



UNIVERSIDAD POLITÉCNICA SALESIANA

SEDE GUAYAQUIL

CARRERA DE BIOMEDICINA

**DISEÑO E IMPLEMENTACIÓN DE UN SISTEMA BIOMÉDICO
PARA LA CLASIFICACIÓN DE ENFERMEDADES
CARDIOPULMONARES MEDIANTE EL ANÁLISIS DE SEÑALES
ACÚSTICAS USANDO MACHINE LEARNING**

Trabajo de titulación previo a la obtención del

Título de Ingeniero Biomédico

AUTORES: Augusto Yáñez Pazmiño & Julio Pino Cruz

TUTOR: Msc. Roberto Bayas Toro

Guayaquil - Ecuador

2024

CERTIFICADO DE RESPONSABILIDAD Y AUTORÍA DEL TRABAJO DE TITULACIÓN

Nosotros, **Augusto Alexander Yánez Pazmiño** con documento de identificación N.º **1251194245** y **Julio Francisco Pino Cruz** con documento de identificación N.º **0956182257**; manifestamos que:

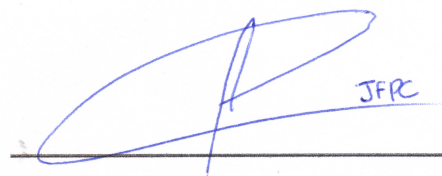
Somos los autores y responsables del presente trabajo; y, autorizamos a que sin fines de lucro la Universidad Politécnica Salesiana pueda usar, difundir, reproducir o publicar de manera total o parcial el presente trabajo de titulación.

Guayaquil, 29 de enero del año 2024.



Augusto Alexander Yánez Pazmiño

1251194245



Julio Francisco Pino Cruz

0956182257

CERTIFICADO DE CESIÓN DE DERECHOS DE AUTOR DEL TRABAJO DE TITULACIÓN A LA UNIVERSIDAD POLITÉCNICA SALESIANA

Nosotros, **Augusto Alexander Yánez Pazmiño** con documento de identificación N.º **1251194245** y **Julio Francisco Pino Cruz** con documento de identificación N.º **0956182257**, expresamos nuestra voluntad y por medio del presente documento cedemos a la Universidad Politécnica Salesiana, la titularidad sobre los derechos patrimoniales, en virtud de que somos autores del trabajo: **DISEÑO E IMPLEMENTACIÓN DE UN SISTEMA BIOMÉDICO PARA LA CLASIFICACIÓN DE ENFERMEDADES CARDIOPULMONARES MEDIANTE EL ANÁLISIS DE SEÑALES ACÚSTICAS USANDO MACHINE LEARNING**, el cual ha sido desarrollado para obtener el título de **Ingeniero Biomédico** de la Universidad Politécnica Salesiana, quedando la Universidad facultada para ejercer plenamente los derechos cedidos anteriormente.

En concordancia con lo manifestado, suscribimos este documento en el momento que hacemos la entrega del trabajo final en formato digital a la Biblioteca de la Universidad Politécnica Salesiana.

Guayaquil, 29 de enero del año 2024.



Augusto Alexander Yánez Pazmiño

1251194245



Julio Francisco Pino Cruz

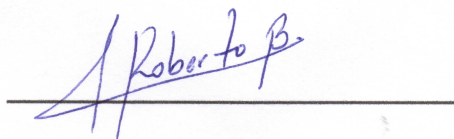
0956182257

CERTIFICADO DE DIRECCIÓN DEL TRABAJO DE TITULACIÓN

Yo, **Roberto Gerardo Bayas Toro**, con documento de identificación N.º **0940622608**, docente de la Universidad Politécnica Salesiana, declaro que bajo mi tutoría fue desarrollado el trabajo de titulación: **DISEÑO E IMPLEMENTACIÓN DE UN SISTEMA BIOMÉDICO PARA LA CLASIFICACIÓN DE ENFERMEDADES CARDIOPULMONARES MEDIANTE EL ANÁLISIS DE SEÑALES ACÚSTICAS USANDO MACHINE LEARNING**, realizado por **Augusto Alexander Yáñez Pazmiño** con documento de identificación N.º **1251194245** y **Julio Francisco Pino Cruz** con documento de identificación N.º **0956182257**, obteniendo como resultado final de trabajo de titulación la opción Proyecto Técnico, que cumple con todos los requisitos determinados por la Universidad Politécnica Salesiana.

Guayaquil, 29 de enero del año 2024.

Atentamente,



Msc. Roberto Gerardo Bayas Toro

0940622608

Resumen

El presente trabajo de investigación se centra en el diseño e implementación de un sistema biomédico para la clasificación de enfermedades cardio pulmonares mediante el análisis de señales acústicas utilizando técnicas avanzadas de machine learning. Este estudio surge de la necesidad de mejorar la precisión en la auscultación cardíaca y pulmonar, especialmente para profesionales de la salud en formación y en etapas tempranas de su carrera. El sistema propuesto busca superar las limitaciones de los métodos tradicionales de diagnóstico, proporcionando una herramienta más precisa y confiable para la detección temprana de afecciones cardio pulmonares.

El desarrollo del sistema incluye el procesamiento y análisis de sonidos cardíacos y pulmonares, la extracción de características relevantes y la implementación de algoritmos de machine learning para la clasificación eficaz de las enfermedades. Se emplearon técnicas de procesamiento de señales para mejorar la calidad de los datos acústicos y se seleccionaron modelos de aprendizaje automático adecuados para la tarea de clasificación. El sistema fue probado y validado con un conjunto de datos etiquetados, demostrando su capacidad para identificar con precisión diversas afecciones cardio pulmonares.

Los resultados obtenidos indican que el sistema biomédico desarrollado es una herramienta valiosa para mejorar la precisión diagnóstica en la práctica médica. Además, este estudio contribuye significativamente al campo de la ingeniería biomédica y la medicina, abriendo nuevas vías para la investigación y el desarrollo de tecnologías de diagnóstico avanzadas. La implementación de este sistema tiene el potencial de mejorar la calidad de la atención médica y de apoyar la formación de profesionales de la salud en la interpretación precisa de sonidos cardíacos y pulmonares.

Abstract

This research work focuses on the design and implementation of a biomedical system for the classification of cardiopulmonary diseases through the analysis of acoustic signals using advanced machine learning techniques. This study arises from the need to improve the precision in cardiac and pulmonary auscultation, especially for health professionals in training and at the early stages of their careers. The proposed system seeks to overcome the limitations of traditional diagnostic methods, providing an accurate and reliable tool for the early detection of cardiopulmonary conditions.

The development of the system includes the collection and analysis of cardiac and pulmonary acoustic signals, the extraction of relevant features, and the implementation of machine learning algorithms for effective diseases classification. Signal processing techniques were employed to improve the quality of the acoustic data and suitable machine learning models were selected for the classification task. The system was tested and validated against a labeled data set, demonstrating its ability to accurately identify various cardiopulmonary conditions.

The results demonstrated that the developed biomedical system is a valuable tool to improve diagnostic accuracy in medical practice. Furthermore, this study contributes significantly to the field of biomedical engineering and medicine, opening new opportunities for research and development of advanced diagnostic technologies. The implementation of this system has the potential to improve the quality of medical care and support the training of health professionals in the accurate interpretation of heart and lung sounds.

Agradecimientos

Dios, por ser mi guía constante e iluminarme en cada desafío y decisión. A mi hermana, la Dra. Paola Yáñez, quien con su entusiasmo y experiencia me ha guiado en la preparación de este proyecto de titulación. Su paciencia, aliento y continuo apoyo han sido inestimables para mí durante este proceso de investigación, haciendo de este viaje académico, una oportunidad de crecimiento personal y profesional. Al personal de la empresa de Servicio y Equipamiento SERQUIP ENGINEERING, por darme la oportunidad de colaborar con ustedes, por confiar en mis ganas de aprender y por permitirme crecer profesionalmente. Gracias a ustedes por compartir sus ideas y recomendaciones en mi trabajo de investigación. Un agradecimiento especial a mi tutor de tesis, el Msc. Roberto Bayas Toro, por sus continuos comentarios e ideas sobre mi investigación. También agradezco al MSc. Danny Barona Valencia y al Ing. Flavio Moreno por sus valiosos comentarios y discusiones fructíferas que me permitieron mejorar el enfoque del proyecto. En particular, me gustaría agradecer al director de la carrera, el Msc. Carlos Aguilera Sarria, por las largas charlas y los valiosos comentarios sobre mi investigación. Estoy en deuda con mis padres, Merid y Augusto, por su apoyo incondicional, amor y guía; mis hermanas, Paola, María José, María Sol, mi sobrina Sophia y a mi perrita una pequeña Snowzer llamada Lula, a todos por su paciencia e inspiración.

Augusto Alexander Yanez Pazmiño

Un agradecimiento muy grande hacia Dios, quien ha iluminado el sendero que he recorrido a lo largo de mi vida, dándome fortaleza en los momentos más difíciles. A mí, familia, quienes me han sabido comprender y ayudar en todo lo que he necesitado, y han sido mi estímulo constante para seguir adelante. A todas las personas que han formado parte de este logro, a mi compañero de tesis Augusto, quien ha aportado también en mi vida con amistad y lealtad, a mis compañeros de curso que han sabido llegar a mí, cada uno de forma de diferente y única, a mis profesores que han dedicado sus ganas y tiempo en impartirnos no solo enseñanzas académicas sino también enseñanzas de vida, al Ingeniero Roberto Bayas quien más que profesor puedo llamar amigo que ha sido partícipe tanto de mi desarrollo humano como profesional.

Julio Francisco Pino Cruz

Dedicatorias

A mi madre, quien con su amor y sabiduría, me ha formado como un hombre de bien, con valores y principios que me han permitido salir adelante en los momentos difíciles. A mi padre, a quien le agradezco por cuidarme y guiarme con sus sabios consejos. A mi hermana, que siempre ha estado junto a mí, brindándome su apoyo incondicional e incluso muchas veces poniéndose en el papel de madre. A mi familia en general por brindarme su apoyo incondicional y compartir conmigo los altibajos de la vida. Cada uno de ustedes ha sido luz en este camino de aprendizaje y crecimiento personal; este logro también es de ustedes.

Augusto Alexander Yanez Pazmiño

A mis padres, que son el pilar fundamental de mi vida, que han fomentado valores buenos, hábitos en mi vida y que me han ayudado a forjarme como persona. A mis hermanos, quienes me han acompañado en mi vida, dándome apoyo y fuerzas cuando más lo necesitaba. A mí mismo, que me he demostrado que gracias a mi esfuerzo y dedicación he podido completar otra meta en mi vida y empezar un nuevo capítulo en la misma.

Julio Francisco Pino Cruz

“Los grandes trabajos no son hechos por la fuerza, sino por la perseverancia.”

Benjamin Johnson

Índice general

1. Introducción	1
1.1. Problema	2
1.2. Objetivo General	3
1.3. Objetivos Específicos	3
2. Fundamentos Teóricos	5
2.1. Auscultación	5
2.2. Estetoscopio	5
2.3. Ingeniería Biomédica	7
2.4. Diagnóstico	8
2.5. Machine Learning	8
2.6. Espectrograma	9
2.7. Cromagrama	10
2.8. Transformada de Fourier de Tiempo Corto	10
2.9. Callbacks	11
2.10. Early Stopping	11
2.11. ReduceLROnPlateau	12
2.12. Capas LSTM	12
2.13. Mel Frequency Cepstral Coefficients (MFCCs)	12
2.14. Redes Neuronales Convolucionales	15
2.15. Aumentación de Datos	16
3. Marco Metodológico	17

3.1. Descripción General	17
3.2. Adquisición de Datos	17
3.2.1. Información sobre el Primer Conjunto de Datos	18
3.2.2. Información sobre el Segundo Conjunto de Datos	20
3.3. Preprocesamiento de Señales	21
3.4. Extracción de Características	24
3.4.1. Transformada de Fourier de Corto Tiempo (STFT)	24
3.4.2. Espectrogramas de Mel	24
3.4.3. Extracción de Características Espectrales	24
3.4.4. Coeficientes Cepstrales en las Frecuencias de Mel (MFCC)	24
3.4.5. Cromagrama	24
3.4.6. Contraste Espectral	24
3.4.7. Características Temporales	25
3.4.8. Normalización de Características	25
3.4.9. Selección de Características	25
3.5. Entrenamiento del Modelo de Clasificación	25
3.5.1. Validación cruzada	25
3.5.2. Configuración de Hiper Parámetros	25
3.5.3. Callbacks	25
3.5.4. Monitoreo del Rendimiento	26
4. Resultados	27
4.0.1. Descripción General	27
4.0.2. Resultados del Primer Código	27
4.0.3. Resultados del Segundo Código	30
4.0.4. Discusión	34
5. Cronograma	35
6. Presupuesto	36

<i>ÍNDICE GENERAL</i>	XII
7. Conclusiones	37
8. Recomendaciones	39
Bibliografía	41
A. Apéndice	44

Índice de figuras

2.1. Estetoscopio de Laennec (Musso & Enz, 2015).	6
2.2. Muestra de un espectrograma (Paz et al., 2020).	9
2.3. Muestra de Croma grama (Benito Gorrón et al., 2017)	10
2.4. Representación de Transformada de Fourier de tiempo corto (Moler, 2022b) . .	11
2.5. Diagrama de secuencia de una capa LSTM.	12
2.6. Potencia espectral de una señal (Franco-Galván, 2017).	14
2.7. Diagrama de flujo de Los Coeficientes Cepstrales en las Frecuencias de Mel. . .	14
4.1. Matriz de confusión.	28
4.2. Curvas de Pérdida etapas iniciales.	29
4.3. Curvas de Pérdida etapas finales.	29
4.4. Ejemplo 1 de los espectrogramas analizados	30
4.5. Ejemplo 2 de los espectrogramas analizados.	30
4.6. Ejemplo 3 de los espectrogramas analizados.	31
4.7. Curvas de pérdida y precisión durante el entrenamiento y la validación.	31
4.8. Histograma de las duraciones de los audios.	32
4.9. Ejemplo de espectrograma 1.	33
4.10. Ejemplo de espectrograma 2.	33
4.11. Ejemplo de espectrograma 3.	33

Índice de tablas

3.1. Descripción del dataset <i>set_a.csv</i>	19
3.2. Descripción del dataset <i>set_b.csv</i>	19
3.3. Descripción de los Datos	21
4.1. Rendimiento del modelo por clase.	28
5.1. Cronograma de Actividades	35
6.1. Presupuesto	36

Abreviaturas

STFT	Transformada de Fourier de Tiempo Reducido
MFCC	Mel Frequency Cepstral Coefficients
CNN	Convolutional neural network
LSTM	Long Short-Term Memory
FFT	Transformada Rápida de Fourier
DFT	Transformada Discreta de Fourier
SSD	Single Shot Multibox Detector
YOLO	You Only Look Once
RNN	Recurrent neural network
DCT	Discrete Cosine Transform

Introducción

La auscultación con estetoscopio es una técnica no invasiva, rápida y de bajo coste, que permite el diagnóstico inicial y seguimiento de enfermedades cardio pulmonares (Williams, 1907). Este examen físico permite a los médicos identificar anormalidades en los sonidos cardíacos y pulmonares y detectar enfermedades como soplos, ritmos cardíacos irregulares o ruidos adicionales (A. D. Rodríguez et al., 2023). Sin embargo, esta técnica representa retos importantes debido a la subjetividad y variabilidad en la interpretación de los sonidos, que está sujeto en gran medida a la experiencia y habilidad de los profesionales de la salud.

En este contexto, algunos estudios sostienen que el entrenamiento brindado a los estudiantes y médicos en formación, no les permiten alcanzar un nivel aceptable de diagnóstico (Haring et al., 2014). Adicionalmente, la interpretación de sonidos cardio pulmonares está sujeta a la subjetividad humana y a la experiencia del profesional, por lo que requiere de práctica continua (Lam et al., 2005). De ahí que, la competencia auscultatoria de los profesionales de la salud y las limitaciones inherentes de los estetoscopios tradicionales, podrían derivar en diagnósticos incorrectos que afecten la calidad de vida de los pacientes.

El objetivo del presente estudio es diseñar e implementar un sistema biomédico innovador para la clasificación de enfermedades cardio pulmonares mediante el análisis de señales acústicas, utilizando técnicas de machine learning. El sistema debe permitir el diagnóstico inicial y seguimiento de enfermedades, mejorando la precisión y confiabilidad de la evaluación médica. Además, el sistema propuesto podrá servir como una plataforma de entrenamiento para los profesionales en formación, que desean mejorar su habilidad de auscultación y diagnóstico de enfermedades cardio pulmonares (A. D. Rodríguez et al., 2023).

Este trabajo introduce una herramienta de diagnóstico avanzada, diseñada específicamente para la auscultación cardíaca y pulmonar e identificación de patologías en pacientes. Por un lado, machine learning es una rama de la Inteligencia Artificial (IA) que se centra en el uso de datos para el desarrollo y modelado de algoritmos que permitan aprender de forma automatizada (Zou, 2019). Por otro lado, la fusión de la información acústica recopilada por estetoscopios digitales con algoritmos de procesamiento de señales ofrece una perspectiva prometedora para la detección de patologías (Mayorga et al., 2011).

El enfoque innovador de este trabajo se enfoca en abordar las limitaciones específicas de los estetoscopios tradicionales, como la interferencia de ruidos ambientales y la variabilidad en la interpretación de sonidos, para ofrecer diagnósticos más precisos y fiables. La implementación del sistema propuesto puede ser de gran utilidad en entornos con recursos limitados, donde el acceso a especialistas y equipos avanzados puede ser restringido (Bank et al., 2016). Además, el desarrollo de esta herramienta representa un avance significativo en la integración de técnicas de machine learning en el campo biomédico y detección temprana y efectivo de enfermedades cardio pulmonares (Kim et al., 2022).

1.1. Problema

En el área de la medicina, la auscultación con estetoscopio es una técnica utilizada en la examinación física de los pacientes, que permite el diagnóstico inicial de enfermedades cardio pulmonares (Williams, 1907). A pesar de su popularidad en la práctica clínica, diferentes estudios han demostrado que médicos y estudiantes en formación enfrentan desafíos significativos en la interpretación de sonidos del corazón y los pulmones (Tavel, 2006).

Algunos estudios señalan que los estudiantes de medicina en formación, no adquieren las habilidades necesarias para distinguir entre sonidos cardio pulmonares normales y anormales (Haring et al., 2014; Roelandt, 2014). De forma similar, un estudio indica que médicos y proveedores de atención médica necesitan de práctica continua de auscultación con estetoscopio para lograr un diagnóstico preciso y eficiente (Lam et al., 2005). Otro estudio advierte que a pesar de que existe guías sobre el manejo y procedimientos para realizar auscultación pulmonar utilizando simuladores, solo la experiencia y la pericia permite a los médicos realizar la

interpretación correcta y precisa de los sonidos cardio pulmonares (Sarkar et al., 2015).

De ahí que, la experiencia auscultatoria subjetiva de médicos y estudiantes en formación y las limitaciones de los estetoscopios tradicionales, destacan la necesidad de desarrollar un sistema biomédico que permita la interpretación de sonidos cardio pulmonares. El sistema debe servir como herramienta inicial para realizar una evaluación preliminar de las enfermedades del corazón y los pulmones, venciendo así las barreras de la experiencia y las habilidades de diagnóstico. Además, el sistema debe superar la variabilidad y subjetividad en la interpretación de los sonidos, conduciendo a un diagnóstico eficiente y efectivo.

Con la finalidad de resolver esta problemática, este trabajo de investigación tiene como objetivo el diseño e implementación de un sistema biomédico que integra tecnologías de machine learning para analizar sonidos del corazón y pulmones y clasificar las enfermedades cardio pulmonares. El sistema propuesto no solo busca el diagnóstico preciso y efectivo de sonidos, que lleve a un tratamiento adecuado y de bajo costo, sino que también pueda ser usado como una herramienta de enseñanza de auscultación cardiaca y pulmonar. Este sistema aspira convertirse en una herramienta útil que mejore la competencia de auscultación y permita el diagnóstico correcto de patologías, que deriven a un tratamiento adecuado.

1.2. Objetivo General

Diseñar e implementar un sistema biomédico innovador para la clasificación de enfermedades cardio pulmonares mediante el análisis de señales acústicas, utilizando técnicas de Machine Learning.

1.3. Objetivos Específicos

- Adquirir datos para el entrenamiento del modelo de Machine Learning.
- Diseñar algoritmos receptores de señales acústicas cardiopulmonares mediante pruebas.

- Seleccionar las características relevantes en las señales acústicas que pueden ser indicativas de enfermedades cardio pulmonares, como frecuencias cardiacas anormales, patrones de sonidos específicos, etc.
- Configurar los algoritmos de Machine Learning adecuados para la clasificación de enfermedades cardio pulmonares.
- Entrenar y ajustar los modelos de Machine Learning, mostrando los resultados de su correcto funcionamiento.

Fundamentos Teóricos

Este capítulo presenta los conceptos claves y la terminología utilizada como base en el proyecto de investigación propuesto. Esta sección facilita la comprensión de las técnicas, teorías, y métodos aplicados en este estudio.

2.1. Auscultación

Se define como la exploración de los parámetros acústicos que producen los distintos órganos y forma parte del examen físico médico, esta se puede realizar simplemente con la auscultación directa (el personal médico escucha sin ninguna herramienta los sonidos emitidos) o mediante la ayuda de artilugios, como el estetoscopio, para su correcta recepción (Williams, 1907). Se utiliza especialmente en estudios cardiológicos del paciente para “escuchar” tonos, los ruidos y los soplos cardiacos; pero también se utiliza en la ciencia de la neumología para estudiar ruidos pulmonares; se discute mucho cuál es el artilugio que presenta una mejor ayuda para realizar la auscultación y según estudios los resultados indicaron que la campana del estetoscopio infantil fue la pieza que proporcionó mejores resultados en todas las arterias donde fue medida la presión arterial (Lamas & Arcuri, 2002).

2.2. Estetoscopio

La historia del estetoscopio tiene sus raíces en el siglo XIX y está estrechamente ligada al médico francés René Laennec (Canal, 2003). Antes de que este dispositivo fuera inventado, los médicos solían examinar los sonidos del cuerpo humano al colocar directamente su oído sobre

el pecho del paciente, una práctica conocida como "percusión directa". No obstante, esta técnica presentaba limitaciones y resultaba incómoda tanto para el médico como para el paciente.

En 1816, mientras se encontraba en el Hospital Necker de París, Laennec se vio enfrentado a la necesidad de auscultar el pecho de una mujer con sobrepeso. Dada la impracticabilidad de la percusión directa en este caso, el médico buscó una alternativa. Rememorando su infancia, cuando solía escuchar sonidos al golpear un extremo de un trozo de madera y colocar el otro extremo en su oído para amplificar los sonidos, Laennec se inspiró en esta experiencia. En consecuencia, ideó enrollar un trozo de papel en forma de cilindro y emplearlo como un tubo para llevar a cabo la auscultación.

Este dispositivo rudimentario, 2.1 que Laennec llamó *estetoscopio*, permitía al médico escuchar los sonidos internos del cuerpo sin tener que colocar su oído directamente en el pecho del paciente. Este invento revolucionó la forma en que se realizaba la auscultación médica. Con el tiempo, el diseño del estetoscopio evolucionó, y hoy en día se utilizan estetoscopios más avanzados con tubos y diafragmas que permiten a los médicos escuchar con mayor claridad los sonidos cardíacos y pulmonares. Esta herramienta se convirtió en un instrumento fundamental para la detección de enfermedades cardíacas y pulmonares, facilitando a los médicos la identificación de soplos cardíacos, ritmos irregulares, y la presencia de sonidos anómalos en los pulmones. El estetoscopio sigue siendo una pieza esencial del equipo médico utilizado en la evaluación clínica y el diagnóstico de diversas afecciones (Musso & Enz, 2015).



Figura 2.1: Estetoscopio de Laennec (Musso & Enz, 2015).

2.3. Ingeniería Biomédica

La Ingeniería Biomédica es una ciencia que aplica utiliza el desarrollo tecnológico para mejorar el ámbito médico (Simini & Vienni, 2016). Su eje principal es el diseño y la invención de equipos médicos, dispositivos como las prótesis, dispositivos de diagnóstico o para terapias. Además, representa un papel vital en la organización de los recursos técnicos vinculados a los sistemas de salud. Fusionando las facultades de las ingenierías con las carencias de los sistemas médicos, esta ciencia de la ingeniería busca optimizar la atención hospitalaria. La Ingeniería Biomédica integra los principios de la ingeniería con el ámbito de la medicina y la fisiología, promoviendo avances significativos en el conocimiento científico y el progreso tecnológico en los campos de la medicina y la biología. Su alcance abarca una amplia gama de actividades, que van desde la aplicación de métodos matemáticos y la realización de investigaciones experimentales hasta el diseño, desarrollo y aplicación clínica de tecnologías innovadoras. Alrededor de 1930, en laboratorios de Europa y Estados Unidos surgieron simultáneamente 3 ramas innovadoras como son la Ingeniería Biomédica, la Biofísica y la Física Médica. En estos entornos, físicos, ingenieros y médicos empleaban métodos analíticos de las ciencias físicas y su habilidad en el manejo de instrumentos para abordar una variedad de problemas planteados por las ciencias de la vida. La instrumentación electrónica a partir de tubos de vacío se empleó por E. Lovett Garceau para amplificar señales eléctricas y el primer sistema de electroencefalógrafo comercial de tres canales fue construido por Alfred Grass en 1935 (Gismondi Glave, 2010). En la década de 1950, estas actividades experimentaron un aumento considerable, y los laboratorios originales evolucionaron para convertirse en departamentos universitarios especializados en Ingeniería Biomédica. El primer programa académico oficial en este campo inició en 1959 como una Maestría en la Universidad Drexel en Estados Unidos. (E. Rodríguez, 2003).

Desde 1970, los estudios de Ingeniería biomédica han ocupado la atención de las más reconocidas universidades latinoamericanas. En el presente, esta ingeniería se expande hacia áreas como la telemedicina, las telecomunicaciones, la electrónica, la informática y otras disciplinas de la ingeniería para resolver desafíos en biología y medicina a distancia (Ariza Nadjar et al., 2013).

2.4. Diagnóstico

Se define como las fases que conlleva el personal médico para identificar una enfermedad, lesión o afección en un paciente dependiendo de sus síntomas o signos vitales. esto se logra utilizando el historial clínico del paciente donde se puede apreciar antecedentes de salud y se confirma la sospecha realizando exámenes tanto físicos como biológicos al afectado, como pueden ser análisis de sangre, pruebas con imágenes y/o biopsias (Lorenzano, 2006).

Hipócrates mismo lo presenta diciendo que el orgullo médico no es tanto el curar las enfermedades, algo que en ocasiones no permite el desarrollo del conocimiento, sino que se cumpla aquello que se pronostica y, como sabemos, no hay pronóstico correcto sin un buen diagnóstico (Cruz Hernández et al., 2012).

2.5. Machine Learning

Este campo, una subdivisión de la inteligencia artificial (IA), se dedica a simular el aprendizaje humano a través del uso de datos y algoritmos con el fin de incrementar su exactitud de manera progresiva. Su origen se remonta a la mitad del siglo XX, marcado por el desarrollo de modelos matemáticos de las redes neuronales del cerebro humano. En 1943, Walter Pitts y Warren McCulloch dieron a conocer un estudio que buscaba representar de manera matemática el razonamiento y la toma de decisiones humanas (Bobadilla, 2021).

Pocos años después, en 1950, Arthur Samuel desarrolló un software capaz de jugar al juego de damas. A pesar de las limitaciones en su capacidad de memoria, este programa podía escoger sus movimientos basándose en la estrategia minimax, la cual considera el peor caso posible en las jugadas del oponente. Esta estrategia fue más tarde conocida como el algoritmo minimax (Hinestroza Ramírez, 2018).

En el mismo año, Alan Turing introdujo una prueba para evaluar la inteligencia de una máquina, que consistía en su habilidad para persuadir a una persona de que era humana (González Fernández, 2007). Sin embargo, Samuel avanzó en sus investigaciones y mejoró sus primeros modelos mediante un enfoque que llamó "aprendizaje de memoria", permitiendo al programa

almacenar y recordar todas las posiciones observadas durante el juego. Esta acumulación de conocimientos previos resultó en un avance notable en la capacidad de jugar damas. En 1952, se introdujo el término "aprendizaje automático" para denominar este fenómeno.

2.6. Espectrograma

Es una herramienta valiosa en el análisis de señales, fue inventado por Joseph Fourier, un matemático francés del siglo XIX. Este método implica la toma de muestras a través de una ventana temporal de tamaño específico, según el tipo de análisis requerido. Posteriormente, se utiliza la Transformada Rápida de Fourier (FFT) para calcular el contenido espectral de las muestras, y los resultados se representan en un gráfico tridimensional. La ventana temporal se desplaza a lo largo del tiempo sobre la señal en análisis, adquiriendo nuevas muestras y repitiendo el proceso de cálculo y representación gráfica, como se ilustra en la figura 2.2 (Paz et al., 2020).

El espectrograma proporciona información crucial en el dominio de la frecuencia, revelando la variación de la energía y la frecuencia de la señal en función del tiempo. Esta técnica encuentra aplicaciones en diversas áreas, como la ingeniería de sonido, la música, la telecomunicación y la medicina (Mascorro & Torres, 2013).

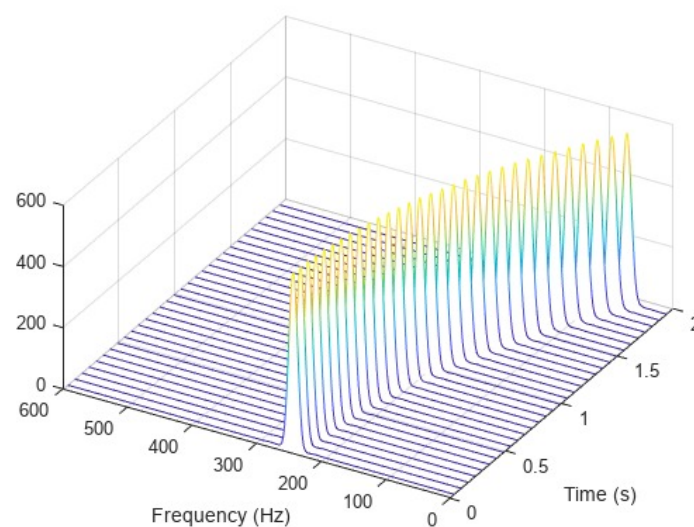


Figura 2.2: Muestra de un espectrograma (Paz et al., 2020).

2.7. Cromagrama

El propósito es condensar los datos del espectrograma en un conjunto compacto de valores, típicamente doce filas. Los cromagramas se fundamentan en la organización de las frecuencias de las notas en la escala cromática. Así, cada fila en el cromagrama refleja los datos espectrales de las frecuencias correspondientes a cada una de las doce notas musicales (una por cada fila) presentes en la escala cromática a lo largo de varias octavas, como se ilustra en la figura 2.3.

La información del espectrograma, y más particularmente la del cromagrama, puede ser utilizada para inferir características de alto nivel de la señal de música, como la tonalidad en la que se encuentra la pieza musical (utilizando, por ejemplo, modelos estadísticos de la frecuencia de aparición de cada nota) o la sucesión de acordes (conjuntos de notas simultáneas) (Benito Gorrón et al., 2017).

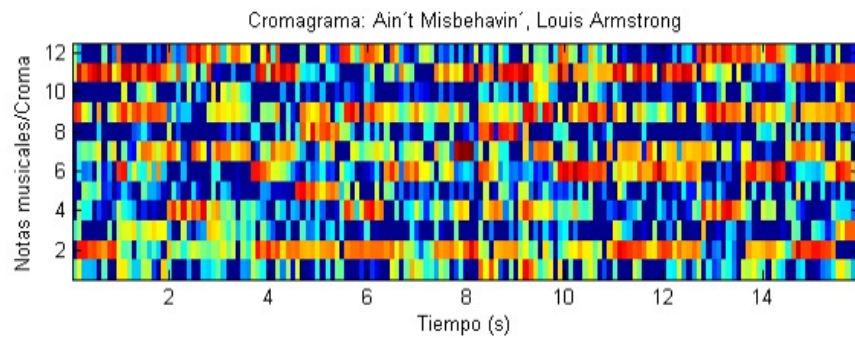


Figura 2.3: Muestra de Cromagrama (Benito Gorrón et al., 2017)

2.8. Transformada de Fourier de Tiempo Corto

Se emplea para examinar las variaciones en el contenido de frecuencia de una señal no estacionaria a medida que transcurre el tiempo. El cuadrado de la magnitud de la Transformada de Fourier de Tiempo Corto (STFT) se denomina la representación tiempo-frecuencia del espectrograma de la señal, y esta se visualiza en la figura 2.4 (Moros Vivas, 2010).

$$K = \left\lceil \frac{Nx - L}{M - L} \right\rceil \quad (2.1)$$

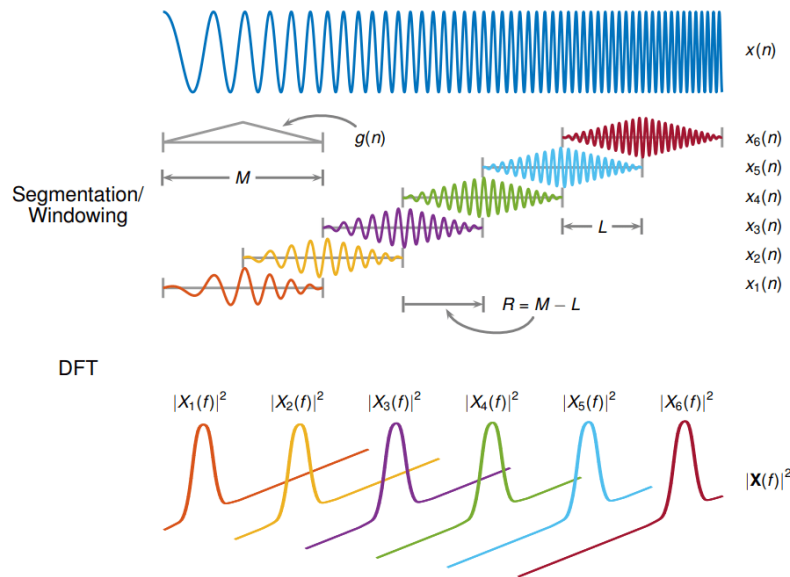


Figura 2.4: Representación de Transformada de Fourier de tiempo corto (Moler, 2022b)

2.9. Callbacks

Los callbacks son funciones pasadas como argumentos a otras funciones. Esta técnica permite la ejecución de la función callback después de que se cumpla una determinada condición o evento, siendo una práctica habitual en el uso de bibliotecas y frameworks, especialmente en el desarrollo de aplicaciones con JavaScript, como JQuery, Angular o Node.js. Estas estructuras son perfectas para crear funcionalidades que necesitan activarse solo después de que ocurra un evento específico, facilitando así el desarrollo de aplicaciones dinámicas y reactivas a las acciones del usuario (Guide, 2020).

2.10. Early Stopping

La técnica de detención temprana en el aprendizaje automático es un método de regularización que tiene como objetivo prevenir el sobreajuste durante la fase de entrenamiento de modelos mediante procesos iterativos, tales como el descenso de gradiente. Esta estrategia implica el seguimiento de determinadas métricas, como la exactitud o el error, a lo largo del proceso de entrenamiento. Al observar que estas métricas ya no experimentan mejoras significativas o comienzan a estabilizarse, la detención temprana actúa deteniendo el entrenamiento anticipadamente. Esto ayuda a asegurar que los modelos resultantes sean capaces de generalizar mejor al

enfrentarse a nuevos datos, evitando así que aprendan y memoricen detalles irrelevantes o ruido presentes en el conjunto de datos de entrenamiento que no contribuyen a una comprensión genuina de los patrones subyacentes (Moler, 2022a).

2.11. ReduceLROnPlateau

Reduce learning rate se utiliza cuando la métrica ha terminado o detenido su desarrollo para cuantificar el porcentaje de desarrollo obtenido y reducir tiempos de aprendizaje futuro (Al-Kababji et al., 2022).

2.12. Capas LSTM

Se trata de una capa de Red Neuronal Recurrente (RNN) diseñada para captar relaciones a largo plazo entre instantes temporales en series temporales y datos secuenciales, tal como se ilustra en la figura 2.5.

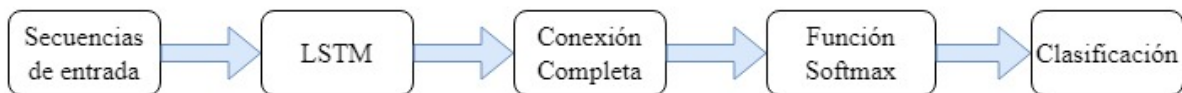


Figura 2.5: Diagrama de secuencia de una capa LSTM.

2.13. Mel Frequency Cepstral Coefficients (MFCCs)

Los Coeficientes Cepstrales de Frecuencias Mel (MFCCs) son una serie de características cruciales para la representación del habla, basadas en cómo el oído humano percibe el sonido. Estos coeficientes surgen de la necesidad de identificar eficazmente los aspectos significativos de una señal de audio en aplicaciones como el reconocimiento automático de voz, eliminando al mismo tiempo elementos innecesarios como el ruido ambiental, las variaciones emocionales, el volumen y el tono, los cuales pueden interferir en el proceso de reconocimiento. Los MFCCs

juegan un rol esencial al filtrar estos aspectos irrelevantes y centrarse en las características sonoras que verdaderamente importan para el reconocimiento, lo que resulta en una eficiencia mejorada en la identificación automática del contenido auditivo. La figura 2.6 muestra la potencia espectral de una señal de audio (Logan et al., 2000).

Los coeficientes cepstrales de frecuencia mel (MFCCs, por sus siglas en inglés) representan un elemento esencial en la identificación automática de voz o de hablantes, introducidos en la década de 1980 por Davis y Mermelstein. Desde su introducción, han mantenido su posición como la técnica líder en el campo.

MFCCs se calculan comúnmente de la siguiente forma:

- Separar la señal en pequeños tramos.
- A cada tramo aplicarle la Transformada de Fourier discreta 2.2 y obtener la potencia espectral de la señal.

$$Xn = \frac{1}{N} \sum_{k=0}^{N-1} x_k e^{\frac{2\pi}{N} K n} \quad n = 0, \dots, N-1 \quad (2.2)$$

- Implementar el conjunto de filtros asociados a la Escala Mel sobre el espectro adquirido en la etapa previa y calcular la suma de las energías en cada filtro. Donde para convertir hercios en mels se usa 2.3:

$$m = 1127,01048 \ln \left(1 - \frac{f}{700} \right) \quad (2.3)$$

- Calcular el logaritmo de las energías correspondientes a cada frecuencia Mel.
- Realizar la transformada de coseno discreta (DCT) 2.4 a los logaritmos obtenidos.

$$F(k) = \frac{\sqrt{2}c(k)}{\sqrt{N}} \sum_{j=0}^{N-1} f(j) \cos \left[\frac{(2j+1)k\pi}{2N} \right] \quad (2.4)$$

$$\text{Donde } c(k) = \frac{1}{\sqrt{2}} \text{ para } k = 0 \text{ para otros numeros enteros hasta } N-1 \quad (2.5)$$

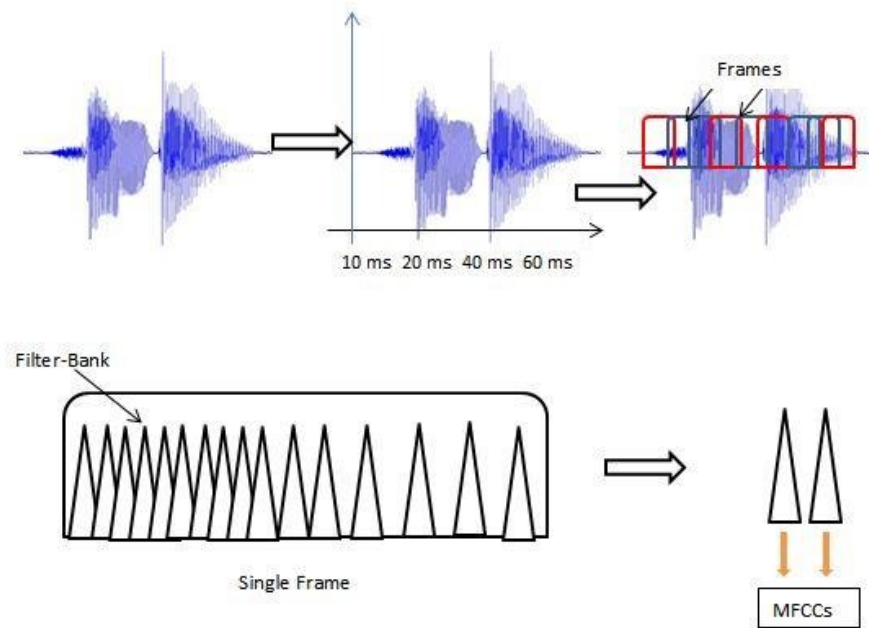


Figura 2.6: Potencia espectral de una señal (Franco-Galván, 2017).

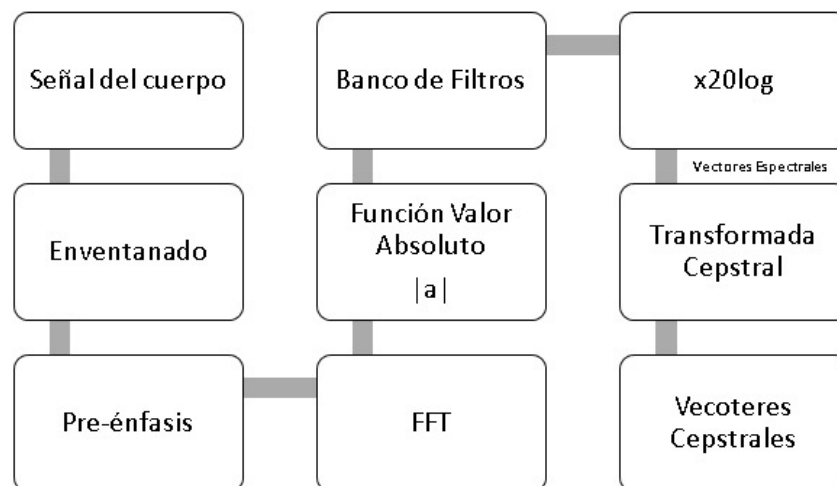


Figura 2.7: Diagrama de flujo de Los Coeficientes Cepstrales en las Frecuencias de Mel.

2.14. Redes Neuronales Convolucionales

El Deep Learning (DL), o aprendizaje profundo, habilita a modelos computacionales formados por múltiples capas de procesamiento para adquirir representaciones de datos en diversos niveles de abstracción, permitiéndoles así aprender de manera autónoma representaciones significativas en extensos conjuntos de datos. Este enfoque ha impulsado significativos progresos en áreas como el reconocimiento de imágenes y videos. Dentro del DL, las Redes Neuronales Convolucionales (CNN) se destacan en numerosas aplicaciones de visión por computadora gracias a su excepcional rendimiento en tareas de reconocimiento e interpretación de imágenes y videos. La efectividad de las CNN se debe a características clave como conexiones locales, pesos compartidos, la técnica de pooling y el empleo intensivo de varias capas. Las CNN buscan extraer características detalladas de las imágenes para luego utilizar estas para identificar o clasificar objetos. Los parámetros de los filtros aprendidos en estas capas se ajustan y optimizan en conjunto con los elementos de clasificación con el fin de reducir el error de clasificación global. El surgimiento de las CNN ha propiciado avances rápidos en el ámbito de la detección de objetos. Por esto nos permitimos mencionar los algoritmos avanzados y más recientes en el campo de detección de objetos con redes neuronales convolucionales: Faster R-CNN, Single Shot Multibox Detector (SSD) y You Only Look Once (YOLO) (Massiris et al., 2018).

Las redes neuronales convolucionales (CNN) se estructuran con múltiples capas de filtros convolucionales de una o más dimensiones. Tras cada capa, se suele incorporar una función que realiza un mapeo causal no lineal (Artola Moreno, 2019).

En el contexto de la clasificación, la CNN experimenta inicialmente una fase de extracción de características, compuesta por neuronas convolucionales y de reducción de muestreo. Al final de la red, se hallan neuronas de perceptrón simples que llevan a cabo la clasificación definitiva basada en las características extraídas. La etapa de extracción de características guarda similitudes con el proceso estimulante observado en las células de la corteza visual. Dicha fase consiste en capas alternas de neuronas convolucionales y neuronas de reducción de muestreo. A medida que los datos avanzan en esta fase, su dimensionalidad disminuye, haciendo que las neuronas en capas más alejadas sean menos sensibles a perturbaciones en los datos de entrada,

aunque, al mismo tiempo, sean activadas por características cada vez más complejas (BODERO et al., 2020).

Al momento de la extracción de características, las neuronas simples del perceptrón son reemplazadas por procesadores en matriz que realizan operaciones sobre los datos que pasan por ellas y la salida de cada una de las neuronas se calcula bajo la fórmula 2.6:

$$Y_j = g \left(b_j + \sum_i K_{ij} \times Y_i \right) \quad (2.6)$$

2.15. Aumentación de Datos

La aumentación de datos comprende un conjunto de técnicas diseñadas para incrementar de manera artificial la cantidad de datos mediante la creación de nuevos puntos de datos a partir de datos preexistentes. Esto implica realizar ajustes mínimos a los datos o emplear modelos de aprendizaje profundo para la generación de nuevos puntos de datos.

Esta se utiliza en el campo del aprendizaje automático y la visión por computadora para mejorar el rendimiento de los modelos al aumentar la diversidad y cantidad de datos disponibles para el entrenamiento. Este enfoque ayuda a prevenir el sobre ajuste del modelo a un conjunto de datos limitado y a mejorar la capacidad del modelo para generalizar a datos nuevos y no vistos (Shorten et al., 2021).

Marco Metodológico

Este capítulo describe de forma sistemática las técnicas y procesos utilizados en el procesamiento y análisis de los datos acústicos. Esta sección también incluye el diseño e implementación de los modelos de machine learning y las estrategias de evaluación.

3.1. Descripción General

Este trabajo utiliza señales acústicas de pacientes extranjeros para desarrollar un sistema biomédico avanzado, aplicando técnicas de machine learning. Se busca mejorar la precisión en el diagnóstico de enfermedades cardio pulmonares mediante un enfoque integral que combina conjuntos de datos específicos, un detallado preprocesamiento de señales y la extracción de características relevantes. La fase de clasificación, respaldada por modelos de aprendizaje automático optimizados, promete contribuir significativamente a la detección temprana y tratamiento efectivo de estas afecciones.

3.2. Adquisición de Datos

Este estudio tiene como objetivo el uso de técnicas de machine learning para clasificar enfermedades cardio pulmonares a partir de conjuntos de datos correspondientes a sonidos cardíacos y pulmonares, incluidos en el Apéndice A. Kaggle, una de las plataformas de Data Science más reconocidas a nivel mundial, fue utilizada para la obtención de los datos, debido a que no existen registros de sonidos pulmonares y cardíacos en el Ecuador. Los conjuntos de datos incluyen grabaciones de alta calidad de pacientes saludables y no saludables de origen extranjero y fueron extraídos con estetoscopios electrónicos. El conjunto de datos contiene grabaciones de

audio del examen de la pared torácica en varios puntos estratégicos. La colocación del estetoscopio en el sujeto fue determinada por el médico especialista que realizó el diagnóstico. Cada grabación fue replicada tres veces correspondientes a varios filtros de frecuencia que enfatizan ciertos sonidos corporales.

La combinación de los dos conjuntos de datos usados en este estudio permite un análisis exhaustivo de las enfermedades cardio pulmonares desde dos perspectivas complementarias. Mientras que uno proporciona una visión profunda de la salud cardíaca, el otro ofrece una perspectiva igualmente importante sobre la salud pulmonar. Juntos, estos conjuntos de datos forman una base sólida para un sistema de clasificación integral que puede identificar una gama más amplia de condiciones cardio pulmonares, mejorando así la precisión del diagnóstico y la eficacia del tratamiento.

3.2.1. Información sobre el Primer Conjunto de Datos

El primer conjunto de datos se centra en la clasificación de sonidos cardíacos, incluyendo categorías como sístoles, extrasístoles, y otros sonidos relacionados con el funcionamiento del corazón. Este conjunto de datos, originalmente creado para un desafío de aprendizaje automático destinado a clasificar sonidos cardíacos, fue recopilado a partir de dos fuentes: (A) del público en general a través de la aplicación iStethoscope Pro para iPhone y (B) de un ensayo clínico en hospitales utilizando el estetoscopio digital DigiScope. Las dos tareas asociadas con este conjunto de datos incluyen la segmentación de sonidos cardíacos normales (S1 y S2) en archivos de audio y la clasificación de los sonidos cardíacos en una de cuatro categorías distintas. Este recurso invaluable proporciona etiquetas y metadatos en los archivos *set_a.csv* y *set_b.csv*, así como información de temporización de referencia en *set_a_timing.csv*. Los archivos de audio, con longitudes variables entre 1 segundo y 30 segundos, son cruciales para el desarrollo y evaluación de métodos de clasificación de sonidos cardíacos.

- **Conjunto de Datos:** Sonidos cardíacos para clasificación de enfermedades.

- **Origen:** Kaggle, recopilado a través de la aplicación iStethoscope Pro y ensayo clínico con DigiScope.
- **Tipo de Señales:** Sonidos Cardíacos.
- **Tareas Asociadas:** Segmentación de S1 y S2, Clasificación de Sonidos Cardíacos.
- **Número de Categorías:** Cuatro categorías distintas de sonidos cardíacos.
- **Disponibilidad:** Público en Kaggle y disponible en Apéndice A. La tabla 3.1 presenta una descripción general del conjunto de datos.

Tabla 3.1: Descripción del dataset *set_a.csv*

dataset	fname	label	sublabel
a	set_a/artifact__201012172012.wav	artifact	
a	set_a/extrahls__201101070953.wav	extrahls	
a	set_a/murmur__201101051104.wav	murmur	
a	set_a/normal__201101070538.wav	normal	
a	set_a/____201106111419.wav		
...			

Tabla 3.2: Descripción del dataset *set_b.csv*

dataset	fname	label	sublabel
b	set_b/Btraining_extrastole.....wav	extrastole	
b	set_b/Btraining_murmur.....wav	murmur	
b	set_b/Btraining_normal.....wav	normal	
b	set_b/Bunlabelledtest.....wav		
b	set_b/Btraining_murmur.....wav	murmur	noisymurmur
b	set_b/Btraining_normal.....wav	normal	noisynormal
...			

3.2.2. Información sobre el Segundo Conjunto de Datos

Por otro lado, el segundo conjunto de datos se centra en la clasificación de sonidos pulmonares, abarcando condiciones como neumonías, taquicardias, bradicardias, entre otras. Se utilizó el conjunto de datos proporcionado por Fraiwana et al (Fraiwana, 2021). que se enfoca en sonidos pulmonares grabados desde la pared torácica utilizando un estetoscopio electrónico. Este conjunto de datos presenta grabaciones de 112 sujetos, abarcando diversas condiciones pulmonares, como asma, insuficiencia cardíaca, neumonía, entre otras. Cada grabación se replicó tres veces con diferentes filtros de frecuencia, permitiendo un análisis detallado de las características acústicas específicas de cada condición. La disponibilidad pública de este conjunto de datos en Mendeley Data facilita su acceso para investigadores y profesionales interesados en el desarrollo de algoritmos de detección de enfermedades pulmonares mediante el análisis de sonidos respiratorios.

- **Conjunto de Datos:** Sonidos pulmonares para la clasificación de enfermedades.
- **Origen:** Jordan University of Science and Technology, Jordan.
- **Tipo de Señales:** Sonidos Pulmonares.
- **Número de Sujetos:** 112 personas (35 saludables y 77 con enfermedades pulmonares).
- **Filtros de Frecuencia:** Tres filtros aplicados en cada grabación para resaltar diferentes características.
- **Disponibilidad:** Público en Mendeley Data y disponible en Apéndice A. La tabla 3.3 presenta una descripción general del conjunto de datos.

Tabla 3.3: Descripción de los Datos

Age	Gender	Location	Sound type	Diagnosis
70	M	P L L	I E W	Asthma
72	F	P R L	I C	Heart Failure + Lung Fibrosis
71	M	P R L	I C E W	Heart Failure + COPD
81	M	P R L	I C B	Pleural Effusion
57	M	A R M	E W	COPD
57	M	P R U	C	Pneumonia
90	M	A R M	C	Asthma and Lung Fibrosis
20	M	P R L	Crep	BRON
50	M	P R L	Crep	Lung Fibrosis
76	F	P R L	Crep	Heart Failure
27	F	P L M	N	N

3.3. Preprocesamiento de Señales

Código 1: Análisis de Sonidos Cardiacos

3.3.0.1. Filtrado de Datos por Duración

En este código, se realiza la limpieza y filtrado de ruido mediante la aplicación de técnicas de procesamiento de señales. El código carga archivos de audio, descarta aquellos con una duración menor a 3 segundos y luego segmenta la señal en intervalos de tiempo específicos, aplicando técnicas de aumento de datos para mejorar la robustez del modelo ante variaciones en tono y velocidad.

3.3.0.2. Normalización de Señales

La duración de las señales se ajusta para asegurar que tengan una longitud consistente y se utilizan técnicas de aumento de datos para diversificar el conjunto de entrenamiento.

3.3.0.3. Segmentación y Ventaneo

La señal se segmenta en intervalos de 3 segundos con un desplazamiento adecuado para garantizar la cobertura total. La técnica de ventaneo se aplica durante la extracción de características para dividir la señal en tramas solapadas, lo que permite una representación más detallada de la señal.

3.3.0.4. Reducción de Dimensionalidad

La dimensión de las características extraídas se mantiene constante y se ajusta mediante relleno o truncamiento, lo que facilita el procesamiento por parte de la red neuronal.

3.3.0.5. Aumento de Datos (Data Augmentation)

El primer código destaca la aplicación de técnicas de aumento de datos para mejorar la diversidad y robustez del conjunto de entrenamiento. Se realiza un cambio de tono y velocidad en las señales, lo que contribuye a mejorar la generalización del modelo ante variaciones en las condiciones de grabación.

3.3.0.6. Ponderación de Clases

Se aplica ponderación de clases para abordar posibles desequilibrios en la distribución de clases en el conjunto de datos. Esto asegura que el modelo se entrene de manera equitativa en todas las clases, mejorando así su capacidad para reconocer diferentes categorías.

3.3.0.7. División de Conjuntos de Datos

El conjunto de datos se divide en conjuntos de entrenamiento y prueba, permitiendo evaluar la capacidad de generalización del modelo en datos no vistos.

3.3.0.8. Codificación de Etiquetas

Se utiliza la codificación de etiquetas (LabelEncoder) para convertir las etiquetas de clase en valores numéricos, facilitando la tarea de clasificación para el modelo.

Código 2: Análisis de Sonidos Pulmonares

3.3.0.9. Limpieza y Filtrado de Ruido

En este segundo código, se realiza la limpieza y filtrado de ruido a través de la creación de espectrogramas a partir de archivos de audio. La función `create_spectrogram` utiliza la transformada de Fourier para calcular el espectrograma, que visualiza cómo varía la intensidad de frecuencia en función del tiempo.

3.3.0.10. Normalización de Señales

Se utiliza la función `adjust_spectrogram_size` que se encarga de ajustar el tamaño de los espectrogramas para que todos tengan la misma longitud.

3.3.0.11. Reducción del Tamaño de Espectrogramas

Se generan los espectrogramas para la señal completa. La función `adjust_spectrogram_size` realiza ventaneo para ajustar el tamaño de los espectrogramas.

3.3.0.12. Información Extraída de Nombres de Archivos

El código extrae información relevante de los nombres de archivos, como el diagnóstico, tipo de sonido, ubicación, edad y género. Esta información se utiliza para etiquetar y asociar cada archivo de audio con datos específicos.

3.3.0.13. Extracción de Características Espectrales

La extracción de características se realiza indirectamente mediante la creación de espectrogramas. Estos proporcionan información detallada sobre la variación de la energía de la señal con respecto al tiempo y la frecuencia.

3.3.0.14. Aumento de Datos (Data Augmentation)

Se procede con la creación de espectrogramas, proporciona una representación diversificada de las señales de audio.

3.4. Extracción de Características

3.4.1. Transformada de Fourier de Corto Tiempo (STFT)

La transformada de Fourier de corto tiempo se utiliza implícitamente en la extracción de características al calcular los espectrogramas de Mel y otros descriptores espectrales.

3.4.2. Espectrogramas de Mel

Se calculan los espectrogramas de Mel utilizando la transformada de Mel, proporcionando una representación visual de cómo varía la energía de la señal con el tiempo en diferentes frecuencias.

3.4.3. Extracción de Características Espectrales

Las características espectrales se extraen utilizando la transformada de Mel y se calculan los espectrogramas de Mel, los coeficientes espectrales en las frecuencias de Mel (MFCC), el cromagrama y el contraste espectral.

3.4.4. Coeficientes Cepstrales en las Frecuencias de Mel (MFCC)

Los MFCCs se obtienen a partir de los espectrogramas de Mel, capturando las características más relevantes del espectro de frecuencia de la señal de audio.

3.4.5. Cromagrama

El cromagrama se extrae para resaltar la información tonal en la señal, mostrando la distribución de energía en diferentes notas musicales a lo largo del tiempo.

3.4.6. Contraste Espectral

El contraste espectral se calcula para medir la diferencia de amplitud entre picos espectrales en diferentes bandas de frecuencia.

3.4.7. Características Temporales

Además de las características espectrales, se incorporan características temporales mediante la segmentación de la señal y la extracción de características en intervalos de tiempo específicos.

3.4.8. Normalización de Características

Ajuste de características en longitud para asegurar consistencia en la entrada del modelo.

3.4.9. Selección de Características

Uso de relleno o truncamiento para asegurar que todas las características tengan la misma longitud.

3.5. Entrenamiento del Modelo de Clasificación

3.5.1. Validación cruzada

Se dividen los conjuntos de datos en 2 partes, una destinada netamente para el entrenamiento de los modelos de machine learning, y la otra que se usará en la etapa de validación.

3.5.2. Configuración de Hiper Parámetros

Los hiperparámetros del modelo, como la tasa de aprendizaje y los parámetros de las capas, son configurados antes del entrenamiento para ajustarse a la naturaleza del problema.

3.5.3. Callbacks

Se utilizan callbacks como Early Stopping y ReduceLROnPlateau durante el entrenamiento del modelo. Early Stopping ayuda a prevenir el sobreajuste al detener el entrenamiento cuando la métrica de validación deja de mejorar, mientras que ReduceLROnPlateau ajusta dinámicamente la tasa de aprendizaje para mejorar la convergencia.

3.5.4. Monitoreo del Rendimiento

Después del entrenamiento, se realiza un análisis detallado del rendimiento del modelo mediante la visualización de curvas de pérdida y precisión en conjuntos de entrenamiento y validación. Esto proporciona información sobre el proceso de aprendizaje y la capacidad de generalización del modelo.

Código 1: Análisis de Sonidos Cardiacos

3.5.4.1. Regularización

Se incorpora regularización en la capa densa final del modelo mediante el uso de la regularización L1 y L2. Esto ayuda a prevenir el sobre ajuste al penalizar coeficientes grandes y favorecer modelos más simples.

3.5.4.2. Entrenamiento de Modelos

El modelo de clasificación se construye utilizando una arquitectura de red neuronal convolucional (CNN) con capas de convolución, normalización por lotes, capas de reducción y capas densas. La clasificación se realiza mediante una capa de salida softmax.

Código 2: Análisis de Sonidos Pulmonares

3.5.4.3. Entrenamiento de Modelos

Se utiliza una arquitectura de red neuronal que incluye capas convolucionales, de normalización por lotes, una capa LSTM y capas densas para realizar la clasificación. La función de activación softmax se utiliza en la capa de salida para la clasificación multiclase.

Resultados

Este capítulo describe los resultados generales obtenidos del análisis de los sonidos pulmonares y cardiacos y la clasificación de las enfermedades cardio pulmonares.

4.0.1. Descripción General

Los resultados obtenidos de ambos códigos demuestran la eficacia de los modelos de machine learning en la clasificación de sonidos cardio pulmonares, alineándose estrechamente con los objetivos y la justificación del proyecto. Estos hallazgos subrayan el potencial de tales sistemas para mejorar la precisión diagnóstica, servir como herramientas educativas y aumentar la accesibilidad a diagnósticos de calidad en multiples entornos médicos.

4.0.2. Resultados del Primer Código

En esta sección se presenta el rendimiento general del modelo, su desempeño por clases, y la interpretación de las métricas obtenidas.

4.0.2.1. Rendimiento General

El modelo alcanzó una precisión de prueba del 84.78 %, indicando una alta capacidad en la clasificación de sonidos cardio pulmonares.

4.0.2.2. Desempeño por Clases

La precisión más alta se observó en la clase 'artifact' (97 %), seguida de 'normal' (90 %) y 'murmur' (84 %). Esto sugiere una eficacia particularmente alta del modelo en la identificación de artefactos y sonidos normales, tal y como se puede ver en la tabla 4.1.

Tabla 4.1: Rendimiento del modelo por clase.

Clase	Precisión	Recall	F1-score	Soporte
Artifact	0.97	0.94	0.96	34
Extrahls	0.67	1.00	0.80	14
Extrastole	0.48	0.72	0.58	18
Murmur	0.97	0.85	0.91	81
Normal	0.88	0.85	0.87	142

La matriz ilustrada en la figura 4.1, destaca la capacidad del modelo para identificar ciertas clases con mayor precisión que otras, lo que es un indicador clave para entender las fortalezas y debilidades del sistema en el contexto clínico.

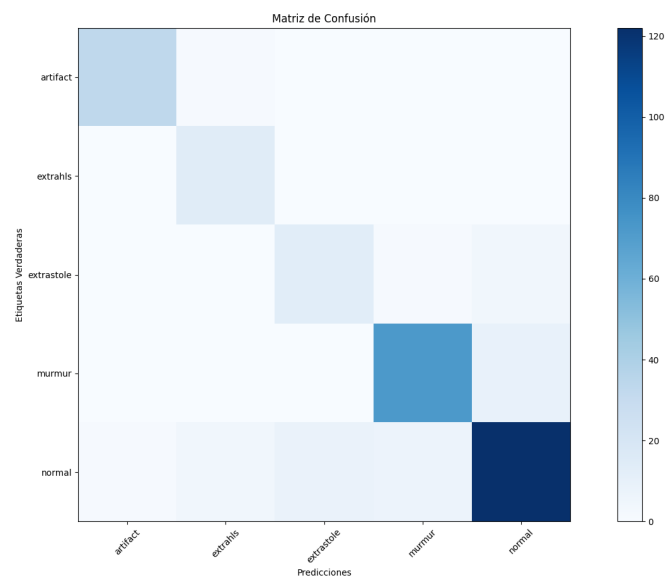


Figura 4.1: Matriz de confusión.

4.0.2.3. Curvas de Pérdida

Las curvas de pérdida en las etapas iniciales y finales del entrenamiento del modelo, ilustradas en las figuras 4.2 y 4.3 respectivamente, reflejan como la pérdida decrece a lo largo del tiempo, indicando un aprendizaje efectivo y una buena generalización.

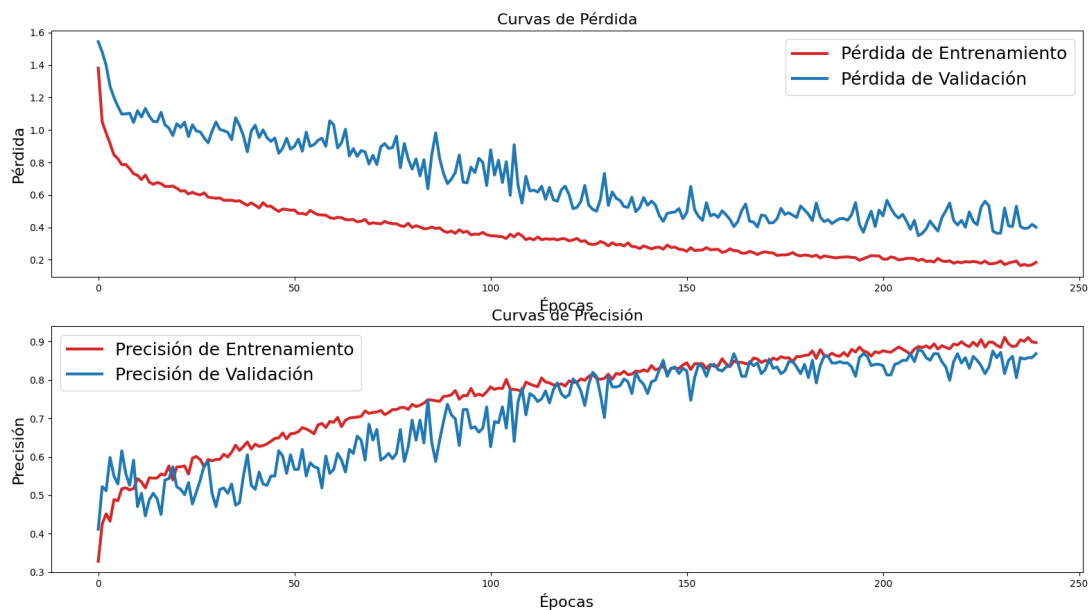


Figura 4.2: Curvas de Pérdida etapas iniciales.

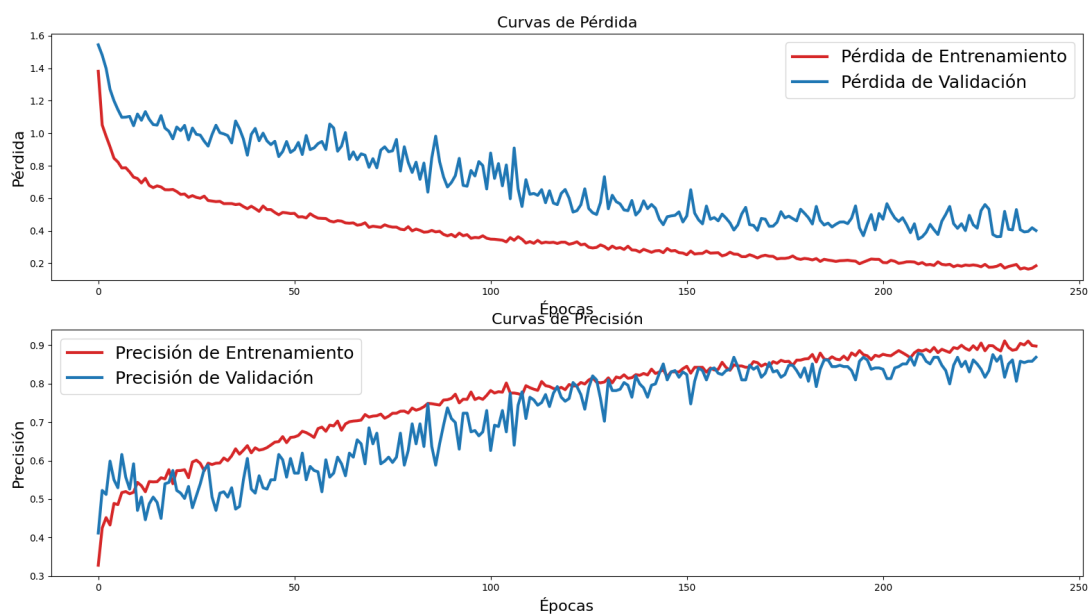


Figura 4.3: Curvas de Pérdida etapas finales.

4.0.2.4. Espectrogramas de Ejemplo

Las figuras 4.4, 4.5 y 4.6, demuestran cómo diferentes tipos de sonidos cardio pulmonares son interpretados y clasificados por el modelo.

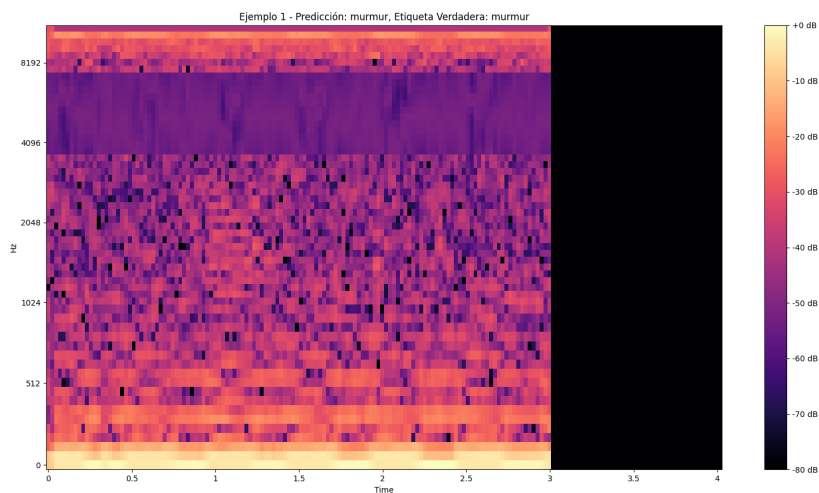


Figura 4.4: Ejemplo 1 de los espectrogramas analizados

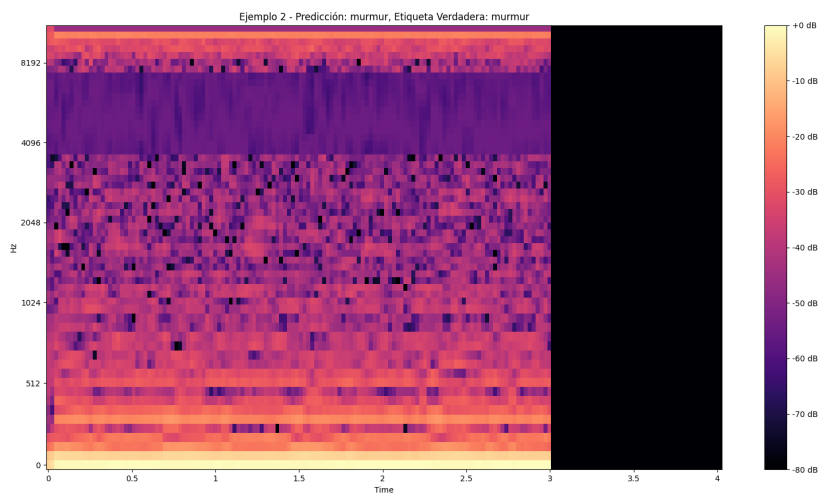


Figura 4.5: Ejemplo 2 de los espectrogramas analizados.

4.0.3. Resultados del Segundo Código

En esta sección, se analiza el rendimiento del modelo, destacando el progreso durante el entrenamiento y validación, así como ejemplos visuales de los datos analizados.

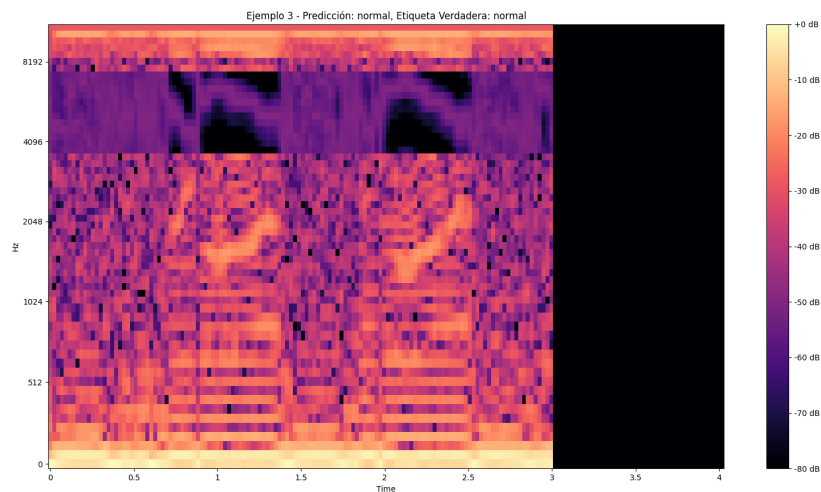


Figura 4.6: Ejemplo 3 de los espectrogramas analizados.

4.0.3.1. Rendimiento del Entrenamiento

La figura 4.7 muestra una mejora constante del modelo, alcanzando una precisión del 100 % en el entrenamiento y 95.24 % en la validación en las últimas épocas.

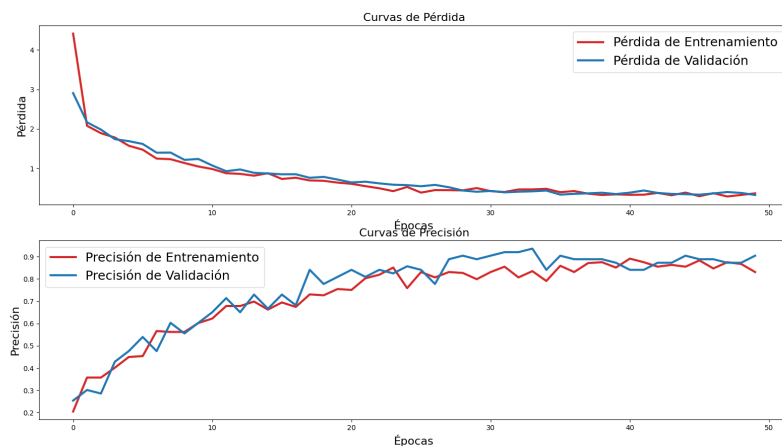


Figura 4.7: Curvas de pérdida y precisión durante el entrenamiento y la validación.

4.0.3.2. Reducción de la Tasa de Aprendizaje

La implementación de 'ReduceLROnPlateau' resultó en una disminución efectiva de la tasa de aprendizaje, mejorando la precisión de validación hasta un 96.83 %.

4.0.3.3. Detención Temprana

El entrenamiento se detuvo en la época 49 debido a la falta de mejora, indicando una convergencia exitosa del modelo.

4.0.3.4. Distribución de Duraciones de Audio

Analizamos la distribución de las duraciones de los audios que se puede ver en la figura 4.8 para comprender mejor el conjunto de datos.

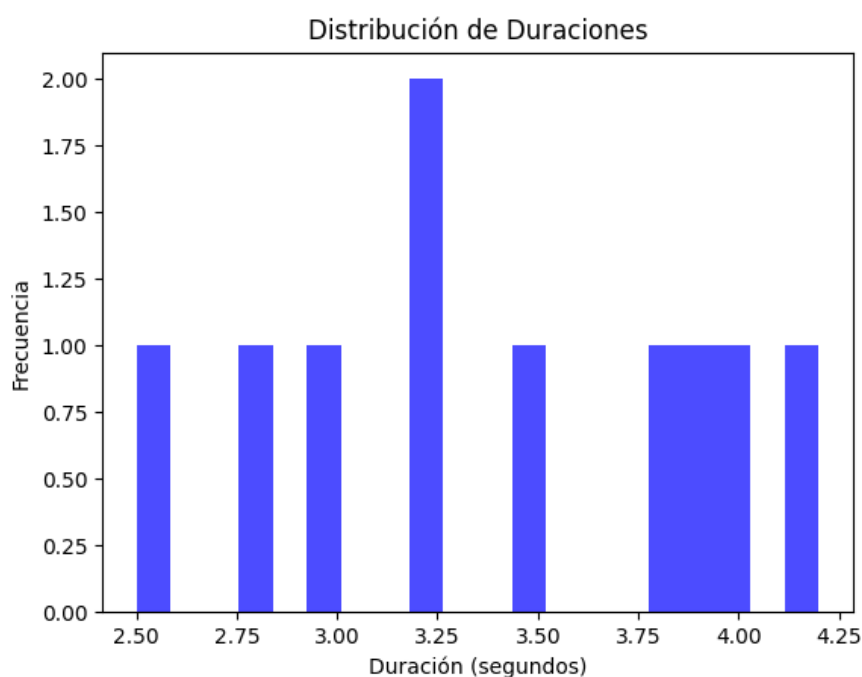


Figura 4.8: Histograma de las duraciones de los audios.

4.0.3.5. Ejemplos de Espectrogramas

Los espectrogramas mostrados en las figuras 4.9, 4.10 y 4.11 corresponden a ejemplos representativos de cómo el modelo procesa y clasifica diferentes tipos de sonidos.

4.0.3.6. Implicaciones para la Práctica Clínica

Estos resultados resaltan la capacidad del modelo para distinguir con precisión entre diferentes sonidos cardio pulmonares, lo que respalda su potencial uso en entornos clínicos.

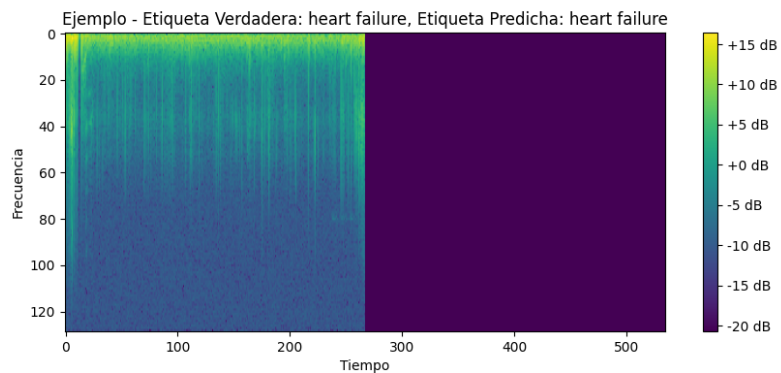


Figura 4.9: Ejemplo de espectrograma 1.

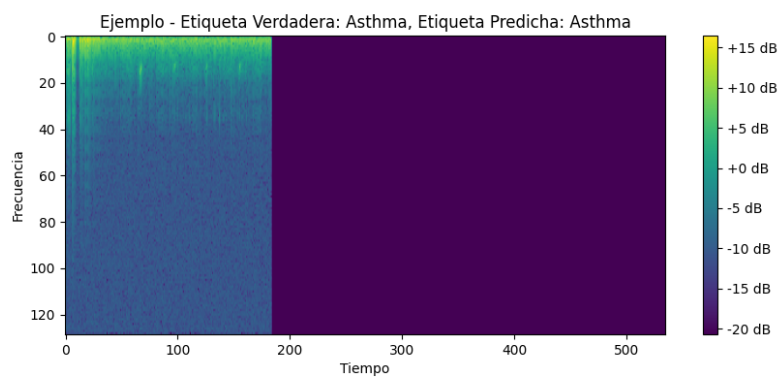


Figura 4.10: Ejemplo de espectrograma 2.

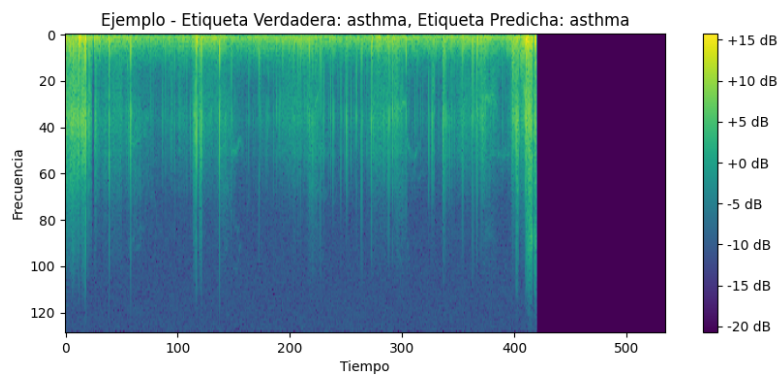


Figura 4.11: Ejemplo de espectrograma 3.

4.0.4. Discusión

4.0.4.1. Mejora en la Precisión Diagnóstica

Esta investigación demuestra un avance significativo en la precisión diagnóstica, crucial en el contexto de enfermedades cardio pulmonares. Los resultados del modelo sugieren una reducción notable en la subjetividad, respaldado por su alta precisión en la clasificación. Sería pertinente discutir cómo este aumento de precisión se compara con las prácticas diagnósticas actuales y su impacto potencial en la reducción de errores diagnósticos.

4.0.4.2. Herramienta Educativa

El desglose detallado de las métricas por clase indica que el sistema podría ser extremadamente valioso en un entorno educativo. Sería enriquecedor explorar cómo los estudiantes de medicina o los profesionales en formación que cursan carreras relacionadas con el sector de la salud pueden utilizar el modelo para familiarizarse con variaciones sutiles en sonidos cardíacos, algo que tradicionalmente requiere años de experiencia para desarrollar.

4.0.4.3. Accesibilidad en Diversos Entornos Médicos

La eficacia del modelo en la clasificación precisa de sonidos cardio pulmonares subraya su potencial para mejorar el diagnóstico en entornos con recursos limitados.

4.0.4.4. Avance en la Tecnología Biomédica

Los resultados demuestran la viabilidad de integrar el machine learning en el campo biomédico, abriendo caminos para futuras investigaciones y aplicaciones clínicas.

4.0.4.5. Respuesta a Necesidades Clínicas Urgentes

Dada la alta prevalencia de enfermedades cardio pulmonares, estos sistemas podrían contribuir significativamente a la detección temprana y el manejo efectivo de estas condiciones.

Cronograma

El desarrollo del sistema biomédico propuesto en este trabajo comprendió una serie de actividades, que iniciaron con la identificación del problema y finalizaron con el análisis de los resultados. La tabla 5.1 presenta un resumen de las actividades realizadas desde octubre del 2023 hasta enero del 2024.

Tabla 5.1: Cronograma de Actividades

No	Tareas por Desarrollar	oct-23				nov-23				dic-23				ene-24			
		1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
1	Propuesta del Proyecto de trabajo de titulación	x	x														
2	Identificación del problema del proyecto	x	x														
3	Elaboración de la Justificación			x	x												
4	Elaboración de objetivos			x	x												
5	Adquisición de datos					x	x										
6	Extracción de características					x	x										
7	Desarrollo de Metodología							x	x	x	x	x	x	x			
8	Diseño final del Proyecto														x	x	
9	Pruebas de fiabilidad														x	x	
10	Análisis de Resultados																x

Presupuesto

El desarrollo del sistema biomédico para la clasificación de enfermedades cardio pulmonares mediante el análisis de señales acústicas, utilizando técnicas de machine learning, supuso una inversión de \$ 2200. La tabla 6.1 describe el desglose de los gastos realizados.

Tabla 6.1: Presupuesto

Elemento	Cantidad	Parcial	Detalle
PC n°1	1	\$450.00	PC de J. Pino Cruz.
PC n°2	1	\$800.00	PC de A. Yanez Pazmiño.
Mano de obra	60 días	\$900.00	Diseño del sistema.
Movilización	5	\$50.00	Transporte a UPS y otros.

Conclusiones

Para la realización de este trabajo, se seleccionó Python como el lenguaje de programación principal, motivado por la riqueza de su ecosistema en análisis de datos y la robustez en el desarrollo de modelos de Machine Learning (ML). La elección de Python facilitó la implementación de algoritmos complejos y el manejo de grandes volúmenes de datos, gracias a su sintaxis clara y concisa.

Para la implementación específica de los modelos, optamos por librerías especializadas que proporcionan herramientas eficientes y especializadas para el análisis de música y audio, así como para el diseño y entrenamiento de modelos de Deep Learning. Librosa fue seleccionada por su capacidad para extraer características relevantes de las señales acústicas, crucial para la preparación de datos en estudios de análisis acústico. TensorFlow y Keras, por otro lado, se eligieron por su flexibilidad y facilidad de uso, lo que permitió un desarrollo ágil de modelos complejos, además de facilitar la experimentación con diferentes arquitecturas y técnicas de optimización.

La elección de CNN en el primer código se justifica por la probada eficacia de estas redes en el análisis y clasificación de imágenes, extendiendo su aplicación a espectrogramas, que transforman las señales acústicas en representaciones visuales bidimensionales. Esta técnica permite capturar patrones locales y características distintivas en los datos acústicos, fundamentales para la clasificación de condiciones cardiopulmonares.

En el segundo código, la combinación de CNN con LSTM (Long Short-Term Memory) refleja una estrategia avanzada para capturar tanto las características espaciales a través de CNN como las dependencias temporales mediante LSTM. Las CNN son eficaces en identificar patrones locales en datos bidimensionales, mientras que las LSTM gestionan secuencias de datos,

permitiendo analizar cómo evolucionan las características acústicas a lo largo del tiempo, lo cual es esencial para entender las señales acústicas complejas asociadas a condiciones cardiopulmonares.

Uno de los principales desafíos del proyecto fue el manejo y preprocesamiento de las señales acústicas para asegurar entradas uniformes y significativas para el modelo. Librosa jugó un papel crucial en este aspecto, facilitando la extracción de características y el preprocesamiento de los datos. Además, se emplearon técnicas de aumento de datos para enriquecer el conjunto de entrenamiento y mejorar la capacidad de generalización del modelo. El desbalance de clases se abordó mediante la ponderación durante el entrenamiento, una funcionalidad soportada eficazmente por TensorFlow y Keras.

Este estudio no solo destaca por su potencial para impactar en el diagnóstico y seguimiento de enfermedades cardiopulmonares, sino también por la cuidadosa selección de herramientas y metodologías. La documentación completa, códigos e instrucciones para continuar con este trabajo se han compartido en el siguiente repositorio de Git Hub [A](#), facilitando así la colaboración y expansión futuras de la investigación.

Recomendaciones

Con base a los análisis realizado y los resultados obtenidos, se recomienda realizar pruebas adicionales con un conjunto de datos cardíacos y pulmonares más amplio y diverso, especialmente recolectado de pacientes ecuatorianos. Esto facilitará la validación clínica y adopción del sistema en diferentes entornos clínicos y demográficos. Paralelamente, se debe considerar la realización de ajustes y la exploración de técnicas de optimización que permitan mejorar la precisión y eficiencia de los algoritmos. En particular, el entrenamiento de los algoritmos con datos atípicos o menos comunes, resultará en un diagnóstico más efectivo y eficiente de enfermedades cardio pulmonares.

Para su integración efectiva en la práctica médica y como herramienta educativa, se debe desarrollar una interfaz de usuario intuitiva y accesible, que pueda ser utilizada fácilmente por profesionales de la salud de diferentes niveles de habilidad técnica. La interfaz debe caracterizarse por ser amigable con el usuario y ser compatible con diferentes navegadores y plataformas móviles. La implementación del sistema propuesto en entornos clínicos existentes debe venir acompañada de capacitación del personal, actualizaciones de software y mantenimiento del mismo.

Es también fundamental asegurar que el sistema cumpla con las regulaciones de privacidad y ética médica, especialmente en lo que respecta al manejo de datos sensibles de los pacientes. Además, se debe considerar el diseño de un formulario que permita al paciente indicar que datos de su historial clínico puedan ser almacenados en el sistema y por cuanto tiempo. Por otro lado, se debe fomentar colaboraciones entre ingenieros biomédicos, médicos, y otros profesionales de la salud para garantizar que el sistema responda adecuadamente a las necesidades clínicas reales. Estas colaboraciones deben extenderse también a las instituciones educativas en

las que se encuentran estudiantes y médicos en formación para desarrollar tecnología que cubra necesidades de enseñanza.

Finalmente, es importante promover la investigación científica en diagnóstico médico entre las futuras generaciones de Biomédicos y médicos. Esto incluye explorar el uso de técnicas avanzadas de machine learning y su soporte al diagnóstico y tratamiento de otras enfermedades. Además, incluir en el p nsum acad mico, asignaturas que integren los avances tecnol gicos con los problemas m s comunes de salud, a fin de contribuir a avances m dicos significativos.

Bibliografía

- Al-Kababji, A., Bensaali, F., & Dakua, S. P. (2022). Scheduling techniques for liver segmentation: Reducelronplateau vs onecyclelr. *International Conference on Intelligent Systems and Pattern Recognition*, 204-212.
- Ariza Nadjar, A., et al. (2013). Elaboración de un libro texto y material de laboratorio para la optativa introducción a la ingeniería Biomédica.
- Artola Moreno, Á. (2019). Clasificación de imágenes usando redes neuronales convolucionales en Python.
- Bank, I., Vliegen, H. W., & Bruschke, A. V. (2016). The 200th anniversary of the stethoscope: Can this low-tech device survive in the high-tech 21st century? *European heart journal*, 37(47), 3536-3543.
- Benito Gorrón, D. d., et al. (2017). *Detección de música en contenidos multimedia mediante ritmo y armonía* [B.S. thesis].
- Bobadilla, J. (2021). *Machine learning y deep learning: usando Python, Scikit y Keras*. Ediciones de la U.
- BODERO, E. M., LOPEZ, M. P., CONGACHA, A. E., CAJAMARCA, E. E., & MORALES, C. H. (2020). Google Colaboratory como alternativa para el procesamiento de una red neuronal convolucional. *Revista Espacios*, 41(07).
- Canal, N. M. (2003). La historia de la medicina en la formación del profesional en medicina: tres casos históricos destacados. *Revista Colombiana de Filosofía de la ciencia*, 4(9), 175-202.
- Cruz Hernández, J., Hernández García, P., Abraham Marcel, E., Dueñas Gobel, N., & Salvato Dueñas, A. (2012). Importancia del método clínico. *Revista Cubana de Salud Pública*, 38(3), 422-437.

- Fraiwan, M. (2021). A dataset of lung sounds recorded from the chest wall using an electronic stethoscope.
- Franco-Galván, C. (2017, mayo). Mel Cepstral frequency coefficients MFCC [Accessed: 2024-1-29].
- Gismondi Glave, G. (2010). Ingeniería biomédica. *Revista Ciencia y Cultura*, (24), 99-118.
- González Fernández, R. (2007). El test de turing: dos mitos, un dogma.
- Guide, I. D. (2020, octubre). Callback: qué son las funciones callback? [Accessed: 2024-1-29].
- Haring, C. M., Cools, B. M., van der Meer, J. W., & Postma, C. T. (2014). Student performance of the general physical examination in internal medicine: an observational study. *BMC Medical education*, 14(1), 1-6.
- Hinestroza Ramírez, D. (2018). El Machine Learning a través de los tiempos, y los aportes a la humanidad.
- Kim, Y., Hyon, Y., Lee, S., Woo, S.-D., Ha, T., & Chung, C. (2022). The coming era of a new auscultation system for analyzing respiratory sounds. *BMC Pulmonary Medicine*, 22(1), 119.
- Lam, M., Lee, T., Boey, P., Ng, W., Hey, H., Ho, K., & Cheong, P. (2005). Factors influencing cardiac auscultation proficiency in physician trainees. *Singapore medical journal*, 46(1), 11.
- Lamas, J. L. T., & Arcuri, E. A. M. (2002).Cuál es la mejor pieza de auscultación para medir la presión arterial? *Temas enferm. actual*, 12-16.
- Logan, B., et al. (2000). Mel frequency cepstral coefficients for music modeling. *Ismir*, 270(1), 11.
- Lorenzano, C. (2006). El diagnóstico médico.
- Mascorro, G. A. M., & Torres, G. A. (2013). Reconocimiento de voz basado en MFCC, SBC y Espectrogramas. *Ingenius*, (10), 12-20.
- Massiris, M., Delrieux, C., & Fernández Muñoz, J. Á. (2018). Detección de equipos de protección personal mediante red neuronal convolucional YOLO. *XXXIX Jornadas de Automática*, 1022-1029.
- Mayorga, P., Druzgalski, C., González, O., Zazueta, A., & Criollo, M. (2011). Expanded quantitative models for assessment of respiratory diseases and monitoring. *2011 Pan American Health Care Exchanges*, 317-322.

- Moler, C. (2022a). Redes neuronales de memoria de corto-largo plazo - MATLAB & Simulink - MathWorks América Latina [Accessed: 2024-1-29].
- Moler, C. (2022b). stft [Accessed: 2024-1-29].
- Moros Vivas, A. M. (2010). Caracterización de señales de precipitación mediante la transformada de Fourier y transformada Wavelet.
- Musso, C. G., & Enz, P. A. (2015). Historia del estetoscopio. Eslabón final de una cadena de ideas y trabajo [Accessed: 2024-1-29].
- Paz, M. E., Friedrich, G. R., & Galasso, C. L. (2020). Procesamiento de señal visualizado sobre un espectrograma. *Elektron: ciencia y tecnología en la electrónica de hoy*, 4(1), 35-39.
- Rodríguez, A. D., González, V. D., & González, V. D. (2023). Papel de la auscultación cardiaca en el siglo XXI: la importancia de la formación. *Revista Española de Urgencias y Emergencias*, 2(3), 180-182.
- Rodríguez, E. (2003). Enseñanza de la Ingeniería Biomédica en Cuba. *Memorias V congreso de la Sociedad Cubana de Bioingeniería, Habana-Cuba*.
- Roelandt, J. (2014). The decline of our physical examination skills: is echocardiography to blame? *European Heart Journal—Cardiovascular Imaging*, 15(3), 249-252.
- Sarkar, M., Madabhavi, I., Niranjan, N., & Dogra, M. (2015). Auscultation of the respiratory system. *Annals of thoracic medicine*, 10(3), 158.
- Shorten, C., Khoshgoftaar, T. M., & Furht, B. (2021). Text data augmentation for deep learning. *Journal of big Data*, 8, 1-34.
- Simini, F., & Vienni, B. (2016). Ingeniería biomédica, interdisciplina y sociedad. *Revista Facultad de Ingeniería. Universidad Central de Venezuela*, 31, 15.
- Tavel, M. E. (2006). Cardiac auscultation: a glorious past—and it does have a future! *Circulation*, 113(9), 1255-1259.
- Williams, C. T. (1907). A lecture on Laennec and the evolution of the stethoscope: delivered before the medical Department of the University of Oxford. *British medical journal*, 2(2427), 6.
- Zou, Q. (2019). Latest machine learning techniques for biomedicine and bioinformatics. *Current Bioinformatics*, 14(3), 176-177.

Apéndice

- Enlace de acceso al repositorio *git*, donde se podrá acceder a los algoritmos desarrollados para la clasificación de enfermedades cardio pulmonares y a su respectiva documentación.

https://github.com/ayanezp/thesis_cardiopulmonarycodes.git

- Link del Dataset of Heartbeat Sounds.

<https://www.kaggle.com/datasets/kinguistics/heartbeat-sounds/data>

- Link del Dataset of Lung Sounds.

<https://www.kaggle.com/datasets/arashnic/lung-dataset>