

Proposal for the Design of an Electronic Auscultation System for the Acquisition and Digital Analysis of Heart Sounds

José Manuel Romero Reyes¹ 

¹Universidad Tecnológica Centroamericana (UNITEC), Honduras, josemanuelrr87@unitec.edu

Abstract— *Cardiac auscultation is a fundamental procedure in clinical practice for the early detection of cardiovascular diseases. However, in primary care services, it faces limitations due to environmental noise and variability in the interpretation of heart sounds. Given this problem, this research seeks to propose the design of a prototype electronic cardiac auscultation device with digital recording and analysis capabilities using MATLAB, to identify cardiac abnormalities that support diagnosis in primary care settings. A two-phase methodology was applied. In the study phase, the electronic stethoscope prototype was built, and the digital processing system was developed based on a recurrent neural network implemented in MATLAB for the classification of heart sounds. In the validation phase, the device was evaluated under simulated conditions with available recordings and by comparison with the diagnosis made by medical personnel through traditional auscultation. The results obtained show that an electronic device is capable of reliably capturing heart sounds and performing basic classification of normal and abnormal recordings in the identification of murmurs. The study showed that a low-cost prototype can be technically viable to complement traditional auscultation and that, with technical improvements and broader clinical validation, these types of devices have the potential to become a diagnostic support tool in primary care.*

Index Terms— *Automatic Classification, Cardiac Auscultation, Electronic Stethoscope, MATLAB, Primary Care, Signal Processing.*

Propuesta de Diseño de Sistema de Auscultación Electrónico para la Adquisición y Análisis Digital de Sonidos Cardíacos

José Manuel Romero Reyes¹ 

¹Universidad Tecnológica Centroamericana (UNITEC), Honduras, josemanuelrr87@unitec.edu

Resumen—La auscultación cardíaca es un procedimiento fundamental en la práctica clínica para la detección temprana de enfermedades cardiovasculares; sin embargo, en servicios de atención primaria, enfrenta limitaciones derivadas del ruido ambiental y la variabilidad en la interpretación de los sonidos cardíacos. Ante esta problemática, la presente investigación busca plantear el diseño de un prototipo de dispositivo electrónico de auscultación cardíaca con capacidad de registro digital y análisis de sonidos cardíacos mediante MATLAB, con el fin de identificar anomalías cardíacas que apoyen al diagnóstico en entornos de atención primaria. Se aplicó una metodología dividida en dos fases. En la fase de estudio se construyó el prototipo de estetoscopio electrónico y se desarrolló el sistema de procesamiento digital basado en una red neuronal recurrente implementada en MATLAB para la clasificación de sonidos cardíacos. En la fase de validación, el dispositivo fue evaluado en condiciones simuladas con registros disponibles y mediante la comparación con el diagnóstico realizado por personal médico a través de auscultaciones tradicionales. Los resultados obtenidos muestran que un dispositivo electrónico es capaz de realizar la captura fiable de sonidos cardíacos y la clasificación básica de registros normales y anormales en la identificación de soplos. El estudio evidenció que un prototipo de bajo costo puede ser técnicamente viable para complementar la auscultación tradicional y que, con mejoras técnicas y validaciones clínicas más amplias, este tipo de dispositivos tienen el potencial de convertirse en una herramienta de apoyo diagnóstico en la atención primaria.

Palabras clave—Estetoscopio Electrónico, Auscultación Cardíaca, Procesamiento de Señales, MATLAB, Atención Primaria, Clasificación Automática.

I. INTRODUCCIÓN

La auscultación ha sido una herramienta diagnóstica fundamental en la práctica médica, permitiendo a los profesionales detectar anomalías cardíacas mediante la escucha de los sonidos del corazón. Sin embargo, a pesar de su valor clínico, su efectividad depende en gran medida de la experiencia, agudeza auditiva y formación del profesional que la realiza. Esta alta dependencia representa una limitación importante para la precisión del diagnóstico. Incluso en contextos clínicos controlados, la auscultación presenta limitaciones en la sensibilidad para la detección de enfermedades cardíacas como las valvulopatías [1].

Estas limitaciones se vuelven más críticas en entornos de atención primaria o zonas de escasos recursos donde no siempre se cuenta con personal especializado o suficientemente capacitado. Actualmente, existe una creciente preocupación sobre la brecha entre la formación médica tradicional y los

desafíos en las interpretaciones de sonidos cardiopulmonares. En el 2024 se realizó un estudio en que se evaluó la capacidad de médicos en formación para interpretar fonocardiogramas y se encontró que la mayoría presentaba dificultades para identificar correctamente los sonidos S1 y S2, los cuales son fundamentales para establecer diagnósticos precisos [2]. Esto refuerza el hecho de que la auscultación tradicional no solo presenta limitaciones técnicas en su aplicación, sino que es vulnerable a errores humanos.

El problema no se limita al componente clínico o humano, existen obstáculos técnicos significativos que afectan la calidad de la captación de sonidos auscultatorios. Uno de los principales es la contaminación por ruido ambiental, especialmente en zonas rurales o centros asistenciales primarios, donde las condiciones no favorecen la adquisición clara de los sonidos cardíacos. Un estudio mostró cómo la calidad de la señal de diferentes modelos de estetoscopios electrónicos se ve comprometida en entornos clínicos ruidosos [3].

La auscultación cardíaca sigue siendo esencial para la detección temprana de enfermedades cardiovasculares, pero enfrenta limitaciones por la falta de personal especializado y el ruido ambiental, lo que dificulta el diagnóstico sobre todo en entornos de atención primaria con recursos limitados. La incorporación de tecnologías digitales, como algoritmos de aprendizaje profundo y estetoscopios electrónicos, ha demostrado mejorar la precisión diagnóstica y facilitar esquemas de teleauscultación, aunque persisten retos relacionados con la calidad de la señal y la clasificación automática [4].

En este contexto, la presente investigación se orienta al diseño y desarrollo de un prototipo de estetoscopio electrónico con capacidad de registro digital y análisis de señales acústicas mediante técnicas de procesamiento de señales. El dispositivo se plantea como una alternativa de bajo costo y aplicable a escenarios clínicos donde los recursos son limitados y se requiere fortalecer la capacidad diagnóstica del personal médico en el primer nivel de atención. Con este fin, también se plantea comparar la precisión diagnóstica frente a la auscultación tradicional, identificar características técnicas para entornos con ruido, determinar técnicas de procesamiento aplicables y analizar la viabilidad de adopción por parte del personal médico.

II. SUSTENTO TEÓRICO

A. La Auscultación Cardíaca

Los sonidos cardíacos se originan a partir del cierre y apertura de las válvulas del corazón para permitir el flujo de sangre unidireccional a través de sus cámaras. El proceso por medio del cual se realiza la evaluación de estos sonidos se conoce como auscultación cardíaca. La auscultación cardíaca consiste en el uso de un instrumento, como el estetoscopio, que permite oír los ruidos que surgen del corazón a través de los tejidos. La auscultación es una técnica muy utilizada en la atención médica ya que permite realizar una revisión rápida de la salud cardíaca del paciente sin la necesidad de exámenes más complejos o métodos más invasivos. El cierre súbito de las válvulas y las turbulencias del flujo sanguíneo generan ondas sonoras que se propagan por las estructuras anatómicas. Estas ondas son las que permiten que los sonidos cardíacos puedan ser percibidos por el examinador. Durante el análisis clínico de los sonidos cardíacos se debe evaluar su frecuencia, intensidad, duración, timbre, localización y sincronía con el ciclo cardíaco [5]. Por esta razón, el entrenamiento y la calidad de los instrumentos usados son muy importantes para hacer estas evaluaciones.

Durante la auscultación cardíaca se pueden percibir diversos sonidos que reflejan los eventos hemodinámicos que ocurren dentro del corazón a través del ciclo cardíaco. En función de su origen y características acústicas, los sonidos pueden ser clasificados en diferentes categorías. Los sonidos normales, conocidos como S1 y S2, corresponden al cierre de las válvulas auriculoventriculares y semilunares respectivamente, reflejando la sístole y diástole del corazón. Por otro lado, sonidos anormales como los soplos cardíacos se producen por turbulencias en el flujo sanguíneo, ocasionadas comúnmente por alteraciones [6], [7]. Su detección es de gran relevancia clínica ya que permiten identificar patologías cardíacas como enfermedades valvulares o defectos septales.

B. El Estetoscopio

El origen de la auscultación moderna se remonta a 1816 con el invento del primer estetoscopio como respuesta a las limitaciones de la auscultación directa, que requería colocar el oído sobre el tórax del paciente. La invención marcó un hito al introducir la auscultación y sentar las bases de una técnica rápidamente adoptada en la medicina clínica. A lo largo del siglo XIX y principios del XX, el estetoscopio evolucionó hacia diseños más ergonómicos y eficientes y la incorporación de materiales más ligeros, membranas acústicas y mejoras en la transmisión de sonido incrementó su utilidad [8].

La evolución tecnológica ha permitido el desarrollo de distintos tipos de estetoscopios, cada uno adaptado a necesidades específicas y con características técnicas que determinan su rendimiento diagnóstico. Los estetoscopios acústicos son los más tradicionales y operan exclusivamente mediante la transmisión mecánica de los sonidos desde el paciente hasta el oído del examinador. La campana o diafragma se coloca sobre la superficie corporal, donde las vibraciones generadas por la

actividad interna ponen en movimiento a la membrana. Este movimiento se transmite a través de una columna de aire en los tubos flexibles hasta las olivas auriculares del examinador, como se observa en la Fig. 1. El diafragma es más eficaz para captar sonidos de alta frecuencia, mientras que la campana se usa para sonidos de baja frecuencia [9].

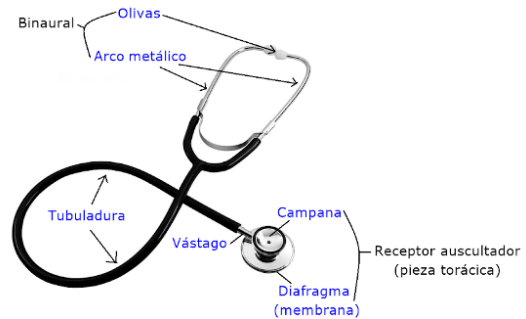


Fig. 1. Partes del Estetoscopio.

Los estetoscopios electrónicos utilizan transductores para convertir las vibraciones acústicas en señales eléctricas. Algunos ejemplos de transductores comúnmente usados son los micrófonos electret, piezoeléctricos o electromecánicos. Las señales eléctricas son amplificadas y filtradas para mejorar la calidad, pudiendo además digitalizarse para su análisis, almacenamiento y transmisión. Algunos modelos incluyen modos de filtrado selectivo para resaltar ciertas frecuencias, así como reducción de ruido ambiental [10].

Aunque ambos tipos de estetoscopios son utilizados en la actualidad, la elección entre un estetoscopio acústico y uno electrónico depende de factores como la experiencia del profesional, las condiciones clínicas y las características del paciente. Estudios han demostrado que los estetoscopios electrónicos mejoran la identificación de sonidos patológicos y la audición en pacientes con obesidad gracias a su capacidad de amplificación y reducción del ruido, especialmente en observadores con menor experiencia. Además, ofrecen ventajas adicionales como el registro digital de sonidos y la posibilidad de realizar teleauscultación [11], [12].

C. Procesamiento de Señales

El procesamiento de señales es una rama de la ingeniería que estudia la representación, transformación y análisis de señales. Una señal puede entenderse como cualquier variación de una magnitud física que transporta información. En el contexto biomédico, las señales fisiológicas requieren de un tratamiento especializado para extraer patrones relevantes, eliminar ruido y mejorar su interpretabilidad [13]. Particularmente, las señales acústicas biomédicas como los fonocardiogramas capturan la actividad mecánica y contienen información diagnóstica que puede complementar hallazgos clínicos obtenidos por otros medios.

El procesamiento de señales sigue un flujo estructurado, ilustrado en la Fig. 2, de etapas que garantizan la integridad de la información y su valor diagnóstico. Aunque la secuencia

puede variar según la aplicación, se pueden reconocer cinco fases principales: adquisición, filtrado y preprocesamiento, extracción de características, segmentación y clasificación.

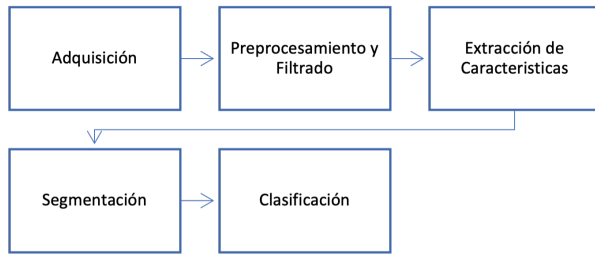


Fig. 2. Etapas del Procesamiento de Señales.

La adquisición convierte la señal analógica en digital, donde la elección del transductor y la frecuencia de muestreo determinan la fidelidad del registro. El preprocesamiento mejora la calidad de la señal mediante filtrado, normalización y eliminación de ruido para optimizar el análisis. Luego, la extracción de características obtiene parámetros temporales, espectrales y estadísticos que permiten reducir la dimensionalidad y facilitar la clasificación. La segmentación divide la señal en partes representativas de eventos específicos, mejorando la precisión en el análisis de sonidos complejos. Finalmente, la clasificación aplica algoritmos estadísticos o de aprendizaje automático, como SVM o redes neuronales, para identificar patrones y asignar diagnósticos con alta sensibilidad y especificidad [14], [15].

III. METODOLOGÍA

La presente investigación fue desarrollada bajo un enfoque mixto, ya que combina la recolección, análisis e interpretación de datos cuantitativos y cualitativos para obtener una visión más amplia y profunda del fenómeno de estudio. El componente cuantitativo permitió medir de forma objetiva variables técnicas asociadas al funcionamiento y desempeño del prototipo para su comparación y evaluación. El componente cualitativo, abordado como estudio de caso, buscó comprender las percepciones, opiniones y experiencias de los potenciales usuarios (médicos de atención primaria) respecto al dispositivo, en un contexto clínico específico como lo es el primer nivel de atención en Honduras. Esta integración metodológica permitió no solo validar parámetros técnicos, sino también valorar la aceptación, facilidad de uso, fortalezas y debilidades del dispositivo en posibles escenarios reales de aplicación.

A. Metodología de Estudio

El método ELECTRE pertenece a la familia de los enfoques de decisión multicriterio como respuesta a las limitaciones de los métodos de suma ponderada tradicionales en problemas complejos de decisión. Este método ha sido ampliamente validado como una herramienta robusta para la selección de alternativas bajo múltiples criterios, con aplicaciones en áreas como la gestión ambiental, la selección de tecnología médica y la optimización de modelos de inteligencia artificial [16], [17]. En esta investigación, se empleó como herramienta para la

selección de alternativas de los componentes de hardware del prototipo y, posteriormente, de las técnicas de procesamiento de señales a implementar. Los pasos a seguir para su realización se ilustran en la Fig. 3.

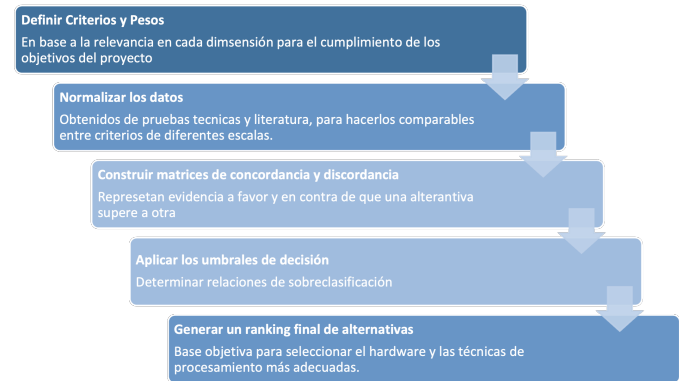


Fig. 3. Esquema de Metodología ELECTRE.

ELECTRE no establece un número mínimo de alternativas, pero se determinó incluir tres opciones como mínimo para evitar inconsistencias en la sobreclasificación. Para los componentes del hardware, las alternativas fueron escogidas con base en su disponibilidad local, su aplicación en otras investigaciones y prototipos de auscultación electrónica y la experiencia previa que se tiene en el uso de estos. La asignación de los pesos para cada criterio se realizó con base en las prioridades de la investigación, siendo las principales la calidad de la señal, la robustez y el costo.

Una vez seleccionados los componentes, se procedió con el diseño y construcción del dispositivo. Se realizó el diseño del esquema eléctrico del dispositivo, considerando las alternativas seleccionadas para cada tipo de componente en la fase anterior. Durante las fases iniciales de prototipado, se realizaron verificaciones preliminares de desarrollo de la calidad de las grabaciones obtenidas. Como criterio de esta evaluación, se estableció que la señal digitalizada debería mantener un SNR lo más cercano a los 10 dB, siendo este considerado como aceptable para identificar S1 y S2 en ambientes con ruido moderado [18].

Para la realización de estas pruebas se establecieron tres situaciones de ruido ambiental diferentes: un ambiente silencioso, un ambiente con ruido medio y un tercero con ruido alto generado por monitores de signos vitales. Los niveles de ruido para cada prueba se establecieron de acuerdo a valores reportados en investigaciones sobre entornos hospitalarios [19]. Se diseñó la distribución espacial de las pruebas, como se observa en la Fig. 4, de manera que se pudiera replicar la exposición al ruido en un ambiente real. El ruido ambiental fue generado con las alarmas de un monitor de signos vitales Comen C50. Las mediciones de nivel sonoro se realizaron con un sonómetro (Q) posicionado a 10 cm del punto de auscultación (P), de manera que se midió el ruido que afecta directamente la captación de sonidos.

Para el desarrollo del algoritmo de clasificación, este fue

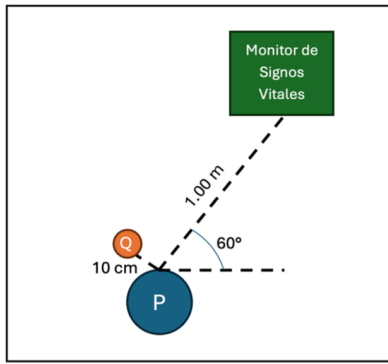


Fig. 4. Configuración del Espacio y Mediciones para Prueba de Configuraciones de Dispositivo de Grabación de Sonidos Cardíacos.

desarrollado en el entorno de MATLAB mediante un proceso estructurado que incluyó la selección de la técnica más adecuada utilizando la metodología ELECTRE. Se evaluaron al menos tres métodos de clasificación basados en las tendencias reportadas en la literatura y en su aplicabilidad al análisis de fonocardiogramas. Dadas las limitaciones de recursos presentes durante la investigación, el estudio se centró en la identificación de soplos cardíacos, debido a su alta relevancia clínica y prevalencia en patologías estructurales, así como la disponibilidad de registros accesibles para el entrenamiento y validación del algoritmo.

B. Metodología de Validación

Con el propósito de validar los resultados de la investigación y comprobar la pertinencia del prototipo desarrollado, se realizó una serie de procedimientos que permitieron evaluar, por un lado, la precisión diagnóstica y la percepción de usabilidad del dispositivo. Para poder evaluar el desempeño diagnóstico del prototipo se realizó un estudio de precisión [20]. Esta metodología se aplicó para comparar el desempeño diagnóstico del prototipo con la evaluación realizada por médicos. Los participantes seleccionados fueron tres médicos capacitados en auscultación cardíaca, de manera que representaran el perfil del personal médico del primer nivel de atención. A partir de los resultados obtenidos de cada método, se calcularon métricas como sensibilidad, especificidad, exactitud y F1-Score. Asimismo, se aplicaron pruebas estadísticas como la prueba de McNemar para evaluar diferencias significativas en las proporciones de aciertos entre métodos.

Las pruebas se realizaron en un Laboratorio de Simulación Avanzada utilizando un simulador SAM como generador estandarizado de sonidos cardíacos. Tanto los médicos como el prototipo realizaron las evaluaciones bajo tres condiciones de ruido ambiental: ambiente silencioso (40 – 50 dBA), ruido por voces humanas (55-60 dBA) y ruido de voces humanas y monitores de signos vitales (70 – 80 dBA) [19]. Todos los médicos emplearon el mismo estetoscopio para evitar sesgos por la variabilidad en el equipo.

Para la prueba se utilizó un monitor de signos vitales Comen C50 y un altavoz con grabaciones de LIBRISPEECH

ASR CORPUS [21] para la creación del ruido ambiental. Las mediciones de ruido ambiental se realizaron desde el punto de auscultación en el simulador SAM. La configuración del espacio y las mediciones realizadas se pueden visualizar en la Fig. 5. Se estableció una muestra de 14 sonidos entre normales y anormales, con una mayor proporción de casos normales para adaptarlos a condiciones comunes de diagnóstico. Dentro de los sonidos anormales se distribuyeron aleatoriamente soplos sistólicos y diastólicos para representar la diversidad de hallazgos en la práctica clínica. Se elaboraron tres listas independientes de sonidos organizadas de manera aleatoria, una para cada situación de ruido ambiental dentro de la prueba, con el propósito de evitar sesgos por predictibilidad.

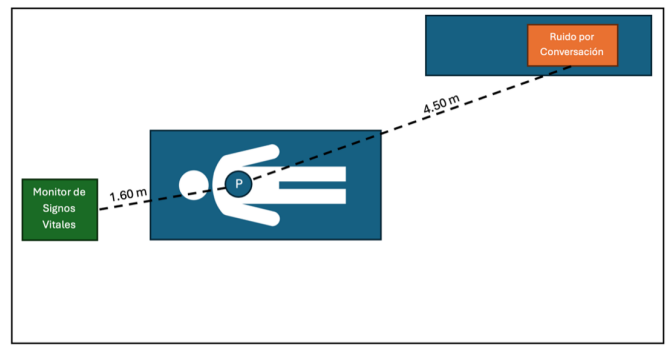


Fig. 5. Configuración del Espacio y Mediciones para Pruebas del Prototipo.

Como segunda parte dentro del proceso de validación, se aplicó una adaptación del *System Usability Scale* (SUS). El SUS es un instrumento estandarizado para medir la usabilidad percibida de sistemas y dispositivos mediante un cuestionario de 10 apartados en escala de Likert de 5 puntos. Su simplicidad y validez psicométrica lo han convertido en una herramienta ampliamente utilizada en la evaluación de la experiencia de usuario. El puntaje final, que varía de 0 a 100, refleja el nivel de facilidad de uso, eficiencia y satisfacción general, siendo 60 el umbral mínimo de aceptabilidad y 80 el indicador de alta usabilidad.

El método SUS fue aplicado a un grupo de profesionales del área de salud que probaron el prototipo. Tras una demostración e interacciones breves, los participantes completaron el cuestionario evaluando aspectos como la facilidad de uso, confianza y disposición para emplear el dispositivo en su práctica clínica. Las respuestas, divididas entre afirmaciones positivas y negativas, se puntuaron para obtener un Índice Global de Usabilidad individual y un promedio grupal.

IV. RESULTADOS Y DISCUSIÓN

A. Selección de Componentes

Para la selección de los componentes del dispositivo de grabación de sonidos cardíacos se aplicó la metodología ELECTRE, considerando criterios técnicos que permitieran equilibrar desempeño y viabilidad. Los elementos principales evaluados fueron el micrófono y el microcontrolador, por ser los más determinantes en la calidad de la adquisición

de la señal. En el caso de los micrófonos, se analizaron tres alternativas: el Electret GY-MAX4466, el micrófono de contacto y el micrófono microelectromecánico (MEMS), seleccionados por su relevancia en investigaciones recientes. Los criterios de evaluación incluyeron la relación señal-ruido (SNR), sensibilidad, banda útil, ruido propio del micrófono, costo y disponibilidad local, con ponderaciones que priorizaran la viabilidad económica y la accesibilidad del componente.

Los resultados mostraron que el micrófono MEMS obtuvo los mejores valores técnicos de SNR y sensibilidad, pero fue superado por el Electret al considerar los factores de costo y disponibilidad local, los cuales son cruciales para la implementación en entornos de atención primaria. El micrófono de contacto, aunque presenta una buena robustez al ruido, obtuvo las puntuaciones más bajas en la mayoría de los criterios. El análisis de concordancia y discordancia confirmó que el Electret superó significativamente al micrófono de contacto y mostró una ventaja moderada sobre el MEMS. Por lo tanto, el micrófono Electret fue seleccionado como la mejor opción, al ofrecer un equilibrio óptimo entre desempeño técnico y viabilidad práctica en contextos de recursos limitados.

En cuanto a los microcontroladores, se evaluaron tres opciones: el ESP32-S3, Arduino Nano y el Raspberry Pi, siguiendo criterios como capacidad de muestreo, resolución ADC, capacidad de procesamiento, facilidad de programación, costo y disponibilidad local. El ESP32-S3 destacó en la mayoría de los parámetros técnicos y obtuvo los mejores puntajes en capacidad de muestreo, resolución y facilidad de programación. Aunque el Raspberry Pi presentó la mayor potencia de cómputo, su costo elevado y menor disponibilidad local lo hicieron menos viable. El Arduino Nano, pese a ser más económico, mostró limitaciones en procesamiento y resolución. El ESP32-S3 fue la opción más equilibrada, al combinar eficiencia técnica, bajo costo y disponibilidad local, lo que lo convierte en el componente ideal para la adquisición y procesamiento del prototipo desarrollado.

B. Diseño del Circuito de Adquisición de Sonidos Cardíacos

Una vez seleccionados los componentes del prototipo, se procedió al diseño y evaluación del circuito eléctrico para la adquisición de sonidos cardíacos. Se probaron tres configuraciones distintas (A, B, C) con el fin de determinar cuál ofrecía la mejor calidad de captura. La configuración A consistió en la conexión directa del micrófono al sistema como referencia básica; la configuración B incorporó un filtro analógico pasa banda de 20-600 Hz para eliminar ruido fuera del rango cardíaco, y la configuración C añadió al filtrado una etapa de amplificación con un amplificador operacional LM358N. Las pruebas se realizaron bajo tres condiciones de ruido ambiental: silencio (44.8 dBA), ruido medio (63.4 dBA) y ruido alto (72.3 dBA).

Los resultados mostraron que la configuración A tuvo valores de SNR bajos en todos los escenarios (-8.89 en silencio y hasta -9.89 dB con ruido medio), evidenciando que la señal sin acondicionamiento queda enmascarada por el ruido. La configuración B mejoró la estabilidad de la señal,

alcanzando -7.52 dB y -7.01 dB en condiciones de ruido medio y alto, confirmando la efectividad del filtrado analógico para eliminar frecuencias indeseadas. Por su parte, la configuración C presentó los valores numéricamente superiores (-4.89 dB en silencio), pero al incorporar la amplificación, también introdujo artefactos que afectaron la claridad de los sonidos cardíacos, lo que redujo la utilidad práctica pese a las mejoras en la relación señal-ruido.

El análisis de espectrogramas complementó los resultados numéricos, mostrando que la configuración A tenía energía difusa en todo el espectro y predominio de ruido de baja frecuencia, mientras que la configuración B (Fig. 6) concentró la energía entre 20 y 200 Hz, permitiendo distinguir los ciclos cardíacos con mayor claridad. En contraste, la configuración C (Fig. 7) presentó picos intensos y discontinuos, con interferencias que dificultaban la interpretación de los sonidos. A partir de estos resultados cuantitativos y cualitativos, se concluyó que la configuración B (Fig. 8) ofrecía el mejor equilibrio entre limpieza, fidelidad y estabilidad de la señal, siendo seleccionada como el diseño más adecuado para el circuito del prototipo.

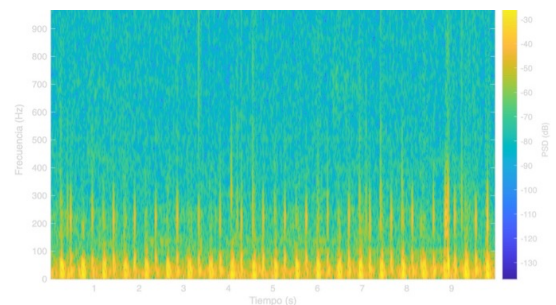


Fig. 6. Espectrograma de la configuración B.

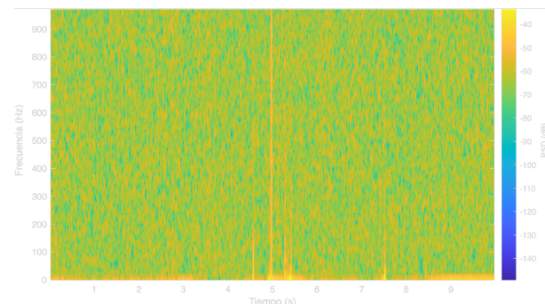


Fig. 7. Espectrograma de la configuración C.

C. Selección de Técnicas de Procesamiento de Sonidos Cardíacos

Para seleccionar las técnicas más adecuadas de procesamiento de señales acústicas cardíacas, se aplicó la metodología ELECTRE considerando ocho alternativas: FFT/STFT, Wavelet, HMM, MFCC-SVM, Random Forest, k-NN, CNN y RNN. Los criterios de evaluación incluyeron

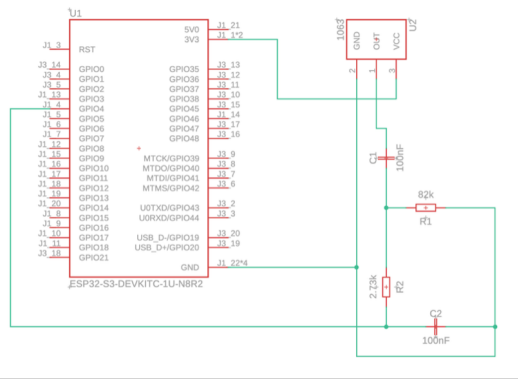


Fig. 8. Circuito de la configuración B.

precisión diagnóstica, robustez al ruido, capacidad de generalización, capacidad de implementación en MATLAB y costo computacional, con ponderaciones respectivas de 0.3, 0.25, 0.2, 0.15 y 0.1. Los resultados mostraron que las redes neuronales profundas alcanzaron los valores más altos en precisión, generalización y robustez, mientras que métodos tradicionales como Wavelet y HMM obtuvieron valores inferiores, evidenciando sus limitaciones en escenarios complejos o con ruido.

Los índices de concordancia confirmaron la superioridad de las técnicas de aprendizaje profundo, ya que las arquitecturas CNN y RNN presentaron mayor consistencia y menor discordancia frente a las demás alternativas. Estos resultados coinciden con estudios previos que destacan el rendimiento de las redes neuronales en tareas de segmentación y clasificación de fonocardiogramas. El ranking final ubicó la RNN como la técnica más adecuada, seguida de la CNN, mientras que Wavelet y HMM mostraron el peor desempeño global, consolidando la elección de modelos de aprendizaje profundo como la opción más eficaz para el análisis de sonidos cardíacos.

D. Desarrollo del Procesamiento Digital y Clasificación de Sonidos Cardíacos

El algoritmo diseñado para la detección y clasificación de sonidos cardíacos se estructuró en una secuencia de etapas que permiten transformar la señal cruda en una decisión diagnóstica automática, esto se puede observar en la Fig. 9. El proceso inicia con la adquisición de la señal cardíaca, la cual pasa por una fase de preprocesamiento para eliminar interferencias y preservar la morfología de los sonidos. Se aplicaron filtros pasa banda de 20-600 Hz, filtros notch para eliminar ruidos eléctricos, un sistema de atenuación de voz de -8 dB y un filtro notch adaptativo para suprimir picos energéticos de alarmas médicas. Además, se incorporó un *denoiser* espectral basado en STFT para reducir el ruido de fondo manteniendo la integridad de los componentes cardíacos.

Posteriormente, la señal preprocesada se sometió a la extracción de características, generando vectores de siete dimensiones por ventana temporal. Estos incluyeron cinco bandas logarítmicas de energía y dos descriptores espectrales

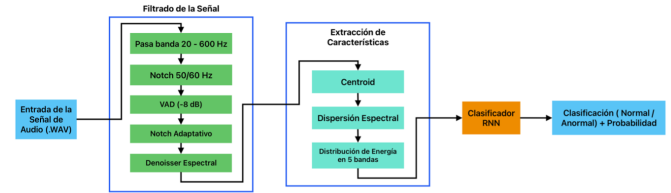


Fig. 9. Diagrama de Bloques de Algoritmo de Análisis de Sonidos Cardíacos.

(centroide y dispersión) que representan la distribución de energía a lo largo del espectro. Estas secuencias se emplearon como entrada a una red neuronal recurrente bidireccional con arquitectura BiLSTM como se observa en la Fig. 10. La red, compuesta por dos capas de 128 unidades cada una, junto con mecanismos de normalización, dropout y softmax ponderado, permitió capturar dependencias temporales y corregir el desbalance entre clases, logrando una adecuada discriminación entre registros normales y patológicos.

El modelo se entrenó con las grabaciones de la base de datos de PhysioNet adaptada por Mathworks [22], dividido en 70% para entrenamiento, 15% para validación y 15% para pruebas. La base de datos proviene de un repositorio de GitHub creado por MathWorks como una adaptación de la base de datos usada en el PhysioNet Computing in Cardiology Challenge 2016. La base contiene un total de 3829 grabaciones, de las cuales 2575 son de personas sanas y 1254 de personas con anomalías cardíacas. Para compensar este desbalance, durante el entrenamiento se implementaron dos mecanismos en la arquitectura de la red. Primero, se asignaron pesos de clase inversamente proporcionales a su frecuencia en la función de pérdida a manera de mitigar las penalizaciones excesivas a la clase minoritaria. En segundo lugar, la capa de salida utilizó un *softmax* ponderado que ajusta las probabilidades de clasificación en función de la distribución real de las clases.

Cada grabación fue muestreada a 2 kHz teniendo así un total de 10,000 muestras por grabación.

Durante el entrenamiento, la exactitud alcanzó un 79% en validación y un 76.5% en pruebas. Los resultados mostraron una sensibilidad del 0.957, especificidad de 0.672 y un F1-Score de 0.76, reflejando buena capacidad de detección de registros patológicos, aunque con cierta tendencia a falsos positivos. Esta tendencia es consecuencia directa de la arquitectura implementada en la red para mitigar el desbalance de muestras en la base de datos. Los mecanismos implementados orientan al clasificador a priorizar la sensibilidad encima de la especificidad. Esto, aunque positivo en contextos de atención primaria donde un falso negativo representa un riesgo clínico mayor a un falso positivo, puede llevar a interpretaciones erróneas. El uso de bases de datos con mejores distribuciones puede ayudar a mejorar el desempeño del clasificador.

E. Desarrollo e Integración del Prototipo

El desarrollo del prototipo incluyó la construcción del circuito principal y el diseño de una carcasa fabricada en PLA mediante impresión 3D. El circuito integró los componentes

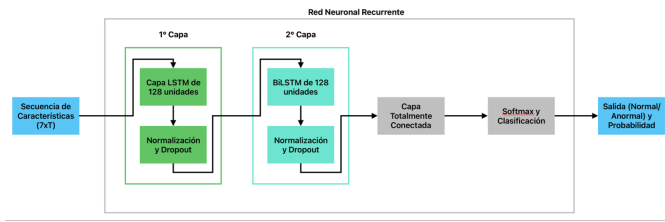


Fig. 10. Diagrama de Bloques de Red Neuronal Recurrente para Clasificación.

seleccionados con base en el análisis ELECTRE, incluyendo el micrófono Electret, un filtro pasa banda analógico de 20-600 Hz y el microcontrolador ESP32-S3. La carcasa fue diseñada para asegurar estabilidad y comodidad durante el uso, incorporando soportes específicos para el micrófono y la placa. A los inicios de la investigación, se planificaba que el micrófono entrara en contacto directo con la piel del paciente, esto con el objetivo de reducir el tamaño del equipo y costos. Sin embargo, durante las pruebas iniciales se observó que la captación directa del micrófono era muy sensible al ruido ambiental, por lo que se incorporó el diafragma de un estetoscopio como interfaz acústica, mejorando la transmisión del sonido y la ergonomía. En la Fig. 11 se puede ver cómo en la señal donde no existe acople acústico, esta se muestra con bastantes picos e irregularidades, los cuales se veían incrementados sobre todo por la presencia de corrientes de aire. En el caso de la señal con acople acústico, mostrado en la Fig. 12, esta muestra picos aislados correspondientes con cada sonido cardíaco captado, lo que se puede entender como una señal más limpia. Esto indica que el acople acústico es un elemento importante a tomar en cuenta al momento de diseñar este tipo de dispositivos, tanto como la calidad de los propios micrófonos.

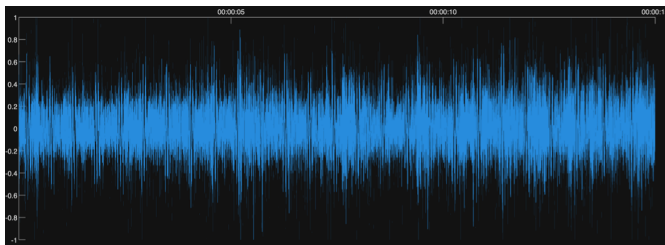


Fig. 11. Señal sin Acople Acústico.

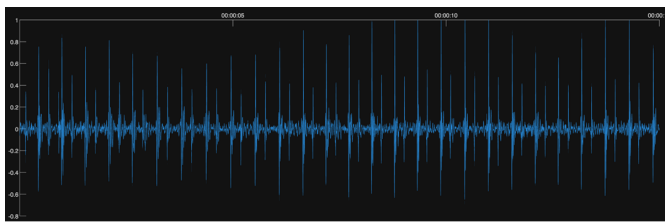


Fig. 12. Señal con Acople Acústico.

El costo total de fabricación del prototipo fue de aproxi-

madamente \$84.88, lo que lo hace entre tres y cinco veces más económico que estetoscopios electrónicos comerciales [23].

El sistema completo opera a través de cinco etapas: captación, filtrado analógico, digitalización, procesamiento digital y clasificación a través de la RNN. El sistema ofrece diversas salidas visuales, incluyendo el espectrograma y la visualización temporal de la señal con el resultado de la clasificación automática, permitiendo una interpretación rápida. Esto se puede observar en la Fig. 13.

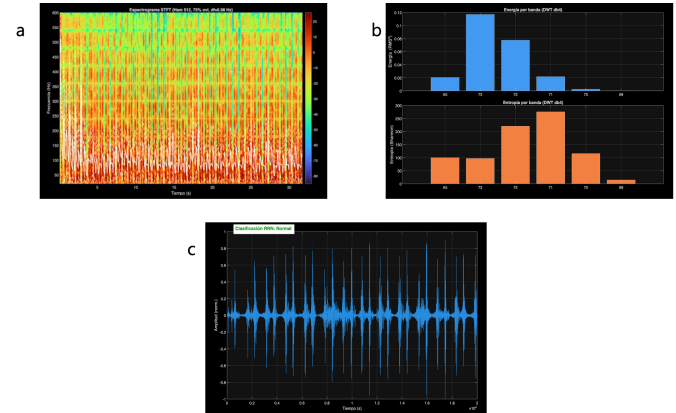


Fig. 13. Salidas del Clasificador. Espectrograma (a). Análisis de Transformada Wavelet Discreta con resultados de energía y entropía (b), grafica de la señal de sonidos cardíacos junto con clasificación diagnóstica de la RNN (c).

F. Estudio de Precisión Diagnostica

El estudio de precisión diagnóstica comparó el desempeño de tres médicos y el dispositivo de captación de sonidos cardíacos junto al algoritmo de clasificación bajo tres condiciones de ruido ambiental. Los niveles de ruido registrados fueron de 43.5 dBA para el ambiente silencioso, 60.7 dBA para el ambiente con ruido por voces y de 79 dBA para el ambiente con ruido por voces y monitores de signos vitales. Los resultados evidenciaron variaciones importantes entre los médicos: el médico 2 mostró el mejor rendimiento con una exactitud global de 0.93, mientras que el médico 1 tuvo resultados irregulares y el médico 3 el peor desempeño, con una exactitud de 0.67. El algoritmo, por su parte, alcanzó un desempeño moderado, destacando su sensibilidad en ambientes silenciosos de 0.8, pero evidenciando vulnerabilidad ante el ruido de voces. El desempeño de los médicos y el algoritmo puede verse en la Fig. 14.

En la comparación entre médicos y algoritmo, se observó que el médico 2 superó significativamente al sistema automático en todas las condiciones. En cambio, frente al médico 3, los resultados fueron similares pero con diferencias en el tipo de errores: el algoritmo mostró un equilibrio entre falsos positivos y negativos, mientras que el médico 3 tendió a evitar falsos negativos a costa de la especificidad. El análisis conjunto a través de la Prueba de McNemar con el consenso médico (Tabla I) reveló que en condiciones sin ruido no hubo diferencias significativas, pero sí bajo escenarios ruidosos, lo

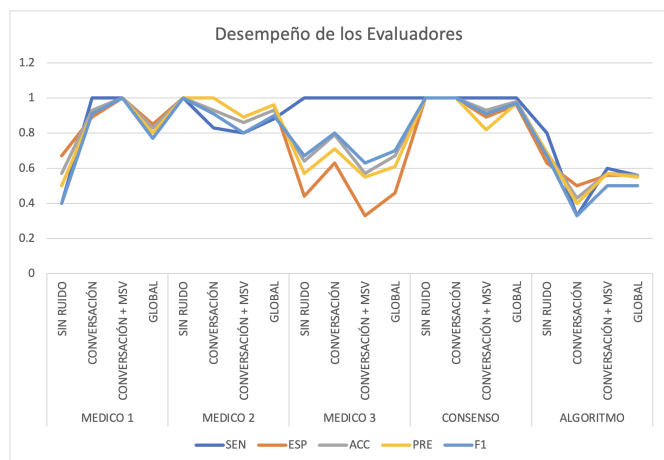


Fig. 14. Evolución Desempeño de los Evaluadores (Médicos y Algoritmos).

que demuestra que el algoritmo aún no iguala el desempeño clínico humano en estos entornos.

TABLA I
PRUEBA DE MCNEMAR ENTRE ALGORITMO Y CONSENSO MÉDICO.

Situación	Valor P	Significancia
Silencio (43.5 dBA)	0.12	No Significativa
Conversación (60.7 dBA)	0.01	Significativa
Conversación + MSV (79 dBA)	0.02	Significativa
Global	0.001	Significativa

Las pruebas de calidad de grabación mostraron una disminución progresiva de la SNR conforme aumenta el ruido ambiental, como se ve en la Fig. 15, pasando de 6.78 dB en silencio a 2.18 dB con ruido combinado. Esta degradación afectó directamente el desempeño del algoritmo, siendo la sensibilidad el parámetro más impactado. No obstante, se observó un comportamiento curioso: en condiciones del ruido por monitores de signos vitales, el rendimiento se recuperó parcialmente en comparación con el ruido por voces. Esto posiblemente porque el ruido por monitores, aunque intenso, presenta un patrón más estructurado y filtrable que las voces humanas.

La caída de precisión ante el ruido de voces representa la limitación más relevante del sistema actual. Para abordarla en versiones futuras se proponen las siguientes alternativas: (1) implementación de algoritmos de cancelación adaptativa de ruido de voz, que estimen y sustraigan la componente de voz en tiempo real sin interrumpir la continuidad temporal de la señal; (2) uso de configuraciones de doble micrófono con técnicas de *beamforming* para la separación espacial de las fuentes de audio, aprovechando la coherencia intercanal entre micrófonos; (3) incorporación de modelos de supresión de voz basados en redes neuronales previos a la etapa de clasificación; (4) reentrenamiento del modelo con datos aumentados que incluyan grabaciones con ruido de voces, mejorando la robustez específica del clasificador ante este tipo de interferencia.

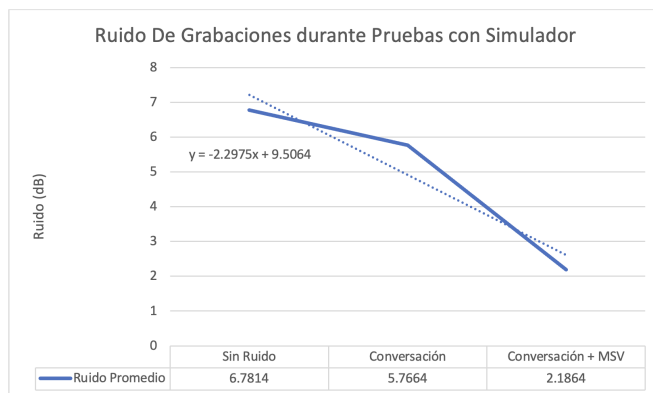


Fig. 15. Evolución del Ruido en Grabaciones durante Pruebas con Simulador.

G. Validación de Usabilidad

La aplicación del cuestionario de usabilidad a cuatro médicos permitió evaluar de forma integral la percepción sobre el prototipo, obteniendo un puntaje promedio de 75.9 puntos, clasificado como una buena usabilidad. Los valores individuales, mostrados en la Fig. 16, varían entre 65 y 85 puntos, mostrando ligeras diferencias entre los participantes, pero una tendencia general positiva hacia su facilidad de uso y aplicabilidad clínica. En general, los médicos consideraron que el dispositivo es útil y pertinente para el entorno clínico, especialmente en servicios de atención primaria, donde podría servir como apoyo diagnóstico para personal con menor experiencia en auscultación. Además, la mayoría coincidió en que el proceso de auscultación y registro de sonidos resultó claro y sencillo.

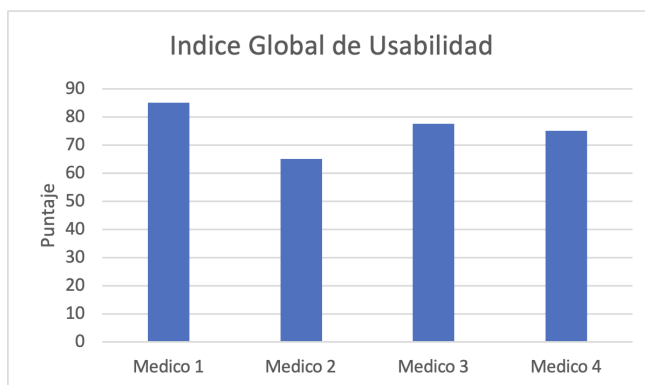


Fig. 16. Índice Global de Usabilidad para Médicos Participantes en Validación.

No obstante, se identificaron áreas que requieren mejoras. Algunos participantes indicaron que el proceso de conexión y análisis de los sonidos podría simplificarse para reducir la curva de aprendizaje, mientras que otros señalaron la necesidad de capacitación técnica básica para un uso óptimo. Esto refleja que la familiaridad del usuario con la tecnología influye directamente en su percepción de facilidad de uso. También se observaron ligeras inconsistencias en el funcionamiento del

dispositivo y en los resultados mostrados, lo que sugiere la necesidad de ajustes que garanticen un desempeño estable y confiable bajo diferentes condiciones clínicas.

Las observaciones adicionales de los médicos incluyeron sugerencias para integrar mediciones de frecuencia cardíaca en tiempo real, mejorar la calidad del audio y optimizar la captación de sonidos en mujeres, donde el tejido mamario o el roce con la ropa puede interferir. Estas últimas destacan como posibles áreas de enfoque en futuras investigaciones y aplicaciones de este tipo de dispositivos. Asimismo, recomendaron el desarrollo de un algoritmo que permita diferenciar tipos de soplos cardíacos, lo cual aumentaría la utilidad diagnóstica del dispositivo, especialmente en pediatría y obstetricia. En conjunto, los resultados reflejan que el prototipo tiene un alto potencial como herramienta accesible y de gran valor clínico, con posibilidad de perfeccionamiento para mejorar su precisión y adaptabilidad.

V. CONCLUSIONES

Se determinaron como características técnicas esenciales para un dispositivo de auscultación electrónica el uso de un micrófono electret, filtros pasa banda de 20-600 Hz, un microcontrolador ESP32-S3 con capacidad de procesamiento y transmisión digital y el uso de acoplamiento acústico con la piel del paciente por medio de un diafragma. Estas configuraciones permitieron obtener un registro robusto a bajo costo y viable en entornos de recursos limitados. Sin embargo, el desempeño del prototipo mostró vulnerabilidades ante ruidos ambientales intensos o generados por el contacto, factores que afectaron la estabilidad de la señal registrada.

Se determinaron como técnicas de procesamiento más efectivas para la identificación y clasificación de anomalías cardíacas la aplicación de redes neuronales. Esta herramienta permitió discriminar entre señales anormales y normales, incluso en condiciones de ruido controlado. No obstante, la precisión disminuyó en registros atípicos o de baja calidad, lo que evidenció la dependencia del sistema respecto a bases de datos amplias y balanceadas para su entrenamiento y las etapas de preparación de la señal dentro del procesamiento.

Se realizó un análisis comparativo entre la clasificación automática y la auscultación tradicional, encontrándose que el algoritmo en MATLAB alcanzó un rendimiento cercano al método clínico convencional en la identificación de soplos cardíacos. El sistema mostró resultados prometedores que refuerzan su posible utilidad para el apoyo diagnóstico. Sin embargo, en registros de baja calidad, el algoritmo mostró inconsistencias, lo que plantea posibles limitaciones bajo ciertos ambientes.

Se diseñó un prototipo de estetoscopio electrónico que integró un micrófono electret, etapas de filtración analógica, un microcontrolador y un módulo de análisis en MATLAB. El dispositivo permitió registrar sonidos cardíacos con fidelidad aceptable y clasificarlos con una precisión superior al 70% en condiciones controladas. No obstante, se identificaron limitaciones relacionadas con la persistencia del ruido y la disminución de la calidad de las señales en ambientes con

mayor contaminación acústica, lo que muestra la necesidad de mejora en filtrado y validación clínica en escenarios reales.

Se analizó la viabilidad de la adopción efectiva del dispositivo electrónico de auscultación cardíaca con análisis automático por parte del personal médico. El análisis permitió evidenciar que existe una disposición favorable hacia la incorporación de esta tecnología, especialmente porque representa una herramienta de apoyo al diagnóstico en escenarios de atención primaria y en contextos con limitaciones de recursos. Asimismo, se identificó que factores como la facilidad de uso, la portabilidad del equipo y la posibilidad de registrar digitalmente los sonidos cardíacos son percibidos como beneficios relevantes que facilitan la aceptación. No obstante, se destacó que la implementación exitosa del dispositivo dependerá en gran medida de la capacitación brindada al personal médico, así como de la disponibilidad de lineamientos claros para su uso en la práctica clínica cotidiana.

AGRADECIMIENTOS

Agradezco al Ing. Yaro Cáceres, asesor de la investigación, por su valiosa orientación y dedicación durante el desarrollo de la investigación. De igual manera, agradezco la colaboración de la Facultad de Ciencias de la Salud de UNITEC, cuyo equipo de doctores contribuyó con su experiencia y recursos en la validación del prototipo.

REFERENCES

- [1] D. Ah, A. S, H. Pa, S. H, R. E, and M. H, "Diagnostic accuracy of heart auscultation for detecting valve disease: A systematic review," *BMJ open*, vol. 13, no. 3, Mar. 2023, Publisher: BMJ Open, ISSN: 2044-6055. DOI: 10.1136/bmjopen-2022-068121. Accessed: Jul. 26, 2025.
- [2] N. Ma et al., "The Evolving Stethoscope: Insights Derived from Studying Phonocardiography in Trainees," *Sensors (Basel, Switzerland)*, vol. 24, no. 16, Aug. 2024, Publisher: Sensors (Basel), ISSN: 1424-8220. DOI: 10.3390/s24165333. Accessed: Jul. 26, 2025.
- [3] M. Fan et al., "Real-world evaluation of the Stethoscope electronic tele-auscultation system," *Biomedical Engineering Online*, vol. 21, no. 1, p. 63, Sep. 2022, ISSN: 1475-925X. DOI: 10.1186/s12938-022-01032-4.
- [4] M. Zhang, M. Li, L. Guo, and J. Liu, "A Low-Cost AI-Empowered Stethoscope and a Lightweight Model for Detecting Cardiac and Respiratory Diseases from Lung and Heart Auscultation Sounds," *Sensors*, vol. 23, no. 5, p. 2591, Feb. 2023, Publisher: MDPI AG, ISSN: 1424-8220. DOI: 10.3390/s23052591. Accessed: Jul. 26, 2025.
- [5] M. Suneja, J. F. Szot, R. F. LeBlond, and D. D. Brown, "Exploración física," in *DeGowin. Examen Diagnóstico, 11e*, New York, NY: McGraw-Hill Education, 2021. Accessed: Aug. 1, 2025. [Online]. Available: accessmedicina.mhmedical.com/content.aspx?aid=1182640686.

- [6] P. T. O’Gara and J. Loscalzo, “Exploración física del aparato cardiovascular,” in *Harrison. Principios de Medicina Interna, 21e*, J. Loscalzo, A. Fauci, D. Kasper, S. Hauser, D. Longo, and J. L. Jameson, Eds., New York, NY: McGraw-Hill Education, 2022. Accessed: Aug. 1, 2025. [Online]. Available: accessmedicina.mhmedical.com/content.aspx?aid=1198286143.
- [7] K. S. Saladin, “SISTEMA CIRCULATORIO: CORAZÓN,” in *Anatomía y Fisiología. La Unidad Entre Forma y Función, 9e*, New York, NY: McGraw Hill, 2022. Accessed: Aug. 1, 2025. [Online]. Available: accessmedicina.mhmedical.com/content.aspx?aid=1187644513.
- [8] M. Choudry, T. S. Stead, R. K. Mangal, and L. Ganti, “The History and Evolution of the Stethoscope,” *Curēus*, vol. 14, no. 8, e28171, Aug. 2022, ISSN: 2168-8184. DOI: 10.7759/cureus.28171. Accessed: Aug. 13, 2025.
- [9] J. J. Seah, J. Zhao, D. Y. Wang, and H. P. Lee, “Review on the Advancements of Stethoscope Types in Chest Auscultation,” *Diagnostics (Basel, Switzerland)*, vol. 13, no. 9, p. 1545, Apr. 2023, ISSN: 2075-4418. DOI: 10.3390/diagnostics13091545.
- [10] S. Leng, R. S. Tan, K. T. C. Chai, C. Wang, D. Ghista, and L. Zhong, “The electronic stethoscope,” *Biomedical Engineering Online*, vol. 14, p. 66, Jul. 2015, ISSN: 1475-925X. DOI: 10.1186/s12938-015-0056-y.
- [11] E. R. Gottlieb, J. M. Aliotta, and D. Tammaro, “Comparison of analogue and electronic stethoscopes for pulmonary auscultation by internal medicine residents,” *Postgraduate Medical Journal*, vol. 94, no. 1118, pp. 700–703, Dec. 2018, ISSN: 1469-0756. DOI: 10.1136/postgradmedj-2018-136052.
- [12] B. Silverman and M. Balk, “Digital Stethoscope-Improved Auscultation at the Bedside,” *The American Journal of Cardiology*, vol. 123, no. 6, pp. 984–985, Mar. 2019, ISSN: 1879-1913. DOI: 10.1016/j.amjcard.2018.12.022.
- [13] E. Vera de Payer, *Teoria de Senales*. Jorge Sarmiento Editor - Universitas, 2020, ISBN: 978-987-572-040-4.
- [14] V. H. Sauchelli, *Teoria de senales y sistemas lineales*. Jorge Sarmiento Editor - Universitas, 2020, ISBN: 978-987-9406-73-1. [Online]. Available: <https://elibro.net/es/lc/unitechn/titulos/175159>.
- [15] A. Fava et al., “Pre-processing techniques to enhance the classification of lung sounds based on deep learning,” *Biomedical Signal Processing and Control*, vol. 92, p. 106009, Jun. 2024, ISSN: 1746-8094. DOI: 10.1016/j.bspc.2024.106009. Accessed: Aug. 6, 2025.
- [16] J. R. Figueira, S. Greco, B. Roy, and R. Słowiński, “An Overview of ELECTRE Methods and their Recent Extensions,” *Journal of Multi-Criteria Decision Analysis*, vol. 20, no. 1-2, pp. 61–85, 2013, eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/mcda.1482>, ISSN: 1099-1360. DOI: 10.1002/mcda.1482. Accessed: Aug. 16, 2025.
- [17] A. Najafi, A. Nemati, M. Ashrafzadeh, and S. Hashemkhani Zolfani, “Multiple-criteria decision making, feature selection, and deep learning: A golden triangle for heart disease identification,” *Engineering Applications of Artificial Intelligence*, vol. 125, p. 106662, Oct. 2023, ISSN: 0952-1976. DOI: 10.1016/j.engappai.2023.106662. Accessed: Aug. 16, 2025.
- [18] Y. Y. Ang, L. R. Aw, V. Koh, and R. X. Tan, “Characterization and cross-comparison of digital stethoscopes for telehealth remote patient auscultation,” *Medicine in Novel Technology and Devices*, vol. 19, p. 100256, Sep. 2023, Publisher: Elsevier BV, ISSN: 2590-0935. DOI: 10.1016/j.medntd.2023.100256. Accessed: Jul. 26, 2025.
- [19] Z. Song, P.-J. Lee, and T. Hampton, “Acoustic environments of intensive care units during the COVID-19 pandemic,” *Applied Acoustics*, vol. 199, p. 109037, Oct. 2022, ISSN: 0003-682X. DOI: 10.1016/j.apacoust.2022.109037. Accessed: Aug. 25, 2025.
- [20] J. Shreffler and M. R. Huecker, “Diagnostic Testing Accuracy: Sensitivity, Specificity, Predictive Values and Likelihood Ratios,” in *StatPearls*, Treasure Island (FL): StatPearls Publishing, 2025. Accessed: Aug. 16, 2025. [Online]. Available: <http://www.ncbi.nlm.nih.gov/books/NBK557491/>.
- [21] V. Panayotov, G. Chen, D. Povey, and S. Khudanpur, “Librispeech: An ASR corpus based on public domain audio books,” in *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, ISSN: 2379-190X, Apr. 2015, pp. 5206–5210. DOI: 10.1109/ICASSP.2015.7178964. Accessed: Sep. 11, 2025.
- [22] MathWorks. “Mathworks/physionet_phonocardiogram.” original-date: 2018-06-12T19:20:36Z, Accessed: Sep. 11, 2025. [Online]. Available: https://github.com/mathworks/physionet_phonocardiogram.
- [23] Medisave. “Littmann CORE Digital Stethoscopes,” Accessed: Sep. 16, 2025. [Online]. Available: <https://www.medisave.net/collections/littmann-core-digital-stethoscopes>.