BEM) como una herramienta computacional rápida para resolver campos acústicos en modelos 3D. Una superficie radiante tridimensional $S_R$ propuesta se modeló por medio de BEM y del método del elemento finito (FEM). Se desarrollaron 4 modelos en el dominio de la frecuencia: 2 con la $S_R$ completa y 2 considerando un plano de simetría a la mitad de $S_R$. Los modelos BEM se validaron con los modelos FEM por medio de contornos de presión a -3 dB y -6 dB, áreas dentro de los contornos, relación de forma elíptica $E_r$ y aproximación elipsoidal focal. Las diferencias promedio en presión y distancia focales fueron 39.875 Pa y 0.4515 mm, respectivamente; las áreas dentro de los contornos mostraron diferencias entre 0.6 $mm^2$ y 2.3 $mm^2$. La $E_r$ focal fue >92 %, mientras que la aproximación volumétrica elipsoidal mostró diferencias entre 0.0817-1.4632 $mm^3$ a -3 dB, y 1.2354-4.1144 $mm^3$ a -6 dB. Los resultados sugieren el uso de BEM para modelar el patrón acústico en medios sin pérdidas durante el diseño de aplicadores biomédicos ultrasónicos reduciendo el tiempo de solución de 22 h (FEM) a 2 min (BEM).

**PALABRAS CLAVE:** aplicador biomédico ultrasónico, método del elemento finito, método del elemento de frontera, modelado de campo acústico, ultrasonido focalizado

**Corresponding author**

TO: Raquel Martínez-Valdez
INSTITUTION: **UNIVERSIDAD AUTÓNOMA DE AGUASCALIENTES**
ADDRESS: **AV. UNIVERSIDAD 940, 20131, AGUASCALIENTES, MÉXICO**
EMAIL: raquel.martinez@edu.uaa.mx

**Received:**

04 September 2024

**Accepted:**

26 November 2024

**Published:**

1 January 2025

## INTRODUCTION

Ultrasounds are mechanical waves with frequencies above 20 kHz that interact with media while propagating through them. This interaction produces phenomena such as reflection, refraction, scattering, attenuation, changes in sound velocity, that may modify the acoustic pattern[1]. The ultrasonic beam pattern is strongly dependent on the geometry of the vibrating element, the operating frequency, the transducer manufacturing, among others[2]. Ultrasonic waves energy can be concentrated in a well-located target or region named focus; this pattern modification is called focused ultrasound (FUS). The focalization can be achieved by different ways such as acoustic lenses, spherically concave transducer, electronic or phase-controlled arrays, etc.[3][4][5].

Focused ultrasound as being a non-ionizing energy and having the capacity to deliver it in a specific region has been under research for clinical applications for the last decades. FUS interaction with biological tissue can cause thermal effects[6], nonthermal effects such as cavitation[4][6], nonlinear behavior like both focus distortion and shift[7][8] and bioeffects[9][10]. In cancer treatment, the main goal is to achieve a thermal effect due to FUS by increasing the target temperature over 60 °C which results in coagulative necrosis of tissue[4][10]. Extracorporeal and intracavitary devices based on FUS are mostly designed to treat malignant tumors in liver, kidney, prostate, breast, osteosarcoma, etc.[3][4][5][11]. FUS in cancer treatment acts as adjuvant method to radiotherapy, immunotherapy and chemotherapy[12]. Yet, FUS effects are being studied in brain malignancies like Parkinson disease, to open the blood brain barrier, in the treatment of varicose veins, and cancer treatment, among others[13][14][15][16].

Computational modeling is a powerful tool used to solve multi-physic scenarios and get an approximate behavior of a real problem. The adequate selection of the physics involved in the problem, the geometry definition, physics configuration and boundary conditions settings influence the numerical solution. The finite element method (FEM) allows acoustic propagation modeling; but it can be challenging when solving high-frequency large geometries, *i.e.*, tridimensional spaces[17][18]. In acoustics, FEM requires both domain and subdomains discretization with a mesh containing frequency-dependent element size; at least 10 elements per wavelength (λ)[19][20][21]. In large or multidomain geometries, this results in an increase of the number of degrees of freedom which implies more computational resources and longer solution time[17][22]. On the other hand, the boundary element method (BEM) allows faster computation of acoustic radiation patterns in large or infinite homogeneous domains, called exterior problems [23][24][25]. BEM modeling requires surface or boundary discretization[21][23] the problem solution is obtained by using the free-space Green's function which satisfies the Sommerfeld radiation condition at infinity [21][26]. However, in multidomain geometries, BEM is limited to obtain the scattering effect over the surface of a subdomain located in the radiation pattern, but not the propagation within it: interior problem $\Omega_i$[23]. Even though BEM can be used to solve interior problems or the acoustic propagation in a finite domain, FEM solution is more approximated to the real problem[23][27]. Meanwhile, when it refers to obtaining the solution of the exterior problem $\Omega_e$, FEM has limitations due to infinite domain discretization[23]; cost-effectiveness comparison between both methods has been analyzed by Harari, I. and Hughes, T.[27]. Nonetheless, both methods present advantages in time-harmonics acoustic model solution whether is the interior problem or the exterior problem. In this sense, Assaad *et al.*[28] combined FEM solution with BEM to compute the radiation pattern of a circular transducer in a 3D geometry and its 2D axisymmetric approach. In order to do so, the authors modeled the piezoelectric vibration by means of FEM, while the load and propagation in water was modeled by BEM. The electrical impedance modulus and the far-field directivity obtained from the 3D model and the axisymmetric model show an identical pattern. Furthermore, the electric impedance of 3D model of a square shape transducer was compared with measurements that showed good agreement, except in the impedance peaks amplitude.

In the biomedical field, Santiago, A. *et al.* compared BEM and FEM methods to estimate the elastic modulus from two generated images including a lesion: one without pressure applied and one after a known pressure was applied[29]. By using the same quantity of elements to mesh the geometries, FEM results showed a higher mean error in elastic modulus estimation of the lesion than BEM results. Computational time to obtain the solutions by FEM was around 6 h compared to 60 s by using BEM[29]. Recently, Shen, F. *et al.*[18] combined BEM and FEM to simulate the transcranial acoustic propagation in 2D and 3D models. For the 2D model, the transducer was a linear array with 512 elements; while for the 3D model, the ultrasonic transducer consisted of 256 source points. Both models were solved at 0.5 MHz. The implemented algorithms corrected the focus shift due to aberration and attenuation caused by the skull bone[18].

In ultrasonic thermal ablation applications, Gélat, P. *et al.*[30] used BEM to model both acoustic and scattered fields of a high intensity focused ultrasound (HIFU) phased array through human ribs for liver and pancreatic tumors treatment. The transducer consisted of 256 circular elements placed over a spherical bowl shape with a focal distance of 18 cm, each element had a 3 mm radius, and a frequency of 1 MHz. The study was performed considering two possible scenarios of beam focusing at 3 cm depth from the rib cage: intercostally (between ribs 10-11) or transcostally (across rib 10). The total acoustic pressure or surface velocity of each element was average to attain a focal peak pressure of 10 MPa in free space. For the intercostal case, for the desired focal pressure, the deposited acoustic pressure in the surface of the ribs reached 2.2 MPa, while it reached 1.8 MPa for the transcostal case. In both cases, that amount of acoustic pressure could lead to both bone and tissue heating. Afterwards, Gélat, P. *et al.*[31] solved the inverse problem by adding constraints to their previous work in order to achieve a desired acoustic pressure on the surface of the rib like geometry. Among the constraints, the element velocity magnitudes were defined in a specific dynamic range. The dimensions of the geometry were modified to work with a frequency of 100 kHz, this way a faster solution could be obtained. The results showed both the velocity magnitude and phase distributions of the 256 elements considering the constraint which could lead to the desired acoustic pressure deposition in the rib like geometry. Furthermore, Van't Wout, E. *et al.*[17] implemented a numerical preconditioner to reduce the converge time of the BEM formulation developed in[30][31] for targeting tissue behind the ribcage. The improved algorithms solved a realistic transcostal simulation of HIFU propagation within 2 h by using a desktop computer with 12 processors and 80 GB RAM [17]. A more complex model for HIFU field radiation through the ribcage based on BEM was reported by Haqshenas, S.R. *et al.*[22]; which included an idealized subdomain with fat layer properties besides two ribs. The HIFU transducer geometry, the operating frequency, focal distance and ribcage position were the same as reported in[30]; the fat layer was placed between 1 mm to 15 mm from the ribs facing the transducer[22]. The reported results showed that the maximum pressure reached at the focus diminished in 15 % considering transmission and scattered field through the ribs compared with the maximum pressure without scatters. After adding the fat layer before the ribs, the focal pressure showed a decay of 50 %. Moreover, Haqshenas, S.R. *et al.*[22], implemented an algorithm to couple the exterior problem solution with the interior problem calculation to model the HIFU beam inside a kidney subdomain surrounded by a fat layer. Numerical modeling results showed that the focal pressure decreased 5 % in the absence of the fat layer, and after considering the fat subdomain, the focal pressure within the kidney decreased 15 %. Computational time for solving the two anatomical models (fat-ribs and fat-layer) varied from 82 min to 40 h approximately in a workstation with 32 processors and 512 GB RAM[22].

## Motivations of the study

The aim of this work is to compare the acoustic pattern of a 3D focused radiating surface in a homogeneous lossless medium by means of both FEM and BEM towards the design of biomedical ultrasonic extracorporeal applicators intended for thermal ablation therapies. The proposed geometry represents a high computational cost, and its shape differs from the typical geometries used in HIFU applications. Therefore, a simplified medium was considered because our priority was to evaluate the performance of both methods in terms of ultrasonic radiation pattern characteristics and solution time. Besides, its potential clinical application could reach areas where a high focal acoustic energy is required at medium or high penetration depths. Additionally, the focal dimension should be capable of attaining targets smaller than 20 mm with an applicator of, relatively, both reduced dimensions and fabrication costs. Nowadays, to our best knowledge, in the ultrasonic ablation field, BEM use has been restricted to liver ablation through the ribcage with 256 elements placed in a spherical bowl applicator and there are few studies oriented to compare BEM and FEM performances in the biomedical field.

## MATERIALS AND METHODS

In order to achieve the goal of this work, four tridimensional time-harmonic acoustic models for a defined radiating surface were developed. Two models consisted of solving the acoustic pattern of the entire radiating surface; while the remaining models considered only half of the geometry, *i.e.* by defining a symmetrical plane. Moreover, FEM and BEM models were performed for each case: entire geometry and symmetric geometry, as described in Table 1. The symmetric cases represent a simplified version of the entire geometry cases that reduce computational resources. The radiating surface $S_R$ proposed in all four cases simulates a mono-element transducer with the shape of a cylindrical section with a radius $R$ = 15 mm, a width $b$ = 6 mm and the length of the chord of the cylindrical boundary $a$ = 21.22 mm (see Figure 1A). The origin coordinates (x, y, z) for all models were set at (x=0, y=0, z=0), as shown in Figure 1A. The models were solved by using the acoustic module of COMSOL Multiphysics 5.4® on a PC with Intel® Xeon® W-2133 CPU @ 3.60 GHz processor and 128 GB RAM.

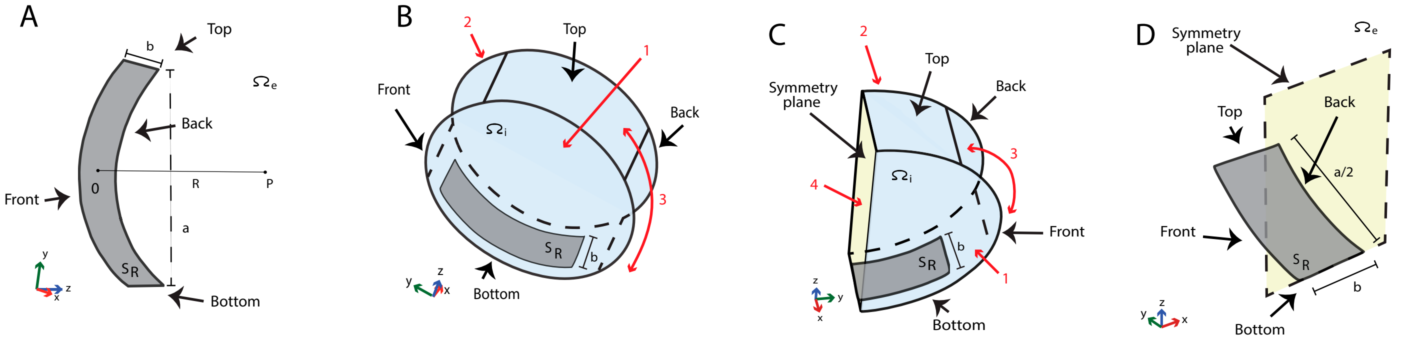


**FIGURE 1. Geometries designed to solve the FEM and BEM models. A) Proposed geometry of the radiating surface, and Case C: entire radiating surface $S_R$ emitting in an infinite void with water acoustic properties to solve the exterior problem, B) Case A: cylindrical domain with entire radiating surface $S_R$ emitting in a finite domain: interior problem, C) Case B: symmetrical geometry obtained from cutting the geometry of case A in half used to solve the interior problem, D) Case D: symmetrical geometry obtained from cutting the geometry of case C in half used to solve the exterior problem. The dimensions of the radiating surface are the same in all four models; the geometrical origin was located at coordinates (x=0, y=0, z=0). The propagation medium was set as water for both interior and exterior problems, *i.e.* FEM and BEM models, respectively.**

TABLE 1. **Model identifier based on geometry type and computational method used to obtain the solution.**

| Geometry | FEM | BEM |
|---|---|---|
| Entire | Case A | Case C |
| Symmetrical | Case B | Case D |

## FEM acoustic radiation modeling

Figures 1B and 1C show the geometries used to model the acoustic beam pattern in the interior problem, where $\Omega_i$ represents the finite domain, *i.e.* cases A and B. The time-harmonic wave equation for pressure acoustic propagation used for FEM modeling is represented by Equation (1) as

$$\nabla^2 p + \left(\frac{\omega^2}{c_s^2}\right) p = 0 \tag{1}$$

where $p$ is the acoustic pressure (Pa), $\omega$ is the angular frequency (rad/s), and $c_s$ is the sound velocity (m/s). The domain consists of a cylinder of 15 mm of radius and a height of 10 mm. The radiating surface center is placed in the middle of the lateral side of the cylinder, *i.e.* there is a 2 mm distance from $S_R$ to each base of the cylinder. The domain acoustic properties were defined as to emulate water which is considered as a lossless medium[32]; the sound velocity was set at 1500 m/s and the density at 1000 kg/m$^3$. The boundary conditions used to solve the model were configured as acoustic impedance, $Z$ (Rayls), in both bases and lateral side of the cylindric domain (boundaries 1, 2 and 3; see Figure 1B). The acoustic impedance is defined by Equation (2) as

$$Z = \rho * c_s \tag{2}$$

where $\rho$ is the medium density (kg/m$^3$); therefore, $Z$ = 1.5 MRayls which corresponds to water to decrease wave reflection[32]. The operating frequency was set at 1 MHz. Besides, the boundary condition of the $S_R$ was set as acoustic pressure with an arbitrary initial value of 500 Pa.

In case B, as the geometry used in the model is half of the geometry in case A, a symmetry boundary condition was configured in the cross-sectional plane of the cylindric domain (boundary 4; see Figure 1C). Boundaries 1, 2 and 3 remained as impedance condition with water acoustic properties. FEM models were solved by using a mesh consisting of tetrahedral elements of quadratic order. A convergence analysis was realized in both cases; therefore, the element size was varied from *λ/6* up to *λ/12* for case A and from *λ/6* up to *λ/14* for case B. The resulting meshes consisted of 7, 942, 317 elements and 63, 823, 707 elements, respectively, for case A, and 3, 946, 491 elements and 50, 797, 081 elements, respectively, for case B. For case A, the solution time for *λ/12* mesh was 1 day 2 h 32 min 15 s (123.74 GB physical memory/192.68 GB virtual memory); for *λ/6* mesh was 14 min 39 s (53.12 GB physical memory/58.34 GB virtual memory). For case B, the solution time for *λ/14* was 14 h 19 min 9 s (123.8 GB physical memory/165.96 GB virtual memory); for *λ/6* was 5 min 55 s (23.22 GB physical memory/25.5 GB virtual memory). Both cases were tried to be solved with finer meshes, but computational resources were not enough.

## BEM acoustic radiation modeling

The geometries depicted in Figures 1A and 1D were used to solve the acoustic beam pattern by means of BEM. The main geometry in the models represents the entire radiating surface or half of it; and the dimensions $b$, $a$ and $R$ are the same as in cases A and B. In this scenario, there is no need to define a finite domain, because BEM is used to solve the exterior problem $\Omega_e$; however, the infinite void (space) that surrounds $S_R$ has the acoustic properties of water. The governing Helmholtz equation defined for BEM modeling is given by Equations (3), (4) and (5) as follows

$$\nabla \cdot \left(-\frac{1}{\rho_c}\nabla p_t\right) - \frac{k_{eq}^2}{\rho_c} p_t = 0 \tag{3}$$

$$p_t = p + p_b \tag{4}$$

$$k_{eq}^2 = \left(\frac{\omega}{c_c}\right)^2 \tag{5}$$

where $p_t$ is the total acoustic pressure (Pa), $k_{eq}$ is the wave number (rad/m), $\rho_c$ is the complex density of the medium (kg/m$^3$), $c_c$ is the speed of sound (m/s) that could have complex values depending on the problem, $p_b$ is the background pressure wave (Pa) and $p$ is the pressure (Pa). For both cases C and D, $c_c = c$, $\rho_c = \rho$, and $p_t = p$; after substituting the previous values in eqs. (3)-(5), the governing equation results in eq. (1). Similarly to cases A and B, the boundary condition of $S_R$ was set as acoustic pressure with an arbitrary value of 500 Pa. In case D, half of the radiating surface is modeled; therefore, in axis $y = 0$, an infinite symmetry plane was configured as shown in Figure 1D. For both cases C and D, the condition at infinity was set as an outgoing wave, in order to satisfy the Sommerfeld radiation condition[21][26]. As well as in FEM models, the operating frequency was set at 1 MHz.

BEM models were solved by using a mesh consisting of triangular elements of quadratic order. A convergence analysis was performed in both cases by varying the boundary element size from $\lambda/6$ up to $\lambda/20$. The resulting meshes for case C consisted of 5, 472 elements and 59, 244 elements, respectively, and 200 elements and 29, 724 elements, respectively, for case D. For case C, the solution time for $\lambda/20$ mesh was 1 min 20 s (3.75 GB physical memory/4.41 GB virtual memory), and for $\lambda/6$ mesh was 10 s (1.61 GB physical memory/2.2 GB virtual memory). For case D, the solution time for $\lambda/20$ mesh was 58 s (2.56 GB physical memory/3.09 GB virtual memory), and for $\lambda/6$ mesh was 12 s (1.31 GB physical memory/1.89 GB virtual memory).

## BEM model validation

After performing the convergency analysis, BEM results were validated with FEM results by using the models solved with a $\lambda$ /10 mesh as reference[19][21]. Then, focus length and width at -3 dB pressure decay, full width at half maximum (FWHM) pressure, and full length at half maximum (FLHM) pressure were calculated. Contours at -3 dB and -6 dB decay and their corresponding areas were obtained in *zy*-, *xy*-, and *zx*- planes for all cases. The focal region produced by high-intensity focused ultrasound (HIFU), typically, has an ellipsoidal shape[4]; however, its shape depends on the radiator geometry and its frequency, among other parameters[2][33]. Therefore, an elliptical shape ratio $E_r$ was calculated in the three planes of each model to obtain the similarity percentage of the contours to an ellipse value; where $E_r$ = 100 % is the ideal[32]. Finally, ellipsoidal approximations of focal volumes at -3 dB and -6 dB were estimated by means of the major and minor axis of the contours at *zy*- and *xy*- planes.

## RESULTS AND DISCUSSION

The convergence analysis was carried out by obtaining the maximum acoustic pressure $P_{max}$ value along *zy*- plane of each model. Then, the relative error was calculated by taking the solution of the model meshed with an element size of λ/10 as reference[32]; where $P_{max}$ started to show stability for each case (see Table 2). For case A, the model was solved for meshes from λ/6 up to λ/12, due to available computational resources, the relative errors obtained were 0.061 % and 0.032%, respectively. For case B, the model was solved for meshes from λ/6 up to λ/14, due to computational resources, the relative errors calculated were 0.021 % and 0.057 %, respectively. For case C, the simulations were done from λ/6 up to λ/20, the relative errors calculated were 0.111 % and 0.064 %, respectively. Finally, for case D, the simulations were done from λ/6 up to λ/20, and the relative errors obtained were 4.141 % and 0.126 %, respectively. From now on, the results presented correspond to the models solved with a λ/10 mesh for all cases.

TABLE 2. **Convergence analysis results for the four cases. Solution from λ/10 mesh models was used as reference for calculating the relative error.**

| Case | Mesh size | No. elements | DOF | Solution time | Relative error [%] |
|---|---|---|---|---|---|
| A | λ/6 | 7942317 | 10673982 | 14 min 39 s | 0.062 |
| | λ /8 | 18862851 | 25293358 | 44 min 1 s | 0.035 |
| | λ /10 | 36888845 | 49391271 | 9 h 2 min 7 s | - |
| | λ /12 | 63823707 | 85374660 | 1 day 2 h 32 min 15 s | 0.032 |
| B | λ /6 | 3946491 | 5314575 | 5 min 55 s | 0.022 |
| | λ /8 | 9394791 | 12616975 | 14 min 34 s | 0.005 |
| | λ /10 | 18402298 | 24664459 | 40 min 4 s | - |
| | λ /12 | 31828411 | 42619267 | 5 h 37 min 9 s | 0.038 |
| | λ /14 | 50797081 | 67965941 | 14 h 19 min 9 s | 0.057 |
| C | λ /6 | 5472 | 14191 | 10 s | 0.111 |
| | λ /8 | 9660 | 24821 | 14 s | 0.051 |
| | λ /10 | 15216 | 38883 | 21 s | - |
| | λ /12 | 21360 | 54403 | 28 s | 0.013 |
| | λ /14 | 29340 | 74525 | 38 s | 0.034 |
| | λ /16 | 37644 | 95445 | 48 s | 0.047 |
| | λ /18 | 48048 | 121627 | 1 min 2 s | 0.059 |
| | λ /20 | 59244 | 149777 | 1 min 20 s | 0.064 |
| D | λ /6 | 200 | 577 | 3 s | 4.141 |
| | λ /8 | 4844 | 12495 | 11 s | 0.014 |
| | λ /10 | 7540 | 19331 | 16 s | - |
| | λ /12 | 10756 | 27467 | 22 s | 0.077 |
| | λ /14 | 14676 | 37363 | 29 s | 0.050 |
| | λ /16 | 18864 | 47925 | 36 s | 0.118 |
| | λ /18 | 23988 | 60831 | 46 s | 0.138 |
| | λ /20 | 29724 | 75267 | 58 s | 0.127 |

Figure 2A depicts the pressure distribution along *z*- axis for all cases and it is observed that the acoustic pressure reaches its maximum near 15 mm. The acoustic pressure difference $\Delta P_{max}$ at its maximum value was 41.101 Pa between cases A and C which correspond to the models with the entire radiating surface, which represents an error of 1.48 % taking as reference the case A (FEM complete geometry). The distance difference $\Delta z$ between $P_{max}$ of both cases was 0.451 mm. For the symmetrical models, cases B and D, $\Delta P_{max}$ was 38.649 Pa that represents an error of 1.37 % taking as reference case B (FEM symmetric geometry) and $\Delta z$ was 0.452 mm. Figures 2B and 2C show the pressure distribution along $y$- axis and $x$- axis at the $z$ position of $P_{max}$, where $P_{max}$ is centered at $x = 0$ and $y = 0$ in accordance with the proposed geometry. Figure 2B depicts that the pressure distribution along $y$- axis is narrower than in the $x$- axis, see Figure 2C. This behavior is strongly related to the geometrical shape of the surface radiator [2][33], presenting an oval shape in $xy$- plane at $z$ position of $P_{max}$ (see Figure 3) with its mayor axis over $x$- axis and differentiating it form the classical HIFU focal shape.

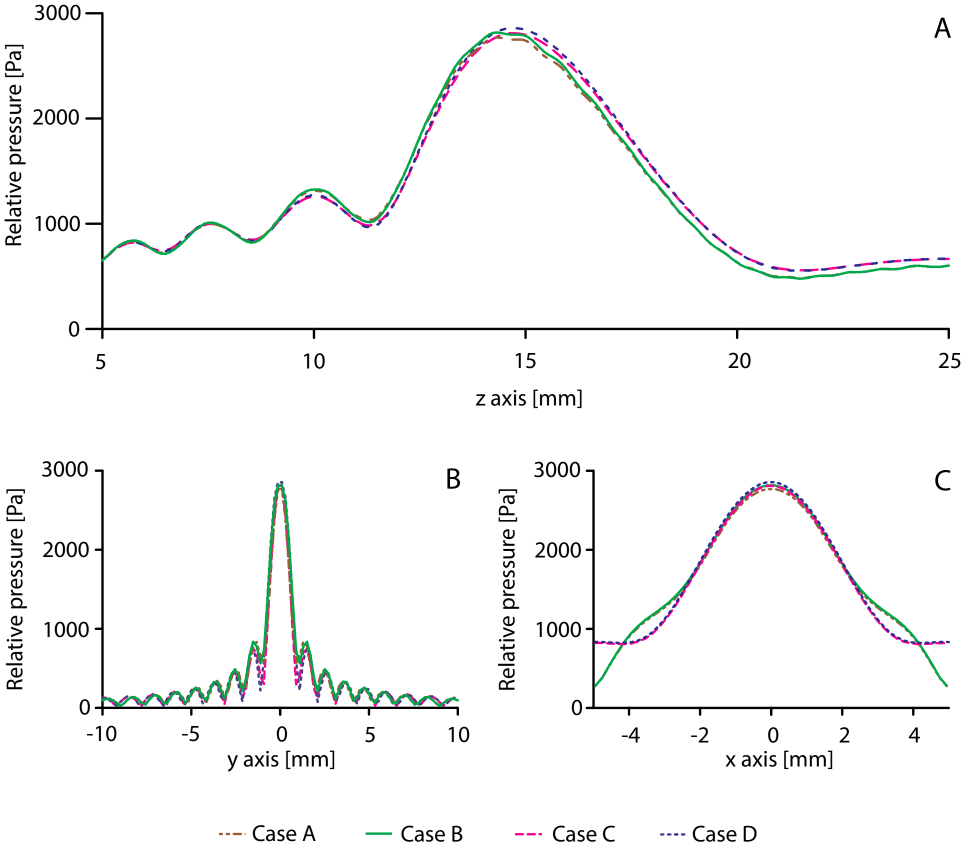


**FIGURE 2. Pressure distribution (A) along *z*- axis, (B) along *y*- axis, and (C) along *x*- axis. All three pressure distributions were taken at the maximum acoustic pressure value $P_{max}$ obtained from *zy*- plane.**

In this sense, it is worth noting that the focal region dimensions in each plane are also different as shown in Figure 3. In Figures 3A-3C depicts the contours at -3 dB $P_{max}$ decay in $zy$-, $xy$-, $zx$- planes for all four cases. In Figure 3A, the contours show an elliptic focal shape, where case A and case B present at slight displacement towards the left (surface radiator location). These displacements are related with the mentioned $\Delta z$ between cases A and C, and cases B and D; while, in Figures 3B and 3C, all contours are centered in $x = 0$ and $y = 0$; *i.e.* no error was found over x and y positions of $P_{max}$. Figures 3D and 3F show the contours at -6 dB pressure decay in $zy$-, $xy$-, $zx$- planes. From Figure 3D, it can be observed an enlarged region towards the left ($S_R$ location) which describes the beam path through focalization. Also, isolated contours can be appreciated for cases A and B (FEM modeling) but not for cases C and D (BEM modeling). In a clinical application, such as FUS thermal ablation, the enlarged area and the appearance of the isolated contours could produce "hot spots" or undesired hot regions[34] at the zone nearby the transducer and its prediction is a fundamental step to avoid patient damage. Figure 4 depicts the pressure distribution along $y$- axis at the position in $z$- axis where the isolated contours are presented ($z$ = 10.82 mm). The pressure peaks reach a maximum value of 1456.03 Pa for case A, and 1445.33 Pa for case B, that represent half decay of $P_{max}$. Focal dimensions are usually measured at half decay of $P_{max}$[35] which could directly relate to the size of the produced thermal lesion. However, the thermal lesion size depends on the exposure time to ultrasound energy, cross-sectional intensity accumulation, tissue absorption coefficient, non-linear effects, frequency, among others[35]. In Figure 3E it can be observed that the focal regions keep an elliptical fashion as in Figure 3B. In Figure 3F, the contour shows an enlargement along $z$- axis from 14 to 18 mm and along $x$- axis which could produce non-uniform thermal lesion, and care must be taken in order to avoid healthy tissue damage.

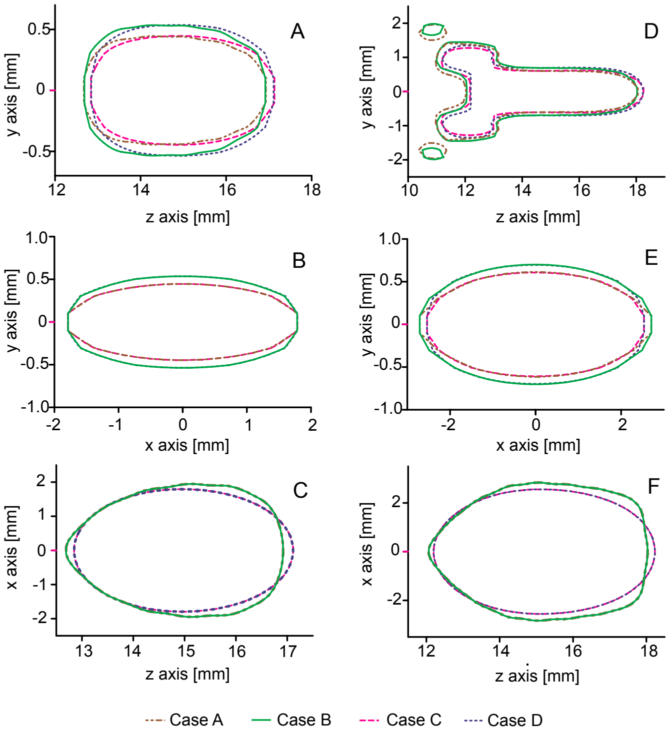


**FIGURE 3.** **Pressure contours at -3 dB and -6 dB decays for all four cases. A) -3 dB contours in *zy*- plane, B) -3 dB contours in *xy*- plane, C) -3 dB contours in *zx*- plane, D) -6 dB contours in *zy*- plane, E) -6 dB contours in *xy*- plane, and F) -6 dB contours in *zx*- plane.**

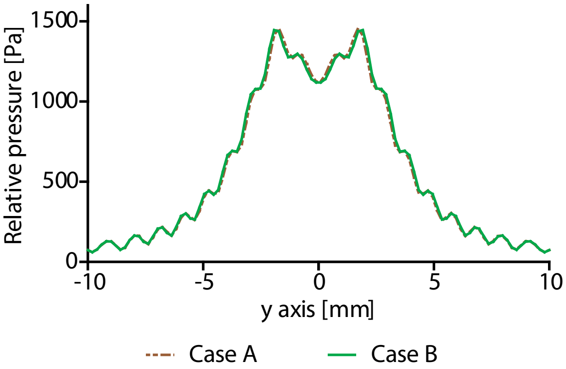


**FIGURE 4.** **Pressure distribution along y- axis at z = 10.82 mm, location of the two isolated contours.**

Table 3 shows the lengths and widths in *zy*-, *xy*- and *zx*- planes calculated at -3 dB and -6 dB (FLHM and FWHM) pressure decays for all four cases. As mentioned before, both solution time and memory resources increase when solving high-frequency large FEM models; therefore, a model simplification must be pursued to reduce computational cost without compromising the approximated solution. Hence, results from cases A and B (FEM) were taken as references to evaluate cases C and D (BEM) respectively. Case B represents the simplification of case A, which could be considered as the more realistic model in this study; as well as Case D represents simplification of case C. Then, the most representative ratio (*i.e.* equal to one when ideal) in length ($\text{length}_{\text{BEM-Case (C or D)}}/\text{length}_{\text{Ref-Case (A or B)}}$) and width ($\text{width}_{\text{BEM-Case (C or D)}}/\text{width}_{\text{Ref-Case (A or B)}}$) were found at -3 dB and -6 dB pressure decays. In both cases (C/A and D/B), the worst ratio values were found at *zx*-plane width (0.92 for C/A and D/B at -3 dB and 0.90 for C/A and D/B at -6 dB) and at *xy*-plane length at -6 dB (0.94 for C/A and D/B), representing errors up to 9.87 %. All other ratios approximate unity with an error minor to 1.7 %. Additionally, it is important to address that both lengths in the *zy*- and *zx*- planes should agree as both distances represent the focus extent in the propagation axis *z*. In the same way, the widths in both *zy*- and *xy*- planes must have close values because they correspond to the focal width along *y*- axis. In consequence, the length in *xy*- planes and the width in *zx*- planes present close values due to its relationship with the dimension in the *x*- axis (see Figure 3 and Table 3).

**TABLE 3. Focal length and width in *zy*-, *xy*- and *zx*- planes calculated at -3 dB and -6 dB (FLHM and FWHM) pressure decays for the four cases.**

| Pressure level | Dimension | Plane | Case A | Case B | Case C | Case D |
|---|---|---|---|---|---|---|
| -3 dB | Length [mm] | *zy* | 4.2586 | 4.2296 | 4.2909 | 4.2631 |
| | | *xy* | 3.5695 | 3.5732 | 3.5743 | 3.5751 |
| | | *zx* | 4.2270 | 4.2270 | 4.265 | 4.2639 |
| | Width [mm] | *zy* | 0.8914 | 1.0707 | 0.8932 | 1.0730 |
| | | *xy* | 0.8914 | 1.0712 | 0.8927 | 1.0724 |
| | | *zx* | 3.8963 | 3.8955 | 3.5872 | 3.5862 |
| -6 dB | FLHM [mm] | *zy* | 6.0302 | 6.0302 | 6.0302 | 5.9296 |
| | | *xy* | 5.4097 | 5.4156 | 5.0728 | 5.0742 |
| | | *zx* | 5.9712 | 5.9704 | 6.0343 | 6.0342 |
| | FWHM [mm] | *zy* | 1.2121 | 1.2121 | 1.2121 | 1.2121 |
| | | *xy* | 1.2270 | 1.4011 | 1.2140 | 1.3933 |
| | | *zx* | 5.6680 | 5.6689 | 5.1098 | 5.1093 |

Table 4 shows the calculated areas within the contours at -3 dB and -6 dB pressure decays for all four models. The areas from cases A and C have in all three -3 dB contours and *xy*- plane in -6 dB contour similar values. This can be due to the geometry type used to model the radiation pattern, *i.e.*, the model includes the entire radiating surface. However, -6 dB contours areas in *zy*- and *zx*- planes present differences of 1.62 mm$^2$ and 2.272 mm$^2$, respectively, due to the isolated contours presence in the *zy*- plane and the wider extension of the focus in *zx*- plane (see Figures 3D-F). As well as before, areas from cases B and D (half of the geometry) show differences of 1.259 mm$^2$ and 2.095 mm$^2$ in the -6 dB contours at *zy*- and *zx*- planes, respectively. In both geometries, entire and half, the larger areas and focus distortions were found at – 6 dB contours FEM solutions in the *zy*- and *zx*- planes.

TABLE 4. **Areas within contours at -3 dB and -6 dB decays of $P_{max.}$**

| Areas within contours | Plane | Case A [mm$^2$] | Case B [mm$^2$] | Case C [mm$^2$] | Case D [mm$^2$] |
|---|---|---|---|---|---|
| -3 dB | *zy* | 3.175 | 3.856 | 3.197 | 3.896 |
| | *xy* | 2.543 | 3.134 | 2.553 | 3.147 |
| | *zx* | 12.837 | 12.844 | 12.086 | 12.327 |
| -6 dB | *zy* | 11.031 | 11.420 | 9.411 | 10.161 |
| | *xy* | 5.089 | 5.988 | 4.932 | 5.794 |
| | *zx* | 26.640 | 26.650 | 24.368 | 24.555 |

Table 5 shows the ellipsoidal shape ratio $E_r$[32] obtained in the four cases. In all cases, the shape ratio calculated is higher than 92 %, except in -6 dB contours of the *zy*- planes due to the presence of the isolated contours in cases A and B, and the enlarged area located between 11-13 mm in *z*- axis (see Figure 3D). This behavior can be avoided by modifying the $S_R$ geometrical characteristics. Even though the typical focal region shape is ellipsoidal[33], there are ultrasonic transducers that could produce different focal shapes such as conical[36], annular[37] or even two simultaneous foci[38]. In addition, it is important to remark that all four cases have a similar percentage of shape ratio in their corresponding planes which approximates to the typical focal shape[33] that was taken as reference.

TABLE 5. **Elliptical shape ratio percentage.**

| Contour level | Plane | Case A [%] | Case B [%] | Case C [%] | Case D [%] |
|---|---|---|---|---|---|
| -3 dB | *zy* | 93.76 | 92.10 | 94.10 | 92.36 |
| | *xy* | 98.16 | 95.90 | 98.16 | 95.67 |
| | *zx* | 99.33 | 99.34 | 99.34 | 99.34 |
| -6 dB | *zy* | 52.03 | 49.42 | 61.00 | 56.62 |
| | *xy* | 97.63 | 99.50 | 98.05 | 95.81 |
| | *zx* | 99.74 | 99.75 | 99.31 | 99.32 |

Table 6 shows the ellipsoidal volume approximations estimated at -3 dB and -6 dB pressure decay levels. For the calculation, half-length of the contour in the *zy*- plane and both half-length and half-width of the contours in the *xy*- plane were considered as the ellipsoid semi-axis. By making this assumption, the enlarged regions in the left side of Figure 3D are ignored; therefore, further analysis in the contribution of these areas should be performed. In the -3 dB contour level, the volume difference between cases A and C was 0.0738 mm$^3$, while between cases B and D was 0.0817 mm$^3$. In the -6 dB contour pressure decay, the volume difference was 1.514 mm$^3$ and 1.2354 mm$^3$ between cases A and C, and B and D, respectively.

TABLE 6. **Ellipsoidal volume approximation at -3 dB and -6 dB decays of $P_{max}$.**

| Contour level | Case A [mm$^3$] | Case B [mm$^3$] | Case C [mm$^3$] | Case D [mm$^3$] |
|---|---|---|---|---|
| -3 dB | 7.0951 | 8.4766 | 7.1689 | 8.5583 |
| -6 dB | 20.9575 | 23.5579 | 19.4435 | 22.3225 |

The larger differences between FEM and BEM areas and volumes, *e.g.* at -6 dB pressure decay contours, may be related to spurious reflections produced by the subdomain cylindrical boundaries and the impedance condition used to limit the interior problem. While BEM solutions account for no-reflections as they were modeled in an infinite or unbounded space. Nevertheless, these differences are reasonable small and validate the use of BEM to model the acoustic pattern of a FUS biomedical applicator prior to solving the interior problem. Additionally, the analyzed data suggest that FEM model needs improving, such as modifying the propagation medium shape and dimensions, and the inclusion of perfectly matched layers (PML). PMLs simulate an infinite domain in the exterior boundaries of the domain and reduces unwanted waves reflections[21][24].

## CONCLUSIONS

Four acoustic propagation models based on the same radiating surface geometry are presented in this work. The models consisted of an entire or a symmetric radiating surface propagating in a lossless medium. The results were obtained by means of FEM and BEM, for the interior $\Omega_i$ and exterior $\Omega_e$ problems solution, respectively. The analyzed data by using the four models solved with a λ/10 mesh show that the average $\Delta P_{max}$ and the average $\Delta z$ reached at the focal spot are 39.875 Pa and 0.4515 mm, respectively. Low values of errors were computed for $P_{max}$ estimated with BEM cases, and taking as references the FEM cases: 1.48 % for entire geometries and 1.37 % for symmetric ones. Isolated contours related to undesirable hot spots that were identified with the FEM models were not estimated using BEM models. This could be due to the computation method of BEM, which has reached better results on the named "exterior problem" than on the "interior one". A future work could be oriented to assess and solve these "erroneous estimations" of BEM.

As for dimensional assessment, almost all of the length and width ratios (BEM/FEM) approximate unity with an error minor to 1.7%, except in *zx*-plane width at -3 dB and -6 dB, and at *xy*-plane length at -6 dB, presenting errors up to 9.87 %. The maximum areas differences between the four cases in each plane at both contours pressure decays vary from 0.6 $mm^2$ to 2.3 $mm^2$; which addresses that symmetric geometries could be used instead of the entire geometries. The $E_r$ was over 92 % in most of the planes, except at -6 dB contours of *zy*- plane, where the average Er was 54.76 %; still, the four models presented a similar beam pattern formation. These results indicate that BEM can be used to solve the acoustic propagation problem in lossless media in a faster and with less computational resources fashion than FEM. In this sense, BEM represents a powerful tool during the design of the beam pattern of ultrasonic radiators intended for thermal ablation therapies. Nevertheless, it is important to mention that BEM presents limitations when solving the interior problem, and FEM continues to be the best method to solve acoustic wave interactions within media. Further work includes geometrical radiator modification to attain deeper penetration distances in a lossless medium, domain sectioning in subdomains and combining meshes size to solve the interior problem, modeling both ultrasonic propagation in biological tissue emulators and tissue heating due to FUS.

## ACKNOWLEDGEMENT

R. Martinez-Valdez thanks the Mexican National Council of Humanities, Sciences and Technologies (CONAHCyT) for both Postdoctoral scholarship and SNII distinction. The authors thank the Autonomous University of Aguascalientes for the opportunity to carry out Dr. R. Martínez-Valdez postdoctoral stay.

## DECLARATION OF INTEREST STATEMENT

The authors report no declaration of interest.

## AUTHOR CONTRIBUTIONS

R. M. V. conceptualization, data curation, formal analysis, investigation, methodology, resources, validation, visualization, writing - original draft, writing - review, and editing; I. B. project administration, resources, supervision, validation, writing - review, and editing.

## REFERENCES

**[1]** C. R. Hill, J. C. Bamber, and G. R. ter Haar, Physical Principles of Medical Ultrasonics, 2nd ed. England: John Wiley & Sons, 2004, doi: https://doi.org/10.1002/0470093978.fmatter

**[2]** D. A. Hutchins and G. Hayward, "The radiated field of ultrasonic transducers," in Physical Acoustics, vol. 19, R. N. Thurston and A. D. Pierce, Eds., United States of America:Academic Press, 1990, pp. 1-80, doi: https://doi.org/10.1016/B978-0-12-477919-8.50007-8

**[3]** C. J. Diederich and K. Hynynen, "Ultrasound technology for hyperthermia," Ultrasound Med. Biol., vol. 25, no. 6, pp. 871-887, 1999, doi: https://doi.org/10.1016/s0301-5629(99)00048-4

**[4]** G. T. Haar and C. Coussios, "High intensity focused ultrasound: Physical principles and devices," Int. J. Hyperth., vol. 23, no. 2, pp. 89-104, 2007, doi: https://doi.org/10.1080/02656730601186138

**[5]** M. Lafond, A. Payne, and C. Lafon, "Therapeutic ultrasound transducer technology and monitoring techniques: a review with clinical examples," Int. J. Hyperth., vol. 41, no. 1, 2024, art. no. 2389288, doi: https://doi.org/10.1080/02656736.2024.2389288

**[6]** C. J. Trujillo Romero and D.-L. Flores, Diagnosis and Treatment of Cancer using Thermal Therapies, 1st ed. Boca Raton, FL, United States of America: CRC Press, 2023, doi: https://doi.org/10.1201/9781003342663

**[7]** E. A. Filonenko and V. A. Khokhlova, "Effect of acoustic nonlinearity on heating of biological tissue by high-intensity focused ultrasound," Acoust. Phys., vol. 47, no. 4, pp. 468-475, 2001, doi: https://doi.org/10.1134/1.1385422

**[8]** J. Gao, S. Cochran, and Z. Huang, "Ultrasound beam distortion and pressure reduction in transcostal focused ultrasound surgery," Appl. Acoust., vol. 76, pp. 337-345, 2014, doi: https://doi.org/10.1016/j.apacoust.2013.06.003

**[9]** S. Paliwal and S. Mitragotri, "Therapeutic opportunities in biological responses of ultrasound," Ultrasonics, vol. 48, no. 4, pp. 271-278, 2008, doi: https://doi.org/10.1016/j.ultras.2008.02.002

**[10]** I. Rivens, C. Jayadewa, P. Mouratidis, and G. Ter Haar, "Histological characterization of HIFU lesions," Int. J. Hyperth., vol. 41, no. 1, 2024, art. no. 2389292, doi: https://doi.org/10.1080/02656736.2024.2389292

**[11]** C. Lafon, D. Melodelima, R. Salomir, and J. Y. Chapelon, "Interstitial devices for minimally invasive thermal ablation by high-intensity ultrasound," Int. J. Hyperthremia, vol. 23, no. 2, pp. 153-163, 2007, doi: https://doi.org/10.1080/02656730601173029

**[12]** M. Dahan, M. Cortet, C. Lafon, and F. Padilla, "Combination of Focused Ultrasound, Immunotherapy, and Chemotherapy: New Perspectives in Breast Cancer Therapy," J. Ultrasound Med., vol. 42, no. 3, pp. 559-573, 2023, doi: https://doi.org/10.1002/jum.16053

**[13]** J. H. Hwang and L. A. Crum, "Current status of clinical high-intensity focused ultrasound," in 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Minneapolis, MN, USA, 2009, pp. 130-133. doi: https://doi.org/10.1109/iembs.2009.5335244

**[14]** O. Ginsburg, C.-H. Yip, A. Brooks, A. Cabanes, et al., "Breast Cancer Early Detection: A Phased Approach to Implementation," Cancer, vol. 126, no. S10, pp. 2379-2393, 2020, doi: https://doi.org/10.1002/cncr.32887

**[15]** B. S. Bachu, J. Kedda, I. Suk, J. J. Green, and B. Tyler, "High-Intensity Focused Ultrasound: A Review of Mechanisms and Clinical Applications," Ann. Biomed. Eng., vol. 49, pp. 1975-1991, 2021, doi: https://doi.org/10.1007/s10439-021-02833-9

**[16]** Y.-H. Hsiao, S.-J. Kuo, H.-D. Tsai, M.-C. Chou, and G.-P. Yeh, "Clinical Application of High-intensity Focused Ultrasound in Cancer Therapy," J. Cancer, vol. 7, no. 3, pp. 225-231, 2016, doi: https://doi.org/10.7150/jca.13906

**[17]** E. Van 't Wout, P. Gélat, T. Betcke, S. Arridge, and P. G. Elat, "A fast boundary element method for the scattering analysis of high-intensity focused ultrasound," J. Acoust. Soc. Am, vol. 138, pp. 2726-2737, 2015, doi: https://doi.org/10.1121/1.4932166

**[18]** F. Shen, F. Fan, F. Li, L. Wang, et al., "An efficient method for transcranial ultrasound focus correction based on the coupling of boundary integrals and finite elements," Ultrasonics, vol. 137, 2024, art. no. 107181, doi: https://doi.org/10.1016/j.ultras.2023.107181

**[19]** P. Burrascano, S. Callegari, A. Montisci, M. Ricci, and M. Versaci, Eds., Ultrasonic Nondestructive Evaluation Systems. Switzerland: Springer, 2015, doi: https://doi.org/10.1007/978-3-319-10566-6

**[20]** S. Marburg, "SIX BOUNDARY ELEMENTS PER WAVELENGTH: IS THAT ENOUGH?," J. Comput. Acoust., vol. 10, no. 01, pp. 25-51, 2002, doi: https://doi.org/10.1142/S0218396X02001401

**[21]** L. L. Thompson, "A review of finite-element methods for time-harmonic acoustics," J. Acoust. Soc. Am., vol. 119, no. 3, pp. 1315-1330, 2006, doi: https://doi.org/10.1121/1.2164987

**[22]** S. R. Haqshenas, P. Gélat, E. van 't Wout, T. Betcke, and N. Saffari, "A fast full-wave solver for calculating ultrasound propagation in the body," Ultrasonics, vol. 110, 2021, art. no. 106240, doi: https://doi.org/10.1016/j.ultras.2020.106240

**[23]** S. Kirkup, "The Boundary Element Method in Acoustics: A Survey," Appl. Sci., vol. 9, no. 8, 2019, art. no. 1642, doi: https://doi.org/10.3390/app9081642

**[24]** S. Preuss, C. Gurbuz, C. Jelich, S. K. Baydoun, and S. Marburg, "Recent Advances in Acoustic Boundary Element Methods," J. Theor. Comput. Acoust., vol. 30, no. 03, 2022, art. no. 2240002, doi: https://doi.org/10.1142/S2591728522400023

**[25]** Y. Liu, “On the BEM for acoustic wave problems,” Eng. Anal. Bound. Elem., vol. 107, pp. 53-62, 2019, doi: https://doi.org/10.1016/j.enganabound.2019.07.002

**[26]** S. Marburg, “Boundary Element Method for Time-Harmonic Acoustic Problems,” in Computational Acoustics. CISM International Centre for Mechanical Sciences, M. Kaltenbacher, Ed., 2018, Switzerland: Springer, pp. 69-158, doi: https://doi.org/10.1007/978-3-319-59038-7_3

**[27]** I. Harari and T. J. R. Hughes, “A cost comparison of boundary element and finite element methods for problems of time-harmonic acoustics,” Comput. Methods Appl. Mech. Eng., vol. 97, no. 1, pp. 77-102, 1992, doi: https://doi.org/10.1016/0045-7825(92)90108-V

**[28]** J. Assaad, A.-C. Hladky, and B. Cugnet, “Application of the FEM and the BEM to compute the field of a transducer mounted in a rigid baffle (3D case),” Ultrasonics, vol. 42, no. 1-9, pp. 443-446, 2004, doi: https://doi.org/10.1016/j.ultras.2003.12.032

**[29]** A. G. Santiago, L. C. Trintinalia, and M. A. Gutierrez, “Boundary element method applied to ultrasound elastography,” Eng. Anal. Bound. Elem., vol. 62, pp. 154-162, 2016, doi: https://doi.org/10.1016/j.enganabound.2015.10.005

**[30]** P. Gélat, G. ter Haar, and N. Saffari, “Modelling of the acoustic field of a multi-element HIFU array scattered by human ribs,” Phys. Med. Biol., vol. 56, no. 17, pp. 5553-5581, 2011, doi: https://doi.org/10.1088/0031-9155/56/17/007

**[31]** P. Gélat, G. ter Haar, and N. Saffari, “Towards the optimisation of acoustic fields for ablative therapies of tumours in the upper abdomen,” J. Phys. Conf. Ser., vol. 457, 2013, art. no. 012002, doi: https://doi.org/10.1088/1742-6596/457/1/012002

**[32]** R. Martínez-Valdez, A. Ramos Fernández, A. Vera Hernandez, and L. Leija Salas, “Design of a low power hybrid HIFU applicator for haemostasis based on acoustic propagation modelling,” Int. J. Hyperth., vol. 32, no. 2, pp. 121-131, 2016, doi: https://doi.org/10.3109/02656736.2015.1112437

**[33]** G. Ter Haar and C. Coussios, “High intensity focused ultrasound: past, present and future.,” Int. J. Hyperthermia, vol. 23, no. 2, pp. 85-7, 2007, doi: https://doi.org/10.1080/02656730601185924

**[34]** H. P. Kok, E. N. K. Cressman, W. Ceelen, C. L. Brace, et al., “Heating technology for malignant tumors: a review,” Int. J. Hyperth., vol. 37, no. 1, pp. 711-741, 2020, doi: https://doi.org/10.1080/02656736.2020.1779357

**[35]** C. R. Hill and G. R. ter Haar, “High intensity focused ultrasound—potential for cancer treatment,” Br. J. Radiol., vol. 68, no. 816, pp. 1296-1303, 1995, doi: https://doi.org/10.1259/0007-1285-68-816-1296

**[36]** D. Melodelima, W. A. N’Djin, H. Parmentier, S. Chesnais, M. Rivoire, and J.-Y. Chapelon, “Thermal Ablation by High-Intensity-Focused Ultrasound Using a Toroid Transducer Increases the Coagulated Volume. Results of Animal Experiments,” Ultrasound Med. Biol., vol. 35, no. 3, pp. 425-435, 2009, doi: https://doi.org/10.1016/j.ultrasmedbio.2008.09.020

**[37]** C. A. Cain and S. Umemura, “Concentric-Ring and Sector-Vortex Phased-Array Applicators for Ultrasound Hyperthermia,” IEEE Trans. Microw. Theory Tech., vol. 34, no. 5, pp. 542-551, 1986, doi: https://doi.org/10.1109/TMTT.1986.1133390

**[38]** S. Umemura and C. A. Cain, “The sector-vortex phased array: acoustic field synthesis for hyperthermia,” IEEE Trans. Ultrason. Ferroelectr. Freq. Control, vol. 36, no. 2, pp. 249-257, 1989, doi: https://doi.org/10.1109/58.19158

**ARTÍCULO DE INVESTIGACIÓN**

VOL. 46 | NO. 1 |**ENERO-ABRIL 2025** | PP 1-18

Revista Mexicana de Ingeniería Biomédica

https://dx.doi.org/10.17488/RMIB.46.1.2

E-LOCATION ID: e1478

# Evaluación de Tubos de Recolección de Muestras de Sangre Utilizando Deep Learning

## Evaluation of Blood Sample Collection Tubes Using Deep Learning

*Ignacio Franco-Alucano*[1], *Julian Aguilar-Duque*[1], *Yolanda Baez-Lopez*[1], *Jorge Limon-Romero* [1],
*María Marcela Solís-Quinteros*[2], *Diego Tlapa*[1] ✉

[1] Universidad Autónoma de Baja California, Facultad de Ingeniería, Arquitectura y Diseño, Baja California - México

[2] Universidad Autónoma de Baja California, Facultad de Contaduría y Administración, Baja California - México

### RESUMEN

La flebotomía es un procedimiento para obtener muestras de sangre principalmente para análisis clínicos en laboratorios. La cantidad de sangre, identificación de tubos y el uso del tubo adecuado son características que el profesional de la salud inspecciona visualmente. Al ser una actividad manual, la posibilidad de error está presente pudiendo tener efectos tanto en la calidad, como en el flujo de trabajo y eficiencia. A pesar del avance de las tecnologías de la industria 4.0, incluida la inteligencia artificial (IA), hay poca evidencia de aplicaciones en laboratorios clínicos. Este estudio tiene como objetivo evaluar la idoneidad de utilizar el aprendizaje profundo o *deep learning* (DL) en la inspección de tubos con muestras de sangre. Particularmente se prueban tres arquitecturas YOLOv5, YOLOv7 y YOLOv8 en la detección de seis clases incluyendo color de tapa y presencia de etiqueta. El mayor desempeño de precisión se presentó con el modelo YOLOv8 obteniendo una precisión de 0.927 en la detección, lo que evidencia una alta capacidad para inspeccionar características importantes en el servicio de flebotomía, siendo DL una alternativa viable para asistir a los profesionales de la salud en actividades de inspección. Trabajo futuro incluye ampliar el número de imágenes de manera balanceada.

**PALABRAS CLAVE:** deep learning, detección de objetos, laboratorio clínico, muestra de sangre, red neuronal convolucional, YOLO

## ABSTRACT

Phlebotomy is a procedure to obtain blood samples, mainly for laboratory clinical analysis. The amount of blood, tube identification, and the use of the appropriate tube are characteristics that the health professional visually inspects. Being a manual activity, the possibility of error is latent and can affect quality, workflow, and efficiency. Despite the advancement of industry 4.0 technologies, including artificial intelligence (AI), there is little evidence of applications in clinical laboratories. This study aims to evaluate the suitability of using deep learning (DL) in inspecting tubes with blood samples. Specifically, three architectures, YOLOv5, YOLOv7, and YOLOv8, are tested to detect six classes, including cap color and the presence of labels. The highest precision performance was presented by the YOLOv8 model, obtaining a precision of 0.927 in detection, which shows a high capacity to inspect important characteristics in the phlebotomy service. Therefore, being DL is a suitable alternative to assist health professionals in inspection activities. Future work includes expanding the number of images in a balanced manner.

**KEYWORDS:** deep learning, object detection, clinical laboratory, blood samples, convolutional neural networks, YOLO

### Autor de correspondencia

DESTINATARIO: **Diego Tlapa**
INSTITUCIÓN:**UNIVERSIDAD AUTÓNOMA DE BAJA CALIFORNIA**
DOMICILIO: **CARRETERA TRANSPENINSULAR ENSENADA-TIJUANA NÚMERO 3917, COLONIA PLAYITAS. ENSENADA, B.C., C.P. 22860, MÉXICO**
CORREO ELECTRÓNICO: diegotlapa@uabc.edu.mx

**Recibido:**

15 Octubre 2024

**Aceptado:**

09 Diciembre 2024

**Publicado:**

31 Enero 2025

## INTRODUCCIÓN

En el campo de la atención médica, los servicios de análisis clínicos son fundamentales para el diagnóstico y tratamiento de pacientes. Al respecto, la flebotomía es el procedimiento médico que consiste en extraer sangre de un paciente para realizar pruebas de laboratorio, transfusiones de sangre o donaciones. La evaluación de una muestra de sangre de un paciente es un paso crítico en la atención médica[1]. Los laboratorios clínicos se rigen por diversos lineamientos sanitarios y estándares, de los cuales, la norma ISO 15189:2022 establece los requisitos de un sistema de gestión de la calidad en laboratorios clínicos, incluyendo lineamientos para procesos pre-examen, durante el examen y post-examen[2]. Algunos lineamientos incluyen el volumen de las muestras, tipo de tubo y sustancia preservativa a utilizar para asegurar que no existe faltante ni excedente de muestras colectadas[2], ya que cualquier desviación es crítica y debe ser corregida[3]. Además, el estándar requiere lineamientos para la identificación correcta de muestras para garantizar la seguridad del paciente, optimizar los procesos de laboratorio y entregar resultados de pruebas para el posterior diagnóstico y seguimiento de enfermedades.

Dentro de las tres fases del periodo analítico (preanalítica, analítica y postanalítica), es en las fases preanalítica y postanalítica donde son más propensas a errores que la fase analítica[4][5][6]. En la fase preanalítica destacan errores como solicitudes incorrectas, mala identificación del paciente o muestras, recolección de muestras inadecuada (volumen insuficiente), uso de contenedores incorrectos y fallos en el manejo, almacenamiento o transporte de la muestra[4]. En cuanto la etapa postanalítica, los errores incluyen validación incorrecta de datos, retrasos en la entrega de resultados críticos, errores en la entrada de datos y tiempos de respuesta excesivos[7][8]. Estos errores potencialmente llevan al rechazo de muestras por parte del laboratorio. En la fase preanalítica, enfermería desempeña un papel crítico, al ser la responsable de la obtención de las muestras de sangre, mientras que en la fase postanalítica es el personal administrativo quien es el responsable de la recopilación y envío de resultados.

En las unidades médicas, el área de flebotomía es uno de los lugares con mayor cantidad de pacientes en un hospital[9], lo que implica grandes flujos de pacientes que pueden experimentar demoras por existencia de cuellos de botella. Al ser el servicio de obtención de muestras un proceso manual, éste conlleva la inspección visual de numerosos recipientes de muestras de sangre y es susceptible a errores humanos, lo que podría dar lugar a diagnósticos erróneos, retrasos en el tratamiento o atención deficiente al paciente. En este contexto, existe una creciente necesidad de sistemas apoyados en tecnologías que puedan ayudar a inspeccionar recipientes con muestras de sangre en laboratorios clínicos, incluyendo la reciente tendencia a la integración de la inteligencia artificial (IA) en el área de hematología[10] y química sanguínea mediante la automatización del análisis, la interpretación de datos, la segmentación precisa de imágenes para aislar células sanguíneas y el análisis de datos clínicos a nivel molecular y del paciente[11]. Sin embargo, persisten desafíos como la generalización de modelos, el desempeño en conjuntos de datos externos, la adopción de arquitecturas más actuales[11], la falta de interpretación, el volumen de datos, así como la confidencialidad de la información[12]. Existen además oportunidades de mejora en la inspección de muestras de sangre para la identificación de errores cuando la muestra de sangre recolectada de un paciente, se etiqueta incorrectamente con información de otro paciente, pudiendo ser difícil para el laboratorio reconocer este tipo de error[13].

En este sentido, mejorar la calidad y la eficiencia en los servicios de salud implica un reto recurrente al que se enfrenta el sector salud[14], por lo que en años recientes se ha incrementado la digitalización de los servicios de salud, conocida como *Healthcare* 4.0[15], que involucra la implementación de variadas tecnologías y con fines distintos incluyendo *deep learning* (DL)[16], *machine learning* (ML)[17], *big data*[18], automatización[19], simulación[20], entre otras. Particularmente, herramientas de IA desempeñan un papel crucial en el campo médico, transformando la manera en que se diagnostican enfermedades, se realizan tratamientos y se gestionan los datos clínicos. Estas tecnologías han demostrado un potencial significativo para mejorar la precisión, la eficiencia y la atención médica en general. Algunas de las actividades clave para el sector médico donde se emplea IA son el análisis de grandes volúmenes de datos y el reconocimiento de patrones u objetos en imágenes médicas, como rayos X[21], tomografías computarizadas[22], resonancias magnéticas[23] y ultrasonidos[24]. DL es una subrama de la IA, la cual ha presenciado avances notables en los últimos años, ofreciendo oportunidades para el análisis de imágenes y las tareas de detección de objetos. Los modelos de DL se componen por múltiples capas que permiten aprender una secuencia en sus distintos niveles[25]. Al respecto, las redes neuronales convolucionales (CNN) procesan imágenes y otras formas de datos con estructura espacial, donde la clasificación se basa en regiones creando un cuadro delimitador alrededor de la región de interés, como la densidad, el color, etc. Los métodos estándares involucran algoritmos de una etapa[26] y dos etapas [27], siendo YOLO *(You Only Look Once)*[28] una familia de algoritmos de una sola etapa que ha mostrado ventajas incluyendo menor tiempo de entrenamiento[29] así como su eficiencia y precisión en la localización y clasificación simultánea de múltiples objetos en tiempo real en imágenes[28].

Este estudio prueba la factibilidad de los algoritmos YOLO para la detección de características relevantes de tubos con muestras de sangre en un laboratorio clínico, centrándose en tres tareas principales: (1) verificar la presencia de identificación en cada recipiente, (2) verificar la presencia de sangre contenida en el recipiente y (3) detectar el adecuado recipiente para la prueba a realizar en función del color de la tapa. Al abordar estas tareas, se busca contribuir a mejorar la eficiencia y confiabilidad del análisis de muestras de sangre, lo que en última instancia mejora la atención al paciente y los flujos de trabajo de laboratorio.

## MATERIALES Y MÉTODOS

Con la finalidad de una evaluación integral de los modelos YOLO, se consideró las arquitecturas YOLOv5, YOLOv7 y YOLOv8 debido a características específicas que aportan cada uno en términos de precisión, velocidad y optimización respecto a sus predecesores. YOLOv5 es reconocido por su balance entre eficiencia y facilidad de implementación, facilitando así la inspección en tiempo real. YOLOv7 incluye mejoras en la velocidad y precisión en diferentes tipos de *hardware*. YOLOv8 incluye avances de arquitectura de red y procesamiento, resultando en mayor adaptabilidad y desempeño en escenarios complejos. Estas diferencias permitieron evaluar cada versión en la identificación precisa y rápida de seis clases de objetos, incluyendo: (i) etiqueta presente en tubos *(label)*, (ii) tubo con llenado insuficiente *(not_ok)*, (iii) tubo con llenado ok *(good)*, (iv) tubo con tapa azul *(blue_cap)* que contienen citrato de sodio y se utilizan para pruebas de coagulación, (v) tubo con tapa morada *(purple_cap)* utilizadas en pruebas de hematología y (vi) tubo con tapa amarilla *(yellow_cap)* con anticoagulante para pruebas de química sanguínea. Se utilizó un conjunto de datos compuesto por 3000 imágenes, desglosadas en 2104 imágenes (70 %) para entrenamiento, 604 imágenes para validación (20 %) y 292 imágenes (10 %) para prueba y se realizó las anotaciones buscando un balance en las clases como se describe en la Tabla 1.

**TABLA 1. Anotaciones por clase.**

| Descripción | Clases | Anotaciones |
|---|---|---|
| Detección de etiqueta presente en | label | 2,378 |
| Tubo con llenado insuficiente | not_ok | 1,918 |
| Tubo con llenado ok | good | 2,580 |
| Tubo con tapa azul | blue_cap | 2,047 |
| Tubo con tapa morada | purple_cap | 1,998 |
| Tubo con tapa amarilla | yellow_cap | 1,725 |

El entrenamiento de los datos empleó el acelerador Tesla T4 incluido en la versión gratuita de Google Colab. Este acelerador, es una unidad de procesamiento gráfico (GPU) diseñada para tareas de alto rendimiento y cuenta con 40 núcleos, una frecuencia de 1.59 GHz y una capacidad de procesamiento de FP16 de 65 TFLOPS, lo cual, lo vuelve una herramienta versátil, potente y accesible. Por su parte, el etiquetado de las imágenes se realizó mediante la segmentación, empleando recuadros delimitadores *"bounding boxes"* en la herramienta *Roboflow,* la cual es una plataforma en línea para la clasificación de imágenes y la gestión de datos[30]. Las imágenes utilizadas presentan una resolución de 640x640 pixeles en formato JPG y se cuentan con acceso libre al conjunto de datos en la siguiente liga: https://doi.org/10.34740/kaggle/dsv/9616793. Los resultados se evaluaron como verdadero positivo (TP, *true positive*), es decir, el número de objetos detectados correctamente; verdadero negativo (TN, *true negative*), es decir, muestras que se rechazan correctamente de la clase; falso positivo (FP, *false positive*), es decir, el número de objetos detectados erróneamente; y falso negativo (FN, *false negative*), es decir, el número de objetos omitidos. Para evaluar las diferentes arquitecturas se utilizaron índices de desempeño como precisión *(precision)*, sensibilidad *(recall)* y precisión media promedio (*m*AP). Precisión, precisión por clase (PRC)[31] o valor predictivo positivo (VPP)[32], mide la proporción de objetos correctamente detectados en relación con los objetos detectados[33]. La ecuación 1 muestra la fórmula utilizada en este estudio.

$$Precision = \frac{TP}{TP + FP} \tag{1}$$

*Recall,* indica la proporción de objetos detectados correctamente en relación con todos los objetos presentes en el conjunto de datos[34], por lo tanto, *recall* es la precisión de las instancias predichas positivamente que describen cuántas se etiquetaron correctamente[35], lo que muestra el nivel de un modelo para predecir la clase positiva cuando la clase real es positiva[36]. En este estudio calculamos *recall* con la ecuación (2).

$$Recall = \frac{TP}{TP + FN} \tag{2}$$

Precisión media promedio (*m*AP) mide en qué porcentaje el algoritmo predice el objeto correctamente de todas las clases individuales[37], siendo una métrica clave para evaluar los algoritmos de detección de objetos[38] y para la comparación entre modelos[39]. La ecuación 3 muestra la descripción matemática de *m*AP, donde APk representa la precisión promedio de las clases y n representa el número de clases[40].

$$mAP = \frac{1}{n}\sum_{k=1}^{k=n} AP_k \tag{3}$$

En conjunto, estos métricos proporcionan una evaluación global del rendimiento del modelo y son ampliamente utilizados en la evaluación de algoritmos de detección de objetos ya que permiten comparar el desempeño del modelo en diferentes aspectos de la detección[41][42]. El estudio también se enfoca en analizar el comportamiento de los modelos en distintas épocas de entrenamiento. Se seleccionaron las épocas 20, 50 y 100 como puntos de referencia para examinar la evolución de los modelos durante el proceso de entrenamiento y evaluar su capacidad de aprendizaje a lo largo del tiempo. Para complementar el resultado, se analizaron las matrices de confusión normalizadas y las gráficas de pérdidas para evaluar y visualizar el rendimiento de un modelo.

### Equipo y Herramientas

El análisis se llevó a cabo en un equipo con un procesador AMD Ryzen 5 4650G con Radeon R7 Renoir, funcionando a una velocidad de 3.70 GHz, con 32 GB de RAM y sistema operativo Windows 11 de 64 bits. Además, se empleó una tarjeta gráfica NVIDIA GeForce GTX 1660 SUPER VENTUS XS OC con 6 GB de memoria dedicada para el procesamiento computacional. Para la captura de imágenes se utilizaron dos dispositivos: una cámara marca Salandens, modelo B0872YBHBV, con una resolución de 1080p, y la cámara de un teléfono móvil Huawei Nova Y9 SE, con una resolución de 108 MP. El entrenamiento de los modelos se realizó utilizando Google Colab en su versión gratuita, la cual ofrece acceso a una GPU T4 con 12.7 GB de RAM, 15 GB de memoria dedicada para la GPU, y un espacio en disco de 112.6 GB.

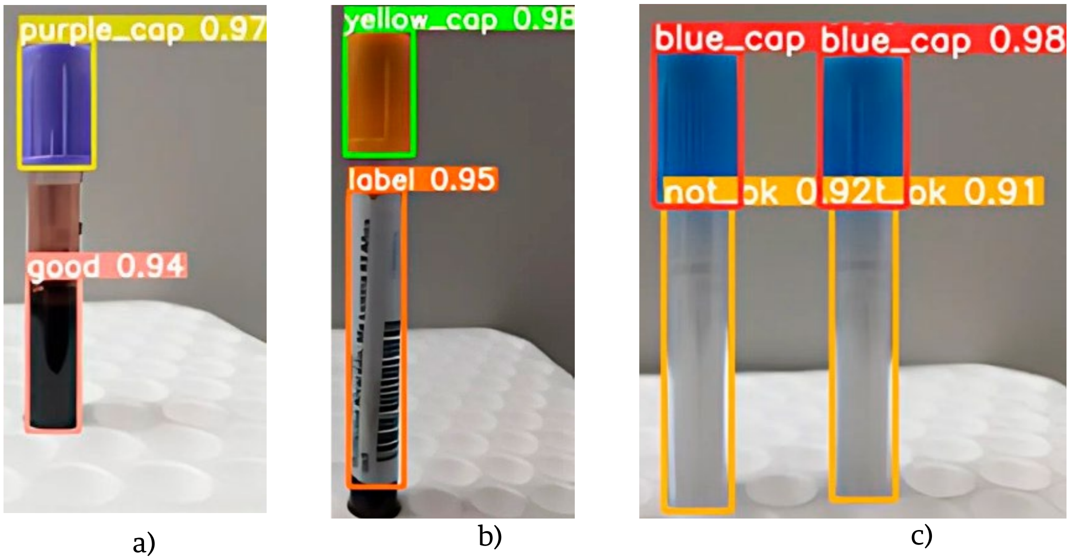
## RESULTADOS Y DISCUSIÓN

El desempeño de los diferentes modelos de YOLO se muestra en la Tabla 2, siendo YOLOv8 la arquitectura que presentó la mayor precisión, 0.927, con 50 épocas mientras que YOLOv5, 0.924, con 100 épocas y YOLOv7, 0.917 con 100 épocas. Respecto a *recall,* YOLOv8 presentó los mayores valores, destacando 0.914 con 100 épocas. Al evaluar la capacidad del modelo para detectar imágenes de distintas clases, YOLOv8 presentó un valor de *m*AP de 0.941 y 0.940 con 100 y 50 épocas respectivamente. En términos de tiempo de entrenamiento, el modelo más rápido fue YOLOv5 con 20 épocas (15 min, 4 s), en contraparte, YOLOv7 con 100 épocas presentó la mayor duración (3 h, 27 min, 44 s).

TABLA 2. **Resultados del rendimiento de los modelos de YOLOv5, YOLOv7 y YOLOv8.**

| Modelo | Épocas | Precisión | Sensibilidad | *m*AP | Tiempo |
|---|---|---|---|---|---|
| YOLOv5 | 20 | 0.909 | 0.897 | 0.919 | 15 min 4 s |
| | 50 | 0.923 | 0.900 | 0.934 | 37 min 45 s |
| | 100 | 0.924 | 0.900 | 0.935 | 1 h 14 min 10 s |
| YOLOv7 | 20 | 0.682 | 0.665 | 0.675 | 42 min 36 s |
| | 50 | 0.903 | 0.852 | 0.856 | 1 h 44 min 24 s |
| | 100 | 0.917 | 0.878 | 0.906 | 3 h 27 min 44 s |
| YOLOv8 | 20 | 0.922 | 0.907 | 0.933 | 34 min 19 s |
| | 50 | 0.927 | 0.909 | 0.940 | 1 h 26 min 13 s |
| | 100 | 0.917 | 0.914 | 0.941 | 2 h 59 min 56 s |

En términos generales, YOLOv8 presenta un mayor desempeño, logrando una mayor precisión y *m*AP en todas las épocas analizadas. Su tiempo de entrenamiento, aunque mayor que YOLOv5, es más corto que YOLOv7, lo cual lo convierte en una opción eficiente en términos de rendimiento y costo computacional. YOLOv5, aunque no alcanza los niveles de YOLOv8, ofrece un desempeño consistente con una buena precisión y *m*AP, y tiempos de entrenamiento significativamente menores, haciéndolo atractivo en escenarios donde los recursos de tiempo y procesamiento son limitados. YOLOv7 comienza con un rendimiento inferior incluyendo precisión y *m*AP bajos con 20 épocas, pero mejora significativamente con más épocas de entrenamiento, aunque su tiempo de entrenamiento es considerablemente mayor que los otros dos modelos, lo cual podría ser un factor limitante dependiendo de la aplicación. La Figura 1 muestra ejemplos de detección de clases empleando YOLOv8.



a) b) c)

FIGURA 1. **(A) Ejemplo de detección de las clases "purple_cap" y "good". (B) Demostración de la detección de las clases "yellow_cap" y "label". (C) Representación de las clases "blue_cap" y "not_ok" empleando YOLOv8.**

El desempeño de las tres arquitecturas por clase al utilizar 50 épocas se muestra en las Tablas 3-5. El resto de resultados con 20 y 100 épocas se agregó como material suplementario al artículo. Particularmente, la Tabla 3 muestra el desempeño de YOLOv5 con 50 épocas, destacando una precisión de hasta 0.980 para tapa morada *"purple_cap"* y una mínima de 0.876 para identificar tubos con tapa amarilla *"yellow_cap"*. *Recall* presenta valores de 0.978 hasta 0.85 para la detección de tubos con tapa morada *"purple_cap"* y tapa amarilla *"yellow_cap"* respectivamente. La clase con mayor *m*AP (0.99) es para tapa morada *"purple_cap"*, mientras que el menor valor fue 0.893 para la clase tapa amarilla *"yellow_cap"*.

TABLA 3. **Resultados del rendimiento del modelo de YOLOv5 con 3000 imágenes entrenado con 50 épocas**

| Clase | Precisión | *Recall* | *m*AP |
|---|---|---|---|
| blue_cap | 0.924 | 0.92 | 0.951 |
| good | 0.934 | 0.915 | 0.938 |
| label | 0.93 | 0.851 | 0.927 |
| not_ok | 0.896 | 0.886 | 0.903 |
| purple_cap | 0.98 | 0.978 | 0.99 |
| yellow_cap | 0.876 | 0.85 | 0.893 |

Por su parte, la Tabla 4 muestra el desempeño de YOLOv7 con 50 épocas, logrando una precisión de hasta 0.966 para tapa morada *"purple_cap"* y una mínima de 0.821 para identificar tubos con tapa amarilla *"yellow_cap"*. *Recall* presentó valores de 0.963 hasta 0.708 para la detección de tubos con tapa morada *"purple_cap"* y tubos con etiqueta "label" respectivamente. La clase con el mayor *m*AP de 0.978 fue tapa morada *"purple_cap"*, mientras que el menor valor fue 0.782 para la clase que reporta los tubos con tapa amarilla *"yellow_cap".*

**TABLA 4. Resultados del rendimiento del modelo de YOLOv7 con 3000 imágenes entrenado con 50 épocas**

| Clase | Precisión | *Recall* | *m*AP |
|---|---|---|---|
| blue_cap | 0.906 | 0.914 | 0.877 |
| good | 0.901 | 0.860 | 0.87 |
| label | 0.944 | 0.708 | 0.826 |
| not_ok | 0.88 | 0.817 | 0.805 |
| purple_cap | 0.966 | 0.963 | 0.978 |
| yellow_cap | 0.821 | 0.847 | 0.782 |

La Tabla 5 muestra los resultados de cada clase utilizando YOLOv8 con 50 épocas, donde se observa una precisión máxima de 0.984 y mínima de 0.883 para identificar los tubos de sangre con tapa morada *("purple_cap")* y los tubos vacíos o que no contienen sangre *("not_ok")* respectivamente. Respecto a *recall,* el valor mayor es de 0.976 para la detección de tubos con tapa morada *("purple_cap"),* mientras que 0.840 fue el valor menor para tubos con tapa amarilla *("yellow_cap").* La clase con mayor *m*AP de hasta 0.993 es la de tapa morada *("purple_cap"),* mientras que el menor valor fue 0.891 para la clase que reporta los tubos que no contienen sangre *("not_ok").*

**TABLA 5. Resultados del rendimiento del modelo de YOLOv8 con 3000 imágenes entrenado con 50 épocas.**

| Clase | Precisión | *Recall* | *m*AP |
|---|---|---|---|
| blue_cap | 0.923 | 0.914 | 0.955 |
| good | 0.933 | 0.921 | 0.942 |
| label | 0.944 | 0.901 | 0.952 |
| not_ok | 0.883 | 0.901 | 0.891 |
| purple_cap | 0.984 | 0.976 | 0.993 |
| yellow_cap | 0.895 | 0.840 | 0.908 |

La Figura 2 muestra la matriz de confusión normalizada para YOLOv8 con 50 épocas donde cada fila representa los porcentajes de instancias reales de cada clase, mientras que las columnas muestran las predicciones del modelo. La diagonal normalizada muestra valores altos en la tasa de acierto, incluyendo *"purple_cap"* con 0.98, mientras *"yellow_cap"* muestra un valor de 0.91, siendo el menor.

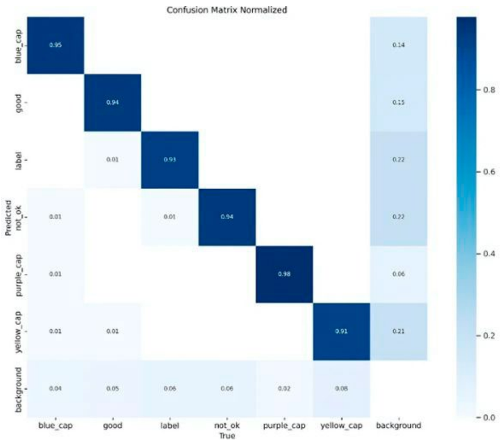


**FIGURA 2. Matriz de confusión normalizada de YOLOv8 con 50 épocas.**

Por su parte, la Figura 3 muestra la gráfica de perdida para YOLOv8 con 50 épocas, destacando cómo la función de pérdida del modelo disminuye a lo largo del tiempo durante el entrenamiento. Una disminución continua en la pérdida indica que el modelo está aprendiendo correctamente a medida que avanza el entrenamiento. La estabilización de la curva o una disminución más lenta después de muchas épocas sugiere que el modelo ha alcanzado su punto de convergencia y que entrenar por más tiempo no mejoraría significativamente su desempeño. Si bien es cierto que extender el tiempo de entrenamiento puede resultar en mejoras marginales en precisión, es necesario considerar el costo computacional y la aplicación práctica para sistemas que trabajan en tiempo real. En conjunto, estas figuras reflejan un buen desempeño de YOLOv8, con una matriz de confusión que revela una alta precisión en la clasificación de los defectos, y una gráfica de pérdida que sugiere que el modelo se ha aprendido eficientemente a lo largo del proceso de entrenamiento.

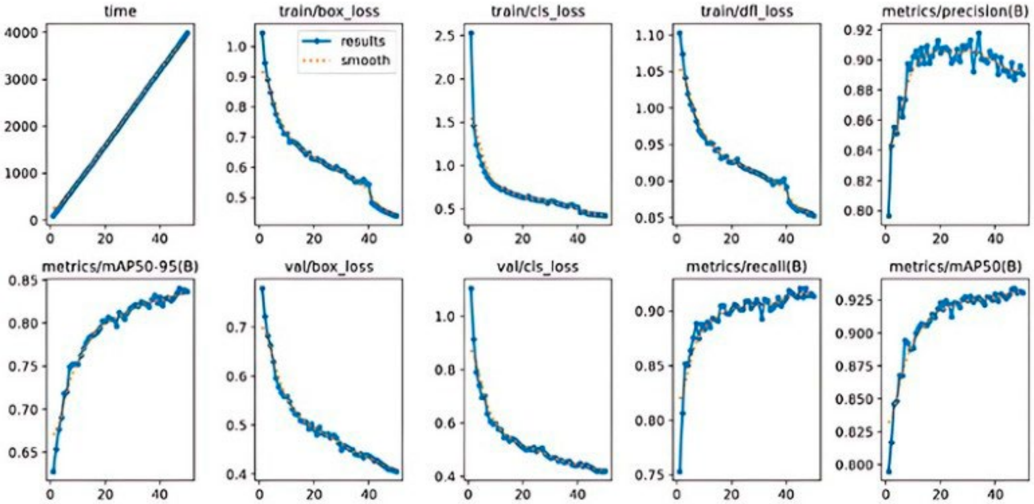


FIGURA 3. **Gráficas de pérdida para YOLOv8 con 50 épocas.**

Adicionalmente, el desempeño de YOLOv8 con 100 épocas se resume en la Tabla 6, destacando una precisión de hasta 0.970 para tapa morada *("purple_cap")* y una mínima de 0.892 para identificar tubos vacíos o que no contienen sangre *("not_ok")*. *Recall* presenta valores de 0.976 hasta 0.877 para la detección de tubos con tapa morada *("purple_cap")* y tapa amarilla *("yellow_cap")* respectivamente. La clase con mayor *m*AP (0.993) es para tapa morada *("purple_cap"),* mientras que el menor valor fue 0.898 para la clase que reporta los tubos que no contienen sangre *("not_ok").* En resumen, YOLOv8 presenta desempeño similar tanto en 50 y 100 épocas.

TABLA 6. **Resultados del rendimiento del modelo de YOLOv8 con 3000 imágenes entrenado con 100 épocas.**

| Clase | Precisión | *Recall* | *m*AP |
|---|---|---|---|
| blue_cap | 0.901 | 0.895 | 0.953 |
| good | 0.928 | 0.928 | 0.943 |
| label | 0.928 | 0.915 | 0.947 |
| not_ok | 0.892 | 0.892 | 0.898 |
| purple_cap | 0.970 | 0.976 | 0.993 |
| yellow_cap | 0.881 | 0.877 | 0.916 |

## Implicaciones para laboratorios clínicos

Los laboratorios clínicos tienen un rol significativo al proveer de información para la toma de decisiones de los profesionales de salud. No obstante, existe un grado de falibilidad en pruebas o dispositivo de laboratorio, por lo que la ocurrencia de errores es factible en las diferentes fases de una prueba de laboratorio (preanalítica, analítica y postanalítica)[43]. Este estudio aborda la evaluación de diferentes modelos de DL para identificar objetos en imágenes de tubos con muestras de sangre de un laboratorio clínico como un medio para asistir al profesional de salud en la inspección de riesgos en la etapa preanalítica. Respecto a precisión, este métrico fluctúa entre 0 y 1, donde valores cercanos a 1 indican una alta tasa de detección de clases de manera correcta[33], siendo YOLOv8 la arquitectura que presentó el mayor valor, 0.927, con 50 épocas. Respecto a *recall,* es decir, la capacidad del modelo de identificar todas las clases en la imagen[44][45][46], el modelo YOLOv8 presentó los mayores valores, destacando 0.914 con 100 épocas. Al evaluar la capacidad del modelo para detectar imágenes de distintas clases[47], YOLOv8 presentó un valor de *m*AP de 0.941 y 0.940 con 100 y 50 épocas respectivamente, demostrando un equilibrio entre la precisión y *recall* en todas las clases durante la validación[45]. Estos resultados indican que el modelo YOLOv8 presentó un mejor desempeño general, seguido por YOLOv5 y YOLOv7.

Estos resultados presentan a los algoritmos de DL como una herramienta factible para asistir al técnico flebotomista en la inspección de características importantes que pueden ser omitidas por la naturaleza repetitiva del proceso y dependiente del factor humano, incluyendo omitir etiqueta de identificación, extraer cantidad no adecuada de sangre, utilizar tubo incorrecto (con o sin preservativo o anticoagulante), entre otros. Particularmente, la identificación errónea del paciente[6] y el consecuente problema para comunicar resultados, afectan a la prestación de servicios de diagnóstico, siendo reconocidos como objetivos para la mejora de la calidad en este sector[4]. Estos errores pueden provocar la retoma de muestra, retraso del análisis, alteración o contaminación de la sangre, teniendo un impacto en el diagnóstico y salud de los pacientes, así como en la calidad y eficiencia del proceso. Aunque se tiende a responsabilizar al personal de atención médica por los errores, en realidad la mayoría de estos problemas surgen debido a la falta de procesos seguros y bien estructurados[5]. A pesar de su importancia, existe poca información sobre el nivel de precisión al realizar actividades de inspección en el sector salud, particularmente en laboratorios clínicos. En sectores industriales, la tasa de errores al realizar inspección visual varía por múltiples factores[48], con tasas entre 20 % y 30 % de manera general[8], mientras que en el sector de construcción va de 19 % a 48 %[49] y de 17.8 % a 29.8 % en sector metalmecánica[50].

En la evaluación de la calidad en la atención médica existen varias herramientas, incluyendo los indicadores de calidad (QIs), los cuales proporcionan medidas objetivas basadas en evidencia para evaluar diferentes aspectos críticos del cuidado de manera consistente[51]. Al respecto, la clasificación de errores en laboratorio según su gravedad, es crucial para identificar áreas prioritarias de mejora en la calidad[4], por lo que, el entendimiento de los QIs permite a los laboratorios clínicos identificar las acciones correctivas y mejoras más adecuadas para la resolución de problemas[52]. Los indicadores han sido agrupados en seis etapas del proceso de pruebas de laboratorio: i) orden de la prueba, ii) identificación del paciente y recolección de la muestra, iii) identificación, preparación y transporte de la muestra, iv) análisis, v) informe de resultados, así como vi) interpretación de resultados y acciones posteriores[53]. A pesar de que la implementación sistemática los QIs en laboratorios puede ser efectiva para reducir los errores, mejorar la seguridad del paciente y cumplir con los requisitos de la norma ISO 15189, en la práctica existen dificultades para mantener la recolección de datos de manera estandarizada y sistemática, así como para fomentar un continuo interés, compromiso y dedicación en todo el personal[52]. Al respecto, los resultados de este estudio contribuyen a brindar información sobre la factibilidad detectar desviaciones y evaluar indicadores correspondientes a las primeras etapas (i, ii e iii), con posibilidad de expansión al resto. En este sentido se ha sugerido que los laboratorios deben crear un mapa de proceso que describa todos los pasos del proceso de prueba desde la orden del médico hasta la entrega del resultado[6].

Las herramientas de IA como DL han apoyado en el diagnóstico médico permitiendo reducir el tiempo para iniciar un tratamiento[54], mediante la detección y clasificación del cáncer[55], diagnósticos mediante electrocardiogramas [56], detección de glaucoma[57] o la clasificación de tumores cerebrales[58], lo que permite reducir el tiempo para iniciar un tratamiento[54]. Particularmente, diferentes versiones de YOLO han sido evaluados en contextos variados como el farmacéutico para la detección de defectos en blíster alcanzando una precisión de 0.974[59], para detectar equipo médico de protección personal logrando un *m*AP de 97.2 %[60], o para detectar máscaras y tubos de ventilación en pacientes logrando una exactitud *(accuracy)* de 93 %[61]. Sin embargo, la evidencia es escasa sobre estudios que aborden DL para la evaluación y análisis de actividades de inspección visual en sector salud y particularmente en laboratorios clínicos. El presente estudio contribuye en la discusión sobre el uso de DL como apoyo al técnico flebotomista en la etapa preanalítica. En este sentido, estudios previos han identificado que el uso de IA puede facilitar la colaboración entre especialistas de la salud[62], permitiendo una mejor la utilización de servicios con el fin de optimizar los recursos[63]. Al respecto, es crucial involucrar al equipo del sector de la salud y centrarse en fortalecer las capacidades locales y adaptar la tecnología disponible a las necesidades específicas de los países latinoamericanos [64]. En la región, la IA se encuentra en fase inicial y muchas áreas carecen de la infraestructura necesaria[65]. Sin embargo, tecnologías como la IA tienen el potencial de mejorar el acceso a servicios, reducir consultas innecesarias, brindar comodidad a los usuarios, reducir la carga de trabajo, disminuir costos[66] o incluso disminuir las readmisiones no planificadas[67], lo cual, tiene un efecto directo a la calidad de atención al paciente[68].

## Inspección visual y aprendizaje profundo (DL)

Con respecto al desempeño de actividades de inspección visual en laboratorios clínicos, existe poca información sobre el grado de precisión o exactitud de los profesionales de salud. En comparación con actividades de inspección en sectores industriales, comúnmente la inspección visual humana es superada por algoritmos de DL[69][70], aunque existen excepciones mostrando resultados mixtos o comparables[71][72] o incluso un rendimiento inferior de los modelos en comparación con los inspectores humanos[73]. El presente estudio destaca una brecha entre la inspección totalmente automatizada y la necesidad continua de participación humana, principalmente en actividades críticas como las asociadas a la salud. En algunos escenarios, los algoritmos comienzan la inspección y los inspectores humanos intervienen en el caso de elementos inciertos o que caen por debajo de un umbral establecido[74][75]. De esta manera, los modelos de DL tienen el potencial de ayudar de manera consistente a los procesos de inspección al minimizar la participación humana y aliviar la fatiga física y mental. La utilización de algoritmos para detectar objetos empleando redes neuronales convolucionales, como YOLO, ofrece la posibilidad de realizar esta tarea de manera automatizada y eficiente. Adicionalmente, las herramientas de DL presentan la capacidad de identificar objetos que pudieran no ser evidentes a simple vista u omitidos debido a condiciones humanas de error. Considerando que hasta dos tercios de decisiones de los profesionales de salud son basadas en resultados de pruebas de laboratorio[76], DL representa un soporte para los profesionales de la salud y en última instancia, un apoyo para mejorar de la calidad de la atención médica.

### Retos Prácticos

Algunos retos identificados incluyen aspectos técnicos para la implementación de sistemas de inspección en laboratorio clínico basados en IA, incluyendo la dependencia al volumen y calidad de conjuntos de datos para su rendimiento[77]. Cuando no se dispone de grandes conjuntos de datos, es necesario realizar un entrenamiento previo[78] o transferencia de aprendizaje[79] como alternativas para trabajar con conjunto de datos pequeños. Otro aspecto y tal vez más relevante es el factor humano, donde la transición de inspección tradicionalmente realizada por un profesional de la salud, a una inspección asistida por tecnología puede representar un reto, principalmente por desconocimiento, falta de infraestructura, resistencia al cambio, entre otros. Adicionalmente, un reto futuro en contextos de laboratorios clínicos es el uso de más arquitecturas de DL, incluyendo modelos de dos etapas como redes neuronales convolucionales basadas en regiones (R-CNN), *Fast* R-CNN[80], *Faster* R-CNN[81] y *Mask* R-CNN[82], que, aunque requieren conjuntos de datos más grandes, pueden aportar una comparativa integral.

### Limitaciones

El estudio no está exento de limitaciones, incluyendo el tamaño del conjunto de datos, que puede considerarse pequeño, por lo que un trabajo futuro incluye su ampliación. El estudio consideró una variabilidad limitada en las condiciones del entorno de las imágenes incluyendo la iluminación y los tipos de tubos, por lo que se proyecta ampliar estas características. Adicionalmente, la incorporación de más y diferentes fondos de imagen *(background)* podrá ayudar a mejorar aún más el desempeño de las arquitecturas evaluadas para detectar objetos con diferentes fondos. Un aspecto importante es el desequilibrio de clases, el cual en este estudio no es un problema crítico, ya que el rango de anotaciones va desde 1,725 a 2,580, como se mostró en Tabla 1, sin embargo, un trabajo futuro será balancear las anotaciones de clase, aumentar el número de imágenes con respecto a las clases a identificar, así como implementar data augmentation para el incremento de imágenes. finalmente, las clases referentes a volumen de sangre, solo consideraron presencia o ausencia de sangre en el tubo, por lo que el siguiente paso será entrenar al modelo en diferentes niveles o cantidades presentes en el tubo.

## CONCLUSIONES

El presente estudio abordó el uso de tecnología asociada a inteligencia artificial como lo es *deep learning,* en la detección de características relevantes al realizar inspecciones en el área de flebotomía de un laboratorio clínico público. Los resultados indican que los modelos YOLO y particularmente YOLOv8 mostraron un buen desempeño en identificar adecuadamente las clases analizadas. Si bien no se cuenta con información sobre el nivel de precisión del personal de salud al realizar actividades de inspección en laboratorios clínicos, estos valores rondan entre 85 y 90 % en otros sectores, por lo que los resultados de este estudio son alentadores y sugieren que los modelos de DL son capaces de detectar características relevantes en la fase preanalítica con una precisión superior. Además, estos resultados abren posibilidades de desarrollo para sistemas de inspección que asistan a los profesionales de salud y complementen la función de inspección que realizan. El estudio brinda información valiosa para futuras investigaciones, incluyendo la inspección de más variables críticas en laboratorios clínicos, la realización de análisis más detallados, así como optimizaciones específicas para cada arquitectura, teniendo en cuenta las necesidades y requisitos específicos del laboratorio clínico. Dada la escasa literatura sobre aplicaciones específicas de DL en el llenado e identificación de muestras de sangre en laboratorios clínicos, los hallazgos respaldan el potencial de estas herramientas para asistir a los profesionales de salud en actividades de inspección visual y con ello contribuir a la mejora de los servicios de salud.

## DECLARACIÓN ÉTICA

El presente trabajo no involucró la participación de seres humanos ni animales. Por lo tanto, no se requirió el consentimiento informado de personas. El protocolo de investigación fue debidamente registrado ante el Instituto de Servicios de Salud Pública del Estado de Baja California.

## AGRADECIMIENTOS

Agradecemos al Consejo Nacional de Humanidades, Ciencias y Tecnologías (CONAHCyT) y a la Universidad Autónoma de Baja California por el apoyo brindado para la realización del presente estudio.

## CONTRIBUCIÓN DE LOS AUTORES

I. F.-A. conceptualización, curación de datos, análisis formal y escritura de manuscrito original; J. A.-D. conceptualización, análisis formal, escritura de manuscrito original y revisión final; Y. B.-L. investigación, metodología, administración de proyecto y validación; J. L.-R. curación de datos, software, validación y revisión final; M. M. S.-Q. investigación, visualización y administración de proyecto; D. T. conceptualización, análisis formal, supervisión, escritura de manuscrito original y revisión final.

## REFERENCIAS

**[1]** R. Singh, A. Sharma, N. Sharma, R. Gupta, “Impact of Adam, Adadelta, SGD on CNN for White Blood Cell Classification,” en 2023 5th International Conference on Smart Systems and Inventive Technology (ICSSIT), Tirunelveli, India, 2023, pp. 1702-1709, doi: **https://doi.org/10.1109/ICSSIT55814.2023.10061068**

**[2]** International Standard Organization, “ISO-15189 Medical laboratories-Requirements for quality and competence,” 2022. [Online]. Available: **https://www.iso.org/standard/76677.html**

**[3]** G. Lima-Oliveira et al., “Laboratory Diagnostics and Quality of Blood Collection,” J. Med. Biochem., vol. 34, no. 3, pp. 288-294, 2015, doi: **https://doi.org/10.2478/jomb-2014-0043**

**[4]** M. Plebani, “The detection and prevention of errors in laboratory medicine,” Ann. Clin. Biochem., vol. 47, no. 2, pp. 101-110, 2010, doi: **https://doi.org/10.1258/acb.2009.009222**

**[5]** A. Aita, A. Padoan, G. Antonelli, L. Sciacovelli, M. Plebani, “Patient safety and risk management in medical laboratories: theory and practical application,” J. Lab. Precis. Med., vol. 2, no. 2, pp. 75-75, 2017, doi: **https://doi.org/10.21037/jlpm.2017.08.14**

**[6]** J. H. Nichols, “Laboratory quality control based on risk management,” Ann. Saudi Med., vol. 31, no. 3, pp. 223-228, 2011, doi: **https://doi.org/10.4103/0256-4947.81526**

**[7]** M. Plebani, “Exploring the iceberg of errors in laboratory medicine,” Clin. Chim. Acta, vol. 404, no. 1, pp. 16-23, 2009, doi: **https://doi.org/10.1016/j.cca.2009.03.022**

**[8]** J. Hickner et al., “Testing process errors and their harms and consequences reported from family medicine practices: A study of the American Academy of Family Physicians National Research Network,” Qual Saf. Health Care, vol. 17, no. 3, pp. 194-200, 2008, doi: **https://doi.org/10.1136/qshc.2006.021915**

**[9]** H. Javadifard et al., “Predicting Patient Waiting Time in Phlebotomy Units Using a Deep Learning Method,” en Proceedings - 2019 Innovations in Intelligent Systems and Applications Conference, ASYU 2019, Izmir, Turkey, 2019, pp. 1-4, doi: **https://doi.org/10.1109/ASYU48272.2019.8946380**

**[10]** B. Fridath, A. Gildas, A. Dooguy, “White Blood Cells Recognition and Classification using Convolutional Neural Network,” en 2023 2nd International Conference on Applied Artificial Intelligence and Computing (ICAAIC), Salem, India, 2023, pp. 145-150, doi: **https://doi.org/10.1109/ICAAIC56838.2023.10140293**

**[11]** J. Wang, “Deep Learning in Hematology: From Molecules to Patients,” Clin. Hematol. Int., vol. 6, no. 4, pp. 19-42, 2024, doi: **https://doi.org/10.46989/001c.124131**

**[12]** R. Chandra, A. K. Geico, “Deep Learning Diagnostics: A Revolutionary Approach to Healthcare Insurance,” NeuroQuantology, vol. 19, no. 12, pp. 745-754, 2021, doi: **https://doi.org/10.48047/nq.2021.19.12.NQ21278**

**[13]** H. Hou, R. Zhang, J. Li, "Artificial intelligence in the clinical laboratory," Clin. Chim. Acta, vol. 559, 2024, art. no. 119724, doi: **https://doi.org/10.1016/j.cca.2024.119724**

**[14]** D. Tlapa et al., "Effects of Lean Interventions Supported by Digital Technologies on Healthcare Services: A Systematic Review," Int. J. Environ. Res. Public Health, vol. 19, no. 15, 2022, art. no. 9018, doi: **https://doi.org/10.3390/ijerph19159018**

**[15]** G. L. Tortorella et al., "Effects of contingencies on healthcare 4.0 technologies adoption and barriers in emerging economies," Technol. Forecast Soc. Change, vol. 156, 2020, art. no. 120048, doi: **https://doi.org/10.1016/j.techfore.2020.120048**

**[16]** M. Ibrahim et al., "Augmentation in Healthcare: Augmented Biosignal Using Deep Learning and Tensor Representation," J. Healthc Eng., vol. 2021, 2021, art. no. 6624764, doi: **https://doi.org/10.1155/2021/6624764**

**[17]** E. Lee et al., "Transforming hospital emergency department workflow and patient care," Interfaces, vol. 45, no. 1, pp. 58-82, 2015, doi: **https://doi.org/10.1287/inte.2014.0788**

**[18]** G. Arcidiacono, A. Pieroni, "The revolution Lean Six Sigma 4.0," Int. J. Adv. Sci. Eng. Inf. Technol., vol. 8, no. 1, pp. 141-149, 2018, doi: **https://doi.org/10.18517/ijaseit.8.1.4593**

**[19]** J. De Mast et al., "Process improvement in healthcare: Overall resource efficiency," Qual. Reliab. Eng. Int., vol. 27, no. 8, pp. 1095-1106, 2011, doi: **https://doi.org/10.1002/qre.1198**

**[20]** D. Marshall et al., "Selecting a dynamic simulation modeling method for health care delivery research - Part 2: Report of the ISPOR dynamic simulation modeling emerging good practices task force," Value Health, vol. 18, no. 2, pp. 147-160, 2015, doi: **https://doi.org/10.1016/j.jval.2015.01.006**

**[21]** A. M. Ismael, A. Şengür, "Deep learning approaches for COVID-19 detection based on chest X-ray images," Expert. Syst. Appl., vol. 164, 2021, art. no. 114054, doi: **https://doi.org/10.1016/j.eswa.2020.114054**

**[22]** J. I. Arreola Minjarez et al., "Detection of COVID-19 Lung Lesions in Computed Tomography Images Using Deep Learning," Rev. Mex. Ing. Biomed., vol. 43, no. 1, pp. 7-18, 2022, doi: **https://doi.org/10.17488/RMIB.43.1.1**

**[23]** R. Indraswari et al., "Brain Tumor Detection on Magnetic Resonance Imaging (MRI) Images Using Convolutional Neural Network (CNN)," en 2022 9th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI), Jakarta, Indonesia, 2022, pp. 367-373, doi: **https://doi.org/10.23919/EECSI56542.2022.9946622**

**[24]** S. Guizani, N. Guizani, S. Gharsallaoui, "A Hybrid CNN-SVM Prediction Approach for Breast Cancer Ultrasound Imaging," en 2023 International Wireless Communications and Mobile Computing( IWCMC), Marrakesh, Morocco, 2023, pp. 1574-1578, doi: **https://doi.org/10.1109/iwcmc58020.2023.10182874**

**[25]** V. H. Benitez Baltazar et al., "Autonomic Face Mask Detection with Deep Learning: an IoT Application," Rev. Mex. Ing. Biomed., vol. 42, no. 2, pp. 160-170, 2021, doi: **https://doi.org/10.17488/RMIB.42.2.13**

**[26]** J. Fan, T. Huo, X. Li, "A review of one-stage detection algorithms in autonomous driving," en 2020 4th CAA International Conference on Vehicular Control and Intelligence (CVCI), Hangzhou, China, 2020, pp. 210-214, doi: **https://doi.org/10.1109/CVCI51460.2020.9338663**

**[27]** Q. Tang et al., "A Two-stage Raman Imaging Denoising Algorithm Based on Deep Learning," en 2022 Asia Communications and Photonics Conference (ACP), Shenzhen, China, 2022, pp. 2096-2099, doi: **https://doi.org/10.1109/ACP55869.2022.10088904**

**[28]** J. Redmon et al., "You only look once: Unified, real-time object detection," en 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 779-788, doi: **https://doi.org/10.1109/CVPR.2016.91**

**[29]** S.-H. Chen et al., "A surface defect detection system for golden diamond pineapple based on CycleGAN and YOLOv4," J. King Saud Univ. - Comput. Inf. Sci., vol. 34, no. 10, pp. 8041-8053, 2022, doi: **https://doi.org/10.1016/j.jksuci.2022.07.018**

**[30]** J. M. Kimeu et al., "Deep learning-based mobile application for the enhancement of pneumonia medical imaging analysis: A case-study of West-Meru Hospital," Inform. Med. Unlocked, vol. 50, 2024, art. no. 101582, doi: **https://doi.org/10.1016/j.imu.2024.101582**

**[31]** S. Shahinfar, P. Meek, G. Falzon, "'How many images do I need?' Understanding how sample size per class affects deep learning model performance metrics for balanced designs in autonomous wildlife monitoring," Ecol. Inform., vol. 57, 2020, art. no. 101085, doi: **https://doi.org/10.1016/j.ecoinf.2020.101085**

**[32]** A. Cruz et al., "Detection of grapevine yellows symptoms in Vitis vinifera L. with artificial intelligence," Comput. Electron. Agric., vol. 157, pp. 63-76, 2019, doi: **https://doi.org/10.1016/j.compag.2018.12.028**

**[33]** D. E. Brown et al., "Neural network methods for diagnosing patient conditions from cardiopulmonary exercise testing data," BioData Min., vol. 15, no. 1, 2022, art. no. 16, doi: **https://doi.org/10.1186/s13040-022-00299-6**

**[34]** H. Wang, Q. Li, D. Zhang, D. Li, Q. Chen, "Key Components of Deep Metric Learning," en 2022 2nd International Conference on Consumer Electronics and Computer Engineering (ICCECE), Guangzhou, China, 2022, pp. 648-651, doi: **https://doi.org/10.1109/ICCECE54139.2022.9712771**

**[35]** V. Gonzalez-Huitron et al., "Disease detection in tomato leaves via CNN with lightweight architectures implemented in Raspberry Pi 4," Comput. Electron. Agric., vol. 181, 2021, art. no. 105951, doi: **https://doi.org/10.1016/j.compag.2020.105951**

**[36]** P. T. Q. Anh, D. Q. Thuyet, Y. Kobayashi, "Image classification of root-trimmed garlic using multi-label and multi-class classification with deep convolutional neural network," Postharvest Biol. Technol., vol. 190, 2022, art. no. 111956, doi: **https://doi.org/10.1016/j.postharvbio.2022.111956**

**[37]** M. Glučina et al., "Automated Detection and Classification of Returnable Packaging Based on YOLOV4 Algorithm," Appl. Sci., vol. 12, no. 21, 2022, art. no. 11131, doi: **https://doi.org/10.1109/ACP55869.2022.10088904**

**[38]** Y. Ke et al., “Empowering Intelligent Home Safety: Indoor Family Fall Detection with YOLOv5,” en 2023 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress (DASC/PiCom/CBDCom/CyberSciTech), Abu Dhabi, United Arab Emirates, 2023, pp. 0942-0949, doi: **https://doi.org/10.1109/DASC/PiCom/CBDCom/Cy59711.2023.10361490**

**[39]** G. P. Bhat, N. G. Cholli, “Effective object detection using Tensorflow facilitated YOLOv3 model,” en 2021 IEEE International Conference on Computation System and Information Technology for Sustainable Solutions (CSITSS), Bangalore, India, 2021, pp. 1-8, doi: **https://doi.org/10.1109/CSITSS54238.2021.9683109**

**[40]** M. L. Francies, M. M. Ata, M. A. Mohamed, “A robust multiclass 3D object recognition based on modern YOLO deep learning algorithms,” Concurr. Comput. Pract. Exp., vol. 34, no. 1, 2022, art. no. e6517, doi: **https://doi.org/10.1002/cpe.6517**

**[41]** R. Ammu, N. Sinha, “Small Segment Emphasized Performance Evaluation Metric for Medical Images,” en 2020 International Conference on Signal Processing and Communications (SPCOM), Bangalore, India, 2020, pp. 1-5, doi: **https://doi.org/10.1109/SPCOM50965.2020.9179617**

**[42]** S. Raj, Y. Gupta, R. Malhotra, “License Plate Recognition System using Yolov5 and CNN,” en 2022 8th International Conference on Advanced Computing and Communication Systems, ICACCS, Coimbatore, India, 2022, pp. 372-377, doi: **https://doi.org/10.1109/ICACCS54159.2022.9784966**

**[43]** S. W. Njoroge, J. H. Nichols, “Risk management in the clinical laboratory,” Ann. Lab. Med., vol. 34, no. 4, pp. 274-278, doi: **https://doi.org/10.3343/alm.2014.34.4.274**

**[44]** R. K. Bania, “Ensemble of deep transfer learning models for real-time automatic detection of face mask,” Multimed. Tools Appl., vol. 82, no. 16, pp. 25131-25153, 2023, doi: **https://doi.org/10.1007/s11042-023-14408-y**

**[45]** I. C. Juanatas, R. A. Juanatas, “Convolution Neural Network Approach for Facial Mask Detection,” en 2023 IEEE 12th Global Conference on Consumer Electronics (GCCE), Nara, Japan, 2023, pp. 1152-1155, doi: **https://doi.org/10.1109/GCCE59613.2023.10315441**

**[46]** L. Kamal, R. J. R. Raj, “Harnessing deep learning for blood quality assurance through complete blood cell count detection,” e-Prime - Adv. Electr. Eng. Electron.Energy, vol. 7, 2024, art. no. 100450, doi: **https://doi.org/10.1016/j.prime.2024.100450**

**[47]** P. Henderson, V. Ferrari, “End-to-end training of object class detectors for mean average precision,” en 13th Asian Conference on Computer Vision, Taipei, Taiwan, 2017, pp. 198-213, doi: **https://doi.org/10.1007/978-3-319-54193-8_13**

**[48]** J. E. See, “Visual Inspection: A Review of the Literature,” Sandia National Laboratories, Albuquerque, New Mexico, Rep. no. SAND2012-8590, 2012. [Online]. Available: doi: **https://www.osti.gov/servlets/purl/1055636**

**[49]** B. A. Graybeal et al., “Visual Inspection of Highway Bridges,” J. Nondestruct. Eval., vol. 21, pp. 67-83, 2002, doi: **https://doi.org/10.1023/A:1022508121821**

**[50]** M. M. Stallard (Voelker), C. A. MacKenzie, F. E. Peters, “A probabilistic model to estimate visual inspection error for metalcastings given different training and judgment types, environmental and human factors, and percent of defects,” J. Manuf. Syst., vol. 48, pp. 97-106, 2018, doi: **https://doi.org/10.1016/j.jmsy.2018.07.002**

**[51]** M. Plebani et al., “Quality indicators in laboratory medicine: A fundamental tool for quality and patient safety,” Clin. Biochem., vol. 46, no. 13-14, pp. 1170-1174, 2013, doi: **https://doi.org/10.1016/j.clinbiochem.2012.11.028**

**[52]** L. Sciacovelli et al., “Quality Indicators in Laboratory Medicine: The status of the progress of IFCC Working Group ‘laboratory Errors and Patient Safety’ project,” Clin. Chem. Lab. Med., vol. 55, no. 3, pp. 348-357, 2017, doi: **https://doi.org/10.1515/cclm-2016-0929**

**[53]** S. Shahangian, S. R. Snyder, “Laboratory medicine quality indicators: A review of the literature,” Am. J. Clin. Pathol., vol. 131, no. 3, pp. 418-431, 2009, doi: **https://doi.org/10.1309/ajcpjf8ji4zldque**

**[54]** P. E. A. Kuriki, F. C. Kitamura, “Artificial Intelligence in Radiology: A Private Practice Perspective From a Large Health System in Latin America,” Semin. Roentgenol., vol. 58, no. 2, pp. 203-207, 2023, doi: **https://doi.org/10.1053/j.ro.2023.01.006**

**[55]** R. Pandian, V. Vedanarayanan, D. N. S. Ravi Kumar, R. Rajakumar, “Detection and classification of lung cancer using CNN and Google net,” Meas.: Sens., vol. 24, 2022, art. no. 100588, doi: **https://doi.org/10.1016/j.measen.2022.100588**

**[56]** I. Shahin, A. B. Nassif, M. B. Alsabek, “COVID-19 Electrocardiograms Classification using CNN Models,” en 14th International Conference on Developments in eSystems Engineering, DeSE 2021, Sharjah, Emiratos Árabes Unidos, 2021, pp. 448-452, doi: **https://doi.org/10.1109/DESE54285.2021.9719358**

**[57]** H. J. Sandoval-Cuellar et al., “Image-based Glaucoma Classification Using Fundus Images and Deep Learning,” Rev. Mex. Ing. Biomed., vol. 42, no. 3, pp. 28-41, 2021, doi: **https://doi.org/10.17488/RMIB.42.3.2**

**[58]** A. U. Haq et al., “MCNN: a multi-level CNN model for the classification of brain tumors in IoT-healthcare system,” J. Ambient. Intell. Humaniz. Comput., vol. 14, no. 5, pp. 4695-4706, 2022, doi: **https://doi.org/10.1007/s12652-022-04373-z**

**[59]** R. Patgiri, V. Ajantha, S. Bhuvaneswari, V. Subramaniyaswamy, “Intelligent Defect Detection System in Pharmaceutical Blisters Using YOLOv7,” en 2024 Second International Conference on Emerging Trends in Information Technology and Engineering (ICETITE), Vellore, India, 2024, pp. 1-7, doi: **https://doi.org/10.1109/ic-ETITE58242.2024.10493735**

**[60]** B. Wu, C. Pang, X. Zeng, X. Hu, “ME-YOLO: Improved YOLOv5 for Detecting Medical Personal Protective Equipment,” Appl. Sci., vol. 12, no. 23, 2022, art. no. 11978, doi: **https://doi.org/10.3390/app122311978**

**[61]** Q. T. Do, J. Chaudri, “Creating Computer Vision Models for Respiratory Status Detection,” en 2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), Glasgow, Reino Unido, 2022, pp. 1350-1353, doi: **https://doi.org/10.1109/EMBC48229.2022.9871978**

**[62]** D. Aguilar Bucheli et al., “Artificial Intelligence in Medical Education: Latin American context,” Metro Ciencia, vol. 31, no. 2, pp. 21-34, 2023, doi: **https://doi.org/10.47464/MetroCiencia/vol31/2/2023/21-34**

**[63]** F. M. Delpino et al., “Emergency department use and Artificial Intelligence in Pelotas: design and baseline results,” Rev. Bras. Epidemiol., vol. 26, 2023, art. no. e230021, doi: **https://doi.org/10.1590/1980-549720230021**

**[64]** R. Garcia Alonso, U. Thoene, D. Davila Benavides, “Digital Health and Artificial Intelligence: Advancing Healthcare Provision in Latin America,” IT Prof., vol. 24, no. 2, pp. 62-68, 2022, doi: **https://doi.org/10.1109/MITP.2022.3143530**

**[65]** F. C. Kitamura et al., “Forging Connections in Latin America to Advance AI in Radiology,” Radiol. Artif. Intell., vol. 4, no. 5, 2022, art. no. e220125, doi: **https://doi.org/10.1148/ryai.220125**

**[66]** M. Álvarez Vega, L. M. Quirós Mora, M. V. Cortés Badilla, “Inteligencia artificial y aprendizaje automático en medicina,” Rev. Med. Sinerg., vol. 5, no. 8, 2020, art. no. e557, doi: **https://doi.org/10.1109/CVCI51460.2020.9338663**

**[67]** C. Fierro, J. Pérez, J. Mora, “Predicting Unplanned Readmissions with Highly Unstructured Data,” 2020, arXiv:2003.11622, doi: **https://doi.org/10.48550/arXiv.2003.11622**

**[68]** M. Jamei et al., “Predicting all-cause risk of 30-day hospital readmission using artificial neural networks,” PLoS One, vol. 12, no. 7, 2017, art. no. e0181173, doi: **https://doi.org/10.1371/journal.pone.0181173**

**[69]** F. Chang et al., “A mobile vision inspection system for tiny defect detection on smooth car-body surfaces based on deep ensemble learning,” Meas. Sci. Technol., vol. 30, no. 12, 2019, art. no. 125905, doi: **https://doi.org/10.1088/1361-6501/ab1467**

**[70]** A. Rachman, R. M. C. Ratnayake, “Machine learning approach for risk-based inspection screening assessment,” Reliab. Eng. Syst. Saf., vol. 185, pp. 518-532, 2019, doi: **https://doi.org/10.1016/j.ress.2019.02.008**

**[71]** J. Aust, D. Pons, “Comparative Analysis of Human Operators and Advanced Technologies in the Visual Inspection of Aero Engine Blades,” Appl. Sci., vol. 12, no. 4, 2022, art. no. 2250, doi: **https://doi.org/10.3390/app12042250**

**[72]** O. Wosner, G. Farjon, A. Bar-Hillel, “Object detection in agricultural contexts: A multiple resolution benchmark and comparison to human,” Comput. Electron. Agric., vol. 189, 2021, art. no. 106404, doi: **https://doi.org/10.1016/j.compag.2021.106404**

**[73]** T. Y. Kim et al., “A Deep Learning Technique for Optical Inspection of Color Contact Lenses,” Appl. Sci., vol. 13, no. 10, 2023, art. no. 5966, doi: **https://doi.org/10.3390/app13105966**

**[74]** K. Y. Chan et al., “Ball bonding inspections using a conjoint framework with machine learning and human judgement,” Appl. Soft. Comput., vol. 102, 2021, art. no. 107115, doi: **https://doi.org/10.1016/j.asoc.2021.107115**

**[75]** I. Rio-Torto et al., “Hybrid Quality Inspection for the Automotive Industry: Replacing the Paper-Based Conformity List through Semi-Supervised Object Detection and Simulated Data,” Appl. Sci., vol. 12, no. 11, 2022, art. no. 5687, doi: **https://doi.org/10.3390/app12115687**

**[76]** M. S. Shaikh et al., “Performance evaluation of a coagulation laboratory using Sigma metrics,” Int. J. Health Care Qual. Assur., vol. 31, no. 6, pp. 600-608, 2018, doi: **https://doi.org/10.1108/ijhcqa-07-2017-0134**

**[77]** M. Cheng et al., “MicroCrack-Net: A Deep Neural Network with Outline Profile-Guided Feature Augmentation and Attention-Based Multiscale Fusion for MicroCrack Detection of Tantalum Capacitors,” IEEE Trans. Aerosp. Electron. Syst., vol. 58, no. 6, pp. 5141-5152, 2022, doi: **https://doi.org/10.1109/TAES.2022.3181117**

**[78]** S. Feng, H. Zhou, H. Dong, “Using deep neural network with small dataset to predict material defects,” Mater. Des., vol. 162, pp. 300-310, 2019, doi: **https://doi.org/10.1016/j.matdes.2018.11.060**

**[79]** D. Wang, D. He, “Channel pruned YOLO V5s-based deep learning approach for rapid and accurate apple fruitlet detection before fruit thinning,” Biosyst. Eng., vol. 210, pp. 271-281, 2021, doi: **https://doi.org/10.1016/j.biosystemseng.2021.08.015**

**[80]** R. Girshick, “Fast R-CNN,” en 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 2015, pp. 1440-1448, doi: **https://doi.org/10.1109/ICCV.2015.169**

**[81]** S. Ren et al., “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks,” IEEE Trans. Pattern. Anal. Mach. Intell., vol. 39, no. 6, pp. 1137-1149, 2017, doi: **https://doi.org/10.1109/TPAMI.2016.2577031**

**[82]** K. He et al., “Mask R-CNN,” en 2017 IEEE International Conference on Computer Vision (ICCV), Venecia, Italia, 2017, pp. 2980-2988, doi: **https://doi.org/10.1109/ICCV.2017.322**

# MATERIAL SUPLEMENTARIO

**TABLA 7.** **Resultados del rendimiento del modelo de YOLOv5 con 3000 imágenes entrenado con *20* épocas.**

| Clase | Precisión | Recall | *m*AP |
|---|---|---|---|
| blue_cap | 0.913 | 0.916 | 0.941 |
| good | 0.93 | 0.909 | 0.923 |
| label | 0.937 | 0.841 | 0.906 |
| not_ok | 0.875 | 0.876 | 0.872 |
| purple_cap | 0.969 | 0.979 | 0.987 |
| yellow_cap | 0.831 | 0.863 | 0.886 |

**TABLA 8.** **Resultados del rendimiento del modelo de YOLOv5 con 3000 imágenes entrenado con *100* épocas.**

| Clase | Precisión | Recall | *m*AP |
|---|---|---|---|
| blue_cap | 0.918 | 0.911 | 0.949 |
| good | 0.934 | 0.922 | 0.94 |
| label | 0.94 | 0.867 | 0.932 |
| not_ok | 0.911 | 0.867 | 0.897 |
| purple_cap | 0.975 | 0.981 | 0.993 |
| yellow_cap | 0.867 | 0.855 | 0.898 |

**TABLA 9.** **Resultados del rendimiento del modelo de YOLOv7 con 3000 imágenes entrenado con *20* épocas.**

| Clase | Precisión | Recall | *m*AP |
|---|---|---|---|
| blue_cap | 0.567 | 0.899 | 0.829 |
| good | 0.722 | 0.666 | 0.694 |
| label | 0.618 | 0.355 | 0.406 |
| not_ok | 0.639 | 0.558 | 0.562 |
| purple_cap | 0.919 | 0.707 | 0.838 |
| yellow_cap | 0.627 | 0.808 | 0.72 |

**TABLA 10.** **Resultados del rendimiento del modelo de YOLOv7 con 3000 imágenes entrenado con *100* épocas.**

| Clase | Precisión | Recall | *m*AP |
|---|---|---|---|
| blue_cap | 0.917 | 0.914 | 0.928 |
| good | 0.919 | 0.899 | 0.902 |
| label | 0.946 | 0.784 | 0.876 |
| not_ok | 0.894 | 0.857 | 0.86 |
| purple_cap | 0.981 | 0.963 | 0.986 |
| yellow_cap | 0.845 | 0.853 | 0.886 |

TABLA 11. **Resultados del rendimiento del modelo de YOLOv8 con 3000 imágenes entrenado con *20* épocas.**

| Clase | Precisión | Recall | *m*AP |
|---|---|---|---|
| blue_cap | 0.928 | 0.914 | 0.947 |
| good | 0.933 | 0.923 | 0.935 |
| label | 0.929 | 0.879 | 0.936 |
| not_ok | 0.893 | 0.892 | 0.89 |
| purple_cap | 0.99 | 0.981 | 0.99 |
| yellow_cap | 0.857 | 0.856 | 0.9 |

TABLA 12. **Resultados del rendimiento del modelo de YOLOv8 con 3000 imágenes entrenado con *100* épocas.**

| Clase | Precisión | Recall | *m*AP |
|---|---|---|---|
| blue_cap | 0.901 | 0.895 | 0.953 |
| good | 0.928 | 0.928 | 0.943 |
| label | 0.928 | 0.915 | 0.947 |
| not_ok | 0.892 | 0.892 | 0.898 |
| purple_cap | 0.97 | 0.976 | 0.993 |
| yellow_cap | 0.881 | 0.877 | 0.916 |

RESEARCH ARTICLE

VOL. 46 | NO. 1 | **JANUARY-APRIL 2025** | PP 1-28

Revista Mexicana de Ingeniería Biomédica

https://dx.doi.org/10.17488/RMIB.46.1.3

E-LOCATION ID: e1480

# Synthetic Data Generation for Pediatric Diabetes Research Using GANs and WGANs

## Generación de Datos Sintéticos para la Investigación de la Diabetes Infantil Usando GANs y WGANs

*Antonio García-Domínguez*[1], *Carlos E. Galván-Tejada*[1], *Rafael Magallanes-Quintanar*[1] ✉, *Miguel Cruz-López*[2], *Miguel Alexander Vázquez-Moreno*[2], *Erika Acosta-Cruz*[3]

[1] Universidad Autónoma de Zacatecas, Unidad Académica de Ingeniería Eléctrica, Zacatecas - México

[2] Centro Médico Nacional Siglo XXI, Instituto Mexicano del Seguro Social, Unidad de Investigación Médica en Bioquímica, Hospital de Especialidades, Ciudad de México - México

[3] Universidad Autónoma de Coahuila, Departamento de Biotecnología, Coahuila - México

## ABSTRACT

Pediatric diabetes research is often constrained by data scarcity, hindering the development of accurate predictive models for clinical applications. This study addresses this limitation by evaluating the effectiveness of Generative Adversarial Networks (GANs) and Wasserstein GANs (WGANs) in generating synthetic datasets that replicate the statistical properties of real pediatric diabetes data. A structured methodology was applied, incorporating preprocessing, model design, and dual evaluation metrics: Jensen-Shannon and Kullback-Leibler divergences for statistical fidelity, and a classification model to assess practical utility. Results demonstrate that both models produce high-fidelity synthetic datasets, with WGANs showing superior performance in capturing complex patterns due to improved training stability. Nonetheless, challenges remain in replicating the inherent variability of pediatric data, influenced by growth and developmental factors. This work highlights the potential of synthetic data to augment pediatric diabetes datasets, facilitating the development of robust and generalizable predictive models. Limitations include the dependency on initial data quality and the specificity of the models to pediatric datasets. By addressing critical gaps in data availability, this study contributes to advancing AI-driven healthcare solutions in pediatric diabetes research.

**KEYWORDS:** Generative Adversarial Networks (GANs), pediatric diabetes, synthetic data generation, Wasserstein GANs

## RESUMEN

La investigación en diabetes pediátrica a menudo está limitada por la escasez de datos, lo que dificulta el desarrollo de modelos predictivos precisos para aplicaciones clínicas. Este estudio aborda esta limitación evaluando la efectividad de las Redes Generativas Antagónicas (GANs) y las Wasserstein GANs (WGANs) para generar conjuntos de datos sintéticos que replican las propiedades estadísticas de los datos reales de diabetes pediátrica. Se aplicó una metodología estructurada que incluye el preprocesamiento, diseño de modelos y métricas de evaluación dual: divergencias de Jensen-Shannon y Kullback-Leibler para evaluar la fidelidad estadística, y un modelo de clasificación para evaluar la utilidad práctica. Los resultados demuestran que ambos modelos generan datos sintéticos de alta fidelidad, siendo las WGANs superiores en la captura de patrones complejos gracias a su estabilidad de entrenamiento mejorada. Sin embargo, persisten desafíos para replicar la variabilidad inherente de los datos pediátricos, influida por el crecimiento y los factores de desarrollo. Este trabajo resalta el potencial de los datos sintéticos para aumentar los conjuntos de datos de diabetes pediátrica, facilitando el desarrollo de modelos predictivos robustos y generalizables. Las limitaciones incluyen la dependencia de la calidad de los datos iniciales y la especificidad de los modelos a los conjuntos de datos pediátricos. Este estudio contribuye a cerrar brechas críticas en la disponibilidad de datos, impulsando soluciones de salud personalizadas basadas en inteligencia artificial para la investigación en diabetes pediátrica.

**PALABRAS CLAVE:** diabetes infantil, generación de datos sintéticos, Redes Generativas Adversarias (GANs), Redes Generativas Adversarias de Wasserstein (WGANs)

**Corresponding author**

TO: Rafael Magallanes-Quintanar
INSTITUTION: UNIVERSIDAD AUTÓNOMA DE ZACATECAS
ADDRESS: JARDÍN JUAREZ 147, CENTRO, ZACATECAS 98000, ZAC, MÉXICO
EMAIL: tiquis@uaz.edu.mx

**Received:**

15 October 2024

**Accepted:**

26 December 2024

**Published:**

07 February 2025

## INTRODUCTION

In recent years, artificial intelligence (AI) has achieved considerable advancements in the healthcare sector, introducing groundbreaking methods for disease diagnosis, prediction, and management. AI-driven models, particularly those utilizing machine learning (ML), have become indispensable tools in enhancing the accuracy of medical diagnoses and in personalizing treatment plans[1][2][3]. These advancements are especially crucial in the management of chronic diseases such as diabetes, where early detection and precise management are vital to improving patient outcomes and quality of life[4][5]. The success of AI in healthcare, however, is heavily dependent on the availability of large, high-quality datasets, which are often difficult to obtain due to various constraints including privacy concerns and the logistical challenges of data collection[6].

Diabetes, a chronic metabolic disorder, requires diligent management and has become a significant global health crisis. It is one of the leading causes of death and disability worldwide, affecting over 537 million adults and contributing to millions of deaths each year[7]. The burden of diabetes is escalating rapidly, with the prevalence expected to rise substantially in the coming decades. The disease is a major contributor to heart disease, stroke, kidney failure, blindness, and lower limb amputations, imposing enormous strain on healthcare systems and economies globally[8]. Without proper management, diabetes can lead to severe and life-threatening complications, making it a critical focus of public health efforts worldwide. Within this global context, diabetes in specific populations, such as children, demands particular attention due to the unique challenges it presents.

Pediatric diabetes, encompassing both type 1 and type 2 diabetes, presents distinct challenges due to its early onset and the requirement for lifelong care. Type 1 diabetes, the most common form in children, is an autoimmune condition that results in the destruction of insulin-producing beta cells in the pancreas, needing continuous insulin therapy for survival[9]. The increasing incidence of pediatric diabetes, particularly type 1 diabetes, is a growing concern, as it often manifests during critical developmental periods, heightening the risk of long-term complications such as cardiovascular disease, neuropathy, and retinopathy[10]. Additionally, the rise of type 2 diabetes in children, largely driven by the global obesity epidemic, adds further complexity to disease management as it combines insulin resistance with the challenges of ongoing growth and development[11].

Effective management and early intervention in pediatric diabetes hinge on the development of accurate and reliable predictive models. These models are essential for identifying at-risk individuals, optimizing treatment strategies, and ultimately improving patient outcomes. However, the scarcity of large, high-quality datasets in pediatric diabetes research presents a significant barrier to the creation of such models. The challenges of collecting comprehensive data in children are compounded by ethical concerns, including the need to protect patient privacy and obtain informed consent, as well as the natural variability in children's growth patterns, which introduces additional complexity into data collection and analysis. As a result, researchers often must work with small, inconsistent datasets that are not sufficient to train robust machine learning models.

To address these challenges, synthetic data generation has emerged as a promising solution. Synthetic data are artificially created datasets that replicate the statistical properties of real-world data, allowing researchers to augment existing datasets or generate new ones when real data are insufficient or difficult to obtain[12][13]. This approach is particularly valuable in pediatric diabetes research, where data limitations can significantly hinder the development of effective AI models. By generating synthetic data, researchers can overcome the constraints imposed by small and inconsistent datasets, thereby enhancing the robustness and generalizability of predictive models. This not only eases the development of more accurate models but also helps in exploring a wider range of scenarios and outcomes that may be underrepresented in real-world data.

Among the various techniques for generating synthetic data, Generative Adversarial Networks (GANs) and their variant, Wasserstein GANs (WGANs), have gained prominence due to their ability to produce realistic and high-quality synthetic data[10]. GANs consist of two neural networks, a generator and a discriminator, which work in tandem to create data that closely mimic real-world observations. The generator is tasked with producing synthetic data, while the discriminator evaluates the realism of the generated data, driving the generator to continuously improve its outputs[14]. WGANs, an extension of traditional GANs, address some of the limitations of GANs by employing the Wasserstein distance as a measure of how well the generated data match the real data. This approach leads to a more stable and effective training process, making WGANs particularly suitable for generating synthetic datasets in complex and high-dimensional domains like medical data[15].

In the context of pediatric diabetes, the application of GANs and WGANs for generating synthetic data is particularly promising. These models are adept at capturing the intricate and multi-dimensional relationships inherent in medical data, making them ideal for creating synthetic datasets that can be used to train AI models. However, generating synthetic data is only one part of the equation, it is equally important to ensure that the synthetic data can serve as a valid substitute for real-world data. The Jensen-Shannon divergence[16] and the Kullback-Leibler divergence [17] evaluate the similarity between the distributions of synthetic and real data[18]. The Jensen-Shannon divergence, due to its symmetric and bounded nature, serves as a general measure of how closely the synthetic data align with the real data, with lower values indicating a closer match. The Kullback-Leibler divergence offers additional insight by highlighting specific areas where the synthetic data may differ from the real data. Additionally, the synthetic data's utility is assessed through a classification model trained on both real and synthetic datasets, using metrics such as accuracy, precision, recall, F1-score, and ROC AUC. This combined approach ensures a comprehensive evaluation of the synthetic data, both in terms of its statistical alignment with real data and its practical applicability in AI model training and medical research.

Building on the importance of evaluating synthetic data with robust metrics, next-generation sequencing (NGS) has significantly contributed to the field of pediatric diabetes by enabling the identification of genetic variants associated with the disease. NGS has been used to uncover key insights into monogenic diabetes and gene-environment interactions, providing a deeper understanding of disease mechanisms and potential therapeutic targets [19]. However, the vast and heterogeneous datasets generated by NGS pose challenges for integration into predictive models, largely due to insufficient annotation and the complexity of multi-omics data[20]. Synthetic data generation, as proposed in this study, offers a complementary approach to address these limitations by augmenting NGS-derived datasets and enhancing their utility for machine learning applications. By employing GANs and WGANs, the synthetic datasets created in this work aim to align closely with real-world data distributions while improving their applicability for predictive modeling in pediatric diabetes research. These challenges underscore the critical need for innovative approaches like synthetic data generation, which not only augment NGS-derived datasets but also enable their integration into machine learning models for predictive analysis.

The importance of generating high-quality synthetic data has been underscored in recent studies, which emphasize the need for advanced synthetic data generation techniques to overcome the inherent limitations of real-world data collection[21][22]. For instance, the application of GANs in generating synthetic diabetes data has been successfully demonstrated in various studies, including work by García-Domínguez *et al.*[4]. This study addresses the critical issue of data scarcity in pediatric diabetes by generating synthetic datasets using both GAN and WGAN methodologies. The quality of the generated data is rigorously evaluated using both the Jensen-Shannon divergence and the Kullback-Leibler divergence to ensure close alignment with the characteristics of real-world data. The aim is to provide valuable insights into effective practices for synthetic data generation in medical research, ultimately supporting the development of AI-driven healthcare solutions that enhance the management and outcomes of pediatric diabetes.

The main contribution of this work lies in the evaluation of GANs and WGANs for generating synthetic data specifically for pediatric diabetes, a domain that remains underexplored despite its clinical importance. Unlike previous studies that focus solely on statistical metrics, this study introduces a dual evaluation approach, combining traditional statistical divergence measures (Jensen-Shannon and Kullback-Leibler) with a supervised classification model. This comprehensive validation not only ensures the fidelity of the generated data but also demonstrates its practical applicability for predictive tasks, addressing a critical gap in the generation and evaluation of synthetic pediatric datasets.

This paper is organized as follows: The “Materials and Methods” section describes the pediatric diabetes dataset and the preprocessing steps for synthetic data generation. It then outlines the architectures of the GAN and WGAN models and the techniques employed in their design. Additionally, the section details the methods used to evaluate the quality of the generated data, including both distributional similarity metrics (Jensen-Shannon and Kullback-Leibler divergences) and the performance of a classification model trained on real and synthetic data. The “Results and Discussion” section presents the outcomes of the synthetic data generation process, including a comparison of the two methods based on these evaluation approaches. Finally, the “Conclusions” section summarizes the key findings, discusses implications for future research, and highlights the potential of synthetic data in enhancing AI-driven healthcare solutions.

## MATERIALS AND METHODS

This section presents the foundational methods and concepts employed in this study to generate and evaluate synthetic data for pediatric diabetes research. It begins with a comprehensive description of the dataset, highlighting the key clinical, biochemical, and genetic variables that are crucial for understanding diabetes in children. This is followed by an introduction to Generative Adversarial Networks (GANs) and their variant, Wasserstein GANs (WGANs), outlining their theoretical basis, mechanisms, and advantages in generating realistic synthetic data that mirrors complex medical datasets. Additionally, the section provides an overview of the evaluation metrics, including the Jensen-Shannon and Kullback-Leibler divergences, which are critical for assessing the quality and fidelity of the synthetic data relative to the original data, as well as the performance metrics derived from a classification model, which evaluate the practical applicability of the synthetic data in AI-driven tasks.

Figure 1 illustrates the methodological framework employed in this study. The process begins with the input of the original pediatric diabetes dataset, which undergoes preprocessing steps such as feature filtering, handling data incompleteness, and normalization to ensure the quality and consistency of the data. Subsequently, synthetic data is generated using two models: Generative Adversarial Networks (GAN) and Wasserstein GANs (WGAN). Finally, the synthetic datasets are evaluated through two complementary approaches: statistical divergence metrics (Jensen-Shannon and Kullback-Leibler) and a supervised classification model (Random Forest), which assesses the practical utility of the generated data. This pipeline provides a structured and comprehensive methodology for generating high-quality synthetic data tailored to the complexities of pediatric diabetes.

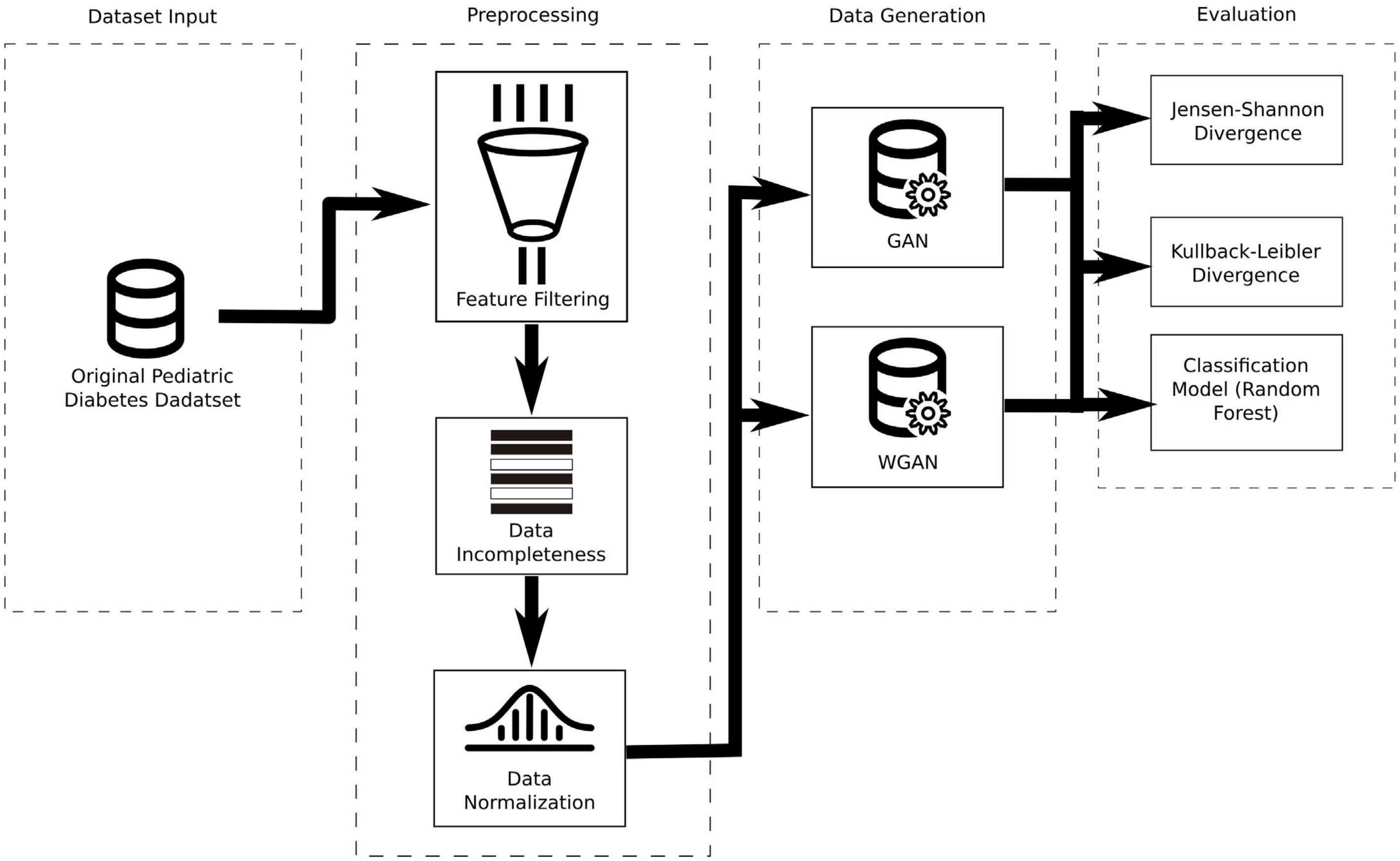


FIGURE 1. **Methodological framework for synthetic data generation and evaluation, including dataset input, preprocessing, data generation (GAN and WGAN), and evaluation through divergence metrics and classification modeling.**

## Original dataset description

The dataset used in this study consists of 834 records related to pediatric diabetes patients. It includes a range of clinical, biochemical, and genetic variables that are crucial for understanding various aspects of diabetes in children. The dataset encompasses 22 features, each representing a different attribute relevant to the diagnosis and management of diabetes. These features are summarized in Table 1.

TABLE 1. **Overview of Features in the Pediatric Diabetes Dataset.**

| Variable | Description |
|---|---|
| Location | Geographic location of the patient or the clinical study site. |
| StudyProject | The project or study identifier. |
| RecordID | A unique identifier for each patient record. |
| Sex | A binary variable indicating the sex of the patient, where '0' represents females (girls) and '1' represents males (boys). |
| Age | The age of the patient, measured in years. |
| BMI | Body Mass Index (BMI), which is a measure of body fat based on height and weight. |
| BMI_Z_Score | The z-score for BMI, which provides a standardized measure of BMI relative to a reference population. |
| BodyWeightStatus | A categorical variable indicating body weight classification, where '0' represents normal weight and '1' represents obesity. |
| SystolicBP | Systolic blood pressure, measured in millimeters of mercury (mmHg). |
| DiastolicBP | Diastolic blood pressure, also measured in millimeters of mercury (mmHg). |
| Glucose | Blood glucose level, measured in milligrams per deciliter (mg/dL). |
| TotalCholesterol | Total cholesterol level, measured in milligrams per deciliter (mg/dL). |
| Triglycerides | Triglyceride level, measured in milligrams per deciliter (mg/dL). |
| HDL_Cholesterol | High-density lipoprotein cholesterol (HDL-C) level, measured in milligrams per deciliter (mg/dL). |
| LDL_Cholesterol | Low-density lipoprotein cholesterol (LDL-C) level, measured in milligrams per deciliter (mg/dL). |
| Insulin | Insulin level, measured in micrograms per milliliter ( g/mL). |
| HOMA_IR | Homeostatic Model Assessment for Insulin Resistance (HOMA-IR), which is used to evaluate insulin resistance. |
| SalivaryAmylase | Salivary amylase activity, measured in units per liter (UI/L). |
| PancreaticAmylase | Pancreatic amylase activity, measured in units per liter (UI/L). |
| TotalAmylase | Total amylase activity, measured in units per liter (UI/L). |
| CNVs_SalivaryAmylase | Copy number variations of the gene associated with salivary amylase production |
| CNVs_PancreaticAmylase | Copy number variations of the gene associated with pancreatic amylase production |

Data corresponds to Mexican patients and were collected at the General Hospital *“Centro Médico Siglo XXI”* of the *Instituto Mexicano del Seguro Social (IMSS).* All participants provided informed consent prior to their inclusion in the study, titled *"Variación y funcionalidad del número de copias del gen de amilasa (AMY1, AMY2) y su asociación con microbiota intestinal en obesidad infantil".* The study adhered to the principles outlined in the Declaration of Helsinki, and the protocol received approval from the Ethics Committee of the *“Instituto Mexicano del Seguro Social”* and the *“Comisión Nacional de Investigación Científica”* (R-2016-785-097).

## Data preprocessing

Data preprocessing is a crucial step in preparing the dataset for synthetic data generation, ensuring that the models employed can effectively learn and generate high-quality synthetic samples that closely resemble real-world data. To achieve this, it is necessary to implement a series of systematic preprocessing steps, including the removal of irrelevant variables, the handling of data incompleteness, and the normalization of continuous variables. These procedures are outlined in detail below.

### *Feature filtering*

In this study, it is necessary to filter the dataset to eliminate features that are not relevant to the objectives. The features 'Location', 'StudyProject', and 'RecordID' are removed, as they do not provide meaningful information for the analysis of pediatric diabetes. 'Location' and 'StudyProject' are administrative variables without clinical, biochemical, or genetic relevance, while 'RecordID' serves merely as a unique identifier with no analytical value. Removing these features reduces dimensionality, thereby simplifying the dataset and enhancing both the efficiency and performance of the models. In the context of machine learning and data science, reducing dimensionality is a critical step that improves model interpretability and mitigates the risk of overfitting by eliminating noise and irrelevant data[23][24].

### *Data incompleteness*

Addressing missing values is essential in preparing medical datasets, where incomplete data can significantly impact the performance and reliability of machine learning models[25]. In this study, data incompleteness is managed through mean imputation, a technique widely employed for handling missing data, particularly when dealing with continuous variables. This approach involves replacing missing values with the mean of the corresponding feature, thereby preserving the overall central tendency and distribution of the data[26].

Mean imputation is applied by calculating the mean of each feature with missing values and substituting this mean for the missing entries[27]. Mathematically, for a given feature $X_j$ with missing values, the imputed value $\hat{x_{ij}}$ for any missing observation $x_{ij}$ is given by Equation (1).

$$\hat{x}_{ij} = \frac{1}{n}\sum_{i=1}^{n} x_{ij} \tag{1}$$

where $n$ represents the total number of observations for which the feature $X_j$ is not missing. This method assumes that the missing data are Missing At Random (MAR), implying that the probability of missingness is not related to the missing values themselves but may be related to the observed data. Mean imputation preserves the mean of the observed data and minimizes the impact on the overall distributional properties, maintaining the variance structure to some extent. However, it may reduce the natural variability of the dataset by introducing repeated values, which is an acceptable trade-off in this context due to the small proportion of missing data[28][29][30].

While other imputation methods, such as median or mode imputation, can handle non-normally distributed data or categorical variables, mean imputation is chosen for its computational efficiency and its ability to maintain the central tendency of the dataset. This ensures that the imputations do not introduce significant bias, allowing the synthetic data generation models to learn from a complete and consistent dataset while preserving its original characteristics as closely as possible.

### *Data normalization*

Data normalization is a fundamental step in preparing datasets for machine learning models, particularly neural networks, which are highly sensitive to the scale of input data[31]. In this study, Min-Max Scaling is employed to normalize the continuous variables, transforming each feature to a fixed range, typically [0, 1]. This normalization technique is defined by Equation (2).

$$X_{scaled} = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{2}$$

where $X$ represents the original feature values, $X_{min}$ is the minimum value, and $X_{max}$ is the maximum value of the feature[32]. Min-Max Scaling is chosen because it preserves the original distribution shape while ensuring that all features contribute equally to the model's learning process. This method is particularly advantageous over other normalization techniques, such as Z-score normalization, which centers the data around the mean and scales based on standard deviation but can be more sensitive to outliers and does not always maintain the relative spacing of values[33][34].

The use of Min-Max Scaling is especially beneficial in the context of Generative Adversarial Networks (GANs) and Wasserstein GANs (WGANs). These models require a stable training process to effectively generate high-quality synthetic data, and normalization helps by ensuring that gradients do not vanish or explode due to disparate feature scales. By standardizing the input data, the models can learn underlying patterns more effectively, thereby improving the robustness and quality of the synthetic data generated[35].

## Generative models

The scarcity and variability of high-quality pediatric diabetes data present significant challenges in developing reliable predictive models. To mitigate these issues, various generative models have been developed to create synthetic datasets that replicate the statistical properties and complex relationships found in real-world data. In this study, Generative Adversarial Networks (GANs) and Wasserstein Generative Adversarial Networks (WGANs) are utilized to generate synthetic data that closely mirrors actual patient data, thereby enhancing the robustness and generalizability of predictive tools for clinical application. These approaches also help address concerns related to data privacy and ethical considerations.

### *Generative Adversarial Networks (GANs)*

Generative Adversarial Networks (GANs), introduced by Goodfellow *et al.* [36], are a powerful class of generative models widely used for creating synthetic data that closely mimics real-world data distributions. GANs consist of two neural networks: a generator *G* and a discriminator *D,* which are trained simultaneously through an adversarial process. As depicted in Figure 2, the generator's role is to produce synthetic data samples by transforming random noise, typically drawn from a prior distribution such as a Gaussian, into data that resembles the real dataset. The discriminator, on the other hand, acts as a binary classifier that attempts to distinguish between genuine data samples and those generated by the generator.

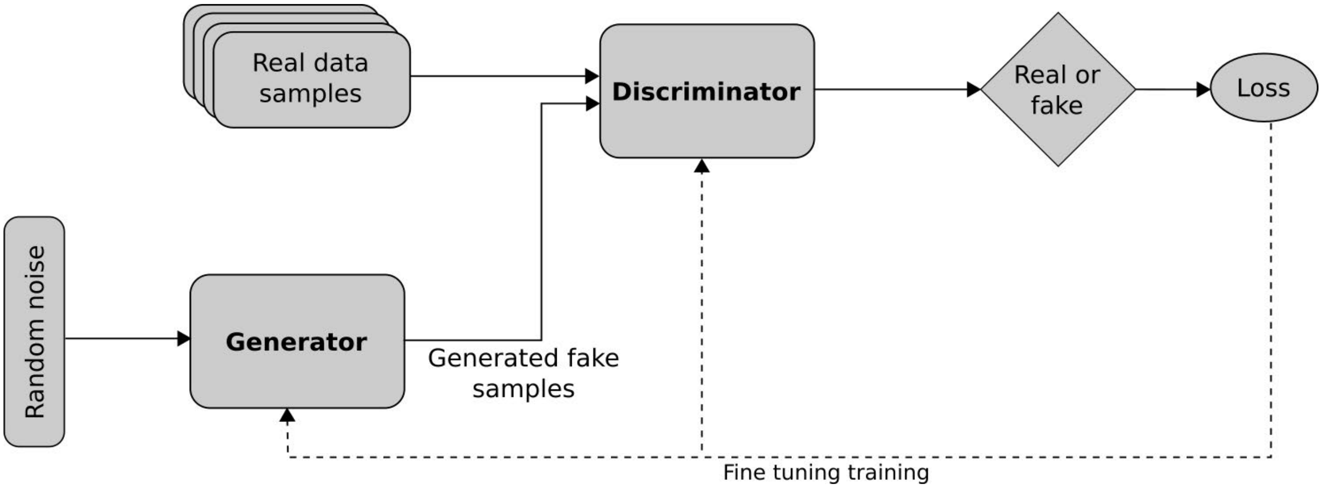


FIGURE 2. **Overview of the Generative Adversarial Network (GAN) Architecture.**

The adversarial training process of GANs is formulated as a minimax game, where the generator aims to "fool" the discriminator by producing data that the discriminator classifies as real, while the discriminator strives to correctly identify whether a given sample is real or fake. The objective function that governs this process is defined by Equation (3).

$$min_G max_D V(D,G) = E_{x \sim p_{data}(x)}[log D(x)] + E_{z \sim p_z(z)}[log(1 - D(G(z)))] \tag{3}$$

where:

- $x$ represents real data samples drawn from the true data distribution $p_{data}(x)$.
- $z$ is a latent variable (noise) sampled from a prior distribution $p_z(z)$.
- $G(z)$ denotes the synthetic sample generated by the generator from noise $z$.
- $D(x)$ provides the probability that a given input $x$ is a real sample.

In this framework, the generator *G* learns to minimize the probability of the discriminator correctly identifying its synthetic outputs as fake, while the discriminator *D* aims to maximize the accuracy of its predictions. This continuous feedback loop drives both networks to improve iteratively: the generator becomes better at creating realistic data, and the discriminator becomes more adept at detecting synthetic samples. The training reaches equilibrium when the discriminator cannot distinguish between real and generated data, outputting a probability close to 0.5 for both[36].

GANs rely on two critical components to function effectively:

1. Adversarial Loss: The loss function quantifies the performance of both the generator and the discriminator. For the discriminator, the loss function can be expressed as shown in Equation (4).

$$L_D = -(E_{x\sim p_{data}(x)}[log\ D(x)] + E_{z\sim p_z(z)}[log(1 - D(G(z)))]) \tag{4}$$

For the generator, the goal is to minimize the loss function as defined in Equation (5).

$$L_G = -E_{z\sim p_z(z)}[log\ D(G(z))] \tag{5}$$

2. Optimization and Stability: Training GANs can be challenging due to issues like mode collapse, where the generator produces limited diversity in outputs, and training instability caused by non-convergence or oscillations. Techniques such as using alternative loss functions (e.g., Wasserstein loss for WGANs [35]), careful tuning of hyperparameters, label smoothing, and architectural adjustments like batch normalization and dropout layers are employed to mitigate these problems [37].

### *Model design and implementation*

To effectively generate synthetic datasets that capture the complex characteristics of pediatric diabetes data for this study, a GAN was carefully designed and implemented with specific architectural choices and training parameters. These were selected to achieve an optimal balance between model complexity, training stability, and the generation of high-quality synthetic data. Table 2 provides a detailed overview of the key architectural components and training parameters used in this study.

TABLE 2. **GAN Architecture and Training Parameters for Synthetic Data Generation.**

| Component | Description |
|---|---|
| Generator | 4 dense layers: Input (150 neurons), Hidden layers (512, 1024, 2048 neurons) with Leaky ReLU activations, Batch Normalization after each hidden layer, Output layer with tanh activation. |
| Discriminator | 5 dense layers: Input layer, Hidden layers (2048, 1024, 512, 256 neurons) with Leaky ReLU activations, Dropout layers (rate = 0.4) after each hidden layer, Output layer with sigmoid activation. |
| Optimizer | RMSprop with learning rate of 0.0001 for both generator and discriminator. |
| Training epochs | 100 epochs with a batch size of 64. |
| Noise addition | Gaussian noise added to inputs of discriminator and final synthetic dataset to maintain diversity. |
| Label smoothing | Applied to real samples (0.9) and fake samples (0.1) to improve training stability. |
| Post-processing | Descaling and adjustment of discrete variables to ensure realistic data values. |

The architecture and training parameters of the GAN model were selected based on established practices in the literature and tailored to meet the requirements of this study. The generator and discriminator were designed with dense layers and activation functions, such as Tanh and Leaky ReLU, which are widely recognized for their ability to improve gradient flow and stabilize adversarial training[36][38]. The RMSprop optimizer, combined with a carefully chosen learning rate, was employed to ensure stable convergence during training[37]. To enhance model robustness and diversity in the generated data, additional techniques such as batch normalization, noise addition, label smoothing, and dropout were integrated. These approaches, supported by prior research, have proven effective in improving training stability and preventing overfitting in GAN architectures[39][40]. Post-processing steps, including descaling and adjustment of discrete variables, were applied to ensure the synthetic datasets reflected realistic and clinically meaningful values. Collectively, these design choices contributed to the successful generation of synthetic datasets that accurately capture the complex characteristics of pediatric diabetes data.

### *Wasserstein Generative Adversarial Networks (WGANs)*

Wasserstein Generative Adversarial Networks (WGANs) are an improved variant of traditional GANs, designed to address common challenges such as instability during training and mode collapse, where the generator fails to capture the full diversity of the real data distribution. Introduced by Arjovsky *et al.*[35], WGANs use the Wasserstein distance (also known as Earth Mover's Distance) as a loss function, which provides more stable gradients for training and leads to improved convergence properties. The Wasserstein distance is defined as shown in Equation (6).

$$W(p_r, p_g) = \ inf_{\gamma \epsilon \Pi(p_r, p_g)} E_{(x,y)\sim\gamma}[||x-y||] \tag{6}$$

where:

- $p_r$ is the distribution of the real data.
- $p_g$ is the distribution of the generated data by the generator $G$.
- $\Pi_{(p_r p_g)}$ represents the set of all joint distributions $\gamma(x,y)$ whose marginals are $p_r$ and $p_g$, respectively.

Using this distance allows WGANs to maintain meaningful gradients throughout the training process, even when there is limited overlap between the real and generated data distributions. Instead of the discriminator outputting a probability, as in traditional GANs, the critic in a WGAN outputs a real-valued score, which estimates the Wasserstein distance between the real and generated data distributions. The generator is trained to minimize this distance, while the critic is trained to maximize it. This formulation leads to more stable training, as it avoids problems such as vanishing gradients that can occur with other types of loss functions[41].

To ensure that the Wasserstein distance is valid, the critic function must satisfy the Lipschitz continuity condition, which is enforced by clipping the weights of the critic to a small range (e.g., [−0.01,0.01]). This weight clipping helps maintain stable gradients and promotes effective training, further reducing the likelihood of mode collapse and enhancing the quality of the synthetic data produced.

While WGANs introduce significant improvements in training dynamics through these changes, their overall architecture in terms of the core components (generator and critic) remains similar to that of a traditional GAN. Both models consist of a generator network that creates synthetic data and a discriminator (or critic in the case of WGAN) that evaluates the data. Therefore, the differences between GAN and WGAN primarily lie in the loss function and training strategies, rather than in the architectural layout itself [41][42]. For this reason, a separate figure illustrating the architecture of WGAN is not included, as it would visually mirror the figure already provided for the standard GAN.

### *Model design and implementation*

To effectively generate synthetic datasets that capture the complex characteristics of pediatric diabetes data for this study, a WGAN was carefully designed and implemented with specific architectural choices and training parameters. The chosen architecture and parameters aim to leverage the advantages of the Wasserstein distance, ensuring stable training and high-quality data generation. The details are shown in Table 3.

TABLE 3. **WGAN Architecture and Training Parameters for Synthetic Data Generation.**

| Component | Description |
|---|---|
| Generator | 3 dense layers: Input (150 neurons), Hidden layers (256, 512 neurons) with Leaky ReLU activations, Batch Normalization after each hidden layer, Output layer with tanh activation. |
| Critic (Discriminator) | 3 dense layers: Input layer, Hidden layers (512, 256 neurons) with Leaky ReLU activations, Dropout layers (rate = 0.4), and a final output layer without activation for Wasserstein loss computation. |
| Optimizer | RMSprop with a learning rate of 0.00005 for both the generator and the critic. |
| Training epochs | 100 epochs with a batch size of 32. |
| Weight Clipping | Applied to the critic with a clip value of 0.01 to enforce the Lipschitz continuity condition. |
| Post-processing | Descaling and adjustment of discrete variables to ensure realistic data values. |

The choice of these parameters and configurations reflects a balance between model complexity, training stability, and the ability to generate high-quality synthetic data. The generator and critic architectures, with three dense layers and Leaky ReLU activations, are specifically designed to capture the intricate patterns in pediatric diabetes data. Batch normalization is incorporated in the generator to stabilize gradient updates, while dropout is applied in the critic to prevent overfitting during adversarial training[37][41].

The RMSprop optimizer with a learning rate of 0.00005 was selected to ensure gradual convergence and prevent instability during the adversarial training process. This is particularly important in the WGAN framework, as the Wasserstein distance requires precise optimization to maintain the Lipschitz continuity condition[35]. Weight clipping with a small range (0.01) enforces this condition, further enhancing the stability and effectiveness of the training process[14].

Additionally, the number of training epochs (100) and the batch size (32) were carefully chosen to balance sufficient training iterations and computational efficiency. These configurations allow both the generator and critic to achieve optimal performance without overfitting the data. Post-processing adjustments, including descaling and the refinement of discrete variable values, ensure that the generated synthetic data remains realistic and clinically meaningful[38][39].

The WGAN framework was chosen for its ability to address limitations in traditional GANs, particularly in stabilizing the training process and generating diverse, high-quality data. By combining the standard GAN approach with the more robust WGAN framework, this study leverages the strengths of each model to produce synthetic datasets that are not only accurate but also diverse and suitable for developing reliable predictive models.

## Evaluation metrics

To rigorously assess the quality of the synthetic datasets generated by the GAN and WGAN models, three complementary evaluation methods are employed: Jensen-Shannon Divergence (JSD), Kullback-Leibler Divergence (KLD), and Predictive Model Evaluation through a Random Forest classifier. These metrics provide a comprehensive analysis of the fidelity, statistical alignment, and utility of the synthetic data by quantifying distributional similarities and assessing their practical application in predictive tasks.

### *Jensen-Shannon Divergence (JSD)*

The Jensen-Shannon Divergence (JSD) is a statistical measure used to quantify the similarity between two probability distributions[43]. As a symmetrized and smoothed version of the Kullback-Leibler divergence, the JSD provides a bounded and interpretable metric for assessing how one distribution diverges from another. It is particularly suitable for evaluating the performance of generative models, such as GANs and WGANs, by measuring the extent to which the synthetic data replicates the distributional characteristics of real-world data[44][45].

Mathematically, the JSD between two discrete probability distributions $P$ and $Q$ is defined as shown in Equation (7):

$$JSD(P \,||\, Q) = \frac{1}{2} KL(P \,||\, M) + \frac{1}{2} KL(Q \,||\, M) \tag{7}$$

where:

- $P$ and $Q$ are the two distributions being compared.
- $M = \frac{1}{2}(P+Q)$ is the average distribution.
- $KL(P \,||\, M)$ and $KL(Q \,||\, M)$ ar re the Kullback-Leibler divergences between each distribution and the mean distribution $M$.

The JSD is symmetric $(JSD(P \,||\, Q) = JSD(Q \,||\, P))$ and is bounded between 0 and 1. A value of 0 indicates that the two distributions are identical, while values closer to 1 indicate greater divergence. This makes JSD a robust metric for comparing distributions, as it remains finite and interpretable even when the distributions have non-overlapping support.

In this study, the JSD is employed to rigorously evaluate the similarity between the real pediatric diabetes dataset and the synthetic datasets generated by the GAN and WGAN models. The JSD is chosen because it provides a comprehensive measure of distributional similarity, capturing both general alignment and specific differences between the real and synthetic data distributions. This is particularly important given the complex nature of pediatric diabetes data, which may contain nuanced patterns and variability that must be accurately reflected in the synthetic data.

The implementation of the Jensen-Shannon Divergence (JSD) in this study is carried out through the following steps:

1. Data Alignment: Both the real and synthetic datasets are aligned to ensure they contain identical feature columns. Any missing values are removed to maintain consistency and comparability across all features.
2. Feature-wise Calculation of JSD: For each feature in the datasets, probability distributions are represented by constructing histograms using a consistent binning strategy. The JSD is computed for each feature by comparing the normalized histograms of real and synthetic data, providing a divergence measure that reflects the degree of similarity between the synthetic and real data distributions for that specific feature.
3. Aggregation Across Features: The JSD values calculated for all features are averaged to produce a single, overall measure of divergence. This mean JSD value captures the overall similarity between the synthetic and real datasets, considering all features simultaneously.

The use of JSD in this study is justified by its capacity to provide a clear and quantitative assessment of the extent to which the synthetic data replicates the distributional characteristics of the original data. By evaluating both global and feature-specific similarities, the JSD offers a robust measure of the performance of the generative models, ensuring that the synthetic datasets produced are suitable for further analysis and modeling within the context of pediatric diabetes research.

### *Kullback-Leibler Divergence (KLD)*

The Kullback-Leibler Divergence (KLD) is a fundamental measure from information theory that quantifies the difference between two probability distributions[17]. It is commonly used to evaluate the performance of generative models by measuring the "cost" or "loss of information" when approximating one probability distribution with another[36][46]. In this study, KLD is employed to assess how closely the synthetic datasets generated by the GAN and WGAN models match the distribution of the real pediatric diabetes data.

Mathematically, KLD between two discrete probability distributions $P$ (real data) and $Q$ (synthetic data) is defined as shown in Equation (8):

$$KL(P \mid\mid Q) = \sum_{i} P(i) log(\frac{P(i)}{Q(i)}) \tag{8}$$

where:

- $P(i)$ represents the probability of event $i$ in the real data distribution.
- $Q(i)$ represents the probability of event $i$ in the synthetic data distribution.

In this study, both JSD and KLD are used to provide a comprehensive evaluation of the synthetic data quality. The combination of these two metrics offers a robust framework for assessing how effectively the generative models replicate the distributional characteristics of the real pediatric diabetes data. While JSD provides a global measure of similarity that is symmetric and bounded, making it suitable for general assessments of distributional overlap, KLD offers a more detailed perspective by identifying specific divergences where the synthetic data may differ significantly from the real data[47][48].

The implementation of KLD follows these steps:

1. Data Preparation: Both the real and synthetic datasets are aligned to ensure they have matching feature

sets, with missing values removed to enable accurate comparison.

2. Feature-wise KLD Calculation: For each feature, probability distributions are estimated by constructing normalized histograms. KLD is computed for each feature by comparing the real and synthetic distributions, providing insight into specific areas where the synthetic data may diverge from the real data.
3. Aggregation Across Features: The KLD values for all features are aggregated to derive an overall measure of divergence. This overall KLD score reflects the general fidelity of the synthetic dataset in representing the real data, considering all features collectively.

The use of both JSD and KLD in this study is justified by their complementary strengths. JSD is robust to extreme discrepancies and provides an overall measure of similarity, making it useful for confirming that the synthetic data generally aligns with the real data. KLD, on the other hand, offers a more granular assessment, pinpointing exact areas of divergence, which is crucial for refining the generative models. Together, these metrics ensure that the synthetic datasets generated are not only globally similar to the real data but also accurately capture its detailed characteristics, thereby validating their use for subsequent analysis and predictive modeling in pediatric diabetes research.

### *Predictive Model Evaluation*

The quality of synthetic datasets is not only defined by their statistical alignment with real data but also by their practical utility in downstream machine learning tasks. To this end, a predictive model evaluation framework is incorporated as an additional metric for assessing the efficacy of the generated synthetic data. This method evaluates how well synthetic datasets support classification tasks, serving as a proxy for their utility in real-world applications.

The Random Forest algorithm is selected as the predictive model due to its inherent robustness, capacity to handle high-dimensional data, and ability to model complex feature interactions[49][50]. Random Forest operates as an ensemble of decision trees, where each tree is constructed using a bootstrap sample of the training data. At each decision node, a random subset of features is considered for splitting, introducing variability that enhances the generalizability of the model. The final prediction is obtained through majority voting (classification) or averaging (regression) across all trees, ensuring resilience to overfitting[51].

Mathematically, the Random Forest Classifier constructs $T$ decision trees, where each tree $t \epsilon T$ minimizes the Gini impurity $I_g$ for classification tasks, as shown in Equation (9).

$$I_g = \sum_{i=1}^{C} p_i(1 - p_i) \tag{9}$$

Here, $C$ denotes the number of classes, and $p_i$, is the proportion of samples belonging to class $i$ in the current node. This criterion guides the selection of optimal splits at each node, ensuring that the trees effectively capture the underlying patterns in the data[52].

By employing Random Forest, this study introduces a pragmatic approach to assess the synthetic data's relevance in predictive modeling. This framework complements statistical measures like Jensen-Shannon Divergence (JSD) and Kullback-Leibler Divergence (KLD) by providing a task-oriented perspective, ensuring a comprehensive evaluation of synthetic data quality.

## RESULTS AND DISCUSSION

This section presents the findings from the experimental processes applied in this study and discusses their implications in the context of pediatric diabetes research. The analysis begins with the outcomes of data preprocessing steps, including feature filtering, management of data incompleteness, and data normalization. It then describes the application of generative models (GANs and WGANs) to create synthetic datasets. The quality and fidelity of these generated datasets are subsequently assessed using the Jensen-Shannon Divergence (JSD), Kullback-Leibler Divergence (KLD), and Predictive Model Evaluation metrics. Together, these approaches provide a comprehensive evaluation of the synthetic data, offering insights into their distributional similarity to real-world data and their utility in downstream machine learning tasks. This combination ensures a robust assessment of their potential applicability for future research.

### Data Preprocessing

#### *Feature filtering*

The preprocessing of the dataset involves several critical steps to ensure its suitability for synthetic data generation and subsequent analysis. The original dataset comprises 834 records with 22 features, including various demographic, clinical, and biochemical variables relevant to pediatric diabetes research. To improve the dataset's quality and focus, three features are removed due to their lack of direct relevance to the study objectives, resulting in a refined dataset containing 19 features. The excluded features are detailed in Table 4.

TABLE 4. **Features Removed During Preprocessing.**

| Feature Name | Reason for removal |
|---|---|
| Location | Lack of direct relevance to study goals |
| StudyProject | |
| RecordID | |

#### *Data Incompleteness*

Handling data incompleteness presents a substantial challenge in the preprocessing stage. Among the 834 records, 462 (approximately 55.4 %) exhibit at least one missing value. To address this issue, an imputation strategy is employed, wherein missing values are replaced with the mean of their respective features. This approach maintains the dataset's integrity by preserving the total number of records and minimizing potential biases that could arise from data omission. Following this imputation process, the dataset retains all 834 records, as summarized in Table 5.

TABLE 5. **Summary of Handling Missing Data.**

| Metric | Value |
|---|---|
| Total number of records in the original dataset | 834 |
| Records with at least one missing value | 465 (55.4 %) |
| Total number of records after imputation | 834 |

These preprocessing steps are vital in preparing the dataset for further analysis. They ensure that the data is clean, complete, and properly scaled, thereby enhancing the reliability and quality of the synthetic data generated.

### *Data normalization*

After the imputation of missing values, a normalization process is applied to ensure consistency in the scale of the dataset's features, which is essential for the effective training of generative models. A Min-Max normalization technique is utilized, scaling each feature to a range between 0 and 1. This standardization mitigates the risk of features with larger numerical ranges disproportionately influencing the model training process, thereby ensuring that all features contribute equitably to the generative modeling efforts.

## Synthetic data generation and assessment

While traditional data augmentation (DA) techniques have proven effective in domains such as images or text, where transformations like rotations or interpolations preserve the semantics of the original data, their application to clinical numerical data presents significant limitations. These techniques often fail to adequately capture the complex relationships between variables or maintain the clinical coherence required in healthcare datasets. For instance, Shorten and Khoshgoftaar[53] highlight how DA techniques can enhance model generalization in image classification but note their limited applicability to more structured and sensitive domains. Similarly, Choi *et al.*[54] demonstrate that GANs outperform DA in generating tabular clinical data by better preserving multidimensional relationships and ensuring consistency.

To address these challenges, this study employs Generative Adversarial Networks (GANs) and Wasserstein GANs (WGANs), which are better suited to model multidimensional distributions and generate synthetic data with higher fidelity and clinical consistency. This approach not only replicates the statistical properties of the original data but also preserves intrinsic relationships between features, a critical requirement in clinical research.

The use of Generative Adversarial Networks (GANs) and Wasserstein Generative Adversarial Networks (WGANs) is crucial for generating synthetic datasets that closely replicate the distributional characteristics of real pediatric diabetes data. Each model's effectiveness in producing realistic synthetic data is assessed based on its ability to preserve the statistical properties of the original dataset and its utility in downstream machine learning tasks. To achieve this, the quality of the synthetic data is evaluated using three complementary metrics: Jensen-Shannon Divergence (JSD), Kullback-Leibler Divergence (KLD), and the performance of a supervised classification model. Together, these methods provide a rigorous assessment of the statistical alignment and practical applicability of the synthetic datasets, ensuring their relevance for research and predictive modeling.

### *Evaluation of GAN-Generated Synthetic Data*

The Generative Adversarial Network (GAN) is utilized to create a synthetic dataset designed to closely replicate the distributional characteristics of the original pediatric diabetes data. A synthetic dataset consisting of 417 records, representing 50 % of the original dataset's size, is generated to maintain a balance between capturing the complexity of the data and ensuring a robust comparison between the real and synthetic datasets.

The effectiveness of the GAN in generating realistic synthetic data is evaluated through a comparative analysis of the global distributions of the real and GAN-generated data. Figure 3 illustrates this comparison by aggregating the normalized values of all features into a single composite measure for each record, providing a comprehensive view of the similarity between the two datasets. This composite measure summarizes the overall alignment between the datasets, capturing both the central tendencies and variability of the original data.

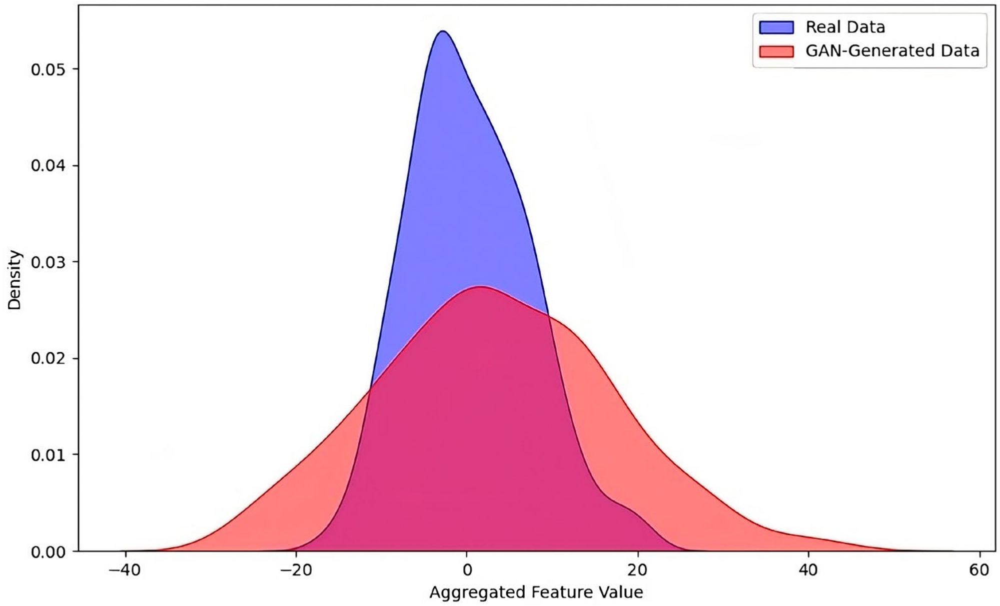


**FIGURE 3.** **Global Distribution Comparison Between Real and GAN-Generated Data.**

In Figure 3, the "Aggregated Feature Value" represents a combined metric derived by summing all normalized feature values for each record, while the probability density indicates the relative frequency of these aggregated values across the dataset. The curve corresponding to the real dataset is depicted in blue, while the red curve represents the GAN-generated dataset. The visualization demonstrates that the synthetic data generated by the GAN aligns well with the general shape of the real data distribution, capturing key patterns and trends, particularly in the central regions where the curves overlap closely.

Discrepancies are observed in the tails of the distribution, where the GAN fails to replicate less frequent and extreme values. This limitation arises from the training instability commonly associated with traditional GAN architectures, particularly when modeling complex, high-dimensional data. While the GAN successfully reproduces the dominant statistical features of the real dataset, these deviations highlight its reduced ability to capture rare patterns that require more stable and fine-tuned training processes.

To further quantify the similarity between the real and GAN-generated data, two divergence metrics are employed: the Jensen-Shannon Divergence (JSD) and the Kullback-Leibler Divergence (KLD). The values of these metrics, presented in Table 6, offer a detailed view of the effectiveness of the GAN in approximating the real data.

**TABLE 6.** **Divergence Metrics for GAN-Generated Synthetic Data.**

| Metric | Value |
|---|---|
| Jensen-Shannon Divergence (JSD) | 0.482 |
| Kullback-Leibler Divergence (KLD) | 1.617 |

As shown in Table 6, the Jensen-Shannon Divergence (JSD) value of 0.482 indicates a substantial degree of similarity between the distributions of the real and GAN-generated data. The relatively low JSD suggests that the GAN effectively captures the overall structure and dominant statistical features of the original dataset. This result confirms that the global patterns, including central tendencies and frequent values, are well-represented in the synthetic data.

TABLE 7. **Classification Performance Comparison between Real and GAN-Generated Synthetic Data.**

| Dataset | Accuracy | Precision | Recall | F1 Score | ROC AUC |
|---|---|---|---|---|---|
| Real data | 0.98 | 0.97 | 0.97 | 0.97 | 0.98 |
| Synthetic data | 0.94 | 0.93 | 0.93 | 0.93 | 0.94 |

The Kullback-Leibler Divergence (KLD) value of 1.617 provides additional insight into the alignment of the two distributions. While this value remains relatively low, it is slightly higher than the JSD, indicating that discrepancies are more pronounced in regions with lower probabilities, such as the tails of the distribution. This reflects the GAN's difficulty in modeling less frequent or extreme values, which aligns with the limitations observed visually in Figure 3. Together, these divergence metrics demonstrate that while the GAN-generated data closely approximates the real data, there is room for improvement in capturing rare patterns and achieving a more precise representation of the complete data distribution.

To complement this evaluation, the quality of the GAN-generated synthetic data is also assessed through its utility in a supervised learning context. A Random Forest classifier is trained separately on the real dataset and the GAN-generated synthetic dataset, using a binary classification task where the target variable indicates the presence or absence of diabetes. The feature Insulin (labeled as Insulin_ugml in the dataset) is used to define this target variable, as insulin levels are clinically recognized as a key indicator for diabetes diagnosis. Based on medical thresholds, records with insulin levels greater than 25 μU/mL are labeled as diabetic (1), while others are labeled as non-diabetic (0).

Key metrics, including accuracy, precision, recall, F1-score, and ROC AUC, are computed to evaluate the performance of the model in each case. The results, summarized in Table 7, provide a point of comparison to understand how well the synthetic data supports predictive modeling relative to the real data.

The performance metrics in Table 7 show that the GAN-generated synthetic data achieves strong and consistent results across all evaluated metrics. Although a slight drop of approximately 4 % is observed in accuracy, precision, recall, and F1-score compared to the real data, this difference is within acceptable limits for synthetic datasets. The marginal decline reflects the challenge of fully replicating the nuanced variability of real-world data, particularly in capturing less frequent patterns or edge cases.

The relatively high precision and recall values indicate that the synthetic data retains the critical predictive features necessary for identifying positive cases while effectively minimizing false positives. This balance is essential for clinical applications, where both sensitivity (recall) and specificity (precision) are critical for reliable predictions. Furthermore, the high ROC AUC score confirms that the model trained on GAN-generated data maintains a strong discriminative capability, demonstrating its ability to distinguish between classes effectively.

These findings highlight that, despite minor discrepancies, the GAN-generated data successfully preserves the underlying relationships and predictive features of the real dataset. This validates the GAN's practical utility for training machine learning models in scenarios where real data are scarce or access is restricted, offering a robust alternative for enhancing predictive modeling in pediatric diabetes research.

However, while these results are promising, they also reflect the inherent challenges of working with pediatric data, which tend to be highly variable due to the ongoing growth and development of children. The diversity in growth rates, physiological changes, and individual responses to health conditions can result in complex patterns that are difficult to replicate fully with synthetic models. These characteristics contribute to the observed differences between the real and synthetic data and underscore the need for careful consideration when generating synthetic datasets in this context.

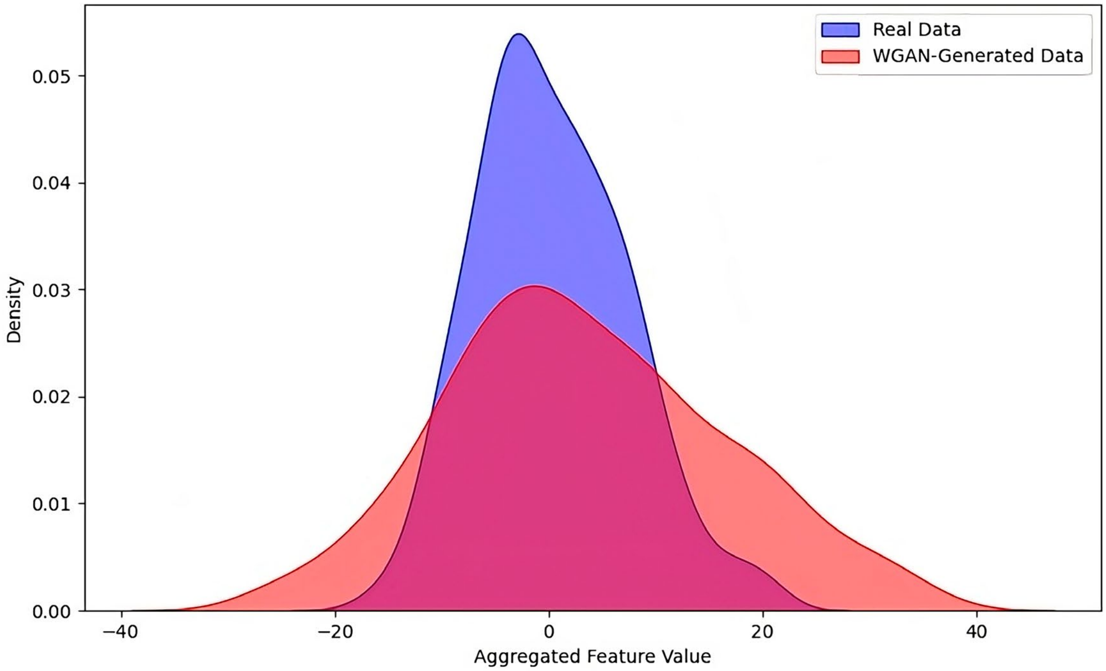


**FIGURE 4.** **Global Distribution Comparison Between Real and WGAN-Generated Data.**

Overall, the combined analysis of the visual comparison, divergence metrics, and supervised learning performance suggests that the GAN is highly effective in generating synthetic data that not only closely matches the general distribution of the real data but also retains the critical predictive features necessary for accurate classification. The comparable performance of the Random Forest classifier on both the real and synthetic datasets reinforces the validity of the synthetic data, demonstrating that it encapsulates the essential patterns and relationships required for supervised learning tasks. This comprehensive evaluation highlights the potential of GAN-generated data to serve as a reliable alternative for training machine learning models when real data is scarce or inaccessible.

### *Evaluation of WGAN-Generated Synthetic Data*

The Wasserstein Generative Adversarial Network (WGAN) is the second method used for generating synthetic data in this study, aimed at producing a dataset that closely replicates the characteristics of the original pediatric diabetes data. The WGAN model generates a synthetic dataset of 417 records, mirroring the size used in the GAN-generated dataset to maintain consistency in comparisons.

The performance of the WGAN is evaluated by comparing the global distributions of the real and WGAN-generated data, as shown in Figure 4. As in the GAN analysis, the normalized values of all features are aggregated into a single composite measure for each record, allowing a direct comparison of the overall alignment between the datasets.

TABLE 8. **Divergence Metrics for WGAN-Generated Synthetic Data.**

| Metric | Value |
|---|---|
| Jensen-Shannon Divergence (JSD) | 0.480 |
| Kullback-Leibler Divergence (KLD) | 1.584 |

This method provides a comprehensive view of how well the WGAN captures the underlying structure and variability of the real data.

In Figure 4, the "Aggregated Feature Value" represents a combined metric derived by summing all normalized feature values for each record, while the probability density indicates the relative frequency of these aggregated values across the dataset. The curve corresponding to the real dataset is depicted in blue, while the red curve represents the WGAN-generated dataset. The visualization shows that the WGAN-generated data aligns closely with the real data distribution, particularly around the central peak and mid-range values, where the curves exhibit almost perfect overlap. This alignment indicates that the WGAN successfully captures the most frequent and dominant statistical patterns in the data, closely approximating the global structure observed in the real dataset.

Compared to the traditional GAN, the WGAN demonstrates a significant reduction in discrepancies at the tails of the distribution, where less frequent and extreme values occur. This improvement reflects the enhanced stability of the WGAN training process, which allows it to model complex patterns with greater accuracy. By reducing deviations in the tails, the WGAN effectively retains both dominant and rare features, preserving the variability essential for realistic synthetic data generation. This capacity makes the WGAN-generated data not only a reliable representation of the real dataset but also a valuable resource for downstream machine learning tasks, particularly in scenarios with limited real data availability.

TABLE 9. **Classification Performance comparison between Real and WGAN-Generated Synthetic Data.**

| Dataset | Accuracy | Precision | Recall | F1 Score | ROC AUC |
|---|---|---|---|---|---|
| Real data | 0.98 | 0.97 | 0.97 | 0.97 | 0.98 |
| Synthetic data | 0.96 | 0.95 | 0.95 | 0.95 | 0.97 |

To further quantify the similarity between the real and WGAN-generated data, the same divergence metrics used for the GAN evaluation, the Jensen-Shannon Divergence (JSD) and the Kullback-Leibler Divergence (KLD), are calculated. The results, presented in Table 8, provide a detailed assessment of the WGAN's performance.

As shown in Table 8, the Jensen-Shannon Divergence (JSD) value of 0.480 is slightly lower than the value obtained with the GAN (0.482), indicating a marginal improvement in the similarity between the WGAN-generated and real data distributions. This lower JSD suggests that the WGAN more effectively captures the global statistical structure of the dataset, aligning closely with the dominant patterns observed in the real data.

The Kullback-Leibler Divergence (KLD) value of 1.584 is also marginally reduced compared to the GAN (1.617), highlighting a better representation of less frequent and extreme values in the tails of the distribution. This improvement reflects the enhanced stability of the WGAN training process, which enables it to model complex, low-probability patterns with greater accuracy. Collectively, these results demonstrate that the WGAN outperforms the traditional GAN in both global alignment and in capturing subtle variations, making it a more reliable approach for generating high-fidelity synthetic data.

Similar to the evaluation conducted for GAN, the quality of the WGAN-generated synthetic data is assessed through its utility in a supervised learning context. This analysis follows the same approach as the previous section, where a Random Forest classifier is trained on both the real dataset and the WGAN-generated synthetic dataset to predict the presence or absence of diabetes based on insulin levels. Key metrics such as accuracy, precision, recall, F1-score, and ROC AUC are calculated to evaluate the model's performance on both datasets. The results are summarized in Table 9.

As shown in Table 9, the Random Forest classifier trained on WGAN-generated synthetic data achieves performance metrics that are closely aligned with those obtained using real data. The accuracy, precision, recall, F1-score, and ROC AUC for the WGAN-generated data show only a slight decrease of approximately 2 % compared to the real dataset. This minor drop highlights the ability of the WGAN to generate high-fidelity synthetic data that preserves critical predictive features and relationships.

When compared to the GAN-generated data (Table 7), the WGAN demonstrates improved performance across all metrics, particularly in recall and ROC AUC. The slight enhancements in recall indicate that the WGAN-generated data better captures edge cases and less frequent patterns, reducing false negatives more effectively. Additionally, the ROC AUC value of 0.97 for the WGAN surpasses the 0.94 obtained with the GAN, showcasing the superior discriminative power of the WGAN-generated data.

The WGAN’s ability to produce synthetic data with enhanced stability and improved representation of subtle and less frequent patterns demonstrates its superiority over traditional GANs. By addressing the limitations observed in earlier models, the WGAN-generated data offer a reliable and realistic alternative for training machine learning models. This is particularly valuable in pediatric research, where access to real-world datasets is often limited, and the quality of synthetic data plays a crucial role in enabling robust predictive modeling.

The results obtained in this study, particularly the superior performance of Wasserstein Generative Adversarial Networks (WGANs) compared to traditional GANs, align with findings from other research on synthetic data generation in healthcare. Goncalves *et al.*[55] conducted a thorough examination of synthetic data generation techniques for cancer patient data, focusing on methods such as probabilistic models, Bayesian networks, and GANs. Their results demonstrated the potential of these techniques to approximate real datasets, achieving a Kullback-Leibler (KL) divergence score of 0.47 for cancer datasets, which is comparable to the results obtained in this study for pediatric diabetes data (KL divergence of 1.584 for WGAN-generated data). Although the divergence in this study is slightly higher, it is crucial to consider the unique challenges posed by pediatric data. Growth and developmental variability introduce significant complexity in modeling, making pediatric data inherently more difficult to capture accurately.

Other studies in synthetic healthcare data generation also reflect similar trends, particularly in the challenges of handling complex datasets. Xu *et al.*[56] applied Conditional GANs (CTGAN) to generate synthetic tabular data and found that CTGAN outperforms Bayesian networks and other GAN-based models on several real-world datasets. Although their focus was not on pediatric data, their results highlight the effectiveness of using conditional GANs for generating structured data patterns. However, their work, like many others, focuses primarily on adult datasets, where the variability present in pediatric populations is not a significant factor. This difference underscores the need for models capable of adapting to more dynamic physiological conditions, such as those found in children with diabetes, where growth and development create a more challenging environment for accurate synthetic data generation.

While the generation of synthetic healthcare data has advanced significantly, much of the focus has been on adult datasets, with limited attention paid to pediatric populations. This study addresses a significant gap in the literature by applying advanced generative models, specifically GANs and WGANs, to pediatric diabetes data. Pediatric datasets present unique challenges due to the high variability in physiological characteristics as children grow and develop. This inherent variability complicates the modeling process and requires more sophisticated approaches to accurately reflect the underlying data distributions. To address these complexities, WGANs were employed due to their stability in training and ability to handle the nuanced data patterns found in pediatric populations. Moreover, the use of Jensen-Shannon Divergence (JSD) and Kullback-Leibler Divergence (KLD) as evaluation metrics ensures that the generated datasets maintain the essential statistical properties necessary for predictive modeling.

To complement the evaluation of these divergence metrics, the quality of the WGAN-generated synthetic data was further assessed through its utility in a supervised learning context. A Random Forest classifier, trained on both real and WGAN-generated synthetic data, revealed that the synthetic data maintained essential predictive features, with classification performance metrics (accuracy, precision, recall, F1 score, and ROC AUC) comparable to the real data. These results demonstrate that the WGAN-generated data can be effectively used in predictive models for pediatric diabetes, showing that it not only captures the global distribution of the real data but also retains the critical features required for accurate classification.

By generating high-quality synthetic datasets for pediatric diabetes, this work can significantly enhance predictive models and personalized treatment approaches, addressing the critical scarcity of real-world pediatric datasets in this domain. To our knowledge, no studies to date have applied generative models specifically to pediatric data in the context of diabetes research, making this contribution particularly novel.

### *Limitations and adaptability*

While this study demonstrates the successful use of GANs and WGANs to generate high-quality synthetic datasets for pediatric diabetes research, certain limitations must be acknowledged. The quality and representativeness of the generated synthetic data depend heavily on the characteristics of the input dataset. Small sample sizes, incomplete data, or datasets with imbalanced classes may hinder the models' ability to capture rare or subtle patterns accurately. This limitation is particularly significant in pediatric datasets, where inherent variability due to growth, development, and physiological changes introduces additional challenges. Ensuring high-quality input data remains a critical prerequisite for achieving reliable synthetic data generation.

Another limitation relates to the specificity of the current methodology to pediatric diabetes data. Although GANs and WGANs are generalizable techniques, their application to other datasets or domains requires adaptation. For instance, datasets with different structures, such as high-dimensional multi-omics data or longitudinal patient records, may require modifications to the preprocessing pipeline, including tailored strategies for handling missing data, normalization techniques, or feature selection. Likewise, adjustments to the model architectures may be necessary to optimize performance for datasets with greater complexity or variability.

Despite these limitations, the proposed methodology provides a solid foundation for generating synthetic datasets in healthcare and beyond. Future research can explore its adaptability to other medical conditions, larger datasets, and alternative domains, such as genomics or environmental monitoring, where data scarcity is also a significant challenge. Additionally, further evaluation across diverse datasets could validate the generalizability of this approach, ensuring its broader applicability in machine learning-driven predictive modeling.

## CONCLUSIONS

This study explored the use of Generative Adversarial Networks (GANs) and Wasserstein GANs (WGANs) to generate synthetic datasets for pediatric diabetes research, addressing critical data scarcity in this domain. The dual approach combining statistical evaluation metrics and supervised classification validated the fidelity and practical utility of the synthetic data. By employing Jensen-Shannon Divergence (JSD) and Kullback-Leibler Divergence (KLD), the statistical alignment of synthetic datasets with real data was confirmed. Specifically, the GAN achieved a JSD of 0.482 and a KLD of 1.617, while the WGAN demonstrated superior performance with a JSD of 0.480 and a KLD of 1.584, indicating improved stability and fidelity in capturing complex data patterns.

The use of a Random Forest classifier to assess performance on classification tasks further demonstrated the capacity of synthetic datasets to retain predictive features essential for machine learning applications. The model trained on real data achieved an accuracy of 0.98, while the GAN-generated data attained 0.94, and the WGAN-generated data achieved 0.96, with similarly high scores for precision, recall, F1 score, and ROC AUC. These results highlight the WGAN's enhanced ability to approximate real dataset performance and reinforce its suitability for predictive modeling.

The results revealed that both GAN and WGAN methodologies successfully captured the complex distribution of pediatric diabetes data, with WGANs offering enhanced stability and improved fidelity. These findings are consistent with the theoretical advantages of the Wasserstein distance in mitigating common challenges in GAN training. Classification performance metrics further underscored the reliability of WGAN-generated data, which achieved metrics close to real data across accuracy, precision, recall, F1 score, and ROC AUC.

Beyond technical evaluations, this work contributes significantly by advancing the application of generative models in pediatric diabetes—a domain characterized by unique challenges, such as high variability in physiological characteristics and ethical constraints in data collection. Unlike prior studies that focused primarily on adult datasets or statistical metrics, this study introduces a comprehensive evaluation framework encompassing both distributional similarity and predictive utility.

Future work should aim to refine generative architectures and incorporate domain-specific knowledge to address the inherent complexities of pediatric datasets more effectively. Potential directions include exploring alternative loss functions, integrating temporal and longitudinal data, and developing hybrid models that merge generative techniques with rule-based systems. Expanding these methodologies to other pediatric conditions could further establish their relevance in healthcare research, paving the way for improved diagnostic tools and personalized treatment strategies.

This study demonstrates the transformative potential of synthetic data in overcoming real-world constraints, enhancing AI-driven solutions for pediatric diabetes management, and supporting broader applications in medical research. By rigorously validating synthetic datasets' fidelity and usability, this work lays the foundation for more robust and generalizable machine learning applications in healthcare.

## ETHICAL STATEMENT

The study protocol was approved by the Ethics Committee of the Mexican Social Security Institute (approval number: R-2016-785-097), in compliance with the guidelines set forth by the National Bioethics Commission (CONBIOETICA-09-CEI-009-20160601).

## CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

## AUTHOR CONTRIBUTIONS

A. G.-D. conceptualization, methodology, supervision, writing - original draft, validation, and funding acquisition; C. E. G.-T. conceptualization, software, visualization, methodology, writing - review & editing, and validation; R. M.-Q. conceptualization, data curation, formal analysis, investigation, writing - review & editing, and validation; M. C.-L. investigation, resources, validation, and data curation; M. A. V.-M. formal analysis, visualization, writing - review & Editing, and validation; E. A.-C. project administration, resources, writing - review & Editing.

## REFERENCIAS

**[1]** D. J. G. Carrasco Ramírez, M. Islam, and I. H. Even, “Machine Learning Applications in Healthcare: Current Trends and Future Prospects,” J. Artif. Intell. Gen. Sci., vol. 1, no. 1, Art. no. 1, 2024, doi: **https://doi.org/10.60087/jaigs.v1i1.33**

**[2]** A. Qayyum, J. Qadir, M. Bilal, and A. Al-Fuqaha, “Secure and Robust Machine Learning for Healthcare: A Survey,” IEEE Rev. Biomed. Eng., vol. 14, pp. 156-180, 2021, doi: **https://doi.org/10.1109/RBME.2020.3013489**

**[3]** M. Javaid et al., “Significance of machine learning in healthcare: Features, pillars and applications,” Int. J. Intell. Netw., vol. 3, pp. 58-73, 2022, doi: **https://doi.org/10.1016/j.ijin.2022.05.002**

**[4]** A. García-Domínguez et al., “Optimizing Clinical Diabetes Diagnosis through Generative Adversarial Networks: Evaluation and Validation,” Diseases, vol. 11, no. 4, 2023, art. no. 134, doi: **https://doi.org/10.3390/diseases11040134**

**[5]** A. García-Domínguez et al., "Diabetes Detection Models in Mexican Patients by Combining Machine Learning Algorithms and Feature Selection Techniques for Clinical and Paraclinical Attributes: A Comparative Evaluation," J. Diabetes Res., vol. 2023, 2023, art. no. 9713905, doi: https://doi.org/10.1155/2023/9713905

**[6]** J. G. Elmore and C. I. Lee, "Data Quality, Data Sharing, and Moving Artificial Intelligence Forward," JAMA Netw. Open, vol. 4, no. 8, 2021, art. no. e2119345, doi: https://doi.org/10.1001/jamanetworkopen.2021.19345

**[7]** H. Sun et al., "IDF Diabetes Atlas: Global, regional and country-level diabetes prevalence estimates for 2021 and projections for 2045," Diabetes Res. Clin. Pract., vol. 183, 2022, art. no. 109119, doi: https://doi.org/10.1016/j.diabres.2021.109119

**[8]** C. W. Tsao et al., "Heart Disease and Stroke Statistics—2022 Update: A Report From the American Heart Association," Circulation, vol. 145, no. 8, pp. e153-e639, 2022, doi: https://doi.org/10.1161/cir.0000000000001052

**[9]** L. C. Stene and J. Tuomilehto, "Epidemiology of Type 1 Diabetes," in Textbook of Diabetes, R.I.G. Holt and A. Flyvbjerg (Eds.)., Pondicherry, India: Wiley, 2024, ch. 4, pp. 41-54, doi: https://doi.org/10.1002/9781119697473.ch4

**[10]** G. Imperatore et al., "Projections of Type 1 and Type 2 Diabetes Burden in the U.S. Population Aged <20 Years Through 2050: Dynamic modeling of incidence, mortality, and population growth," Diabetes Care, vol. 35, no. 12, pp. 2515-2520, 2012, doi: https://doi.org/10.2337/dc12-0669

**[11]** P. Bjornstad et al., "Youth-onset type 2 diabetes mellitus: an urgent challenge," Nat. Rev. Nephrol., vol. 19, no. 3, pp. 168-184, 2023, doi: https://doi.org/10.1038/s41581-022-00645-1

**[12]** J. Jordon et al., "Synthetic Data -- what, why and how?," 2022, arXiv:2205.03257, doi: https://doi.org/10.48550/arXiv.2205.03257

**[13]** V. C. Pezoulas et al., "Synthetic data generation methods in healthcare: A review on open-source tools and methods," Comput. Struct. Biotechnol. J., vol. 23, pp. 2892-2910, 2024, doi: https://doi.org/10.1016/j.csbj.2024.07.005

**[14]** I. Goodfellow et al., "Generative Adversarial Networks," 2014, arXiv:1406.2661, doi: https://doi.org/10.48550/arXiv.1406.2661

**[15]** M. Arjovsky, S. Chintala, L. Bottou, "Wasserstein Generative Adversarial Networks," in Proceedings of the 34th International Conference on Machine Learning, Sydney, Australia, 2017, pp. 214-223, doi: https://dl.acm.org/doi/10.5555/3305381.3305404

**[16]** M. L. Menéndez, J. A. Pardo, L. Pardo, and M. C. Pardo, "The Jensen-Shannon divergence," J. Frankl. Inst., vol. 334, no. 2, pp. 307-318, 1997, doi: https://doi.org/10.1016/S0016-0032(96)00063-4

**[17]** S. Kullback and R. A. Leibler, "On Information and Sufficiency," Ann. Math. Stat., vol. 22, no. 1, pp. 79-86, 1951, doi: https://doi.org/10.1214/aoms/1177729694

**[18]** P. A. Apellániz et al., "Synthetic Tabular Data Validation: A Divergence-Based Approach," IEEE Access, vol. 12, pp. 103895-103907, 2024, doi: https://doi.org/10.1109/ACCESS.2024.3434582

**[19]** X. Donath et al., "Next-generation sequencing identifies monogenic diabetes in 16% of patients with late adolescence/adult-onset diabetes selected on a clinical basis: a cross-sectional analysis," BMC Med., vol. 17, no. 1, 2019, art. no. 132, doi: https://doi.org/10.1186/s12916-019-1363-0

**[20]** Y. Park, D. Heider, and A.-C. Hauschild, "Integrative Analysis of Next-Generation Sequencing for Next-Generation Cancer Research toward Artificial Intelligence," Cancers, vol. 13, no. 13, 2021, art. no. 3148, doi: https://doi.org/10.3390/cancers13133148

**[21]** P. S. Paladugu et al., "Generative Adversarial Networks in Medicine: Important Considerations for this Emerging Innovation in Artificial Intelligence," Ann. Biomed. Eng., vol. 51, pp. 2130-2142, 2023, doi: https://doi.org/10.1007/s10439-023-03304-z

**[22]** A. Aggarwal, M. Mittal, and G. Battineni, "Generative adversarial network: An overview of theory and applications," Int. J. Inf. Manag. Data Insights, vol. 1, no. 1, 2021, art no. 100004, doi: https://doi.org/10.1016/j.jjimei.2020.100004

**[23]** C. O. S. Sorzano, J. Vargas, and A. Pascual, "A survey of dimensionality reduction techniques", 2014, arXiv:1403.2877, doi: https://doi.org/10.48550/arXiv.1403.2877

**[24]** G. T. Reddy et al., "Analysis of Dimensionality Reduction Techniques on Big Data," IEEE Access, vol. 8, pp. 54776-54788, 2020, doi: https://doi.org/10.1109/ACCESS.2020.2980942

**[25]** P. C. Austin, I. R. White, D. S. Lee, and S. van Buuren, "Missing Data in Clinical Research: A Tutorial on Multiple Imputation," Can. J. Cardiol., vol. 37, no. 9, pp. 1322-1331, 2021, doi: https://doi.org/10.1016/j.cjca.2020.11.010

**[26]** T. Emmanuel et al., "A survey on missing data in machine learning," J. Big Data, vol. 8, no. 1, 2021, art. no. 140, doi: https://doi.org/10.1186/s40537-021-00516-9

**[27]** R. S. Somasundaram and R. Nedunchezhian, "Evaluation of Three Simple Imputation Methods for Enhancing Preprocessing of Data with Missing Values," Int. J. Comput. Appl., vol. 21, no. 10, pp. 14-19, 2011, doi: https://doi.org/10.5120/2619-3544

**[28]** R. J. A. Little and D. B. Rubin, Statistical Analysis with Missing Data, 3rd ed., Hoboken, NJ, USA: Wiley, 2019, doi: https://doi.org/10.1002/9781119482260

**[29]** C. K. Enders, Applied Missing Data Analysis, 2nd ed., Guilford Publications, 2022.

**[30]** S. van Buuren, Flexible Imputation of Missing Data, 2nd ed., NY, USA: CRC Press, 2018, doi: https://doi.org/10.1201/9780429492259

**[31]** A. Géron, Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems, 2nd ed., O'Reilly Media, Inc., 2019.

**[32]** S. G. K. Patro and K. K. Sahu, "Normalization: A Preprocessing Stage," Int. Adv. Res. J. Sci. Eng. Technol., pp. 20-22, 2015, doi: https://doi.org/10.17148/IARJSET.2015.2305

**[33]** D. Singh and B. Singh, "Investigating the impact of data normalization on classification performance," Appl. Soft Comput., vol. 97, 2020, art. no. 105524, doi: https://doi.org/10.1016/j.asoc.2019.105524

**[34]** L. Rahmad Ramadhan and Y. Anne Mudya, "A Comparative Study of Z-Score and Min-Max Normalization for Rainfall Classification in Pekanbaru," J. Data Sci., vol. 2024, 2024, art. no. 4. [Online]. Available: https://iuojs.intimal.edu.my/index.php/jods/article/view/446

**[35]** M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein GAN," 2017, arXiv:7101.07875, doi: https://doi.org/10.48550/arXiv.1701.07875

**[36]** I. Goodfellow et al., "Generative Adversarial Nets," in Advances in Neural Information Processing Systems, 2014. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2014/file/5ca3e9b122f61f8f06494c97b1afccf3-Paper.pdf

**[37]** T. Salimans et al., "Improved Techniques for Training GANs," in Advances in Neural Information Processing Systems, 2016. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2016/file/8a3363abe792db2d8761d6403605aeb7-Paper.pdf

**[38]** A. Radford, L. Metz, and S. Chintala, "Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks," 2016, arXiv:1511.06434, doi: https://doi.org/10.48550/arXiv.1511.06434

**[39]** S. Jenni and P. Favaro, "On Stabilizing Generative Adversarial Training With Noise," in 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 2019, pp. 12137-12145, doi: https://doi.org/10.1109/CVPR.2019.01242

**[40]** X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks", in Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics, Sardinia, Italy, 2010. [Online]. Available: https://proceedings.mlr.press/v9/glorot10a/glorot10a.pdf

**[41]** I. Gulrajani et al., "Improved Training of Wasserstein GANs," in NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems, Long Beach, California, USA, 2017. [Online]. Available: https://dl.acm.org/doi/10.5555/3295222.3295327

**[42]** I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. MIT Press, 2016.

**[43]** J. Lin, "Divergence measures based on the Shannon entropy," IEEE Trans. Inf. Theory, vol. 37, no. 1, pp. 145-151, 1991, doi: https://doi.org/10.1109/18.61115

**[44]** M. Lucic et al., "Are GANs Created Equal? A Large-Scale Study," in NIPS'18: Proceedings of the 32nd International Conference on Neural Information Processing Systems, Montreal, Canada, 2018. [Online]. Available: https://dl.acm.org/doi/10.5555/3326943.3327008

**[45]** A. Borji, "Pros and cons of GAN evaluation measures," Comput. Vis. Image Underst., vol. 179, pp. 41-65, 2019, doi: https://doi.org/10.1016/j.cviu.2018.10.009

**[46]** S. Arora et al., "Generalization and Equilibrium in Generative Adversarial Nets (GANs)," in ICML'17: Proceedings of the 34th International Conference on Machine Learning - Volume 70, Sydney, Australia, 2017, pp. 224-232. [Online]. Available: https://dl.acm.org/doi/10.5555/3305381.3305405

**[47]** M. S. M. Sajjadi et al., "Assessing Generative Models via Precision and Recall," in NIPS'18: Proceedings of the 32nd International Conference on Neural Information Processing Systems, Montreal, Canada, 2018. [Online]. Available: https://dl.acm.org/doi/10.5555/3327345.3327429

**[48]** S. Nowozin, B. Cseke, and R. Tomioka, "f-GAN: Training Generative Neural Samplers using Variational Divergence Minimization," in NIPS'16: Proceedings of the 30th International Conference on Neural Information Processing Systems, Barcelona, España, 2016. [Online]. Available: https://dl.acm.org/doi/10.5555/3157096.3157127

**[49]** L. Breiman, "Random Forests," Mach. Learn., vol. 45, no. 1, pp. 5-32, 2001, doi: https://doi.org/10.1023/A:1010933404324

**[50]** A. Liaw and M. Wiener, "Classification and Regression by randomForest," R News, vol. 2, pp. 18-22, 2002. [Online]. Available: https://www.r-project.org/doc/Rnews/Rnews_2002-3.pdf

**[51]** T. K. Ho, "Random decision forests," in Proceedings of 3rd International Conference on Document Analysis and Recognition, Montreal, QC, Canada, 1995, pp. 278-282, doi: https://doi.org/10.1109/ICDAR.1995.598994

**[52]** F. Pedregosa et al., "Scikit-learn: Machine Learning in Python," J. Mach. Learn. Res., vol. 12, pp. 2825-2830, 2011. [Online]. Available: https://dl.acm.org/doi/10.5555/1953048.2078195

**[53]** C. Shorten and T. M. Khoshgoftaar, "A survey on Image Data Augmentation for Deep Learning," J. Big Data, vol. 6, no. 1, 2019, art. no. 60, doi: https://doi.org/10.1186/s40537-019-0197-0

**[54]** E. Choi et al., "Generating Multi-label Discrete Patient Records using Generative Adversarial Networks," 2018, arXiv:1703.06490, doi: https://doi.org/10.48550/arXiv.1703.06490

**[55]** A. Goncalves et al., "Generation and evaluation of synthetic patient data," BMC Med. Res. Methodol., vol. 20, no. 1, 2020, art. no. 108, doi: https://doi.org/10.1186/s12874-020-00977-1

**[56]** L. Xu, et al., "Modeling Tabular data using Conditional GAN," 2019, arXiv:1907.00503, doi: https://doi.org/10.48550/arXiv.1907.00503

**RESEARCH ARTICLE**

VOL. 46 | NO. 1 | **JANUARY-APRIL 2025** | PP 1-13

Revista Mexicana de Ingeniería Biomédica

https://dx.doi.org/10.17488/RMIB.46.1.1484

E-LOCATION ID: e1484

# *Allium cepa*: A Natural Enhancer of Wound Closure and Cell Viability in O-Carboxymethyl Chitosan Films

## *Allium cepa*: Potenciador Natural del Cierre de Heridas y Viabilidad Celular en Películas de O-Carboximetil Quitosano

*Laura Valencia-Gómez*[1], *Claudia Rodríguez-González*[1], *Manuel Aguilar Valenzuela*[1], *Faviola Miranda Cabrera* [1], *Kenia Rodríguez López*[1], *Juan Hernández Paz*[1], *Hortensia Reyes Blas*[1], *Jesús Sáenz Villela*[1], *Imelda Olivas Armendáriz*[1]

[1] Universidad Autónoma de Ciudad Juárez, Chihuahua - México

## ABSTRACT

Wound healing and cicatrization after tissue damage are essential for the body's recovery process. Regenerative medicine has been focused on finding biocompatible polymeric materials reinforced with natural active substances, which have healing, anti-inflammatory, and antimicrobial properties. This study aimed to enhance O-carboxymethyl chitosan films with *Allium cepa* extracts, endemic to Chihuahua, Mexico. The extract's properties were analyzed, and the results show the saponin, flavonoid, and terpenoid contents and an antioxidant activity of 70 %. Cicatrization *in vitro* was also studied, and the results show that a 15 mg/mL concentration of the *Allium cepa* extracts improved cellular proliferation and migration in fibroblasts, which was further confirmed in a cellular viability study of the O-carboxymethyl chitosan films loaded with 7 and 20 wt. % of the extract, which showed a higher percentage of cellular viability after three days. The study's results suggest that low extract concentrations can be used as an active ingredient in polymeric biomaterials to aid skin cicatrization.

**KEYWORDS:** *Allium cepa*, O-carboxymethyl chitosan, tissue regeneration, wound healing

## RESUMEN

La cicatrización de heridas y el daño en los tejidos son fundamentales para la recuperación del organismo. La medicina regenerativa se ha centrado en desarrollar materiales poliméricos biocompatibles, enriquecidos con principios activos naturales, que posean propiedades cicatrizantes, antiinflamatorias y antimicrobianas. Este estudio tuvo como objetivo mejorar el desempeño de películas de O-carboximetil quitosano utilizando extractos de *Allium cepa*, una planta endémica de Chihuahua, México. Se analizaron las propiedades del extracto, encontrando un contenido significativo de saponinas, flavonoides y terpenoides, así como una actividad antioxidante del 70 %. La cicatrización *in vitro* indica que una concentración de 15 mg/mL de extracto de *Allium cepa* potencia la proliferación y migración celular en fibroblastos. Esta mejora se confirmó en un estudio de viabilidad celular de las películas de O-carboximetil quitosano cargadas con 7 y 20 wt. % mg del extracto, las cuales mostraron un mayor porcentaje de viabilidad celular después de tres días. Los hallazgos sugieren que concentraciones bajas del extracto pueden ser utilizadas como ingredientes activos en biomateriales poliméricos, contribuyendo así a la cicatrización de la piel.

**PALABRAS CLAVE:** *Allium cepa*, O-carboximetil quitosano, ingeniería tisular, cicatrización

**Corresponding author**

TO: Imelda Olivas Armendáriz
INSTITUTION: UNIVERSIDAD AUTÓNOMA DE CIUDAD JUÁREZ
ADDRESS: AV. DEL CHARRO NO. 450 NTE. COL. PARTIDO ROMERO, CIUDAD JUÁREZ, CHIHUAHUA, C.P. 32310, MÉXICO.
EMAIL: iolivas@uacj.mx

**Received:**

21 October 2024

**Accepted:**

21 January 2025

**Published:**

19 March 2025

## INTRODUCTION

In recent years, there has been a growing interest in discovering new materials that can help heal skin wounds. This interest stems, in part, from the limitations of conventional medical products, which are often composed of inorganic materials that may pose the environment and health risks over time. Consequently, researchers are increasingly exploring alternative solutions, particularly natural bioactive substances with healing, anti-inflammatory, and antimicrobial properties.

Chitosan, a natural cationic polysaccharide derived from deacetylation of chitin, has garnered significant attention as a biomaterial. Its exceptional properties, including biocompatibility, bioadhesiveness, antimicrobial, and cicatrization capabilities, make it an ideal candidate for wound healing and tissue regeneration[1]. Among its derivatives, O-carboxymethyl chitosan (OCMC) has been widely studied due to its water solubility and enhanced performance[2]. For example, in 2002, Chen and his colleagues demonstrated that OCMC promotes fibroblast collagen proliferation and secretion[3]. Additionally, Weng *et al.* reported that hydrogels composed of oxidized dextran and an OCMC derivative accelerated the healing of skin wounds in mice[4]. Natural extracts have also been extensively investigated for their potential therapeutic applications. One such extract, derived from *Allium cepa*, has been used in traditional medicine due to its rich composition of bioactive compounds. *Allium cepa* exhibits diverse pharmacological effects, including antioxidant, anti-inflammatory, antimicrobial, anticarcinogenic, antidiabetic, and cardiovascular benefits[5]. The plant primarily consists of water, lipids, carbohydrates, fibers, potassium, sulfur, phosphorous, calcium, and vitamins such as C, E, B6, folic acid, glutamic acid, arginine, lysine, and leucine. Furthermore, its secondary metabolites (phenolic acids, terpenes, alkaloids, phenylpropanoids, and flavonoids) play a vital role in its medicinal properties[6].

This paper presents a study on the incorporation of *Allium cepa* extract into OCMC films to create a bioactive material for skin regeneration. The innovative combination leverages the bioactive compounds of *Allium cepa* and the inherent properties of OCMC, offering a promising solution for wound healing applications.

## MATERIALS AND METHODS

### *Allium cepa* extract

The bulbs of *Allium cepa* were sourced from Delicias, Chihuahua, Mexico. The bulbs were sectioned and frozen for 24 hours before being lyophilized and ground into a fine powder. A 10 g sample of the powdered *Allium cepa* was mixed with 130 ml of 80% methanol (Methanol Certified ACS, Fisher Chemical) and stirred for two hours at room temperature. The resulting mixture was centrifuged at 3400 rpm for 10 minutes to collect the supernatant. This process was repeated three times to obtain a concentrated supernatant. To remove the solvent, the extract was evaporated at room temperature, yielding the *Allium cepa* extract (ACE). The chemical composition of the extract was analyzed using Fourier-transform infrared spectroscopy (FT-IR). Samples were prepared by cutting and drying the material and analyzed at room temperature in the spectral range of 4000 to 500 $cm^{-1}$. Each spectrum was recorded using 100 scans at a resolution of 16 $cm^{-1}$.

### Metabolite quantitative identification

The Salkowski test was conducted to identify the presence of saponins. In a Pyrex test tube, 5 mL of ACE was combined with 2 mL of chloroform and stirred for 15 seconds. Subsequently, 3 mL of sulfuric acid was added, and the mixture was shaken for 10 minutes. The presence of saponins was indicated by the formation of foam on the surface, accompanied by a reddish-brown coloration. The saponin content was classified as follows, abundant with a foam height >14 mm; moderate content, foam height between 10 and 14 mm; low content, foam height <10 mm. Flavonoid identification involved mixing 3 mL of ACE with 15 drops of a 50% NaOH solution. The mixture was left to stand until a yellow coloration appeared, after which concentrated HCl was added dropwise until the solution became clear. The intensity of the yellow color in the alkaline medium reflected the flavonoid concentration. An intense yellow color indicates a high concentration of flavonoids, while a fainter color indicates a lower concentration. When HCl is added at high concentrations, a rapid and noticeable disappearance of yellow color further confirmed the flavonoid presence. Alkaloid identification was performed using Dragendorff reagent. For each test, 3 mL of liquid extract was used, with the test conducted in triplicate. The reagent was added dropwise until a precipitate formed. The color of precipitate indicated alkaloid concentration: a reddish precipitate indicated a high concentration, while a cream-colored precipitate indicated a low or moderate concentration. The criteria for the qualitative assessment of these metabolites are summarized in Table 1[7].

TABLE 1. **Defined Criteria for Qualitative Metabolite Testing.**

| Criteria | Yardstick |
|---|---|
| Abundant presence | +++ |
| Moderate presence | ++ |
| Low presence | + |
| Absence | - |

## Quantification of antioxidant activity

A 150 μmol/L methanolic dilution of DPPH was prepared. Aliquots of 100 μL of methanolic ACE at a 75 mg/mL concentration were collected in triplicate, and 200 μL of the prepared DPPH dilution was added to each. The mixture was incubated in the dark at room temperature for 30 minutes. After the incubation time, the absorbance was measured at 517 nm using a Benchmark Plus Microplate Reader spectrophotometer at various time intervals. A blank was prepared by mixing an equal volume of DPPH solution with methanol.

The antioxidant activity of the ACE was assessed by calculating the percentage of DPPH radical inhibition at a concentration, using the following equation 1:

$$\%\ \text{Inhibition} = \frac{\text{Ac} - \text{As}}{\text{Ac}} \times 100 \tag{1}$$

Ac= Control's absorbance t=0

As= Sample's absorbance t=30 minutes

## Cell viability

ACE at concentrations of 15, 50, and 75 mg/mL was cultured with 50,000 3T3 murine fibroblast cells in 24-well plates, and 3 mL of DMEM containing 10 % fetal bovine serum (FBS) and 3 % antibiotics was added. The culture was incubated at 37°C and 5% $CO_2$ for 24 and 72 hours. Afterward, the medium was removed from the wells, and 200 μL of DMEM and 50 μL of MTT solution (5 mg/mL in 1X PBS) were added. The plate was incubated at 37 °C and 5 % $CO_2$ for 1 hour. After incubation, the solution was removed from the wells, and 50 μL of Dimethyl Sulfoxide (DMSO) was added to each well to dissolve the formazan crystals synthesized by the cells. The absorbance readings were taken using a Benchmark Plus Microplate Reader spectrophotometer at 570 nm. The tests were performed in triplicate, and images were captured during the evaluation. The process was conducted using a Carl Zeiss Optical Microscope. The percentage of cell viability is given by Equation 2:

$$\text{Cell viability } (\%) = \frac{\text{Absorbance of sample}}{\text{Absorbance of control}} \times 100 \tag{2}$$

## Cicatrization activity *in vitro*

50,000 3T3 fibroblasts were added to 6-well plates with 3 ml of DMEM containing 2 % FBS. The plates were then incubated at 37 $^{0}$C with 5 % CO2 until the cells attached. Once the cells covered the entire surface, a wound was made under sterile conditions with a micropipette. Next, 3 ml of DMEM mixed with ACE at varying concentrations was added to each condition. The plates were then incubated at 37 $^{0}$C with 5% CO2. Images were taken at 0, 2, 4, 24, and 26 hours using a Zeiss Axio Vert A1 microscope. This entire process was performed in triplicate for extracts. The concentration of mitomycin used was 10 μg/mL to prevent cell migration. Wound closure was calculated using the following equation 3:

$$\text{Percentage of wound closure } (\%) = \frac{\text{At0} - \text{At}}{\text{At0}} \times 100 \tag{3}$$

At0 represents the wound length at time zero, while At represents the wound length later in the analysis.

Cell adhesion was observed and assessed using confocal microscopy. After culturing the cells for up to 72 hours, the cell nuclei were stained blue by incubating them in a PBS solution containing 4′,6-Diamidine-2′-phenylindole dihydrochloride (DAPI) (Invitrogen™, Thermo Fisher Scientific, Waltham, MA, USA). Samples were incubated in the staining solution for 1 hour and then observed using fluorescence microscopy (Axio Vert. A1, Carl Zeiss), with the ZEN Blue software used to capture the images.

## OCMC synthesis

OCMC was produced using Valencia-Gomez's method[1]. Firstly, 50 mL of isopropanol (Sigma-Aldrich) was placed under continuous stirring at room temperature. Then, 5 g of chitosan (Medium Molecular Weight Sigma-Aldrich) was added and stirred at a constant speed for 2 hours at room temperature. After that, 20 mL of a 40 % w/v solution of sodium hydroxide NaOH (J.T. Baker) was added every 15 minutes to 80 mL. The mixture was stirred for 30 more minutes. Then, 100 mL of a monochloroacetic acid solution (Sigma-Aldrich, 99 %) was added, previously prepared by dissolving 25.5 g of monochloroacetic acid in 100 mL of isopropanol. The reaction continued for another 24 hours at room temperature and constant stirring. Then 200 mL of ethyl alcohol was added to the mixture, and the resulting solid was filtered out. The solid was then washed six times with methanol (Methanol Certified ACS, Fisher Chemical), using 200 mL portions between each wash to eliminate the NaCl salt formed as a byproduct of the reaction. The resulting OCMC was placed inside a laminar flow hood at room temperature until completely dry.

### OCMC/ACE Film

Two grams of OCMC were mixed in 100 milliliters of deionized water (w/v) while stirring constantly at room temperature for 20 minutes. Following this, 1 milliliter of glycerin (Sigma-Aldrich) was added under constant stirring for another 10 minutes. Finally, 2 milligrams of genipin (Challenge Bioproducts Co., Ltd.) were added. Immediately, 25 milliliters of the solution were poured into each petri dish and left at room temperature until the solvent had evaporated entirely. For the OCMC films, the ACE was added after the OCMC had been added to the solution. The freeze-dried extract concentrations used for functionalization were 7 and 20 % by weight (Table 2). A film containing 20 wt. % of quercetin was used as a control, as it is one of the most relevant flavonoids in ACE.

TABLE 2. **Concentrations of film components**

| Film | OCMC (wt. %) | ACE (wt. %) | Quercetin (wt.%) |
|---|---|---|---|
| OCMC | 100 | 0 | 0 |
| OCMC/ACE/7 | 93 | 7 | 0 |
| OCMC/ACE/20 | 80 | 20 | 0 |
| OCMC/Q | 80 | 0 | 20 |

### Film characterization: FT-IR and SEM

The FT-IR characterization of the films was conducted at room temperature in the range of 4000 to 500 $cm^{-1}$. The samples were cut and dried before analysis using a Thermo Nicolet 6700 FT-IR spectrometer with germanium glass, in ATR mode. All spectra were recorded using 100 scans and 16 cm-1. SEM (Scanning Electron Microscopy) is a valuable technique for surface analysis. The test utilized a Hitachi Field Emission Scanning Electron Microscope (FESEM) with an accelerating voltage of 5 KV. The sample size used for this technique was 2 X 2 cm.

### OCMC/ACE film cell viability

Films were cut into 0.5 cm × 0.5 cm squares and sterilized under UV light for 15 minutes. The sterile films were placed in 24-well plates, and 50,000 3T3 fibroblasts were added to each well, along with 3 mL of DMEM containing antibiotics and FBS. The plates were incubated at 37 °C with 5 % $CO_2$. After 24 and 72 hours of incubation, the medium was removed, and 200 μL of DMEM and 50 μL of MTT solution were added. The plate was incubated for 1 hour, after which the solution was removed, and 50 μL of DMSO was added to dissolve the formazan crystals. Absorbance was measured at 570 nm using a Benchmark Plus microplate reader. The analysis was conducted in triplicate, using wells containing only cells without material as control. The cell viability percentage was calculated using Equation 2. The cells adhered to the films were fixed with a 3 % glutaraldehyde solution. The films were washed three times with deionized water, and hematoxylin was added for 6.5 minutes. After this time, the hematoxylin was completely removed, and the films were washed three times with deionized water, followed by acid alcohol, and finally washed with deionized water again. Following this, eosin was applied to completely cover the films, and the samples were left to rest for 6.5 minutes. Additional washes were then performed with 70 %, 96 %, and 100 % ethanol solutions. The fibroblasts were observed using a Zeiss AXIO Vert.A1 microscope.

## RESULTS AND DISCUSSION

### *Allium cepa* extract characterization: FT-IR

Figure 1 illustrates the infrared spectrum of the obtained extract. The bands at 3300 and 2928 $cm^{-1}$ are attributed to the presence of O-H and C-H groups, respectively[8]. Additionally, a band appears at 1618 $cm^{-1}$, which may correspond to the C=C stretching of the aromatic ring in combination with the C-C stretching attributed to the phenyl group. This suggests the presence of high levels of polyphenolic components[9]. The band at 1648 $cm^{-1}$ is attributed to amide I, while the bands at 1405 and 1740 $cm^{-1}$ are assigned to the asymmetric deformation of $CH_3$ and the stretching of the carbonyl group (C=O), respectively. The band at 1340 $cm^{-1}$ is due to C-O stretching in combination with phenyl ring stretching, while the band at 1255 $cm^{-1}$ is assigned to amide III (random coil conformation) in proteins[10]. The stretching band at 1105 $cm^{-1}$ is attributed to C-OH groups, while the bands at 1025, 985, and 868 $cm^{-1}$ are related to -$CH_2OH$, $OCH_3$, and CH groups, respectively[10]. These results suggest that ACE and the powder contain phenolic compounds, alkaloids, and terpenoids.

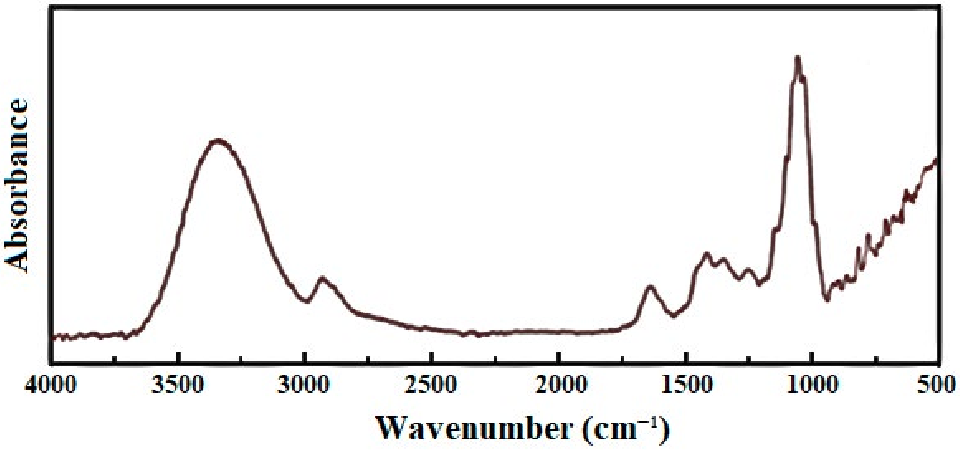


**FIGURE 1.** Infrared spectrum of ACE.

### Qualitative analysis to identify saponins, flavonoids, and alkaloids in the ACE

Previous studies have indicated that certain plant families, such as monocotyledons including *Dioscoreaceae, Agavaceae*, and *Liliaceae* (including *Allium* plants), contain saponins[11], which are high-molecular-weight glycosides consisting of a sugar moiety attached to a triterpene or steroidal aglycone. Saponins have properties that are beneficial for skin regeneration, as they can act as antimicrobial, healing, and anti-inflammatory agents[11]. Table 3 summarizes the identification test results for saponins, alkaloids, and flavonoids. The qualitative test to identify alkaloids in the ACE yielded positive results, indicating a high concentration of alkaloids. R. Pérez *et al.*[12] have mentioned various benefits of alkaloids, including their considerable antioxidant action and antimicrobial, cytotoxic, and anti-inflammatory activity.

**TABLE 3.** Results of Metabolites in ACE.

| Metabolite | ACE |
|---|---|
| Saponins | ++ |
| Flavonoids | +++ |
| Alkaloids | +++ |

### Antioxidant activity by free radical DPPH

The ACE was studied for its antioxidant effect using the 2,2-diphenyl-1-picrylhydrazyl (DPPH) free radical method. This method uses a stable free radical with a maximum absorbance of 517 nm. Some plant derivatives can act as hydrogen donors, preventing or delaying the development of degenerative diseases. Such derivatives include tannins, flavonoids, lignans, quinones, phenolic acids, catechins, and anthocyanins[13].

During the test, the degree of discoloration is an indication of the antioxidant potential of the extract. This study examined the free radical scavenging potential of the ACE at different time intervals ranging from 1 to 5 hours. The final result of the test is depicted in Figure 2. According to Figure 2, the ACE exhibits a high antioxidant effect from the first hour, with an antioxidant activity percentage of 66.02 %. This percentage increases significantly in the fifth hour when the percentage of antioxidant activity is 70.86 %.

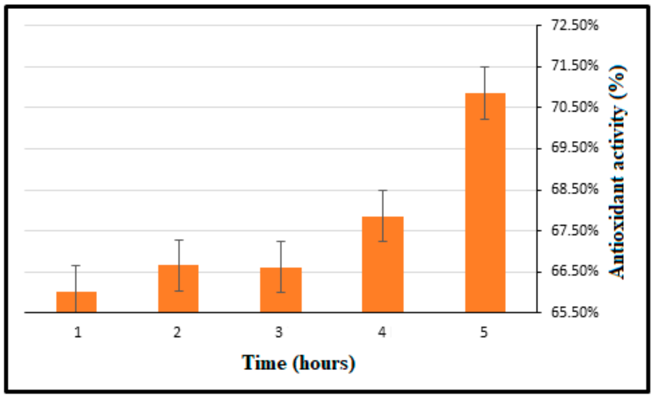


FIGURE 2. **Percentage of antioxidant activity of ACE at a concentration of 75 mg/L.**

## Cell viability and *in vitro* wound closure of ACE

Figure 3 shows cell proliferation at different concentrations of ACE seeded with fibroblasts on the first and third days of culture. No significant difference in cell numbers was observed between these periods of time across the different concentrations of ACE. However, cell viability percentages greater than 100 % were observed. The cell viability study revealed that the different concentrations of extract maintained a viability percentage above 70 %, classifying it as non-toxic according to ISO 10993-5:2009 - Biological Evaluation of Medical Devices.

Figure 3 (II) presents images obtained after 72 hours of culture of fibroblast cells seeded with different concentrations of ACE. The images, captured using fluorescent staining (DAPI), show the nuclear material of viable cells in blue[14]. DAPI-stained nuclei display a well-defined outline, with shapes ranging from rounded to elongated, which is indicative of healthy cells, including those in the process of mitosis[15].

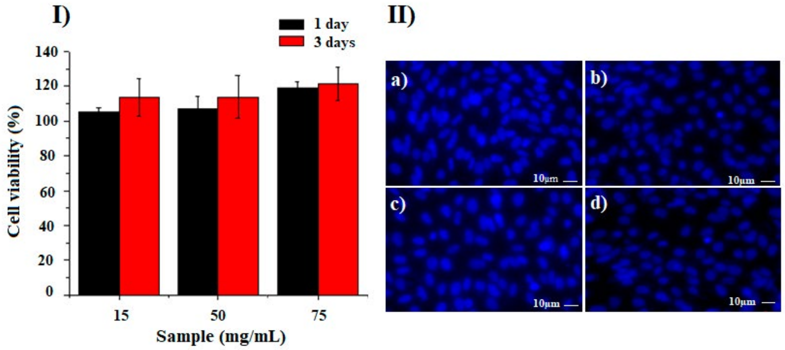


FIGURE 3. **Cell viability percentage at 24 and 72 hours, and II) Cell morphology at 72 hours of a) 15 mg/mL of ACE, b) 50 mg/mL of ACE, and c) 75 mg/mL of ACE seeded with fibroblastic cells.**

The effect of the extract on cell healing was evaluated *in vitro* using the wound healing assay. Figure 4 shows that the extract had no significant effect on wound closure at 2 hours, but after 4 hours, cell numbers were increased in the sample containing 15 mg/mL compared to the control. The wound was completely closed at 24 hours. While the control group showed noticeable wound closure after 24 hours, it still had a partially open wound after 26 hours. The percentage decrease in the wound area was calculated for each concentration and incubation time. The calculations showed that at 4 hours, the closure percentage was very similar in the presence of 15 mg/mL of extract (25 %) and in the control, which had no extract (18 %). At 24 hours, the highest closure percentage was observed with the sample containing 15 mg/mL of extract (90 %), compared to 50 and 75 mg/mL. The wound was first completely closed in the presence of the extract at 15 mg/mL, while in the absence of the extract, full closure occurred at 26 hours. These results demonstrate that the extract is not toxic to cells at concentrations up to 75 mg/mL.

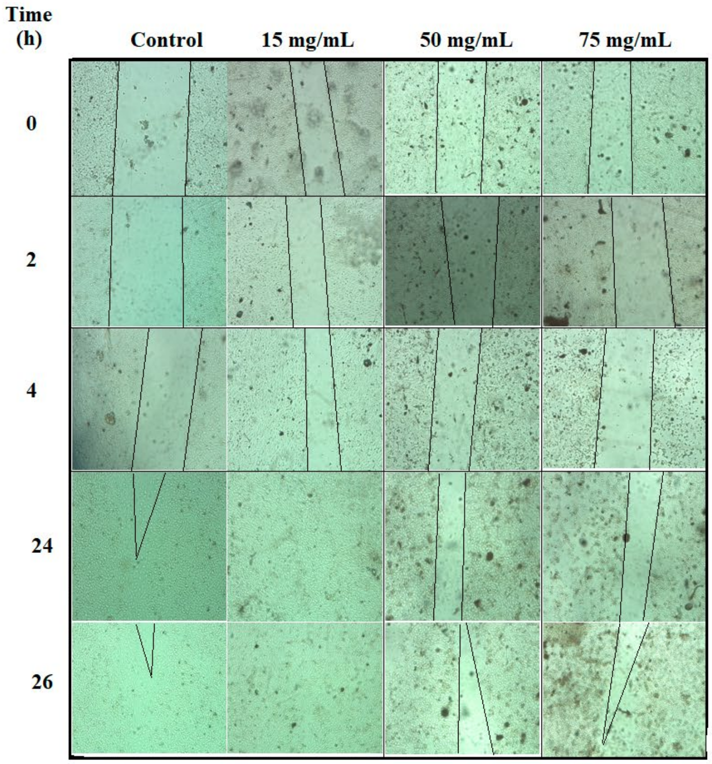


**FIGURE 4. Tear test with fibroblasts for 26 hours, a) control, b) 15 mg/mL of ACE, c) 50 mg/mL of ACE, and d) 75 mg/mL of ACE.**

Significant amounts of saponins, flavonoids, carbohydrates, and, to a lesser extent, alkaloids were found in ACE. Various studies have shown that flavonoids and phenolic acids can facilitate different stages of wound healing by stimulating collagen synthesis, cell proliferation, and angiogenesis[16][17][18]. This is due to their redox potential, which allows them to act as oxygen scavengers and reducing agents, thus reducing free radical stress and preventing oxidative damage, thereby offering the therapeutic potential to accelerate the healing process. Additionally, they modulate signaling pathways, activating endogenous mechanisms that increase antioxidant enzyme levels and progressively promote fibroblast growth, leading to faster wound closure.

During the healing process, complex cellular and molecular mechanisms regulate the inflammatory, proliferative, and maturation phases. Therefore, accelerated healing is attributed to the synergistic action of the various bioactive compounds present in extracts such as ACE[19]. Thus, the results obtained in this study indicate that the use of a concentration of 15 mg/mL, may reflect the optimum combined effect of all the molecules present in ACE resulting in a faster wound closure process, likely due to its antioxidant, antibacterial, and anti-inflammatory properties. However, results obtained at higher concentrations, as reported in other studies, could be the result of an imbalance between free radical generation and antioxidant activity, leading to excessive oxidative stress in cells or the inhibition of signaling pathways involved in cell migration, factors that could slow or inhibit wound closure[17]. Additionally, other studies have reported cytotoxic effects of high saponin concentrations in certain cell lines[16].

### Chemical characterization of O-carboxymethyl chitosan films with ACE

In Figure 5, the spectrogram displays bands corresponding to the stretching vibrations of O-H groups around 3350, at 2928 cm-1 to C-H stretching. Other bands can be seen around 1600 and 1410 $cm^{-1}$, which can be attributed to the asymmetric and symmetric stretching of C=O groups. Bands around 1320, 1050 $cm^{-1}$, and 850 $cm^{-1}$ are also observed, related to C-H, C-O, and C-H groups, respectively[1][9]. No significant differences were observed between the spectra, regardless of the ACE concentration. This may be because the characteristic bands of the extract overlap with the absorption bands present in the OCMC.

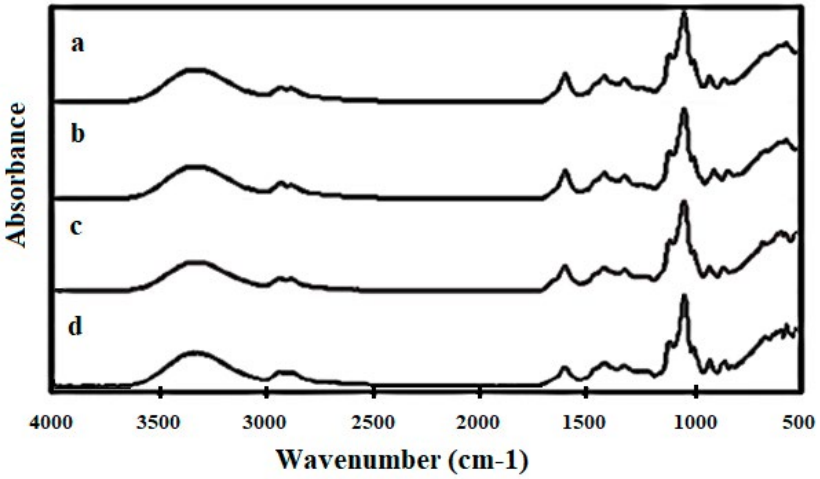


**FIGURE 5. Spectrogram obtained for the films of a) OCMC, b) OCMC/ACE/7, c) OCMC/ACE/20, and d) OCMC/Q.**

### Morphological surface analysis by SEM

Four SEM images were taken using a scanning electron microscope as shown in Figure 6. The surface of different concentrations of ACE and commercial quercetin were observed. The results demonstrate that the OCMC films without ACE and those containing 7 wt. % of ACE and 20 wt.% of quercetin had a smooth surface without any irregularities. However, the film containing 20 wt. % of ACE displayed an irregular surface with extract particles.

### Cell viability activity of films

Figure 7 shows the cell viability results for four synthesized films. The t-student test shows a statistically significant difference in cell viability after 24 hours between the OCMC film without ACE and the OCMC film with 20 wt. % of ACE and 20 wt. % of commercial quercetin. After 72 hours, there is a statistically significant difference in cell viability between the OCMC film with commercial quercetin and the other three films. Based on these results, it can be inferred that ACE or commercial quercetin improves cell viability in the first 24 hours. Still, after 72 hours, there is no significant difference in cell viability between the OCMC film with or without ACE. However, the presence of quercetin at 72 hours does not have a positive effect on cell viability. It is important to note that none of the films were found to be cytotoxic.

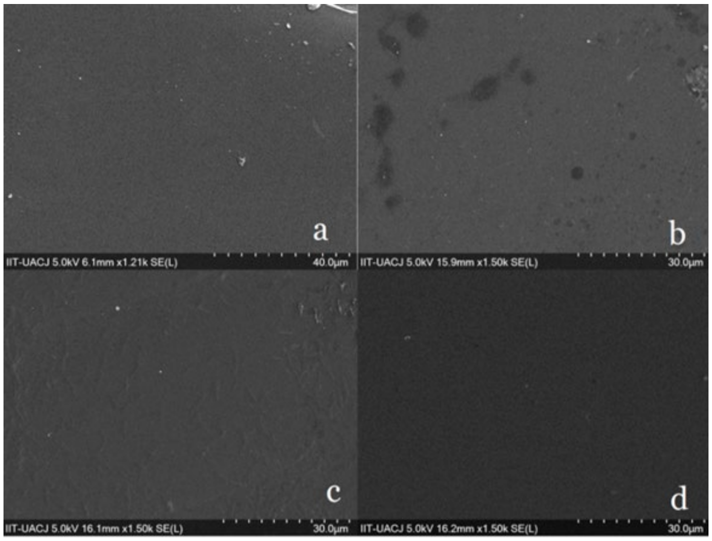


**FIGURE 6. SEM images were obtained from the surface morphology of the films of a) OCMC, b) OCMC/ACE/7, c) OCMC/ACE/20, and c) OCMC/Q.**

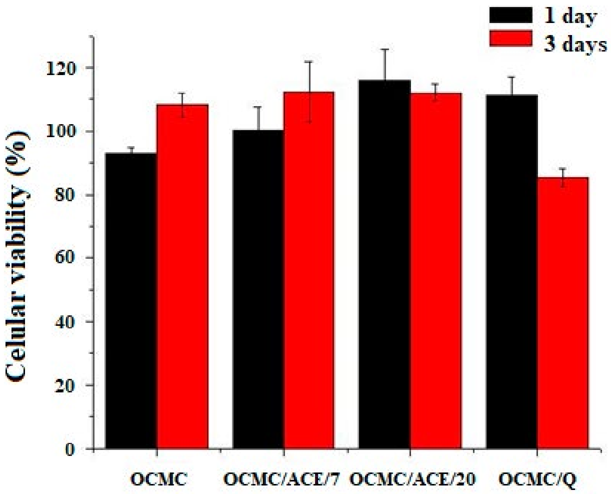


**FIGURE 7. The cytotoxicity of films synthesized from OCMC and ACE or commercial quercetin in 3T3 fibroblasts was measured through the MTT assay.**

Eosin and hematoxylin were used to stain cultured fibroblasts and observe their morphology. Figure 8 displays the cells cultured with synthesized film samples after 72 hours. The images show cells with an elongated shape, exhibiting pink cytoplasm and dark purple nuclei in all the samples. Uniform fibroblast growth is visible in the OCMC film (Figure 8a) and the OCMC film with 7 wt. % of ACE (Figure 8b). However, in the OCMC film with 20 wt .% of ACE (Figure 8c) and the OCMC film with 20 wt. % of commercial quercetin (Figure 8c), blank spots can be observed where cells do not grow. Furthermore, Figure 8b shows many junctions between cells, indicating that small concentrations of ACE can benefit cell proliferation.

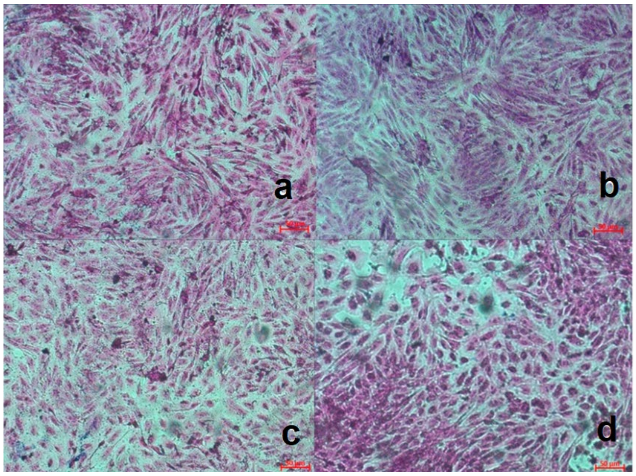


FIGURE 8. **Observation of fibroblasts under an optical microscope after 72 hours of incubation. The fibroblasts were stained with hematoxylin and eosin and were seeded in four different conditions: a) OCMC, b) OCMC/ACE/7, c) OCMC/ACE/20, and d) OCMC/Q.**

Previous studies have reported that the cellular viability of fibroblasts depends on both the incubation time and the concentration of EPA or quercetin. Low or moderate concentrations can promote fibroblast proliferation, while high concentrations may inhibit it. An increase in quercetin concentration is associated with a decrease in mTOR protein levels, a protein that plays a crucial role in cellular biological processes such as proliferation and apoptosis. Additionally, quercetin regulates the biological behavior of fibroblasts by inhibiting the expression of the migration-associated protein α-SMA and the production of collagen types I and III[20][21].

## CONCLUSIONS

The results of this study demonstrate that *Allium cepa* extract at a concentration of 15 mg/mL significantly accelerated wound healing by promoting wound closure within 24 hours and enhancing fibroblast proliferation and migration during the assay. In contrast, higher concentrations of the extract (50 and 75 mg/mL) required more than 26 hours to achieve wound closure. The cell viability study revealed that the extract maintained a viability percentage greater than 70 %, classifying it as non-toxic. Additionally, films containing 7 % and 20 % (wt.) of the extract exhibited cellular viability exceeding 100 %, indicating enhanced cell growth and activity. These findings highlight the potential of *Allium cepa* extract and its polymeric films as promising materials for supporting tissue regeneration and wound healing applications.

## AUTHOR CONTRIBUTIONS

L. E. V. G. conceptualization, methodology, validation, data curation, and writing—original draft; C. A. R.-G. methodology, investigation, and validation; M. A. V. methodology and investigation; F. M. C. methodology and investigation; K. R. L. methodology and investigation; J. F. H.-P. conceptualization, validation, and formal analysis; H. R. B. methodology and visualization; J. S. V. data curation and review & editing; I. O.-A. project administration, conceptualization, methodology, validation, resources, supervision, writing—review & editing.

## REFERENCES

**[1]** L.-E. Valencia-Gómez, et al., “Characterization and evaluation of a novel O-carboxymethyl chitosan films with Mimosa tenuiflora extract for skin regeneration and wound healing,” J. Bioact. Compat. Polym., vol. 35, no. 1, pp. 39-56, 2020, doi: https://doi.org/10.1177/0883911519885976

**[2]** L. E. Valencia-Gómez, et al., “Comparative Study of the Antibacterial, Biodegradable, and Biocompatibility Properties of Composite and Bi-Layer Films of Chitosan/Gelatin Coated with Silver Particles,” Materials, vol. 16, no. 8, 2023, art. no. 3000, doi: https://doi.org/10.3390/ma16083000

**[3]** X.-G. Chen, Z. Wang, W.-S. Liu, H.-J. Park, “The effect of carboxymethyl-chitosan on proliferation and collagen secretion for normal and keloid skin fibroblasts,” Biomaterials, vol. 23, no. 23, pp. 4609-4614, 2002, doi: https://doi.org/10.1016/S0142-9612(02)00207-7

**[4]** L. Weng, A. Romanov, J. Rooney, W. Chen, “Non-cytotoxic, in situ fellable hydrogels composed of N-carboxymethyl chitin and carboxymethyl chitosan films for potential use in wound care application,” Biomaterials, vol. 29, no. 29, pp. 3905-3913, 2008, doi: https://doi.org/10.1016/j.biomaterials.2008.06.025

**[5]** S. Sanyal, P. K. Vemula, S. Law, “Investigating the therapeutic potential of Allium cepa extract in combating pesticide exposure induced ocular damage,” Exp. Eye Res., vol. 42, 2024, art. no. 109816, doi: https://doi.org/10.1016/j.exer.2024.109816

**[6]** H.-S. Lee, et al., “Anti-inflammatory effects of Allium cepa L. peel extracts via inhibition of JAK-STAT pathway in LPS-stimulated RAW264.7 cells,” J. Ethnopharmacol., vol. 317, 2023, art. no. 116851, doi: https://doi.org/10.1016/j.jep.2023.116851

**[7]** D. E. Garcia, F. Ojeda, “Evaluación de los principales factores que influyen en la composición fitoquímica de Morus alba (Linn.). I. Análisis cualitativo de metabolitos secundarios,” Pastos y Forrajes, vol. 26, no. 4, pp. 335-343, 2003.

**[8]** X. Lu, et al., “Determination of total phenolic content and antioxidant capacity of onion (Allium cepa) and shallot (Allium oschaninii) using infrared spectroscopy,” Food Chem., vol. 129, no. 2, pp. 637-644, 2011, doi: https://doi.org/10.1016/j.foodchem.2011.04.105

**[9]** P. Sharma, et al., “Green Synthesis of Silver Nanoparticle Capped with Allium cepa and Their Catalytic Reduction of Textile Dyes: An Ecofriendly Approach,” J. Environ. Polym. Degrad., vol. 26, no. 5, pp. 1795-1803, 2017, doi: https://doi.org/10.1007/s10924-017-1081-7

**[10]** X. Lu, “Determination of Antioxidant Content and Antioxidant Activity in Foods using Infrared Spectroscopy and Chemometrics: A Review,” Crit. Rev. Food. Sci. Nutr., vol. 52, no. 10, pp. 853-875, 2012, doi: https://doi.org/10.1080/10408398.2010.511322

**[11]** A. Bhattacharya, “High-temperature stress and metabolism of secondary metabolites in plants. Effect of High Temperature on Crop Productivity and Metabolism of Macro Molecules,” in Effect of High Temperature on Crop Productivity and Metabolism of Macro Molecules, Cambrige, United State: Elsevier, 2019, ch. 5, pp. 391-484, doi: https://doi.org/10.1016/B978-0-12-817562-0.00005-7

**[12]** R. M. Pérez, et al., “Actividad antioxidante de los alcaloides de Bocconia arborea. Estudio sobre seis métodos de análisis,” Ars Pharm., vol. 44, no. 1, pp. 5-21, 2002, doi: https://revistaseug.ugr.es/index.php/ars/article/view/5122

**[13]** S. Baliyan, et al., “Determination of Antioxidants by DPPH Radical Scavenging Activity and Quantitative Phytochemical Analysis of Ficus religiosa,” Molecules, vol. 27, no. 4, 2022, art. no. 1326, doi: https://doi.org/10.3390/molecules27041326

**[14]** M.A. Rojas-Yañez, et al., “Composite scaffolds of chitosan/polycaprolactone functionalized with protein of Mytilus californiensis for bone tissue regeneration,” AIMS Mater. Sci., vol. 9, no. 3, pp. 344-358, 2022, doi: https://doi.org/10.3934/matersci.2022021

**[15]** N. Arrieta-Sandoval, et al., “Effect of Ag2S-BSA nanoparticle size on 3T3 fibroblast cell line cytotoxicity,” J. Nanopart. Res., vol. 22, 2020, art. no. 106, doi: https://doi.org/10.1007/s11051-020-04834-6

**[16]** A. Areesanan, et al., “Pharmacological in vitro profiling of Buddleja officinalis flower extracts in the context of dry eye disease,” Biomed. Pharmacother., vol. 181, 2024, art. no. 117685, doi: https://doi.org/10.1016/j.biopha.2024.117685

**[17]** E. Yadav, et al., “Antioxidant and anti-inflammatory properties of Prosopis cineraria based phenolic rich ointment in wound healing,” Biomed. Pharmacother., vol. 108, pp. 1572-1583, 2018, doi: https://doi.org/10.1016/j.biopha.2018.09.180

**[18]** A. Joshi, V. K. Joshi, D. Pandey, S. Hemalatha, “Systematic investigation of ethanolic extract from Leea macrophylla: implications in wound healing,” J. Ethnopharmacol., vol. 191, pp. 95-106, 2016, doi: http://dx.doi.org/10.1016/j.jep.2016.06.034

**[19]** G. Hernández-Pasteur, et al., “Evaluación in vitro de la Actividad Cicatrizante y Antibacteriana de Extractos de Buddleja cordata Kunth y Vismia baccifera (L.) Triana & Planch,” Rev. fitotec. mex., vol. 42, no. 2, pp. 93-99, 2019. [Online]. Available: https://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S0187-73802019000200093

**[20]** Y. Cao, et al., “Quercetin inhibits fibroblasts proliferation and reduces surgery-induced epidural fibrosis via the autophagy-mediated PI3K/Akt/mTOR pathway,” Bioengineered, vol. 13, no. 4, pp. 9973-9986, 2022, doi: https://doi.org/10.1080/21655979.2022.2062530

**[21]** C. Chittasupho, et al., “Effects of Quercetin and Curcumin Combination on Antibacterial, Antioxidant, In Vitro Wound Healing and Migration of Human Dermal Fibroblast Cells,” Int. J. Mol. Sci., vol. 23, no. 1, 2022, art. no. 142, doi: https://doi.org/10.3390/ijms23010142

**REVIEW ARTICLE**
VOL. 46 | NO. 1 | **JANUARY-APRIL 2025** | PP 1-18

Revista Mexicana de Ingeniería Biomédica

https://dx.doi.org/10.17488/RMIB.46.1.1472

E-LOCATION ID: e1472

# Emerging Technologies as a Support for Proprioceptive Rehabilitation: a Scoping Review

## Tecnologías Emergentes como Apoyo en la Rehabilitación Propioceptiva: una Revisión del Alcance

*Uziel Trujillo Colón*[1], *José Luis Hernández Hernández*[1], *Eduardo De La Cruz Gámez*[2], *Rayma Ireri Maldonado Astudillo*[1], *Ricardo Salazar*[3]

[1] Universidad Autónoma de Guerrero, Guerrero - México
[2] Tecnológico Nacional de México/Instituto Tecnológico de Acapulco, Guerrero -México
[3] Universidad Autónoma de Guerrero / CONAHCYT, Guerrero - México

## ABSTRACT

Proprioceptive training encompasses interventions aimed at enhancing proprioceptive function to improve motor function performance. Three types of interventions are considered: Movement Training (MT); Somatosensory Stimulation Training (SST), and Force Reproduction Training (FRT). This study analyzes the potential of emerging technologies, such as exoskeletons, mechanical devices, Artificial Intelligence (AI), Virtual Reality (VR), the Internet of Things (IoT), and sensors, highlighting their application in proprioceptive therapies, with particular emphasis on MT, SST, and FRT. A total of 107 articles published in scientific journals were reviewed, of which 30 complied with inclusion criteria: 1) Implementation of proprioceptive intervention therapy; 2) use of technology; 3) publication after 2019, and 4) written in the English language. Of the studies analyzed, 43 % employed AI, indicating its increasing adoption, while IoT was the least utilized technology, with only 3 %. It is concluded that emerging technologies plays a crucial role in proprioceptive rehabilitation by enabling the analysis of data before and after surgical procedures, real-time pattern assessment, and the classification of sensory signals. Moreover, it offers alternatives to traditional measurement methods.

**KEYWORDS:** convolutional neural networks, exoskeletons, mechanical devices, therapies

## RESUMEN

El entrenamiento propioceptivo representa cualquier intervención de la función propioceptiva que ayude a mejorar el desempeño de la función motora. Se consideran tres tipos de intervenciones: Entrenamiento de Movimiento (EM); Entrenamiento de Estimulación Somatosensorial (EES) y Entrenamiento de Reproducción de Fuerza (ERF). Este estudio analiza el alcance de las tecnologías emergentes, como los exoesqueletos, dispositivos mecánicos, Inteligencia Artificial (IA), Realidad Virtual (VR), el Internet de las Cosas (IdC) y sensores, destacando su aplicación en las terapias propioceptivas, con énfasis en el EM, EES, y ERF. Se revisaron 107 artículos publicados en revistas científicas, de los cuales 30 cumplieron los criterios de inclusión: 1) Implementación de terapia de intervención propioceptiva; 2) uso de tecnología; 3) publicación posterior al año 2019, y 4) redacción en inglés. De los estudios analizados, el 43 % empleó IA, mostrando su creciente adopción, mientras que el IdC fue la tecnología menos utilizada, con un 3 %. Se concluye que las tecnologías emergentes son fundamentales en la rehabilitación propioceptiva, al permitir el análisis de información antes y después de procedimientos quirúrgicos, la evaluación de patrones en tiempo real, y la clasificación de señales sensoriales. Además, ofrecen alternativas efectivas frente a métodos tradiciones de medición.

**PALABRAS CLAVE:** dispositivos mecánicos, exoesqueletos, redes neuronales convolucionales, terapias

**Corresponding author**

TO: José Luis Hernández Hernández
INSTITUTION:UNIVERSIDAD AUTÓNOMA DE GUERRERO
ADDRESS: TERCERA CALLE NO. 81, COL. VILLA MODERNA, CHILPANCINGO GUERRERO, C.P. 39070, MÉXICO.
EMAIL: joseluis.hernandez@itchilpancingo.edu.mx

**Received:**

30 September 2024

**Accepted:**

14 February 2025

**Published:**

20 March 2025

## INTRODUCTION

Proprioception depends on sensory signals from muscle spindles, skin, and joint receptors. It refers to the body's ability to perceive limb position, detect movement (whether passive or active), and recognize the forces exerted[1]. Several factors can negatively the impact proprioceptive function, including typical aging[2], sport injuries[3], motor disorders[4], strokes[5], or numerous neurological and orthopedic conditions such as Parkinson's disease[6], focal dystonia[7], and sensory neuropathies[8]. Consequently, individuals with compromised proprioception often experience a deterioration in their quality of life. In this context and considering the critical role of proprioception in motor control, it has been proposed that therapies targeting motor function restoration should emphasize the enhancement of proprioceptive abilities. Thus, physical therapy rehabilitation is regarded as the most effective approach for this purpose[9].

In 2021, an analysis explored the potential of integrated Virtual Reality (VR) into physical therapies[10]. In the year 2022, a review was conducted on strokes, their rehabilitation, and the role of robotic technology, including its adoption and the barriers encountered during its implementation[11].

Currently, a wide range of therapies is available to enhance the proprioceptive system, and these should be administered by specialized professionals, such as Kinesiotherapists, Physiotherapists, or Traumatologists. Treatment modeling indicates that clinical assessment and data acquisition are followed by interpretation, which leads to diagnosis and prognosis. For interventions, it is sometimes necessary employ specialized medical instruments, such as goniometers, inclinometers, and scales, during evaluations[1]. Based on the latter, this article reviews the literature on various methods of proprioceptive intervention and the different technologies that can support rehabilitation therapies. It also examines the application of emerging technologies and identifies potential niches where these innovations could be utilized in the near future.

## MATERIALS AND METHODS

This section outlines the process of gathering and analyzing relevant literature. A comprehensive search was conducted in indexed scientific journals using specific parameters to ensure the quality and relevance of the selected articles. Following the collection of articles, a detailed analysis and classification were performed based on the following key aspects: 1) Proprioception therapies applied in the studies; 2) technologies utilized in the intervention and evaluation, and 3) anatomical areas targeted for rehabilitation.

### Literature search

The literature search for related works primarily involved scientific articles published in biomedical journals indexes, including PubMed, Google Scholar, MDPI, Science Direct, Web of Science, and Scopus. The parameters used in the searches included proprioception, physical rehabilitation therapy, movement training interventions, strength reproduction, somatosensory stimulation interventions, and technology in proprioceptive therapies. The following four criteria have been met for inclusion in the study: 1) the use of at least one proprioceptive intervention therapy; 2) the incorporation of some type of technology; 3) publication of the study must be from 2019 or later, and 4) the information consulted must be in English language. A total of 107 articles were reviewed, of which 30 met all the inclusion criteria.

## Analysis of parameters

To analyze the various related works according to the previously established criteria, a classification was performed based on the type of intervention: Movement Training (MT); Somatosensory Stimulation Training (SST), and Force Reproduction Training (FRT). The following categories were then considered: Artificial Intelligence (AI); mechanical devices; Virtual Reality (VR); exoskeletons; sensors, and IoT, with each of the related works was classified into one of these categories. Finally, the focus area of the rehabilitation work is categorized into: shoulders, arms, knees, feet, hands, gait re-education, upper limbs, and lower limbs, this serving as the third classification. The complete flow of the literature search and parameter selection are illustrated in Figure 1.

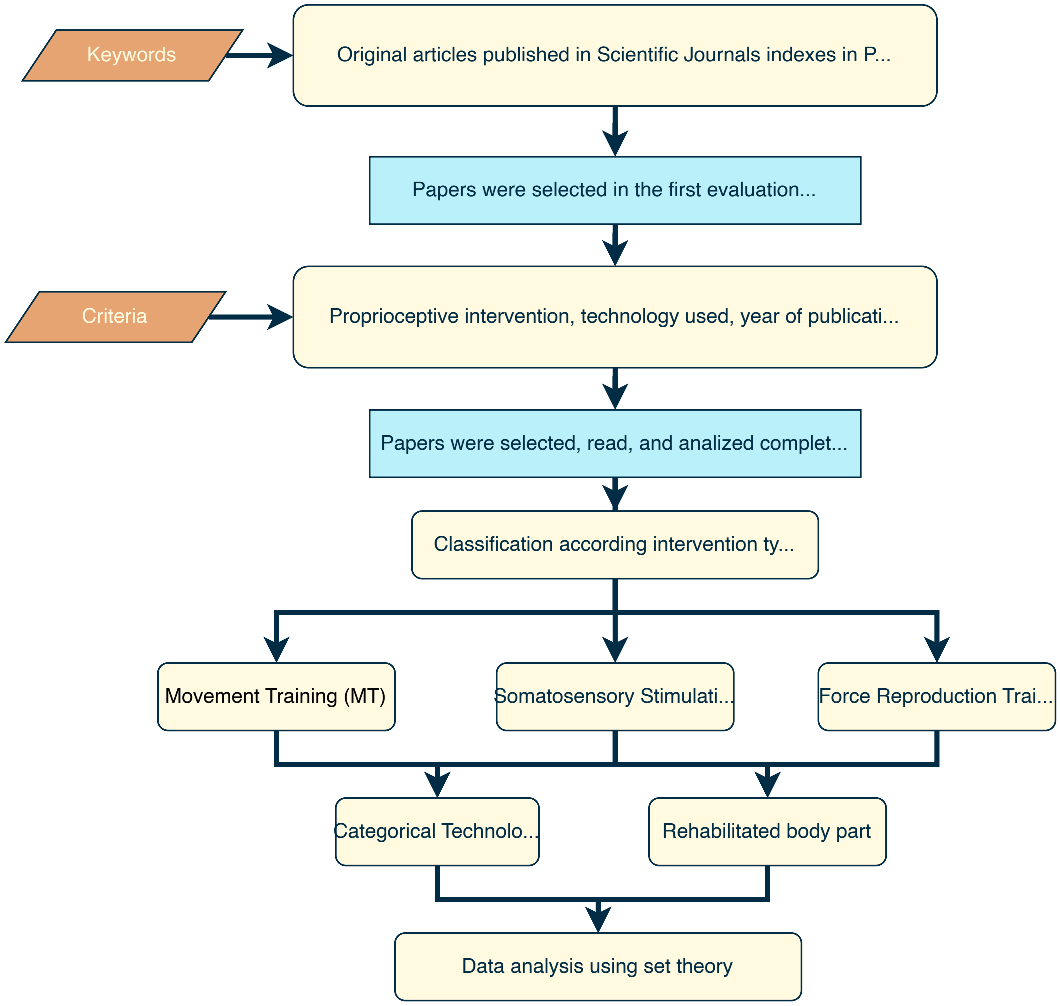


FIGURE 1. **Search and criteria for related works.**

Proprioceptive training refers to any intervention designed to enhance proprioceptive function, with the aim of recovering motor function and performance[9]. In clinical practice, there are various methods for proprioceptive interventions, including Movement Training (MT), Somatosensory Stimulation Training (SST), and Force Reproduction Training (FRT)[12], as shown in Figure 2. Additionally, various auxiliary technologies can be utilized at different phases of the evaluations.

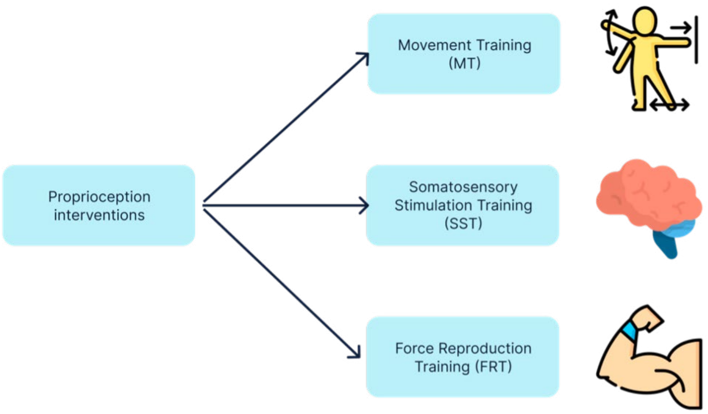


FIGURE 2. **Types of proprioceptive interventions.**

## RESULTS AND DISCUSSION

### a) Movement Training (MT)

Movement Training (MT) focuses on improving both the active and passive movement of a person's joints. Passive movement involves reaching a target at a specific position, and is assessed through a positional error test to determine whether the target has been successfully reached. Active movement, on the other hand, involves moving a body part from one position to another and assesses the accuracy of the Joint Position Sense (JPS). This assessment is conducted by measuring the movement threshold.

The compendium of Table 1 summarizes various studies utilizing interventions based on MT. These studies suggest the use of several technologies, including Convolutional Neural Network (CNN)[13], mechanical devices[14], Deep Neural Networks (DNN)[15], exoskeletons[16], linear regression[17], sensors[18] and Virtual Reality (VR)[19]. These technologies have been applied to different aspects of the rehabilitation or improvement of proprioceptive movement functioning, such as the elbow, shoulder, knee, hip, arms, upper limbs, lower limbs, gait re-education, and body.

TABLE 1. **Related works for proprioceptive interventions in MT (Continue in the next page).**

| Ref. | Objective | Input data | Outcomes | Category | Applications |
|---|---|---|---|---|---|
| [13] | Develop a cost-efficient system for monitoring home-based rehabilitation using advanced technologies | Motion capture | The MLP algorithm effectively classified ROM and compensatory movements with accuracy in 89 % | AI | Elbow, shoulder, standing upright |
| [20] | Evaluate the validity and reliability of deep learning-based motion capture, such as DeepLabCut, for proprioception assessment compared to the gold-standard 3D optoelectronic system | Motion capture | DeepLabCut show a similar performance to that of Vicon to measure knee position | AI | Knee |

TABLE 1. **Related works for proprioceptive interventions in MT (Continue in the next page).**

| Ref. | Objective | Input data | Outcomes | Category | Applications |
|---|---|---|---|---|---|
| [21] | Validity OpenPose for knee range of motion assessment after TKA against radiography and goniometry | Motion capture, images, and goniometry | OpenPose showed higher ICC and narrower 95 % limits of agreement for ROM extension than goniometry | AI | Knee |
| [22] | Develop a computer vision-based method for real-time gait detection in patients using rehabilitation exoskeletons | Images | GaitPoseNet demonstrated a good percentage of correct keypoints, 95.77 % | AI, Exoskeleton | Hip, knee |
| [17] | Quantify lower limb proprioception and its correlation with clinical characteristics in SCA3 patients | Sensors | Lower limb proprioception in SCA3 patients was significantly impaired compared to the healthy control group, $p < 0.05$ | Mechanical device | Lower limbs |
| [14] | Assess the impact of proprioceptive exercises on disease activity in postmenopausal women with RA | Sensors | The proprioceptive exercise program improved ankle function during gait in postmenopausal women with rheumatoid arthritis | Mechanical device | Gait re-education |
| [23] | Develop a framework for 3D user pose estimation from egocentric videos, utilizing proprioception as a key signal | Video | Experimental results show that the proposed framework significantly outperforms state-of-the-art methods | AI | Body |
| [24] | Evaluate shoulder proprioception after reverse shoulder arthroplasty for irreparable humeral fractures | Sensors | The results showed significant improvement in shoulder proprioception at 3, 6, and 12-months post-surgery | Mechanical device | Shoulder |
| [25] | The objective is real-time, on device rehabilitation assessments, overcoming the limitations of existing models such as PoseNet | Images | DeepRehab uses ResNet101, enabling more precise keypoint identification, particularly in the lower limbs | AI | Body |
| [26] | Develop a stimulation approach to assess the performance of the rehabilitation exoskeleton designed by the authors | Sensors | The exoskeleton improved gait training in stroke patients, enhancing kinematics, proprioception, metabolism, and muscle activation | Exoskeleton | Gait re-education |
| [27] | Evaluate the feasibility of predicting the KAM using anatomical landmarks that can be obtained from 2D video analysis | Video | KAM can be predicted with 95 % accuracy using a NN and video-based landmark positions | AI | Knee |

TABLE 1. **Related works for proprioceptive interventions in MT (Continue from previous page).**

| Ref. | Objective | Input data | Outcomes | Category | Applications |
|---|---|---|---|---|---|
| [15] | Propose a deep learning framework for automated evaluation of rehabilitation exercise quality | Kinect v2 | The framework generates movement quality scores with 87 % accuracy, closely matching actual quality scores | AI | Body |
| [28] | Introduce an innovative and user-friendly system using M-IMUs for evaluating and rehabilitating patients with proprioceptive disorders | Sensors | Results in healthy subjects show that the VR game is easy to use and that it effectively captures the subject's attention | Sensors, VR | Arms, knee |
| [29] | Develop a real-time 3D proprioception system for soft bodies using computer vision and deep learning techniques | Motion capture | Experiments showed the methods high accuracy in 3D shape detection (dH: ≈1 mm, relative error: ≤1 %) | AI | Body |
| [19] | Evaluate the effect of VR-based proprioceptive training on postural stability in workers performing task at height | Sensors | The study provides evidence that VR can effectively improve postural stability and reduce fall risk in workers at height | Mechanical device, VR | Postural stability |
| [30] | Develop and assess a low-cost, portable system for measuring knee proprioception using inertial sensors | Sensors | Significant effects were observed for absolute error in 15-25° and 35-45°, and for variable error in 35-45° | Sensors | Knee |
| [31] | Investigate whether reward feedback can generate a lasting recalibration of vision and proprioception compared to error feedback | Handle of the robotic manipulator | Reward feedback reduced visuo-proprioceptive errors but did not improve retention compared to error-based feedback | Mechanical device | Arms |
| [2] | Determine the impact of movement speed and distance on upper limb in young and older adults | Handle of the robotic manipulator | The study found that faster speeds and longer distances increased proprioceptive error in both groups | Exoskeleton, VR | Upper limbs |

The reviewed studies highlight a growing focus on the application of AI-based technologies, particularly Convolutional Neural Networks (CNN) and Deep Neural Networks (DNN). These technologies process input data from videos, images, or motion capture systems to generate 2D or 3D kinematic models[13][29], enabling precise digitization and mapping of the human body structure. This capability has proven valuable in evaluating parameters such as range of motion (ROM), achieving accuracies of 89 % and a Mean Per Joint Pose Error (MPJPE) of 57.4 %. In certain instances, these tools are specifically designed to assess the accuracy (95 %) of keypoint detection used by AI to estimate human posture or, alternatively, to classify movement patterns. On the other hand, the second most commonly used technology was exoskeletons. Their primary advantage lies in the integration of mechanical components, sensors, actuators and, in many cases, proprietary software, which allows the unification of these elements in a single device. Notable examples include KINARM, with an approximate commercial value of $ 189,000.00

USD[31], and BEAR-H1, which is currently under development[26]. Compare the performance of the exoskeleton with AI models, highlighting that significant discrepancies continue to exist between these technologies[22]. Specifically, the reported activation values were as follows: left hip (4.8); right hip (8.36); left knee (11.89), and right knee (17.75). Other studies focus on muscle activation, enhancing movement though targeted exercises, which improved the proprioceptive sense from 11.6 to 37.8[26].

In contrast, while mechanical devices can be costly, they are generally user-friendly. Some options available on the market include, for example, the Pro-Kin priced at approximately $ 22,499.00 USD[17], and the Human Norm II with an estimated cost of $ 79,995.00 USD[24]; these devices provide users with specialized tools that enable them to target specific areas and address deficiencies, such as those associated with total shoulder arthroplasty, Spinocerebellar Ataxia (SCA3), Rheumatoid Arthritis (RA), among other conditions. The studies analyzed include proprioception assessments utilizing these devices, reporting a range of results. Notable findings include significant values of $p < 0.005$, reductions in the Disease Activity Score (DAS) from 4.7 to 4.2[14], and improvements in the ROM of the treated areas, with an average of 31.1° ± 2.0, $p = 0.72$.

VR technologies are often integrated with recreational games[28], enabling users to perceive rehabilitation processes not as monotonous exercises, but as engaging activities. As games, these systems foster motivation, encouraging users to consistently improve and gradually progress in their recovery. The results of the reviewed studies indicate faster proprioception recovery in the affected areas. Another application of these systems involves their integration with external devices, such as balance platforms, to measure postural stability[19]. This integration allows for the acquisition of parameters such as the Center of Pressure (COP) and the evaluation of postural stability. In this regard, the use of sensors such as M-IMU is a low-cost alternative compared to exoskeletons and mechanical devices. The authors employed elastic bands positioned in areas requiring proprioceptive evaluation[30]. This technology facilitates the measurement of absolute and relative error, demonstrating significant effects in flexion and abduction movements within the ranges of 15° - 25° and 35° - 45°.

Categories, their primary applications, and input data were analyzed and represented in a network diagram (see Figure 3). Node sizes vary, indicating the relative importance or frequency of the use of each term.

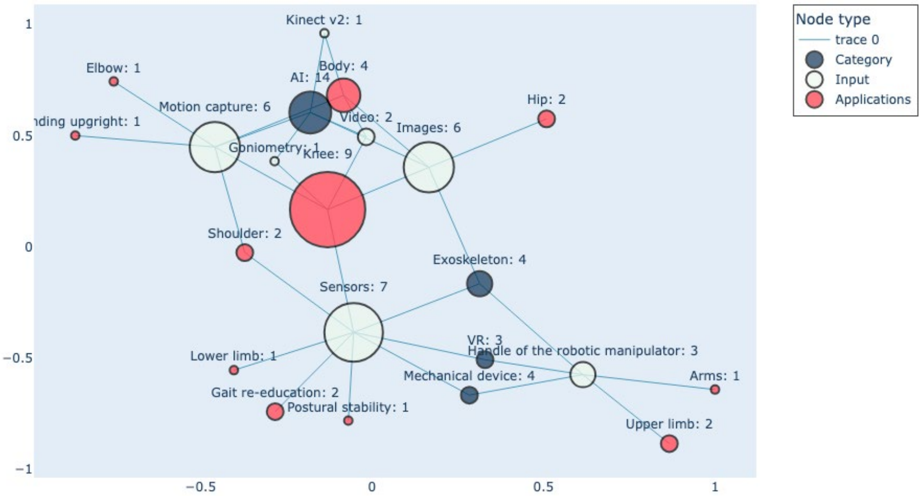


FIGURE 3. **Network of related works in MT considering Category, Input data, and main Applications.**

Rehabilitation therapies for MT revealed a predominant trend toward the adoption of Artificial Intelligence (AI), with 14 studies focusing on this approach. In contrast, niche opportunities were observed in the use of sensors, featured in one study, and Internet of Things (IoT) devices, which were not addressed in any study. In terms of input data, sensors were the most frequently employed, appearing in seven studies, followed by images and motion capture, each utilized in six studies. Regarding niche opportunities, robotics were the focus of three studies, while traditional devices were addressed in one study. Finally, trends in applications primarily targeted the knee (nine studies) and the whole body (four studies), whereas niche applications included standing upright and the elbow, with one study each.

### b) Somatosensory Stimulation Training (SST)

Somatosensory Stimulation Training (SST) involves the application of a variety of methods, including vibrations, thermal stimulation, magnetic stimulation, electrical stimulation, or acupuncture[32][33]. To assess somatic proprioceptive sensitivity, evaluations usually consist of skin tests; for example, graphesthesia, which measures the ability to recognize handwriting on the skin, is assessed by having the patient blindfolded while the examiner writes numbers or letters on the palm of the hand or on the fingerprints using a relatively blunt instrument. Another example is the assessment of paresthesia, which involves evaluating abnormal sensations such as tingling or burning in the absence of specific stimuli, this evaluation is conducted through laboratory tests and physical examinations[34].

The technology employed for somatosensory stimulation is less varied compared to MT and typically includes electronic modules, exoskeletons and mechanical devices. These are commonly used elements in SST. Table 2 presents some related works associated with this type of intervention.

TABLE 2. **Related works for proprioceptive interventions in SST (Continue in the next page).**

| Ref. | Objective | Input data | Outcomes | Category | Applications |
|---|---|---|---|---|---|
| [35] | Introduce and validate a new portable device for sensorimotor rehabilitation, combining MT and SST | Electromyography | The sensory training significantly improved two-point discrimination (TPD) on stimulated skin ($p$=0.047) | IoT | Arm |
| [36] | Investigate how the proprioceptive pathway processes muscle spindle signals to support various computational objectives | Musculoskeletal model | Models optimized for limb kinematics tasks best predicted neuronal activity in the cuneate nucleus | AI, Mechanical device | Shoulder, elbow |
| [32] | Investigate the proprioceptive system's role in action recognition, beyond its traditional function in representing bodily posture | Musculoskeletal model | Directional selective units were found only in models for action recognition, not in those for trajectory decoding | AI | Arm |

TABLE 2. **Related works for proprioceptive interventions in SST (Continue from previous page).**

| Ref. | Objective | Input data | Outcomes | Category | Applications |
|---|---|---|---|---|---|
| [18] | Evaluate the accuracy of perceiving different aspects of artificial proprioceptive feedback using a vibrotactor array | Interpretation of subjects' responses | The results showed over 90 % accuracy in recognizing proprioception components during single and dual tasks | Mechanical device | Forearm |
| [37] | Determine whether non-visual proprioceptive training can improve sensorimotor function in individuals who have suffered a stroke | Handle of the robotic manipulator | Seventy-three percent of stroke participants responded to the 2-day somatosensory training, improving wrist proprioceptive acuity by 30 % | Exoskeleton | Wrist |
| [38] | Present a new standardized psychometric platform for collecting data on somatosensory sensations evoked by neuroprosthesis | Software | The study found the platform easy to use and valuable for sensing electrical stimulation effects | Software | Arm, hand |
| [39] | Determine whether combining advanced and conventional rehabilitation therapy is more effective than treadmill training for improving proprioception and balance in subacute stroke patients | Sensors, software | The study showed a 43 % improvement in proprioceptive sensory changes through multisensory training | Mechanical device | Lower limb |

Mechanical devices utilized in SST play a crucial role in rehabilitation by integrating advanced sensors capable of measuring and adjusting the patient's responses. Researchers combined IA with mechanical devices to investigate how proprioception processes signals from the muscle spindle[36]. This approach has resulted in the development of optimized models for kinematic tasks involving the upper limbs, achieving a 40 % increase in Explained Variance (EV). Other studies combined mechanical devices with strategically placed sensors for targeted evaluation[18]. Vibratory stimulations were subsequently applied to the muscle spindles, with sensors capturing and analyzing the emitted signals. By processing this data, the ROM was measured with 99 % accuracy, and the JPS with 91.7 %.

AI has the potential to significantly enhance somatosensory training by providing personalized and adaptive feedback. In the reviewed studies, AI was employed to generate a dataset representing the proprioceptive trajectories of characters, neural networks models were designed and trained, achieving accuracies of 97 %. Furthermore, the Temporal Convolutional Network (TCN) demonstrated superior performance, with an accuracy of 98.86 % [32].

Today, it is possible to construct devices that utilize electronic modules, microcontrollers, or microcomputers (such as, ESP32, Arduino, Raspberry Pi) to combine motor training with pattern recognition and myoelectric signals. These devices employ training sensors to transmit data, enabling the integration of output devices such as screens or monitors. The advantage of these technologies includes their low energy consumption and cost. Integration with other open source tools is usually simple, and the materials used for construction allow for flexible shaping and adapting different projects, ensuring ease-of-use based on specific needs. Research highlights wearable devices that integrate MT and SST with rehabilitation therapies[35]. The analysis demonstrated that proprioceptive exercises with these devices significantly improved Two-Point Discrimination (TDT) in the stimulated areas, achieving a p-value of 0.0021.

Exoskeletons, such as the WristBot, with an estimated market value of $ 17,831.00 USD[37], play a central role in studies investigating the efficacy of non-visual proprioceptive training in improving sensorimotor function among stroke survivors. The device facilitates movement execution through handle manipulation, effectively stimulating and enhancing sensorimotor function. Results indicate that 73 % of participants achieved a 30.2 % reduction in difference threshold and an average 22 % decrease in tracking error. A key advantage of the WristBot lies its all-in-one design, integrating essential components for regular measurement, analysis, and therapy evaluation based on the targeted body area. However, a notable limitation is its substantial size, which necessitates considerable space for installation. Figure 4 illustrates the relationship among the proposed technological categories, input data, and SST applications in the reviewed studies.

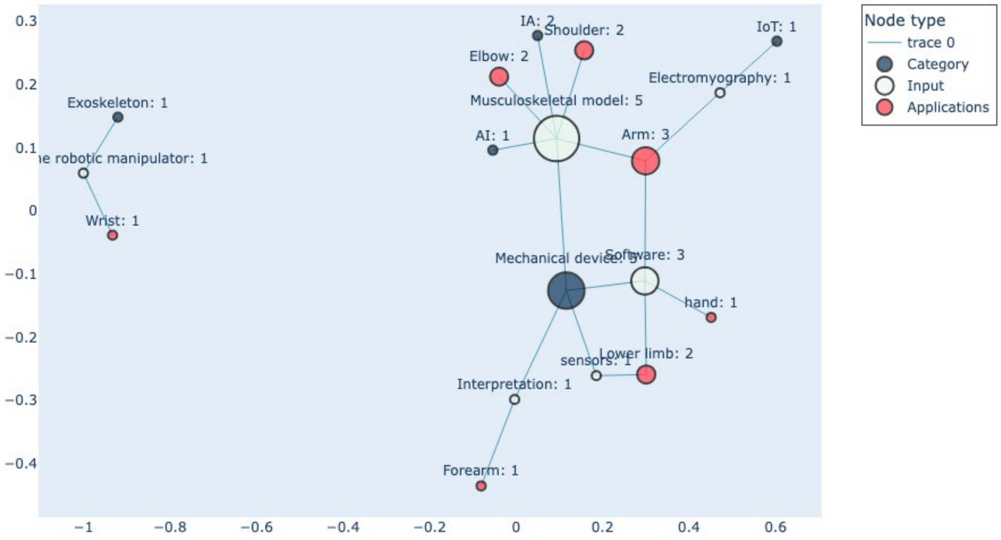


FIGURE 4. **Network of related works in SST considering Category, Input data and main Applications.**

A clear trend emerges toward the use of mechanical devices, which account for a total of five studies, while emerging low-cost niches such as IoT are represented by a single study within the technological categories. Additionally, musculoskeletal models, featured in five studies, emerged as the preferred type of data input, closely associated with mechanical devices. These are followed by proprietary software, utilized in three studies. Opportunities for further exploration were identified in the use of electromyography tests, which were featured in only one study.

Finally, the majority of on SST primarily focus on applications involving the arms, elbows, hands, and wrist. However, a notable gap remains in research addressing the lower limbs.

### c) Force Reproduction Training (FRT)

Tension and strength are commonly used to measure proprioception and are often evaluated during Force Reproduction Training (FRT). Contralateral limb matching has become the preferred method for assessing force reproduction. This method usually involves applying the Maximum Voluntary Isometric Contraction (MVIC) measurement force and attempting to replicate it. The coincidence forces can occur either in the same limb or in the contralateral limb[40]. Within the clinical environment, the dynamometer (such as grip, hand, or traction dynamometers) is the most commonly used device for conducting force reproduction assessments[41]. In this context, emerging technologies utilized to assess FRT predominantly include exoskeletons and mechanical devices, has gained popularity in recent years. The main advantage of these technologies is their ability to evaluate multiple areas of the body with a single device. However, their main disadvantages are the high costs of acquisition, maintenance, administration, and the space required for their use. Table 3 presents a selection of studies focused on FRT.

TABLE 3. **Related works for proprioceptive interventions in FRT**

| Ref. | Objective | Input data | Outcomes | Category | Applications |
|---|---|---|---|---|---|
| [5] | Investigate the relationship between proprioceptive deficits and motor functions in chronic stroke survivors | Handle of the robotic manipulator | Stroke participants had greater errors and variability in the passive position matching test | Mechanical device | Arm |
| [42] | Create a simulation framework for a lower-limb model with CNS signals and proprioceptive feedback | Images | Incorporating proprioceptors enhances simulation realism and enables more accurate analysis of human movement mechanisms | Software | Leg |
| [43] | Develop a bionic knee joint structure based on tensegrity for high-efficiency rehabilitation knee exoskeleton | Images | Simulations and experiments validated the NTZNN controllers' performance and noise resistance in noisy environments | IA | Knee |
| [16] | Evaluating the effectiveness of ML and DL methods in detecting stroke presence using kinematic data from a robotic APM task | Handle of the robotic manipulator | ML and DL models outperformed the traditional cutoff scoring technique in stroke classification | IA, exoskeleton | Upper limb |
| [44] | Investigate how gender and grip types affect the accuracy of grip force reproduction in healthy individuals | Dynamometer | Women are more accurate or consistent in detecting pinch force changes than men | Electronic device | Finger |

AI is revolutionizing strength training by integrating with technologies such as computer vision and wearable sensors to enhance proprioception. Recent studies highlight the use of neural networks (NN) combined with biomechanical devices to improve knee tensegrity[43], These studies involve torque testing and represent various rehabilitation states (-40 N, -150 N, or 0 N), achieving improved knee flexion in older adults, with ranges reaching 130°-140°. Another study demonstrates the use of AI in conjunction with exoskeletons, focusing on a robotic arm that assists patients in performing kinematic movements[16]. The data is processed by a DNN, comparing technological applications against traditional methods. The DL models outperformed conventional cut-off scoring techniques for stroke patients, achieving an accuracy of 86 % and a logistic regression score of 86.6 %.

It should be emphasized that the visualization of information in a user-friendly and comprehensible manner is crucial in the development of technologies, as it facilitates the rapid integration of processed data from sensors, AI, and mechanical devices. Among the reviewed works, a framework stands out that simulates a musculoskeletal model of the lower limbs, processing signals from the nervous system[42]. This study performs a Golgi evaluation, enabling the detection of force tensions ranging from 0.4 N - 7.2 N. The ability to present this data clearly and accessibly is essential for accurate interpretation and decision-making in applications related to proprioception and strength training. In this regard, dynamometers have undergone significant evolution through the integration of advanced technologies such as digital sensors, wireless connectivity, and proprietary software, enabling more precise measurements and real-time force analysis. Research has concentrated on using these devices to measure pinch strength in older adults, particularly to assess MVIC in both men and women[44]. The findings suggest that women exhibit greater accuracy and sensitivity in detecting changes in pinch strength compared to men. This difference may have important implications for designing proprioceptive training programs and strength assessments tailored to specific populations.

Figure 5 illustrates the relationship among the suggested technological categories, the processed input data, and their main applications across different areas of the body. This visual representation reveals how different technologies, including AI, electronic devices, exoskeletons, and software, interact with the data generated by the body during exercises. Additionally, it details how each technology is applied to specific areas, enabling a more precise and personalized analysis of strength.

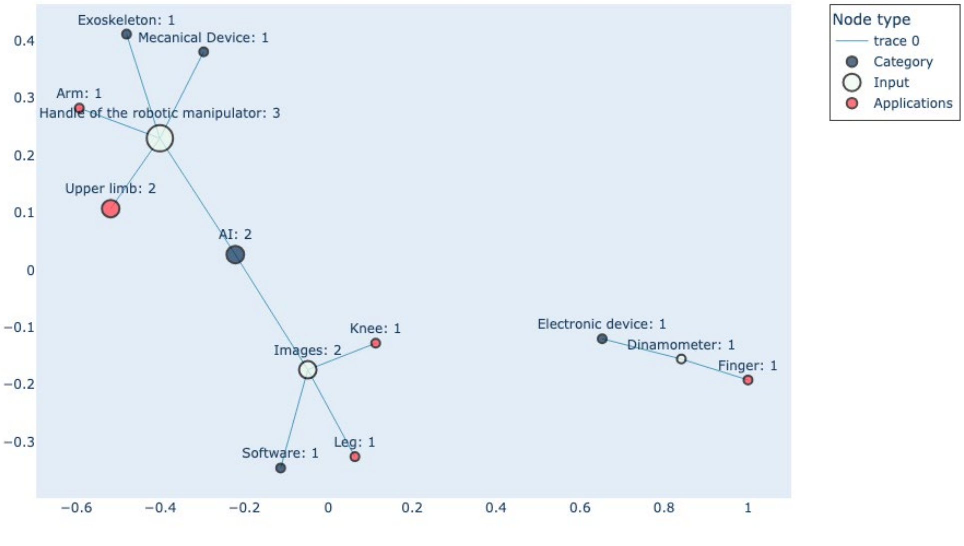


FIGURE 5. **Network of related works in FRT considering Category, Input data, and main Applications**

Studies focused on improving proprioception in strength training remain limited, although there is a growing trend in the use of AI. However, VR represents an area of opportunity that remains underexplored. The primary sources of data input in the studies derive from robotic devices, which enable real-time measurement and analysis of forces and movements. The applications identified in the studies mainly focus on the upper extremities, including the fingers and arms, as well as the knees and legs—key areas for enhancing proprioception in strength training.

### d) Gaps and limitations detected

Emerging technologies applied to proprioceptive rehabilitation exhibit significant potential. However, it is essential to recognize their current limitations and take proactive steps toward addressing these. Several studies have identified the following gaps:

- AI: Despite being the most widely used technology, there is a need to improve the accuracy of various elements, such as classification, keypoint detection, ROM, JRE and MAE

- Exoskeletons: while showing promise for rehabilitation and the improvement of proprioception, they tend to be costly, limiting their accessibility for many users. Additionally, their size and weight can make them challenging to use in certain environments or for specific patients, especially those with reduced mobility. The setup and operation of these devices may also require specialized personnel and increases complexity and operational costs, rendering their implementation in low-budget clinical or rehabilitation settings more difficult

- Mechanical devices: It is necessary to reduce the cost of these devices, as they are effective for specific needs but require a broader range of applicable situations

- VR: users with ophthalmic issues may experience difficulties when using VR devices, as these can affect vision or exacerbate certain visual conditions. The majority of studies utilizing VR employ games, prioritizing patient-friendliness rather than increasing the difficulty of movement

- Sensors: although these devices are typically low-cost, they require the support of the remaining technologies to perform proprioceptive assessment. Examples include actuators, elastic bands, microcontrollers, AI, and proprietary software, among others

Based on the aforementioned analysis, the gaps and limitations identified in MT, SST, and FRT indicate that, despite the existence of various supporting technologies, further work is needed in real-world validation, particularly with patients who have diverse conditions and needs. While these technologies have applications across different parts of the body, the reviewed studies highlight limited research in key areas such as the fingers, wrist, ankle, hip, and elbow. These underexplored body areas present an opportunity to expand the application of these technologies in proprioceptive rehabilitation and training.

### e) Challenges

Despite technological advancements across various fields, continuous improvements are essential in each of the emerging technologies identified. The main challenges include:

1. AI: enhancing the precision of its algorithms, exploring different techniques, and reducing the amount of information required to generate the dataset

2. Exoskeletons, Mechanical devices, and VR: principally focusing on reducing costs, improving ergonomics, and refining their design
3. Sensors: focusing on improving accuracy, simplifying calibration and configuration processes, and reducing energy consumption

Thus, the challenges for MT, SST, and FRT may lie in the integration of multiple technologies, such as AI with sensors or electronic devices with software. The goal is to reduce the costs associated with more complex solutions, such as exoskeletons, mechanical devices, or VR.

### f) Trends

As illustrated in Figures 3-5, IA is a contemporary approach that addresses various proprioceptive interventions from multiple perspectives, proving to be particularly useful when budget constraints are present by offering a wide range of solutions. On the other hand, home devices (computers, telephones, tablets) are accessible tools that are capable of processing substantial amounts of information, making the use of artificial vision algorithms (CNN, DNN) feasible for image processing. It is evident that current technological access, combined with the ease of acquisition and integration with smart tools, facilitates the development of low-cost solutions that do not require specialized spaces or complex structures, this presenting a promising area for future exploration.

IoT can make communication possible among devices developed by different teams, storing information in the cloud, and constantly generating data regarding the use of the device. This opens a world of possibilities for communication among different device architectures that were initially incompatible or designed for a single purpose.

## CONCLUSIONS

This article aims to provide an overview of emerging technologies and their applications within proprioceptive training. There is a significant opportunity for growth in MT, SST, and FRT, as the studies presented are limited, underscoring the need for more in-depth exploration. It is recommended to conduct comprehensive analyses of each type of rehabilitation therapy, focusing on specific areas of application (e.g., arm, leg, hip) or particular techniques (e.g., ROM, JPS, JRE). This approach will enable a more focused and detailed examination.

Technology has been employed to enhance proprioception in various ways, encompassing both the pre- and post-operative periods. In the field of MT, studies have been conducted to assess various aspects of mobility, such as Range of Motion (ROM), Joint Position Sense (JPS), Joint Reproduction Error (JRE), Percentage of Correct Keypoints (PCK), Mean Absolute Error (MAE), Knee Adduction Moment (KAM), and Center of Pressure (COP). A similar approach is observed in SST, where Two-Point Discrimination (TPD), accuracy, and usability have been evaluated. Last, FRT evaluates the absolute error, knee torque, and Maximum Voluntary Isometric Contraction (MVIC). All of these studies employed various emerging technologies applied to different areas of the body.

AI is a key technological trend that complements other technologies in rehabilitation therapies. Due to its cost-effectiveness, ease-of-implementation, and flexibility, AI is increasingly adopted in therapies such as Movement Training (MT), Somatosensory Stimulation (SST), and Force Reproduction Testing (FRT). AI is utilized for tasks such as posture classification, keypoint detection, signal and pattern analysis, recognition, and correlation detection. In addition, when adequate space and budget are available, mechanical devices, exoskeletons, or virtual reality can serve as a viable and effective alternative, tailored to the specific needs of each patient. Thus, devices that operate with IoT or sensors are presented as a low-cost alternative that must be integrated with other technologies to serve as a viable solution for proprioceptive rehabilitation.

It is possible to affirm that emerging technologies currently play a crucial role in Medicine. They serve as a support throughout various stages of rehabilitation therapies, allowing for the development of modern and innovative alternatives that enhance and elevate an individual's quality of life.

It can be concluded that there currently exists a greater technological demand oriented towards MT, followed by SST and finally FRT. The review found that multiple factors must be assessed when selecting the appropriate support technology for rehabilitation, including: the type of intervention required (MT, SST, or FRT), input data, application, rehabilitation site (clinic or home), budget constraints, available space (suitability and adaptation of the area), and whether specialized supervision is needed. Therefore, the final selection will depend on what the user can afford and the specialist's recommendation.

## ACKNOWLEDGMENTS

We would like to express our gratitude to the Universidad Autónoma de Guerrero for providing the facilities and resources necessary for conducting postgraduate studies in Innovation Sciences. We also extend our thanks to CONAHCyT for grant number 421452, which supported this research.

## DECLARATION OF CONFLICT OF INTEREST

The authors declares that they have no financial or personal relationship with organizations or individuals that could unnaturally influence or interfere in the work.

## CONTRIBUTIONS OF THE AUTHORS

U. T. C. conceptualization, investigation, methodology, formal analysis, visualization and writing of the original draft; J. L. H. H. conceptualization, methodology, validation, supervision; E. D. G. conceptualization, methodology, references, supervision; R. I. M. A. conceptualization, supervision, methodology; R. S. conceptualization, methodology, validation, visualization, supervision. All authors reviewed and approved the final version of the manuscript.

## REFERENCES

**[1]** A.-V. Bruyneel, "Evaluación de la propiocepción: pruebas de estatestesia y cinestesia," EMC - Kinesiter. - Med. Fis., vol. 37, no. 4, pp. 1-11, 2016, doi: https://doi.org/10.1016/S1293-2965(16)78903-1

**[2]** D. T. Tulimieri and J. A. Semrau, "Aging increases proprioceptive error for a broad range of movement speedand distance estimates in the upper limb," Front. Hum. Neurosci., vol. 17, 2023, art. no. 1217105, doi: https://doi.org/10.3389/fnhum.2023.1217105

**[3]** M. L. Alfonso Mora, et al., "Métodos de evaluación de la propiocepción en deportistas. Revisión de la Literatura," Rev. Digi. Act. Fís. Deport., vol. 4, no. 1, pp. 69-82, 2018. doi: https://doi.org/10.31910/rdafd.v4.n1.2018.415

**[4]** A. E. Khrulev, et al., "Modern Rehabilitation Technologies of Patients with Motor Disorders at an Early Rehabilitation of Stroke (Review)," Sovrem Tekhnologii Med., vol. 14, no. 6, pp. 64-78, doi: https://doi.org/10.17691/stm2022.14.6.07

**[5]** A. Cherpin, et al., "A preliminary study on therelationship between proprioceptive deficits and motor functions in chronic stroke patients," 2019 IEEE 16th International Conference on Rehabilitation Robotics (ICORR), Toronto, Canada, 2019, pp. 465-470, doi: https://doi.org/10.1109/ICORR.2019.8779447

**[6]** V. Suglia et al., "A Serious Game for the Assessment of Visuomotor Adaptation Capabilities during Locomotion Tasks Employing an Embodied Avatar in Virtual Reality," Sensors, vol. 23, no. 11, 2023, art. no, 5017, doi: https://doi.org/10.3390/s23115017

**[7]** L. Avanzino and M. Fiorio, "Proprioceptive dysfunction in focal dystonia: From experimental evidence to rehabilitation strategies," Front. Hum. Neurosci., vol. 8, 2014, art. no. 1000, doi: https://doi.org/10.3389/fnhum.2014.01000

**[8]** H. Shibasaki et al., "A new form of congenital proprioceptive sensory neuropathy associated with arthrogryposis multiplex," J. Neurol., vol. 251, no. 11, pp. 1340-1344, 2004, doi: https://doi.org/10.1007/s00415-004-0539-4

**[9]** J. E. Aman, N. Elangovan, I. L. Yeh, and J. Konczak, "The effectiveness of proprioceptive training for improving motor function: A systematic review," Front. Hum. Neurosci., vol. 8, 2015, art. no. 1075, doi: https://doi.org/10.3389/fnhum.2014.01075

**[10]** A. Vladimirovich Zakharov, et al., "Proprioception in Immersive Virtual Reality," in Proprioception, 2021, p. 668, doi: https://doi.org/10.5772/intechopen.96316

**[11]** A. Sidarta, et al., "Current clinical practice in managing somatosensory impairments and the use of technology in stroke rehabilitation," PLoS One, vol. 17, no. 8, 2022, art. no. e0270693, doi: https://doi.org/10.1371/journal.pone.0270693

**[12]** K. Valdes, K. C. Manalang, and C. Leach, "Proprioception: An evidence-based review," J. Hand Ther., 2023, art. no. 00142-4, doi: https://doi.org/10.1016/j.jht.2023.09.015

**[13]** C. Mennella, U. Maniscalco, G. De Pietro, and M. Esposito, "A deep learning system to monitor and assess rehabilitation exercises in home-based remote and unsupervised conditions," Comput. Biol. Med., vol. 166, 2023, art. no. 107485, doi: https://doi.org/10.1016/j.compbiomed.2023.107485

**[14]** P. Aleixo, T. Atalaia, J. V. Patto, and J. Abrantes, "The Effect of a Proprioceptive Exercises Programme on Disease Activity and Gait Biomechanical Parameters of Post-Menopausal Women with Rheumatoid Arthritis," in Rheumatoid Arthritis, 2022, doi: https://doi.org/10.5772/intechopen.99462

**[15]** Y. Liao, A. Vakanski, and M. Xian, "A Deep Learning Framework for Assessing Physical Rehabilitation Exercises," IEEE Trans. Neural Syst. Rehabil. Eng., vol. 28, no. 2, pp. 468-477, 2020, doi: https://doi.org/10.1109/tnsre.2020.2966249

**[16]** D. Hossain, S. H. Scott, T. Cluff, and S. P. Dukelow, "The use of machine learning and deep learning techniques to assess proprioceptive impairments of the upper limb after stroke," J. Neuroeng. Rehabil., vol. 20, no. 1, 2023, art. no. 15, doi: https://doi.org/10.1186/s12984-023-01140-9

**[17]** X. H. Liu, et al., "Impaired Lower Limb Proprioception in Spinocerebellar Ataxia Type 3 and Its Affected Factors," Front. Neurol., vol. 13, 2022, art. no. 833908, doi: https://doi.org/10.3389/fneur.2022.833908

**[18]** L. Vargas, H. H. Huang, Y. Zhu, and X. Hu, "Static and dynamic proprioceptive recognition through vibrotactile stimulation," J. Neural. Eng., vol. 18, no. 4, 2021, art. no. 046093, doi: https://doi.org/10.1088/1741-2552/ac0d43

**[19]** M. Cyma-Wejchenig, J. Tarnas, K. Marciniak, and R. Stemplewski, "The influence of proprioceptive training with the use of virtual reality on postural stability of workers working at height," Sensors, vol. 20, no. 13, 2020, art. no. 3731, doi: https://doi.org/10.3390/s20133731

**[20]** M. Van Den Bogaart, et al., "Validity of deep learning based motion capture using DeepLabCut to assess proprioception," Gait Posture, vol. 106, no. 1, 2023, art. no. S212, doi: https://doi.org/10.1016/j.gaitpost.2023.07.255

**[21]** Y. Saiki, et al., "Reliabitily and validity of pose estimation algorithm for measurement of knee range of motion after total knee arthroplasty," Bone Joint Res., vol. 12, no. 5, pp. 313-320, 2023, doi: https://doi.org/10.1302/2046-3758.125.bjr-2022-0257.r1

**[22]** Y. Wang, Z. Pei, C. Wang, and Z. Tang, "Depth-aware pose estimation using deep learning for exoskeleton gait analysis," Sci Rep, vol. 13, no. 1, 2023, art. no. 22681, doi: https://doi.org/10.1038/s41598-023-50207-z

**[23]** W. Su, Y. Liu, S. Li, and Z. Cai, "Proprioception-Driven Wearer Pose Estimation for Egocentric Video," in 2022 26th International Conference on Pattern Recognition (ICPR), Montreal, Canada, 2022, pp. 3728-3735, doi: https://doi.org/10.1109/ICPR56361.2022.9956092

**[24]** Y. Guler, et al., "Shoulder Proprioception Following Reverse Total Shoulder Arthroplasty for Unreconstructable Upper Third Fractures of the Humerus: 2-Year Outcomes 2022," Indian J. Orthop., vol. 56, no. 12, pp. 2245-2252, 2022, doi: https://doi.org/10.1007/s43465-022-00769-3

**[25]** Dos Santos Melício, et al., "DeepRehab: Real Time Pose Estimation on the Edge for Knee Injury Rehabilitation," in Artificial Neural Networks and Machine Learning - ICANN 2021, Bratislava, Slovakia, 2021, doi: https://doi.org/10.1007/978-3-030-86365-4_31

**[26]** F. Mo, et al., "A simulation-based framework with a proprioceptive musculoskeletal model for evaluating the rehabilitation exoskeleton system," Comput. Methods Programs Biomed., vol. 208, 2021, art. no. 106270, doi: https://doi.org/10.1016/j.cmpb.2021.106270

**[27]** M. A. Boswell, et al., "A neural network to predict the knee adduction moment in patients with osteoarthritis using anatomical landmarks obtainable from 2D video analysis," Osteoarthritis Cartilage, vol. 29, no. 3, pp. 346-356, 2021, doi: https://doi.org/10.1016/j.joca.2020.12.017

**[28]** M. Lapresa, et al., "A Smart Solution for Proprioceptive Rehabilitation through M-IMU Sensors," in 2020 IEEE International Workshop on Metrology for Industry 4.0 & IoT, Roma, Italy, 2020, pp. 591-595, doi: https://doi.org/10.1109/MetroInd4.0IoT48571.2020.9138193

**[29]** R. Wang, et al., "Real-time Soft Body 3D Proprioception via Deep Vision-based Sensing," IEEE Robot Autom. Lett., vol. 5, no. 2, pp. 3382-3389, 2020, doi:

https://doi.org/10.1109/LRA.2020.2975709

**[30]** A. L. Rahlf, et al., “Validity and Reliability of an Inertial Sensor-Based Knee Proprioception Test in Younger vs. Older Adults,” Front. Sports Act. Living., vol. 1, 2019, art. no. 27, doi: https://doi.org/10.3389/fspor.2019.00027

**[31]** I. A. Kuling, A. J. de Brouwer, J. B. J. Smeets, and J. R. Flanagan, “Correcting for natural visuo-proprioceptive matching errors based on reward as opposed to error feedback does not lead to higher retention,” Exp. Brain Res., vol. 237, no. 3, pp. 735-741, 2019, doi: https://doi.org/10.1007/s00221-018-5456-3

**[32]** K. J. Sandbrink, et al., “Contrasting action and posture coding with hierarchical deep neural network models of proprioception,” Elife, vol. 12, 2023, art. no. e81499, doi: https://doi.org/10.7554/elife.81499

**[33]** L. Winter, Q. Huang, J. V. L. Sertic, and J. Konczak, “The Effectiveness of Proprioceptive Training for Improving Motor Performance and Motor Dysfunction: A Systematic Review,” Front. Rehabil. Sci., vol. 3, 2022, art. no. 830166, doi: https://doi.org/10.3389/fresc.2022.830166

**[34]** P. Mehta, et al., “Paresthesias and dysesthesias,” in Imaging Acute Neurologic Disease: A Symptom-Based Approach, Cambridge University Press, 2014, pp. 332-346, doi: https://doi.org/10.1017/CBO9781139565653.022

**[35]** M. Buist, et al., “Novel Wearable Device for Mindful Sensorimotor Training: Integrating Motor Decoding and Somatosensory Stimulation for Neurorehabilitation,” IEEE Trans. Neural Syst. Rehabil. Eng., vol. 32, pp. 1515-1523, 2024, doi: https://doi.org/10.1109/TNSRE.2024.3379996

**[36]** A. Marin Vargas, A. Bisi, A. S. Chiappa, C. Versteeg, L. E. Miller, and A. Mathis, “Task-driven neural network models predict neural dynamics of proprioception,” Cell, vol. 187, no. 7, pp. 1745-1761.e19, 2024, doi: https://doi.org/10.1016/j.cell.2024.02.036

**[37]** I. L. Yeh, et al., “Effects of a robot-aided somatosensory training on proprioception and motor function in stroke survivors,” J. Neuroeng. Rehabil., vol. 18, no. 1, 2021, art. no. 77, doi: https://doi.org/10.1186/s12984-021-00871-x

**[38]** G. Valle, et al., “A Psychometric Platform to Collect Somatosensory Sensations for Neuroprosthetic Use,” Front. Med. Technol., vol. 3, 2021, art. no. 619280, doi: https://doi.org/10.3389/fmedt.2021.619280

**[39]** C. Lim, “Multi-sensorimotor training improves proprioception and balance in subacute stroke patients: A randomized controlled pilot trial,” Front. Neurol., vol. 10, 2019, art. no. 157, doi: https://doi.org/10.3389/fneur.2019.00157

**[40]** G. Dover and M. E. Powers, “Reliability of Joint Position Sense and Force-Reproduction Measures During Internal and External Rotation of the Shoulder,” J. Athl. Train., vol. 38, no. 4, pp. 304-310, 2003.

**[41]** G. Coskun, B. Talu, and A. Cools, “Proprioceptive force-reproduction of the rotator cuff in healthy subjects before and after muscle fatigue,” Isokinet. Exerc. Sci., vol. 26, no. 3, pp. 175-181, 2018, doi: https://doi.org/10.3233/IES-173206

**[42]** H. Zhang, et al., “A Framework of a Lower Limb Musculoskeletal Model with Implemented Natural Proprioceptive Feedback and Its Progressive Evaluation,” IEEE Trans. Neural Syst. Rehabil. Eng., vol. 28, no. 8, pp. 1866-1875, 2020, doi: https://doi.org/10.1109/tnsre.2020.3003497

**[43]** C. Xu, et al., “An advanced bionic knee joint mechanism with neural network controller,” Front. Neurorobot, vol. 17, 2023, art. no. 1178006, doi: https://doi.org/10.3389/fnbot.2023.1178006

**[44]** L. Li, et al., “Effect of pinch types on pinch force sense in healthy adults,” Front. Hum. Neurosci., vol. 16, 2022, art. no. 990431, doi: https://doi.org/10.3389/fnhum.2022.990431

**REVIEW ARTICLE**
VOL. 46 | NO. 1 | **JANUARY-APRIL 2025** | PP 1-19

Revista Mexicana de Ingeniería Biomédica

https://dx.doi.org/10.17488/RMIB.46.1.1465

E-LOCATION ID: **e1465**

# Scientific and Technical overview about Artificial Proprioception in Prosthetics

## Panorama Científico y Técnico sobre Propiocepción Artificial en Prótesis

*Octavio Diaz-Hernandez*[1]

[1] Universidad Nacional Autónoma de México, Escuela Nacional de Estudios Superiores Unidad Juriquilla, Querétaro- México

## ABSTRACT

Proprioception is the body's ability to perceive its position and movement, which plays a crucial role in motor control, and its loss following amputation presents significant challenges for prosthesis users. Artificial Proprioception is an innovation that enhances sensory feedback and motor control in prosthetic devices. This review presents a comprehensive overview of current research and technological developments in Artificial Proprioception, focusing on sensory feedback mechanisms, neural interface systems, and the integration of biomechatronic technologies. With a growing interest in restoring sensory feedback for amputees, this work explores key innovations such as electrotactile and vibrotactile stimulation, artificial intelligence, and neural interfaces that enable a more natural and intuitive prosthetic control. The methodology included reviewing studies from databases like Scopus, Web of Science, and PubMed on proprioceptive feedback in prosthetics in recent years. It evaluates research related to sensory feedback, amputation levels, neural interfaces, and technological advancements, classifying papers by feedback mechanisms. The paper concludes by discussing potential future developments, including more advanced, user-centered prosthetic devices that address the sensory needs of amputees and improve their quality of life.

**KEYWORDS:** artificial proprioception, biomechatronic devices, prosthetics, rehabilitation technology, sensory feedback

## RESUMEN

La propiocepción es la capacidad del cuerpo para percibir su posición y movimiento, que desempeña un papel crucial en el control motor, y su pérdida tras una amputación plantea importantes retos a los usuarios de prótesis. La propiocepción artificial es un avance innovador para mejorar la respuesta sensorial y el control motor de las prótesis. Esta revisión presenta una visión global de la investigación actual y los avances tecnológicos en Propiocepción Artificial, centrándose en los mecanismos de retroalimentación sensorial, los sistemas de interfaz neural y la integración de la biomecatrónica. Con un interés creciente en la restauración de la retroalimentación sensorial para amputados, este trabajo explora innovaciones clave como la estimulación electrotáctil y vibrotáctil, la inteligencia artificial y las interfaces neurales que permiten un control protésico más natural e intuitivo. La metodología incluyó la revisión de estudios de bases de datos como Scopus, Web of Science y PubMed sobre retroalimentación propioceptiva en prótesis en los últimos años. Se evalúa la investigación relacionada con la retroalimentación sensorial, los niveles de amputación, las interfaces neurales y los avances tecnológicos, analizando los artículos por mecanismos de retroalimentación. El artículo concluye con un debate sobre posibles desarrollos futuros, incluidos dispositivos protésicos más avanzados y centrados en el usuario que aborden las necesidades sensoriales de los amputados y mejoren su calidad de vida.

**PALABRAS CLAVE:** dispositivos biomecatrónicos, propiocepción artificial, prótesis, retroalimentación sensorial, tecnología de rehabilitación

**Corresponding author**

TO: Octavio Diaz-Hernandez
INSTITUTION: UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
ADDRESS: BLV. JURIQUILLA 3001, QUERÉTARO, C.P. 76230, MÉXICO.
EMAIL: octavio.diaz@unam.mx

**Received:**

15 September 2024

**Accepted:**

24 February 2025

**Published:**

08 April 2025

# INTRODUCTION

## Normal sensory feedback

Humans can perform daily tasks like opening doors with a lock, navigating obstacles in a hallway, and operating a car efficiently because touch, proprioception, and vision all contribute to the closed-loop motor control system. People can benefit from knowing the anatomical and physiological basis of the tactile, proprioceptive, and visual sensory systems and how they affect movement control and limit human motor skill performance in all these skill performance scenarios. Sensory information's role in regulating action is fundamental to all motor control theories. Out of all the senses, touch, proprioception, and vision play significant roles in the motor control of abilities. Touch and proprioception are considered senses of the somatic sensory system in the study of human sensory physiology, while vision is the sense related to the visual sensory system[1]. The sense and awareness of one's own body's location and motion is known as proprioception. Proprioception is one of our fundamental senses that is frequently disregarded. However, it provides sensory data regarding movement properties like direction, location in space, velocity, and muscle activation to the central nervous system. Proprioceptive feedback is essential in closed-loop models of movement control because it can help adjust while moving when proprioceptive information is used to facilitate closed-loop control.

## Human Proprioception and enhanced feedback

Proprioception involves several essential physiological elements; proprioceptors are specialized sensory receptors in muscles, tendons, and joints crucial to this “sixth sense.” Golgi-tendon organs, found at the junction of muscles and tendons, check tension, while muscle spindles by sensing variations in muscle length and tension. Joint receptors provide information on joint angle and movement. Through sensory neurons, these receptors send signals to the central nervous system, where the brain combines the data to construct a coherent perception of movement and position inside the body[2]. People can access two general forms of performance-related information (feedback) when they execute a motor skill, which will "tell" them something about the result of the performance or the reason behind it. One is the sensory-perceptual data obtained from executing a task naturally, known as task-intrinsic feedback. The sensory systems (proprioception included) can deliver this kind of feedback. But there is a second type of feedback, which is “Task-intrinsic,” and the name used in the literature is “enhanced feedback”[3]; even though other names for this kind of input have been proposed, such as task-extrinsic feedback and external feedback.

## Amputation and prosthetics

Amputation of a limb is a tragic occurrence that negatively affects the individual's health and quality of life. The lack of comprehensive and up-to-date global data complicates figuring out the exact number of amputees worldwide. However, some estimates can be provided based on available studies and data. According to the World Health Organization (WHO), it is estimated that there are more than forty million people in the world living with an amputation. This number includes amputations caused by accidents, disease, medical complications, and armed conflict. In some countries, the figures may be more specific. For example, In the United States, it is estimated that around two million people are living with an amputation, and approximately 185,000 amputations are performed each year. In Europe, the number of amputations varies between countries, but it is estimated that there are hundreds of thousands of people living with amputations[4].

In Mexico, Amputations are a significant health issue, primarily because of the country's high diabetes and trauma injury rates. With almost twelve million affected, Mexico has one of the highest rates of diabetes worldwide, according to the Mexican Diabetes Federation[5]. Amputations are often caused by diabetic complications such as foot ulceration and peripheral vascular disease. According to estimates, the risk of having an amputation is up to fifteen times higher in individuals with diabetes than in those without the condition. Between 70,000 and 100,000 amputations are thought to be conducted in Mexico each year, the majority of which are connected to diabetes and its consequences[6][7].

### The importance of proprioceptive feedback in amputees

Since all sensory feedback is lost in an amputee, proprioception has been suggested to be one of the most important senses for movement and the ability to perform specific tasks. Considering that it is possible to provide "augmented feedback," it is possible to give the prosthesis wearer a series of controlled proprioceptive stimuli through technological means, that is, to obtain artificially generated proprioceptive feedback. In 1999, the importance of providing sensory feedback in upper limb prosthetics was addressed[8]. This is how Artificial Proprioception has been proposed to be a breakthrough in the rehabilitation of amputees. For this reason, Artificial Proprioception has been proposed as a means of retrieving somatosensory feedback that is helpful in the performance of tasks through an electronic system that causes the brain to perceive important information from the activities performed. One of the most relevant issues is the recovery gait control of individuals with lower limb prostheses[9]. In previous work, we studied and defined Artificial Proprioception[10], and a method was proposed. However, in this review, we have investigated the trend of science and technology worldwide to give an overview of the current state of artificial proprioception.

The review's main aim is to find relevant or comparable work on Artificial Proprioception worldwide, as well as information on the technology employed, advantages noted, and actions taken to enhance the quality of life of prosthesis users.

## MATERIALS AND METHODS

### Eligibility criteria

The studies will be related to sensory feedback generated by external devices that have been developed around the world. The keywords used in the searches are detailed below. The population of interest is amputees using upper and lower limb prostheses, regardless of socket type or amputation level, e.g., transfemoral, transtibial, transhumeral, or transradial. In addition, information is sought on the technologies being used worldwide or what has been new in recent years, such as artificial intelligence or actuators to carry out sensory feedback.

### Search strategy

Data sources such as Scopus, Web of Science, and PubMed will be used for searches. The search terms are: a) Proprioception, b) Sensory feedback, and c) Prosthesis, with which the basic searches are done with the different search engines. The documents to search are in English, and we will retrieve records from the last 20 years.

### Study selection

The Selection Process for the studies found will include a review of titles, abstracts, full text, methodology, findings, conclusions, and future work. Finally, Mendeley will be the reference manager.

### Data extraction

The variables extracted from the included studies will be extracted from the documents related to proprioceptive feedback in the prosthesis, e.g., by level of amputation, proprioceptive feedback mechanisms, neural interface systems, and advancements in prosthetic control. Figure 1 provides a guide to extracting information for analysis, showing the two ways of classifying information related to the document's general and research data.

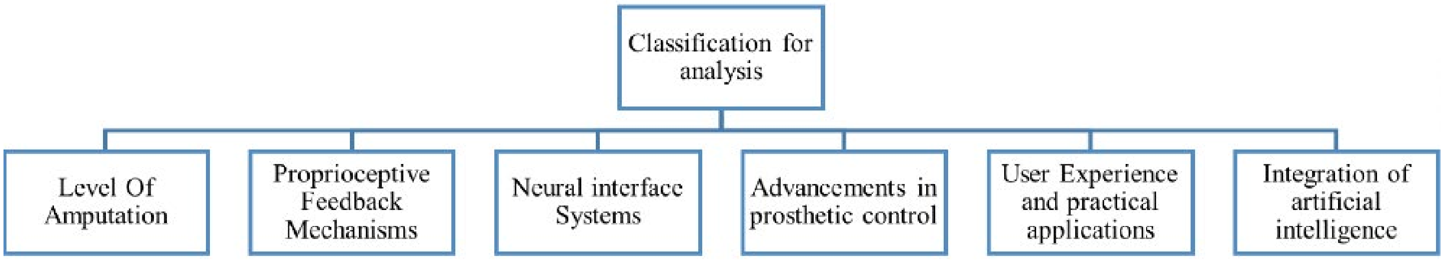


FIGURE 1. **A proposed data extraction classification according to general information and related to the theme of the review: Proprioception feedback in prosthesis.**

## RESULTS AND DISCUSSION

While searching for studies in Scopus, PubMed, and Web of Science, other databases, such as Nature or SpringerLink, produced repetitive results. The keywords used were proprioception, feedback, and prosthesis. Figure 2 shows a flowchart of the study selection.

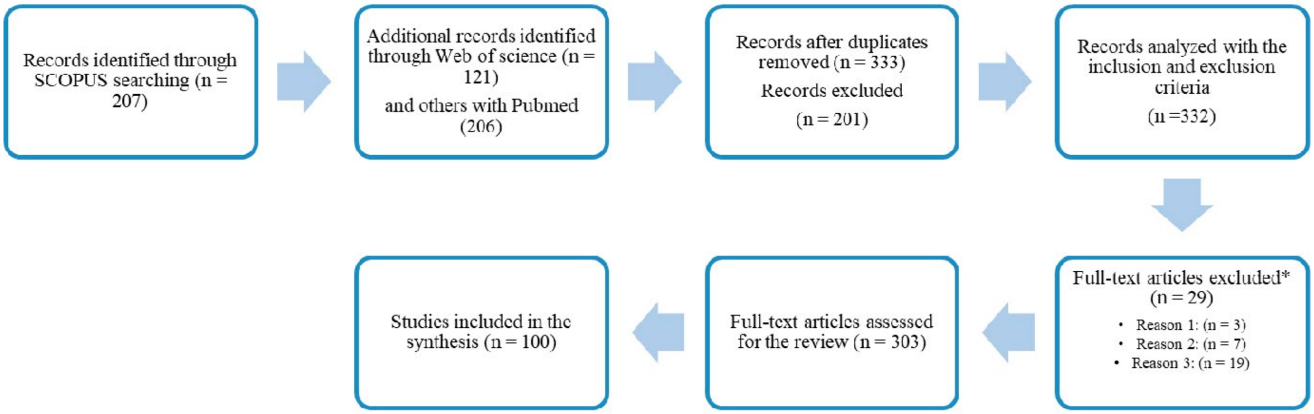


FIGURE 2. **Flowchart of study selection. * Reason 1: animal-related studies; Reason 2: other types of prostheses; Reason 3: different areas of study.**

We found 207 from SCOPUS, 206 from PubMed, and 121 from the Web of Science. Five hundred thirty-four documents were registered; however, 202 duplicates were removed, leaving 332 works for the database gathered in Mendeley.

## Analysis related to proprioceptive feedback in prosthesis

The studies were grouped into key themes: advances in upper and lower limb prosthetics, proprioceptive feedback mechanisms, neural interface systems, advancements in prosthetic control, user experience and practical applications, and the integration of artificial intelligence

## Upper limb

To transmit hand aperture or wrist rotation angle during sequential prosthesis control, a study by Dideriksen J., Siebold E., *et al.* (2024)[11] describes and assesses a feedback system with four vibration motors incorporated in the prosthesis socket. A functional task involving the manipulation of delicate objects with different compliance (with vibrotactile and/or visual or neither) was conducted by ten non-disabled and two amputee volunteers. The results for the amputee participants were similar because all participants perceived the vibrotactile feedback as helpful, dependable, and simple to perceive and utilize. However, the researchers noticed that it took longer to use the vibrotactile feedback than the visual. In conclusion, even when visual feedback is not entirely available, the proprioceptive feedback in this paper offers a valuable way to support object manipulation.

Research by Federico Masiero *et al.* (2024)[12] presents an example of a human-machine interface (HMI) that uses permanent magnets implanted in amputees' remaining muscles to control robotic limb prosthetics; they have called it a “pyrokinetic interface.” This interface uses the selected vibrations created by carefully regulated magnetic fields from external coils to activate muscle-tendon proprioceptors. A problem solved was the real-time tracking of several moving magnets under vibration. The outcomes demonstrate the feasibility of a system that can monitor and move several magnets in three dimensions, producing highly effective torsional vibrations at frequencies that provide the impression of movement.

The study of Yichen Han *et al.* (2023)[13] addresses the need for enhanced proprioceptive feedback in upper-limb prostheses, focusing on the position and movement of a prosthetic wrist. An electrotactile scheme was developed to encode these proprioceptive cues, and an experimental platform was designed for testing. Preliminary experiments determined sensory and discomfort thresholds, followed by two main proprioceptive feedback experiments: position sense and movement sense. The study demonstrates that the electrotactile stimulation scheme can effectively provide proprioceptive feedback for the position and movement of a prosthetic wrist.

Matthieu Guémann and Christophe Halgand *et al.* (2022)[14] evaluate vibrotactile feedback for myoelectric control of virtual elbow in prosthetic users, comparing performance in healthy subjects and transhumeral amputees. They interviewed sixteen healthy participants and seven transhumeral prosthesis users who performed myoelectric control of a virtual arm under four different conditions of feedback: Vision alone (VIS), Vibration alone (VIB), Vision plus vibration (VIS + VIB), and No feedback (NO). The study measured reach accuracy through angular errors during discrete and continuous movements. With extended training, the effectiveness of VIB alone is expected to improve, potentially reducing reliance on vision for closed-loop prosthesis control.

Marasco PD, Hebert JS, Sensinger JW, *et al.* (2021)[15] presented a complex neurorobotic touch feedback system for the prosthetic hand that involves advanced modifications and precise engineering to provide comprehensive sensory feedback. They investigated how the neurorobotic fusion of these sensory modalities in bionic upper limbs can enhance functional integration with the user's neural and sensory systems. Participants in this study previously underwent targeted reinnervation for proximal limb amputation and were habituated to myoelectric prosthesis usage. Each participant was fitted with an experimental prosthesis using modified commercial components integrated with sensors. Small, robust robotic four-bar haptic touch tactors were installed in the prosthetic socket and transmitted proportional pressure and contact transients (tap detection) from the hand to the appropriate touch percept sites in the reinnervated skin. These tactors generate up to 10 N forces with a 10-ms latency closed-loop position control, ensuring rapid and accurate feedback. Monitoring and analyzing brain activity revealed that participants exhibited brain activation patterns similar to those observed with natural limb use, indicating a more intrinsic and natural control over the bionic limbs. This enhanced sensory integration led to improved motor control and more intuitive use of the prosthetic limbs, with participants reporting a more natural and seamless interaction with their bionic limbs. The findings suggest that the neurorobotic fusion of sensory feedback promotes intrinsic brain behaviors, potentially leading to significant advancements in the design and functionality of bionic limbs, thereby improving the quality of life for prosthetic users. Developing a neurorobotic interface that integrates touch, kinesthesia, and movement feedback mechanisms was central to the study, involving individuals using bionic upper limbs.

Shiyong Su *et al.* (2023) conducted a thorough investigation into the brain mechanisms underpinning the integration of sensory feedback in myoelectric prosthesis control[16]. Their study, which comprised fifteen participants doing standard prosthesis control tasks, showed that visual feedback is crucial for manipulating blocks and grasping force control. The study also demonstrated the importance of tactile feedback for proprioceptive location perception tasks. Confirmed by concurrent EEG recordings and behavioral evaluations, these results offer vital insights into the significance of multisensory integration for efficient prosthesis control.

The study by Nikita Piliugin *et al.* (2024)[17] emphasizes how the modulation of PNS parameters—specifically stimulation frequency and pulse width—affects evoked sensations, including their intensity, naturalness, and impact on phantom limb pain suppression. The authors employed a combination of site mapping and impulse mapping techniques with two transhumeral amputees, each equipped with implanted cylindrical electrodes on the median nerve. These methods allowed the researchers to establish correlations between electrode stimulation zones and perceived sensations and evaluate how different spectral parameters of stimulation influenced the quality of feedback. This study contributes to the growing literature on neuromodulation in neuroprosthetics, offering a novel approach to mapping and analyzing evoked sensory feedback through behavioral data and advanced computational methods. However, its small sample size and variability among subjects signal the need for further research to generalize findings and refine PNS techniques.

In 2022, Enzo Romero and Dante A. Elias[18] published a conference paper that presented the conceptual design of a haptic palmar-finger feedback system for a transradial myoelectric upper limb prosthesis that allows an amputated person to acquire the sensations related to force-gripping, object-sliding, and pressure of the prosthetic fingers. The designed system has a monitoring unit arranged on the prosthetic hand and an actuation unit embedded in a bracelet around the user's forearm; they use tree vibration modules. It was a design exercise, with no manufacturing nor patient test.

The study of Marco Gallone and Michael D. Naish (2022)[19] examines the development and evaluation of a head-worn Wearable Haptic Feedback Device (WHFD) designed to transmit sensory information from upper-limb prostheses. The suggested WHFD is a skullcap worn on the head that has 30 vibratory units stitched into it. A 14-week study involving 18 participants explored the learning process associated with interpreting haptic patterns conveying joint proprioception. The study compared three different haptic stimulation methods, revealing insights into the effectiveness and potential of each approach for enhancing sensory feedback in prosthetic users. Participants in this work demonstrated significant learning and improved ability to interpret the haptic information throughout the study. The spatiotemporal stimulation group showed a slight advantage in interpreting the haptic patterns compared to the other groups. The spatiotemporal stimulation refers to a sensory feedback technique that integrates spatial and temporal patterns to encode information through haptic signals.

In this context, vibratory actuators activate in a sequential, sweeping pattern, where the sequence's duration and direction convey specific data, such as the magnitude and motion of a joint. This method allows complex proprioceptive information to be represented effectively, leveraging the brain's ability to interpret dynamic patterns, as demonstrated in applications like prosthetic control.

Trujillo *et al.* (2022)[20] proposed a skin-stretching actuator to transmit proprioceptive information to a person with an amputation to provide feedback, and they evaluated the efficacy of two fixation methods using a longitudinal skin-stretching haptic device. One is neoprene foam, and the second is a double-sided adhesive tape. Ten participants without amputation were interviewed for the study, and the proprioceptive information was transmitted using a skin-stretching actuator. Two quantitative surveys were conducted based on the mirror box to measure the effectiveness of the fixation methods. They found no statistically significant differences between neoprene foam and double-sided adhesive tape. However, neoprene foam was preferred due to its non-adhesive nature and was perceived as a more natural stimulus. These authors suggest foam is a viable option to transmit sensory feedback more naturally. As another university proposal, this conference paper didn’t evaluate the device with prosthetic users or amputees but suggested further research should focus on long-term studies and testing with amputee participants to confirm these findings.

In 2022, Magbagbeola *et al.*[21] investigated how vibration patterns can improve the perception of tactile information in prosthetic limbs, aiding in the long-term use of prosthetics and neuropathic pain management. The researchers employed a deep-learning algorithm to categorize the dissipation of vibration artifacts in Electromyographic (EMG) signals. Using two vibration motors, four different texture patterns were applied to seven participants in the experiment; each pattern was repeated three times. After post-processing, each participant's unseen data was effectively classified using a Recurrent Neural Network (RNN) to identify the artifact features across equidistantly separated EMG electrodes. By enhancing the precision and usability of sensory feedback in prosthetic devices, this effort may lead to a higher rate of long-term adoption. Nevertheless, the study was on non-disabled persons.

Cha *et al.* (2022)[22] presented a closed-loop control system for robotic prosthetic hands, combining EMG-based intention recognition with proprioceptive feedback to enhance control. This study explores the power of a robotic prosthetic hand by combining intention recognition via Electromyography (EMG) classification with sensory feedback through a rule-based haptic device. A Convolutional Neural Network (CNN) model was designed to classify EMG signals from multiple channels, achieving over 97 % accuracy in recognizing user intentions across ten different grip states. The integrated system, which merges the CNN-based EMG classification with the haptic feedback device, was evaluated on able-bodied subjects and demonstrated high accuracy in both intention recognition and sensory feedback.

Battaglia *et al.'s* (2017)[23] research addresses the challenge of restoring hand functionality in upper limb amputees using myoelectric prostheses, which often lack intuitive control and haptic feedback. To improve user experience, the authors introduce the Rice Haptic Rocker, a device designed to provide proprioceptive feedback through skin stretch, integrated with the Pisa/IIT Soft Hand. The results showed that the device is a feasible tool for enhancing proprioceptive feedback in prosthetic hands, improving task performance requiring object size discrimination.

The study of Mulvey *et al.* (2014)[24] investigates how the perceptual incorporation of an artificial limb can improve the manual control of prosthetic devices. It studies explicitly whether transcutaneous electrical nerve stimulation (TENS) can simplify the perceptual embodiment of artificial limbs. Findings discovered that combining visual, tactile, and TENS stimuli significantly heightened the intensity of perceptual embodiment, with the most noticeable effect occurring when all three stimuli were used together. Additionally, the strength of this effect augmented over time. The study concluded that TENS can modestly enhance the sensation of embodiment in artificial limbs.

Papalos *et al.*[25] wrote a review in 2023 about proprioceptive feedback in upper limb prostheses using non-invasive approaches, emphasizing the uses and difficulties related to proprioceptive restitution in upper limb prostheses. This work establishes that when an artificial stimulus is given to a user, and it comes from the same sensory system and modality as the missing information, it is said to be homo-modal feedback (e.g., transmitting touch with devices that provide pressure feedback). On the other hand, hetero-modal feedback requires a sensory channel that differs from the one used physiologically (producing angular movement through hearing, for example) or uses the same channel but modifies the input stimulus's modality (producing limb position through vibration instead of skin stretch, for example). It turns out hetero-modal stimulation is less intuitive than homo-modal techniques, making it a valuable option for feedback restitution.

Lecompte *et al.* (2024)[26] provide a vision focused on proprioceptive feedback approaches for upper-limb myoelectric prostheses. It deals with numerous methods for incorporating proprioception into prosthetic devices and explains the significance of this capacity. The techniques enlist the most common so far in the literature: a) skin stretching, b) auditory input, c) electrotactile stimulation, d) kinesthetic illusions, e) direct brain stimulation, f) vibrotactile stimulation, and g) intracortical microstimulation. The document emphasizes the downsides and restrictions of these techniques, mostly the struggle of obtaining real-time, non-invasive, anatomically consistent feedback that closely resembles sensory input from the environment. For example, the ineffectiveness of embodiment systems to perfectly imitate sensory integration may intensify the cognitive load necessary during the use, as well as uncomfortable sensations and even distress.

### Lower Limb

Petersen *et al.* (2023)[27] explored the relationship between somatosensory impairments and functional performance in individuals with lower-limb amputation. Despite significant differences in balance and gait between amputees and able-bodied controls, the study found that existing clinical measures were insufficient to differentiate between levels of sensory impairment within the amputee group. This suggests that more sophisticated and challenging metrics are necessary to accurately assess sensory impairments' effects on functional abilities.

Canton Leal *et al.* (2022)[28] developed an innovative haptic feedback system called HapticLink, designed to enhance balance and proprioception in individuals with lower-limb amputations. The system, which uses force sensors and vibration motors to convey weight distribution information, was assessed with promising results. This development highlights the potential of haptic feedback technology to improve the quality of life for amputees by providing them with enhanced sensory feedback and greater control over their prosthetic limbs.

Di Zubiena *et al.* (2022)[29] focused their research on the static characterization of a novel stretchable strain sensor created through 3D printing, aiming to restore proprioception in lower-limb amputees. The sensor, which combines an elastomeric material with a metal alloy sensitive to deformation, demonstrated excellent sensitivity, repeatability, and response to strain. These characteristics make it a promising candidate for developing wearable proprioceptive devices that could significantly enhance balance and gait stability in amputees. Also, Zubiena et al. (2021)[30] conducted a Finite Element Modeling (FEM) analysis to investigate the potential application of an elastomeric strain sensor for restoring proprioception in transtibial prostheses. The study found areas of maximum deformation within the prosthesis during gait, which could be best for sensor placement. These findings are encouraging, as they support the future development of proprioceptive feedback devices that could improve balance, gait stability, and overall mobility for lower-limb amputees.

Gardetto *et al.* (2021)[31] presented a case series demonstrating the effectiveness of Targeted Sensory Reinnervation (TSR) in reducing phantom limb pain and improving proprioception in patients with lower-limb amputations. The surgical technique involved rewiring sensory nerves and pairing them with a specialized prosthetic device that provided sensory feedback from the prosthesis. Remarkably, the intervention resulted in significant pain reduction, with some patients becoming completely pain-free and others able to discontinue pain medication. This study provides compelling evidence for the benefits of TSR, particularly when combined with advanced prosthetic technologies, in enhancing the quality of life for amputees.

Foster *et al.* (2020)[32] examined the accuracy and precision of foot placement during a targeted stepping task in Individuals with Unilateral Transtibial Amputation (IUTAs). The study revealed that these individuals exhibited reduced accuracy and precision compared to able-bodied controls, particularly with their intact limbs. This finding suggests that the disruption of sensory information and the characteristics of prosthetic components may contribute to difficulties in dynamic balance and foot placement during everyday activities.

Charkhkar *et al.* (2020)[33] investigated how sensory neuroprostheses affected the balance of amputees who had lost limbs. The device mimics sensory feedback equivalent to plantar pressure beneath prosthetic feet by placing non-penetrating cuff electrodes around the remaining nerves. According to the results obtained from two transtibial amputees, the neuroprostheses significantly improved postural stability when ocular and intact leg somatosensory inputs were disrupted. This represents a breakthrough in prosthetic technology since it implies that neuroprostheses may enhance balance and lower the risk of falls.

Christie *et al.* (2019)[34] investigated the temporal perception of stimulation-induced sensations in amputees, mainly focusing on how these sensations synchronize with visual cues. The study found that stimulation-induced sensations can be perceived as synchronous with vision, similar to natural somatosensation when timed correctly. This research provides valuable insights into the design of sensory neuroprostheses, emphasizing the importance of temporal alignment in creating sensations that feel natural to the user.

Coker *et al.* (2019)[35] used computational modeling to compare the stimulation artifacts produced by different Peripheral Nerve Interfaces (PNIs) used in prosthetic limbs. Their findings indicated that micro-channel sieve electrodes generated fewer artifacts than other configurations, such as thin-film Transverse Intrafascicular Multichannel Electrodes (tfTIMEs). This reduction in artifacts is crucial for achieving concurrent sensory feedback and motor control in neuroprosthetics.

Plauché *et al.* (2016)[36] proposed that providing feedback based on the Center Of Pressure (COP) under the prosthetic foot can enhance proprioception and improve phase sensing in above-knee amputees. Their work introduces a device that delivers vibrotactile feedback derived from the COP of the prosthesis, aiming to restore proprioception and enhance phase awareness. Experiments conducted with novice users (non-disabled individuals) of a transfemoral prosthetic leg showed that the device significantly reduced variability in stride length, step width, and trunk sway during treadmill walking. This suggests that the haptic device effectively improves gait stability in users.

Plauché *et al.* (2016)[36] proposed that providing feedback based on the Center Of Pressure (COP) under the prosthetic foot can enhance proprioception and improve phase sensing in above-knee amputees. Their work introduces a device that delivers vibrotactile feedback derived from the COP of the prosthesis, aiming to restore proprioception and enhance phase awareness. Experiments conducted with novice users (non-disabled individuals) of a transfemoral prosthetic leg showed that the device significantly reduced variability in stride length, step width, and trunk sway during treadmill walking. This suggests that the haptic device effectively improves gait stability in users.

Yang *et al.* (2012)[37] studied a real-time feedback system called the Lower Extremity Ambulatory Feedback System (LEAFS), designed to improve gait symmetry in individuals with transtibial amputation. The system provides auditory feedback to correct asymmetries in gait, with promising results for rehabilitation gait asymmetries, with promising rehabilitation results. LEAFS is a wearable wireless gadget that uses the stance time symmetry ratio between the right and left limbs to generate real-time aural feedback. The outcomes were inconsistent; two individuals had notable increases in gait symmetry. The results imply that LEAFS may help people with transtibial amputations achieve better gait symmetry, despite the small sample size. The study highlights an inconsistency in outcomes among the three participants, explicitly noting that while two subjects showed marked improvements in gait symmetry and trunk sway, the third did not demonstrate any objective enhancements. This inconsistency is attributed to individual differences, such as residual limb sensitivity or adaptation capabilities. For instance, the third subject exhibited numbness in the intact limb, potentially limiting their ability to respond to the feedback system. Such variability underscores the need for further investigation with larger sample sizes to account for individual factors influencing the effectiveness of the LEAFS system.

Ghiami *et al.* (2024)[38] report research on the sensorimotor parameters related to powered lower limb prostheses through walking movements of individuals with ankle amputation. Based on a review of 29 articles, the study outlines how amputees struggle to feel sensations through disintegrated nerves. This work focuses on prosthetic knees as the most capable component of a mechanically passive prosthesis for minimum energy expenditure walking. It focuses on proprioception created by integrated mechatronic systems to assist nervous rehabilitation and improve movement accuracy. This review includes our work from 2023[10].

## Proprioceptive feedback mechanisms

A substantial part of the research applications is developing and refining proprioceptive feedback mechanisms, and in this section, we highlight some of the most representative ones. Early studies, such as those by Wall and Kentala (2005)[39], explored vibrotactile feedback to aid postural control in patients with deficits. Farrell *et al.*[40] examined the effects of static friction and backlash on the control of powered prostheses, emphasizing the importance of feedback in extended physiological proprioception. Later research delved into more sophisticated feedback systems. Kuchenbecker *et al.* (2007, 2009)[41][42] and Blank *et al.* (2008)[43] investigated the effects of visual and proprioceptive feedback on human control of targeted movements and virtual hand prostheses, respectively. These studies laid the groundwork for understanding how sensory cues can enhance prosthetic control.

In recent years, research has continued to evolve with studies like those by Plauche *et al.* (2016)[36] and Wendelken *et al.*[44], which presented advanced haptic feedback systems for prosthetic leg users and the restoration of motor control and sensation in amputees using Utah Slanted Electrode Arrays. Lima and Hammond[45] further advanced the field by examining simultaneous rotary skin stretch and vibrotactile stimulation for proprioceptive feedback, and the findings showed that participants could identify the dial angle when skin stretch feedback was provided. Additionally, Mablekos-Alexiou *et al.* (2015)[46] suggested a biomechatronic system that uses Extended Physiological Proprioception (EPP), a form of subconscious sensory feedback, to present a novel method of operating multi-joint prostheses. This architecture activates an implanted micro servo actuator, which offers similar control capabilities without the functional and aesthetic limitations of older approaches that rely on Bowden cables and cineplasty. The authors imply that this method will be more acceptable to users and could be a basis for more sophisticated and highly controllable multi-degree-of-freedom prosthetic devices.

## Neural interface systems

Neural interface systems represent another critical area of research, and in this section, we highlight some of the most representative ones. Weber *et al.* (2012)[47] highlighted essential considerations for interfacing the somatosensory system to restore touch and proprioception. Ramos-Murguialday *et al.* (2012)[48] explored brain-computer interface-based neuroprostheses with proprioceptive feedback, providing a foundation for integrating neural signals into prosthetic control. Studies by Gaunt *et al.* (2009)[49] and Tabot *et al.* (2015)[50] examined the microstimulation of primary afferent neurons and the restoration of tactile and proprioceptive sensation through brain interfaces. This study investigates the long-term stability of intracortical microstimulation (ICMS) as a method for providing sensory feedback in upper limb neuroprostheses. Tabot experimented with non-human primates that could detect ICMS, which remained stable over the years, even with extensively used electrodes. These findings suggest that ICMS could be a viable and reliable approach for restoring somatosensation in neuroprosthetic devices, potentially improving their usability and effectiveness. Other recent research, such as that by Srinivasan *et al.* (2021)[51], explored the implementation of regenerative agonist-antagonist myoneural interfaces for preserving joint function and perception in above-knee amputations. This line of research underscores the potential for advanced neural interfaces to improve prosthetic devices' functionality and user experience. All these studies contributed to the understanding of how neural stimulation can be used to enhance sensory feedback in prosthetics.

## Advancements in prosthetic control

The research also highlights significant advancements in prosthetic control mechanisms, and in this section, we highlight some of the most representative ones. Kuiken *et al.* (2007)[52] introduced targeted reinnervation, a technique to redirect cutaneous sensation, enhancing sensory feedback in amputees. Subsequent studies, such as those by Li and Kuiken (2008)[53] and Akhtar *et al.* (2014)[54], focused on modeling prosthetic limb rotation control and proprioception with numerous degrees of freedom via passive mechanical skin stretching, respectively. Brown *et al.* (2015)[55] and Schiefer *et al.* (2018)[56] explored the regulation of grip force control using a myoelectric prosthesis with low impedance and the improvement of object identification tasks through artificial tactile and proprioceptive feedback. These studies demonstrate the ongoing efforts to refine control mechanisms to enhance prosthetics' functionality and user experience. Rouse *et al.* (2011)[57] introduce the Osseo-Magnetic Link (OML), which is a unique control system intended to maintain in prosthetic devices, notably for humeral or wrist rotation. The OML system places sensors in the prosthetic socket to measure magnetic field vectors and implant a magnet within the residual bone. With this configuration, people can rotate their bones voluntarily to operate a prosthetic rotator.

## User experience and practical applications

A key theme in the research is the focus on user experience and practical applications, and in this section, we highlight some of the most representative ones. For example, Cuberovic *et al.* (2019)[58] emphasized the long-term home use of sensory-enabled prostheses, displaying real-world applicability and user adaptation. Marasco *et al.*[59] showed that illusory movement perception could enhance prosthetic hands' ability to manage their motor function, highlighting the importance of creating intuitive and effective user interfaces. Studies like those by Sienko *et al.* (2018)[60] and Bates *et al.* (2020)[61] investigated the potential mechanisms of sensory augmentation systems on human balance control and the technological advances in prosthesis design and rehabilitation. These studies contribute to a better understanding of how prosthetic devices can be optimized for user comfort and effectiveness.

## Integration of artificial intelligence

Integrating Artificial Intelligence (AI) into prosthetic systems represents an innovative area of research; in this section, we highlight some of the most representative ones. For example, Luu *et al.* (2022)[62] explored how AI enables real-time and intuitive control of prostheses via nerve interfaces. The neuroprosthetic system presented in this paper uses a Recurrent Neural Network RNN-based artificial intelligence agent to decode movement intent for amputees in real-time from peripheral nerve signals. Experiments with three human amputees showed that technology allows intuitive control of a prosthetic hand with up to 97-98 % accuracy in individual finger and wrist movements. The AI agent's long-term resilience was verified over 16 months, and its real-time performance was verified using assessments of reaction time and information throughput. The results show how AI-enabled nerve technology can be used to create the next generation of prosthetic hands that are intuitive and dexterous. Other recent studies, such as those by Vargas *et al.*[63] and Berger *et al.*[64], have explored closed-loop control of prosthetic fingers via evoked proprioceptive information and the use of AI for texture recognition based on multi-sensory integration of proprioceptive and tactile signals. These studies highlight the transformative potential of AI in the field of prosthetics. This topic proves the potential for AI to significantly enhance the functionality and user experience of prosthetic devices by providing more natural and responsive control mechanisms applying sensory feedback.

## Our work in Mexico

In 2019, the Universidad Nacional Autónoma de México (UNAM) started offering a bachelor's degree in orthotics and prosthetics with a curriculum aligned with the country's needs. Amongst the university's goals are research, design, and technological development in this transdisciplinary area. Regarding sensory feedback for prosthesis users, we have progressed with the Artificial Proprioception proposal in work published in 2023[10]. We were trying to fill a truly relevant perspective in the current state of the art about the proprioceptive feedback applied to prostheses. The aim was and still is to integrate biomechatronic devices that imitate the sensory feedback lost due to limb loss. In our paper of 2023, we focus the application on the rehabilitation of amputees, active or passive, because proprioceptive feedback constitutes a significant obstacle in this field.

## Discussion: interpretation of the main findings

### *Neural Interface Systems*

One of the most significant and essential prosthetic advancements is neural interface systems that directly interface with the user's nervous system to provide sensory feedback. These systems are critical for the user to feel the sensation used to manage and command prosthetic devices more naturally and intuitively. For example, Weber *et al. (*2012)[47] emphasized important considerations for interfacing with the somatosensory system to restore touch and proprioception, such as electrode design and the need for selective activation of specific neural pathways and drawing on experience from other neuroprosthetic systems. In addition, Srinivasan *et al.* (2017, 2019, and 2021)[51][65][66][67][68] developed and published the use of regenerative agonist-antagonist myoneural interfaces to maintain joint function and perception in above-knee amputations. This is a massive leap into what is described as "natural-like" since the approach not only restores partial levels of natural sensations but also maintains the structural integrity of the residual limb, making good use of a prosthetic device.

### *Advancements in prosthetic control mechanisms*

Prosthetic limb control has reached a new dimension by incorporating sophisticated feedback mechanisms that send actual sensations to the user in real time. Kuiken *et al.* (2027, 2008, and 2011)[52][53][57] showcased targeted reinnervation, a method through which cutaneous sensation is rerouted to enhance sensory input for amputees. The control method has been crucial in optimizing the functionality of myoelectric prostheses to enable users to execute skills with increased precision and confidence.

Further studies, as conducted by Plauche *et al.* (2016)[36], for instance, introduced a haptic feedback system for above-knee prosthetic leg users that substantially enhanced gait stability through the provision of vibrotactile feedback related to the center of pressure under the prosthetic foot. The innovation underlines that sensory feedback is crucial for upper-limb prosthetics and lower-limb devices, in which the ability to maintain balance and gait is essential.

### *Artificial Intelligence and Machine Learning in prosthetics*

AI integration has opened new frontiers in prosthetic systems' functional and user experience-based development. Artificial intelligence allows for real-time control over prostheses intuitively and exquisitely responsive to the user's intentions. For example, Luu *et al.* (2022)[62] demonstrated how an AI-based system could use Neural Networks to decode movement intention from peripheral nerve signals with high accuracy and enhance prosthetic hand control. This AI-driven approach makes the prosthetic device more dexterous, self-improving, and adjusted to the user's needs. The application of AI in this matter does much good, as it can learn and improve with time, hence giving out prosthetic devices that are more considerate of human movement.

### *Proprioceptive feedback mechanisms*

Effective proprioceptive feedback mechanisms have been central in prosthetics. The position sense and movement of the prosthesis are part of activities involving fine motor skills. Research by Dideriksen *et al.* (2020, 2023, and 2024)[11][69][70] introduced a feedback system that uses vibrotactile stimulation to convey proprioceptive information, and this indeed enhances the manipulation of objects with the prosthetic hand.

Besides that, electrotactile feedback has also been explored as a method to enhance the proprioceptive input from upper-limb prostheses. Using a developed electrotactile stimulation scheme that could effectively convey information about the flexion-extension position of a prosthetic wrist, Han *et al.*[13] demonstrated how such technology could enhance the precision of prosthetic control.

### *Innovative feedback devices and user experience*

There is a growing interest in developing feedback devices that provide users with more natural and intuitive sensory feedback. Such examples include the Rice Haptic Rocker, designed by Battaglia *et al.* (2017 and 2019)[23][71][72], which integrated skin stretch feedback into a myoelectric prosthesis and significantly improved the perception of object sizes. This device stands out as an ideal example of how novel feedback mechanisms can substantially enhance functionality and usability for prosthetic devices.

Marasco *et al.* (2021)[15] introduced a neurorobotic system that combines the sense of touch with prosthetic control, allowing users to feel a more seamless and natural interaction with their prosthetic limbs. This system's added benefit includes enhanced motor control in users while providing an improved sense of embodiment, whereby the prosthetic limb feels like a natural body extension.

### *Implications for clinical practice*

The advances in prosthetics related to artificial proprioception have profound implications for clinical practice by facilitating rehabilitation and improving the quality of life for individuals with limb loss. There are areas where the platform could improve clinical outcomes when it develops further. The most critical clinical implication would be the possibility of more personalized prosthetic solutions. For example, advanced neural interface systems, AI-driven prosthetics, and customized feedback mechanisms like vibrotactile and electrotactile feedback will allow clinicians to better match prosthetics with each patient's needs. These technologies restore sensory feedback to the user and enable better motor control and balance, allowing for smoother movement with reduced cognitive load required to control prosthetic devices.

## CONCLUSIONS

Advancing artificial proprioception for prosthetics represents a leap into the future, where biomedical engineering will play an essential role in ensuring a more excellent livelihood for amputees. The stride of research has been underpinned by high underpinning in collaboration across disciplines, technological innovation, and raising the quality of life for prosthetics users. At the heart of this advancement has been embedding proprioception and other forms of sensory feedback into the prosthetic device, which allows users to regain a sense of normalcy in their daily activities by extending functionality, control, and overall usability.

The development in this field fundamentally highlights the role played by neural interface systems, involving direct interaction with the user's nervous system to restore sensation and proprioception. These systems have shown great promise in enhancing the intuitiveness and responsiveness of prosthetic limbs and providing users with more significant, more natural control over their movements. In the same way, integrating AI and machine learning into prosthetic systems has opened new perspectives for real-time, personalized control to make the prosthetic device more adaptable to each user's specific needs.

The counterpoint for lower-limb devices besides upper-limb prosthetics was when the proprioceptive feedback mechanisms are essential to maintain balance and gait. Vibrotactile feedback systems and more recent electrotactile stimulations strongly enhanced stability and functionality in lower-limb prosthetics, enhancing mobility in amputees and instilling more confidence in the activities.

Surgical advances must incorporate proprioceptive feedback and increased control to maximize the functional utility of prosthetic use. These findings highlight the necessity for continued innovation in more natural prostheses to improve life quality and rehabilitation.

## ACKNOWLEDGMENTS

The author thanks the support to DGAPA of UNAM, given by the projects PAPIIT TA100622 and TA100225, during the realization of this project. I sincerely thank Dr. Jesús Manuel Dorador González for his valuable support in reviewing the English language of this manuscript, whose contribution was essential in improving the clarity and quality of the text.

## AUTHOR CONTRIBUTIONS

O. D.-H. conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology, project administration, software, resources, supervision, validation, visualization, writing-original draft, and writing-review & editing.

## REFERENCES

**[1]** R. A. Magill and D. I. Anderson, "Sensory components of motor control," in Motor Learning and Control: Concepts and Applications, 12th ed, New York, NY: McGraw Hill, 2021. [Online]. Available: **https://accessphysiotherapy.mhmedical.com/content.aspx?bookid=3082&sectionid=256573379**.

**[2]** V. E. Abraira and D. D. Ginty, "The sensory neurons of touch," Neuron, vol. 79, no. 4, pp. 618–639, 2013, doi: **https://doi.org/10.1016/J.NEURON.2013.07.051**

**[3]** R. A. Magill and D. I. Anderson, "Augmented feedback," in Motor Learning and Control: Concepts and Applications, 12th ed., New York, NY: McGraw Hill, 2021. [Online]. Available: **https://accessphysiotherapy.mhmedical.com/content.aspx?bookid=3082&sectionid=256574674**

**[4]** WHO standards for prosthetics and orthotics, World Health Organization and USAID, Geneva, 2017. [Online]. Available: **https://www.who.int/publications/i/item/9789241512480**

**[5]** Federación Mexicana de Diabetes. "Estadísticas en México." Federación Mexicana de Diabetes, A. C. Accessed: Jun. 30, 2024. [Online]. Available: **https://fmdiabetes.org/estadisticas-en-mexico/**

**[6]** I. J. Ascencio-Montiel, "10 years analysis of diabetes-related major lower extremity amputations in México," Arch. Med. Res., vol. 49, no. 1, pp. 58-64, 2018, doi: **https://doi.org/10.1016/j.arcmed.2018.04.005**

**[7]** M. A. Chahrour et al., "Major lower extremity amputations in a developing country: 10-Year experience at a tertiary medical center," Vascular, vol. 29, no. 4, pp. 574-581, 2021, doi: **https://doi.org/10.1177/1708538120965081**

**[8]** R. R. Riso, "Strategies for providing upper extremity amputees with tactile and hand position feedback - Moving closer to the bionic arm," Technol. Health Care, vol. 7, no. 6, pp. 401–409, 1999, doi: **https://doi.org/10.3233/thc-1999-7604**

**[9]** S. Raspopovic, G. Valle, and F. M. Petrini, "Sensory feedback for limb prostheses in amputees," Nat. Mater., vol. 20, no. 7, pp. 925–939, 2021, doi: **https://doi.org/10.1038/s41563-021-00966-9**

**[10]** O. Diaz-Hernandez and I. Salinas-Sanchez, "Towards artificial proprioception in prosthetic devices," Int. J. Med. Sci., vol. 10, no. 1, pp. 1–5, 2023, doi: **https://doi.org/10.14445/23939117/ijms-v10i1p101**

**[11]** J. Dideriksen, E. Siebold, S. Dosen, and M. Markovic, "Investigating the benefits of multivariable proprioceptive feedback for upper-limb prostheses," IEEE Trans. Med. Robot. Bionics, vol. 6, no. 2, pp. 757–768, 2024, doi: **https://doi.org/10.1109/TMRB.2024.3385983**

**[12]** F. Masiero, E. La Frazia, V. Ianniciello, and C. Cipriani, "Generating frequency selective vibrations in remote moving magnets," Adv. Intell. Syst., vol. 6, no. 6, 2024, art. no. 2300751, doi: **https://doi.org/10.1002/aisy.202300751**

**[13]** Y. Han et al., "Substitutive proprioception feedback of a prosthetic wrist by electrotactile stimulation," Front. Neurosci., vol. 17, 2023, art. no. 1135687, doi: **https://doi.org/10.3389/FNINS.2023.1135687**

**[14]** M. Guémann et al., "Sensory substitution of elbow proprioception to improve myoelectric control of upper limb prosthesis: experiment on healthy subjects and amputees," J. Neuroeng. Rehabil., vol. 19, 2022, art. no. 59, doi: **https://doi.org/10.1186/S12984-022-01038-Y**

**[15]** P. D. Marasco et al., "Neurorobotic fusion of prosthetic touch, kinesthesia, and movement in bionic upper limbs promotes intrinsic brain behaviors," Sci. Robot., vol. 6, no. 58, 2021, art. no. eabf3368, doi: **https://doi.org/10.1126/scirobotics.abf3368**

**[16]** S. Su et al., "Neural evidence for functional roles of tactile and visual feedback in the application of myoelectric prosthesis," J. Neural Eng., vol. 20, no. 1, 2023, art. no. 016038, doi: **https://doi.org/10.1088/1741-2552/ACAB32**

**[17]** N. Piliugin et al., "Unraveling Sensory Restoration: Decoding Complex Behavioral Data through Peripheral Nerve Stimulation Analysis," in 12th Int. Winter Conf. Brain-Comput. Interface (BCI), Gangwon, Corea, 2024, pp. 1-4, doi: **https://doi.org/10.1109/BCI60775.2024.10480475**

**[18]** E. Romero and D. A. Elias, "Design of a haptic palmar-finger feedback system for upper limb myoelectric prosthesis model PUCP-Hand," in 2022 Int. Conf. Elect. Comput. Energy Technol. (ICECET), Prague, Czech Republic, 2022, pp. 1-6,2022, doi: **https://doi.org/10.1109/ICECET55527.2022.9872575**

**[19]** M. Gallone and M. D. Naish, "Scalp-Targeted Haptic Proprioception for Upper-Limb Prosthetics," in 2022 Int. Conf. Rehabil. Robot. (ICORR), Rotterdam, Netherlands, 2022, pp. 1-6, doi: **https://doi.org/10.1109/ICORR55369.2022.9896551**

**[20]** E. Trujillo-Trujillo, K. Acuna-Condori, and M. G. S. P. Paredes, "Evaluation of the skin fixation method of a haptic device to provide proprioceptive information for a sensory feedback prosthesis," in 2022 IEEE ANDESCON, Barranquilla, Colombia, 2022, pp. 1-6, doi: **https://doi.org/10.1109/ANDESCON56260.2022.9989931**

**[21]** M. Magbagbeola, M. Miodownik, S. Hailes, and R. C. V. Loureiro, "Correlating vibration patterns to perception of tactile information for long-term prosthetic limb use and continued rehabilitation of neuropathic pain," in 2022 Int. Conf. Rehab. Robot. (ICORR), Rotterdam, Netherlands, 2022, pp. 1-6, doi: **https://doi.org/10.1109/ICORR55369.2022.9896412**

**[22]** H. Cha et al., "Study on intention recognition and sensory feedback: control of robotic prosthetic hand through emg classification and proprioceptive feedback using rule-based haptic device," IEEE Trans. Haptics, vol. 15, no. 3, pp. 560–571, 2022, doi: **https://doi.org/10.1109/TOH.2022.3177714**

**[23]** E. Battaglia et al., "The rice haptic rocker: skin stretch haptic feedback with the Pisa/IIT softhand," in 2017 IEEE World Haptics Conference (WHC), Munich, Germany, 2017, pp. 7-12, doi: **https://doi.org/10.1109/WHC.2017.7989848**

**[24]** M. Mulvey, H. Fawkner, and M. I. Johnson, "An investigation into the perceptual embodiment of an artificial hand using transcutaneous electrical nerve stimulation (TENS) in intact-limbed individuals," Technol. Health Care, vol. 22, no. 2, pp. 157–166, 2014, doi: **https://doi.org/10.3233/THC-140780**

**[25]** E. D. Papaleo et al., "Integration of proprioception in upper limb prostheses through non-invasive strategies: a review," J. Neuroeng. Rehabil., vol. 20, no. 1, 2023, art. no. 118, doi: **https://doi.org/10.1186/S12984-023-01242-4**

**[26]** O. Lecompte, S. Achiche, and A. Mohebbi, "A review of proprioceptive feedback strategies for upper-limb myoelectric prostheses," IEEE Trans. Med. Robot. Bionics., vol. 6, no. 3, pp. 930–939, 2024, doi: **https://doi.org/10.1109/TMRB.2024.3407532**

**[27]** B. A. Petersen, P. J. Sparto, and L. E. Fisher, "Clinical measures of balance and gait cannot differentiate somatosensory impairments in people with lower-limb amputation," Gait Posture, vol. 99, pp. 104–110, 2023, doi: **https://doi.org/10.1016/J.GAITPOST.2022.10.018**

**[28]** J. M. Canton Leal, J. V. Gyllinsky, A. A. Arredondo Zamudio, and K. Mankodiya, "HapticLink: a force-based haptic feedback system for single and double lower-limb amputees," in 2022 44th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC), Glasgow, Scotland, United Kingdom, 2022, pp. 4226-4229, doi: **https://doi.org/10.1109/EMBC48229.2022.9871460**

**[29]** F. C. Gattinara Di Zubiena, L. D'Alvia, Z. Del Prete, and E. Palermo, "A static characterization of stretchable 3D-printed strain sensor for restoring proprioception in amputees," in 2022 IEEE International Conference on Flexible and Printable Sensors and Systems (FLEPS), Vienna, Austria, 2022, pp. 1-4, doi: **https://doi.org/10.1109/FLEPS53764.2022.9781497**

**[30]** F. C. G. Di Zubiena et al., "FEM deformation analysis of a transtibial prosthesis fed with gait analysis data: A preliminary step towards restoring proprioception in amputees," in

2021 IEEE International Workshop on Metrology for Industry 4.0 & IoT (MetroInd4.0&IoT), Rome, Italy, 2021, pp. 493-498, doi: **https://doi.org/10.1109/MetroInd4.0IoT51437.2021.9488482**

**[31]** A. Gardetto et al., "Reduction of phantom limb pain and improved proprioception through a TSR-based surgical technique: a case series of four patients with lower limb amputation," J. Clin. Med., vol. 10, no. 17, 2021, art. no. 4029, doi: **https://doi.org/10.3390/JCM10174029**

**[32]** R. J. Foster et al., "Individuals with unilateral transtibial amputation exhibit reduced accuracy and precision during a targeted stepping task," J. Biomech., vol. 105, 2020, art. no. 109785, doi: **https://doi.org/10.1016/J.JBIOMECH.2020.109785**

**[33]** H. Charkhkar, B. P. Christie, and R. J. Triolo, "Sensory neuroprosthesis improves postural stability during Sensory Organization Test in lower-limb amputees," Sci. Rep., vol. 10, no. 1, 2020, art. no. 6984, doi: **https://doi.org/10.1038/s41598-020-63936-2**

**[34]** B. P. Christie et al., "Visuotactile synchrony of stimulation-induced sensation and natural somatosensation," J. Neural. Eng., vol. 16, no. 3, 2019, art. no. 036025, doi: **https://doi.org/10.1088/1741-2552/AB154C**

**[35]** R. A. Coker, E. R. Zellmer, and D. W. Moran, "Micro-channel sieve electrode for concurrent bidirectional peripheral nerve interface. Part B: Stimulation," J. Neural. Eng., vol. 16, no. 2, 2019, art. no. 026002, doi: **https://doi.org/10.1088/1741-2552/aaefab**

**[36]** A. Plauche, D. Villarreal, and R. D. Gregg, "A haptic feedback system for phase-based sensory restoration in above-knee prosthetic leg users," IEEE Trans. Haptics., vol. 9, no. 3, pp. 421–426, 2016, doi: **https://doi.org/10.1109/TOH.2016.2580507**

**[37]** L. Yang et al., "Utilization of a lower extremity ambulatory feedback system to reduce gait asymmetry in transtibial amputation gait," Gait Posture, vol. 36, no. 3, pp. 631–634, 2012, doi: **https://doi.org/10.1016/J.GAITPOST.2012.04.004**

**[38]** A. Ghiami Rad and B. Shahbazi, "A systematic investigation of sensorimotor mechanisms with intelligent prostheses in patients with ankle amputation while walking," J. Mech. Behav. Biomed. Mater., vol. 151, 2024, art. no. 106357, doi: **https://doi.org/10.1016/j.jmbbm.2023.106357**

**[39]** C. Wall and E. Kentala, "Control of sway using vibrotactile feedback of body tilt in patients with moderate and severe postural control deficits," J. Vestib. Res., vol. 15, no. 5–6, pp. 313–325, 2005, doi: **https://doi.org/10.3233/VES-2005-155-607**

**[40]** T. R. Farrell, R. F. Weir, C. W. Heckathorne, and D. S. Childress, "The effects of static friction and backlash on extended physiological proprioception control of a powered prosthesis," J. Rehabil. Res. Dev., vol. 42, no. 3, pp. 327–341, 2005, doi: **https://doi.org/10.1682/jrrd.2004.05.0052**

**[41]** K. J. Kuchenbecker, N. Gurari, and A. M. Okamura, "Effects of visual and proprioceptive motion feedback on human control of targeted movement," in 2007 IEEE 10th Int. Conf. Rehab. Robot., Noordwijk, Netherlands, 2007, pp. 513-524, doi: **https://doi.org/10.1109/ICORR.2007.4428474**

**[42]** N. Gurari, K. J. Kuchenbecker, and A. M. Okamura, "Stiffness discrimination with visual and proprioceptive cues," in Proc. 3rd Joint EuroHaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, Salt Lake City, UT, USA, 2009, pp. 121–126, doi: **https://doi.org/10.1109/WHC.2009.4810845**

**[43]** A. Blank, A. M. Okamura, and K. J. Kuchenbecker, "Effects of proprioceptive motion feedback on sighted and non-sighted control of a virtual hand prosthesis," in 2008 Symposium on Haptics Interfaces for Virtual Environment and Teleoperator Systems, Reno, NV, USA, 2008, pp. 141–142, doi: **https://doi.org/10.1109/HAPTICS.2008.4479933**

**[44]** S. Wendelken et al., "Restoration of motor control and proprioceptive and cutaneous sensation in humans with prior upper-limb amputation via multiple Utah Slanted Electrode Arrays (USEAs) implanted in residual peripheral arm nerves," J. Neuroeng. Rehab., vol. 14, no. 1, 2017, art. no. 121, doi: **https://doi.org/10.1186/S12984-017-0320-4**

**[45]** B. Lima and F. L. Hammond, "Haptically-displayed proprioceptive feedback via simultaneous rotary skin stretch and vibrotactile stimulation," in 2023 32nd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN), Busan, Republic of Korea, 2023, pp. 241–247, doi: **https://doi.org/10.1109/RO-MAN57019.2023.10309652**

**[46]** A. Mablekos-Alexiou, G. A. Bertos and E. Papadopoulos, "A biomechatronic Extended Physiological Proprioception (EPP) controller for upper-limb prostheses," in 2015 IEEE/RSJ Int. Conf. Intel. Robots Syst. (IROS), Hamburg, Germany, 2015, pp. 6173-6178, doi: **https://doi.org/10.1109/IROS.2015.7354257**

**[47]** D. J. Weber, R. Friesen, and L. E. Miller, "Interfacing the somatosensory system to restore touch and proprioception: essential considerations," J. Mot. Behav., vol. 44, no. 6, pp. 403–418, 2012, doi: **https://doi.org/10.1080/00222895.2012.735283**

**[48]** A. Ramos-Murguialday et al., "Proprioceptive feedback and brain computer interface (BCI) based neuroprostheses," PLoS One, vol. 7, no. 10, 2012, art. no. e47048, doi: **https://doi.org/10.1371/JOURNAL.PONE.0047048**

**[49]** R. A. Gaunt, J. A. Hokanson, and D. J. Weber, "Microstimulation of primary afferent neurons in the L7 dorsal root ganglia using multielectrode arrays in anesthetized cats: Thresholds and recruitment properties," J. Neural. Eng., vol. 6, no. 5, 2009, art. no. 055009, doi: **https://doi.org/10.1088/1741-2560/6/5/055009**

**[50]** G. A. Tabot, S. S. Kim, J. E. Winberry, and S. J. Bensmaia, "Restoring tactile and proprioceptive sensation through a brain interface," Neurobiol. Dis., vol. 83, pp. 191–198, 2015, doi: **https://doi.org/10.1016/J.NBD.2014.08.029**

**[51]** S. S. Srinivasan et al., "Agonist-antagonist myoneural interfaces in above-knee amputation preserve distal joint function and perception," Ann. Surg., vol. 273, no. 3, pp. E115–E118, 2021, doi: **https://doi.org/10.1097/SLA.0000000000004153**

**[52]** T. A. Kuiken et al., "Redirection of cutaneous sensation from the hand to the chest skin of human amputees with targeted reinnervation," Proc. Natl. Acad. Sci. USA, vol. 104, no. 50, pp. 20061–20066, 2007, doi: **https://doi.org/10.1073/pnas.0706525104**

**[53]** G. Li and T. A. Kuiken, "Modeling of prosthetic limb rotation control by sensing rotation of residual arm bone," IEEE Trans. Biomed. Eng., vol. 55, no. 9, pp. 2134–2142, 2008, doi: **https://doi.org/10.1109/TBME.2008.923914**

**[54]** A. Akhtar, et al., "Passive mechanical skin stretch for multiple degree-of-freedom proprioception in a hand prosthesis," in 9th International Conference, EuroHaptics 2014, Versailles, France, 2014, pp. 120–128, doi: **https://doi.org/10.1007/978-3-662-44196-1_16**

**[55]** J. D. Brown et al., "An exploration of grip force regulation with a low-impedance myoelectric prosthesis featuring referred haptic feedback," J. Neuroeng. Rehabil., vol. 12, no. 1, 2015, art. no. 104, doi: **https://doi.org/10.1186/S12984-015-0098-1**

**[56]** M. A. Schiefer et al., "Artificial tactile and proprioceptive feedback improves performance and confidence on object identification tasks," PLoS One, vol. 13, no. 12, 2018, art. no. e0207659, doi: **https://doi.org/10.1371/JOURNAL.PONE.0207659**

**[57]** E. J. Rouse, D. C. Nahlik, M. A. Peshkin, and T. A. Kuiken, "Development of a model osseo-magnetic link for intuitive rotational control of upper-limb prostheses," IEEE Trans. Neural. Syst. Rehabil. Eng., vol. 19, no. 2, pp. 213–220, 2011, doi: **https://doi.org/10.1109/TNSRE.2010.2102365**

**[58]** I. Cuberovic et al., "Learning of artificial sensation through long-term home use of a sensory-enabled prosthesis," Front. Neurosci., vol. 13, 2019, art. no. 853, doi: **https://doi.org/10.3389/FNINS.2019.00853**

**[59]** P. D. Marasco et al., "Illusory movement perception improves motor control for prosthetic hands," Sci. Transl. Med., vol. 10, no. 432, 2018, art. no. eaao6990, doi: **https://doi.org/10.1126/SCITRANSLMED.AAO6990**

**[60]** K. H. Sienko et al., "Potential mechanisms of sensory augmentation systems on human balance control," Front. Neurol., vol. 9, 2018, art. no. 944, doi: **https://doi.org/10.3389/FNEUR.2018.00944**

**[61]** T. J. Bates, J. R. Fergason, and S. N. Pierrie, "Technological advances in prosthesis design and rehabilitation following upper extremity limb loss," Curr. Rev. Musculoskelet. Med., vol. 13, no. 4, pp. 485–493, 2020, doi: **https://doi.org/10.1007/s12178-020-09656-6**

**[62]** D. K. Luu et al., "Artificial intelligence enables real-time and intuitive control of prostheses via nerve interface," IEEE Trans. Biomed. Eng., vol. 69, no. 10, pp. 3051–3063, 2022, doi: **https://doi.org/10.1109/TBME.2022.3160618**

**[63]** L. Vargas, H. H. Huang, Y. Zhu, and X. Hu, "Closed-loop control of a prosthetic finger via evoked proprioceptive information," J. Neural. Eng., vol. 18, no. 6, 2021, art. no. 066029, doi: **https://doi.org/10.1088/1741-2552/ac3c9e**

**[64]** C. C. Berger, S. Coppi, and H. H. Ehrsson, "Synchronous motor imagery and visual feedback of finger movement elicit the moving rubber hand illusion, at least in illusion-susceptible individuals," Exp. Brain. Res., vol. 241, pp. 1021–1039, 2023, doi: **https://doi.org/10.1007/s00221-023-06586-w**

**[65]** S. Srinivasan et al., "On prosthetic control: a regenerative agonist-antagonist myoneural interface," Sci. Robot., vol. 2, no. 6, 2017, art. no. eaan2971, doi: **https://doi.org/10.1126/scirobotics.aan2971**

**[66]** S. S. Srinivasan, M. Diaz, M. Carty, and H. M. Herr, "Towards functional restoration for persons with limb amputation: A dual-stage implementation of regenerative agonist-antagonist myoneural interfaces," Sci. Rep., vol. 9, 2019, art. no. 1981, doi: **https://doi.org/10.1038/S41598-018-38096-Z**

**[67]** H. M. Herr et al., "Reinventing extremity amputation in the era of functional limb restoration," Ann. Surg., vol. 273, no. 2, pp. 269–279, 2021, doi: **https://doi.org/10.1097/SLA.0000000000003895**

**[68]** S. S. Srinivasan, S. Gutierrez-Arango, A. C. Teng and H. M. Herr, "Neural interfacing architecture enables enhanced motor control and residual limb functionality postamputation," Proc. Natl. Acad. Sci. USA, vol. 118, no. 9, 2021, art. no. e2019555118, doi: **https://doi.org/10.1073/PNAS.2019555118**

**[69]** M. A. Garenfeld et al., "Closed-loop control of a multifunctional myoelectric prosthesis with full-state anatomically congruent electrotactile feedback," IEEE Trans. Neural Syst. Rehab. Eng., vol. 31, pp. 2090–2100, 2023, doi: **https://doi.org/10.1109/TNSRE.2023.3267273**

**[70]** M. A. Garenfeld et al., "Amplitude versus spatially modulated electrotactile feedback for myoelectric control of two degrees of freedom," J. Neural Eng., vol. 17, no. 4, 2020, art. no. 046034, doi: **https://doi.org/10.1088/1741-2552/aba4fd**

**[71]** E. Battaglia et al., "Skin stretch haptic feedback to convey closure information in anthropomorphic, under-actuated upper limb soft prostheses," IEEE Trans. Haptics, vol. 12, no. 4, pp. 508–520, 2019, doi: **https://doi.org/10.1109/TOH.2019.2915075**

**[72]** M. Rossi et al., "HapPro: a wearable haptic device for proprioceptive feedback," IEEE Trans. Biomed. Eng., vol. 66, no. 1, pp. 138–149, 2019, doi: **https://doi.org/10.1109/TBME.2018.2836672**

**RESEARCH ARTICLE**

VOL. 46 | NO. 1 | **JANUARY-APRIL 2025** | PP 1-17

Revista Mexicana de Ingeniería Biomédica

https://dx.doi.org/10.17488/RMIB.46.1.1487

LOCATION ID: e1487

# Comparison of Machine Learning Models for Identification of Depressive Patients through Motor Activity

## Comparación de Modelos de Aprendizaje para la Identificación de Pacientes Depresivos por Medio de Actividad Motora

*Gerardo Neftalí Rivera-Rojas*[1], *Carlos Eric Galván-Tejada*[1], *Jorge Isaac Galván-Tejeda*[1],
*José M. Celaya-Padilla*[1], *Erika Acosta-Cruz* [2]

[1] Universidad Autónoma de Zacatecas, Zacatecas - México

[2] Universidad Autónoma de Coahuila, Coahuila - México

## ABSTRACT

The present study aims to evaluate various classification algorithms for data pertaining to subjects diagnosed with depression and non-depressive subjects. To this end, the data obtained from the "depresjon" dataset proposed by Garcia-Ceja, E., *et al* were analyzed. This dataset comprises motor activity recorded by the Actiwatch device (Cambridge Neurotechnology Ltd, England, model AW4). Predictions were made using various machine learning models, including synthetic data. Subsequently, metrics such as specificity, sensitivity, and precision were compared. The results highlight the best features of the data and the best machine learning model (using an ensemble model) for classifying potential depressive episodes in activity during the afternoon and night, with a precision of 96.6 %, sensitivity of 100 %, and specificity of 93.33 %.

**KEYWORDS:** data mining, data analysis, depression, machine learning and motor activity

## RESUMEN

El presente estudio tiene como objetivo evaluar diversos algoritmos de clasificación de datos pertenecientes a sujetos diagnosticados con depresión y sujetos no depresivos. Para ello, se analizaron los datos obtenidos del *dataset "depresjon"* propuesto por Garcia-Ceja, E., *et al*, el cual se compone de la actividad motora captada por el dispositivo *Actiwatch (Cambridge Neurotechnology Ltd, England, model AW4).* Mediante distintos modelos de aprendizaje automático se realizaron predicciones incluyendo datos sintéticos. Posteriormente, se compararon métricas como especificidad, sensibilidad y precisión. Los resultados muestran las mejores características de los datos, así como el mejor modelo de aprendizaje automático (mediante modelo de ensamble) para realizar la clasificación de posibles episodios depresivos en la actividad durante la tarde y la noche, con una precisión del 96.6 %, una sensibilidad del 100 % y una especificidad del 93.33 %.

**PALABRAS CLAVE:** actividad motora, análisis de datos, aprendizaje automático, depresión, minería de datos

**Corresponding author**

TO: **CARLOS ERIC GALVÁN-TEJADA**
INSTITUTION: **UNIVERSIDAD AUTÓNOMA DE ZACATECAS**
ADDRESS: **JARDÍN JUÁREZ #147, CENTRO HISTÓRICO,**
**ZACATECAS, ZAC. C.P. 98000, MÉXICO.**
EMAIL: **ericgalvan@uaz.edu.mx**

**Received:**

05 November 2024

**Accepted:**

24 February 2025

**Published:**

24 April 2025

## INTRODUCTION

Portable electronic devices known as "wearables" have been utilized in the field of healthcare over the last decade. These devices not only accurately capture data, but have also found applications in several areas such as medicine, psychology, rehabilitation, and intervention for various psychological disorders[1]. They are particularly valuable in non-invasive research, providing greater precision with measurements ranging from seconds to entire weeks. Additionally, the implementation of these devices enables continuous monitoring of subjects under study without the need for direct observation by the experimenter, resulting in cost savings in research and minimal disruption to the daily activities of the subjects.

Major Depressive Disorder (MDD), a condition affecting approximately 280 million people worldwide[2], is characterized by persistent sadness, loss of interest in previously enjoyable activities, inability to perform daily activities for at least two weeks, decreased energy, changes in appetite, alterations in circadian rhythms, among other clinical features[3]. Due to the reduction in daily activities, various mental illnesses have been studied using electronic devices[4], and different machine learning models[5][6]. As a result, the use of different devices for recording markers has become increasingly common in current research. These markers prove useful when analyzing variables related to morbid processes, including but not limited to blood pressure[7], blood oxygenation[8], motor activity[9], and other parameters to identify different physical and mental health alterations.

Recent studies indicate that motor activity during different times of the day can strongly correlate with depression. The use of wearables has highlighted a key clinical feature: reduced movement, linked to symptoms such as drowsiness, insomnia, decreased interest in physical activity, and unwarranted fatigue. Additionally, sociodemographic characteristics have established connections between depression and risk groups. However, the diversity in parenting styles, social skills, cognitive assets, and external variables complicates psychological research, making it difficult to find well-supported, generalizable relationships for accurate differential diagnosis[10]. Thus, providing precise prognoses and treatments requires careful consideration of these diverse factors.

Accurate diagnosis of affective disorders is crucial, since these conditions can lead to suicidal ideation and irritability, endangering both the affected individuals and their caregivers[11]. Research has explored the impact on quality of life[12] and the potential comorbidities with cardiovascular, metabolic, and cancer-related diseases[13]. While MDD is treatable with medication and psychotherapy, early intervention significantly improves outcomes. Thus, certain tools are essential for the timely detection of depressive episodes. Existing tests, such as the Beck Depression Inventory (BDI)[11], the Diagnostic and Statistical Manual (DSM-V)[12], the International Classification of Diseases (ICD-10)[14], and the Montgomery-Asberg Depression Rating Scale (MADRS)[13], rely on patient-reported information and are subject to potential human error. Additionally, technical requirements, frequent updates by healthcare institutions, and the willingness of at-risk individuals can hinder timely treatment in some populations.

Hence, the implementation of new technologies is necessary to enable early detection through comprehensive, harmless, and non-invasive monitoring for individuals, thereby reducing medical costs[15]. This approach offers the adaptation of treatment based on captured data as time progresses. The monitoring of motor activity in medical and psychiatric fields proves advantages in abnormal behaviors identification detecting specific periods of movement that can provide insights into recognizing behavioral patterns associated with diseases such as dementia (cita), depression[15], anxiety (cita), and schizophrenia[16]. In conditions like anxiety and depression, it has been demonstrated that affected patients tend to reduce daytime activity, increase nighttime activity, and engage in activities related to certain diagnostic criteria outlined by the DSM-V[12]. On the other hand, patients diagnosed with different types of bipolar disorder exhibit increased energy during recognizable periods accompanied by periods in which activity significantly decreases. Therefore, monitoring motor activity serves as a good indicator for differentiation between individuals with a depressive episode or bipolarity, as they show discrepancies compared to healthy individuals[17].

Motor signals have been acquired using diverse methodologies, including different types of accelerometers. It has been suggested that assorted devices of this nature are viable for behavioral analysis in sports, disease prevention, and management[18], enabling efficient and precise measurement of treatment progress. Once the information is obtained, varied methods of statistical processing and/or classification can assist in the pharmaco-psychological treatment by providing quantitative insights that complement the work of the professional team. Enrique Garcia-Ceja *et al.*[6] conducted a study where data on the motor activity of 23 patients with depression (the "condition" group) were obtained, along with sociodemographic characteristics such as age, gender, presence of melancholy, type of illness (unipolar or bipolar), patient type (outpatient or hospitalized), level of education, marital status, employment status (employed, unemployed, or subsidized by a government program), as well as scores on the depression scale (MADRS). Additionally, 32 healthy subjects (the "control" group) were included, from whom only age and gender characteristics were obtained.

Different machine learning algorithms were employed to classify depressive and non-depressive signals using the mentioned data. The study demonstrated that these models could analyze motor behavior to classify subjects. Additionally, classifying conditions based on sociodemographic characteristics significantly impacts psychological diagnosis and can guide tailored treatments. This article addresses both motor and sociodemographic characteristics to classify depressive and healthy subjects, as well as different types of depression within the "condition" group. Classification methods through machine learning algorithms will be used, involving data mining processes to minimize variables affecting predictors' performance and maximize the impact of variables highly associated with depression[18], aiming to maintain control and generate useful data for classification. The data were segmented by different time periods: early morning (00:00 - 06:00 hs), morning (06:01 - 12:00 hs), afternoon (12:01 - 18:00 hs), and night (18:01 - 24:00 hs).

The structure of the article is as follows: Materials and Methods, Results, Discussion, and Conclusion. These sections describe the processes implemented for classifying various parameters in the control and condition groups.

## MATERIALS AND METHODS

### Pre-processing data

The data mining process commenced with the organization and cleaning of data extracted from the dataset available at: https://datasets.simula.no//depresjon/. As depicted in, this dataset contains information regarding the motor activity (counts per minute) of patients diagnosed with some form of depression, as well as information about healthy subjects. The records were initially stored individually in separate files, each containing details such as one-minute time intervals, date of data acquisition, and motor activity. Each record was analyzed individually, and data that were deemed non-informative (exhibiting abnormal distribution, auto-correlation, or suggesting little relevance of characteristics to the explanation of the phenomenon in question) were excluded.

The records exhibited varying duration, spanning up to 20 days. Consequently, the minimum number of records per subject was considered as a threshold for extracting statistical features and subsequent training and testing of machine learning methods. The dataset was divided into 52 subsets, corresponding to each of the 4 daily periods (6 hours per period). This division involved selecting 13 continuous days of data for each subject from the aforementioned groups, ensuring that all subjects commenced and concluded their records at the same minute. The resulting dataset comprises 15 subjects from the condition group and 15 subjects from the control group, aiming to minimize potential data imbalance.

## Descriptive statistics

The statistical treatment of the data began with normality tests applied to each of the 52 subsets (4 subsets for each of the 13 days), using the Shapiro-Wilk test (see Algorithm 1), revealing a non-normal distribution $p < 0.05$ of the data in most time periods analyzed in both groups. Subsequently, descriptive statistical tests were conducted to identify differences in movements between groups and generate potentially informative data. the parameters minimum value (Min), maximum value (Max), and mean (Mean) equations (2 and 3), were computed using the R function summary, while the calculation of the standard deviation (SD) was performed using the (stats) package equation (4), Additionally, the skewness (SK) and kurtosis (KURT) were calculated with the (e1071) library equations (5 and 6) were calculated for each subset Table 1. Other libraries were employed for specific purposes, such as (ggplot2) for generating graphs, (dplyr) and (tidyr) for data manipulation and preprocessing, (caret) for training and testing machine learning models, and (ape) for advanced statistical analyses.

Based on relevant research on movement and mental disorders [19], maximum values for the groups were imputed using the metric of mean + 2 standard deviations. This was done to maintain expected ranges of movement in accordance with literature reports[20], thereby reducing the chances of misclassification due to data acquisition errors or potential sensor failures.

**TABLE 1.** **"Dataset created from the Depresjon dataset."**

| Periods | Number of observations | Features |
|---|---|---|
| Early Mornings | 4680 | Min, Max, Mean, SD, SK, KURT |
| Morning | 4680 | Min, Max, Mean, SD, SK, KURT |
| Afternoon | 4680 | Min, Max, Mean, SD, SK, KURT |
| Night | 4680 | Min, Max, Mean, SD, SK, KURT |

```
Algorithm: Shapiro-Wilk Test Formula
Input:
  Sample data $X = (x_1,x_2...x_n)$
Output:
  Test statistic W and p-value
```

Steps:

1. Sort the data in ascending order: $x_{(1)} \leq x_{(2)} \ldots \leq x_{-(n)}$.
2. Calculate the coefficients $a_i, b_i$ *and* $c_i$ *for* $i = 1,2,\ldots n$.
3. Calculate the test statistic

$$W = \frac{\left(\sum_{i=1}^{n} a_i x_{(i)}\right)^2}{\left(\sum_{i=1}^{n}(x_{(i)} - \bar{x}\right)^2} \tag{1}$$

where $\bar{x}$ is the sample mean.

4. Calculate the expected value E(W) and the variance Var (W) under normality for the given sample size.
5. Calculate the p-value by comparing W to the distribution of W under the null hypothesis.

The formulas used to obtain the metrics mentioned in Table 1 are presented below

$$U = \min(U_1, U_2) = R - \frac{n_1 \cdot (n_1 + 1)}{2} \tag{2}$$

where:

$U_1$ and $U_2$ are the sums of ranks for the two samples,
$R$ is the sum of ranks for the entire sample,
$n_1$ and $n_2$ are these the sample sizes of the two groups.

$$mean = \frac{\sum_{i=1}^{n} x_i}{n} \tag{3}$$

where:

$x_i$ represent the value of the sample
$n$ is the total number of observations.

$$SD = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \bar{x})^2}{N - 1}} \tag{4}$$

where:

$x_i$ represents the values in the sample
$\bar{x}$ is the sample mean
$N$ is the sample size

$$SK = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^3}{\left(\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}\right)^3} \tag{5}$$

where:

$x_i$ represents the values in the sample
$\bar{x}$ is the sample mean
$n$ is the sample size

$$KURT = \frac{\frac{\sum_i (x_i - \bar{x})^4}{n}}{\sigma^4} - 3 \tag{6}$$

where:
$x_i$ represents the values in the sample
$\bar{x}$ is the sample mean
$n$ is the sample size
$\sigma$ is standard deviation

## Feature selection

After organizing and obtaining metrics for each of the 52 subsets, a total of 312 features were collected. These features were then analyzed to extract relevant characteristics, achieved through multi-variable feature selection using genetic algorithms with the Galgo package[21], below is a brief description of the classification algorithms used for feature extraction through the Galgo library. Forward selection was performed using various methods: Support Vector Machines (SVM) an algorithm that maximizes a specific mathematical function with respect to a set of information, through key concepts: the separation of the hyperplane, maximizing the margin of the hyperplane, soft of the margin, and the Kernel function[22], the algorithm learn by example to assign labels to objects[23].

Neural Networks (NNET): Neural networks consist of an input layer with several nodes, internal hidden layers, and an output layer. Each node is associated with weights, activation functions such as: Linear, Tanh, ReLU, Sigmoid, to mention a few, and thresholds, connecting to the next layer until reaching the output layer where the classification results from the information in the first layer, typically, the data undergoes a training phase in which weights and thresholds are adjusted to provide a more accurate classification. On the other hand, there are various algorithms such as "Adeline, perceptron, and backpropagation" that enable the adjustment of parameters to achieve optimal classification[24].

k-Nearest Neighbors (KNN) supervised machine learning algorithm that can be used to solve both classification and regression problems, is a non-parametric classification method[25], through the calculation of distances between the data points with respect to others[26], assigning labels and iterating until well-differentiated groups are found based on distances such as Euclidean or Manhattan.

Random Forest Algorithm (RF) is a supervised classification algorithm which classifies the data by constructing a number of Classifiers (decision trees) with an aim to achieve a higher accuracy of prediction[27], the Random forest uses "Adaboost and Bootstrapping" techniques to construct multiple classifiers[28]. This algorithm has been applied in economics, medicine, commerce, and the financial sector in recent decades, as high levels of accuracy have been reported in classification tasks with large amounts of data[29]. Furthermore, this technique allows the construction of multiple classifiers that cater to specific issues, thereby minimizing errors in predictions[30].

## Training and test phase

The use of machine learning algorithms previously mentioned in the analysis of human movement through wearable devices like the Actiwatch is motivated by the need to process and classify complex temporal data patterns effectively[6]. Human motor activity, as recorded by wearable sensors, often exhibits high variability due to individual differences, behavioral routines, and external influences[20]. Additionally, these data can be noisy, non-linear, and sometimes chaotic, making traditional statistical approaches insufficient for distinguishing between relevant patterns and random fluctuations.

By implementing these algorithms, we aim to improve the accuracy and reliability of classification models that distinguish between depressive and non-depressive subjects based on movement data. Furthermore, analyzing the most relevant features influencing these models can provide insights into the behavioral differences underlying these conditions. The combination of SVM, KNN, and Neural Networks allows for a comprehensive assessment of motor activity, balancing interpretability, efficiency, and performance in detecting meaningful patterns in wearable sensor data[3].

Classifications were conducted by applying the various algorithms mentioned earlier separately to identify subjects with depression and healthy subjects. This process was carried out in three stages. In the first stage, the motor activity data were partitioned into 75 % for the training phase and the remaining 25 % for the testing stage, with a k=5 cross-validation applied.

The Support Vector Machine (SVM) was implemented using the *svm()* function from the (e1071) package. By default, the kernel function used is the Radial Basis Function (RBF). The cost parameter defaults to 1, and the gamma parameter is automatically computed as equation 7:

$$\gamma = \frac{1}{number\ of\ predictors} \tag{7}$$

The Neural Network (NN) classifier was trained using the *train()* function from the (caret) package, leveraging the "nnet" method from the nnet package. The neural network consists of a single hidden layer by default, with the number of neurons set to (input features + output classes)/2. The activation function used for hidden layers is the sigmoid function, and the output layer applies the softmax function for classification tasks. The weight decay (decay) parameter, which prevents overfitting, defaults to 0, and the maximum number of iterations (maxit) defaults to 100. The optimization is performed using backpropagation with a gradient-based method.

The Random Forest (RF) classifier implemented using the "rf" method in train(), which internally calls the randomForest function from the randomForest package. By default, the number of trees (ntree) is set to 500, and the number of randomly selected predictors per split (mtry) is set to the square root of the total number of predictors for classification tasks. The model aggregates multiple decision trees and uses majority voting to improve predictive performance while reducing overfitting.

The K-Nearest Neighbors (KNN) model was implemented using the "knn" method in train(), which relies on the knn() function from the (class) package. The number of neighbors (k) is tuned automatically by (caret). The Euclidean distance metric is used to measure similarity between data points.

The second stage involved generating synthetic random data by bootstrap method, a statistical technique used to estimate the distribution of statistics by resampling with replacement from the original data. It allows for assessing the variability of a statistic without relying on strong assumptions about the underlying population distribution[31], considering the key characteristics described, and matching the number of subjects in the generated dataset to n=30.

In the third and final stage, proportional partitioning of sociodemographic data was performed similarly to the motor activity data. Feature selection was implemented using Galgo[21]. Subsequently, various machine learning methods were applied, and the results were validated using k=5 cross-validation. Finally the Stacked Ensemble Model[32] integrates the predictions from the SVM, NN, RF, and KNN classifiers. The stacking process involves generating out-of-fold (OOF) predictions from the base models and using them as inputs for a meta-learner, which in this case was a logistic regression model (glm method) from the stats package. The goal of the stacked model is to leverage the strengths of each individual classifier and improve overall predictive accuracy.

## RESULTS AND DISCUSSION

The top 10 features were extracted in order of relevance: "maximum value night 1", "maximum value afternoon 2", "maximum value night 8", "maximum value afternoon 1", "maximum value afternoon 3", "maximum value afternoon 5", "maximum value afternoon 10", "maximum value afternoon 8", "maximum value morning 4", and "maximum value morning 2". These features exhibited classification accuracy ranging from 96.6 % to 100 % for distinguishing between groups using different algorithms.

Several predictions were made using the models mentioned earlier. Finally, to assess the performance of each classification model based on metrics such as True Positives (TP) (subjects with depression correctly classified), True Negatives (TN) (control subjects correctly classified), False Positives (FP) (control subjects incorrectly classified), and False Negatives (FN) (subjects with depression incorrectly classified). Sensitivity (7), specificity (8), and accuracy (9).

$$Sensitivity = \frac{TP}{TP + FN} \tag{8}$$

$$Specificity = \frac{TN}{TN + FP} \tag{9}$$

$$Accuracy = \frac{TP}{TP + FP} \tag{10}$$

$$MCC = \frac{TN \times TP - FN \times FP}{\sqrt{}(TP + FP)(TP + FN)(TN + FP)(TN + FN)} \tag{11}$$

The Mann-Whitney U (2) tests were conducted, contrasting the maximum values of the groups during each of the days considered in the present study. In the following tables 2 and 3, W-statistic values and the p-value for each of the different periods over the 13 days are shown. Significant differences are noted between the control and condition groups, primarily during the morning (Figure 1), afternoon, and night periods (Figure 2).

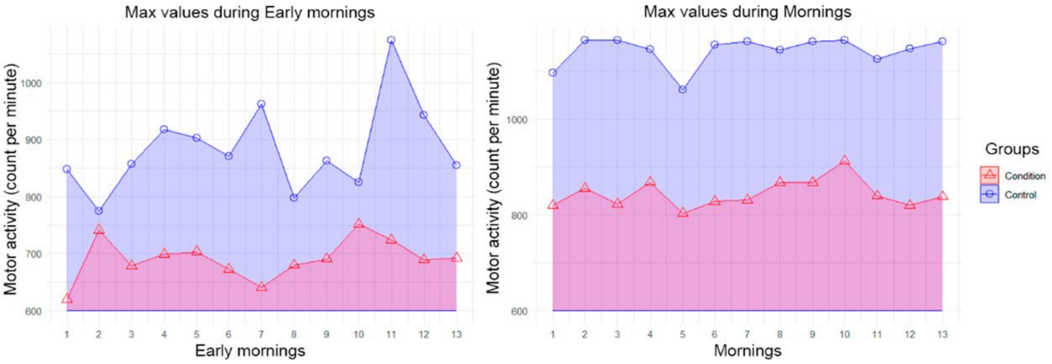


**FIGURE 1. Maximum values of motor activity over 13 days (early morning and morning).**

**TABLE 2. Statistical results of maximum motor activity comparison between different groups during early mornings and mornings.**

| Early mornings | W-value | P-value | Mornings | W-value | P-value |
|---|---|---|---|---|---|
| 1 | 160.5 | 0.04597 | 1 | 210 | 1.108e-05 |
| 2 | 117 | 0.8633 | 2 | 210 | 6.362e-06 |
| 3 | 151 | 0.1087 | 3 | 210 | 1.117e-05 |
| 4 | 174 | 0.009906 | 4 | 210 | 1.36e-05 |
| 5 | 165 | 0.02776 | 5 | 196 | 0.0002472 |
| 6 | 160 | 0.04736 | 6 | 210 | 1.07e-05 |
| 7 | 182 | 0.003716 | 7 | 210 | 8.625e-06 |
| 8 | 135 | 0.353 | 8 | 210 | 8.625e-06 |
| 9 | 143 | 0.2028 | 9 | 210 | 1.07e-05 |
| 10 | 155 | 0.07057 | 10 | 210 | 6.362e-06 |
| 11 | 210 | 2.83e-05 | 11 | 197 | 0.0001651 |
| 12 | 164 | 0.02796 | 12 | 210 | 1.36e-05 |
| 13 | 151.5 | 0.1072 | 13 | 210 | 1.07e-05 |

The condition group shows a tendency toward lower motor activity compared to the control group (Figures 3-4). This finding aligns with clinical criteria described in diagnostic manuals such as the DSM-5[12] and ICD-10[14], which highlight psychomotor retardation, lethargy, and reduced engagement in previously enjoyed activities as key symptoms of depression. Individuals with depression often experience diminished motivation, fatigue, and an overall reduction in voluntary movement, which could contribute to the lower motor activity observed in the experimental group.

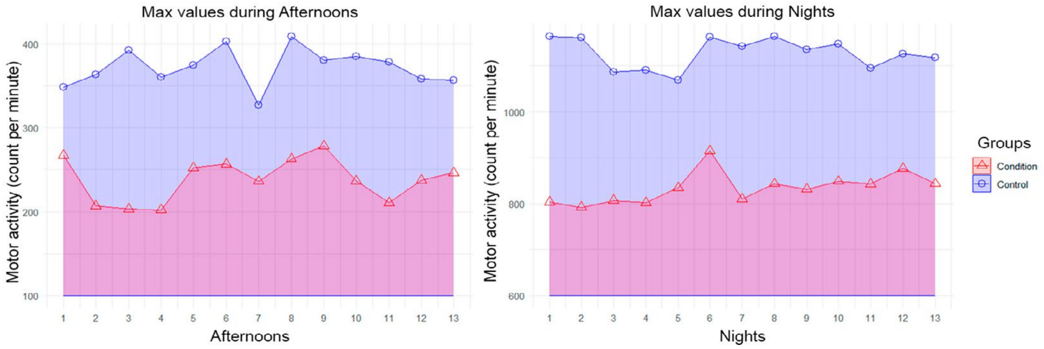


**FIGURE 2. Maximum values of motor activity over 13 days (Afternoon and night).**

TABLE 3. **Statistical results of maximum motor activity comparison between different groups during afternoons and nights.**

| Afternoons | W-value | P-value | Nights | W-value | P-value |
|---|---|---|---|---|---|
| 1 | 160.5 | 0.04597 | 1 | 210 | 1.108e-05 |
| 2 | 117 | 0.8633 | 2 | 210 | 6.362e-06 |
| 3 | 151 | 0.1087 | 3 | 210 | 1.117e-05 |
| 4 | 174 | 0.009906 | 4 | 210 | 1.36e-05 |
| 5 | 165 | 0.02776 | 5 | 196 | 0.0002472 |
| 6 | 160 | 0.04736 | 6 | 210 | 1.07e-05 |
| 7 | 182 | 0.003716 | 7 | 210 | 8.625e-06 |
| 8 | 135 | 0.353 | 8 | 210 | 8.625e-06 |
| 9 | 143 | 0.2028 | 9 | 210 | 1.07e-05 |
| 10 | 155 | 0.07057 | 10 | 210 | 6.362e-06 |
| 11 | 210 | 2.83e-05 | 11 | 197 | 0.0001651 |
| 12 | 164 | 0.02796 | 12 | 210 | 1.36e-05 |
| 13 | 151.5 | 0.1072 | 13 | 210 | 1.07e-05 |

Additionally, the Mann Whitney test was conducted with the averages of motor activity under the same conditions for both groups. The following tables 4-5 display the statistical values as well as the p-values for each of the subsets.

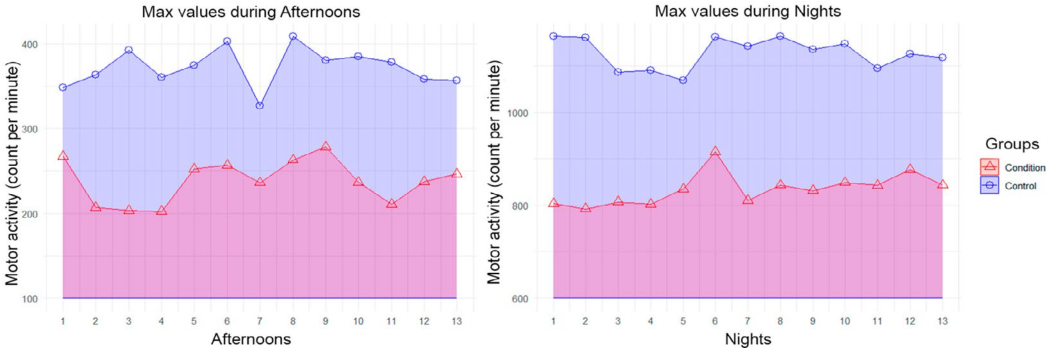


FIGURE 3. **Average values of mean motor activity over 13 days (Early morning and morning).**

TABLE 4. **Statistical results of mean motor activity comparison between different groups during early mornings and mornings.**

| Early mornings | W-value | P-value | Mornings | W-value | P-value |
|---|---|---|---|---|---|
| 1 | 130 | 0.4806 | 1 | 170 | 0.01804 |
| 2 | 99 | 0.5897 | 2 | 202 | 0.0002222 |
| 3 | 116 | 0.9009 | 3 | 208 | 8.104e-05 |
| 4 | 129 | 0.5068 | 4 | 169 | 0.02016 |
| 5 | 131 | 0.4552 | 5 | 152 | 0.1057 |
| 6 | 137 | 0.3194 | 6 | 182 | 0.004201 |
| 7 | 136 | 0.34 | 7 | 172 | 0.01438 |
| 8 | 99 | 0.5897 | 8 | 210 | 8.625e-06 |
| 9 | 103 | 0.7089 | 9 | 184 | 0.003223 |
| 10 | 89 | 0.34 | 10 | 181 | 0.004785 |
| 11 | 164 | 0.03436 | 11 | 131 | 0.4552 |
| 12 | 139 | 0.2807 | 12 | 166 | 0.02789 |
| 13 | 105 | 0.7715 | 13 | 174 | 0.01138 |

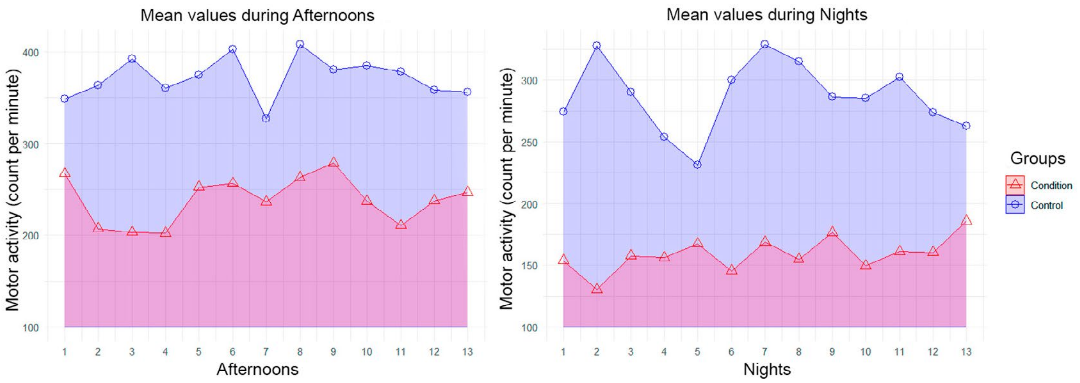


**FIGURE 4.** **Average values of mean motor activity over 13 days (Afternoon and night).**

**TABLE 5.** **Statistical results of mean motor activity comparison between different groups during afternoons and nights.**

| Afternoons | W-value | P-value | Nights | W-value | P-value |
|---|---|---|---|---|---|
| 1 | 146 | 0.171 | 1 | 179 | 0.006178 |
| 2 | 168 | 0.0225 | 2 | 209 | 6.811e-05 |
| 3 | 195 | 0.000669 | 3 | 175 | 0.01011 |
| 4 | 185 | 0.002817 | 4 | 156 | 0.07443 |
| 5 | 161 | 0.04644 | 5 | 136 | 0.34 |
| 6 | 158 | 0.06191 | 6 | 186 | 0.002457 |
| 7 | 152 | 0.1057 | 7 | 172 | 0.01438 |
| 8 | 176 | 0.008957 | 8 | 182 | 0.004201 |
| 9 | 147 | 0.1584 | 9 | 163 | 0.03805 |
| 10 | 172 | 0.01438 | 10 | 174 | 0.01138 |
| 11 | 179 | 0.006178 | 11 | 167 | 0.02507 |
| 12 | 153 | 0.09702 | 12 | 155 | 0.08143 |
| 13 | 157 | 0.06793 | 13 | 139 | 0.2807 |

We conducted within-group comparisons across different time periods. Table 6 presents the significant differences in maximum motor activity values exclusively within the control group during the early morning, morning, and night periods. Meanwhile, Table 7 displays the significant differences in average activity within the control group during the early morning, the significant differences in maximum activity during the early morning for the condition group. These findings suggest variations in activity levels across different times of the day, highlighting potential temporal patterns in motor behavior.

**TABLE 6.** **Statistical results of the maximum motor activity comparison within the control group during early morning and morning periods.**

| Early mornings | W-value | P-value | Mornings | W-value | P-value | Nights | W-value | P-value |
|---|---|---|---|---|---|---|---|---|
| 2 vs 11 | 64 | 0.02456 | 2 vs5 | 150 | 1.81E-02 | 1 vs 4 | 142.5 | 3.84E-02 |
| 3 vs 11 | 66 | 0.0312 | 4 vs 5 | 150 | 1.81E-02 | 1 vs 5 | 150 | 1.81E-02 |
| 6 vs 11 | 66 | 0.0312 | 5 vs 10 | 75 | 1.81E-02 | 4 vs 8 | 82.5 | 3.84E-02 |
| 8 vs 11 | 65 | 0.02771 | - | - | - | 5 vs 8 | 75 | 1.81E-02 |
| 10 vs 11 | 66 | 0.03117 | - | - | - | - | - | - |
| 11 vs 13 | 166 | 0.01509 | - | - | - | - | - | - |

**TABLE 7. Statistical results of the mean motor activity comparison within the control group during early morning period and the maximum motor activity comparison within the condition group during early morning period.**

| Control group mean motor activity | | | Condition group maximum motor activity | | |
|---|---|---|---|---|---|
| Early mornings | W-value | P-value | Early mornings | W-value | P-value |
| 7 vs 11 | 64 | 0.04644 | 2 vs 7 | 156 | 0.044 |
| 8 vs 11 | 62 | 0.03805 | 7 vs 10 | 68 | 0.03931 |
| 11 vs 13 | 162 | 0.04206 | - | - | - |

Regarding the section on training, testing, and validation of different classification methods, considering the key features chosen through "forward selection features". The resulting features were as follows: "maximum value night 1", "maximum value afternoon 2", "maximum value afternoon 1", "maximum value night 8", "maximum value afternoon 3", "maximum value afternoon 5", "maximum value afternoon 10", "maximum value afternoon 8", "maximum value morning 4", and "maximum value morning 2". Below are the various performances of each individual algorithm, as well as the final model using 3 and 10 variables extracted with Galgo (Figure 5).

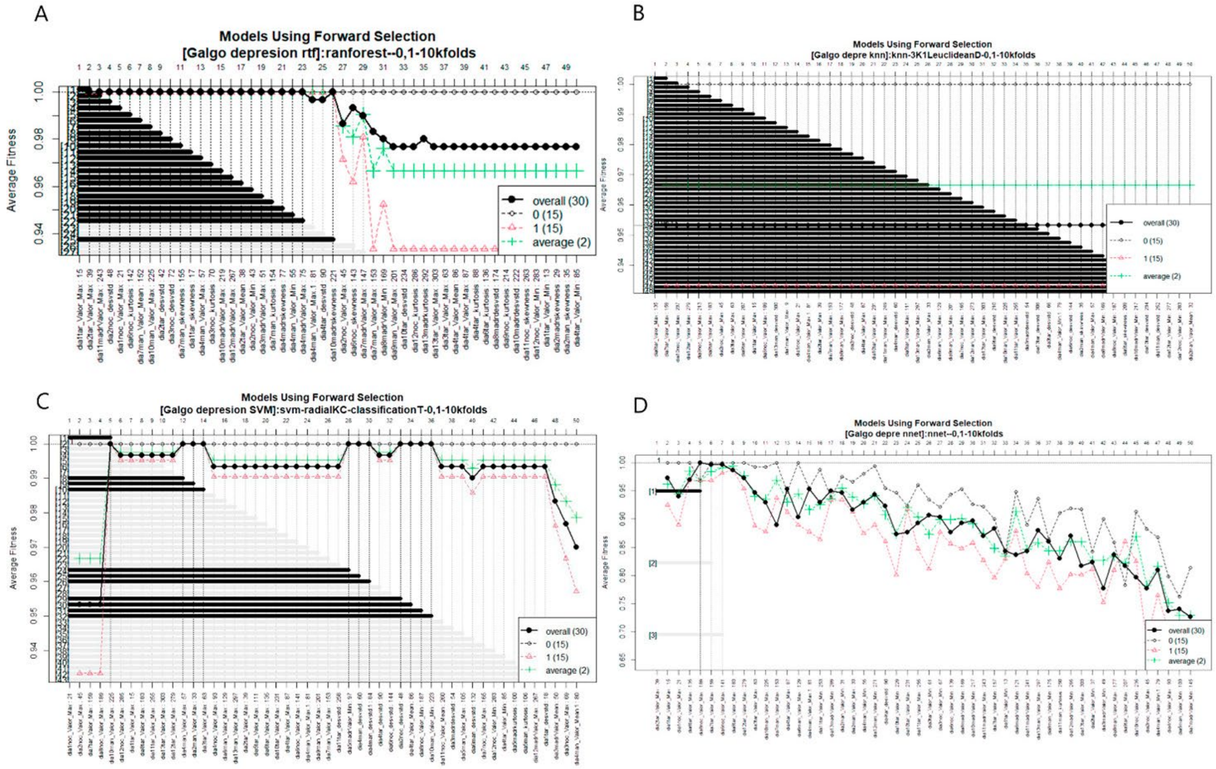


**FIGURE 5. Feature selection using forward selection with the Galgo package. The graphs illustrate the average fitness of the different classification algorithms involved: A) Random Forest, B) KNN, C) SVM, and D) Neural Networks. Additionally, they show the accuracy of the models for classifying depressive and non-depressive subjects, as well as the number and names of the most relevant variables for classification. The results indicate that after selecting 10 features, the accuracy of the different algorithms begins to decline.**

TABLE 8. **Evaluation metrics for different classification methods using 3 most relevant features.**

| Model | Sensitivity | Specificity | Accuracy | MCC |
|---|---|---|---|---|
| RF | 100.00 % | 100.00 % | 100.00 % | 1 |
| KNN | 100.00 % | 100.00 % | 100.00 % | 1 |
| SVM | 100.00 % | 100.00 % | 100.00 % | 1 |
| NNET | 100.00 % | 100.00 % | 100.00 % | 1 |
| STACK | 100.00 % | 100.00 % | 100.00 % | 1 |
| RF Synthetic data | 96.66 % | 100.00 % | 93.30 % | 0.707 |
| KNN synthetic data | 96.66 % | 100.00 % | 93.30 % | 1 |
| SVM synthetic data | 80.20 % | 77.00 % | 83.00 % | 0.707 |
| NNET synthetic data | 96.66 % | 100.00 % | 93.33 % | 0.707 |
| STACK synthetic data | 96.66 % | 100.00 % | 93.33 % | 1 |

The decrease in accuracy, sensitivity, and specificity is observed during the "blind" test with synthetic data n=30 as seen on table 8. While the evaluation metrics of different classification methods vary, it is evident that the final model, based on predictions from previous models, achieves an accuracy of 96.66 %.

TABLE 9. **Evaluation metrics for different classification methods using the top 10 most relevant features.**

| Model | Sensitivity | Specificity | Accuracy | MCC |
|---|---|---|---|---|
| RF | 100.00 % | 100.00 % | 100.00 % | 1 |
| KNN | 100.00 % | 100.00 % | 100.00 % | 1 |
| SVM | 100.00 % | 100.00 % | 100.00 % | 1 |
| NNET | 100.00 % | 100.00 % | 100.00 % | 1 |
| STACK | 100.00 % | 100.00 % | 100.00 % | 1 |
| RF Synthetic data | 96.66 % | 100.00 % | 93.30 % | 0.707 |
| KNN synthetic data | 96.66 % | 100.00 % | 93.30 % | 1 |
| SVM synthetic data | 80.20 % | 77.00 % | 83.00 % | 0.707 |
| NNET synthetic data | 96.66 % | 100.00 % | 93.33 % | 0.707 |
| STACK synthetic data | 96.66 % | 100.00 % | 93.33 % | 1 |

The decrease in accuracy, sensitivity, and specificity is observed during the "blind" test with synthetic data n=30 table 9. The evaluation metrics of different classification methods vary, it is noteworthy that the final model, based on predictions from previous models, achieves 100% accuracy. However, something alarming is the drastic decrease in the performance of the SVM classification model.

In this study, motor activity data from the dataset published by García-Ceja[6] were analyzed to compare different classification methods using machine learning techniques and to obtain useful data through data mining practices, also considering previous research and the findings reported by Rodríguez *et al.*[3]. In terms of descriptive statistics, significant differences ($p < 0.05$) were found between the movement of healthy individuals and those affected by depression, mainly during the day, afternoon, and night periods. Several neurobiological and psychological mechanisms may explain why the control group exhibits greater motor activity than the condition group. First, depression is associated with dysregulation of neurotransmitter systems, particularly dopaminergic and serotonergic pathways, which play a crucial role in motivation, reward processing, and motor function. Reduced dopamine levels, especially in the mesolimbic and mesocortical pathways, can lead to decreased goal-directed behavior and physical activity. Similarly, serotonin dysfunction has been linked to fatigue and reduced psychomotor speed[12], hypoactivity in the prefrontal lobe (dorsal medial prefrontal cortex [dmPFC], ventral medial prefrontal cortex [vmPFC], and dorsal lateral prefrontal cortex [dlPFC], ventral lateral prefrontal cortex [vlPFC], orbital frontal cortex [OFC])[33] regions involved in movement initiation and executive function, has been observed in individuals with depression. This reduced neural activity may contribute to slower movement, decreased exploration, and a general lack of physical engagement. Additionally, hyperactivity of the hypothalamic-pituitary-adrenal (HPA) axis, leading to chronic stress and elevated cortisol levels, has been shown to negatively affect energy levels and contribute to fatigue, further reducing motor activity[14].

From a behavioral perspective, individuals with depression may experience anhedonia, a decreased ability to experience pleasure, leading to reduced engagement in activities that typically require movement or physical effort. This aligns with the behavioral inhibition system (BIS) theory, which suggests that increased sensitivity to negative stimuli in depression leads to avoidance behavior and decreased motor output[34].

The results reported in this study regarding the predictive capacity of movements recorded through accelerometers in conjunction with the application of different machine learning algorithms suggest that models like RF, KNN, and SVM are well-performance tools that can achieve accuracy above 96 % in the classification of mental illnesses, similar to what Vahia *et al.*[4], mention in their research on monitoring systems related to MDD. However, the implementation of ensemble models provides the possibility of making predictions with stronger support[35]. While this article mentions 100 % accuracy with the final ensemble model, it is recommended to reproduce the model, if possible, with a more representative sample, and conduct power tests.

## CONCLUSIONS

The objectives of this study were the statistical analysis of the data provided by the dataset proposed by Garcia *et al.*[6], as well as the comparison of the performance of different machine learning models for the classification between depressive and non-depressive subjects based on the motor activity during daily activities over days. Contrasting the findings of this work with those mentioned earlier, it is concluded that motor activity is a viable parameter for identifying depressive behaviors. The significant differences between healthy and depressive subjects indicate a significant reduction in movements by those affected by this condition, a possible effect of apathy on human behavior.

Regarding the performances of the different machine learning models, although the algorithms showed performances greater than 96 % with the data extracted from the original dataset, these metrics significantly decreased when subjected to new synthetic blind data. However, in both cases, the stacked model showed an improvement in the classification of different groups, achieving an accuracy of 100 % considering 10 relevant features for the different prediction methods. Nevertheless, it is recommended to increase the sample size, as having so few data points can lead to overfitting issues.

## ETHICAL STATEMENT

This publication is based on the works of Garcia-Ceja[6] The acquisition of the original dataset was part of the Introducing Mental health through Adaptive Technology (INTROMAT) project, funded by the Norwegian Research Council (259293/o70).

## AUTHOR CONTRIBUTION

G. N. R.-R. conceptualization, data curation, formal analysis, investigation, methodology, visualization, writing - original draft, writing - review, and editing; C. E. G.-T. conceptualization, investigation, methodology, writing - original draft, writing - review, and editing; J. I. G.-T. supervision; J. M. C.-P. data curation, supervision; E. A. C. supervisión.

## REFERENCIAS

**[1]** M. E. Morris and A. Aguilera, "Mobile, social, and wearable computing and the evolution of psychological practice," Prof. Psychol. Res. Pract., vol. 43, no. 6, pp. 622-626, 2012, doi: https://doi.org/10.1037/a0029041

**[2]** C. A. Woody et al., "A systematic review and meta-regression of the prevalence and incidence of perinatal depression," J. Affect Disord, vol. 219, pp. 86-92, 2017, doi: https://doi.org/10.1016/j.jad.2017.05.003

**[3]** . G. Rodríguez-Ruiz et al., "Comparison of night, day and 24 h motor activity data for the classification of depressive episodes," Diagnostics (Basel), vol. 10, no. 3, 2020, art. no. 162, doi: https://doi.org/10.3390/diagnostics10030162

**[4]** I. V. Vahia and D. D. Sewell, "Late-life depression: a role for accelerometer technology in diagnosis and management," Amer. J. Psychiatry, vol. 173, no. 8, pp. 763-768, 2016, doi: https://doi.org/10.1176/appi.ajp.2015.15081000

**[5]** A. Abd-Alrazaq et al., "Wearable artificial intelligence for anxiety and depression: scoping review," J. Med. Internet Res., vol. 25, 2023, art. no. e42672, doi: https://doi.org/10.2196/42672

**[6]** E. Garcia-Ceja et al., "Depresjon: a motor activity database of depression episodes in unipolar and bipolar patients," in Proc. 9th ACM Multimedia Syst. Conf., Amsterdam, Netherlands , 2018, pp. 472-477, doi: https://doi.org/10.1145/3204949.3208125

**[7]** A. Massaro et al., "Decisional support system with artificial intelligence oriented on health prediction using a wearable device and big data," in 2020 IEEE Int. Workshop Metrol. Ind. 4.0 & IoT, Roma, Italy, 2020, pp. 718-723, doi: https://doi.org/10.1109/MetroInd4.0IoT48571.2020.9138258

**[8]** D. Wang, J. Weng, Y. Zou, and K. Wu, "EmoTracer: a wearable physiological and psychological monitoring system with multi-modal sensors," in Adjunct Proceedings of the 2022 ACM International Joint Conference on Pervasive and Ubiquitous Computing and the 2022 ACM International Symposium on Wearable Computers, New York, NY, USA, 2022, pp. 444-449, doi: https://doi.org/10.1145/3544793.3560409

**[9]** P. M. Singh and P. Sathidevi, "Design and implementation of a machine learning-based technique to detect unipolar and bipolar depression using motor activity data," in Smart Trends in Computing and Communications, vol. 286, 2021, pp. 99-107, doi: https://doi.org/10.1007/978-981-16-4016-2 10

**[10]** C. Iribarne, V. Renner, C. Pérez, and D. L. de Guevara, "Trastornos del Ánimo y Demencia. Aspectos clínicos y estudios complementarios en el diagnóstico diferencial," Rev. Méd. Clín. Las Condes, vol. 31, no. 2, pp. 150-162, 2020, doi: https://doi.org/10.1016/j.rmclc.2020.02.001

**[11]** G. Jackson-Koku, "Beck Depression Inventory," Occup. Med., vol. 66, no. 2, pp. 174-175, 2016, doi: https://doi.org/10.1093/occmed/kqv087

**[12]** J. Morrison. DSM-5® Guía para el diagnóstico clínico. (2015). [Online]. Available: https://cdn.website-editor.net/50c6037605bc4d1e9286f706427108e6/files/uploaded/DSM-5%2520guia%2520para%2520el%2520diagnstico%2520clinico%2520-%2520James%2520Morrison.pdf.

**[13]** J. Davidson et al., "The Montgomery-Åsberg Depression Scale: reliability and validity," Acta Psychiatr. Scand., vol. 73, no. 5, pp. 544-548,1986, doi: https://doi.org/10.1016/j.csbj.2024.07.005

**[14]** World Health Organization, 2015, "International statistical classification of diseases and related health problems," distributed by World Health Organization, https://iris.who.int/handle/10665/246208

**[15]** S. Saeb et al., "The relationship between clinical, momentary, and sensor-based assessment of depression," in 2015 9th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth), Istanbul, Turquía, 2015, doi: https://doi.org/10.4108/icst.pervasive-health.2015.259034

**[16]** Z. Y. Wee et al., "Actigraphy studies and clinical and biobehavioural correlates in schizophrenia: a systematic review," J. Neural. Transm., vol. 126, no. 5, pp. 531-558, 2019, doi: https://doi.org/10.1007/s00702-019-01993-2

**[17]** J. O. Berle et al., "Actigraphic registration of motor activity reveals a more structured behavioural pattern in schizophrenia than in major depression," BMC Res. Notes, vol. 3, no. 1, 2010, art. no. 149, doi: https://doi.org/10.1186/1756-0500-3-149

**[18]** R. Ramakrishnan et al., "Accelerometer measured physical activity and the incidence of cardiovascular disease: Evidence from the UK Biobank cohort study," PLoS Med., vol. 18, no. 1, 2021, art. no. e1003487, doi: https://doi.org/10.1371/journal.pmed.1003487

**[19]** B. Helgadóttir, Y. Forsell, and Ö. Ekblom, "Physical activity patterns of people affected by depressive and anxiety disorders as measured by accelerometers: a cross-sectional study," PLoS One, vol. 10, no. 1, 2015, art. no. e0115894, doi: **https://doi.org/10.1371/journal.pone.0115894**

**[20]** A. C. Cote et al., "Evaluation of wearable technology in dementia: a systematic review and meta-analysis," Front. Med., vol. 7, 2021, art. no. 501104, doi: **https://doi.org/10.3389/fmed.2020.501104**

**[21]** V. Trevino and F. Falciani, "GALGO: an R package for multivariate variable selection using genetic algorithms," Bioinformatics, vol. 22, no. 9, pp. 1154-1156, 2006, doi: **https://doi.org/10.1093/bioinformatics/btl074**

**[22]** W. S. Noble, "What is a support vector machine?," Nat. Biotechnol., vol. 24, no. 12, pp. 1565-1567, 2006, doi: **https://doi.org/10.1038/nbt1206-1565**

**[23]** B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A training algorithm for optimal margin classifiers," in Proc. COLT92 5th. Annu. Workshop Comput. Learn. Theory, Pittsburgh, PA, USA, 1992, pp. 144-152, doi: **https://doi.org/10.1145/130385.130401**

**[24]** A. Krogh, "What are artificial neural networks?," Nat. Biotechnol., vol. 26, no. 2, pp. 195-1977, 2008, doi: **https://doi.org/10.1038/nbt1386**

**[25]** G. Guo et al., "KNN model-based approach in classification," in On The Move to Meaningful Internet Systems 2003: CoopIS, DOA, and ODBASE, Sicily, Italy, vol. 2888, 2003, pp. 986-996, doi: **https://doi.org/10.1007/978-3-540-39964-3_62**

**[26]** B. Mahesh et al., "Machine learning algorithms-a review," Int. J. Sci. Res., vol. 9, no. 1, pp. 381-386, 2020, doi: **https://doi.org/10.21275/ART20203995**

**[27]** M. Gashler, C. Giraud-Carrier, and T. Martinez, "Decision tree ensemble: Small heterogeneous is better than large homogeneous," in 2008 7th Int. Conf. Mach. Learn. Appl., San Diego, CA, USA, 2008, pp. 900-905, doi: **https://doi.org/10.1109/ICMLA.2008.154**

**[28]** A. B. Shaik and S. Srinivasan, "A brief survey on random forest ensembles in classification model," in Proc. 2018 Int. Conf. Innov. Comput. Commun., vol. 56, Nov. 2018, pp. 253-260, doi: **https://doi.org/10.1007/978-981-13-2354-6_27**

**[29]** L. Pietruczuk, L. Rutkowski, M. Jaworski, and P. Duda, "How to adjust an ensemble size in stream data mining?," Inf. Sci., vol. 381, pp. 46-54, 2017, doi: **https://doi.org/10.1016/j.ins.2016.10.028**

**[30]** S. Chetlur et al., Dec. 2014, "ccuDNN: Efficient primitives for deep learning," 2014, arXiv:1410.0759, doi: **https://doi.org/10.48550/arXiv.1410.0759**

**[31]** A. C. Davison and D. V. Hinkley, Bootstrap methods and their application (Cambridge Series in Statistical and Probabilistic Mathematics). New York, NY, USA: Cambridge University press, 1997, doi: **https://doi.org/10.1017/CBO9780511802843**

**[32]** Y. Xie et al., "Stacking ensemble learning models for daily runoff prediction using 1D and 2D CNNs," Expert Syst. Appl., vol. 217, 2023, art. no. 119469, doi: **https://doi.org/10.1016/j.eswa.2022.119469**

**[33]** I. M. Lin et al., "Prefrontal Lobe and Posterior Cingulate Cortex Activations in Patients with Major Depressive Disorder by Using Standardized Weighted Low-Resolution Electromagnetic Tomography," J. Pers. Med., vol. 11, no. 11, 2021, art. no. 1054, doi: **https://doi.org/10.3390/jpm11111054**

**[34]** S. Wang, F. Leri and S. J. Rizvi, "Anhedonia as a central factor in depression: Neural mechanisms revealed from preclinical to clinical evidence," Prog. Neuro-Psychopharmacol. Biol. Psychiatry vol. 110, 2021, art. no. 110289, doi: **https://doi.org/10.1016/j.pnpbp.2021.110289**

**[35]** J. Chen, A. Zeb, Y. A. Nanehkaran, and D. Zhang, "Stacking ensemble model of deep learning for plant disease recognition," J. Ambient Intell. Human. Comput., vol. 14, pp. 12359-12372, 2023, doi: **https://doi.org/10.1007/s12652-022-04334-6**

**ARTÍCULO DE INVESTIGACIÓN**

VOL. 46 | NO. 1 | **ENERO-ABRIL 2025** | PP 1-28

Revista Mexicana de Ingeniería Biomédica

https://dx.doi.org/10.17488/RMIB.46.1.1490

LOCATION ID: e1490

# Algoritmo de Delineación Especifico basado en el EWMA de dos Etapas para la Estimación de la Presión Arterial no Invasiva

## Specific Delineation Algorithm Based on Two-Stage EWMA for Noninvasive Estimation of Blood Pressure

*Maria Laura Nuñez Ccallo*[1], *Ronny Vilavila Contreras*[1], *Jorge Eusebio Rendulich Talavera*[1], *Erasmo Sulla Espinoza*[1]

[1] Universidad Nacional de San Agustín de Arequipa - Perú

[2] Institute of Electrical and Electronics Engineers - Perú

## RESUMEN

El ruido dinámico fuerte, debido a su variabilidad e intensidad, impide que los métodos convencionales de detección de picos en señales ECG y PPG basados en umbrales fijos funcionen correctamente en dispositivos portátiles. La inflexibilidad de estos umbrales fijos resulta en una baja sensibilidad y valor predictivo positivo. Por lo tanto, en este trabajo se propone un algoritmo de delineación específico con umbral adaptativo basado en el Modelo de Suavizado Exponencial Ponderado (EWMA) de dos etapas, enfocándose en la flexibilidad, precisión, robustez frente al ruido dinámico fuerte y baja carga computacional. El algoritmo propuesto demostró un desempeño robusto en condiciones de alto SNR (24 dB y 18 dB), alcanzando una sensibilidad y un valor predictivo positivo del 100 %. En condiciones de ruido moderado (12 dB), el algoritmo mantuvo una alta sensibilidad del 99.39 % y un valor predictivo positivo del 98.18 %, con una tasa de error de delineación (DER) del 2.43 %. Incluso en condiciones de bajo SNR (6 dB), el algoritmo superó significativamente a los enfoques basados en umbrales fijos, en comparación con más del 50 % en métodos convencionales. Además, se validó el algoritmo utilizando un modelo matemático para estimar la presión arterial basado en el tiempo de tránsito del pulso, con señales provenientes de la base de datos MIMIC. Los resultados mostraron un error medio de -1.422 mmHg para la presión arterial sistólica (SBP) y 0.577 mmHg para la presión arterial diastólica (DBP), con desviaciones estándar de 4.668 mmHg y 2.888 mmHg, respectivamente, cumpliendo con los estándares de la Asociación para el Avance de la Instrumentación Médica (AAMI).

**PALABRAS CLAVE:** algoritmo, ABP, ECG, EWMA, PPG

## ABSTRACT

Dynamic noise, due to its variability and intensity, prevents conventional peak detection methods in ECG and PPG signals based on fixed thresholds from performing effectively in wearable devices. The inflexibility of these fixed thresholds results in low sensitivity and positive predictive value. Therefore, this study proposes a specific delineation algorithm with an adaptive threshold based on the Two-Stage Exponentially Weighted Moving Average (EWMA) model, focusing on flexibility, precision, robustness against strong dynamic noise, and low computational load. The proposed algorithm demonstrated robust performance under high SNR conditions (24 dB and 18 dB), achieving 100 % sensitivity and positive predictive value. Under moderate noise conditions (12 dB), the algorithm maintained a high sensitivity of 99.39 % and a positive predictive value of 98.18 %, with a delineation error rate (DER) of 2.43 %. Even under low SNR conditions (6 dB), the algorithm significantly outperformed fixed-threshold-based approaches, which exhibited error rates exceeding 50 %. Furthermore, the algorithm was validated using a mathematical model to estimate blood pressure based on pulse transit time, with signals from the MIMIC database. The results showed a mean error of -1.422 mmHg for systolic blood pressure (SBP) and 0.577 mmHg for diastolic blood pressure (DBP), with standard deviations of 4.668 mmHg and 2.888 mmHg, respectively, meeting the standards of the Association for the Advancement of Medical Instrumentation (AAMI).

**KEYWORDS:** algorithm, ABP, ECG, EWMA, PPG

### Autor de correspondencia

DESTINATARIO: **ERASMO SULLA ESPINOZA**
INSTITUCIÓN: **UNIVERSIDAD NACIONAL DE SAN AGUSTÍN DE AREQUIPA**
DOMICILIO: **SANTA CATALINA NRO. 117. AREQUIPA C.P. 04000, PERÚ.**
CORREO ELECTRÓNICO: **esullae@unsa.edu.pe**

**Recibido:**

06 Noviembre 2024

**Aceptado:**

03 Febrero 2025

**Publicado:**

16 Abril 2025

## INTRODUCCIÓN

Las enfermedades cardiovasculares abarcan diversos trastornos que afectan al corazón y a los vasos sanguíneos, siendo reconocidas como una de las principales causas de mortalidad global, según la Organización Mundial de la Salud (OMS)[1]. Estas enfermedades requieren una monitorización continua, constante y de manera no invasiva de la presión arterial para facilitar una evaluación precisa de la salud cardiovascular. La medición no invasiva es esencial para identificar tempranamente posibles problemas de salud, reduciendo la incomodidad y los riesgos asociados a los métodos invasivos, lo cual contribuye a que el cuidado médico sea más seguro y orientado al bienestar del paciente[2][3][4]. Los avances en tecnología y la miniaturización de dispositivos electrónicos han hecho posible integrar sensores sofisticados en dispositivos portátiles, como relojes inteligentes, el *OMRON Heart-Guide* inteligente, utiliza un manguito inflable integrado en la correa para realizar mediciones precisas de la presión arterial de forma no invasiva; el *Galaxy Watch 3* y *Active 2*, el brazalete Aktiia, utiliza la fotopletismografía (PPG) combinada con algoritmos avanzados para estimar la presión arterial de manera continua, el *Huawei Watch D*, equipado con un mini manguito inflable que permite medir la presión arterial con alta precisión y el Heartisans Watch, que utiliza tecnología de fotopletismografía (PPG)[5]. Estos dispositivos ilustran cómo la tecnología en dispositivos portátiles está evolucionando para proporcionar soluciones prácticas y precisas para la monitorización de la presión arterial[5]. Para que la monitorización continua de la presión arterial (PA) en dispositivos portátiles sea viable, es esencial que se realice de manera no invasiva[6]. La estimación no invasiva de la presión arterial a través de dispositivos portátiles representa un avance significativo en la monitorización de la salud cardiovascular, ofreciendo de una manera más cómoda, continua y precisa de medir la presión arterial, lo que puede llevar a una mejor gestión de las enfermedades cardiovasculares, incrementar la calidad de vida de los pacientes y reducir los costos de atención médica[6]. Además, es fundamental que estos dispositivos portátiles, debido a sus componentes miniaturizados y limitaciones de energía, utilicen algoritmos de baja carga computacional que sean simples y eficientes. Esto no solo optimiza el consumo de energía y los costos, sino que también extiende la duración de la batería y mejora la experiencia del usuario al mantener la precisión y funcionalidad de los dispositivos. Por lo tanto, la implementación de algoritmos de baja carga computacional, que logren un balance entre simplicidad y eficiencia, es crucial para el éxito y adopción de estos dispositivos portátiles en la monitorización de la salud.

En la actualidad, uno de los métodos más usados para estimar la presión arterial no invasiva, es el tiempo de llegada del pulso sanguíneo (PAT), basándose en los picos del ECG y PPG[7][8][9]. En este contexto, el pico R en el ECG y el punto de inflexión máximo en el PPG son especialmente útiles debido a su prominencia y facilidad de localización, permitiendo estimaciones confiables de la presión arterial sistólica (SBP) y diastólica (DBP)[10][11][12]. Uno de los puntos más importantes para elaborar un correcto algoritmo de detección del PAT es la delineación. La delineación es la identificación precisa de puntos específicos en las señales cardíacas, como los picos del ECG, PPG, entre otras[13][14]. En los últimos años, la delineación de picos del ECG y PPG ha adquirido una importancia significativa para el análisis del ritmo cardíaco y el diagnóstico de arritmias[11][14]. Estos algoritmos generalmente se dividen en dos etapas: “Preprocesamiento de la señal” y “Detección de picos”[15][16]. En la primera etapa, se optimizan los picos R del ECG y los máximos del PPG mediante técnicas como el filtrado digital y las transformadas wavelets[17][18][19][20][21][22]. El filtrado digital es particularmente beneficioso, ya que mitiga ruidos de alta intensidad, como artefactos musculares, interferencias eléctricas y desplazamientos de la línea base, los cuales pueden causar detecciones erróneas de los picos de ECG y PPG[14]. Además, el filtrado digital es eficaz para reducir el ruido de baja frecuencia sin comprometer la información útil de la señal, lo que mejora significativamente la calidad de las señales ECG y PPG y facilita una detección más precisa de los picos R en el ECG y picos máximos en el PPG[16][23]. Sin embargo, este tipo de filtrado no es suficiente para eliminar por completo el ruido dinámico fuerte que ocurre en personas en movimiento. En la segunda etapa, se utilizan comúnmente métodos como el método de detección por umbrales, la desviación estándar móvil, la derivada de la señal y redes neuronales[24][25][26][27][28][29]. El método de detección por umbrales compara la amplitud de la señal con un valor límite, lo que lo hace fácil de entender e implementar, ya que no requiere cálculos computacionalmente intensivos, por lo que resulta ideal para aplicaciones con recursos limitados. Aunque es sensible al ruido, su efectividad aumenta al ajustar adecuadamente el umbral y aplicar técnicas de preprocesamiento[24][25]. El método de la desviación estándar móvil es útil para adaptarse a cambios abruptos y variaciones en el ruido, lo que lo hace más sensible de detectar picos nítidos. Sin embargo, puede presentar un retraso en la detección del pico real debido al uso de una ventana móvil. A pesar de esta limitación, es efectivo para mejorar la detección en condiciones de señales variables[26]. La derivada de la señal también se adapta bien a cambios abruptos y variaciones en el ruido, haciéndola capaz de detectar picos nítidos. No obstante, este método puede ser susceptible al ruido de alta frecuencia, lo que puede afectar su precisión. A pesar de esta limitación, sigue siendo una técnica eficaz para la detección de picos en señales con alta variabilidad[27]. Las redes neuronales han demostrado un alto rendimiento, incluso en señales con ruido e interferencia, pero en su entrenamiento requieren una gran cantidad de datos con mayor poder de carga computacional[28][29].

Por lo tanto, para cumplir nuestras expectativas y objetivos, el método de detección por umbrales es el más adecuado ya que brinda cierto grado de flexibilidad, simplicidad y robustez e ideal para aplicaciones con recursos limitados. Los principales tipos de métodos de detección por umbrales que se utilizan en la detección de picos R del ECG y picos máximos del PPG son los umbrales fijos, los umbrales móviles y los umbrales adaptativos. Los umbrales fijos, en la detección de picos en señales de ECG y PPG son susceptibles a falsas detecciones debido al ruido dinámico fuerte en la señal, lo cual requiere ajustes manuales para su corrección. Por otro lado, los umbrales móviles pueden presentar un desfase entre el pico real y el pico detectado debido al uso de la ventana móvil. La falta de adaptabilidad compromete la precisión en diferentes segmentos de la señal, destacando la necesidad de utilizar umbrales adaptativos para mejorar la precisión en condiciones de ruido de línea base[30]. Los umbrales adaptativos utilizados en la detección de picos en señales ECG y PPG muestran una capacidad para ajustarse dinámicamente a las variaciones de la señal, lo que mejora la precisión de la detección y disminuye la incidencia de resultados falsos positivos y negativos, también pueden ajustarse automáticamente para compensar las fluctuaciones de la señal[24][31][32]. En síntesis, la delineación específica de señales ECG y PPG es desafiante debido al ruido dinámico fuerte y la inflexibilidad de los umbrales fijos, resaltando la importancia de los umbrales adaptativos para mejorar la fiabilidad de los resultados en entornos ambulatorios con recursos limitados para dispositivos portátiles.

Para desarrollar un algoritmo de delineación especifico con umbral adaptativo que sea flexible, preciso, robusto y no requiera de una recalibración constante, se considera varios métodos de predicción de series temporales[31][33][34][35][36][37][38][39][40][41]. En diversas investigaciones se han empleado métodos reconocidos de series temporales como: El Promedio Móvil Integrado Autorregresivo (ARIMA) es flexible para el análisis de series temporales, pero requiere que los datos sean estacionarios, lo que implica transformaciones previas en muchos casos y su demanda computacional es relativamente alta[34][37]. Los métodos Autorregresivos Condicionalmente Heterocedásticos Generalizados (GARCH) son excelentes para modelar series con varianza cambiante, pero son complejos, computacionalmente intensivos y no se enfocan directamente en la detección de picos, siendo más adecuados para series financieras y predicción de volatilidad[38]. La Autorregresión Vectorial (VAR) captura la dinámica de múltiples series, pero es complejo y requiere que los datos sean estacionarios[39]. Las Redes Neuronales pueden manejar relaciones no lineales complejas y múltiples series, pero son computacionalmente intensivas y difíciles de ajustar[40]. Finalmente, el promedio móvil ponderado exponencial (EWMA), es simple, flexible, computacionalmente ligero a comparación de otros métodos y permite ajustar el factor de suavización[35]. Este modelo es una generalización del suavizado exponencial de *Brown*, que permite el ajuste de múltiples parámetros. Como *Zi Hao*[15] que emplea un algoritmo de umbral adaptativo basado en el modelo de suavizado exponencial de *Brown*, que incorpora la selección de características morfológicas y un método de corrección de errores, logrando una alta precisión en ambientes con ruido dinámico fuerte. Este enfoque destaca por su sensibilidad, flexibilidad, robustez y baja carga computacional, lo que lo hace ideal para condiciones de ruido. La elección del modelo EWMA se debe a su simplicidad y capacidad para ajustar el factor de suavización, equilibrando así la sensibilidad y flexibilidad en la detección de picos en dispositivos portátiles[34][35].

Jaehong Yu[33], compara diferentes métodos de Promedio Móvil Ponderado Exponencial (EWMA) para evaluar su eficacia en procesos de pronóstico autoiniciales, una técnica útil cuando hay pocas observaciones históricas. Jaehong Yu, examina cómo cada método (modelo EWMA simple (ES), modelo EWMA de doble suavizado (ES doble), modelo EWMA de triple suavizado (ES triple) y modelo EWMA de dos etapas) se desempeña en la predicción de series temporales con datos insuficientes y patrones cambiantes. Encuentra que el modelo de EWMA de dos etapas supera a los otros en términos de manejo de datos complejos y ruidosos, gracias a su capacidad para ajustar dinámicamente el sesgo y la deriva en las series temporales. Jaehong Yu, sugiere que este modelo es el más adecuado para pronósticos autoiniciales y recomienda futuras investigaciones sobre los parámetros de suavizado óptimos. Sin embargo, no se ha considerado su uso para la detección de picos en señales ECG y PPG para una monitorización de presión arterial. Jaehong Yu, tiene un enfoque más general en el pronóstico de series temporales, comparando varios métodos EWMA, mientras que este trabajo se enfoca en un modelo específico de EWMA de dos etapas para adaptarse a las necesidades de un algoritmo de delineación en entornos ruidosos, aplicado específicamente a la monitorización de presión arterial en dispositivos portátiles. Este trabajo propone un algoritmo de delineación específico de picos R del ECG y los puntos de inflexión máximos del PPG, utilizando un método de umbral adaptativo basado en un Modelo de Suavizado Exponencial Ponderado (EWMA) de dos etapas. Los resultados del estudio demuestran que el modelo EWMA de dos etapas mantiene una alta sensibilidad y valor predictivo positivo, incluso en condiciones de alto ruido dinámico fuerte, superando de manera notable tanto al modelo EWMA simple (Es) como al método de umbral fijo. La validación del algoritmo se realizó utilizando las bases de datos QT y MIMIC, en combinación con un modelo matemático de estimación de la presión arterial basado en el tiempo de llegada del pulso. Los resultados obtenidos demuestran que el algoritmo es eficaz para una monitorización precisa y continua de la presión arterial, destacándolo como una solución confiable para aplicaciones en dispositivos portátiles. En la Tabla 1, podemos observar donde se sintetizan las ventajas, limitaciones y contribuciones específicas de los métodos existentes y del modelo propuesto. Este trabajo es parte del desarrollo de investigación en el área de ingeniería biomédica de la Universidad Nacional de San Agustín de Arequipa[42][43][44][45] que busca mejorar el proceso de asistencia médica.

TABLA 1. **Comparación de Métodos para la Detección de Picos en Señales Biomédicas.**

| Método/Modelo | Ventajas | Limitaciones | Referencias |
|---|---|---|---|
| **Umbral fijo** | Fácil implementación; baja carga computacional. | Sensible al ruido dinámico fuerte; requiere ajustes manuales. | [24][25] |
| **Umbral móvil** | Se adapta a cambios locales en la señal; mejora la detección en segmentos variables. | Puede presentar desfase debido al uso de ventanas móviles; limitado en condiciones de alto ruido. | [26][27] |
| **Umbral adaptativo** | Ajuste dinámico a variaciones de la señal; reduce falsos positivos y negativos. | Mayor carga computacional que los métodos anteriores. | [24][30][32] |
| **Redes neuronales** | Alta precisión incluso con señales ruidosas; capacidad de manejar relaciones no lineales complejas. | Requiere grandes cantidades de datos para entrenamiento; alta carga computacional. | [28][29] |
| **Modelo EWMA simple (ES)** | Computacionalmente ligero; adecuado para datos estacionarios. | Menor capacidad de adaptación en condiciones de ruido dinámico fuerte. | [34][35] |
| **Modelo EWMA de dos etapas** | Alta sensibilidad y precisión en condiciones de alto ruido dinámico; computacionalmente eficiente. | Requiere ajustes iniciales en los parámetros de suavizado. | [31][33][34][35] |

## MATERIALES Y MÉTODOS

### Algoritmo de delineación especifico

En esta sección, se describe la delineación específica de las señales necesarias para el estudio, con el objetivo de realizar un análisis estadístico posterior sobre la estimación de la presión arterial. Se delinearon dos señales fisiológicas ECG, PPG con el fin de analizar la relación entre distintas definiciones del tiempo de tránsito del pulso (PAT), presión arterial sistólica (SBP) y diastólica (DBP)[46]. La Figura 1 muestra las señales utilizadas, sus puntos fiduciales y las diferencias de tiempo consideradas[46], en este trabajo se utilizará solo un punto: *"R-peak"*, el cual representa el pico más destacado de la señal ECG. Este pico, que puede ser positivo o negativo según la derivación del ECG, es el más utilizado debido a la facilidad de su delineación[12]. Para la señal PPG se ve cinco puntos, aunque los tres primeros son los más frecuentemente utilizados en la literatura, en este trabajo solo se utilizara un punto, el "Máximo" que es la parte superior del pulso. La señal ABP permite determinar los valores máximos y mínimos de la presión arterial en cada latido: el valor máximo corresponde a la presión arterial sistólica (SBP) y el mínimo a la presión arterial diastólica (DBP) de cada pulso. En la Tabla 2 se presentan las principales variables temporales no invasivas obtenidas de las señales ECG y PPG[46]. Este trabajo se centra específicamente en el punto de referencia "Rm", derivado del pico R en la señal ECG y del pico máximo en la señal PPG.

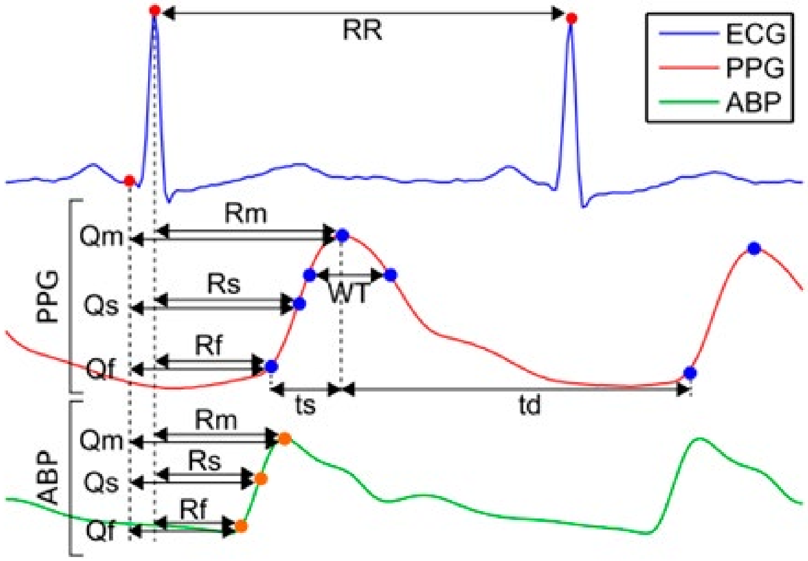


FIGURA 1. **Punto de referencia Rm, pico R del ECG y Máximo del PPG para el algoritmo de delineación especifico [46].**

TABLA 2. **Variables temporales no invasivas definidas a partir de las señales de ECG y PPG [46].**

| Variable | Puntos de referencia | Origen |
|---|---|---|
| Qf | Inicio Q - Pie | ECG-PPG |
| Qs | Inicio Q - Pendiente | ECG-PPG |
| Qm | Inicio Q - Máximo | ECG-PPG |
| Rf | Pico R - Pie | ECG-PPG |
| Rs | Pico R - Pendiente | ECG-PPG |
| Rm | Pico R - Máximo | ECG-PPG |
| WT | Inicio WT - Final WT | PPG |
| ts | Pie - Máximo | PPG |
| td | Máximo - Pie | PPG |

Durante la adquisición de datos, se seleccionaron señales sin procesar provenientes de la base de datos MIMIC[47]. Esta base de datos es una recopilación de múltiples parámetros fisiológicos de pacientes en unidades de cuidados intensivos (UCI), que incluye 72 registros completos de pacientes con grabaciones de al menos 20 horas, muchas de ellas con más de 40 horas de duración. Las señales de ECG, PPG y ABP han sido muestreadas a una frecuencia de 125 Hz, lo que garantiza una resolución adecuada para su análisis[14]. El desarrollo del algoritmo se llevó a cabo utilizando el *software MATLAB* R2021b[48]. El algoritmo propuesto se divide en dos etapas: “Preprocesamiento de la señal” y “Detección de picos”, como se observa en la Figura 2. En la primera etapa “Preprocesamiento de la señal", se emplea un filtro FIR que elimina la interferencia de la línea de base y ajusta la señal para optimizar la detección de picos. En la segunda etapa “Detección de picos”, se aplican técnicas de umbral adaptativo basadas en un modelo de promedio móvil ponderado exponencial (EWMA) de dos etapas, permitiendo así identificar los picos de presión arterial sistólica (SBP) y diastólica (DBP) correspondientes a un ciclo cardíaco válido.

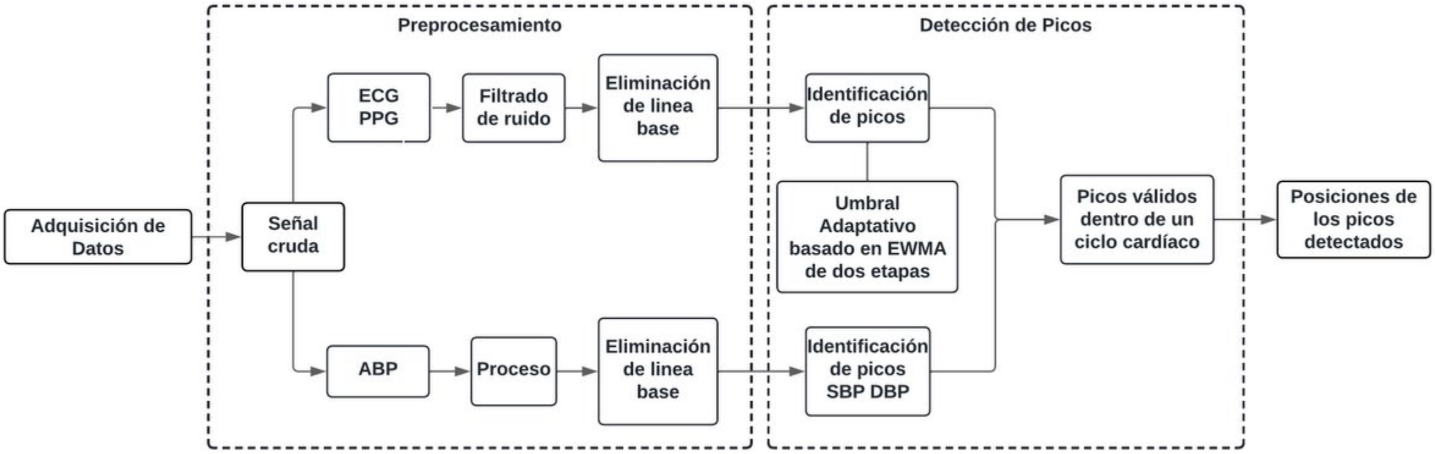


FIGURA 2. **Diagrama de bloques del algoritmo de delineación específico.**

## Preprocesamiento de la señal

El preprocesamiento de las señales de ECG, PPG y ABP incluye técnicas de filtrado digital diseñadas para mejorar la identificación de los picos R y minimizar el ruido. Las señales se obtienen directamente de la base de datos MIMIC, la cual proporciona datos con una frecuencia de muestreo de 125 Hz, eliminando la necesidad de realizar un proceso adicional de submuestreo o diezmado. Cabe destacar que, en configuraciones de adquisición diferentes, las señales originales pueden registrarse a frecuencias más altas (por ejemplo, 1 kHz) y luego reducirse a 125 Hz mediante un filtro *antialiasing* seguido de un proceso de diezmado adecuado. Aunque este procedimiento no se realiza en este trabajo, entendemos la importancia de mencionarlo como contexto adicional, ya que explica cómo podrían generarse señales con frecuencia de muestreo de 125 Hz a partir de datos iniciales en 1 kHz en otros estudios.

El filtrado digital consta de dos etapas secuenciales: primero, se aplica un filtro FIR de paso alto con ventana de *Hamming* y una frecuencia de corte de 0.016 Hz, con una longitud de 101 coeficientes. Este filtro elimina las fluctuaciones de baja frecuencia asociadas con movimientos, respiración y cambios en la línea base. Luego, se utiliza un filtro FIR de paso bajo, también con ventana de *Hamming*, con una frecuencia de corte de 40 Hz, con una longitud de 251 coeficientes para suprimir el ruido de alta frecuencia, como interferencias electromagnéticas y ruido muscular. La selección de 40 Hz se basa en los estándares establecidos por la *Society for Cardiological Science & Technology* (SCST), que indican que la información relevante en estas señales se encuentra por debajo de este límite.

Este filtrado asegura que las ondas QRS del ECG, esenciales para detectar los picos R y fundamentales en la estimación del tiempo de tránsito del pulso y presión arterial, se conserven mientras se eliminan los ruidos[49]. El resultado obtenido tras este proceso de filtrado, como se muestra en la Figura 3, proporciona una señal más limpia y lista para la siguiente etapa del análisis. Este enfoque garantiza que se mantenga la integridad de las señales esenciales mientras se minimiza el impacto del ruido y las interferencias.

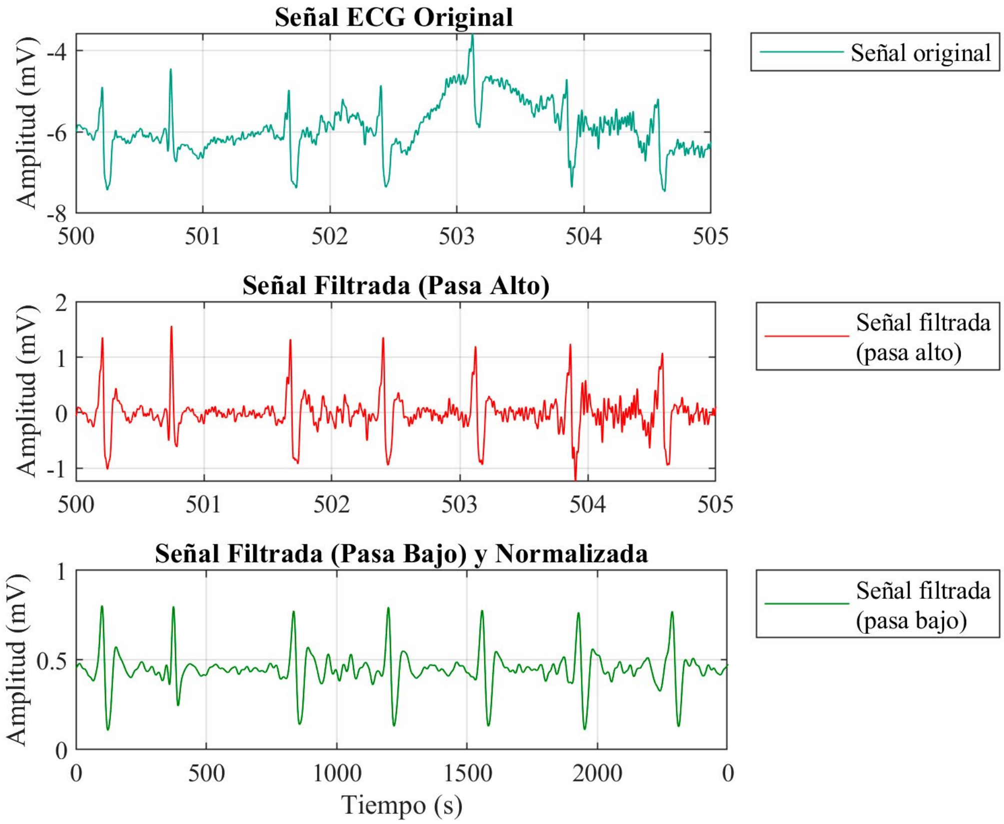


FIGURA 3. **Señal ECG original y señal ECG procesada usando filtros FIR.**

## Detección de picos

En esta etapa, se realiza un ajuste de rango y normalización de las señales de ECG y PPG, transformándolas a un rango estándar predefinido entre 0 y 1. Este ajuste homogeniza la escala de amplitud de las señales, reduciendo el impacto de variaciones entre registros individuales, como las causadas por diferencias en los dispositivos de adquisición o las características fisiológicas de cada paciente. Este proceso mejora la precisión en la detección de los picos R en el ECG y los puntos máximos en el PPG al facilitar la aplicación de umbrales adaptativos, los cuales se ajustan dinámicamente a las características locales de cada señal. Esto permite reducir la probabilidad de errores, como falsos positivos causados por fluctuaciones que superan un umbral absoluto y falsos negativos causados cuando un pico real no alcanza el umbral predefinido. Tras la normalización, se aplica un método de umbrales adaptativos para identificar los picos R en el ECG y los puntos de inflexión máximos en el PPG, que corresponden a momentos clave en el ciclo cardíaco[24][30][32]. En la señal de presión arterial, se detectan los picos sistólicos y diastólicos, que representan la presión máxima y mínima, respectivamente. Estos picos se filtran si exceden los límites preestablecidos, definidos entre 90-250 mmHg para la presión sistólica (SBP) y 40-120 mmHg para la presión diastólica (DBP), asegurando que los valores sean coherentes con un ciclo cardíaco válido. Este proceso no altera el ciclo cardíaco válido, sino que garantiza que los valores utilizados para el cálculo de las características del ciclo sean consistentes con eventos fisiológicos reales, como se muestra en la Figura 4. La correcta identificación de estos picos es fundamental para calcular la presión arterial sistólica (SBP) y diastólica (DBP), como se ilustra en la Figura 5.

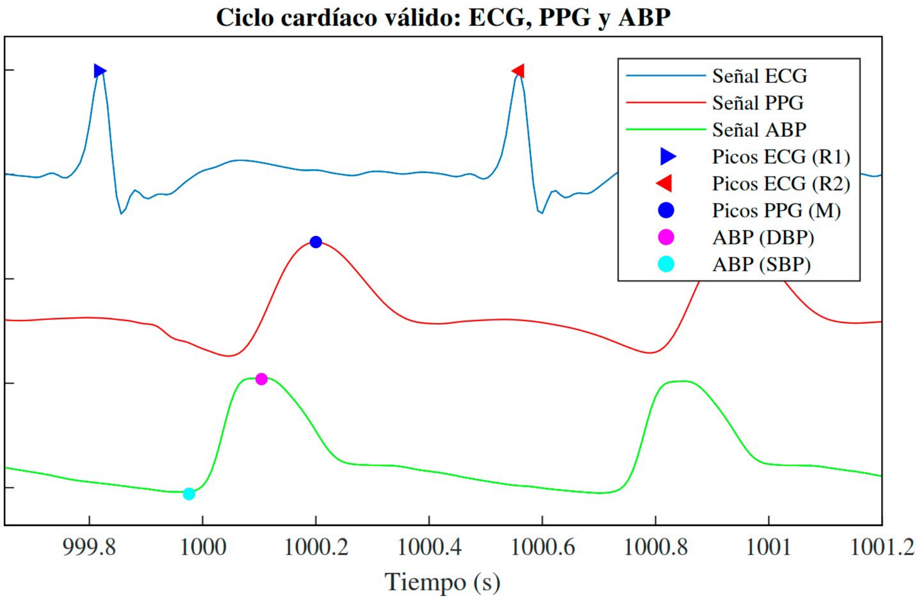


**FIGURA 4.** Ciclo cardiaco válido ECG, PPG y ABP.

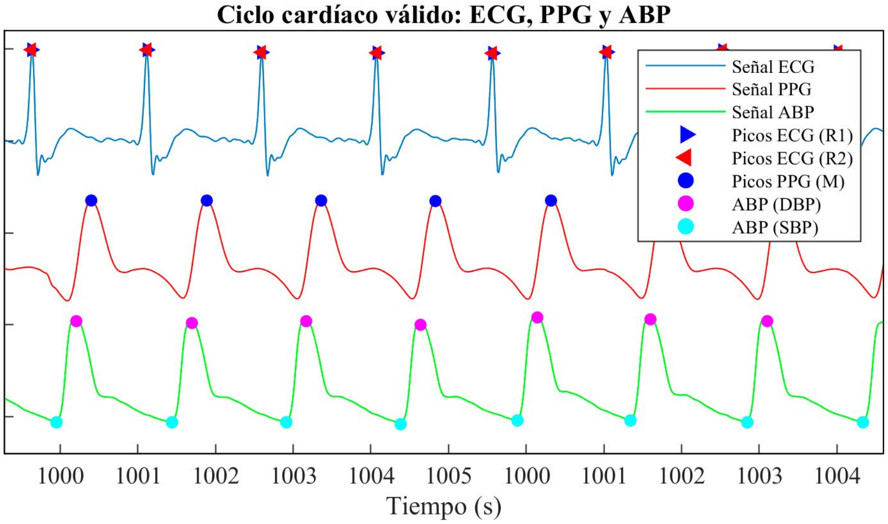


**FIGURA 5.** Señales normalizadas ECG, PPG y ABP del registro 039 m.mat.

## Promedio Móvil Ponderado Exponencial (EWMA)

El Promedio Móvil Ponderado Exponencial (EWMA) es ampliamente utilizado para suavizar señales y reducir el ruido, ajustándose a características específicas de las señales procesadas. Aunque los modelos básicos de EWMA asumen ausencia de estacionalidad, este enfoque ha sido adaptado en este estudio para trabajar con señales biomédicas como ECG y PPG, que son inherentemente estacionales debido a sus ciclos fisiológicos. En este contexto, el modelo de dos etapas incorpora un término de derivada ($d\hat{y}_t$), que ajusta el suavizado a cambios dinámicos, permitiendo capturar tanto las tendencias de largo plazo como las fluctuaciones rápidas y recurrentes. Este diseño asegura que el modelo no solo atenúe las variaciones aleatorias, sino que también preserve los patrones cíclicos propios de las señales biomédicas, como los complejos QRS en ECG y los puntos de inflexión en PPG. Las ecuaciones (8) a (12) detallan cómo el modelo maneja estas características, integrando factores de ajuste que permiten la detección precisa de eventos incluso en señales ruidosas. Este enfoque adaptativo demuestra ser una solución efectiva para manejar tanto la estacionalidad como las variaciones rápidas en señales biomédicas.

Esto es crucial en el procesamiento de señales biomédicas, donde la precisión en la detección de patrones es esencial, como se muestra en la ecuación (1). La ecuación (1) ilustra esta propiedad, donde los pesos disminuyen exponencialmente a medida que retrocedemos en el tiempo. Esta característica es esencial en el procesamiento de señales biomédicas, donde la precisión y la capacidad de detectar patrones dinámicos son fundamentales para garantizar resultados confiables y robustos.

Disminuye exponencialmente el ponderado de $\alpha_i$ a medida que lo devolvemos en el tiempo.

$$\alpha_{i+1} = \lambda\alpha_i = \lambda^2\alpha_{i-1} = \cdots = \lambda^{n+1}\alpha_{i-n} \tag{1}$$

Donde $\propto$: Es el peso asignado a cada dato en la serie temporal.

Donde $\lambda$: Es el factor de suavizado exponencial, con un valor en el rango 0< lambda<1.

El modelo EWMA incluye todas las observaciones previas, asignando ponderaciones que decrecen exponencialmente con el tiempo. Esto permite dar mayor relevancia a los datos más recientes mientras se atenúan progresivamente las contribuciones de las observaciones más antiguas. Luego, en la ecuación (2), la sumatoria de los ponderadores se ajusta para que su valor total equivalga a la unidad, cumpliendo con la restricción de normalización:

$$\sum_{i=1}^{\infty} \alpha_i = \alpha_1 \sum_{i-1}^{\infty} \lambda^i = 1 \tag{2}$$

Para $|\lambda|<1$, el valor de $\alpha_1$=1-$\lambda$

Ahora, conectamos esos términos de nuevo en la ecuación (3). Para el $\sigma^2_{n-1}$ estimado:

$$\begin{gathered} \sigma_{n-1}^2 = \sum_{i=1}^{n-1} \alpha_i\, r_{n-i-1}^2 = \alpha_1 r_{n-2}^2 + \lambda\alpha_1 r_{n-3}^2 + \cdots + \lambda^{n-3}\alpha_1 r_1^2 \\ \sigma_{n-1}^2 = (1-\lambda)(r_{n-2}^2 + \lambda r_{n-3}^2 + \cdots + \lambda^{n-3} r_1^2) \end{gathered} \tag{3}$$

Y el $\sigma^2_n$ estimado de la ecuación (4) se puede expresar así:

$$\begin{gathered} \sigma_n^2 = (1-\lambda) r_{n-1}^2 + \lambda\sigma_{n-1}^2 \\ \sigma_n^2 = (1-\lambda) r_{n-1}^2 + \lambda\big((1-\lambda) r_{n-2}^2 + \lambda\sigma_{n-2}^2\big) \\ \sigma_n^2 = (1-\lambda)\big(r_{n-1}^2 + \lambda r_{n-2}^2 + \lambda^2 r_{n-3}^2 + \cdots + \lambda^{k+1} r_{n-k}^2\big) + \lambda^{k+2}\sigma_{n-k} \end{gathered} \tag{4}$$

En nuestros experimentos utilizamos $\lambda$=0.95 para priorizar la sensibilidad a los cambios recientes en las fluctuaciones dinámicas de la señal, lo cual resulta especialmente útil en el contexto de nuestra investigación.

### Modelo de Suavizado Exponencial Ponderado (EWMA) Simple (Es)

En 1956, se propuso por primera vez un modelo simple (Es)[35], diseñado para pronosticar demandas en sistemas de control de inventario. En el modelo EWMA único, los datos de series de tiempo $\gamma_1$, $\gamma_2$,...,$\gamma_t$ no tienen tendencia ni patrón estacional, y el factor de nivel en el momento $t, \hat{\gamma}_t$ viene dada por:

$$\widehat{\gamma_t} = \alpha\gamma_t + (1-\alpha)\hat{\gamma}t - 1, \tag{5}$$

Donde $\hat{\gamma}_1 = \gamma_1$ y $\alpha$ es un parámetro de suavizado entre 0 y 1. La ecuación (5) se puede reescribir de la siguiente manera:

$$\widehat{\gamma_t} = \sum_{i=0}^{t} \alpha(1-\alpha)^i \gamma_{t-i} \tag{6}$$

Como se ve en la ecuación (6) puede considerarse un promedio ponderado de las observaciones actuales y pasadas. Debido a que el multiplicador de la observación pasada es (1-$\alpha$), que es menor que uno para cualquier valor de $\alpha$, le da más peso a la observación actual que a las observaciones pasadas, el pronóstico para $s$ tiempos por delante a partir del momento $t$ se determina:

$$\gamma^*_{t+s} = \widehat{\gamma_t} \tag{7}$$

En la Ecuación (7), el EWMA simple (ES) pronostica las observaciones futuras como un valor constante, es posible que este método no produzca un rendimiento de predicción satisfactorio en muchas situaciones reales porque los datos de la serie temporal cambian dinámicamente con el tiempo.

### Modelo de Suavizado Exponencial Ponderado (EWMA) de dos etapas

Promedio Móvil Ponderado Exponencialmente (EWMA) de dos etapas, fue propuesto por Ryu y Han que se asemeja al modelo de suavización exponencial doble[33]. Este modelo es útil para pronosticar series temporales, ya que permite ajustar los niveles y tendencias de la serie con mayor precisión.

El factor de nivel $t$, se calcula:

$$\widehat{\gamma_t} = \alpha\gamma_t + (1-\alpha)\hat{\gamma}t - 1, \tag{8}$$

Dónde $\hat{\gamma}_1 = \gamma_1$ es, el factor de ajuste en el tiempo $t$, $d\hat{\gamma}_t$ . La ecuación (8) se puede reescribir de la siguiente manera:

$$d\widehat{\gamma_t} = \beta d\gamma_t + (1-\beta)d\hat{\gamma}t - 1 \tag{9}$$

Donde $d\gamma_t$ es una diferencia de primer orden entre el tiempo $t$ y $t$-1. La ecuación (9) se puede definir de la siguiente manera:

$$d\widehat{\gamma_t} = \gamma_t - \gamma_{t-1}. \tag{10}$$

Además, $d\hat{\gamma}_2 = \gamma_2 - \gamma_1$. Finalmente, estimamos los factores de deriva del tiempo *t* mediante,

$$\Delta\widehat{\gamma_t} = rd\gamma_t + (1-r)\Delta\hat{\gamma}t - 1, \tag{11}$$

Donde $\Delta\gamma_2 = \gamma_2 - \gamma_1$ . El pronóstico final para *s* tiempos posteriores a partir del tiempo actual *t* se puede calcular mediante la siguiente ecuación (12):

$$\gamma^*_{t+s} = \widehat{\gamma_t} + d\widehat{\gamma_t} + s \times \Delta\widehat{\gamma_t}. \tag{12}$$

Este método implica factores ajustados para aliviar los efectos de patrones cambiantes repentinos e inesperados. Por lo tanto, el método EWMA de dos etapas supera al método simple (ES)[33].

## Métricas de validación del algoritmo

La validación del algoritmo se realizó con la base de datos pública QT[46] y mediante un modelo matemático para estimar la presión arterial usando el tiempo de llegada de pulso con la base de datos MIMIC, como se ve en la Figura 6[8][46].

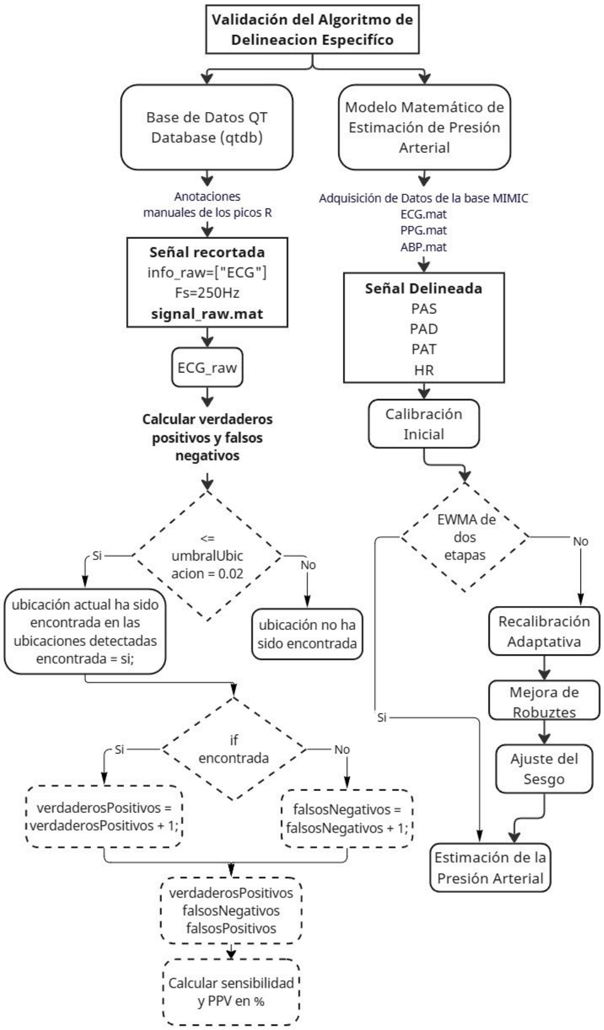


FIGURA 6. **Diagrama de flujo que muestra las dos modalidades de validación del algoritmo de delineación especifico EWMA de dos etapas.**

## Base de datos QT Database (qtdb)

El algoritmo propuesto se validó con la base de datos QT, está base de datos contiene 105 registros de 15 minutos, seleccionados de bases de datos como MIT-BIH y la Sociedad Europea de Cardiología ST-T[50]. Los registros, anotados manualmente por cardiólogos, incluyen entre 30 y 100 latidos, con ondas P, complejos QRS y ondas T[50]. La base de datos QT consiste en grabaciones de ECG de dos derivaciones, muestreadas a 250 Hz[14].

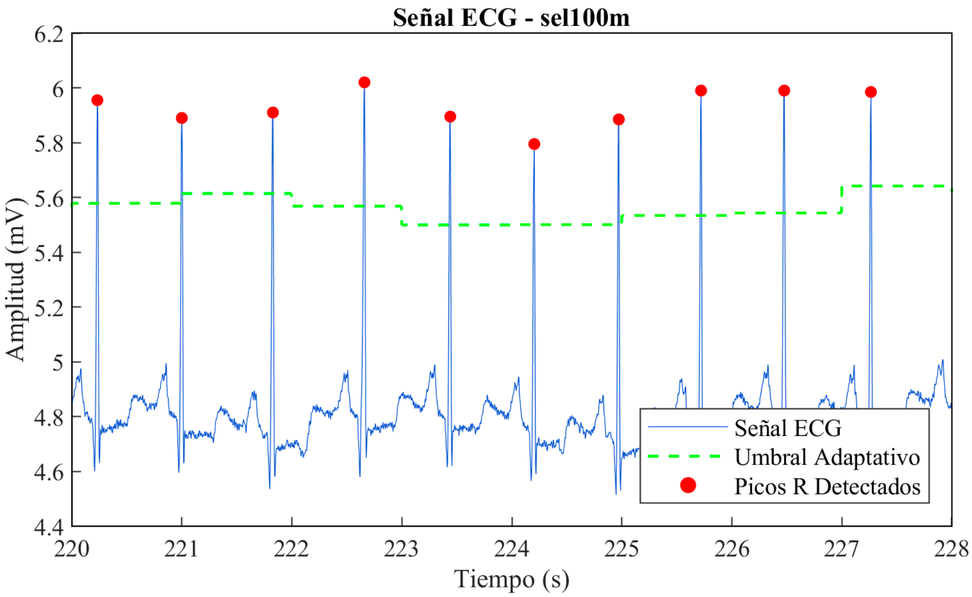


FIGURA 7. **Delineación del pico R del ECG del registro sel100m de la base de datos QT.**

Como se muestra en la Figura 7, el algoritmo demuestra una alta precisión en la detección de verdaderos positivos (TP), puntos registrados en la base de datos QT, sin generar falsos negativos (FN) ni falsos positivos (FP). Las métricas de evaluación indican una sensibilidad y un Valor Predictivo Positivo del 100 % (ecuaciones 13 y 14), lo que respalda la eficacia y aplicabilidad de nuestra solución.

$$Se = \frac{TP}{TP + FN} = 100\ \% \tag{13}$$

$$PPV = \frac{TP}{TP + FP} = 100\ \% \tag{14}$$

## Modelo matemático de estimación de presión arterial

Para este estudio, se seleccionaron seis registros de la base de datos MIMIC, una base de datos de acceso libre que contiene información anónima de más de 40,000 pacientes en unidades de cuidados críticos entre 2001 y 2012. Los criterios de inclusión consideraron registros con señales simultáneas de ECG, PPG y ABP, con una duración mínima de 20 minutos y calidad suficiente para la delineación precisa, evaluada mediante inspección visual. Los criterios de exclusión incluyeron registros que no contaran con las tres señales necesarias, señales con un nivel de ruido alto que comprometiera la precisión de la delineación, y registros incompletos o con interrupciones significativas. Estas señales fueron empleadas para validar el algoritmo de delineación basado en EWMA de dos etapas en la estimación de presión arterial[47][51][52].

La ecuación de Moens-Korteweg describe la relación entre la velocidad de la onda del pulso (PWV) y la distensibilidad arterial, o módulo elástico incremental de la pared arterial.

$$C = \frac{L}{PTT} = \sqrt{\frac{E.h}{\rho.2R}} \tag{1}$$

En la ecuación (1), "*C*" representa la velocidad de la onda, "*L*" es la distancia arterial recorrida por la onda de presión, "PTT" Tiempo de Tránsito del Pulso y el aumento de volumen de la sangre en el punto periférico de medición, "$\rho$" es la densidad del fluido, "R" es el radio interno del vaso, "*E*" es el módulo de elasticidad de la pared (módulo de *Young*), y "*h*" es el espesor del vaso [53][54].

La relación entre PAT y PTT se obtiene a través de esta ecuación (2). La medición indirecta del PTT se realiza generalmente a través del Tiempo de Llegada del Pulso (PAT), calculado como el retardo entre el pico R del ECG y el punto de inflexión máximo del PPG como se muestra en la Figura 8.

$$PAT = PEP + PTT \tag{2}$$

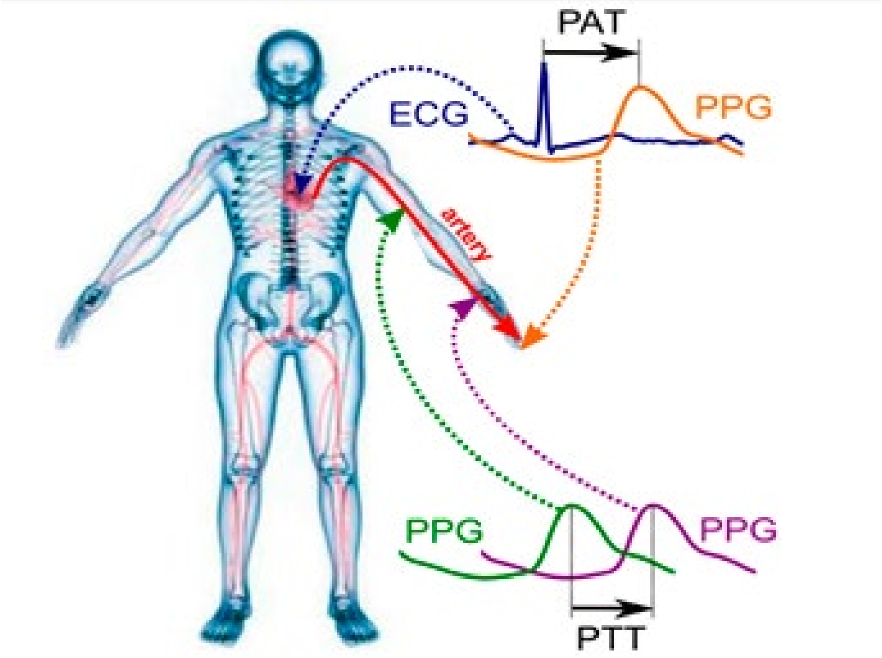


**FIGURA 8. Medición del del Tiempo de Llegada del Pulso (PAT) entre el pico R del ECG y el punto de inflexión máximo del PPG[8].**

El modelo matemático de estimación de presión arterial se basa en estudios preliminares que utilizan la ecuación de Moens-Korteweg y se fundamenta en investigaciones previas ampliamente reconocidas, como las de S. Cattivelli[10][55] y otros autores como J. M. Bote[11][14]. Este modelo permite estimar la presión arterial sistólica (SBP) y diastólica (DBP) a partir del tiempo de llegada del pulso sanguíneo (PAT) y la frecuencia cardíaca instantánea (HR), ajustando

Las ecuaciones (3) y (4) representan el mismo modelo matemático, aunque están expresadas con notaciones diferentes según las investigaciones de referencia. La ecuación (3) describe el modelo en términos de $PAT$ y $HR$, como propuesto originalmente por Cattivelli:

$$\begin{aligned} SBP &= \alpha_1.PAT + \beta_1.HR + \delta_1 \\ DBP &= \alpha_2.PAT + \beta_2.HR + \delta_2 \end{aligned} \tag{3}$$

Por otro lado, la ecuación (4) utiliza una notación compacta adoptada por otros investigadores[11], donde las variables se agrupan de manera más general:

$$\alpha.R_m + \beta.HR + \delta \tag{4}$$

En este caso, $Rm$ equivale al tiempo de llegada del pulso $PAT$ descrito en la ecuación (3). La correspondencia entre estas variables se presenta claramente en la Tabla 3, la cual justifica esta equivalencia y destaca cómo el modelo ha sido adaptado en distintos contextos.

Donde:

- Pico $R$ del $ECG$ y el punto de inflexión máximo del $PPG$=$PAT$=$R_m$.
- Frecuencia cardiaca = $HR$
- Parámetros fisiológicos desconocidos = $\alpha$, $\beta$, $\delta$

TABLA 3. **Variable temporal definida a partir de los puntos de referencia delineados en las señales ECG, PPG Y ABP. La primera columna recoge la variable que es definida mediante los puntos de referencia de la segunda columna[12].**

| Variable | Puntos de Referencia | Origen |
|---|---|---|
| Rm | Pico R - Máximo | ECG-PPG |

La calibración inicial se realiza la primera vez que el sistema es utilizado, mediante un procedimiento de mínimos cuadrados descrito en la ecuación (5). Los parámetros se agrupan en la matriz $\theta$ tal que:

$$\theta = \begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \\ c_1 & c_2 \end{bmatrix}$$

Los valores observados de SBP y DBP ($\gamma_{1N}$) y $PAT$ ($X_{1N}$) se organizan:

$$\gamma_{1:N} = \begin{bmatrix} SBP_{i1} & DBP_{i1} \\ \vdots & \vdots \\ SBP_{iN} & DBP_{iN} \end{bmatrix}$$

$$X_{1:N} = \begin{bmatrix} PAT_{i1} & HR_{i1} & 1 \\ \vdots & \vdots & \vdots \\ PAT_{iN} & HR_{iN} & 1 \end{bmatrix} \tag{5}$$

La calibración minimiza la siguiente ecuación para determinar:

$$\theta_N = [X_{1:N}^T X_{1:N}]^{-1} X_{1:N}^{*} \gamma_{1:N} \tag{6}$$

Inversa de la matriz:

$$P_N = [X_{1:N}^{*} \gamma_{1:N}]^{-1} \tag{7}$$

Este proceso asegura que el modelo se ajuste a las características.

La recalibración suele requerir que el usuario mida su presión arterial sistólica (SBP) y diastólica (DBP) con dispositivos oscilométricos. Estudios previos[7][10][54][55][56][57][58][59] muestran que la recalibración puede mejorar la precisión, pero su frecuencia debe balancearse: recalibraciones frecuentes aumentan la exactitud, pero son incómodas, mientras que intervalos largos reducen la precisión[14]. Nuestro algoritmo de delineación específico, basado en un Promedio Móvil Ponderado Exponencial (EWMA) de dos etapas, reduce significativamente la dependencia de recalibraciones manuales frecuentes, al adaptarse automáticamente a variaciones fisiológicas en entornos ambulatorios. Sin embargo, este modelo no elimina completamente la necesidad de recalibración a largo plazo, lo que será objeto de estudios futuros enfocados en estrategias de recalibración automática para mantener una alta precisión sostenida en el tiempo.

## RESULTADOS Y DISCUSIÓN

### Evaluación frente al ruido

La base de datos MIT-BIH *Noise Stress Test Database* (nstdb) se utiliza para evaluar el rendimiento del algoritmo en condiciones de ruido, con diferentes tipos e intensidades, contiene 12 registros de ECG y 3 registros de ruido, cada uno con una duración de media hora, que son representativos de registros de ECG ambulatorios. Para crear estos registros, se tomaron dos grabaciones limpias, 118 y 119, de la base de datos de arritmias del MIT-BIH, a las cuales se les añadieron cantidades calibradas de ruido. Por lo tanto, aunque la base MIT-BIH forma parte del origen de los registros de la base QT, en este caso se utiliza específicamente para pruebas de ruido dinámico debido a las alteraciones calibradas de los registros 118 y 119, con sus respectivas versiones modificadas. "118 em" y "119 em"[46].

Para nuestras pruebas, nos enfocaremos en el registro "118 em", ya que este registro presenta un nivel de ruido significativamente mayor en comparación con el registro "119 em". Podremos evaluar la capacidad de nuestro algoritmo de delineación específico de umbral adaptativo basado en el Modelo de Suavizado Exponencial Ponderado (EWMA) de dos etapas para la monitorización continua de presión arterial no invasiva en entornos con niveles elevados de ruido dinámico. Esto nos permitirá asegurar la flexibilidad, precisión, exactitud y robustez de nuestro algoritmo frente a condiciones de ruido. Las relaciones señal/ruido (SNR), nos indica el nivel de la señal en comparación con el ruido de fondo y a mayor SNR, menor es el ruido relativo, durante los segmentos ruidosos de 24dB, 18dB y 12dB. Las métricas de evaluación como la sensibilidad S (%), el valor predictivo positivo PPV (%), la tasa de error DER (%) y la exactitud AC (%), nos detallarán los resultados obtenidos sin el algoritmo EWMA (umbral fijo), con el algoritmo EWMA simple (Es) y con el EWMA de dos etapas en los diferentes niveles de SNR.

### Señal / ruido (SNR) para 12dB

La Figura 9 ilustra el desempeño del algoritmo de delineación específico basado en el Modelo de Suavizado Exponencial Ponderado (EWMA) de dos etapas aplicado al registro "118 em" con un nivel de ruido de 12 dB. En esta figura, se observa cómo el algoritmo detecta los picos R con alta precisión, incluso en presencia de ruido, destacando su eficacia frente a métodos comparativos. Los resultados cuantitativos presentados en la Tabla 4 respaldan visualmente estas observaciones. El algoritmo de EWMA de dos etapas logra una sensibilidad del 99,39 %, superando al EWMA simple (88,54 %) y al enfoque sin EWMA con umbral fijo (43,24 %). Su valor predictivo positivo (PPV) alcanza el 98,18 %, en comparación con el 84,96 % del EWMA simple y el 79,56 % del umbral fijo. Además, su exactitud general (AC) es del 97,59 %, frente al 76,55 % del EWMA simple y al 38,92 % del enfoque sin EWMA. En términos de error, el EWMA de dos etapas registra la menor tasa de error (2,43 %), destacándose frente al 26,03 % del EWMA simple y al 124,88 % sin EWMA. Estos resultados confirman las ventajas del algoritmo EWMA de dos etapas, no solo por su capacidad de mantener un desempeño robusto en condiciones de ruido moderado, sino también por su precisión y eficiencia en comparación con métodos alternativos. La figura también resalta el impacto visual de los picos detectados en el contexto de una señal ruidosa, proporcionando evidencia visual del rendimiento del algoritmo.

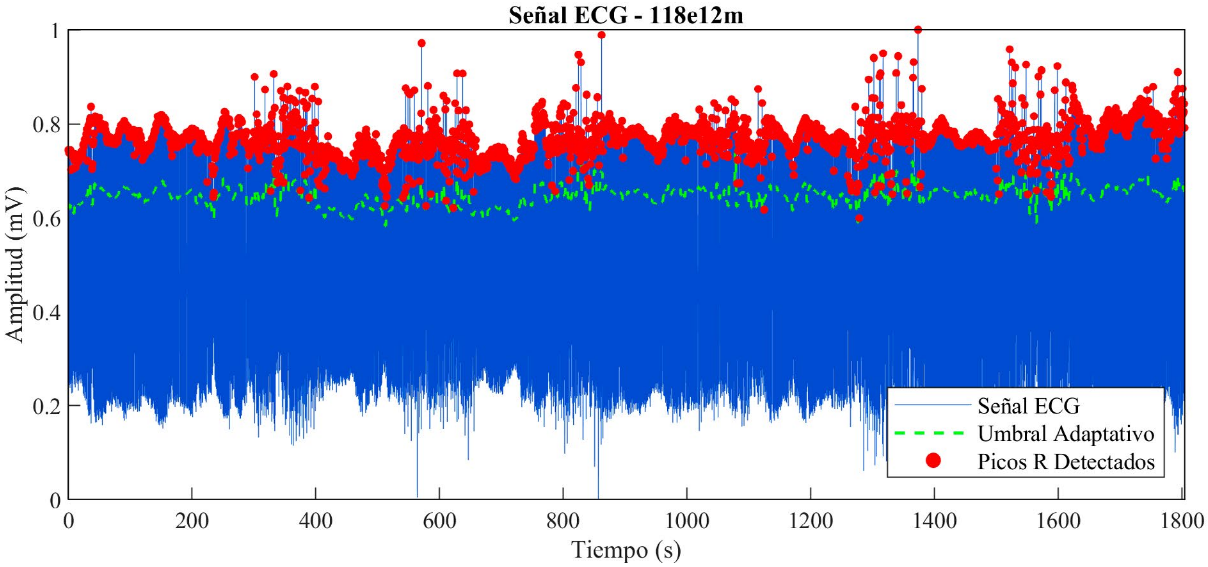


**FIGURA 9. Rendimiento del Algoritmo de Delineación Basado en EWMA de Dos Etapas con SNR de 12 dB.**

**TABLA 4. Evaluación de la relación señal/ruido (SNR) en el registro "118 em" con 12dB, comparando el uso del algoritmo de delineación específico EWMA simple (Es), EWMA de dos Etapas y sin EWMA (umbral fijo).**

| 118em 12dB | | | |
|---|---|---|---|
| SNR | Sin EWMA | Con EWMA Simple (ES) | Con EWMA de dos Etapas |
| S (%) | 43.240 | 88.543 | 99.385 |
| PPV (%) | 79.564 | 84.962 | 98.179 |
| DER (%) | 124.880 | 26.032 | 2.4284 |
| AC (%) | 38.917 | 76.546 | 97.586 |

### Señal / ruido (SNR) para 18dB

La Figura 10 muestra cómo el algoritmo de delineación basado en el Modelo de Suavizado Exponencial Ponderado (EWMA) de dos etapas mantiene un desempeño óptimo en condiciones de ruido moderado, con un SNR de 18 dB. En este entorno, el algoritmo alcanza una sensibilidad perfecta del 100 %, superando al EWMA simple (98,86 %) y al enfoque sin EWMA (92,98 %). En términos de valor predictivo positivo (PPV), el algoritmo de dos etapas también alcanza un 100 %, mientras que el EWMA simple y el enfoque sin EWMA logran un 95,18 % y un 96,98 %, respectivamente. Asimismo, su tasa de error (DER) es del 0 %, significativamente mejor que el 5,92 % del EWMA simple y el 10,35 % del método con umbral fijo. Finalmente, la exactitud global (AC) del algoritmo es también del 100 %, en comparación con el 94,15 % del EWMA simple y el 90,36 % del enfoque sin EWMA.

Estos resultados, presentados en la Tabla 5, destacan las ventajas del algoritmo de EWMA de dos etapas no solo en términos de precisión y sensibilidad, sino también en su capacidad de manejar ruido moderado sin comprometer su desempeño. La Figura 10 refuerza visualmente la capacidad del algoritmo para delinear picos R en una señal ECG bajo estas condiciones, demostrando su robustez frente al ruido.

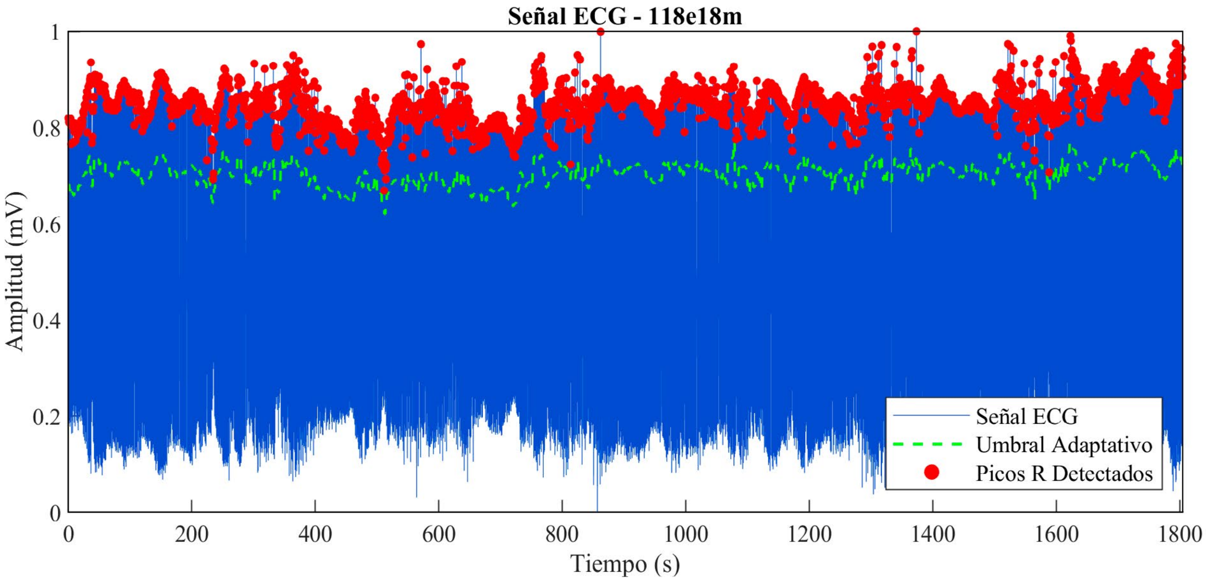


**FIGURA 10.** **Rendimiento del Algoritmo de Delineación Basado en EWMA de Dos Etapas con SNR de 18 dB.**

**TABLA 5.** **Evaluación de la relación señal/ruido (SNR) en el registro "118 em" con 18dB, comparando el uso del algoritmo de delineación específico EWMA simple (Es), EWMA de dos Etapas y sin EWMA (umbral fijo).**

| 118em 18dB | | | |
|---|---|---|---|
| SNR | Sin EWMA | Con EWMA Simple (ES) | Con EWMA de dos Etapas |
| S (%) | 92.976 | 98.859 | 100 |
| PPV (%) | 96.978 | 95.182 | 100 |
| DER (%) | 10.348 | 5.917 | 0 |
| AC (%) | 90.358 | 94.147 | 100 |

## Señal / ruido (SNR) para 24dB

La Figura 11 ilustra el desempeño del algoritmo de delineación específico basado en el Modelo de Suavizado Exponencial Ponderado (EWMA) de dos etapas en condiciones de ruido bajo, con un SNR de 24 dB. En este entorno, el algoritmo alcanza una sensibilidad perfecta del 100 %, destacándose frente al EWMA simple (99,74 %) y al enfoque sin EWMA (99,69 %). En términos de valor predictivo positivo (PPV), el algoritmo de dos etapas también logra un 100 %, superando al EWMA simple (97,55 %) y al enfoque sin EWMA (99,47 %). Asimismo, su tasa de error (DER) es del 0 %, siendo significativamente mejor que el 2,71 % del EWMA simple y el 0,83 % del método con umbral fijo. Finalmente, en cuanto a exactitud global (AC), el algoritmo de EWMA de dos etapas alcanza el 100 %, en comparación con el 97,30 % del EWMA simple y el 99,17 % del enfoque sin EWMA. Estos resultados, resumidos en la Tabla 6, evidencian las mejoras significativas del algoritmo EWMA de dos etapas en entornos con bajo ruido. La Figura 11 refuerza visualmente la capacidad del algoritmo para delinear picos R en una señal ECG con alta precisión, destacando su robustez y exactitud en estas condiciones ideales.

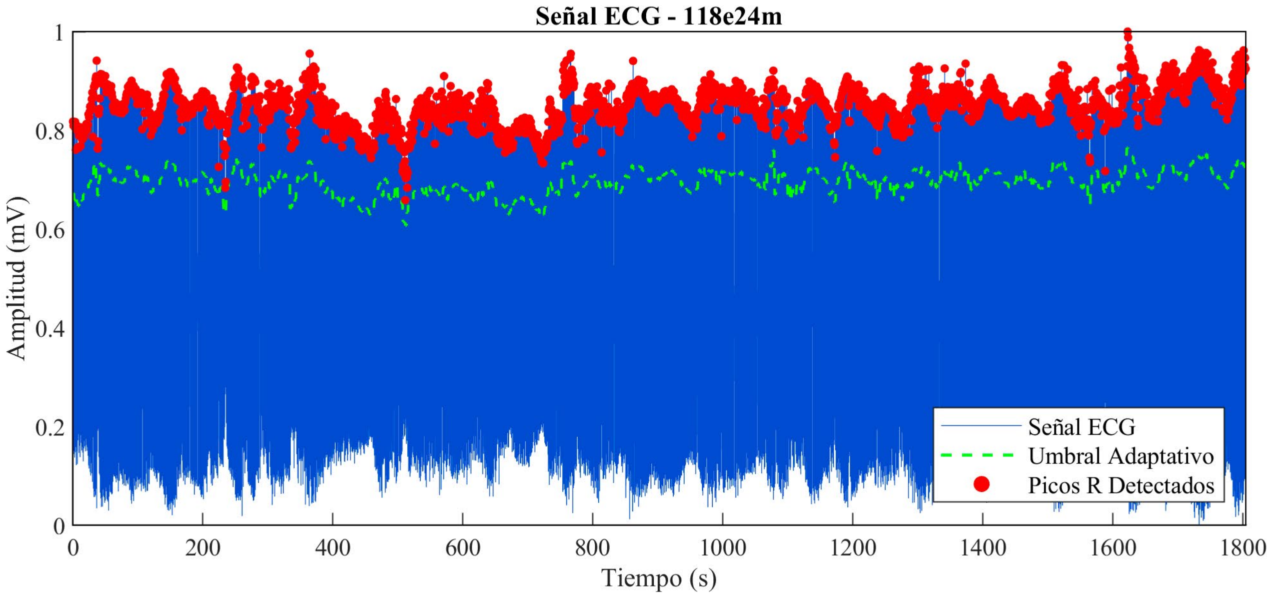


**FIGURA 11. Rendimiento del Algoritmo de Delineación Basado en EWMA de Dos Etapas con SNR de 24 dB.**

**TABLA 6. Evaluación de la relación señal/ruido (SNR) en el registro "118 em" con 24dB, comparando el uso del algoritmo de delineación específico EWMA simple (Es), EWMA de dos Etapas y sin EWMA (umbral fijo).**

| 118em 24dB | | | |
|---|---|---|---|
| SNR | Sin EWMA | Con EWMA Simple (ES) | Con EWMA de dos Etapas |
| S (%) | 99.693 | 99.737 | 100 |
| PPV (%) | 99.474 | 97.553 | 100 |
| DER (%) | 0.832 | 2.705 | 0 |
| AC (%) | 99.170 | 97.302 | 100 |

### Evaluación mediante el Modelo de Estimación de la presión arterial

El modelo matemático de estimación de presión arterial fue aplicado al registro 221m de la base de datos MIMIC[46], como se muestra en la Figura 12. Esta figura presenta una comparación entre las presiones arteriales sistólica (SBP) y diastólica (DBP) reales y estimadas, obtenidas mediante nuestro algoritmo de delineación específico basado en un modelo EWMA de dos etapas. Los puntos representados corresponden a los valores de SBP y DBP delineados en el tiempo, mostrando la precisión del modelo al aproximar los valores reales. Los resultados cuantitativos derivados de este análisis se resumen en la Tabla 7. El error medio (ME) para la SBP fue de 1.318 mmHg y para la DBP de 0.150 mmHg, lo cual indica una alta precisión en las estimaciones. Asimismo, la desviación estándar (S.D.) fue de 5.930 mmHg para la SBP y de 3.279 mmHg para la DBP, mientras que el error cuadrático medio (RMSE) fue de 6.074 mmHg para la SBP y de 3.283 mmHg para la DBP.

Estos resultados destacan la robustez del algoritmo de EWMA de dos etapas en la estimación no invasiva de la presión arterial, cumpliendo con los estándares establecidos por la Asociación para el Avance de la Instrumentación Médica (AAMI). La precisión observada en las estimaciones de SBP y DBP refuerza la aplicabilidad de este enfoque en contextos clínicos y ambulatorios.

Para facilitar la interpretación de la Figura 12, hemos optimizado su presentación incrementando la resolución, ajustando el tamaño de las etiquetas, garantizando una visualización más clara de los datos y su análisis.

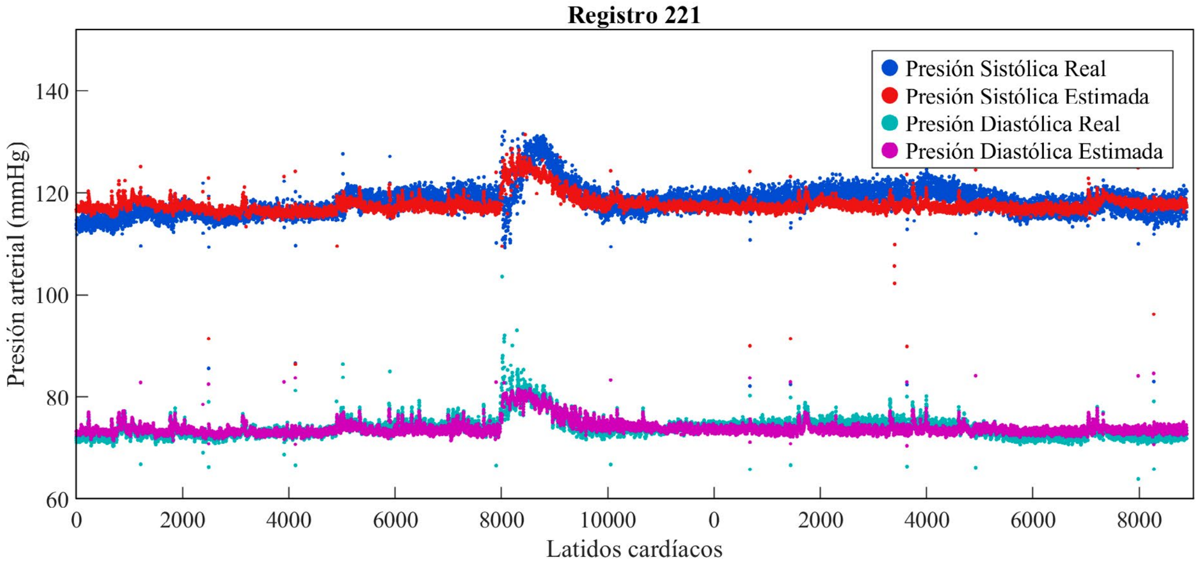


**FIGURA 12. SBP y DBP Delineadas y Estimadas Utilizando EWMA de Dos Etapas para el Registro 221m.**

**TABLA 7. Error medio (ME), desviación estándar (S.D) y error medio cuadrático (RMSE) usando el algoritmo de delineación especifico EWMA de dos etapas para el registro 221m.**

| Registro 221m | | | |
|---|---|---|---|
| PA | ME | S. D | RMSE |
| SBP | 1.318 | 5.930 | 6.074 |
| DBP | 0.150 | 3.279 | 3.283 |

En este estudio, se analizaron 6 registros provenientes de la base de datos MIMIC, cada uno compuesto por tres señales biomédicas: ECG, PPG y ABP. Estas señales fueron utilizadas para la delineación de los picos R en el ECG, los puntos de inflexión máximos en el PPG y los picos sistólicos y diastólicos en el ABP, respectivamente. Los resultados obtenidos para estos registros se presentan en la Tabla 8, destacando el desempeño del algoritmo de delineación específico en diferentes condiciones de señal. Estos resultados permiten analizar el rendimiento del modelo en cada registro, mostrando variaciones en precisión y consistencia según el caso.

TABLA 8. **Error medio (ME), desviación estándar (S.D) y error medio cuadrático (RMSE) usando el algoritmo de delineación especifico EWMA de dos etapas para cada uno de los registros.**

| | SBP | | | DBP | | |
|---|---|---|---|---|---|---|
| Registro | ME | S. D | RMSE | ME | S. D | RMSE |
| 212m | -0.561 | 4.108 | 4.146 | -0.906 | 3.287 | 3.409 |
| 216m | -2.305 | 3.290 | 4.017 | 0.492 | 2.089 | 2.146 |
| 221m | 1.318 | 5.930 | 6.074 | 0.150 | 3.279 | 3.283 |
| 225m | -1.123 | 5.429 | 5.543 | 0.455 | 4.262 | 4.286 |
| 444m | -3.589 | 6.135 | 7.107 | 1.317 | 1.995 | 2.391 |
| 3109737m | 0.469 | 4.378 | 4.402 | 1.527 | 2.809 | 3.197 |

En la Tabla 9, se presentan los valores promedio del error medio (ME), desviación estándar (SD) y error medio cuadrático (RMSE) obtenidos en los seis registros analizados utilizando el algoritmo de delineación específico EWMA de dos etapas. Para la presión arterial sistólica (SBP), se observó un (ME) promedio de -0.9651 mmHg, con una (SD) de 4.8783 mmHg y un (RMSE) de 5.2149 mmHg. Para la presión arterial diastólica (DBP), los valores promedios fueron un (ME) de 0.5058 mmHg, una (SD) de 2.9536 mmHg y un (RMSE) de 3.1186 mmHg. Estos resultados cumplen con los estándares de la Asociación para el Avance de la Instrumentación Médica (AAMI), que establece un límite máximo de 5 mmHg para el ME y de 8 mmHg para la S.D.

TABLA 9. **Promedio del error medio (ME), desviación estándar (S.D) y error medio cuadrático (RMSE) usando el algoritmo de delineación especifico EWMA de dos etapas para los 6 registros.**

| Promedio | | | | | |
|---|---|---|---|---|---|
| SBP | | | DBP | | |
| ME | S. D | RMSE | ME | S. D | RMSE |
| -0.9651 | 4.8783 | 5.2149 | 0.5058 | 2.9536 | 3.1186 |

## Justificación del uso del algoritmo EWMA de dos etapas

La elección del algoritmo EWMA de dos etapas frente a otras técnicas más avanzadas, como redes neuronales o métodos híbridos, se fundamenta en varias razones específicas al contexto de este estudio y la aplicación en dispositivos portátiles. Una de las principales razones es su baja carga computacional. Los dispositivos portátiles, como relojes inteligentes y pulseras de monitoreo, están limitados en términos de capacidad de procesamiento y consumo energético. El modelo EWMA de dos etapas es computacionalmente ligero, lo que lo hace adecuado para estos entornos. En contraste, técnicas avanzadas como las redes neuronales requieren un procesamiento intensivo y grandes conjuntos de datos para el entrenamiento, lo que puede ser impráctico para aplicaciones con recursos limitados. Otra ventaja clave del modelo EWMA de dos etapas es su robustez frente al ruido dinámico fuerte.

Aunque las técnicas avanzadas, como los métodos híbridos o las redes neuronales, pueden manejar ruido extremo, el modelo EWMA de dos etapas ha demostrado un desempeño altamente competitivo. Su capacidad para ofrecer una detección precisa de los picos R en el ECG y los puntos máximos en el PPG, incluso en entornos de ruido dinámico fuerte, lo convierte en una solución robusta para escenarios de monitorización ambulatoria. Además, el algoritmo EWMA destaca por su simplicidad y flexibilidad. Es sencillo de implementar y ajustar, lo que facilita su integración en dispositivos portátiles con recursos limitados. Esta simplicidad no solo optimiza el diseño del sistema, sino que también mejora la reproducibilidad de los resultados en diferentes escenarios y bases de datos, una característica crucial en el desarrollo de algoritmos aplicados.

Por último, el algoritmo EWMA de dos etapas ofrece un balance entre precisión y practicidad. Aunque métodos más avanzados, como las redes neuronales, pueden ofrecer una mayor precisión, su alta carga computacional y las necesidades de entrenamiento dificultan su implementación en dispositivos con restricciones de recursos. En contraste, el modelo EWMA de dos etapas logra un equilibrio óptimo entre precisión, robustez y eficiencia computacional, adaptándose perfectamente a los requisitos de la monitorización continua en tiempo real. Este balance entre simplicidad, eficiencia computacional y robustez respalda la elección del modelo EWMA de dos etapas como la solución más adecuada para este trabajo. Los resultados y las validaciones obtenidos en las bases de datos QT y MIMIC refuerzan esta decisión, mostrando un desempeño competitivo incluso en condiciones de ruido dinámico fuerte.

## Comparación del desempeño del algoritmo EWMA de dos etapas frente a Métodos Alternativos bajo Diferentes Niveles de SNR

En este estudio, se compararon los resultados obtenidos con tres enfoques: el algoritmo de EWMA de dos etapas (Tabla 10), el EWMA simple (Es) (Tabla 11) y el enfoque sin EWMA (umbral fijo) (Tabla 12), bajo distintos niveles de SNR (24 dB, 18 dB, 12 dB, 6 dB, 0 dB y -6 dB). Los resultados revelan diferencias significativas en la robustez de cada método frente al ruido. El algoritmo de EWMA de dos etapas mostró un rendimiento sobresaliente, manteniendo una sensibilidad perfecta (100 %) en condiciones de bajo ruido (24 dB y 18 dB). Su desempeño comenzó a deteriorarse a partir de un SNR de 12 dB, alcanzando una sensibilidad del 62,64 % en condiciones extremas de ruido (-6 dB).

**TABLA 10. Evaluación de la relación señal/ruido (SNR) en el registro "118 em" con 24dB, 18dB, 12dB, 06dB, 00dB y -06Db con el algoritmo de delineación específico EWMA de dos Etapas.**

| EWMA DE DOS ETAPAS | | | | | | |
|---|---|---|---|---|---|---|
| | 118em 24dB | 118em 18dB | 118em 12dB | 118em 06dB | 118em 00dB | 118em -06dB |
| S (%) | 100 | 100 | 99.385 | 87.05 | 71.115 | 62.643 |
| PPV (%) | 100 | 100 | 98.179 | 85.401 | 74.723 | 67.375 |
| DER (%) | 0 | 0 | 2.4284 | 27.304 | 55.627 | 72.805 |
| AC (%) | 100 | 100 | 97.586 | 75.774 | 57.325 | 48.063 |
| TP | 2278 | 2278 | 2264 | 1983 | 1620 | 1427 |
| FN | 0 | 0 | 14 | 295 | 658 | 851 |
| FP | 0 | 0 | 42 | 339 | 548 | 691 |

TABLA 11. **Evaluación de la relación señal/ruido (SNR) en el registro "118 em" con 24dB, 18dB, 12dB, 06dB, 00dB y -06Db con el algoritmo de delineación específico EWMA simple (Es).**

| EWMA SIMPLE (Es) | | | | | | |
|---|---|---|---|---|---|---|
| | 118em 24dB | 118em 18dB | 118em 12dB | 118em 06dB | 118em 00dB | 118em -06dB |
| S (%) | 99.693 | 99.693 | 98.200 | 72.256 | 4.302 | 2.502 |
| PPV (%) | 100.000 | 100.000 | 98.982 | 86.769 | 19.444 | 10.795 |
| DER (%) | 0.308 | 0.308 | 2.832 | 46.547 | 513.095 | 509.848 |
| AC (%) | 99.693 | 99.693 | 97.219 | 65.085 | 3.651 | 2.073 |
| TP | 2271 | 2271 | 2237 | 1646 | 98 | 57 |
| FN | 7 | 7 | 41 | 632 | 2180 | 2221 |
| FP | 0 | 0 | 23 | 251 | 406 | 471 |

TABLA 12. **Evaluación de la relación señal/ruido (SNR) en el registro "118 em" con 24dB, 18dB, 12dB, 06dB, 00dB y -06dB sin EWMA (umbral fijo).**

| SIN EWMA (UMBRAL FIJO=0.62) | | | | | | |
|---|---|---|---|---|---|---|
| | 118em 24dB | 118em 18dB | 118em 12dB | 118em 06dB | 118em 00dB | 118em -06dB |
| S (%) | 99.254 | 89.947 | 34.987 | 12.687 | 2.678 | 1.536 |
| PPV (%) | 99.735 | 97.386 | 78.445 | 50.000 | 16.095 | 10.386 |
| DER (%) | 1.015 | 13.498 | 167.323 | 394.118 | 668.865 | 755.193 |
| AC (%) | 98.993 | 87.827 | 31.918 | 11.258 | 2.350 | 1.357 |
| TP | 2261 | 2049 | 797 | 289 | 61 | 35 |
| FN | 17 | 229 | 1481 | 1989 | 2217 | 2243 |
| FP | 6 | 55 | 219 | 289 | 318 | 302 |

En comparación, el EWMA simple (Es) presentó un rendimiento aceptable en niveles bajos de ruido, con una sensibilidad del 99,69 % en 24 dB y 18 dB, pero su rendimiento disminuyó drásticamente en entornos ruidosos, llegando a una sensibilidad del 4,30 % a 0 dB y 2,51 % a -6 dB. El enfoque sin EWMA (umbral fijo) mostró aún menos resiliencia, con una caída de la sensibilidad a 12,69 % en 6 dB y 1,54 % en -6 dB. El valor predictivo positivo (PPV) siguió una tendencia similar. El EWMA de dos etapas alcanzó un PPV perfecto del 100 % en condiciones de bajo ruido, pero descendió progresivamente a 67,38 % en -6 dB. El EWMA simple también mantuvo un alto PPV en entornos de bajo ruido, pero experimentó una fuerte disminución en ruido severo, con solo 10,80 % en -6 dB.

El umbral fijo mostró una caída considerable en el PPV, alcanzando 10,39 % a -6 dB. En cuanto a la tasa de error (DER %), el EWMA de dos etapas mostró una tasa de error del 0 % en condiciones de bajo ruido, pero aumentó a 47,94 % a -6 dB. El EWMA simple presentó un incremento aún más pronunciado, pasando de 0,31 % en 24 dB y 18 dB a 509,85 % en -6 dB, mientras que el umbral fijo mostró el peor desempeño, con una tasa de error de hasta 755,19 % en -6 dB. La exactitud (AC %) mostró que el EWMA de dos etapas mantuvo una excelente exactitud en condiciones de bajo ruido (100 %), pero disminuyó a 48,21 % en -6 dB. Tanto el EWMA simple como el umbral fijo experimentaron una mayor degradación, con solo 2,07 % y 1,36 %, respectivamente, en -6 dB.

Finalmente, al observar las tasas de verdaderos positivos (TP), falsos negativos (FN) y falsos positivos (FP), el EWMA de dos etapas mostró un rendimiento destacado, sin falsos negativos hasta un SNR de 12 dB. Sin embargo, a medida que aumentó el ruido, también lo hicieron los falsos positivos, alcanzando 691 a -6 dB. Tanto el EWMA simple como el umbral fijo experimentaron problemas más severos con los FN y FP en niveles altos de ruido, con hasta 2221 FN y 2243 FN, y 302 FP y 471 FP, respectivamente, en -6 dB. En conclusión, el algoritmo EWMA de dos etapas muestra una mejor resistencia al ruido que el EWMA simple y el umbral fijo, manteniendo un rendimiento superior en términos de sensibilidad, PPV y tasa de error en niveles de ruido moderados a altos, demostrando su efectividad para la delineación en diversas condiciones, como se muestra en la Tabla 10.

## CONCLUSIONES

En esta investigación, se ha diseñado un algoritmo de delineación específico con umbral adaptativo basado en el Modelo de Suavizado Exponencial Ponderado (EWMA) de dos etapas. Las ventajas de este algoritmo incluyen su simplicidad, ya que no requiere recalibraciones constantes para mejorar la precisión de la presión arterial, su baja carga computacional, robustez y la capacidad de ajustar el factor de suavización. Esto permite equilibrar la sensibilidad y flexibilidad de los umbrales fijos, mejorando la detección de picos, lo cual es crucial para identificar picos en señales ECG y PPG en dispositivos portátiles.

El algoritmo EWMA de dos etapas ha demostrado un buen desempeño y robustez en presencia de ruido, siendo efectivo en condiciones de SNR como 24 dB, 18 dB, 12 dB, 6 dB, 0 dB y -6 dB. En comparación con el algoritmo EWMA simple (Es) y el enfoque sin EWMA (umbral fijo), el EWMA de dos etapas muestra un rendimiento mucho más sólido y consistente en diferentes niveles de SNR, especialmente en entornos con alto ruido, como en 6 dB, donde los otros enfoques enfrentan grandes dificultades. Esto convierte al EWMA de dos etapas en la mejor opción para aplicaciones que requieren alta fiabilidad en condiciones de salud. Los resultados preliminares respaldan la eficacia de nuestro algoritmo, con una sensibilidad del 100 % y un Valor Predictivo Positivo del 100 % en alto SNR, lo que sugiere su aplicabilidad para la estimación precisa de la presión arterial no invasiva. Además, se obtuvieron resultados satisfactorios utilizando un modelo matemático de estimación evaluado en 5 registros de la base de datos MIMIC, que incluyen señales ECG, PPG y mediciones de presión arterial. Los resultados mostraron un error medio de -1.422 mmHg para la presión arterial sistólica (SBP) y 0.577 mmHg para la presión arterial diastólica (DBP), con una desviación estándar (s.d.) de 4.668 mmHg para la SBP y 2.888 mmHg para la DBP. Estos valores cumplen con los estándares de la Asociación para el Avance de la Instrumentación Médica (AAMI), que establece límites de 5 mmHg para el error medio (MAE) y 8 mmHg para la desviación estándar (S.D.).

El algoritmo de delineación específico EWMA de dos etapas es una opción prometedora para estimar la presión arterial de manera no invasiva con alta precisión, adaptándose bien a los cambios rápidos en entornos de monitorización ambulatoria para dispositivos portátiles. Continuamos trabajando en superar las limitaciones actuales de medición y hacer que estas mediciones sean más accesibles tanto en entornos clínicos como ambulatorios.

## AGRADECIMIENTOS

Los autores desean expresar su más profundo y sincero agradecimiento a la Universidad Nacional de San Agustín de Arequipa por el invaluable apoyo brindado a lo largo de esta investigación. Su compromiso y dedicación han sido fundamentales para el desarrollo y éxito de este proyecto. Este trabajo forma parte del proyecto de investigación “Plataforma de monitoreo integral e inalámbrico para la detección y seguimiento efectivo de enfermedades cardiovasculares en pacientes geriátricos mediante sensores ópticos no invasivos”, el cual ha sido financiado por la

Universidad Nacional de San Agustín de Arequipa bajo el contrato número IBA-IB-63-2020-UNSA.

## CONTRIBUCIÓN DE LOS AUTORES

L. N. conceptualización, curación de datos, análisis formal, *software*, visualización, investigación, metodología, administración de proyectos, redacción y validación; R. V. curación de datos, análisis formal, investigación, *software*, visualización y redacción; E. S. supervisión, revisión, adquisición de financiación y provisión de recursos; J. R. supervisión, revisión, adquisición de financiación y provisión de recursos.

## CONFLICTO DE INTERÉS

Los autores declaran que no existe ningún conflicto de intereses en relación con este trabajo de investigación.

## REFERENCIAS

**[1]** M. Kachuee, M. M. Kiani, H. Mohammadzade, M. Shabany, "Cuffless blood pressure estimation algorithms for continuous health-care monitoring," IEEE Trans. Biomed. Eng., vol. 64, no. 4, pp. 859-869, 2017, doi: https://doi.org/10.1109/TBME.2016.2580904

**[2]** W. H. Lin et al., "Towards accurate estimation of cuffless and continuous blood pressure using multi-order derivative and multivariate photoplethysmogram features," Biomed. Signal Process. Control, vol. 63, 2021, art. no. 102198, doi: https://doi.org/10.1016/j.bspc.2020.102198

**[3]** O. Litvinova et al., "Patent landscape review of non-invasive medical sensors for continuous monitoring of blood pressure and their validation in critical care practice," Front. Med., vol. 10, 2023, art. no. 1138051, doi: https://doi.org/10.3389/fmed.2023.1138051

**[4]** M. Salem et al., "Non-invasive data acquisition and IoT solution for human vital signs monitoring: Applications, limitations and future prospects," Sensors, vol. 22, no. 17, 2022, art. No. 6625, doi: https://doi.org/10.3390/s22176625

**[5]** Y. Wang et al., "Recent advancements in flexible and wearable sensors for biomedical and healthcare applications," J. Phys. D: Appl. Phys., vol. 55, no. 13, 2022, art. no. 103002, doi: https://doi.org/10.1088/1361-6463/ac3c73

**[6]** A. Medina et al., "Continuous blood pressure estimation in wearable devices using photoplethysmography: A review," Int. J. Emerg. Technol. Adv. Eng., vol. 12, no. 10, pp. 104-113, 2022, doi: https://doi.org/10.46338/ijetae1022_12

**[7]** M. Y. M. Wong, E. Pickwell-MacPherson, Y. T. Zhang, J. C. Y. Cheng, "The effects of pre-ejection period on post-exercise systolic blood pressure estimation using the pulse arrival time technique," Eur. J. Appl. Physiol., vol. 111, no. 1, pp. 135-144, 2011, doi: https://doi.org/10.1007/s00421-010-1626-0

**[8]** T. Vajanarat, A. Lek-uthai, "A comparison of cuff-less blood pressure estimation between pulse arrival time and pulse transit time using photoplethysmography," en 17th Int. Conf. Electr. Eng./Electron., Comput., Telecommun. Inf. Technol. (ECTI-CON), Phuket, Tailandia, 2020, pp. 13-16, doi: https://doi.org/10.1109/ECTI-CON49241.2020.9158100

**[9]** J. Cano, "Desarrollo de un método de estimación de la presión arterial combinando fotopletismografía y electrocardiografía," Tesis de grado, Dept. Ing. Electron., Universidad Politécnica de Valencia, Valencia, España, 2020. [En línea]. Disponible en: http://hdl.handle.net/10251/150068

**[10]** F. S. Cattivelli, H. Garudadri, "Noninvasive cuffless estimation of blood pressure from pulse arrival time and heart rate with adaptive calibration," en Proc. Int. 6th Workshop Wearable Implantable Body Sensor Netw., Berkeley, CA, USA, 2009, pp. 114-119, doi: https://doi.org/10.1109/BSN.2009.35

**[11]** J. M. Bote, J. Recas, R. Hermida, M. Díaz-Vicente, "Implementación de técnicas de estimación de la presión arterial a partir de ECG y PPG," en I Jornadas de Computación Empotrada y Reconfigurable, Jornadas SARTECO (CEDI), Salamanca, España, 2016, pp. 551-559.

**[12]** J. M. Bote et al., "A modular low complexity ECG delineation algorithm for real-time embedded systems," IEEE J. Biomed. Health Inform., vol. 22, no. 2, pp. 429-441, 2018, doi: https://doi.org/10.1109/JBHI.2017.2671443

**[13]** S. Nurmaini et al., "Robust electrocardiogram delineation model for automatic morphological abnormality interpretation," Sci. Rep., vol. 13, 2023, art. no. 13736, doi: https://doi.org/10.1038/s41598-023-40965-1

**[14]** J. Bote, "Algoritmos de monitorización de señales biomédicas en tiempo real para dispositivos portátiles de bajo consumo," Tesis de Doctorado, Fac. Inform., Univ. Complutense de Madrid, Madrid, España, 2021. [En línea]. Disponible en: https://hdl.handle.net/20.500.14352/5570

**[15]** Z. Hao, X. Zhang, Z. Lai, "Adaptive R-peak detection algorithm based on brown exponential smoothing model," IEEE Access, vol. 10, pp. 114355-114363, 2022, doi: https://doi.org/10.1109/ACCESS.2022.3218308

**[16]** X. Wan et al., "Electrocardiogram baseline wander suppression based on the combination of morphological and wavelet transformation based filtering," Comput. Math. Methods Med., vol. 2019, no. 1, 2019, art. no. 196156, doi: https://doi.org/10.1155/2019/7196156

**[17]** F. R. Hashim et al., “Electrocardiogram noise cancellation using wavelet transform,” J. Fundam. Appl. Sci., vol. 9, no. 3S, pp. 131-140, 2017, doi: https://doi.org/10.4314/jfas.v9i3s.11

**[18]** Md. A. Kabir, C. Shahnaz, “Denoising of ECG signals based on noise reduction algorithms in EMD and wavelet domains,” Biomed. Signal Process. Control, vol. 7, no. 5, pp. 481-489, 2012, doi: https://doi.org/10.1016/j.bspc.2011.11.003

**[19]** Y. Hong, Y. Lian, “A Memristor-based continuous-time digital FIR filter for biomedical signal processing,” IEEE Trans. Circuits Syst. I Reg. Papers, vol. 62, no. 5, pp. 1392-1401, 2015, doi: https://doi.org/10.1109/TCSI.2015.2403033

**[20]** J. Pan, W. J. Tompkins, “A real-time QRS detection algorithm,” IEEE Trans. Biomed. Eng., vol. BME-32, no. 3, pp. 230-236, 1985, doi: https://doi.org/10.1109/TBME.1985.325532

**[21]** J. A. Van Alste, T. S. Schilder, “Removal of base-line wander and power-line interference from the ECG by an efficient FIR filter with a reduced number of taps”, IEEE Trans. Biomédica. Ing., vol. 32, no. 12, pp. 1052-1060, 1985, doi: https://doi.org/10.1109/TBME.1985.325514

**[22]** J. Park, S. Lee, U. Park, “R peak detection method using wavelet transform and modified Shannon energy envelope,” J. Healthc. Eng., vol. 2017, 2017, art. no. 4901017, doi: https://doi.org/10.1155/2017/4901017

**[23]** K. Krishnakumar, S. Latha, “Different methods to remove baseline wandering noise from the ecg signal: a research perspective,” en Proc. 3rd Int. Conf. Artif. Intell. Smart Energy (ICAIS), Coimbatore, India, 2023, doi: https://doi.org/10.1109/ICAIS56108.2023.10073914

**[24]** T. N. Nguyen, T. H. Nguyen, B. V. Ngo, “R peak determination using a WDFR algorithm and adaptive threshold,” Appl. Comput. Sci., vol. 18, no. 3, pp. 19-30, 2022, doi: http://dx.doi.org/10.35784/acs-2022-18

**[25]** T. Rodrigues, S. Samoutphonh, H. Silva, A. Fred, “A low-complexity R-peak detection algorithm with adaptive thresholding for wearable devices,” en 2020 25th International Conference on Pattern Recognition (ICPR), Milan, Italia, 2021, doi: https://doi.org/10.1109/ICPR48806.2021.9413245

**[26]** M. Shahbakhti et al., “Fusion of EEG and eye blink analysis for detection of driver fatigue,” IEEE Trans. Neural Syst. Rehabil. Eng., vol. 31, pp. 2037-2046, 2023, doi: https://doi.org/10.1109/TNSRE.2023.3267114

**[27]** J. Arteaga-Falconi, H. Al Osman, A. El Saddik, “R-peak detection algorithm based on differentiation,” en Proc. 2015 IEEE 9th Int. Symp. Intell. Signal Process. (WISP), Siena, Italia, 2015, doi: https://doi.org/10.1109/WISP.2015.7139147

**[28]** PCP Chao et al., “Machine learning leading cuffless PPG blood pressure sensors into the next stage,” IEEE Sens. J., vol. 21, no.11, pp. 12498-12510, 2021, doi: https://doi.org/10.1109/JSEN.2021.3073850

**[29]** H. Tan et al., “Heartbeat-aware convolutional neural network for R-peak detection of wearable device ECG data,” J. South. Med. Univ., vol. 42, no. 3, pp. 375-383, 2022, doi: https://doi.org/10.12122/j.issn.1673-4254.2022.03.09

**[30]** L. Khriji, A. M. Al-Busaidi, "New adaptive thresholding-based ECG R-peak detection technique," en Proc. IEEE Middle East Conf. Biomed. Eng. (MECBME), Túnez, Túnez, 2018, doi: https://doi.org/10.1109/MECBME.2018.8402423

**[31]** A. Hamed, S. Saeid, M. Karim, “A novel signal segmentation method based on standard deviation and variable threshold,” Int. J. Comput. Appl., vol. 34, no. 2, pp. 27-34, 2011. [En línea]. Disponible en: https://www.ijcaonline.org/archives/volume34/number2/4073-5860/

**[32]** M. Niaz, M. Khan, “Pan-Tompkins: a robust approach to detect R-peaks in ECG signals,” en Proc. IEEE Int. Conf. Bioinf. Biomed. (BIBM), Las Vegas, NV, USA, 2022, pp. 2905-2912, doi: https://doi.org/10.1109/BIBM55620.2022.9995552

**[33]** J. Yu, S. B. Kim, J. Bai, S. W. Han, “Comparative study on exponentially weighted moving average approaches for the self-starting forecasting,” Appl. Sci., vol. 10, no. 20, 2020, no. art. 7351, doi: https://doi.org/10.3390/app10207351

**[34]** P. R. Winters, “Forecasting sales by exponentially weighted moving averages,” Manage. Sci., vol. 6, no. 3, pp. 324-342, 1960. [En línea]. Disponible en: https://www.jstor.org/stable/2627346

**[35]** T. Andrasto et al., “Method EWMA (Exponentially Weighted Moving Average) as a filter to fine and remove noise on time series data,” en 11th Engineering International Conference: Applied Green Technology for Environment Conservation Through Continous Engineering (EIC 2022)., Indonesia, 2023, art. no. 012012, doi: https://doi.org/10.1088/1755-1315/1203/1/012012

**[36]** A. D. Little, R. G. Brown. (1956). Exponential smoothing for predicting demand. [En línea]. Disponible: https://www.industrydocuments.ucsf.edu/docs/jzlc0130

**[37]** G. T. Wilson, “Book review: time series analysis: forecasting and control, 5th edition, by George E. P. Box, Gwilym M. Jenkins, Gregory C. Reinsel and Greta M. Ljung, 2015. Published by John Wiley and Sons Inc., Hoboken, New Jersey, pp. 712. ISBN: 978-1-118-67502-1,” J. Time Ser. Anal., vol. 37, pp. 709-711, 2016, doi: https://doi.org/10.1111/jtsa.12194

**[38]** V. Akgiray, “Conditional heteroscedasticity in time series of stock returns: evidence and forecasts,” J. Bus., vol. 62, no. 1, pp. 55-80, 1989, doi: http://doi.org/10.1086/296451

**[39]** M. W. Watson, “Chapter 47 Vector autoregressions and cointegration,” in Handbook of Econometrics, 1994, pp. 2843-2915, doi: https://doi.org/10.1016/S1573-4412(05)80016-9

**[40]** Y. LeCun, Y. Bengio, G. Hinton, “Deep learning,” Nature, vol. 521, págs. 436-444, 2015, doi: https://doi.org/10.1038/nature14539

**[41]** P. Jarrín, J. Xavier, “Comparación entre varios métodos de pronósticos basados en series de tiempo para predecir la demanda de placas digitales en empresas del sector gráfico quiteño desde el año 2009 hasta el año 2015,” Tesis de maestría, Facultad de Ciencias Administrativas, Escuela Politécnica Nacional, Ecuador, 2015. [En línea]. Disponible en: http://bibdigital.epn.edu.ec/handle/15000/17016

**[42]** R. Huamani et al., “Implementation of a real-time 60 Hz interference cancelation algorithm for ECG signals based on ARM Cortex M4 and ADS1298,” en

Proc. IEEE 24th Int. Congress on Electron. Elect. Eng. Comput. (INTERCON), Cusco, Peru, 2017, doi: https://doi.org/10.1109/INTERCON.2017.8079725

**[43]** E. Supo et al., "PRD as an indicator proposal in the evaluation of ECG signal acquisition prototypes in real patients," Technol. Innov. Andean Ind., Barranquilla, Colombia, 2022, doi: https://doi.org/10.1109/ICAIS56108.2023.10073914

**[44]** C. Huisa et al., "PCG heart sounds quality classification using neural networks and SMOTE Tomek links for the think health project," en Proc. Data Anal. Manag. Lect. Notes Netw. Syst., A. Khanna, Z. Polkowski, O. Castillo, Eds. 2023, vol. 572, pp. 803-811, doi: https://doi.org/10.1007/978-981-19-7615-5_65

**[45]** R. Sulla, J. Talavera, E. Supo, A. Montoya, "Non-invasive glucose monitor based on electric bioimpedance using AFE4300," en Proc. IEEE 26th Int. Conf. Electron., Elect. Eng. Comput. (INTERCON), Lima, Peru, 2019, doi: https://10.1109/INTERCON.2019.8853561

**[46]** J. M. Bote, J. Recas, R. Hermida, "Evaluation of blood pressure estimation models based on pulse arrival time," Comput. Elect. Eng., vol. 84, 2020, no. art. 106616, doi: https://doi.org/10.1016/j.compeleceng.2020.106616

**[47]** A. Johnson, T. Pollard, R. Mark, 2016, "MIMIC-III Clinical Database (1.4 version)," PhysioNet, doi: https://doi.org/10.13026/C2XW26

**[48]** I. Silva, G. Moody, "An open-source toolbox for analysing and processing PhysioNet databases in MATLAB and Octave", J. Open Res. Softw., vol. 2, no. 1, 2014, no. art. e27, doi: https://doi.org/10.5334/jors.bi

**[49]** A. H. Sayed, Fundamentals of Adaptive Filtering, Hoboken, NJ, USA: Wiley, 2003.

**[50]** P. Laguna, R. Mark, A. Goldberg, G. Moody, "A database for evaluation of algorithms for measurement of QT and other waveform intervals in the ECG," en Conf. Comput. Cardiol. 1997, Lund, Suecia, 1997, pp. 673-676, doi: https://doi.org/10.1109/CIC.1997.648140

**[51]** A. Johnson et al., "MIMIC-III, a freely accessible critical care database," Scientific Data, vol. 3, 2016, no. art. 60035, doi: https://doi.org/10.1038/sdata.2016.35

**[52]** A. Goldberger et al., "PhysioBank, PhysioToolkit, and PhysioNet : components of a new research resource for complex physiologic signals," Circulation, vol. 101, no. 23, pp. e215-e220, 2000, doi: https://doi.org/10.1161/01.CIR.101.23.e215

**[53]** D. J. Hughes, C. F. Babbs, L. A. Geddes, J. D. Bourland, "Measurements of Young's modulus of elasticity of the canine aorta with ultrasound," Ultrason. Imag., vol. 1, no. 4, pp. 356-367, 1979, doi: https://doi.org/10.1016/0161-7346(79)90028-2

**[54]** W. Chen et al., "Continuous estimation of systolic blood pressure using the pulse arrival time and intermittent calibration," Med. Biol. Eng. Comput., vol. 38, pp. 569-574, 2000, doi: https://doi.org/10.1007/BF02345755

**[55]** P. K. Baheti, H. Garudadri, "Heart rate and blood pressure estimation from compressively sensed photoplethysmograph," en Proc. 4th Int. ICST Conf. on Body Area Netw., Los Angeles, CA, USA, 2009, doi: http://dx.doi.org/10.4108/ICST.BODYNETS2009.6023

**[56]** C. C. Y. Poon, Y. T. Zhang, "Cuff-less and Noninvasive Measurements of Arterial Blood Pressure by Pulse Transit Time," en Proc. 27th Annu. IEEE Eng. Med. Biol. Conf., vol. 6, Shanghai, China, 2006, pp. 5877-5880, doi: https://doi.org/10.1109/IEMBS.2005.1615827

**[57]** C. P. Chua, C. Heneghan, "Continuous Blood Pressure Monitoring using ECG and Finger Photoplethysmogram," en Proc. 2006 Int. Conf. IEEE Eng. Med. Biol. Soc., NY, USA, 2006, pp. 5117-5120, doi: https://doi.org/10.1109/IEMBS.2006.259612

**[58]** P. Fung et al., "Continuous noninvasive blood pressure measurement by pulse transit time," en Proc. 26th Ann. Int. Conf. IEEE Eng. Med. Biol. Soc., San Francisco, CA, USA, 2004, pp. 738-741, doi: https://doi.org/10.1109/IEMBS.2004.1403264

**[59]** A. A. Awad et al., "How Does the Plethysmogram Derived from the Pulse Oximeter Relate to Arterial Blood Pressure in Coronary Artery Bypass Graft Patients?," Anesth. Analg., vol. 93, no. 6, pp. 1466-1471, 2001, doi: https://doi.org/10.1097/00000539-200112000-00022