



EDUCACIÓN

SECRETARÍA DE EDUCACIÓN PÚBLICA



TECNOLÓGICO
NACIONAL DE MÉXICO

Tecnológico Nacional de México

**Centro Nacional de Investigación
y Desarrollo Tecnológico**

Tesis de Maestría

**Reconocimiento de Arritmias Cardíacas Mediante un
Autoencoder Profundo Espectral**

presentada por

Ing. Enrique Quezada Próspero

como requisito para la obtención del grado de
Maestro en Ciencias de la Computación

Director de tesis

Dr. Dante Mújica Vargas

Codirector de tesis

Dr. Nimrod González Franco

Cuernavaca, Morelos, México. Julio de 2025.



Cuernavaca, Mor., 02/Junio/2025

OFICIO No. DCC/148/2025

Asunto: Aceptación de documento de tesis

CENIDET-AC-004-M14-OFICIO

CARLOS MANUEL ASTORGA ZARAGOZA
SUBDIRECTOR ACADÉMICO
PRESENTE

Por este conducto, los integrantes de Comité Tutorial de **ENRIQUE QUEZADA PRÓSPERO** con número de control **M23CE074**, de la Maestría en Ciencias de la Computación, le informamos que hemos revisado el trabajo de tesis de grado titulado **"Reconocimiento de arritmias cardiacas mediante un autoencoder profundo espectral"** y hemos encontrado que se han atendido todas las observaciones que se le indicaron, por lo que hemos acordado aceptar el documento de tesis y le solicitamos la autorización de impresión definitiva.

ATENTAMENTE

Excelencia en Educación Tecnológica®
"Conocimiento y Tecnología al Servicio de México"

Dr. Dante Mújica Vargas
Director de tesis

Dr. Noé Alejandro Castro Sánchez
Revisor 1

Dr. Nimrod González Franco
Codirector de tesis

Dr. Juan Gabriel González Serna
Revisor 2

C.c.p. Depto. Servicios Escolares.
Expediente / Estudiante



2025
Año de
**La Mujer
Indígena**

Interior Internado Palmira S/N, Col. Palmira,
C. P. 62490, Cuernavaca, Morelos Tel. 01 (777) 3627770, ext. 3201,
e-mail: dcc_cenidet@tecnm.mx tecnm.mx | cenidet.tecnm.mx

cenidet
Centro Nacional de Investigación y Desarrollo Tecnológico





Centro Nacional de Investigación y Desarrollo tecnológico
Subdirección Académica

Cuernavaca Mor, 04/junio/2025

Oficio No. SAC/119/2025

Asunto: Autorización de impresión de tesis

ENRIQUE QUEZADA PRÓSPERO
CANDIDATO AL GRADO DE MAESTRO
EN CIENCIAS DE LA COMPUTACIÓN
P R E S E N T E

Por este conducto, tengo el agrado de comunicarle que el Comité Tutorial asignado a su trabajo de tesis titulado **"Reconocimiento de arritmias cardiacas mediante un autoencoder profundo espectral"**, ha informado a esta Subdirección Académica, que están de acuerdo con el trabajo presentado. Por lo anterior, se le autoriza a que proceda con la impresión definitiva de su trabajo de tesis.

Esperando que el logro del mismo sea acorde con sus aspiraciones profesionales, reciba un cordial saludo.

ATENTAMENTE

Excelencia en Educación Tecnológica®

"Conocimiento y Tecnología al Servicio de México"

CARLOS MANUEL ASTORGA ZARAGOZA
SUBDIRECTOR ACADÉMICO



c.c.p. Departamento de Ciencias Computacionales
Departamento de Servicios Escolares

CMAZ/lmz



2025
Año de
La Mujer
Indígena

Interior Internado Palmira S/N, Col. Palmira,
C. P. 62490, Cuernavaca, Morelos Tel. 01 (777) 3627770, ext. 4104,
e-mail: acad_cenidet@tecnm.mx tecnm.mx | cenidet.tecnm.mx

cenidet
Centro Nacional de Investigación
y Desarrollo Tecnológico



Dedicatoria

Dedico esta tesis principalmente a mis padres Teresa y Jorge por siempre apoyarme e impulsarme en la realización de mis sueños y por no dejar que me rinda; y a todos a los que me acompañaron en esta travesía, incluyendo:

A mi abuela, por cuidarme y siempre ver lo mejor para mí.

A mi hermana Lilly, por escucharme y darme consejos cuando más los necesité.

A Maria, por haberme apoyado e inspirado a iniciar esta aventura y compartir este sueño, dandome su apoyo incondicional.

A mi camarada Orvil, por acompañarme en el inicio de esta aventura y hacer que no fuera tan complicada y pesada, y sobre todo, por haber estado y escuchado en momentos complicados.

A Cheetah, por haber sido la mejor mascota y sacarme una sonrisa siempre que llegaba a Tapachula.

A todos los reales que siempre creyeron en mí y han estado conmigo desde siempre.

A TØP, porque a pesar de que probablemente nunca lo sabrán, jamás habrán palabras suficientes para agradecerles por haberme salvado con su música y por darme la fuerza para seguir adelante en innumerables ocasiones.

En general, extendiendo esta dedicatoria a todas las personas que me acompañaron en este camino y me dieron su apoyo cuando lo necesité, gracias por todo.

“Algunas veces debes sangrar para saber que estás vivo y tienes alma, pero hace falta que aparezca alguien para mostrarte cómo” - TØP.

Que esta tesis sirva como prueba de que la determinación, esfuerzo y sobre todo, el apoyo de quienes creen en nosotros, pueden hacer posible cualquier meta, a pesar de las dificultades.

Sigo agradeciendo con hechos.

Agradecimientos

Agradezco al Consejo Nacional de Humanidades, Ciencias y Tecnologías (CONAHCYT) por el apoyo económico brindado durante mis estudios de maestría y al Tecnológico Nacional de México / Centro Nacional de Investigación y Desarrollo Tecnológico (CENIDET), por brindar las instalaciones y permitirme realizar los estudios de maestría.

Expreso mi más profundo reconocimiento y gratitud a mi director de tesis, el Dr. Dante Mújica Vargas, cuya inigualable guía, exigencia y vasto conocimiento fueron pilares fundamentales en la realización de este trabajo. Su liderazgo, visión y compromiso con la excelencia académica han sido una fuente de inspiración constante, permitiéndome crecer tanto profesional como personalmente. Sus invaluable aportes y orientación marcaron la diferencia en el desarrollo de esta investigación, dejando una huella imborrable en mi formación.

A mi Codirector, el Dr. Nimrod González Franco, le agradezco sus valiosos consejos y apoyo en este proceso. De igual manera, agradezco al comité revisor, conformado por el Dr. Gabriel González Serna y el Dr. Noé Alejandro Castro Sánchez, por sus observaciones y esfuerzo en el desarrollo de esta investigación. Extiendo también mi gratitud a los maestros Manuel Matuz Cruz y Christian García Aquino por iniciarme en este mundo y por su confianza, amistad, dedicación y ayuda para el desarrollo de este trabajo. A su vez, también agradezco a todos los integrantes del clan *Left Hand Path* por su apoyo incondicional brindado en este recorrido.

Finalmente, quiero expresar mi reconocimiento a todas aquellas personas que siempre creyeron en mí, apoyándome y animándome a lo largo de esta aventura. Su respaldo fue un pilar esencial en esta etapa de mi vida. Siempre los consideraré mi otra familia.

GRACIAS!

Resumen

Las arritmias cardíacas representan un desafío significativo en el ámbito de la salud, ya que su detección temprana es crucial para prevenir complicaciones graves. La clasificación automática de señales electrocardiográficas (ECG) ha cobrado relevancia en los últimos años debido al creciente interés en el desarrollo de herramientas computacionales eficientes para el diagnóstico de arritmias. Si bien existen múltiples enfoques para la clasificación de señales ECG, muchos de ellos dependen de arquitecturas profundas como redes convolucionales (CNN) o redes neuronales recurrentes (RNN), las cuales requieren grandes volúmenes de datos y una alta capacidad computacional para su entrenamiento. En contraste, los *Autoencoders* profundos han demostrado ser una alternativa efectiva debido a su capacidad para la extracción de características de datos complejos. En este contexto, el presente trabajo propone un modelo basado en un *Autoencoder* profundo combinado con técnicas de estimación espectral, con el objetivo de mejorar la clasificación de señales ECG mediante una arquitectura de bajo costo computacional y alto rendimiento.

Para evaluar el desempeño del modelo, se realizaron múltiples experimentos. Inicialmente, se analizaron señales en crudo, lo que evidenció limitaciones en la capacidad del modelo para discriminar patrones debido a la falta de características distintivas. Posteriormente, se incorporó un preprocesamiento al aplicar un filtro pasa-bajos, logrando una mejora en las métricas de clasificación, aunque con restricciones inherentes a la representación en el dominio del tiempo. Con el fin de optimizar el desempeño del modelo, se llevó a cabo una sintonización exhaustiva de hiperparámetros, lo que permitió incrementar significativamente la capacidad de clasificación.

Uno de los avances más relevantes del trabajo fue la transformación de las señales al dominio de la frecuencia mediante estimaciones espectrales, lo que permitió representar con mayor precisión las características discriminantes de cada clase. Entre las técnicas evaluadas, el espectro de Covarianza se destacó como la mejor opción, logrando métricas de calidad superiores al 95 %. Este enfoque permitió al modelo alcanzar un rendimiento robusto y competitivo en comparación con otras técnicas reportadas en la literatura.

Además, se evaluó la capacidad de generalización del modelo aplicándolo a una base de datos distinta, lo que permitió validar su estabilidad ante variaciones en la adquisición de señales ECG. Los resultados demostraron que el modelo mantuvo un desempeño elevado, demostrando ser una alternativa viable para la clasificación automática de arritmias en entornos clínicos.

Finalmente, se comparó el desempeño del modelo con arquitecturas avanzadas del estado del arte, demostrando que la propuesta logra una clasificación eficiente con una estructura más simple y menor costo computacional. Este hallazgo refuerza su aplicabilidad en sistemas de monitoreo embebidos y dispositivos de bajo costo.

Palabras clave: ECG, Costo computacional, Filtro pasa-bajos, Espectros, *Autoencoder*

Abstract

Cardiac arrhythmias pose a significant challenge in the healthcare field, as their early detection is crucial for preventing severe complications. The automatic classification of electrocardiographic (ECG) signals has gained relevance in recent years due to the growing interest in developing computational tools for arrhythmia diagnosis. While multiple approaches exist for ECG signal classification, many rely on deep architectures such as convolutional neural networks (CNNs) or recurrent neural networks (RNNs), which require large datasets and high computational power for training. In contrast, deep Autoencoders have proven to be an effective alternative due to their ability to extract meaningful features from complex data without the need for explicit labels. In this context, this work proposes a model based on a deep Autoencoder combined with spectral estimation techniques to enhance ECG signal classification through a low computational cost and high-performance architecture.

To evaluate the model's performance, multiple experiments were conducted. Initially, raw signals were analyzed, revealing limitations in the model's ability to discriminate patterns due to a lack of distinctive features. Subsequently, a preprocessing step based on filtering was introduced, improving classification metrics, although with inherent constraints related to time-domain representation. To optimize the model's performance, an exhaustive hyperparameter tuning process was carried out, significantly enhancing classification capability.

One of the most relevant advances of this work was the transformation of signals into the frequency domain using spectral estimations, allowing for a more precise representation of discriminative features for each class. Among the evaluated techniques, the Covariance spectrum proved to be the most effective, achieving accuracy and sensitivity metrics above 95 %. This approach enabled the model to reach a robust and competitive performance compared to other techniques reported in the literature.

Additionally, the model's generalization capability was assessed by applying it to a different dataset, validating its stability against variations in ECG signal acquisition. The results demonstrated that the model maintained high performance, establishing itself as a viable alternative for automatic arrhythmia classification in clinical environments.

Finally, the model's performance was compared with state-of-the-art architectures, such as BiLSTM, demonstrating that the proposed approach achieves efficient classification with a simpler structure and lower computational cost. This finding reinforces its applicability in embedded monitoring systems and low-cost devices.

Keywords: ECG, Computational cost, Low-pass filter, Spectrum, Autoencoder

Índice General

1	Introducción	1
1.1.	Planteamiento del problema	2
1.1.1.	Delimitación del problema	2
1.1.2.	Complejidad del problema	3
1.2.	Objetivos	3
1.2.1.	Objetivo general	3
1.2.2.	Objetivos específicos	3
1.3.	Alcances y limitaciones	3
1.3.1.	Alcances	3
1.3.2.	Limitaciones	4
1.4.	Justificación	4
1.5.	Organización de la tesis	4
2	Marco teórico	5
2.1.	Electrocardiograma	5
2.2.	Arritmias cardíacas	6
2.2.1.	Contracción Prematura Auricular	6
2.2.2.	Contracción Prematura Ventricular	6
2.2.3.	Bloqueo de Rama Izquierda	6
2.2.4.	Bloqueo de Rama Derecha	7
2.3.	<i>Dataset</i>	7
2.4.	Filtros Pasa-bajo	8
2.5.	Estimaciones Espectrales	9
2.5.1.	<i>FFT</i>	9
2.5.2.	<i>Periodogram</i>	9
2.5.3.	<i>Welch</i>	10
2.5.4.	<i>AR Burg</i>	10
2.5.5.	<i>Yule-Walker</i>	11
2.5.6.	<i>Covariance</i>	12
2.5.7.	Aprendizaje profundo	13
2.6.	Discusión	16
3	Estado del arte	17
3.1.	Antecedentes	17
3.2.	Trabajos relacionados	19
3.3.	Discusión del estado del arte	35
4	Metodología de Solución	42
4.1.	Preprocesamiento	42
4.2.	Modificación de <i>Autoencoder</i>	43
4.2.1.	Descripción General de la Arquitectura	43
4.2.2.	Descripción matemática	44
4.2.3.	Clasificación basada en las características codificadas	44

4.2.4.	Justificación de la arquitectura	46
4.2.5.	Consideraciones para mejoras	47
4.3.	Entrenamiento del modelo	47
4.4.	Clasificación	49
4.4.1.	Transformación y normalización de la señal	49
4.4.2.	Codificación de características latentes	49
4.4.3.	Clasificación mediante capa completamente conectada	49
4.4.4.	Evaluación del modelo	50
4.5.	Discusión	50
5	Experimentación y Resultados	51
5.1.	Entorno de desarrollo	51
5.2.	Gestión del <i>Dataset</i>	51
5.3.	Métricas de evaluación	53
5.4.	Experimentación y Resultados	54
5.4.1.	Experimento 1: Entrenamiento del <i>Autoencoder</i> con hiperparámetros por defecto/Señales en crudo	54
5.4.2.	Experimento 2: Entrenamiento del <i>Autoencoder</i> con hiperparámetros por defecto/Señales Filtradas	56
5.4.3.	Experimento 3: Sintonización de hiperparámetros/Señales filtradas	58
5.4.4.	Experimento 4: Entrenamiento con espectros	59
5.4.5.	Experimento 5: Comparación con otros modelos del estado del arte	61
5.4.6.	Experimento 6: Evaluación del modelo propuesto con otra base de datos	61
5.4.7.	Experimento 7: Entrenamiento del modelo con señales sintéticas y evaluación con señales reales	64
5.4.8.	Experimento 8: Entrenamiento con momentos de tiempo-frecuencia	65
5.5.	Discusión	67
6	Conclusiones	68
6.1.	Objetivos y alcances logrados	68
6.2.	Productos	69
6.3.	Conclusión	70
	Referencias	71
	Anexos	76
	A Productos	76

Indice de Figuras

2.1. Diferentes ondas del ECG [4].	5
2.2. Muestras de las diferentes clases de la base de datos MIT-BIH	7
2.3. Proceso de filtrado	8
2.4. Estimaciones espectrales obtenidas a partir de una muestra de ritmo cardíaco normal	13
2.5. Arquitectura base de un <i>Autoencoder</i>	14
2.6. Ejemplo de <i>Autoencoders</i> apilados.	15
3.1. Modelo de Chauffeur modificado propuesto en [25].	17
3.2. Diagrama de bloques del modelo propuesto en [2].	18
3.3. Modelo de red neuronal propuesto en [28].	18
3.4. Diagrama de flujo del método propuesto en [1].	19
3.5. Diagrama de bloques del proceso de agrupamiento propuesto en [46].	20
3.6. Diagrama de flujo del modelo propuesto en [16].	20
3.7. Esquema de clasificación CardioNet para arritmias cardíacas utilizado en [30].	21
3.8. Modelo híbrido para la clasificación de arritmias cardíacas de 2 derivaciones utilizado en [34].	22
3.9. Arquitectura de un <i>Autoencoder</i> [8].	22
3.10. Diagrama de bloques del sistema propuesto en [27].	23
3.11. Arquitectura de la red propuesta en [47].	23
3.12. Diagrama de la metodología propuesta en [19].	24
3.13. Modelo de red convolucional propuesto en [24].	24
3.14. Modelos propuestos en [48].	25
3.15. Pasos principales de la metodología propuesta en [20].	26
3.16. Diagrama de la metodología propuesta en [31].	26
3.17. Esquema de la metodología propuesta en [39].	27
3.18. Diagrama de flujo del proceso del modelo propuesto en [33].	28
3.19. Diagrama de la metodología propuesta en [13].	29
3.20. Proceso de descomposición y reconstrucción MOWPT empleado en [18].	29
3.21. Framework del modelo propuesto en [29].	30
3.22. Diagrama de la metodología propuesta en [17].	31
3.23. Diagrama de la metodología propuesta en [38].	31
3.24. Diagrama de bloques propuesto en [32].	32
3.25. Diseño metodológico resumido utilizado en [3].	33
3.26. Arquitectura de la red propuesta en [45].	33
3.27. Diagrama esquemático de la estrategia propuesta en [6].	34
3.28. Modelo de la metodología general propuesto en [9].	35
4.1. Metodología de solución propuesta	42
4.2. Arquitectura propuesta para realizar la tarea de clasificación	43
5.1. Histograma de las instancias por clase	52
5.2. Matriz de confusión de la fase de prueba del entrenamiento realizado con señales en crudo	55

5.3. Matriz de confusión de la fase de prueba del entrenamiento realizado con señales filtradas	57
5.4. Matriz de confusión de la fase de prueba del entrenamiento realizado con hiperparámetros sintonizados	58
5.5. Comparación de métricas del experimento 4	60
5.6. Muestras de la base de datos Physionet CinC 2017	62
5.7. Valores estadísticos de validación cruzada para la evaluación del modelo propuesto sobre la base de datos Physionet CinC 2017	63
5.8. Arquitectura de la GAN utilizada	64
5.9. Comparación entre señales sintéticas y originales	65
5.10. Momentos tiempo-frecuencia extraídos a partir de las señales filtradas	66
6.1. Primera página del Artículo 1	76
6.2. Constancia de participación en el congreso WITCOM 2024	77
6.3. Primera página del artículo 2	78
6.4. Primera página del artículo 3	79
6.5. Comprobante de presentación del poster	80

Indice de Tablas

3.1. Resumen del estado del arte.	36
5.1. Resumen cuantitativo de la base de datos	52
5.2. Cantidad de datos para entrenamiento y prueba	52
5.3. Estructura de una matriz de confusión	53
5.4. Métricas del experimento 1	55
5.5. Métricas del experimento 2	57
5.6. Métricas del experimento 3	59
5.7. Métricas del experimento 5	61
5.8. Métricas del experimento 6	63
5.9. Métricas del experimento 7	65
5.10. Métricas del experimento 8	66
6.1. Objetivos específicos	68
6.2. Alcances	69
6.3. Limitaciones	69

1 | Introducción

Las arritmias cardíacas (AC) son alteraciones de la secuencia de contracciones y relajaciones del corazón. Existen varios tipos dependiendo de la condición en que ocurren; entre las más comunes se encuentran: fibrilación auricular (FA), fibrilación ventricular (FV), flutter auricular (FAR), taquicardias supraventriculares paroxísticas (TSVP), bloqueo aurículoventricular (BA), entre otras. Las AC representan una de las principales causas de muerte a nivel mundial según la OMS. La prevalencia de estas anomalías es relativamente alta: cerca del 3 % de la población mundial padece esta condición [14].

El interés por la clasificación automática de las arritmias cardíacas crece significativamente debido a la gravedad de esta condición. Partiendo de lo anterior, en el área de aprendizaje automático se desarrollan técnicas de detección para el apoyo a los especialistas. Estas técnicas se basan en redes neuronales convolucionales (CNN) [42], redes neuronales recurrentes (RNN) [43], máquinas de soporte vectorial (SVM) [21], entre otras. Sin embargo, muchas de las técnicas mencionadas requieren una gran cantidad de datos de entrenamiento para que, en aplicaciones reales, alcancen un rendimiento deseable. Esto implica que el entrenamiento de estos modelos consuma tanto tiempo como recursos computacionales.

La motivación principal de esta investigación radica en la necesidad de mejorar la eficiencia en la clasificación automática de arritmias cardíacas. Los algoritmos de aprendizaje profundo, en particular los *Autoencoders*, demuestran ser herramientas efectivas para la extracción de características relevantes de datos complejos [12]. Este trabajo propone un enfoque innovador basado en un *Autoencoder* profundo modificado, entrenado con características espectrales derivadas de señales ECG. La metodología incluye la transformación de las señales ECG del dominio del tiempo al dominio de la frecuencia mediante diversas técnicas de estimación espectral. Estas transformaciones facilitan una interpretación más precisa de los datos y permiten la extracción de características clave, esenciales para la clasificación de las señales.

Una vez obtenidas las características espectrales, se modifica la arquitectura del *Autoencoder* tradicional para integrar una capa completamente conectada en su núcleo. Esta modificación permite al *Autoencoder* no solo reducir la dimensionalidad de los datos, sino también clasificar las señales ECG en distintas categorías de arritmias. La capa completamente conectada utiliza las características extraídas durante la fase de codificación para realizar la tarea de clasificación con mayor precisión. La relevancia de esta investigación radica en su potencial para mejorar la práctica clínica, al proporcionar herramientas avanzadas para el diagnóstico de arritmias a los especialistas. La combinación de técnicas de estimación espectral con algoritmos de aprendizaje profundo constituye un enfoque innovador que aporta significativamente al campo de la cardiología. Además, la reducción del costo computacional del modelo facilita su implementación en sistemas embebidos de bajo a medio costo, ampliando su aplicabilidad.

En resumen, la propuesta de este trabajo se centra en el diseño de una arquitectura de bajo costo computacional que sea competitiva en comparación con otras técnicas reportadas en la literatura. Para ello, se explora el uso de *Autoencoders*, redes especializadas en la abstracción de características que, en su configuración estándar, no permiten realizar clasificación. No obstante, al incorporar una capa totalmente conectada en su estructura central, es posible extender su funcionalidad para esta tarea. Adicionalmente, el modelo se entrena utilizando estimaciones espectrales de las señales ECG, las cuales contienen las características más relevantes de la señal original. Con este enfoque, se busca obtener resultados eficientes en la clasificación de arritmias mediante una arquitectura optimizada y de menor costo computacional en comparación con modelos del estado del arte. Las principales

aportaciones de este trabajo son:

- Un enfoque híbrido basado en un *Autoencoder* con una capa totalmente conectada apilada para la clasificación de señales ECG.
- Demostración de la efectividad del dominio espectral de señales ECG al extraer características relevantes conduciendo a una mejora en el rendimiento de la clasificación.
- La compresión de características de datos a una forma abstracta para reducir la dimensionalidad de datos y el costo computacional de un modelo clasificador a la hora de realizar los respectivos análisis.

1.1. Planteamiento del problema

Existen trabajos en el estado del arte que ya realizan la tarea de clasificación de arritmias cardíacas, sin embargo, varios de estos modelos utilizan tipos de datos que son complicados analizar, como por ejemplo, imágenes en el caso de una CNN. En el antecedente de este trabajo [2], se buscó mejorar el rendimiento de estos modelos al utilizar una red neuronal recurrente para realizar una clasificación binaria. Como este modelo se especializa en analizar señales de tiempo, para entrenarlo se tuvieron que realizar dos transformaciones a las señales: Primero, la obtención de un espectrograma a partir de señales en crudo, y después, la conversión de ese espectrograma a series de tiempo-frecuencia, ya que el modelo propuesto es incapaz de analizar imágenes. El objetivo del trabajo se cumplió, ofreciendo mejores resultados, sin embargo, el modelo es computacionalmente costoso al ser entrenado con series de tiempo.

En este trabajo se propone otro enfoque: el utilizar un modelo que obtenga resultados deseados esperando tener menor costo computacional, utilizando solo una transformación de las señales y realizando una clasificación multiclase. Esto se puede lograr al tratar las señales en crudo de otra manera, ya que, al ser de naturaleza caótica, es necesario aplicar transformaciones para extraer las características más significativas. Partiendo de lo anterior, teniendo una señal en crudo, se obtendrá su estimación espectral, la cuál tendrá las características más significativas. Posteriormente se utilizará para entrenar un *Autoencoder* profundo modificado, una red neuronal más simple que será capaz de realizar una clasificación multiclase.

Hipótesis: Utilizar estimaciones espectrales para entrenar un *Autoencoder* profundo modificado con una capa apilada totalmente conectada, podrá realizar una clasificación multiclase entre distintos tipos de arritmias y un ritmo cardíaco normal, mejorando el rendimiento del modelo y teniendo menor costo computacional que otras técnicas en el estado del arte.

1.1.1. Delimitación del problema

En este trabajo de investigación, se desarrollará el reconocimiento del ritmo cardíaco normal y de cuatro tipos de arritmias cardíacas, condición que hace que el ritmo cardíaco se altere, esto con el objetivo de determinar si los datos procesados corresponden a un latido normal o a uno anormal causado por dicha condición. Dentro de los modelos de aprendizaje profundo que se encuentran en la literatura, se hará una incursión exploratoria al usar un *Autoencoder* profundo para realizar el reconocimiento de arritmias cardíacas, modificando su arquitectura original al apilar una capa totalmente conectada tipo *softmax* para que pueda realizar tareas de clasificación. El *Autoencoder*

se entrenará con 6 tipos de estimaciones espectrales para realizar una comparación entre ellas, las cuáles se obtendrán al transformar señales en crudo de un ECG.

1.1.2. Complejidad del problema

La complejidad radica en que las señales ECG en crudo son de difícil análisis debido a su naturaleza caótica. Para entrenar modelos computacionales, es necesario que estos extraigan características significativas de los datos de entrada. Esto se logra al transformar los datos originales a un dominio que sea compatible con el modelo a entrenar y que a este se le haga más fácil analizar dichas características, un ejemplo de ello es obtener la estimación espectral de las señales en crudo.

Este tipo de transformaciones permiten realizar una reducción de dimensionalidad, conservando las características más significativas de la señal. Por otra parte, los *Autoencoders* son redes neuronales diseñadas para reconstruir datos y extraer características, por lo tanto no tienen la capacidad de realizar la tarea de clasificación. Para ello, es necesario modificar la arquitectura original de un *Autoencoder* para que sea capaz de realizar una clasificación.

1.2. Objetivos

1.2.1. Objetivo general

Diseñar un modelo capaz de identificar entre arritmias cardíacas y un ritmo cardíaco normal utilizando una arquitectura modificada de un *Autoencoder* profundo entrenado con estimaciones espectrales obtenidas a partir señales ECG en crudo.

1.2.2. Objetivos específicos

- Analizar y comprender las señales ECG, explorando las diversas características que se pueden extraer de ellas.
- Profundizar en los conceptos relacionados con la estimación espectral. Además, se abordará la arquitectura de los *Autoencoders*, con especial atención al concepto de apilamiento.
- Rediseñar la estructura de un *Autoencoder* mediante la incorporación de una capa totalmente conectada para realizar la tarea de clasificación.
- Evaluar el rendimiento del modelo desarrollado utilizando métricas mencionadas en el estado del arte, así como aquellas que se desarrollen a lo largo de esta investigación. Este análisis incluirá la consideración de los diferentes espectros previamente mencionados.
- Desarrollar un sistema computacional que integre todos los módulos del procesamiento de señales ECG y el modelo diseñado.

1.3. Alcances y limitaciones

1.3.1. Alcances

- Se considerarán los espectros *AR-Burg*, *Covariance*, *Yule-walker*, *Fast Fourier Transform*, *Periodogram* y *Welch*.

- Se hará una comparación entre distintas estimaciones espectrales.
- Se considerará el dominio espectral de las señales.
- Se buscará mantener un costo computacional bajo con el modelo propuesto.
- Se buscará obtener un rendimiento alto en la clasificación multiclase.

1.3.2. Limitaciones

- Se utilizarán las bases de datos mencionadas en este documento, así como alguna otra que se encuentre en el desarrollo de esta investigación que contenga señales de ECG de arritmias cardíacas.
- Se utilizarán únicamente estimaciones espectrales de las señales.
- Se centrará únicamente en el reconocimiento de 4 tipos de arritmias y 1 ritmo cardíaco normal.

1.4. Justificación

Una arritmia cardíaca se caracteriza por tener un comportamiento irregular del corazón que afecta el ritmo y duración de la señal, por lo que cada arritmia es clasificable según su comportamiento. La prevalencia de esta condición es relativamente alta, aproximadamente un 3% de la población mundial la padece. Existen trabajos previos en los cuáles se obtienen resultados deseables, sin embargo, proporcionan modelos muy pesados que necesitan de equipos de rendimiento alto para poder ser implementados. Para apoyar al sector médico, es necesario desarrollar una solución que permita ejecutar estos sistemas en equipos de rendimiento medio.

A partir de lo anterior, en este trabajo se propone una arquitectura de bajo costo computacional capaz de detectar entre diferentes tipos de arritmias cardíacas, además del ritmo normal del corazón. Para obtener resultados competentes en comparación a otras técnicas, se propone transformar las señales a otro dominio en donde la extracción de características significativas sea más fácil.

1.5. Organización de la tesis

El resto de esta tesis está organizada de la siguiente manera: El Capítulo 2 contiene el marco teórico, dentro de este se encuentran los conceptos básicos que se tomaron en cuenta para el desarrollo de este trabajo, incluyendo las técnicas de preprocesamiento empleadas y una breve descripción de la arquitectura del *Autoencoder*. En el Capítulo 3 se realiza una revisión del estado del arte contemplando principalmente tres núcleos temáticos: trabajos que usan enfoques similares, trabajos que realizan clasificación ECG y trabajos referentes de análisis de señales. Dentro del Capítulo 4 se detalla la metodología propuesta para el desarrollo de este trabajo, incluyendo la formalización matemática del funcionamiento del modelo propuesto. El Capítulo 5 detalla la experimentación realizada y los resultados obtenidos. Finalmente, en el Capítulo 6 se presentan las conclusiones obtenidas, el trabajo futuro considerado y los productos obtenidos. Adicionalmente se incluyen referencias y anexos.

2 | Marco teórico

En esta sección se presentan los conceptos que se tomaron en cuenta para el desarrollo de este trabajo. Se ofrece una visión clara y detallada acerca de los datos, herramientas y métodos utilizados para el desarrollo de esta investigación.

2.1. Electrocardiograma

Un electrocardiograma (ECG) es un gráfico en el que se estudian las variaciones de voltaje en relación con el tiempo. Es una prueba diagnóstica asequible, segura y sencilla de realizar, que proporciona una gran cantidad de información con relación al estado del corazón. Consiste en registrar en un formato especialmente adaptado la actividad de la corriente eléctrica que se está desarrollando en el corazón durante un tiempo determinado. La actividad eléctrica del corazón recogida por el ECG se observa en forma de un trazado que presenta diferentes deflexiones que se corresponden con el recorrido de los impulsos eléctricos a través de las diferentes estructuras del corazón [10].

Fundamentalmente se utiliza para detectar trastornos del ritmo cardíaco y en el diagnóstico de las situaciones que cursan con un aporte insuficiente de sangre al corazón. Permite diferenciar el ritmo normal del corazón de cualquier tipo de arritmia cardíaca, condición que hace que el corazón lata a una frecuencia anormal. El ECG presenta una línea guía llamada línea isoelectrica o basal, que se puede identificar fácilmente como la línea horizontal existente entre cada latido. Los latidos cardíacos quedan representados como las diferentes oscilaciones de la línea basal en forma de ángulos, segmentos, ondas e intervalos, constituyendo una imagen característica que se repite con una frecuencia regular a lo largo del tiempo. En un ECG normal, cada complejo consta de una serie de deflexiones que se alternan con la línea basal [36].

Realizando la lectura de izquierda a derecha, se distinguen la onda *P*, la cuál representa el momento en que las aurículas se están contrayendo y enviando sangre hacia los ventrículos; el segmento *P – R*, es el momento en que las aurículas terminan de vaciarse y se produce una relativa desaceleración en la transmisión de la corriente eléctrica a través del corazón; el complejo *QRS*, representa el momento en que los ventrículos se contraen y expulsan su contenido sanguíneo; el segmento *ST*, su elevación o descenso en relación con la línea basal puede significar insuficiencia en el riego del corazón y finalmente la onda *T*, representa el momento en que el corazón se encuentra en un período de relajación, una vez que ha expulsado la sangre que se hallaba en los ventrículos. En la Figura 2.1 se puede

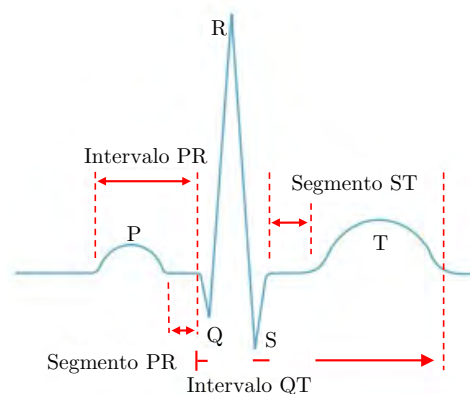


Figura 2.1: Diferentes ondas del ECG [4].

Como se mencionó anteriormente, el ECG se utiliza para monitorear el ritmo cardíaco en el cuál se espera que este ultimo sea normal. Sin embargo existen ocasiones en que el ritmo cardíaco es alterado por diferentes causas, dando lugar a condiciones cardíacas a las que se les conoce como arritmias cardíacas.

2.2. Arritmias cardíacas

Las arritmias cardíacas son alteraciones en la frecuencia o el ritmo del latido del corazón, las cuales pueden manifestarse como latidos demasiado rápidos (taquicardia), demasiado lentos (bradicardia) o irregulares. Estas alteraciones pueden originarse por múltiples causas, incluyendo anomalías estructurales del corazón, desequilibrios electrolíticos, isquemia miocárdica o alteraciones en el sistema de conducción eléctrica del corazón [22].

El sistema de conducción eléctrica del corazón está compuesto por el nodo sinoauricular (SA), el nodo auriculoventricular (AV), el Haz de His y las fibras de Purkinje, los cuales trabajan en conjunto para coordinar la contracción del músculo cardíaco [37]. Cuando se produce una alteración en este sistema, la conducción del impulso eléctrico puede verse afectada, generando arritmias de diversa naturaleza.

Su diagnóstico de estas condiciones cardiacas es fundamental en la práctica clínica, ya que pueden ser asintomáticas o manifestarse con síntomas que van desde palpitaciones leves hasta complicaciones severas como síncope o muerte súbita cardíaca [12]. Para este trabajo, se destacan cuatro tipos específicos de arritmias de interés clínico: la contracción prematura auricular (APC), la contracción prematura ventricular (PVC), el bloqueo de rama izquierda (LBBB) y el bloqueo de rama derecha (RBBB).

2.2.1. Contracción Prematura Auricular

La contracción prematura auricular es un tipo de extrasístole que se origina en las aurículas antes de que ocurra el latido cardíaco normal. Se produce cuando un foco ectópico dentro de las aurículas genera un impulso antes de la activación normal del nodo SA, provocando una despolarización anticipada. Esta condición es relativamente común y, en la mayoría de los casos, benigna, aunque puede estar relacionada con patologías subyacentes.

2.2.2. Contracción Prematura Ventricular

Las contracciones prematuras ventriculares son impulsos ectópicos originados en los ventrículos antes de que ocurra el siguiente latido normal. A diferencia de las APCs, en las PVCs la despolarización ventricular se produce sin el estímulo del nodo AV, lo que genera un latido prematuro con una morfología anormal.

2.2.3. Bloqueo de Rama Izquierda

El bloqueo de rama izquierda se produce cuando la conducción del impulso eléctrico a través de la rama izquierda del Haz de His está retardada o bloqueada, lo que provoca una despolarización ventricular asincrónica. Esto se traduce en una activación retardada del ventrículo izquierdo con respecto al derecho.

2.2.4. Bloqueo de Rama Derecha

El RBBB puede ser un hallazgo incidental en pacientes asintomáticos, pero también puede ser un marcador de disfunción del miocardio y mayor riesgo de insuficiencia cardíaca. En pacientes con síntomas, puede requerir estudio con ecocardiografía y pruebas de perfusión miocárdica.

El estudio de las arritmias cardíacas requiere bases de datos confiables para su clasificación. En este trabajo, se utiliza la base de datos MIT-BIH por su validación en estudios previos. A continuación, se describen sus características y la estrategia de uso en el modelo propuesto.

2.3. Dataset

Los datos empleados en el desarrollo de este trabajo provienen de la base de datos MIT-BIH, disponible en el repositorio de PhysioNet [44]. Esta base de datos incluye muestras de las siguientes clases: ritmo cardíaco normal, contracción prematura ventricular, contracción prematura auricular, bloqueo de rama izquierda y bloqueo de rama derecha. Las señales están registradas a una frecuencia de 360 Hz, lo que equivale a 360 características por segundo. Se seleccionó esta base de datos debido a su uso frecuente en la literatura, lo que proporciona una línea base confiable para comparar los resultados obtenidos. La Figura 2.2 ilustra la representación gráfica de las diferentes muestras.

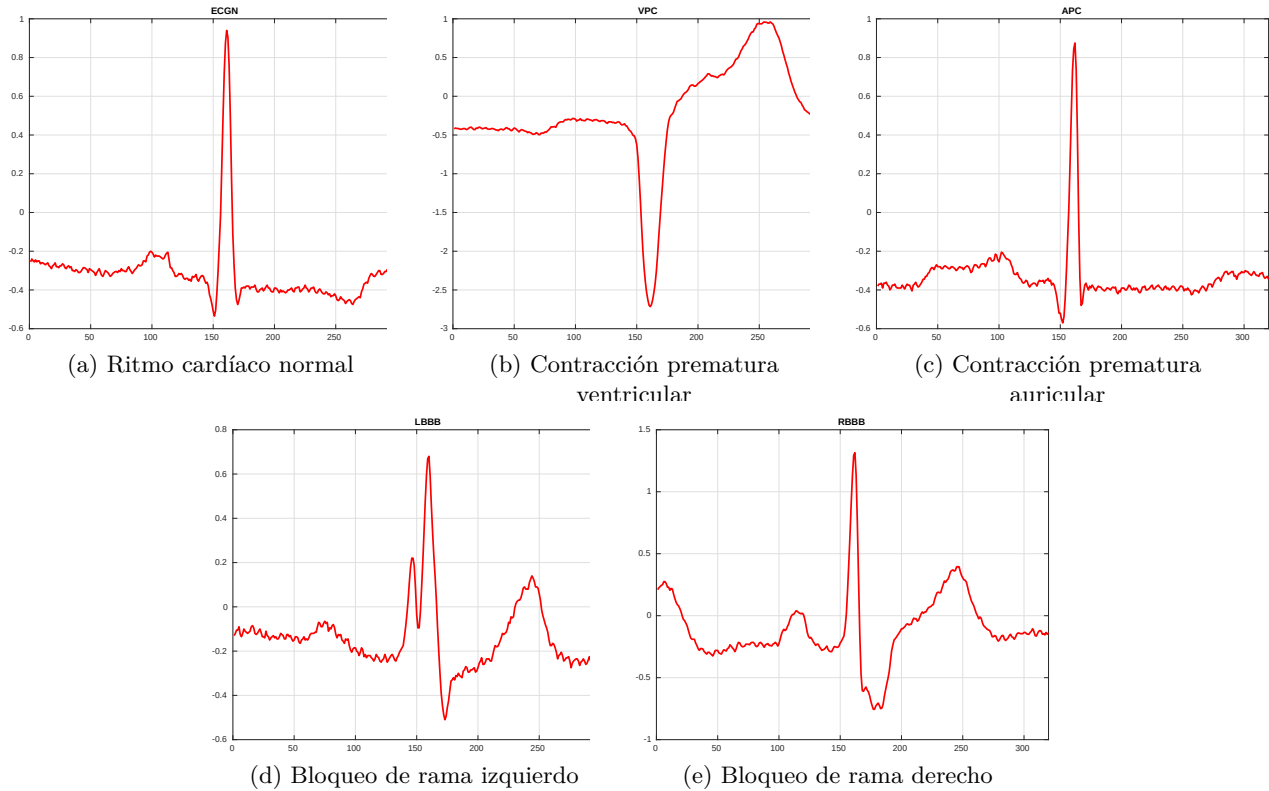


Figura 2.2: Muestras de las diferentes clases de la base de datos MIT-BIH

Las señales ECG suelen estar contaminadas por ruido debido a diversos factores, como movimientos musculares e interferencias eléctricas. Por esta razón, es necesario, en ocasiones, aplicar filtros digitales para eliminar dichas interferencias. En este trabajo, y mediante técnicas de análisis de señales, se decidió emplear filtros pasa-bajos, dado que las principales interferencias se encuentran en frecuencias altas.

2.4. Filtros Pasa-bajo

Los filtros pasa-bajo permiten el paso de señales de baja frecuencia mientras atenúan las frecuencias más altas. Son herramientas fundamentales en el procesamiento de señales, ya que eliminan el ruido de alta frecuencia, suavizan las fluctuaciones y preservan los componentes esenciales de baja frecuencia [15]. El filtro pasa-bajos se define matemáticamente mediante la siguiente ecuación:

$$\varphi[n] = \alpha \cdot x[n] + (1 - \alpha) \cdot y[n - 1] , \quad (2.1)$$

donde $\varphi[n]$ es la señal filtrada en el tiempo n , $x[n]$ es la señal de entrada en el tiempo n , $y[n - 1]$ es la señal filtrada en el tiempo $n - 1$ y α es el factor de suavizado, que determina la frecuencia de corte del filtro y se calcula mediante:

$$\alpha = \frac{2\pi f_c}{2\pi f_c + f_s} , \quad (2.2)$$

donde f_c es la frecuencia de corte y f_s es la frecuencia de muestreo de la señal original. Utilizando herramientas como la transformada rápida de Fourier, se aplicó un filtro con una frecuencia de corte de 35 Hz para eliminar el ruido de alta frecuencia y suavizar las señales. Este proceso se muestra en la Figura 6.3 utilizando una muestra de un ritmo cardíaco normal como referencia.

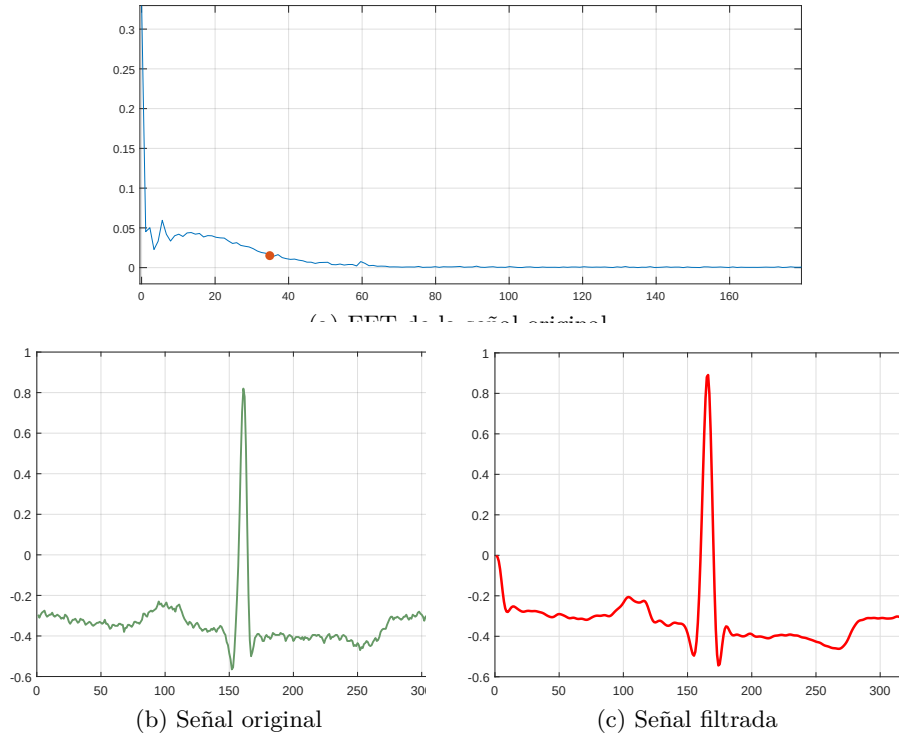


Figura 2.3: Proceso de filtrado

Como se puede observar, la señal filtrada está mas suavizada, evidencia de la eliminación de ruido. Una vez que las señales fueron filtradas, pueden aplicarse otros métodos de preprocesamiento para la extracción de características.

2.5. Estimaciones Espectrales

Para poder mejorar la calidad de los resultados en la clasificación, es necesario aplicar un pre-procesamiento a las señales ECG para la extracción de las características más relevantes. Para ello, se optó por utilizar los siguientes espectros:

2.5.1. *FFT*

La Transformada Rápida de Fourier (FFT) es un algoritmo eficiente para calcular la Transformada Discreta de Fourier (DFT) de una secuencia de datos, que es una herramienta fundamental en el análisis espectral y el procesamiento de señales. Su objetivo es convertir una señal de su representación en el dominio del tiempo al dominio de la frecuencia, revelando las frecuencias que componen la señal y sus amplitudes. Divide recursivamente la secuencia de datos en subconjuntos más pequeños, calcula las DFT de estos subconjuntos y luego combina los resultados para obtener la DFT completa [40] [26] [35]. El espectro FFT de una señal está definido por:

$$X(n) = \frac{1}{N} \sum_{m=0}^{N-1} X(m) \cdot e^{(j2\pi mn/N)}, \quad (2.3)$$

donde $X(m)$ es el coeficiente k th del DFT, que representa la amplitud y fase del componente de frecuencia k , $x(n)$ es el valor de la señal de ECG en el tiempo discreto n y N es la duración de la señal del ECG.

2.5.2. *Periodogram*

Es un algoritmo para estimar la densidad espectral de potencia (PSD) de una señal calculando el cuadrado de la magnitud de la transformada de Fourier de la señal. El espectro del periodograma proporciona información sobre cómo se distribuye la energía de la señal en el dominio de la frecuencia [11]. El periodograma proporciona un vínculo crítico entre las cantidades que se pueden medir directamente desde la señal de longitud finita y aquellas que no se pueden medir directamente. La resolución espectral del espectro del periodograma está determinada por la longitud de la señal y la frecuencia de muestreo [40]. La PSD usando periodograma $\wp(f)$ se obtiene mediante:

$$\wp(f) = \int_0^T \chi(\tau) \cdot e^{-j2\pi f\tau} d\tau, \quad (2.4)$$

en donde $\chi(\tau)$ es la función de auto-correlación. Dado que la energía de una señal analógica $x(t)$ es igual a la magnitud de la señal, se puede utilizar un método directo utilizando una integración en el tiempo definida por:

$$\mathcal{E} = \int_{-\infty}^{\infty} |x(t)|^2 dt. \quad (2.5)$$

Entonces, teniendo en cuenta el teorema de Parseval, se puede suponer lo siguiente:

$$\mathcal{E} = \int_{-\infty}^{\infty} |X(f)|^2 df, \quad (2.6)$$

donde $|X(f)|^2$ representa la función de densidad de energía sobre la frecuencia f . Luego, la PSD se calcula como la magnitud al cuadrado de FFT de la forma de onda de interés $x(t)$:

$$\wp_{Per}(f) = |X(f)|^2. \quad (2.7)$$

2.5.3. Welch

El método Welch es una técnica utilizada para estimar la densidad espectral de potencia de una señal a partir de su función de auto-correlación. Es una mejora del método del periodograma que reduce la varianza de la estimación de PSD al promediar los segmentos superpuestos de la señal en lugar de calcular el periodograma directamente sobre toda la señal. El método de Welch divide la señal en segmentos superpuestos de longitud fija [15]. Antes de calcular la Transformada de Fourier de cada segmento, se aplica una ventana a cada uno para minimizar el efecto de las discontinuidades en los extremos de los segmentos. Los periodogramas de los segmentos se promedian para obtener una estimación final de la PSD. El promedio reduce la varianza de la estimación de PSD, lo que da como resultado un espectro más suave y menos propenso a variaciones aleatorias. El método Welch es útil para señales que varían en el tiempo o para señales ruidosas, ya que promediar los segmentos ayuda a reducir el efecto del ruido y proporciona una estimación de PSD más confiable y fluida [41]. El espectro de Welch se estima mediante:

$$\wp_{Welch}(\omega, L, D, M) = \frac{\frac{1}{M} \sum_{m=0}^{M-1} \left(\left| \sum_{n=0}^{L-1} w_L[n] \cdot x[n + m \cdot D] \cdot e^{-j\omega n} \right|^2 \right)}{\sum_{n=0}^{L-1} w_L^2[n]}, \quad (2.8)$$

donde $x[n]$ denota la señal original en tiempo discreto que se desea analizar; ω corresponde a la frecuencia angular, expresada en radianes por muestra, en la cual se evalúa la densidad espectral, L representa la longitud de cada segmento en el que se divide la señal, D indica el solapamiento entre segmentos consecutivos, M es el número total de segmentos en que se ha particionado la señal para realizar el proceso de promediado, $w_L[n]$ corresponde a la ventana de análisis aplicada a cada segmento de longitud L , el término $\sum_{n=0}^{L-1} w_L^2[n]$ actúa como un factor de normalización que compensa la energía introducida por la aplicación de la ventana y la expresión $\sum_{n=0}^{L-1} w_L[n] \cdot x[n + m \cdot D] \cdot e^{-j\omega n}$ representa la transformada discreta de Fourier (DFT) aplicada al segmento m -ésimo de la señal, una vez ponderado por la ventana correspondiente. El proceso completo de suma y promediado sobre los M segmentos permite obtener una estimación suavizada y de menor varianza de la densidad espectral de potencia de la señal analizada.

2.5.4. AR Burg

Es conocido por estimar la densidad espectral de potencia autoregresiva (PSD) minimizando los errores de predicción hacia adelante y hacia atrás y estimando el coeficiente de reflexión [41]. Los errores de predicción hacia adelante y hacia atrás se calculan con:

$$\hat{\Psi}_{f,o}(n) = x(n) + \sum_{i=1}^o \hat{\Omega}_{o,i} \cdot x(n-i) \quad \forall n = o+1, \dots, N, \quad (2.9)$$

$$\hat{\Psi}_{b,o}(n) = x(n-o) + \sum_{i=1}^o \hat{\Omega}_{o,i} \cdot x(n-o+i) \quad \forall n = o+1, \dots, N. \quad (2.10)$$

Los parámetros autoregresivos se calculan mediante:

$$\hat{\Omega}_{o,i} = \begin{cases} \hat{\Omega}_{o-1,i} + \hat{\kappa}_o \cdot \hat{\Omega}_{o-1,o-i} & i = 1, \dots, o-1 \\ \hat{\kappa}_o & i = o. \end{cases} \quad (2.11)$$

El método de Burg refleja la estimación recursiva en orden de k_o dado que se han calculado los coeficientes AR para el orden $p-1$. El coeficiente de reflexión se estima mediante:

$$\hat{\kappa}_o = \frac{-2 \sum_{n=o+1}^N \hat{\Psi}_{f,o-1}(n) \cdot \hat{\Psi}_{b,o-1}(n-1)}{\sum_{n=o+1}^N \left[\left| \hat{\Psi}_{f,o-1}(n) \right|^2 + \left| \hat{\Psi}_{b,o-1}(n-1) \right|^2 \right]}. \quad (2.12)$$

Los errores de predicción en las ecuaciones 2.9 y 2.10 satisfacen las siguientes expresiones recursivas:

$$\hat{\Psi}_{f,o}(n) = \hat{\Psi}_{f,o-1}(n) + \hat{\kappa}_o \cdot \hat{\Psi}_{b,o-1}(n-1), \quad (2.13)$$

$$\hat{\Psi}_{b,o}(n) = \hat{\Psi}_{b,o-1}(n-1) + \hat{\kappa}_o \cdot \hat{\Psi}_{f,o-1}(n). \quad (2.14)$$

Finalmente, la PSD obtenida por este modelo se define como:

$$\hat{\phi}_{Burg}(f) = \frac{\hat{\Psi}_{f,o} + \hat{\Psi}_{b,o}}{\left| 1 + \sum_{l=1}^o \hat{\Omega}_o(l) \cdot e^{-j2\pi fl} \right|^2}. \quad (2.15)$$

2.5.5. Yule-Walker

En este método, los parámetros autoregresivos (AR) se calculan minimizando una estimación de la potencia del error de predicción. Su objetivo es el mismo que AR Burg, encontrar los coeficientes del modelo AR que mejor se ajustan a una señal de entrada, minimizando el error cuadrático medio. Una vez aplicado el algoritmo se obtienen los coeficientes del modelo AR que definen el comportamiento de la señal de entrada en el dominio de la frecuencia [52] [41]. Los parámetros AR son calculados por:

$$\hat{\Phi} = \frac{1}{N} \sum_{n=-\infty}^{\infty} \left| x(n) + \sum_{l=1}^o \gamma(l) \cdot x(n-l) \right|^2. \quad (2.16)$$

A partir de 2.16, el gradiente complejo puede ayudar a minimizar el poder de error de predicción estimado mediante la diferenciación, este proceso está definido por:

$$\frac{1}{N} \sum_{n=-\infty}^{\infty} \left(x(n) + \sum_{l=1}^o \gamma(l) \cdot x(n-l) \right) \cdot x(n-i) = 0 \quad i = 0, \dots, o. \quad (2.17)$$

El proceso anterior puede ser representado como una matriz por:

$$\hat{r}_o + \hat{R}_o \cdot \hat{\gamma} = 0, \quad (2.18)$$

en donde

$$\hat{r}(l) = \begin{cases} \frac{1}{N} \sum_{n=0}^{N-1-l} x(n) \cdot x(n+l) & l = 0, 1, \dots, o \\ \hat{r}(-l) & l = (-o+1), (-o+2), \dots, -1 \end{cases}. \quad (2.19)$$

A partir de la representación matricial, los parámetros AR se obtienen mediante:

$$\hat{\gamma} = -\frac{\hat{r}_o}{\hat{R}_o}. \quad (2.20)$$

Luego, la varianza del ruido blanco $\hat{\Theta}^2$ se encuentra como $\hat{o}_{min}$, que viene dada por:

$$\hat{\Theta}^2 = \hat{\sigma}_{min} = \frac{1}{N} \sum_{n=-\infty}^{\infty} \left| x(n) + \sum_{l=1}^o \hat{a}(l)x(n-l) \right|^2. \quad (2.21)$$

El resultado final de 2.16 es:

$$\hat{\Theta}^2 = \hat{r}(0) + \sum_{l=1}^o \hat{\gamma}(l) \cdot \hat{r}(-l). \quad (2.22)$$

Y finalmente, a partir de 2.21 y 2.22, la PSD para Yule-Walker se estima mediante:

$$\hat{\phi}_{YW}(f) = \frac{\hat{\Theta}^2}{|1 + \sum_{l=1}^o \hat{\gamma}(l) \cdot e^{-j2\pi fl}|^2}. \quad (2.23)$$

2.5.6. *Covariance*

El espectro de covarianza es una técnica utilizada en el análisis espectral, su objetivo es estimar la densidad espectral de potencia de una señal a partir de su matriz de covarianza, permitiendo analizar cómo se distribuye la energía de la señal en el dominio de la frecuencia y comprender mejor su contenido espectral [41]. Para las series temporales complejas, la potencia del error de predicción se estima mediante:

$$\hat{\Phi} = \frac{1}{N-o} \sum_{n=o}^{N-1} \left| x(n) + \sum_{l=1}^o \gamma(l) \cdot x(n-l) \right|^2. \quad (2.24)$$

A partir de lo anterior, la predicción del error se puede minimizar utilizando el gradiente complejo; este proceso está definido por:

$$c_o + C_o \cdot \hat{\gamma} = 0, \quad (2.25)$$

en donde

$$c(j, l) = \frac{1}{N-o} \sum_{n=o}^{N-1} x(n-j) \cdot x(n-l). \quad (2.26)$$

A partir de 2.25, las estimaciones de los parámetros AR se calculan mediante:

$$\hat{\gamma} = -\frac{c_o}{C_o}. \quad (2.27)$$

La varianza del ruido blanco se estima como:

$$\hat{\Theta}^2 = c(0, 0) + \sum_{l=1}^o \hat{\gamma}(l) \cdot c(0, l). \quad (2.28)$$

Y luego, a partir de 2.27 y 2.28 la PSD se puede estimar mediante:

$$\hat{\phi}_{Cov}(f) = \frac{\hat{\Theta}^2}{|1 + \sum_{l=1}^o \hat{\gamma}(l) \cdot e^{-j2\pi fl}|^2}. \quad (2.29)$$

donde $\hat{\Theta}^2$ es la varianza del ruido blanco, o es el orden del modelo y $\hat{\gamma}_o$ son los parámetros AR del espectro de covarianza.

Para un mayor entendimiento, la Figura 2.3 muestra una representación gráfica de cada espectro mencionado obtenidos a partir de una muestra ECG de ritmo cardíaco normal.

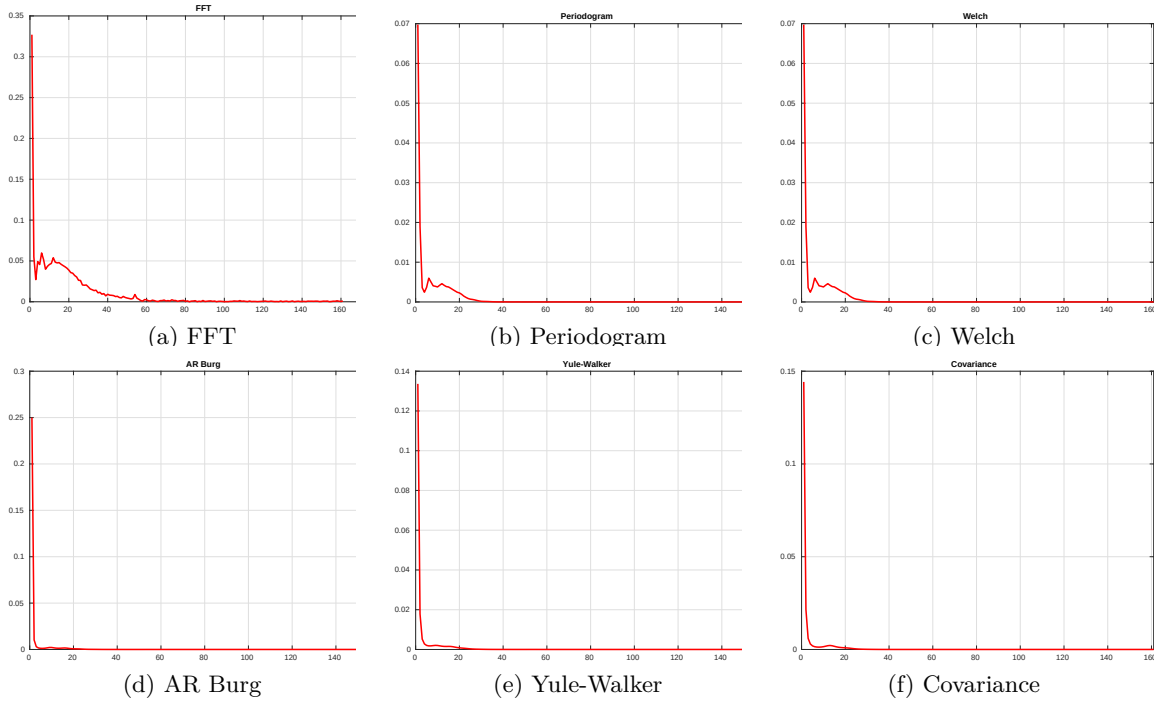


Figura 2.4: Estimaciones espectrales obtenidas a partir de una muestra de ritmo cardíaco normal

Como se puede observar, cada una de estas técnicas proporciona una perspectiva única sobre la distribución de frecuencias en las señales ECG, ofreciendo información valiosa para la extracción de características y la clasificación de arritmias. La transformación al dominio de la frecuencia constituye un paso fundamental en el preprocesamiento de las señales ECG, ya que permite resaltar las características más relevantes para el diagnóstico automático de arritmias cardíacas. En este contexto, el eje x representa las frecuencias (Hz) que componen la señal original, mientras que el eje y indica la magnitud de cada una de las frecuencias detectadas. Una vez preprocesados, los datos pueden ser empleados para entrenar algoritmos de aprendizaje profundo, optimizando el desempeño de los modelos y los resultados obtenidos.

2.5.7. Aprendizaje profundo

El aprendizaje profundo es una derivación del *Machine Learning* que usa una estructura jerárquica de redes neuronales artificiales, que se constituyen de una forma similar a la estructura neuronal del cerebro humano. Este paradigma permite abordar el análisis de datos de forma no lineal. El aprendizaje profundo puede aprender a realizar tareas de clasificación directamente desde imágenes, texto o sonido. Los modelos se entrenan utilizando un gran conjunto de datos etiquetados y arquitecturas de redes neuronales que contienen muchas capas.

Estos modelos ayudan a que los productos electrónicos cumplan con las expectativas del usuario, y que es crucial para las aplicaciones críticas para la seguridad. El concepto de aprendizaje profundo se teorizó por primera vez en la década de 1980, sin embargo, en ese entonces no se contaba con la tecnología suficiente para expresar su potencial, ahora está tomando mucha más importancia

debido a que:

- Requiere grandes cantidades de datos etiquetados. Tareas como la auto-conducción pueden requerir millones de datos. Mientras más material se tenga, mejores resultados se obtendrán.
- Requiere una gran potencia de cálculo. Esto incluye componentes computacionales especializados, como lo son las unidades de procesamiento gráfico (GPU), dentro de estos componentes se encuentran los que tienen arquitecturas paralelas que son eficientes para llevar a cabo técnicas de aprendizaje profundo [23].

Autoencoders

Los *Autoencoders* conforman la base de este trabajo de investigación. Por lo tanto, en este apartado se abordara su estructura, arquitectura y funcionamiento, así como cualquier concepto importante para el desarrollo del tema.

Los *Autoencoders* son redes neuronales de propagación hacia adelante y no recurrentes, pertenecientes al paradigma del aprendizaje no supervisado. El objetivo principal de los *Autoencoders* es reducir la dimensionalidad de los datos y minimizar las diferencias entre los datos codificados y decodificados [51]. Con base en lo anterior, la arquitectura utilizada en este trabajo se fundamenta en un *Autoencoder*, al cual se le realizarán las configuraciones necesarias para obtener un modelo capaz de llevar a cabo la tarea de clasificación. La Figura 2.5 muestra la arquitectura base de un *Autoencoder*.

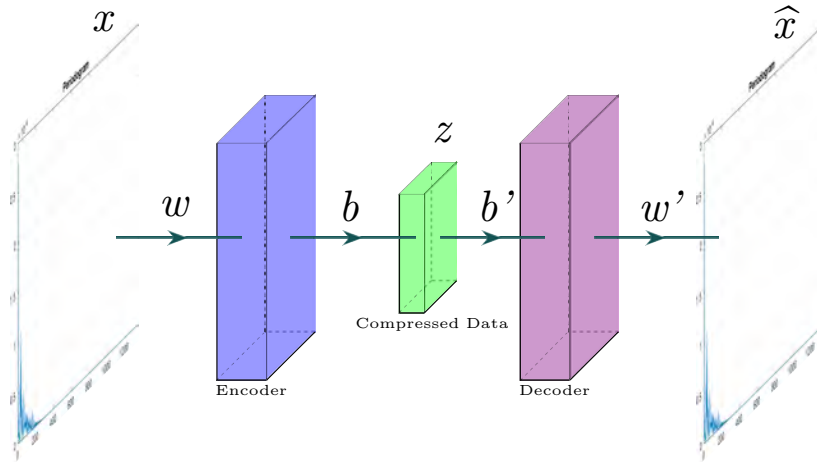


Figura 2.5: Arquitectura base de un *Autoencoder*

Típicamente los *Autoencoders* son simétricos con respecto a la capa central y se componen de 2 partes:

- Desde la capa de entrada hasta la capa central, ambas incluidas, se les denomina codificador.
- Desde la capa central, hasta la capa de salida, se les denomina decodificador.

Partiendo de lo anterior, desde un punto de vista matemático, el proceso de codificación y decodificación está definido por:

$$z = \sigma(\mathcal{W} \cdot x + b), \quad (2.30)$$

$$\hat{x} = \sigma'(\mathcal{W}' \cdot z + b'), \quad (2.31)$$

donde z representa los datos codificados, $\hat{x}$ corresponde a los datos decodificados, σ y σ' son las funciones de activación, W y W' los pesos de la entrada y salida, y b y b' los sesgos de los datos codificados y decodificados, respectivamente. Como se mencionó anteriormente, el *Autoencoder* es una arquitectura perteneciente al paradigma del aprendizaje no supervisado; en este sentido, se debe considerar la ecuación 2.32 para minimizar el error de reconstrucción.

$$\ell(\mathbf{x}, \hat{\mathbf{x}}) = \|\mathbf{x} - \sigma'(\mathcal{W}' \cdot \sigma(\mathcal{W} \cdot \mathbf{x} + \mathbf{b}) + \mathbf{b}')\|^2. \quad (2.32)$$

Entonces, la pérdida global se obtiene mediante:

$$\mathcal{J}(\Theta) = \sum_{\mathbf{x}} Lr(\mathbf{x}, \hat{\mathbf{x}}), \quad (2.33)$$

en donde $\Theta = (\mathcal{W}, \mathbf{b}, \mathcal{W}', \mathbf{b}')$, que resume los parámetros inherentes.

Aunado a lo explicado anteriormente, la arquitectura base de un *Autoencoder* puede modificarse dependiendo de la tarea a realizar. Este ajuste es necesario porque un *Autoencoder* simple, por sí solo, no es capaz de realizar tareas de clasificación. Partiendo de esta premisa, es posible integrar modelos clasificadores en su capa central, aprovechando la capacidad de la capa codificadora para la extracción de características. En la siguiente sección se detallarán las modificaciones realizadas para continuar con el desarrollo de este trabajo.

Apilamiento de *Autoencoders*

Como ya se mencionó anteriormente, la capa central de un *Autoencoder* extrae características fundamentales de la entrada, sin embargo, se puede mejorar este proceso al pasar como entrada las características extraídas de una capa central a un modelo clasificador, de esta forma, se obtiene una red más compleja que puede aprovechar las características extraídas para realizar tareas de clasificación. En la Figura 2.6 se presentan ejemplos de *Autoencoders* Apilados.

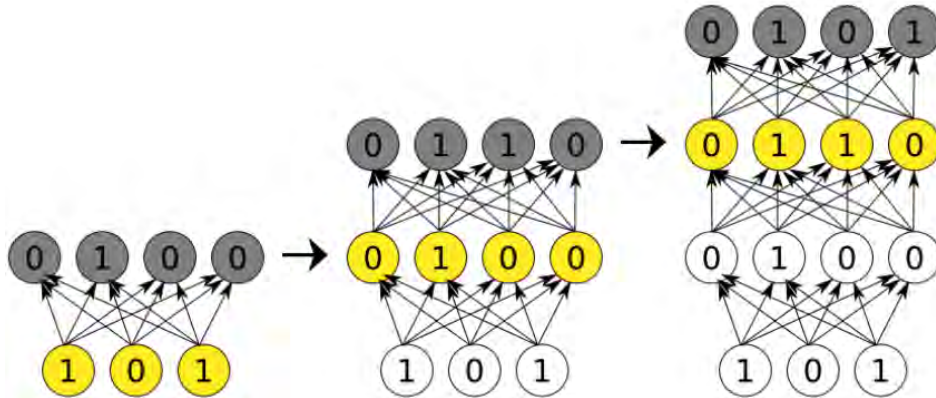


Figura 2.6: Ejemplo de *Autoencoders* apilados.

2.6. Discusión

En este capítulo se abordaron los conceptos clave para el estudio del reconocimiento de arritmias cardíacas con un modelo de inteligencia artificial. Primero, se describe el electrocardiograma (ECG), que es una herramienta esencial para medir la actividad del corazón y detectar irregularidades en su ritmo. Luego, se menciona la base de datos MIT-BIH, utilizada en este trabajo porque contiene registros de diferentes tipos de arritmias y ritmos normales.

También se abordan técnicas para mejorar la calidad de las señales, como el uso de filtros para eliminar el ruido. Además, se presentan métodos para transformar las señales y analizar su información en el dominio de la frecuencia, lo que ayuda a identificar patrones importantes.

Por último, se introdujo el concepto de aprendizaje profundo y los *Autoencoders*, que son modelos diseñados para extraer información relevante de las señales. En este caso, se ha modificado su estructura para que, además de procesar los datos, también puedan clasificar diferentes tipos de arritmias. Con estas herramientas, se busca desarrollar un modelo eficiente y preciso para ayudar en la detección de problemas cardíacos.

3 | Estado del arte

3.1. Antecedentes

Los siguientes trabajos que se presentan a continuación preceden a la investigación actual, ya que utilizan técnicas similares y servirán como referencia para el desarrollo de este trabajo.

Conducción autónoma de un vehículo simulado mediante un modelo de red neuronal convolucional Recurrente [25].

En este trabajo se propuso un modelo de red neuronal híbrido basado en una combinación de capas convolucionales y recurrentes para realizar el control de un vehículo simulado en un sistema operativo robótico. Para obtener la información que se procesó, el vehículo estuvo dotado de sensores visuales y espaciales que permitieron la percepción del entorno. El modelo utilizado fue el llamado Chauffeur, el cual fue modificado sustituyendo su parte convolucional por un bloque llamado convolución distribuida en el tiempo. Este modelo permitió evaluar la correcta predicción de dirección de un vehículo en condiciones climáticas no favorables. El modelo propuesto fue comparado con otros modelos de la literatura, incluido el modelo base de Chauffeur. Los resultados mostraron una mejoría del 11.4% en la autonomía del vehículo en escenarios libres y un 8.9% en escenarios con obstáculos. El modelo propuesto se muestra en la Figura 3.1.

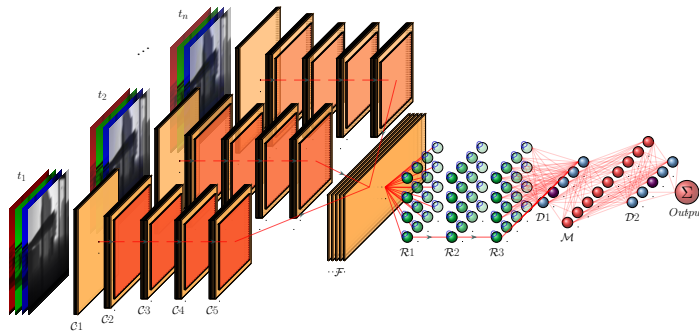


Figura 3.1: Modelo de Chauffeur modificado propuesto en [25].

Clasificación de la fibrilación auricular en señales electrocardiográficas utilizando una red neuronal recurrente profunda entrenada con momentos de tiempo-frecuencia [2].

En este trabajo se propuso un método eficiente para la clasificación de la fibrilación utilizando una red neuronal híbrida bidireccional. La metodología de este trabajo se dividió en 3 módulos: transformación al dominio de tiempo-frecuencia, extracción de características y reducción de dimensionalidad. Para realizar las pruebas, se utilizó la base de datos Physionet Cinc 2017. Se realizaron numerosos experimentos con el objetivo de sintonizar el modelo y de compararlo con otros en la literatura. El modelo propuesto se evaluó con las métricas derivadas de la matriz de confusión, teniendo como resultado un desempeño mayor al 90%, reduciendo el costo computacional hasta en un 90% en comparación con otros. El modelo de la red neuronal propuesta se muestra en la Figura 3.2.

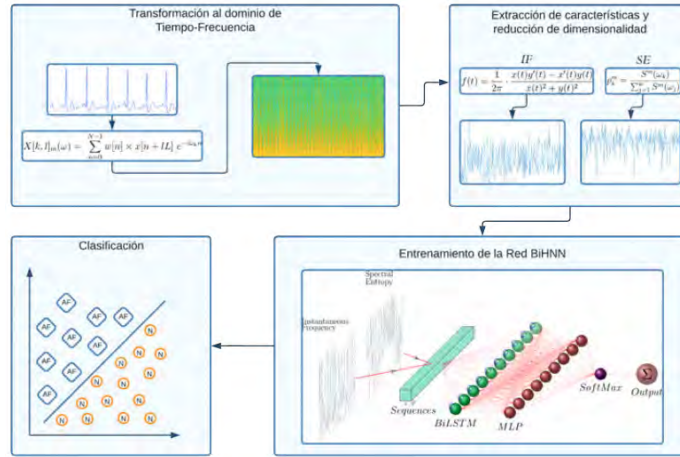


Figura 3.2: Diagrama de bloques del modelo propuesto en [2].

Detección de anomalías en mamografías utilizando la red neuronal convolucional AlexNet [28].

En este trabajo se presentó una metodología computacional que permitió detectar anomalías en mamografías utilizando la red neuronal convolucional AlexNet, ya que tuvo un nivel alto de abstracción. Para la recolección de datos, se hizo uso de la base de datos MIAS, la cual contiene información referente al tipo de anomalía. La metodología propuesta se dividió en 2 partes: pre-procesamiento de las imágenes y diseño y caracterización de la red para la detección de anomalías. Se realizaron diferentes pruebas con diferentes configuraciones para comparar resultados. Los resultados del mejor modelo lograron una efectividad del 94 % al clasificar las mamografías como sanas o anormales, además de tener un área bajo la curva (AUC) de 0.925. La metodología de solución propuesta se muestra en la Figura 3.3.

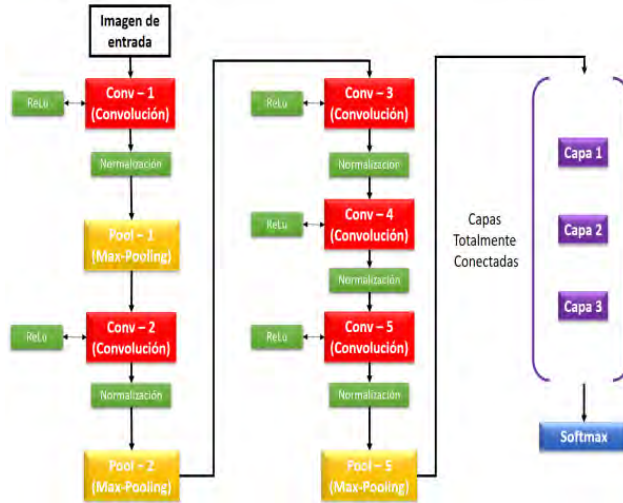


Figura 3.3: Modelo de red neuronal propuesto en [28].

Red neuronal pulsante para el reconocimiento de imágenes de cerebros sanos y no sanos [7].

En este trabajo se desarrolló un análisis en la clasificación de una neurona SNN representada por el modelo Izhikevich, usada en el reconocimiento de imágenes cerebrales obtenidas de imágenes de resonancia magnética, demostrando que la precisión de la clasificación en el modelo propuesto alcanzó valores comparables a otras redes neuronales artificiales (ANN) como MLP. El modelo propuesto alcanzó una exactitud de 80 %.

3.2. Trabajos relacionados

En esta sección se presentan brevemente los trabajos más relevantes en relación al uso de diferentes variantes de redes neuronales y de técnicas que se aplican a la tarea de detección, clasificación y reconocimiento de arritmias cardíacas.

Un enfoque de aprendizaje profundo para la detección en tiempo real de la fibrilación auricular [1].

En este trabajo se propuso un modelo combinado de redes neuronales convolucionales y recurrentes para extraer características significativas de segmentos de los intervalos RR para clasificarlos como fibrilación auricular o ritmo normal. El modelo fue entrenado y validado sobre 3 diferentes bases de datos incluyendo un total de 89 muestras. El modelo logró una sensibilidad de 98.98 % y una especificidad del 96.95 %. El modelo propuesto en este trabajo se muestra en la Figura 3.4.

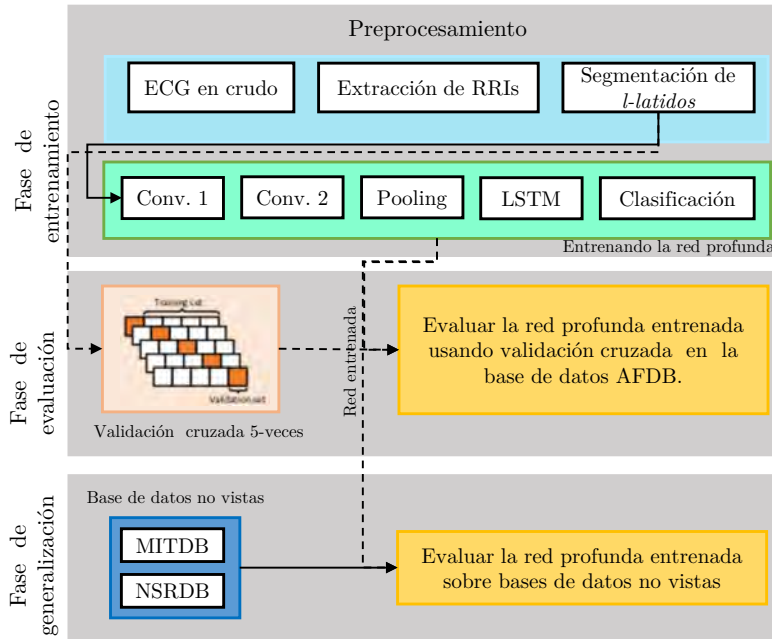


Figura 3.4: Diagrama de flujo del método propuesto en [1].

Agrupamiento de Señales ECG con Rasgos Aprendidos Usando *Autoencoder* Profundo [46].

En este trabajo se propuso un algoritmo basado en *Autoencoders* convolucionales como extractor de rasgos no supervisado para hallar grupos de señales electroencefalográficas, como apoyo para el especialista médico para facilitar el diagnóstico de la condición de epilepsia. Se diseñaron 3 *Autoencoders* con señales de entrada 4096×1 , 2048×2 y 768×3 para analizar el efecto de la longitud de la señal sobre la representación latente generada por los *Autoencoders*. La representación latente se llevó a un espacio bidimensional donde se obtuvo la media y la desviación estándar para visualizarla y operar sobre ellas los algoritmos de agrupamiento. Los resultados mostraron un margen de error de 0.

se muestra en

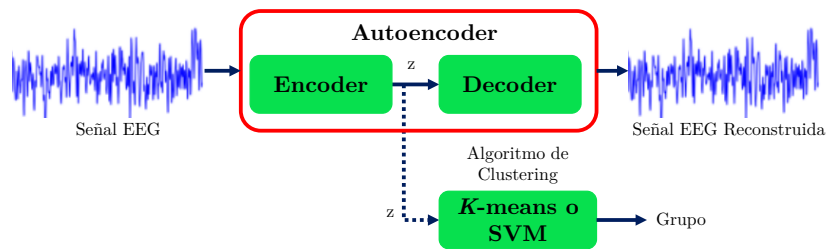


Figura 3.5: Diagrama de bloques del proceso de agrupamiento propuesto en [46].

Detección de fibrilación auricular usando una red neuronal convolucional 1d [16].

Este trabajo propuso un método de detección de fibrilación auricular basado en una arquitectura de red neuronal convolucional 1D para aumentar la precisión de la detección y reducir la complejidad de la red. La red propuesta se desarrolló al investigar el impacto de los componentes principales de un bloque de convolución en la precisión de detección, además de utilizar la búsqueda en cuadrícula para obtener hiperparámetros óptimos de la CNN. Con todo esto, se obtuvo una red simple pero efectiva. El resultado del modelo propuesto en este trabajo dio un valor de F1-score de 78.2 %. La metodología propuesta

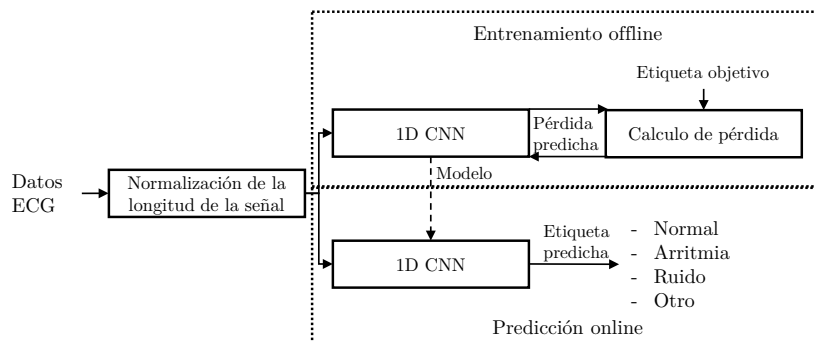


Figura 3.6: Diagrama de flujo del modelo propuesto en [16].

CardioNet: un sistema eficiente para la clasificación de arritmias usando *Transfer learning* [30].

Este artículo presentó un método novedoso de clasificación de los latidos del corazón a partir de ECG mediante aprendizaje profundo. Se diseñó un sistema automatizado llamado *CardioNet* que empleó el principio de aprendizaje por transferencia para una clasificación más rápida y sólida de los latidos del corazón para la detección de arritmias. Utilizó una arquitectura previamente entrenada de *DenseNet* que se entrenó en un conjunto de datos de *ImageNet* de millones de imágenes. Los pesos obtenidos durante el entrenamiento de *DenseNet* se utilizaron para ajustar el aprendizaje de *CardioNet* en el conjunto de datos de ECG, lo que dio como resultado un sistema único que proporcionó entrenamiento y pruebas más rápidos. El conjunto de datos de ECG se preparó mediante un proceso de aumento para proporcionar un aprendizaje integral de la morfología de los latidos del corazón en presencia de variaciones intra-clase. El modelo propuesto logró una exactitud de clasificación alta de 98.92 %, superando a otras técnicas y mostrando ser robusto en la clasificación entre un ritmo cardíaco normal y una arritmia. La arquitectura propuesta se muestra en la Figura 3.7.

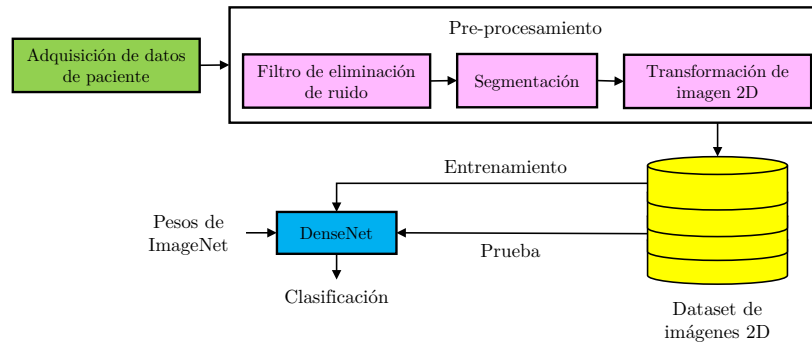


Figura 3.7: Esquema de clasificación CardioNet para arritmias cardiacas utilizado en [30].

Modelo híbrido basado en redes neuronales, sistemas difusos tipo 1 y 2 para clasificación de arritmias cardíacas de 2 derivaciones [34].

En este trabajo se presentó un enfoque usando métodos de inteligencia computacional para construir un modelo híbrido como método clasificador para arritmias cardíacas de 2 derivaciones. El modelo mencionado se construyó usando redes neuronales artificiales y lógica difusa y fue entrenado con datos obtenidos de la base de datos MIT-BIH [44]. Las señales de ECG fueron pre-procesadas para mejorar los resultados de la clasificación. El modelo estuvo formado por 2 módulos básicos, y dichos módulos estuvieron conformados por 3 diferentes clasificadores basados en: algoritmo KNN difuso, perceptrón multicapa con descenso del gradiente y perceptrón multicapa con retropropagación de gradiente. Las salidas de los clasificadores fueron combinadas usando un sistema difuso para la integración de resultados. El modelo logró una tasa de clasificación del 92.9 % para el módulo 1 y 92.7 % para el módulo 2. La representación del modelo propuesto se muestra en la Figura 3.8.

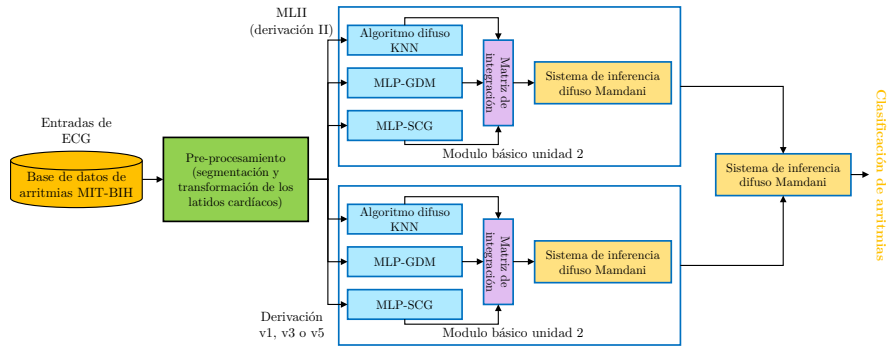


Figura 3.8: Modelo híbrido para la clasificación de arritmias cardíacas de 2 derivaciones utilizado en [34].

Diagnóstico de la fibrilación auricular mediante el análisis inteligente de registros intracardíacos [8].

Este trabajo tuvo como objetivo ofrecer una herramienta de apoyo a los especialistas médicos para el diagnóstico de la fibrilación auricular a partir de datos intracardíacos capturados por marcapasos de individuos que padecían la enfermedad. Para lograr el objetivo, se exploró el uso de un *Autoencoder* variacional, un tipo de red neuronal, la cual tuvo como objetivo aprender una representación simplificada de los datos de entrada, reduciéndolos a menores dimensiones para después reconstruirlos nuevamente a la dimensión original. El modelo dio como resultado una precisión del 97.12 %, demostrando que fue capaz de competir con los mejores clasificadores para series temporales reportados hasta el año 2019. La arquitectura de un *Autoencoder* se muestra en la Figura 3.9.

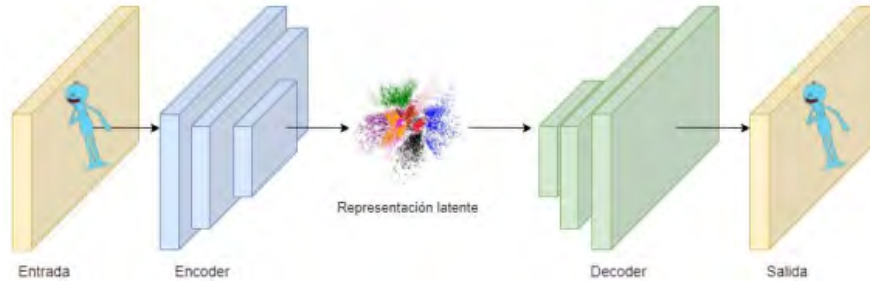


Figura 3.9: Arquitectura de un *Autoencoder* [8].

Clasificación de arritmias cardíacas utilizando análisis de componentes principales y redes neuronales pre-alimentadas en segmentos cortos de electrocardiograma [27].

En este trabajo se presentó el diseño y desarrollo de un sistema que permitió la clasificación de arritmias cardíacas mediante la utilización de análisis de componentes principales y redes neuronales pre-alimentadas. Para la evaluación del sistema se consideraron 3 registros ECG de prueba, concretamente 207, 223 y 106 de la base de datos MIT-BIH. Posteriormente, los resultados obtenidos se concentraron en una matriz de confusión, dando una exactitud general del 80.1 % al distinguir entre los 14 tipos de señales ECG en estudio. El diagrama de bloques del sistema se muestra en la Figura 3.10.

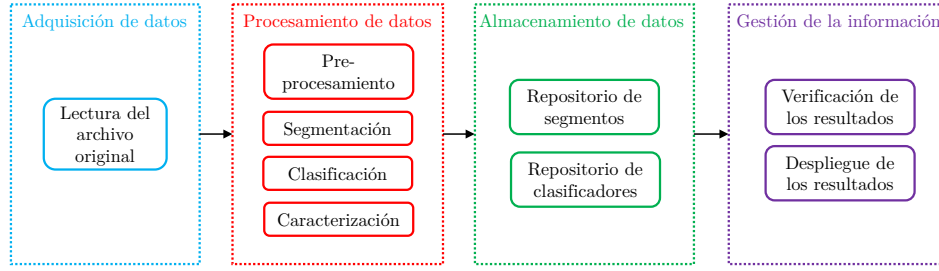


Figura 3.10: Diagrama de bloques del sistema propuesto en [27].

Enfoque novedoso de aprendizaje profundo para la clasificación de fibrilación auricular basado en una red neuronal convolucional y modificada de Elman [47].

En este trabajo se desarrolló un enfoque novedoso para la detección automática de la fibrilación auricular basado en una red neuronal de 11 capas. La estructura de la red estuvo principalmente apilada por una red neuronal convolucional y una red neuronal modificada Elman, mientras realizaba automáticamente la clasificación de señales de un extremo a otro. Para demostrar la superioridad del modelo, se utilizó un conjunto de datos específicos de fibrilación auricular. Se muestran los resultados de la clasificación de las señales de fibrilación auricular.

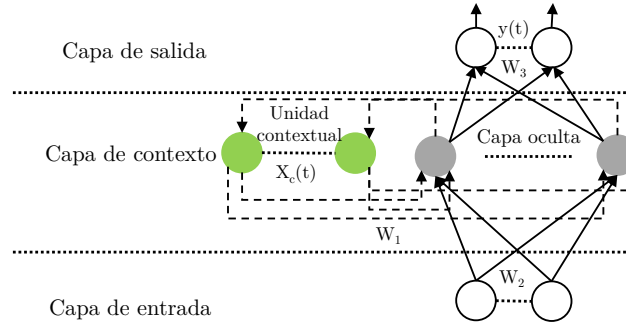


Figura 3.11: Arquitectura de la red propuesta en [47].

Clasificación de arritmias cardíacas mediante funciones basadas en transformadas Q-wavelet sintonizables y en un clasificador de máquina de soporte vectorial [19].

Este trabajo propuso una técnica novedosa y eficiente de clasificación de señales ECG para ritmos cardíacos normales y 7 tipos de arritmias. El modelo propuesto utilizó una Q-wavelet sintonizable basada en características de señales ECG, las cuales fueron adquiridas desde diferentes registros ECG de la base de datos MIT-BIH. Para la extracción de características, cada latido del ECG se descompuso hasta el sexto nivel de la transformada Q-wavelet. Se seleccionaron coeficientes aproximados del sexto nivel como características de cada latido del ECG. Para la clasificación, se utilizaron características de 14.878 latidos de ECG para el entrenamiento y 26.219 latidos de ECG para fines de prueba. El modelo logró una exactitud de 99.27 %, una sensibilidad de 96.22 % y una especificidad de 99.58 %. El modelo propuesto se muestra en la Figura 3.12.

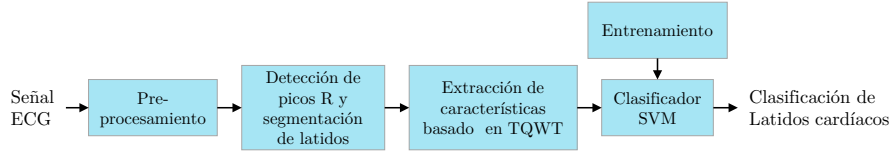


Figura 3.12: Diagrama de la metodología propuesta en [19].

Red neuronal convolucional 1D para la detección de latidos ventriculares [24].

En este trabajo se desarrolló un enfoque novedoso para detectar latidos ventriculares basado en el uso de redes neuronales convolucionales en 1D. La entrada de este algoritmo fueron las señales en crudo del electrocardiograma, sin realizar pre-procesamiento ni extracción de características. La salida del algoritmo fue filtrada usando una combinación de filtros lineales y no lineales, produciendo el resultado. La metodología de evaluación siguió el modelo inter-paradigma del paciente, donde el algoritmo fue entrenado y evaluado utilizando subconjuntos independientes.

El desempeño del modelo propuesto se evaluó en 2 tareas: la detección QRS y la clasificación de los latidos. La precisión del 96.5% y un valor de 85.8% y un valor de 85.8% en la Figura 3.14.

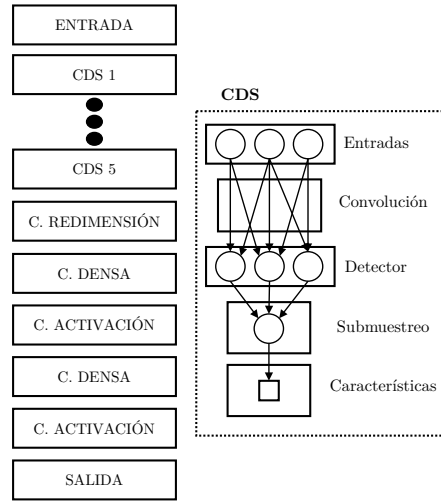


Figura 3.13: Modelo de red convolucional propuesto en [24].

Detección de la fibrilación auricular y clasificación ECG basado en una red CNN-BiLSTM [48].

En este trabajo se propuso un sistema que usa 2 diferentes *frameworks* de la combinación de una red neuronal convolucional y una red LSTM para clasificar el ritmo cardíaco normal, fibrilación auricular y otras señales ruidosas. Se utilizó la base de datos MIT-BIH. El resultado de este trabajo demostró que la cascada de 2 redes neuronales profundas tiene un mayor desempeño que la concatenación de ellas, logrando un F1-score de 82%. Los modelos utilizados se muestran en la Figura 3.14.

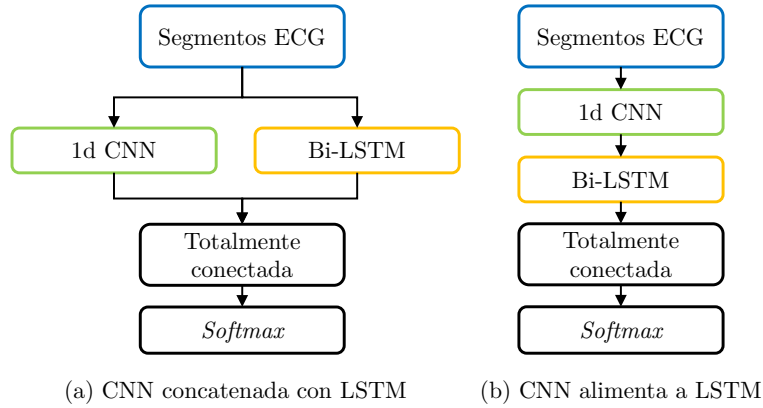


Figura 3.14: Modelos propuestos en [48].

En la Figura 3.14a, CNN o LSTM se consideraron un extractor de características que se implementó para capturar la información espacial y la característica temporal, respectivamente. Se concatenó la característica espacial y la característica temporal generadas directamente por diferentes modelos de aprendizaje profundo. Siguió una capa completamente conectada para predecir la salida. Luego, la salida de la capa completamente conectada se pasó a una función *softmax* y se transformó como la probabilidad de pertenecer a las clases correspondientes.

En la Figura 3.14b, LSTM recibió la salida de CNN y extrajo características de una manera más profunda. La salida de la porción de CNN, que es el vector de características espaciales, se envió a LSTM para extraer la información temporal. La salida de la parte de LSTM se pasó a una capa completamente conectada para predecir la clase. La función *softmax* se aplicó para transformar la salida en forma de probabilidad.

Método de ingeniería de características basado en aprendizaje profundo y K-NN para la clasificación de latidos de ECG y la detección de arritmias [20].

El objetivo de este trabajo fue clasificar los latidos cardíacos para la detección de arritmias. Para lograr esto, se propuso un novedoso método basado en aprendizaje profundo y KNNs. Las características extraídas por el modelo propuesto fueron clasificadas con diferentes clasificadores, así como árboles de decisión, máquinas de soporte vectorial con diferentes *kernels* y bosques aleatorios. Para separar los datos de entrenamiento y prueba, se utilizó validación cruzada. La principal ventaja del método propuesto fue su bajo tiempo de cálculo en comparación con el entrenamiento de modelos de aprendizaje profundo desde cero y su alta precisión en comparación con los modelos tradicionales de aprendizaje automático.

La solidez e idoneidad del método propuesto para la extracción de características se demuestra con los valores de las métricas, teniendo 99.77% para exactitud, 99.75% para precisión y 99.3% para sensibilidad. La metodología propuesta se muestra en la Figura 3.15.

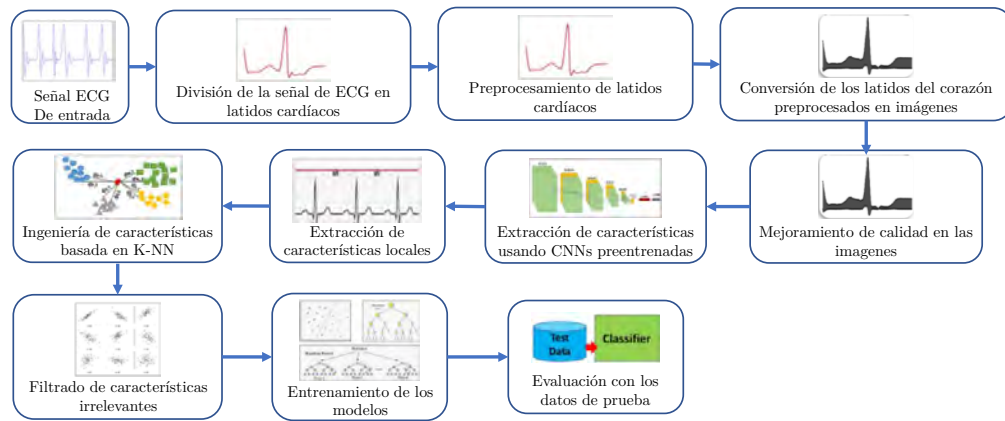


Figura 3.15: Pasos principales de la metodología propuesta en [20].

Reconocimiento automático de arritmias desde señales de ECG usando diferentes modelos de red de métodos de funciones con memoria a largo plazo [31].

En este trabajo se presentaron nuevos métodos de extracción de características empleando modelos de red de memoria a largo plazo (LSTM), los cuales ayudaron en la detección de ritmos cardíacos desde señales ECG. Las señales de ECG se descompusieron en 45 características, las cuales fueron usadas como una secuencia de entrada hacia un modelo simple LSTM. La base de datos empleada para entrenar este modelo fue MIT-BIH. El modelo propuesto ayudó a clasificar 5 distintos ritmos cardíacos con resultados de: precisión de 96.73 %, exactitud de 99.37 %, especificidad de 99.14 %, F1-score de 95.77 % y sensibilidad de 94.80 %. El diagrama del sistema se muestra en la Figura 3.16.

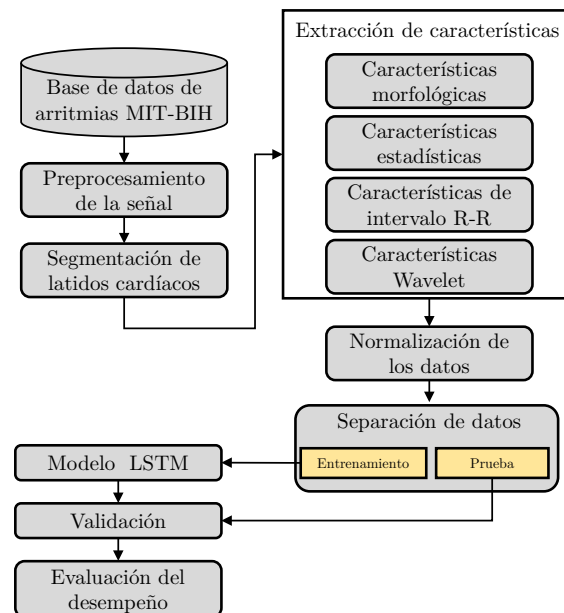


Figura 3.16: Diagrama de la metodología propuesta en [31].

Un método novedoso híbrido de aprendizaje profundo con algoritmo de búsqueda de *cuckoo* para la clasificación de enfermedades arritmias mediante señales de ECG [39].

Este trabajo presentó un enfoque híbrido eficiente para la clasificación de muestras de ECG en clases de arritmia para detectar anomalías en los latidos del corazón. Se pretendió realizar la clasificación al eliminar el ruido inherente de las señales en la fase de preprocesamiento usando la transformada discreta *wavelet* (DWT). El complejo QRS jugó un papel importante en la identificación de la señal. Por lo tanto, se determinaron la posición y amplitud de los picos R para detectar el complejo QRS. Los vectores de características del complejo QRS se optimizaron aún más con el algoritmo de optimización de búsqueda *cuckoo* (CS), además de eliminar señales de ruido utilizando DWT. El vector de la máquina de soporte vectorial contuvo la mejor información de entrenamiento utilizada para entrenar la red neuronal de retropropagación (FFBPNN) para proponer la variante DWT + CS + SMS-FFBPNN. Para realizar el entrenamiento se utilizó la base de datos MIT-BIH. El modelo logró una precisión de 98.247 % y una sensibilidad de 95.68 %. La representación del modelo propuesto se muestra en la Figura 3.17.

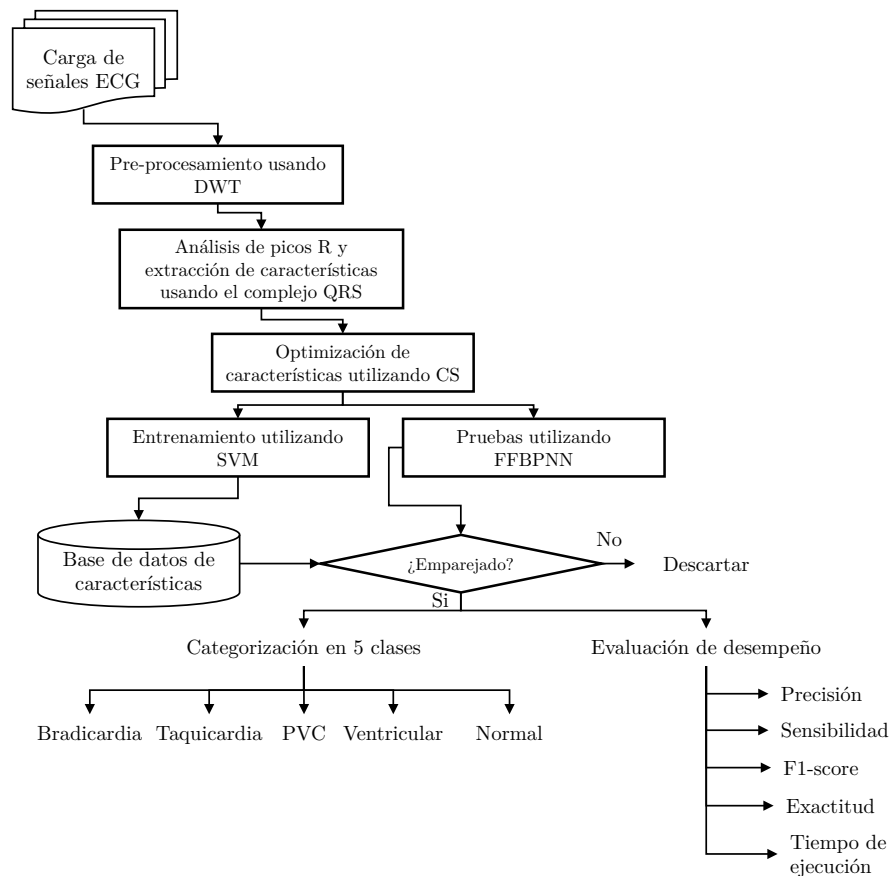


Figura 3.17: Esquema de la metodología propuesta en [39].

Los resultados del desempeño del modelo fueron: F1-score de 0.806, FAF-score de 0.914 para la fibrilación auricular. F-block de 0.879 para bloques, FPC y FST de 0.801 y 0.742 respectivamente. Estos resultados demuestran un buen desempeño que tiene potenciales aplicaciones prácticas.

Una clasificación mejorada de arritmias cardíacas usando un intervalo RR [33].

En este trabajo se propuso un enfoque mejorado de clasificación de arritmias cardíacas basado en el intervalo RR. Se utilizó la transformada discreta de *wavelet* y filtros para eliminar el ruido de alta frecuencia y la desviación de la línea base del ECG sin procesar. Después, el ECG procesado fue segmentado tras la determinación de la región QRS. Se extrajeron las características principales del intervalo RR para clasificar distintos tipos de arritmias. Los clasificadores vecino más cercano (KNN), máquina de soporte vectorial (SVM), árboles de decisión (DT), naive bayes (NB) y bosques aleatorios (RF) fueron utilizados. El mejor modelo fue SVM, con un kernel gaussiano, el cual logró una sensibilidad del 99.28 %, una especificidad del 99.63 % y una exactitud del 99.51 %. El modelo propuesto se muestra en la Figura 3.18.

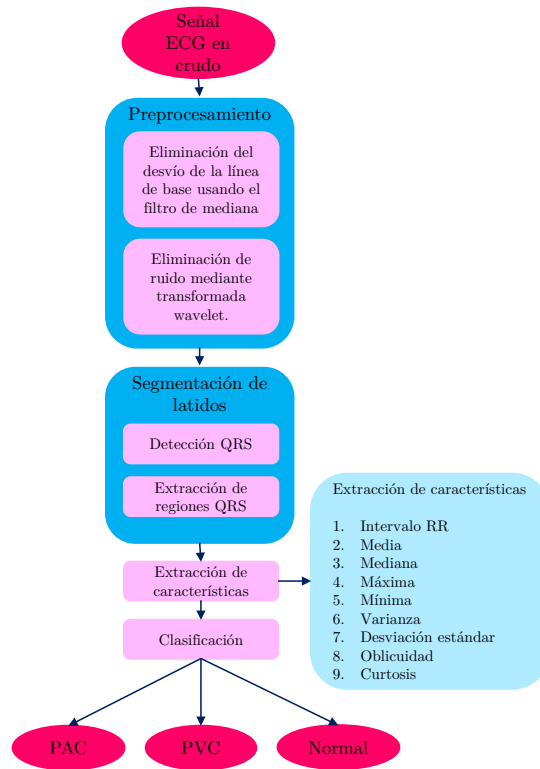


Figura 3.18: Diagrama de flujo del proceso del modelo propuesto en [33].

Clasificación automática de arritmias cardíacas usando una combinación de redes profundas residuales y memoria a largo plazo bidireccional [13].

En este trabajo se propuso un nuevo método para la clasificación automática de arritmias basado en redes neuronales profundas. Se entrenaron 2 modelos de redes neuronales, concretamente una red neuronal de módulos residuales convolucionales y una red neuronal bidireccional de memoria a largo plazo para extraer características de señales de ECG en crudo. Las características extraídas fueron concatenadas para formar un vector de características, el cual fue entrenado para realizar la clasificación final. El diagrama de flujo del sistema propuesto se muestra en la Figura 3.19.

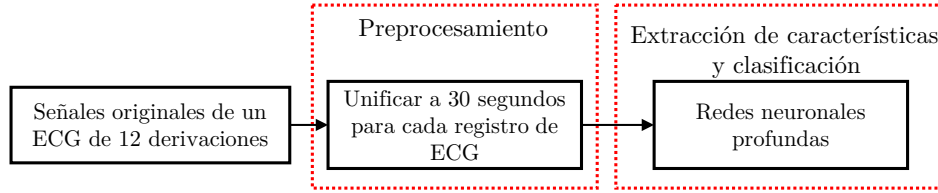


Figura 3.19: Diagrama de la metodología propuesta en [13].

Clasificación precisa de arritmias de ECG utilizando redes de aprendizaje profundo de compresión rápida mejoradas MOWPT [18].

En este trabajo se propuso un clasificador de ECG inteligente basado en un método de clasificación preciso que utilizó redes neuronales convolucionales residuales de compresión rápida (FCResNet). En este método, la transformada de paquete wavelet de superposición máxima (MOWPT), que proporciona un patrón de pavimentación de escala de tiempo integral y posee la propiedad de invariancia de tiempo, se utilizó para descomponer las señales de ECG originales en muestras de subseñales de diferentes escalas. Posteriormente, las muestras de los cinco tipos de arritmia se utilizaron como entrada para FCResNet, de modo que se identificaron y clasificaron los tipos de arritmia del ECG. El modelo propuesto solucionó sustancialmente los problemas de baja eficiencia computacional, difícil convergencia y degradación del modelo. Las pruebas del modelo mostraron un resultado de 98.79 % ,

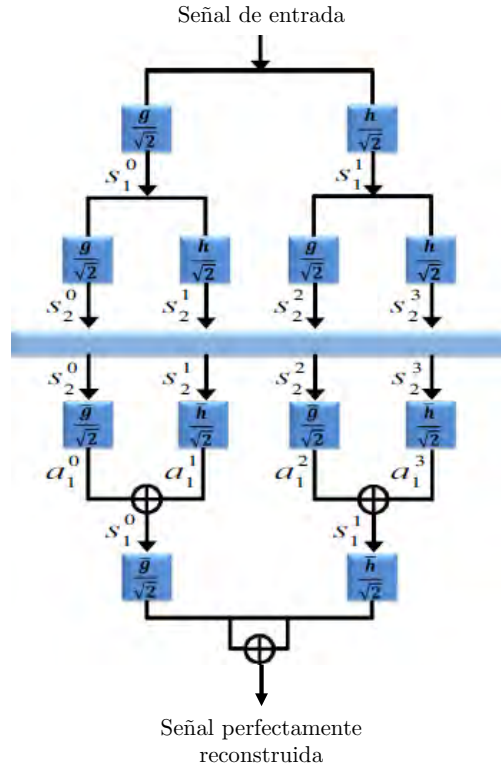


Figura 3.20: Proceso de descomposición y reconstrucción MOWPT empleado en [18].

Detección robusta de fibrilación auricular a partir de ECG de corta duración mediante redes neuronales convolucionales [29].

Este trabajo propuso mejorar el método ya existente de identificación de fibrilación auricular usando transformadas discretas de Fourier (DWT) junto con una red neuronal convolucional 1D para clasificar las clases de: Ritmo sinusal normal, Fibrilación auricular y No Fibrilación auricular. Los datos que se utilizaron para entrenar el modelo estuvieron conformados por 3 datasets públicos y 1 dataset de un hospital de Indonesia. La robustez del modelo fue evaluada con base en varios datos de validación con un patrón invisible de 4 datasets.

Los resultados generados demostraron que la 1D-CNN superó a otros enfoques, logrando un rendimiento satisfactorio con una alta capacidad de generalización. Los resultados fueron: Exactitud de 99.98 %, sensibilidad de 99.91 %, especificidad de 99.91 %, precisión de 99.99 % y F1-score de 99.95 %. Estos resultados demostraron que este enfoque podía ayudar al diagnóstico de fibrilación auricular y al autocontrol del paciente para mejorar la detección temprana y el tratamiento de la FA. El método de solución propuesto se muestra en la Figura 3.21

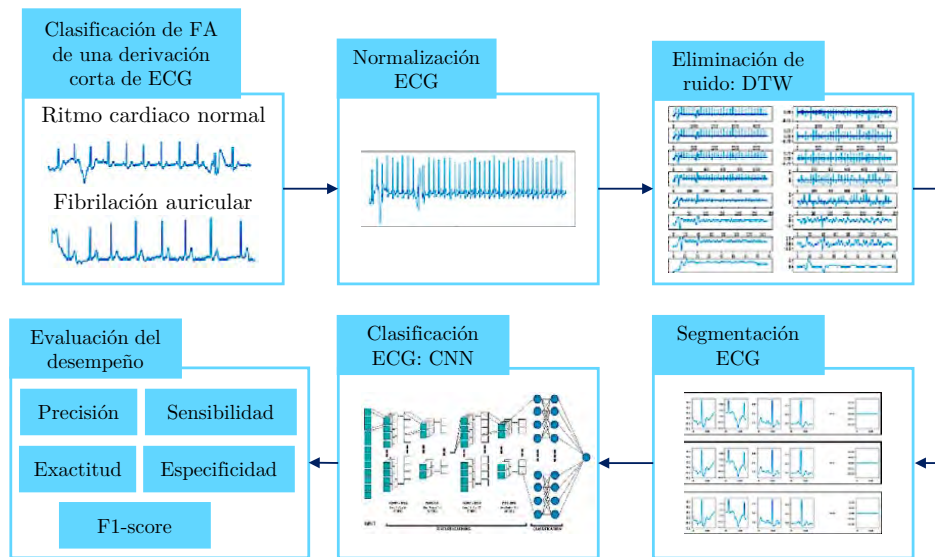


Figura 3.21: Framework del modelo propuesto en [29].

Clasificación de arritmias mediante aprendizaje profundo y aprendizaje automático con funciones extraídas del procesamiento de señales basado en formas de onda [17].

Este trabajo presentó un método de análisis de señales basado en formas de onda (WBSP) para la clasificación de arritmias cardíacas. Al realizar WBSP, primero se filtraron las señales de ECG, se buscaron mínimos locales y se eliminó la desviación inicial. Posteriormente, se ajustaron las señales de ECG procesadas con gaussianas y se extrajeron los parámetros. Luego, se aprovecharon los productos de WBSP para lograr la clasificación de arritmias con los clasificadores propuestos basados en aprendizaje automático y aprendizaje profundo. Se utilizó la base de datos MIT-BIH para validar el modelo, alcanzando una exactitud de 98.8 % y una sensibilidad de 96.3 %. La metodología propuesta se muestra en la Figura 3.22.

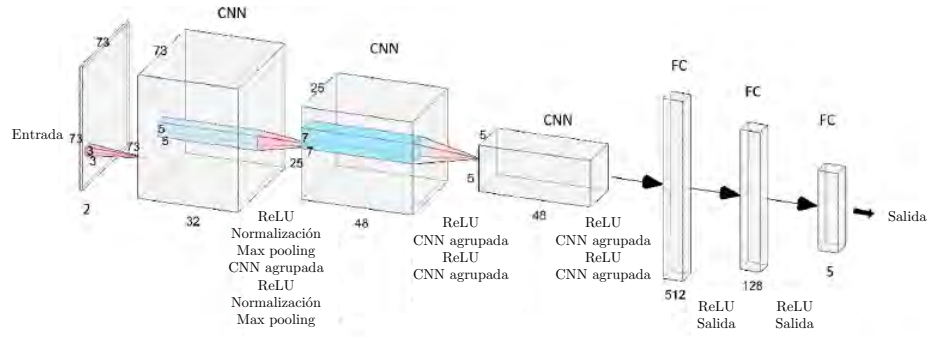


Figura 3.22: Diagrama de la metodología propuesta en [17].

Clasificación de arritmias cardíacas mediante redes neuronales convolucionales y optimización por enjambre de partículas [38].

Este trabajo propuso la implementación de una arquitectura híbrida entre redes neuronales convolucionales y el algoritmo metaheurístico de optimización por enjambre de partículas. El metaheurístico se encargó de optimizar el número y tipo de capas de la red neuronal a través de la minimización en la pérdida durante el entrenamiento. Los resultados demostraron que el modelo fue capaz de encontrar una adecuada arquitectura de capas de una red neuronal convolucional y, con ello, obtener una precisión del 97 % en la clasificación de estas arritmias. La metodología empleada se muestra en la Figura 3.23.

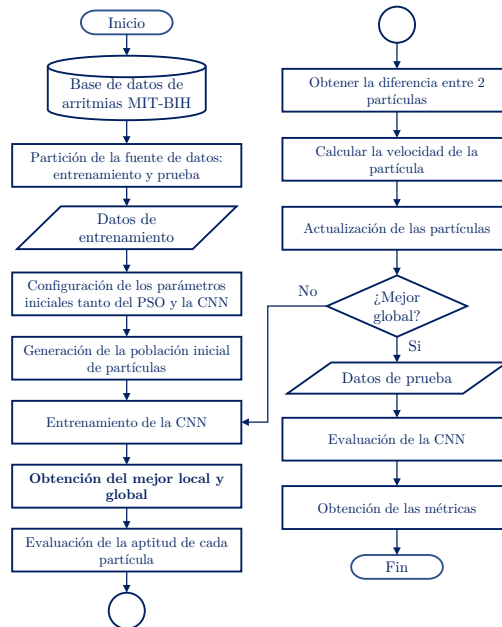


Figura 3.23: Diagrama de la metodología propuesta en [38].

Detección automatizada de fibrilación auricular usando una red híbrida CNN-LSTM sobre bases de datos ECG desbalanceadas [32].

En este trabajo se propuso un modelo híbrido novedoso utilizando pérdida focal, una versión mejorada de pérdida de entropía cruzada, para tratar con datos de entrenamiento desbalanceados. Las características ECG inicialmente extraídas por una red neuronal convolucional fueron proporcionadas como entrada a una red LSTM para una memorización dinámica temporal y, de este modo, se obtuvo una clasificación más precisa en los cuatro tipos de ritmo de ECG: normal, fibrilación auricular, aleteo auricular y ritmo de la unión AV. El modelo fue entrenado con la base de datos MIT-BIH y alcanzó una sensibilidad de 97.87 % y una especificidad de 99.29 % utilizando una estrategia de validación cruzada. En la Figura 3.24 se muestra la metodología empleada.

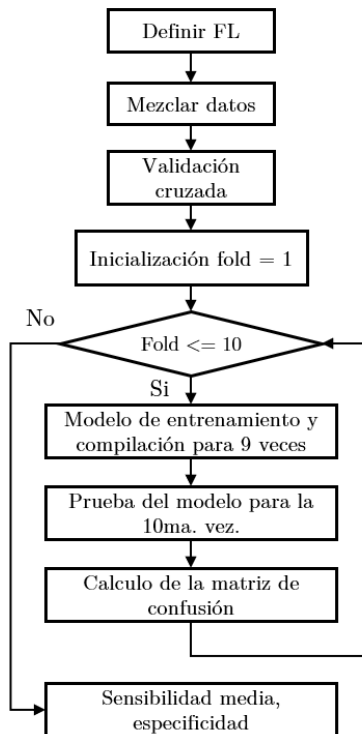


Figura 3.24: Diagrama de bloques propuesto en [32].

Sistema de detección de fibrilación auricular a partir de procesamiento de señales de ECG e inteligencia artificial [3].

Este trabajo propuso un sistema de detección de fibrilación auricular basado en inteligencia artificial con el fin de abordar el problema desde una aproximación preventiva y así disminuir la ocurrencia de eventos adversos. Se entrenaron varios modelos de IA con datos pre-procesados de señales ECG para posteriormente evaluarlos y validarlos con otro conjunto de señales mediante métricas de evaluación correspondientes. Se eligió el mejor modelo (MLP), el cual tuvo una exactitud superior al 92 % y fue embebido en un microprocesador y probado con un simulador de pacientes. La metodología utilizada se muestra en la Figura 3.25.

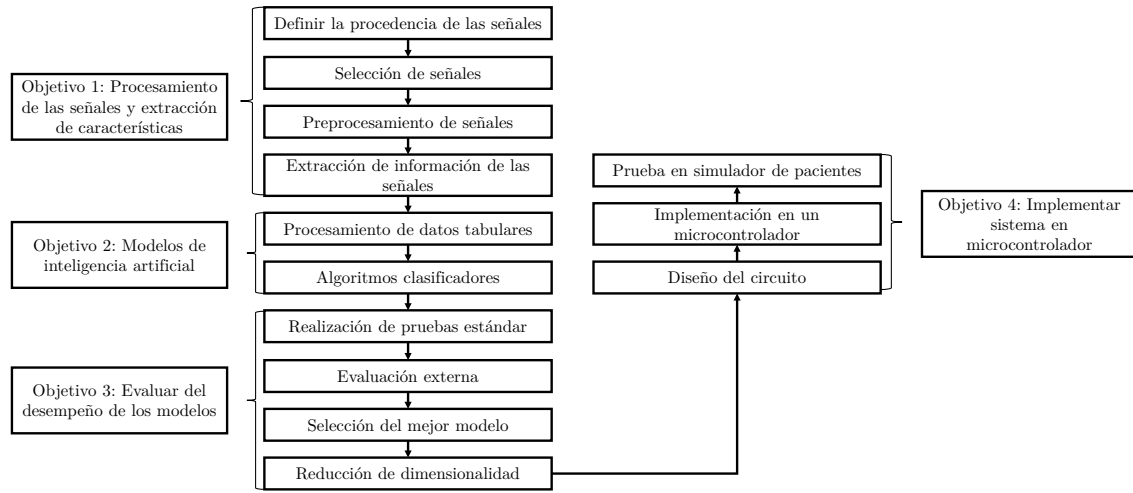


Figura 3.25: Diseño metodológico resumido utilizado en [3].

***MultiFusionNet*: Detección de fibrilación auricular con redes neuronales profundas [45].**

En este trabajo se propuso una técnica denominada *MultiFusionNet*, la cual es una red neuronal profunda que usó un método de fusión multiplicativa para combinar 2 redes neuronales profundas sobre diferentes fuentes de conocimiento. De esta forma, esta técnica pudo explotar las características relevantes extraídas para mejorar la utilización del modelo de aprendizaje profundo en los datos sin procesar. Los experimentos mostraron que este enfoque ofreció la clasificación de AF más precisa y superó a los algoritmos publicados en ese entonces que utilizaban características extraídas o datos sin procesar por separado. Los resultados generados correspondieron a un F1-score de 90 %. En la

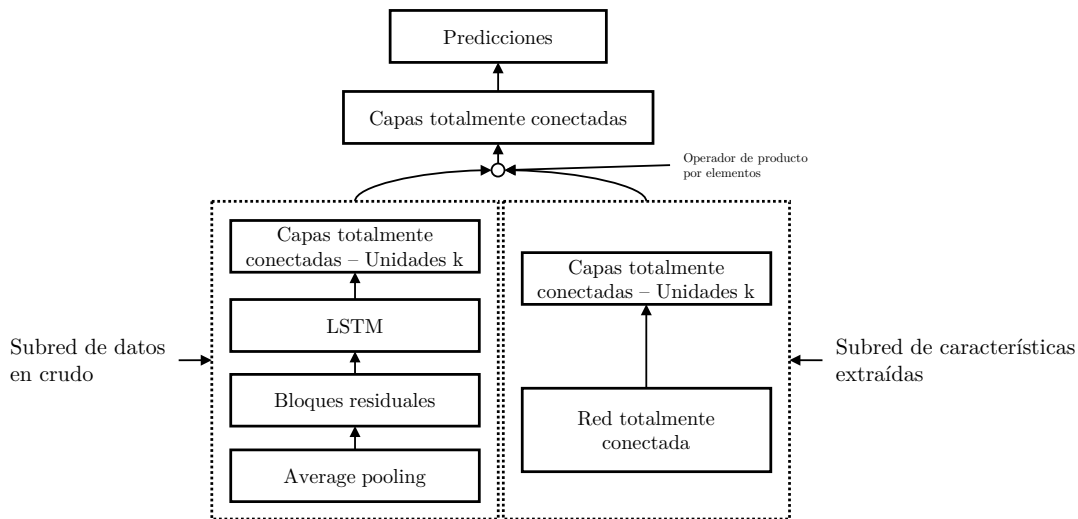


Figura 3.26: Arquitectura de la red propuesta en [45].

Un nuevo método de aumento de datos para mejorar las redes neuronales profundas para la detección de la fibrilación auricular [6].

En este trabajo se propuso una estrategia novedosa para el aumento de los datos basada en duplicación, concatenación y remuestreo de los episodios ECG para equilibrar el número de muestras entre diferentes categorías, así como incrementar la diversidad de las muestras. Se entrenó una red LSTM con los datos aumentados. Los resultados del modelo demostraron ser superiores a otras técnicas de aumento de datos, como lo son permutación y deslizamiento de ventana. La red también fue sometida al sistema de evaluación del desafío CinC 2017, teniendo un F1-score cercano al ganador (92 % vs 93 %). El modelo propuesto para el aumento de datos se muestra en la Figura 3.27.

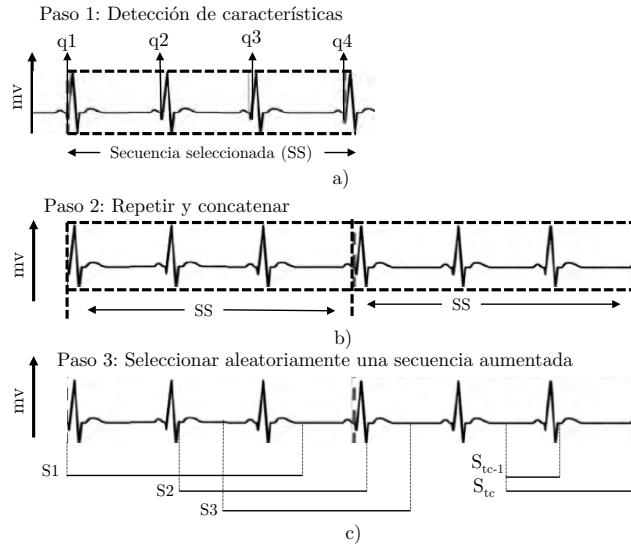


Figura 3.27: Diagrama esquemático de la estrategia propuesta en [6].

Detección de arritmia cardíaca mediante un enfoque de aprendizaje profundo y representación de frecuencia de tiempo de señales de ECG [9].

El objetivo principal de este estudio fue crear un modelo automático de aprendizaje profundo capaz de clasificar precisamente señales de ECG en 3 categorías: arritmia cardíaca, fallo del corazón congestivo y ritmo normal sinusal. Para lograr esto, las señales obtenidas de la base de datos MIT-BIH y BIMC fueron pre-procesadas y segmentadas antes de ser utilizadas por el modelo propuesto. Los modelos ResNet 50 y AlexNet fueron correctamente sintonizados y configurados para obtener resultados óptimos. El modelo propuesto generó los siguientes resultados: Exactitud de 99.2 %, especificidad de 99.6 % y sensibilidad, precisión y F1-score de 99.2 %. En la Figura 3.28 se muestra la metodología empleada.

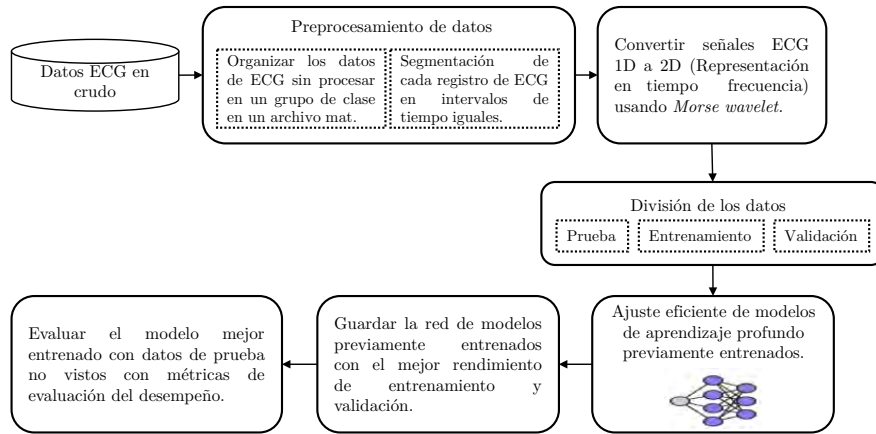


Figura 3.28: Modelo de la metodología general propuesto en [9].

3.3. Discusión del estado del arte

Los enfoques basados en redes neuronales profundas, como CNN y RNN, han sido ampliamente utilizados en la clasificación de arritmias debido a su capacidad para extraer características automáticamente y capturar patrones complejos en señales ECG. Sin embargo, estos modelos requieren grandes volúmenes de datos para su entrenamiento y tienen un alto costo computacional, lo que dificulta su implementación en entornos con recursos limitados. Por otro lado, métodos tradicionales como SVM y k-NN han sido empleados por su menor demanda computacional y facilidad de implementación, aunque dependen de la extracción manual de características, lo que puede afectar su capacidad de generalización en datos con ruido o variabilidad entre pacientes.

Algunos trabajos han explorado enfoques híbridos que combinan técnicas tradicionales de extracción de características con modelos de aprendizaje profundo, buscando mejorar la representación de las señales ECG. Estos métodos han mostrado ventajas en casos con datos limitados o señales de calidad variable, pero pueden requerir preprocesamiento adicional y ajustes específicos, aumentando su complejidad y costo computacional. En general, la literatura refleja un esfuerzo por equilibrar precisión, eficiencia y capacidad de generalización, destacando la necesidad de metodologías que optimicen la clasificación sin depender exclusivamente de grandes volúmenes de datos o arquitecturas computacionalmente costosas.

Tabla 3.1: Resumen del estado del arte.

Título	Objetivo	Algoritmo/ Técnica	Resultados	Utilidad
Método de ingeniería de características basado en aprendizaje profundo y KNN para la clasificación de latidos de ECG y la detección de arritmias [20].	Clasificar latidos cardíacos para la detección de arritmias.	KNN, arboles de decisión, SVM y bosques aleatorios.	Exactitud de 99.77 %, precisión de 99.75 % y sensibilidad de 99.3 %.	Comparación entre varios modelos clasificadores para arritmias cardíacas. para obtener el de menor costo computacional.
Sistema de detección de fibrilación auricular a partir de procesamiento de señales de ECG e inteligencia artificial [32].	Desarrollar un sistema que estime de forma temprana una fibrilación auricular en personas adultas.	SVM, regresión logística, RF, MLP.	Precisión superior a 92 % y modelo embebido a un microprocesador.	Preprocesamiento de señales para eliminación de ruido.
Red convolucional 1D para detectar latidos ventriculares [24].	Desarrollar un clasificador para la detección de latidos de origen ventricular.	CNN.	Sensibilidad de 99 % y tasa predictiva positiva de 96.5 % para detección QRS y sensibilidad de 85.8 % y tasa de predictividad positiva de 64.5 % para detección de latidos ventriculares.	Uso de filtros para mejorar la calidad y clasificación en la salida de una red neuronal.
MultiFusionNet: Detección de fibrilación auricular con redes neuronales profundas [45].	Discriminar entre fibrilación auricular y ritmo cardíaco normal.	Redes neuronales profundas.	F1-score de 90 %.	Procesamiento de señales ECG en crudo.

Título	Objetivo	Algoritmo/ Técnica	Resultados	Utilidad
Reconocimiento automático de arritmias a partir de señales ECG usando diferentes modelos de características de redes neuronales de memoria a largo plazo [31].	Detectar anomalías en el ritmo cardíaco.	LSTM.	96.73 % de precisión, 99.37 % de exactitud, 99.14 % de especificidad, 99.77 % de F1-score y 94.89 % de sensibilidad.	Utilización de diferentes métodos para extracción de características de señales.
Detección automatizada de fibrilación auricular usando una red híbrida CNN-LSTM sobre bases de datos ECG desbalanceadas [32].	Detectar fibrilación auricular, tratar datos desbalanceados.	CNN-LSTM.	Sensibilidad de 97.87 % y especificidad de 99.29 %.	Utilización de un modelo híbrido para tratar el desbalance de datos.
Clasificación de arritmias cardíacas mediante redes neuronales convolucionales y optimización por enjambre de partículas [38].	Detectar y clasificar diferentes arritmias cardíacas.	Redes neuronales convolucionales, Optimización por enjambre de partículas.	Se clasificaron 5 tipos de arritmias cardíacas, siendo los mejores valores para la categoría N, con 97 % de precisión, 100 % para sensibilidad y 99 % para F1-score.	Preprocesamiento de señales y clasificación de arritmias usando una CNN.
Clasificación automática de arritmias cardíacas usando una combinación de redes profundas residuales y memoria a largo plazo bidireccional [13].	Desarrollar un método confiable para la clasificación automática de arritmias cardíacas.	Módulos residuales convolucionales, LSTM bidireccional.	80.6 % de F1-score, 91.4 % de F_{AF} -score, 87.9 % de F_{Block} por bloque, 80.1 % de F_{PC} y 74.2 % de F_{ST} .	Utilización de 2 modelos de redes neuronales profundas para extracción de características de señales ECG.

Título	Objetivo	Algoritmo/ Técnica	Resultados	Utilidad
Diagnóstico de la fibrilación auricular mediante el análisis inteligente de registros intracardíacos [8].	Desarrollar un sistema capaz de detectar distintos tipos de arritmias para apoyo al personal clínico.	<i>Autoencoder</i> Variacional (VAE).	Se logró una precisión del 97.12 %.	Análisis de la característica del punto de cambio de señales de arritmias cardíacas con un <i>Autoencoder</i> variacional.
Detección de arritmias cardíacas mediante un enfoque de aprendizaje profundo y representación de frecuencia de tiempo de señales de ECG [9].	Clasificar 3 tipos de señales ECG.	AlexNet y ResNet.	Exactitud de 99.2 %, especificidad de 99.6 % y sensibilidad, precisión y F1-score de 99.2 %.	Pre-procesamiento y segmentación de señales ECG.
Clasificación de arritmias cardíacas utilizando análisis de componentes principales y redes neuronales pre-alimentadas en segmentos cortos de ECG [27].	Desarrollar un sistema que promueva la brevedad en el análisis de la monitorización electrocardiográfica ambulatoria.	PCA, Perceptrón multicapa.	Exactitud de 80.1 %.	Uso de PCA sobre señales ECG.
Un nuevo método de aumento de datos para mejorar las redes neuronales profundas para la detección de la fibrilación auricular [6].	Aumento de muestras de señales ECG.	RNN-LSTM.	F1-score de 92 %.	Aumento de datos en datasets de señales ECG para entrenamiento de modelos de aprendizaje profundo.

Título	Objetivo	Algoritmo/ Técnica	Resultados	Utilidad
Un enfoque de aprendizaje profundo para la detección en tiempo real de la fibrilación auricular [1].	Extraer características y clasificar arritmias cardíacas.	CNN y RNN.	Sensibilidad de 98.98 % y especificidad de 96.95 %.	Detección de fibrilación auricular en tiempo real.
Modelo híbrido basado en redes neuronales, sistemas difusos tipo 1 y 2 para clasificación de arritmias cardíacas de 2 derivaciones [34].	Clasificar arritmias cardíacas de 2 derivaciones.	Lógica difusa, KNN, MLP.	Tasa de clasificación de 92.9 % para el módulo 1 y 92.7 % para el módulo 2.	Preprocesamiento de señales ECG para obtención de características y prueba entre diferentes clasificadores.
Una clasificación mejorada de arritmias cardíacas usando un intervalo RR [33].	Clasificación de arritmias cardíacas.	DWT, KNN, SVM, DT, NB, RF.	El mejor resultado fue para SVM con una Exactitud de 99.51 %, sensibilidad de 99.28 % y especificidad de 99.63 %.	Uso de filtros para remover ruido de alta frecuencia en señales ECG.
Clasificación precisa de arritmias de ECG utilizando redes de aprendizaje profundo de compresión rápida mejoradas MOWPT [18].	Clasificación de arritmias cardíacas.	CNN, MOWPT, FCResNet.	Exactitud de 98.79 %.	Utilización de un modelo con bajo costo computacional para la clasificación de arritmias cardíacas.
Agrupamiento de señales ECG con rasgos aprendidos usando un <i>Autoencoder</i> profundo [46].	Hallar grupos de señales ECG como apoyo al especialista médico para facilitar el diagnóstico de epilepsia.	<i>Autoencoder</i> , k-means, SVM.	Margen de error de 0.2 % para el mejor de los casos y 1.47 % para el peor.	Extracción de características usando un <i>Autoencoder</i> .

Título	Objetivo	Algoritmo/ Técnica	Resultados	Utilidad
Detección de la fibrilación auricular y clasificación ECG basado en una red CNN-BiLSTM [48].	Clasificar ritmos cardíacos normales y anormales.	CNN, RNN- BiLSTM.	F1-score de 82 %.	Comparación de modelos para la clasificación de ritmos cardíacos.
Detección robusta de fibrilación auricular a partir de ECG de corta duración mediante redes neuronales convolucionales [29].	Mejorar un modelo existente de clasificación de FA, identificar FA.	DWT, CNN.	Exactitud de 99.98 %, sensibilidad de 99.91 %, especificidad de 99.91 %, precisión de 99.99 % y F1-score de 99.95 %.	Utilización de la base de datos MIT-BIH complementada con muestras de un hospital. Normalización, eliminación de ruido y segmentación de señales ECG.
Un método novedoso híbrido de aprendizaje profundo con algoritmo de búsqueda de <i>cuckoo</i> para la clasificación de enfermedades de arritmias mediante señales de ECG [39].	Clasificación y detección de anomalías en los latidos del corazón.	DWT, algoritmo <i>cuckoo</i> , SVM, FFBPNN.	Precisión de 98.247 % y sensibilidad de 95.68 %.	Eliminación del ruido en señales de ECG utilizando la transformada discreta de onda.
Enfoque novedoso de aprendizaje profundo para la clasificación de fibrilación auricular basado en una red neuronal convolucional y modificada de Elman [47].	Detectar de forma automática arritmias cardíacas.	CNN.	Exactitud de 97.4 %, sensibilidad de 97.9 % y especificidad de 97.1 %.	Utilización de una arquitectura apilada entre una CNN y una red neuronal Elman para la clasificación de fibrilación auricular.

Título	Objetivo	Algoritmo/ Técnica	Resultados	Utilidad
Clasificación de arritmias mediante aprendizaje profundo y aprendizaje automático con funciones extraídas del procesamiento de señales basado en formas de onda [47].	Clasificar arritmias cardíacas.	WBSP.	Exactitud de 98.8 % y sensibilidad de 96.3 %.	Extracción de características de señales ECG en pre-procesamiento.
CardioNet: un sistema eficiente para la clasificación de arritmias usando <i>Transfer learning</i> [30].	Clasificación de latidos del corazón.	DenseNet, ImageNet.	Exactitud de 98.92 %.	Utilización de <i>Transfer learning</i> para entrenamientos de modelos de aprendizaje profundo.
Detección de fibrilación auricular usando una red neuronal convolucional 1d [16].	Detectar la fibrilación auricular.	CNN.	F1-score de 78.2 %.	Utilización de un algoritmo para normalizar las señales de ECG.
Clasificación de arritmias cardíacas mediante funciones basadas en transformadas <i>Q-wavelet</i> sintonizables y en un clasificador de máquina de soporte vectorial [19].	Clasificación de señales ECG.	<i>Q-wavelet</i> , SVM.	Exactitud de 99.27 %, sensibilidad de 96.22 %, especificidad de 99.58 %.	Descomposición de las señales de ECG en 6 niveles de la onda Q para extracción de características.

4 Metodología de Solución

Para cumplir con el objetivo de esta investigación, se propone la metodología mostrada en la Figura 4.1, la cuál abarca desde el preprocesamiento de las señales ECG, hasta la transformación de las señales al dominio espectral mediante distintas técnicas de estimación espectral. Posteriormente, se emplea una arquitectura basada en un *Autoencoder*, cuya función principal es la codificación de características representativas y su posterior clasificación mediante una capa totalmente conectada. La inclusión de esta última permite extender la funcionalidad del *Autoencoder* tradicional. A continuación se dan más detalles de la metodología propuesta.

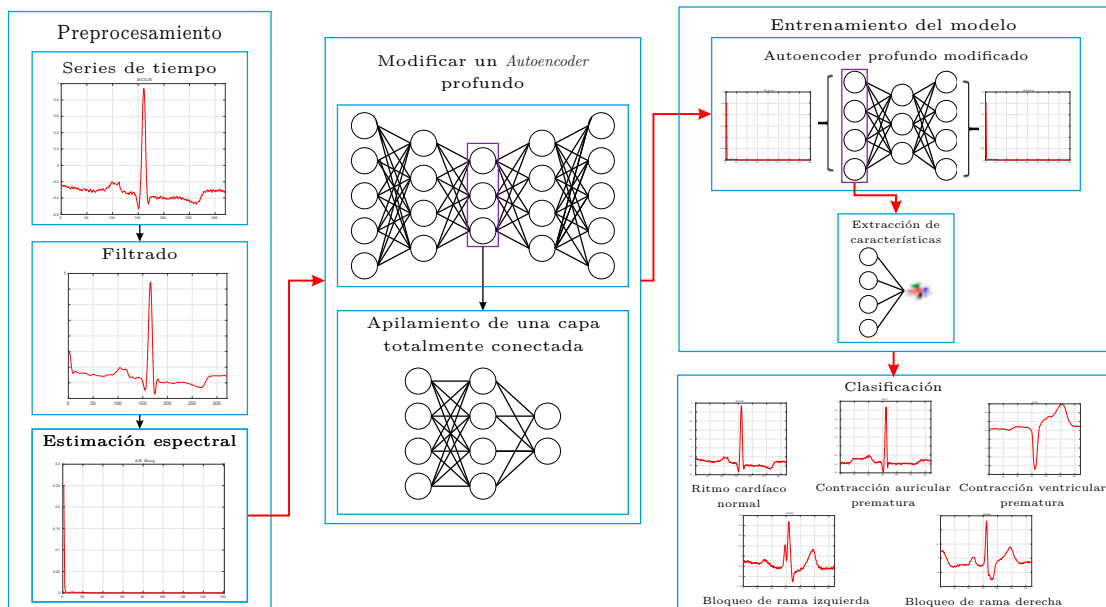


Figura 4.1: Metodología de solución propuesta

4.1. Preprocesamiento

El preprocesamiento es una etapa crucial para mejorar la calidad de las señales ECG antes de su análisis. En esta fase, las señales en crudo son sometidas a diferentes procedimientos para eliminar ruido y resaltar las características más relevantes. Se aplica un filtro pasa-bajo con frecuencia de corte de 35 Hz, eliminando interferencias de alta frecuencia provenientes de fuentes externas, como el ruido muscular o interferencias eléctricas.

Posteriormente, las señales son transformadas al dominio de la frecuencia utilizando seis métodos de estimación espectral: *FFT*, *Periodograma*, *Welch*, *AR Burg*, *Yule-Walker* y *Covariance*. Cada uno de estos métodos proporciona una representación única de la señal ECG, permitiendo extraer información relevante sobre la distribución de frecuencia de cada clase de arritmia. La conversión al dominio espectral permite reducir la dimensionalidad de los datos y resaltar patrones característicos que pueden mejorar la capacidad de clasificación del modelo.

Una vez obtenidas las estimaciones espectrales, los datos son normalizados para estandarizar las características y evitar sesgos en el entrenamiento del modelo. Esta normalización garantiza que todas las características tengan la misma escala, facilitando la convergencia del modelo y evitando dominancias de ciertos valores en la optimización.

4.2. Modificación de *Autoencoder*

Dado que los *Autoencoders* no están diseñados para realizar tareas de clasificación, fue necesario modificar su arquitectura para adaptarlo a este propósito. En este trabajo, se tomó como base la arquitectura tradicional de un *Autoencoder* y se le añadió una capa totalmente conectada en la sección del codificador. Esta modificación permite que las características extraídas sean utilizadas directamente como entrada de una función *softmax*, la cual es responsable de la clasificación en múltiples categorías. La Figura 4.2 muestra la arquitectura propuesta, donde se pueden identificar sus principales componentes y su funcionamiento general.

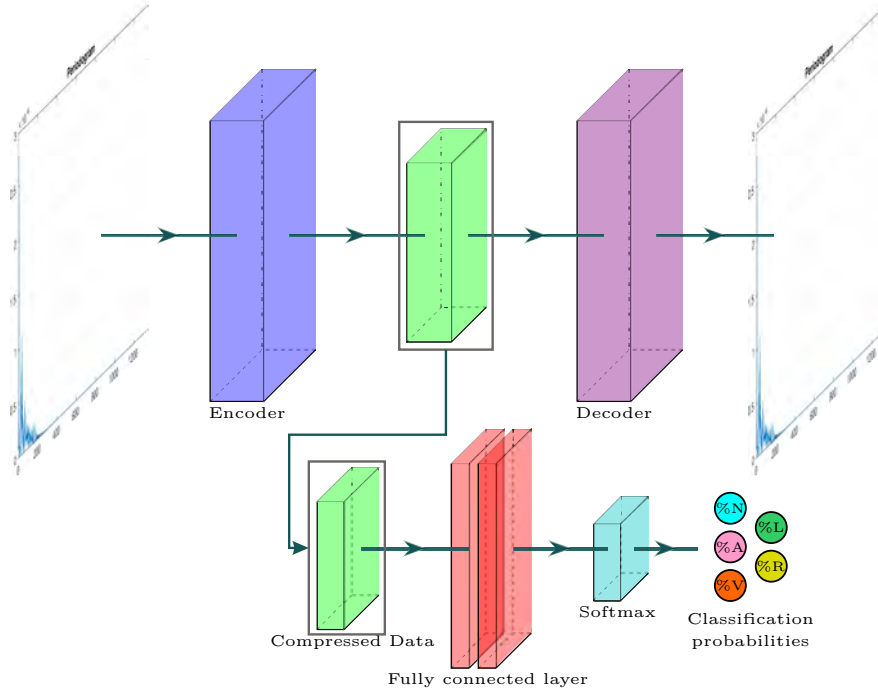


Figura 4.2: Arquitectura propuesta para realizar la tarea de clasificación

Esta modificación permite que el modelo no solo aprenda una representación comprimida de la señal ECG, sino que también la utilice para realizar inferencias en la clasificación de distintos tipos de arritmias. A continuación se da la definición detallada del modelo propuesto.

4.2.1. Descripción General de la Arquitectura

La arquitectura está compuesta por los siguientes módulos:

- **Codificador:** Extrae características esenciales de las señales de entrada y las comprime en un espacio latente de menor dimensión.
- **Capa intermedia de datos comprimidos:** Representación latente de la señal, que conserva la información relevante para la clasificación.
- **Capa totalmente conectada:** Se encarga de proyectar las características latentes hacia el espacio de clasificación.
- **Función *softmax*:** Asigna probabilidades a cada clase de arritmia basada en los valores obtenidos en la capa anterior.

- **Decodificador:** Se usa únicamente durante el entrenamiento para reconstruir la señal original y mejorar la capacidad de representación del *Autoencoder*.

4.2.2. Descripción matemática

Como se puede observar en la Figura 4.2, se trata de una arquitectura simple que no cuenta con muchas capas ocultas, lo cual garantiza un menor costo computacional en comparación con otros modelos. Los componentes que conforman esta arquitectura son: una capa de codificación y decodificación con el mismo tamaño que los datos de entrada; desde el punto de vista matemático, esto se define como $x = \{x_i \mid i = 0, \dots, n\}$ y $\hat{x} = \{\hat{x}_j \mid j = 0, \dots, n\}$, respectivamente, donde n es la longitud de los datos de entrada. Además, incluye una capa oculta de ψ neuronas, equivalentes a las características que el *Autoencoder* extraerá, representada por z , este proceso se define como:

$$z = \sigma(W \cdot x + b) , \quad (4.1)$$

en donde x es la señal de entrada, W son los pesos iniciales, b son los sesgos iniciales y σ es una función de activación. El propósito de esta transformación es reducir la dimensionalidad de la señal original y extraer características relevantes que faciliten la clasificación. En esta arquitectura, la representación latente está compuesta por un vector de características z , el cual preserva la información más discriminante de la señal ECG y se define mediante: $z = \{z_l \mid l = 1, \dots, \psi\}$.

4.2.3. Clasificación basada en las características codificadas

Una vez extraídas las características representativas, estas se utilizan como entrada para una capa totalmente conectada, la cual proyecta la información en el espacio de clasificación. La salida de esta capa es:

$$y = \hat{w} \cdot z + \hat{b} , \quad (4.2)$$

donde W_s y b_s son los pesos y sesgos de la capa de clasificación y y es el vector de activaciones previo a la función *softmax*. Esta capa contiene 5 neuronas, dado que se trabaja con cinco clases distintas. Con base en lo anterior, la clasificación multiclase es realizada por la función de activación *softmax* debido a que es ideal para la clasificación multiclase. Esta recibirá las activaciones previamente obtenidas y las convertirá en probabilidades, Este proceso se define como:

$$\sigma(z)_j = \frac{e^{x_j}}{\sum_{k=1}^K e^{z_K}} , \quad (4.3)$$

donde $\sigma(z)_j$ representa la salida de la neurona j tras la aplicación de la función *softmax*, z_j es la entrada a la neurona j , y K es el número total de clases en el problema de clasificación. Con lo anterior explicado, el proceso de codificación y clasificación realizado por un *Autoencoder* modificado se desarrolla en los Algoritmos 1 y 2.

Algorithm 1: Codificación y decodificación

Entrada: X , max_epoch , α

Salida: z

$\Theta \leftarrow random([0, 1])$

for $epoch = 1$ to max_epoch **do**

$z \leftarrow \sigma(W \cdot x + b)$

$\hat{x} \leftarrow \sigma'(W' \cdot z + b')$

$\ell(x, \hat{x}) \leftarrow \|x - \hat{x}\|^2$

$\nabla W \leftarrow \frac{\partial \ell}{\partial W}$

$\nabla b \leftarrow \frac{\partial \ell}{\partial b}$

$\nabla W' \leftarrow \frac{\partial \ell}{\partial W'}$

$\nabla b' \leftarrow \frac{\partial \ell}{\partial b'}$

$W \leftarrow W - \alpha \cdot \nabla W$

$b \leftarrow b - \alpha \cdot \nabla b$

$W' \leftarrow W' - \alpha \cdot \nabla W'$

$b' \leftarrow b' - \alpha \cdot \nabla b'$

Regresar: z

El algoritmo anterior corresponde al proceso básico que realiza un *Autoencoder*, considerando las etapas de codificación y decodificación. Este algoritmo recibe como entrada el parámetro X , que hace referencia a una señal ECG; max_epoch , que indica el número de épocas de entrenamiento; y α , que representa la tasa de aprendizaje. Además, se inicializan aleatoriamente los parámetros de Θ , que incluyen los pesos y sesgos correspondientes (W, b, W', b').

Partiendo de estos datos, se inicia el proceso de codificación y extracción de características mediante un bucle de tamaño equivalente al número de épocas, donde se calcularán z y $\hat{x}$, correspondientes a las etapas de codificación y decodificación. En este proceso se emplean las funciones de activación σ (*satlin*) en la capa de codificación y $\hat{\sigma}$ (*purelin*) en la capa de decodificación para calcular los valores respectivos; posteriormente, se minimiza el error de reconstrucción ℓ entre x y $\hat{x}$.

Adicionalmente, los pesos se actualizan mediante retropropagación y gradiente descendente, ajustando los pesos y sesgos del *Autoencoder* para reducir la pérdida de reconstrucción en cada iteración. Al finalizar este proceso, se retorna z , que corresponde a las características extraídas (codificadas) de cada valor X , las cuales se utilizarán en la capa de clasificación para inferir las clases correspondientes. El proceso descrito anteriormente se muestra en el Algoritmo 2.

Algorithm 2: Proceso de clasificación

Entrada: z , max_epoch , α

Salida: $\hat{y}$

$W_s, b_s \leftarrow random([0, 1])$

for $epoch = 1$ to max_epoch **do**

$\hat{y} \leftarrow softmax(W_s \cdot z + b_s)$

$\ell_{class} \leftarrow cross_entropy(y, \hat{y})$

$\nabla W_s \leftarrow \frac{\partial \ell}{\partial W_s}$

$\nabla b_s \leftarrow \frac{\partial \ell}{\partial b_s}$

$W_s \leftarrow W_s - \alpha \cdot \nabla W_s$

$b_s \leftarrow b_s - \alpha \cdot \nabla b_s$

Regresar: $\hat{y}$

Como se observa en el algoritmo anterior, tras la codificación de características, se pasa a la etapa de clasificación en la capa central. Este algoritmo requiere tres parámetros de entrada: las características extraídas z , el número deseado de épocas de entrenamiento max_epochs y la tasa de aprendizaje α . Además, es necesario inicializar los pesos y sesgos.

A continuación, se ejecuta un bucle de tamaño max_epochs en el cual se emplean las características z extraídas previamente, los pesos W_s , los sesgos b_s y una función de activación *softmax* para generar la probabilidad de pertenencia a las clases para cada señal analizada. A partir de aquí, se utiliza la entropía cruzada para calcular la discrepancia entre las predicciones del modelo y las etiquetas reales. Tras este proceso, se ajustan los pesos y sesgos en la capa de clasificación para minimizar el error de clasificación, utilizando nuevamente la retropropagación y el gradiente descendente.

Al finalizar el bucle correspondiente, se retorna el parámetro $\hat{y}$, que representa la predicción realizada por el modelo. Luego de modificar la arquitectura, es necesario evaluar el desempeño del modelo mediante entrenamientos, pruebas y validación de los resultados utilizando métricas de calidad.

4.2.4. Justificación de la arquitectura

El diseño de la arquitectura se basó en la necesidad de optimizar tanto la extracción de características como la clasificación de señales ECG con un bajo costo computacional. En comparación con modelos más complejos como redes LSTM o CNNs profundas, esta arquitectura ofrece una alternativa eficiente que conserva la información relevante mediante la representación latente del *Autoencoder*. Además, la inclusión de una capa totalmente conectada permite evitar la necesidad de modelos adicionales de clasificación, simplificando la estructura del modelo sin comprometer su precisión. Otro aspecto clave de la arquitectura propuesta es la separación del entrenamiento en dos fases:

- **Primera etapa:** Entrenamiento del *Autoencoder* tradicional, siguiendo el proceso correspondiente de ajustar pesos y sesgos para minimizar el error en la reconstrucción de la señal de entrada y extraer características más discriminantes.

- **Segunda etapa:** Se incorpora una capa de clasificación correspondiente a un perceptrón simple el cuál es entrenado con las características extraídas para mejorar la clasificación.

Los puntos claves de este enfoque son:

- **Reducción de dimensionalidad:** La capa oculta del *Autoencoder* comprime los datos de entrada a una menor dimensión, permitiendo realizar cálculos y análisis más rápido.
- **Menor costo computacional:** Al no depender de modelos profundos, el entrenamiento y la inferencia son más rápidos y eficientes.
- **Capacidad de generalización:** La arquitectura es flexible y puede adaptarse a distintas bases de datos sin modificaciones estructurales significativas.
- **Integración de clasificación con extracción de características:** En lugar de utilizar un *Autoencoder* solo como un extractor de características previo a otro modelo, la clasificación se realiza en la misma arquitectura, evitando dependencias externas.

4.2.5. Consideraciones para mejoras

Si bien la arquitectura propuesta logra un balance entre eficiencia y precisión, es posible mejorarla mediante:

- **Ajuste de hiperparámetros:** Explorar diferentes configuraciones en el número de neuronas en la capa oculta y la tasa de aprendizaje.
- **Regularización y técnicas de dropout:** Para evitar sobreajuste y mejorar la capacidad de generalización.

En conclusión, la arquitectura propuesta representa un diseño eficiente y optimizado para la clasificación de señales ECG. Su enfoque basado en la transformación espectral, reducción de dimensionalidad mediante *Autoencoder* y clasificación con *softmax* permite una clasificación precisa con un menor costo computacional en comparación con modelos más complejos.

4.3. Entrenamiento del modelo

Para garantizar que el modelo alcance su máximo rendimiento, se llevó a cabo un proceso exhaustivo de sintonización de hiperparámetros. El entrenamiento se realizó utilizando la base de datos MIT-BIH, empleando una división de datos de 90 % para entrenamiento y 10 % para prueba. La optimización se llevó a cabo mediante el algoritmo de Estimación del momento adaptativo (Adam) con *backpropagation*, asegurando que la red ajustara sus pesos de manera eficiente a lo largo del entrenamiento. El proceso de propagación del error mediante el algoritmo Adam se muestra en el Algoritmo 3.

Algorithm 3: Optimización mediante el algoritmo Adam

Entrada: $\Theta_0, \alpha, \beta_1, \beta_2, \epsilon, max_epochs$ **Salida:** Θ

```
 $m_0 \leftarrow 0$  ▷ Inicializar primer momento  
 $v_0 \leftarrow 0$  ▷ Inicializar segundo momento  
 $t \leftarrow 0$   
  
for  $epoch = 1$  to  $max\_epoch$  do  
   $t \leftarrow t + 1$   
   $g_t \leftarrow \nabla_{\Theta} f(\Theta_{t-1})$  ▷ Calcular gradiente  
   $m_t \leftarrow \beta_1 \cdot m_{t-1} + (1 - \beta_1) \cdot g_t$  ▷ Actualizar primer momento  
   $v_t \leftarrow \beta_2 \cdot v_{t-1} + (1 - \beta_2) \cdot g_t^2$  ▷ Actualizar segundo momento  
   $\hat{m}_t \leftarrow \frac{m_t}{1 - \beta_1^t}$  ▷ Corrección del sesgo del primer momento  
   $\hat{v}_t \leftarrow \frac{v_t}{1 - \beta_2^t}$  ▷ Corrección de sesgo del segundo momento  
   $\Theta_t \leftarrow \Theta_{t-1} - \alpha \cdot \frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \epsilon}$  ▷ Actualizar parámetros
```

Regresar: Θ_t

El algoritmo anterior corresponde un método de ajuste de parámetros basado en el descenso de gradiente adaptativo, combinando las ventajas de Momentum y RMSprop para mejorar la estabilidad y eficiencia del entrenamiento. En cada iteración, se calcula el gradiente de la función de pérdida $g_t = \nabla_{\Theta} f(\Theta_{t-1})$, determinando la dirección en la que deben actualizarse los parámetros del modelo. Para lograr un ajuste más preciso, Adam introduce dos momentos: el primer momento, m_t que corresponde a un promedio móvil exponencial del gradiente actual y los gradientes anteriores, y el segundo momento v_t que estima la varianza del gradiente acumulado. Estos momentos permiten adaptar la magnitud de la actualización en función de la variabilidad de los gradientes, reduciendo fluctuaciones bruscas y acelerando la convergencia.

Dado que en las primeras iteraciones los momentos tienden a estar sesgados hacia valores cercanos a cero, Adam aplica una corrección del sesgo, definiendo $\hat{m}_t = \frac{m_t}{1 - \beta_1^t}$ y $\hat{v}_t = \frac{v_t}{1 - \beta_2^t}$, asegurando valores más estables en las primeras etapas del entrenamiento. Una vez corregidos los momentos, la actualización de los parámetros Θ_t se realiza mediante la ecuación:

$$\Theta_t = \Theta_{t-1} - \alpha \cdot \frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \epsilon}, \quad (4.4)$$

donde α es la tasa de aprendizaje, que controla la magnitud el ajuste en cada iteración, y ϵ es un pequeño valor que evita divisiones por 0. Este proceso se repite durante un número determinado de épocas max_epochs o hasta que la función de pérdida alcance un criterio de convergencia. Gracias a su capacidad de ajustar dinámicamente la tasa de aprendizaje para cada parámetro, Adam ofrece una optimización eficiente y estable, siendo una de las técnicas más utilizadas en el entrenamiento de redes neuronales profundas.

Aunado a lo anterior, se implementó validación cruzada para evaluar la estabilidad del modelo y evitar sobreajuste. Adicionalmente, se compararon los resultados de diferentes espectros para determinar cuál proporcionaba la mejor representación para la clasificación de arritmias.

4.4. Clasificación

La fase de clasificación se basa en la extracción de características latentes a partir de representaciones espectrales de señales ECG por medio del *Autoencoder* y posteriormente su procesamiento mediante una capa completamente conectada con activación *softmax*. A diferencia de un *Autoencoder* tradicional, que solo busca reconstruir la entrada, en este caso se modifica su estructura para permitir la clasificación multiclase. Este proceso se compone de varias etapas, que incluyen la transformación de la señal, la codificación de características, la propagación a través de una capa de clasificación y la obtención de la predicción final.

4.4.1. Transformación y normalización de la señal

Cada señal ECG preprocesada, representada como un vector en el dominio temporal definido por $x \in \mathbb{R}^n$, es transformada al dominio de la frecuencia mediante una de las seis técnicas de estimación espectral implementadas en la metodología. Los espectros obtenidos se normalizan para asegurar que todas las entradas del modelo se encuentren en la misma escala, evitando sesgos en el entrenamiento. Lo anterior se define como:

$$x' = \frac{x - \mu}{\sigma}, \quad (4.5)$$

en donde x' es la señal normalizada, μ es la media de la señal y σ es la desviación estándar.

4.4.2. Codificación de características latentes

La señal transformada x se introduce en el *Autoencoder*, el cual extrae un conjunto de características relevantes y las representa en un espacio de menor dimensionalidad. La transformación de entrada a representación latente está definida por:

$$z = \sigma(W_e \cdot x' + b_e), \quad (4.6)$$

en donde z es el vector de características latentes, W_e y b_e son los pesos y sesgos de la capa de codificación y $\sigma(\cdot)$ representa la función de activación utilizada en la capa oculta.

4.4.3. Clasificación mediante capa completamente conectada

Una vez obtenida la representación latente z , esta se propaga a través de una capa totalmente conectada que proyecta los datos en un espacio de clasificación de K dimensiones, donde K es el número total de clases. La salida de esta capa está definida por:

$$y = W_c \cdot z + b_c, \quad (4.7)$$

donde y es el vector de activaciones previas a la clasificación y W_c y b_c son los pesos y sesgos de la capa de clasificación. Para convertir estas activaciones en probabilidades, se aplica la función de activación *softmax*, la cuál se define como:

$$\hat{y}_j = \frac{e^{y_j}}{\sigma_{k=1}^K e^{y_k}} , \quad (4.8)$$

en donde $\hat{y}_j$ representa la probabilidad de que la señal ECG pertenezca a la clase j , e^{y_j} es la exponencial del valor de activación de la neurona j y $\sigma_k^K = 1e^{y_k}$ es la suma de todos los exponenciales, lo que garantiza una salida de 1.

4.4.4. Evaluación del modelo

Para evaluar el desempeño del modelo, se realiza un análisis basado en métricas de clasificación comúnmente utilizadas en el estado del arte. Estas métricas permiten cuantificar la precisión del modelo al asignar una clase a cada señal ECG y evaluar su capacidad para diferenciar correctamente entre distintos tipos de arritmias. La evaluación se lleva a cabo comparando las predicciones obtenidas con las etiquetas reales del conjunto de prueba, lo que proporciona una medida objetiva del rendimiento del modelo. Además, se registra el tiempo de inferencia con el propósito de analizar la eficiencia computacional y su viabilidad para su implementación en entornos clínicos en tiempo real.

4.5. Discusión

Este capítulo presenta la metodología propuesta para la clasificación automática de arritmias cardíacas, estableciendo un flujo de procesamiento que abarca desde la preparación de los datos hasta su transformación y posterior análisis en una arquitectura basada en un *Autoencoder*. El preprocesamiento de las señales ECG se centra en la aplicación de un filtro pasa-bajo para reducir el ruido de alta frecuencia y garantizar que la información relevante se conserve antes de ser utilizada en el modelo. Posteriormente, se implementa una transformación al dominio de la frecuencia mediante seis métodos distintos de estimación espectral, con el objetivo de representar las señales de manera más estructurada y capturar patrones que puedan facilitar la diferenciación entre clases y posteriormente comparar la eficiencia de cada método.

A partir de estas representaciones espectrales, la metodología propone una arquitectura basada en un *Autoencoder* modificado, diseñado para extraer características relevantes de las señales ECG y optimizar su representación en un espacio de menor dimensionalidad. Esta arquitectura incorpora una capa totalmente conectada en la fase de codificación, permitiendo que las características latentes sean utilizadas directamente en un proceso de clasificación multiclase mediante una función *softmax*. Además, se establecen los criterios de evaluación del modelo mediante métricas de calidad utilizadas en el estado del arte. En conjunto, la metodología propuesta proporciona una estructura clara y eficiente para el procesamiento de señales ECG, asegurando que la arquitectura implementada sea capaz de manejar las representaciones espectrales y facilitar la clasificación de arritmias de manera óptima.

5 | Experimentación y Resultados

En esta sección se detallan las actividades realizadas y los logros alcanzados durante el desarrollo de este trabajo. Se describen los datos utilizados, los métodos de preprocesamiento aplicados, las modificaciones implementadas en la arquitectura del modelo de *Autoencoder*, la experimentación realizada, así como los resultados obtenidos y las métricas de evaluación empleadas para medir el rendimiento del modelo.

5.1. Entorno de desarrollo

Como se mencionó previamente, el objetivo de este trabajo fue desarrollar una arquitectura eficiente en términos de recursos computacionales, capaz de realizar la clasificación de señales ECG. Partiendo de esta premisa, no fue necesario contar con un equipo de cómputo con componentes dedicados, como GPU's, para realizar los entrenamientos, aunque estos componentes pueden acelerar los procesos si están disponibles. Para el desarrollo del código, se optó por utilizar un software que proporcionara las herramientas necesarias para la creación de la red. A continuación, se presenta un resumen del equipo de cómputo y las herramientas utilizadas:

Hardware

- Procesador Intel Core i5-12400f de 6 núcleos.
- 16 Gb de memoria RAM.
- Disco de estado solido de 500Gb.

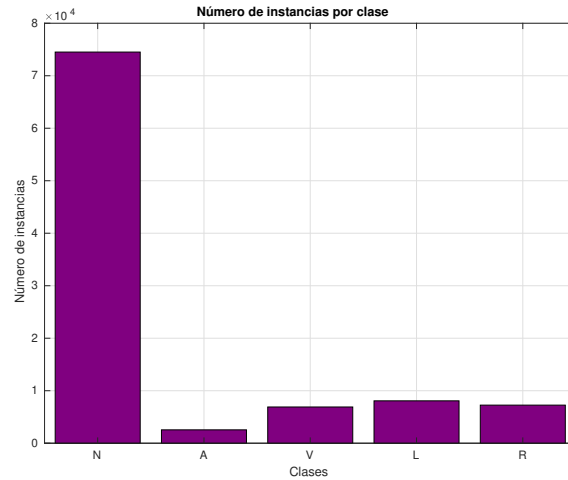
Software

- Sistema operativo Ubuntu 22.04.4 LTS.
- MATLAB 2024a

Tras establecer el entorno de desarrollo adecuado para implementar el modelo de clasificación, el siguiente paso consistió en preparar los datos necesarios para su entrenamiento y validación. Este proceso incluyó la selección de una base de datos adecuada, el análisis de su composición y características, así como su división estratégica en conjuntos de entrenamiento y prueba. Estas acciones son fundamentales para garantizar que el modelo pueda ser evaluado de manera rigurosa y bajo condiciones que permitan medir su desempeño en escenarios representativos de la realidad clínica.

5.2. Gestión del *Dataset*

La experimentación correspondiente se llevó a cabo utilizando la base de datos MIT-BIH, debido a su frecuente utilización en la literatura, marcando una línea base para la posterior comparación de resultados con investigaciones previas. Esta base de datos contiene muestras de diferentes clases de arritmias cardíacas y señales de ritmo cardíaco normal. Las señales están registradas a una frecuencia de muestreo de 360 Hz. Se realizó un análisis exploratorio de los datos y se verificó el número de instancias por clase; este proceso se ilustra en la Figura 5.1.



• Figura 5.1: Histograma de las instancias por clase

Como se observa, existe un desbalance notorio entre las clases, los datos cuantitativos se muestran en la Tabla 5.1.

Tabla 5.1: Resumen cuantitativo de la base de datos

Clase	Instancias
Normal	74512
Contracción Prematura Auricular	2546
Contracción Prematura Ventricular	6902
Bloqueo de Rama Izquierda	8072
Bloqueo de Rama Derecha	7257

Para realizar un balanceo de clases, se optó por hacer un recorte de datos a la cantidad de instancias de la clase menor, teniendo un total de 2546 instancias por clase. Una vez realizado el proceso anterior, se dividió el dataset en 90 % de los datos para entrenamiento y 10 % para pruebas, tal como se muestra en la Tabla 5.2.

Tabla 5.2: Cantidad de datos para entrenamiento y prueba

Clase	Datos de Entrenamiento	Datos de Prueba
Normal	2291	255
Contracción Prematura Auricular	2291	255
Contracción Prematura Ventricular	2291	255
Bloqueo de Rama Izquierda	2291	255
Bloqueo de Rama Derecha	2291	255

Una vez se realizó la separación de los datos en conjuntos de entrenamiento y prueba, se procedió a evaluar el modelo en diferentes condiciones, para ello se utilizaron métricas de calidad estándar utilizadas en el estado del arte.

5.3. Métricas de evaluación

Para evaluar el resultado del modelo propuesto y comparar con otros modelos en el estado del arte, se contemplan las métricas mencionadas en la literatura: exactitud, precisión, sensibilidad, f1-score, coeficiente de matthews y coeficiente de kappa. A continuación se detallan las métricas consideradas:

- **Matriz de confusión:** La matriz de confusión es una herramienta utilizada en el campo de la estadística y el aprendizaje automático para evaluar el desempeño de un modelo de clasificación. Proporciona una forma visual de observar el rendimiento del algoritmo, mostrando las predicciones correctas e incorrectas en comparación con las observaciones reales [36]. Su estructura es la siguiente:

Tabla 5.3: Estructura de una matriz de confusión

VP	FN
FP	VN

en donde VP son los datos clasificados correctamente en la clase positiva, FN son los datos clasificados erróneamente como negativos, FP son los valores clasificados erróneamente como positivos y VN son los datos clasificados correctamente como negativos.

- **Exactitud:** Es la proporción de todas las predicciones correctas respecto al total de predicciones. Es útil cuando las clases están equilibradas, pero puede ser inconsistente si hay una gran desproporción entre las clases [12]. Se calcula mediante:

$$Exactitud = \frac{VP + VN}{VP + VN + FP + FN}. \quad (5.1)$$

- **Precisión:** Es el porcentaje de verdaderos positivos sobre el total de predicciones positivas. Es útil cuando el costo de los falsos positivos es alto [12] [15]. Se calcula mediante:

$$Precisión = \frac{VP}{VP + FP}. \quad (5.2)$$

- **Sensibilidad:** Es el porcentaje de verdaderos positivos sobre el total de positivos reales. Es útil cuando el costo de los falsos negativos es alto [40]. Se calcula mediante:

$$Sensibilidad = \frac{VP}{VP + FN}. \quad (5.3)$$

- **F1-score:** Es la media armónica de la precisión y la sensibilidad. Es útil para un equilibrio entre precisión y sensibilidad, especialmente con clases desbalanceadas [12]. Se calcula mediante:

$$F1 - score = \frac{Exactitud * Recall}{Exactitud + Recall}. \quad (5.4)$$

- **Coeficiente de Matthews:** Es una medida de la calidad de la clasificación binaria que toma en cuenta verdaderos y falsos positivos y negativos. Es útil cuando las clases están desbalanceadas y se busca una métrica equilibrada que refleje todas las categorías [12]. Se calcula mediante:

$$Coeficiente\ de\ Matthews = \frac{(VP * VN) - (FP * FN)}{\sqrt{(VP + FP) * (VP + FN) * (VN + FP) * (VN + FN)}}. \quad (5.5)$$

- **Coefficiente de Kappa:** Mide la concordancia entre dos evaluadores, teniendo en cuenta la concordancia que podría ocurrir por azar. Es útil para evaluar la fiabilidad de la clasificación y para comparaciones inter-evaluadores [26]. Se calcula mediante:

$$\text{Coeficiente de Kappa} = \frac{p_o - p_e}{1 - p_e}, \quad (5.6)$$

en donde p_o es la proporción observada de precisión y p_e es la proporción esperada de acuerdo por azar. Posterior a la explicación de las métricas a utilizar, es necesario evaluar el desempeño del modelo propuesto, para ello es necesario realizar múltiples entrenamientos y pruebas en busca de los mejores resultados. Aunado a lo anterior, para la comparación de modelos con el estado del arte, se tomó en cuenta el tiempo de inferencia para evaluar el costo computacional de los modelos.

5.4. Experimentación y Resultados

En esta sección se presentan los experimentos realizados utilizando la arquitectura propuesta, así como detalles adicionales acerca del entorno de desarrollo y experimentación y de los datos utilizados. Tomando como base la metodología de solución mostrada en la Figura 4.1, se condujeron una serie de experimentaciones con la finalidad de probar el desempeño del modelo propuesto usando diferentes criterios y bajo diferentes condiciones, evaluando su evolución. Los objetivos de la serie de experimentaciones se detallan en los siguientes puntos:

- Evaluar el rendimiento del modelo al ser entrenado con hiperparámetros por defecto y utilizando señales en crudo.
- Evaluar el rendimiento del modelo al ser entrenado con hiperparámetros por defecto y utilizando señales filtradas.
- Evaluar el rendimiento del modelo al ser realizar una sintonización hiperparamétrica utilizando señales filtradas.
- Evaluar el rendimiento del modelo al ser entrenado con una sintonización hiperparamétrica utilizando estimaciones espectrales
- Evaluar el rendimiento del modelo al aplicarlo a otras bases de datos.
- Comparar el rendimiento del modelo con otros modelos neuronales usados en el estado del arte.

5.4.1. Experimento 1: Entrenamiento del *Autoencoder* con hiperparámetros por defecto/Señales en crudo

Condiciones del experimento

El Experimento 1 se diseñó con el objetivo de evaluar el desempeño inicial del modelo utilizando señales ECG en crudo, sin aplicar técnicas de preprocesamiento. Esto permite establecer una línea base para los resultados y comparar la influencia de futuros ajustes, como el filtrado de ruido o la extracción de características. Para este entrenamiento, se emplearon los hiperparámetros por defecto del modelo de *Autoencoder*, definidos de la siguiente manera:

- 100 épocas de entrenamiento.
- Capa oculta de 20 neuronas.
- Regularización L2 = 0.001.
- Regularización de dispersión = 1.
- Proporción de dispersión = 0.5.
- Función *logsig* para la codificación y decodificación.

Datos Utilizados

Se utilizó el conjunto de datos balanceado previamente descrito, compuesto por cinco clases de ritmos cardíacos (ritmo normal, contracción auricular prematura, contracción ventricular prematura, bloqueo de rama izquierda y bloqueo de rama derecha), con 2291 instancias para entrenamiento y 255 para pruebas en cada clase. Las señales fueron procesadas en su formato original (señales en crudo) y divididas en los conjuntos mencionados. Una vez mencionado lo anterior, se ejecutó el experimento y se obtuvieron resultados. La Figura 5.2 se muestra la matriz de confusión de prueba del entrenamiento con datos en crudo y la Tabla 5.4 muestra las métricas correspondientes al experimentos, obtenidas a partir de la matriz de confusión.

		Test					
Output Class	1	132 10.4%	35 2.7%	45 3.5%	20 1.6%	23 1.8%	51.8% 48.2%
	2	30 2.4%	130 10.2%	35 2.7%	35 2.7%	25 2.0%	51.0% 49.0%
	3	20 1.6%	0 0.0%	174 13.6%	20 1.6%	41 3.2%	68.2% 31.8%
	4	0 0.0%	30 2.4%	20 1.6%	175 13.7%	30 2.4%	68.6% 31.4%
	5	20 1.6%	0 0.0%	40 3.1%	10 0.8%	185 14.5%	72.5% 27.5%
		65.3% 34.7%	66.7% 33.3%	55.4% 44.6%	67.3% 32.7%	60.9% 39.1%	62.4% 37.6%
		Target Class					
		1	2	3	4	5	

Figura 5.2: Matriz de confusión de la fase de prueba del entrenamiento realizado con señales en crudo

Tabla 5.4: Métricas del experimento 1

Exactitud	Precisión	F1-score	Sensibilidad	Especificidad	Kappa	Matthews
0.624	0.630	0.622	0.624	0.906	0.148	0.532

El rendimiento observado en el Experimento 1 fue significativamente limitado, como lo demuestra los valores de la Tabla 5.4. Las principales limitaciones derivan del uso de señales en crudo, las cuales contienen ruido y artefactos que dificultan la discriminación efectiva entre las clases por parte del modelo. Este enfoque no incorpora técnicas de preprocesamiento, lo que impide la extracción de características distintivas que podrían facilitar la clasificación.

La exactitud general del modelo fue baja, con un alto porcentaje de falsos positivos y negativos distribuidos entre todas las clases. Este comportamiento se explica por varios factores clave:

- **Ausencia de Preprocesamiento:** Las señales en crudo están contaminadas por ruido inherente al proceso de adquisición, como interferencias eléctricas o movimientos musculares. Estos elementos introducen variabilidad no deseada que afecta la capacidad del modelo para aprender patrones relevantes.
- **Hiperparámetros por Defecto:** El uso de configuraciones iniciales genéricas para los hiperparámetros limita la capacidad del modelo de adaptarse a las características específicas de los datos. Esto puede llevar a un sobreajuste o subajuste dependiendo de la complejidad de las señales.
- **Falta de Representaciones Más Complejas:** Al trabajar directamente con señales en el dominio del tiempo, el modelo no cuenta con información más rica, como las frecuencias predominantes u otros indicadores espectrales que suelen ser más representativos para identificar arritmias.

En conclusión, los hallazgos del Experimento 1 subrayan la necesidad de técnicas avanzadas de preprocesamiento y ajustes en los hiperparámetros para mejorar el rendimiento del modelo, aspectos que se abordan en las siguientes fases experimentales.

5.4.2. Experimento 2: Entrenamiento del *Autoencoder* con hiperparámetros por defecto/Señales Filtradas

Condiciones del experimento

Para el Experimento 2, se llevó a cabo un entrenamiento y prueba del modelo utilizando sus hiperparámetros por defecto, tomando como entrada señales filtradas. Este procedimiento se realizó con el propósito de evaluar el rendimiento del modelo al analizar datos previamente filtrados de ruido y comparar los resultados con los obtenidos anteriormente. La configuración del modelo fue la misma que en el experimento anterior para garantizar una comparación justa.

Datos Utilizados

Se utilizó el conjunto de datos balanceado previamente descrito, con 2291 instancias para entrenamiento y 255 para pruebas en cada una de las cinco clases de ritmos cardíacos. Las señales fueron filtradas con un filtro de paso bajo con frecuencia de corte de 35 Hz para eliminar ruido de alta frecuencia y suavizar las señales, garantizando entradas más limpias al modelo. La Figura 5.3 muestra la matriz de confusión correspondiente a la prueba del entrenamiento con las señales filtradas y la Tabla 5.5 los valores de las métricas correspondientes.

		Test					
Output Class	1	205 16.1%	0 0.0%	25 2.0%	25 2.0%	0 0.0%	80.4% 19.6%
	2	18 1.4%	192 15.1%	15 1.2%	15 1.2%	15 1.2%	75.3% 24.7%
	3	12 0.9%	20 1.6%	201 15.8%	10 0.8%	12 0.9%	78.8% 21.2%
	4	5 0.4%	10 0.8%	5 0.4%	230 18.0%	5 0.4%	90.2% 9.8%
	5	10 0.8%	20 1.6%	30 2.4%	40 3.1%	155 12.2%	60.8% 39.2%
	82.0% 18.0%	79.3% 20.7%	72.8% 27.2%	71.9% 28.1%	82.9% 17.1%	77.1% 22.9%	
		Target Class					

Figura 5.3: Matriz de confusión de la fase de prueba del entrenamiento realizado con señales filtradas

Tabla 5.5: Métricas del experimento 2

Exactitud	Precisión	F1-score	Sensibilidad	Especificidad	Kappa	Matthews
0.771	0.777	0.769	0.771	0.943	0.284	0.716

Los valores presentados en la Tabla 5.5 refleja una mejora notable en comparación con los resultados obtenidos en el Experimento 1. Las métricas de evaluación muestran incrementos significativos, lo que indica que el modelo pudo capturar mejor las características discriminativas al trabajar con señales filtradas. Sin embargo, las señales siguen en el dominio del tiempo, lo que limita la capacidad del modelo para identificar patrones más complejos. Los puntos claves de este experimento son:

- **Exactitud global:** Incremento del rendimiento con valores que superan en más de 15 % al experimento anterior.
- **Distribución de errores:** Reducción de falsos positivos y negativos, aunque persisten ciertos errores en la clasificación de clases con características similares.

El uso de señales filtradas mejoró considerablemente el desempeño del modelo en comparación con el Experimento 1. Esta mejora se atribuye a la reducción de ruido y artefactos presentes en las señales en crudo, lo que facilitó el aprendizaje de patrones más relevantes por parte del modelo. Sin embargo, las señales en el dominio del tiempo no contienen información rica en características espectrales, lo que podría limitar el rendimiento máximo alcanzable con esta metodología.

Para superar estas limitaciones, en el siguiente experimento se exploró la sintonización de hiperparámetros, buscando optimizar el modelo para maximizar su capacidad de generalización al trabajar con señales filtradas.

5.4.3. Experimento 3: Sintonización de hiperparámetros/Señales filtradas

Condiciones del experimento

En el Experimento 3 se implementó un proceso exhaustivo de sintonización de hiperparámetros con el objetivo de optimizar el rendimiento del modelo. Este proceso se llevó a cabo mediante rigurosa experimentación, utilizando un enfoque iterativo de prueba y error. En cada iteración, se evaluó el impacto de los ajustes realizados en los hiperparámetros clave, tomando como base métricas de rendimiento como la precisión, sensibilidad y exactitud global. Después de la experimentación correspondiente, los mejores hiperparámetros fueron:

- 30 épocas de entrenamiento.
- Capa oculta de 60 neuronas.
- Tasa de aprendizaje = 0.01.
- Regularización L2 = 0.0001.
- Regularización de dispersión = 4.
- Proporción de dispersión = 0.3.
- Función *satlin* para codificación y *pureln* para la decodificación.

Datos Utilizados

El conjunto de datos balanceado empleado en el Experimento 2 fue utilizado nuevamente, compuesto por 2291 instancias para entrenamiento y 255 para prueba en cada una de las cinco clases de ritmos cardíacos. Las señales fueron filtradas previamente para eliminar ruido de alta frecuencia, asegurando datos de entrada de mayor calidad. En la Figura 5.4 se muestra la matriz de confusión del experimento 3 y en Tabla 5.6 se muestran los resultados de los hiperparámetros sintonizados.

		Test					
Output Class	1	205 16.1%	0 0.0%	25 2.0%	25 2.0%	0 0.0%	80.4% 19.6%
	2	18 1.4%	192 15.1%	15 1.2%	15 1.2%	15 1.2%	75.3% 24.7%
	3	6 0.5%	11 0.9%	231 18.1%	2 0.2%	5 0.4%	90.6% 9.4%
	4	5 0.4%	10 0.8%	5 0.4%	230 18.0%	5 0.4%	90.2% 9.8%
	5	5 0.4%	5 0.4%	5 0.4%	5 0.4%	235 18.4%	92.2% 7.8%
	85.8% 14.2%	88.1% 11.9%	82.2% 17.8%	83.0% 17.0%	90.4% 9.6%	85.7% 14.3%	
		Target Class					
		1	2	3	4	5	

Figura 5.4: Matriz de confusión de la fase de prueba del entrenamiento realizado con hiperparámetros sintonizados

Tabla 5.6: Métricas del experimento 3

Exactitud	Precisión	F1-score	Sensibilidad	Especificidad	Kappa	Matthews
0.857	0.858	0.856	0.857	0.964	0.554	0.822

La Tabla 5.6 muestra un incremento significativo en el rendimiento del modelo, teniendo como puntos claves:

- **Exactitud global:** Mejora aproximada del 8 % en comparación con el Experimento 2.
- **Distribución de errores:** Reducción considerable de falsos positivos y negativos, con mejoras notables en clases con características similares.

El proceso sistemático de sintonización de hiperparámetros permitió identificar configuraciones óptimas que maximizaron la capacidad del modelo para adaptarse a las características específicas de las señales ECG filtradas. La inclusión de una capa oculta con mayor número de neuronas incrementó la capacidad del modelo para aprender representaciones más complejas, mientras que ajustes precisos en la regularización L2 y la proporción de dispersión minimizaron riesgos de subajuste o sobreajuste.

Es importante destacar que la selección de estos hiperparámetros no se realizó de manera arbitraria. Por el contrario, cada configuración fue evaluada rigurosamente en múltiples iteraciones para determinar su impacto en el desempeño del modelo. Este enfoque metódico asegura que los resultados obtenidos reflejan el uso eficiente de los recursos del modelo bajo las condiciones establecidas.

A pesar de estas mejoras, las señales en el dominio del tiempo presentan limitaciones inherentes en términos de riqueza informativa. Por ello, en el siguiente experimento, se evaluaron características espectrales para capturar patrones más representativos y mejorar aún más el desempeño del modelo.

5.4.4. Experimento 4: Entrenamiento con espectros

Condiciones del experimento

En el Experimento 4, se buscó aprovechar las características espectrales de las señales ECG para mejorar el rendimiento del modelo. Este enfoque se basó en la transformación de las señales del dominio del tiempo al dominio de la frecuencia, utilizando distintas técnicas de estimación espectral, las cuáles fueron extraídas a partir de las señales en crudo y corresponden con las mencionadas en 2.5. Cada espectro fue tratado de manera independiente, y los entrenamientos se realizaron utilizando la configuración de hiperparámetros sintonizados obtenida en el Experimento 3. Esto aseguró condiciones controladas para evaluar el impacto de cada técnica espectral en el rendimiento del modelo.

Datos Utilizados

Se emplearon los mismos conjuntos de entrenamiento y prueba balanceados utilizados en los experimentos previos, con 2291 instancias por clase para entrenamiento y 255 para prueba. Las señales fueron previamente filtradas y posteriormente transformadas al dominio de la frecuencia mediante cada una de las técnicas espectrales mencionadas, evaluando cada uno de las estimaciones espectrales. La Figura 5.5 muestra una comparación entre los resultados obtenidos por cada espectro, tomando en cuenta las matrices de confusión correspondientes.

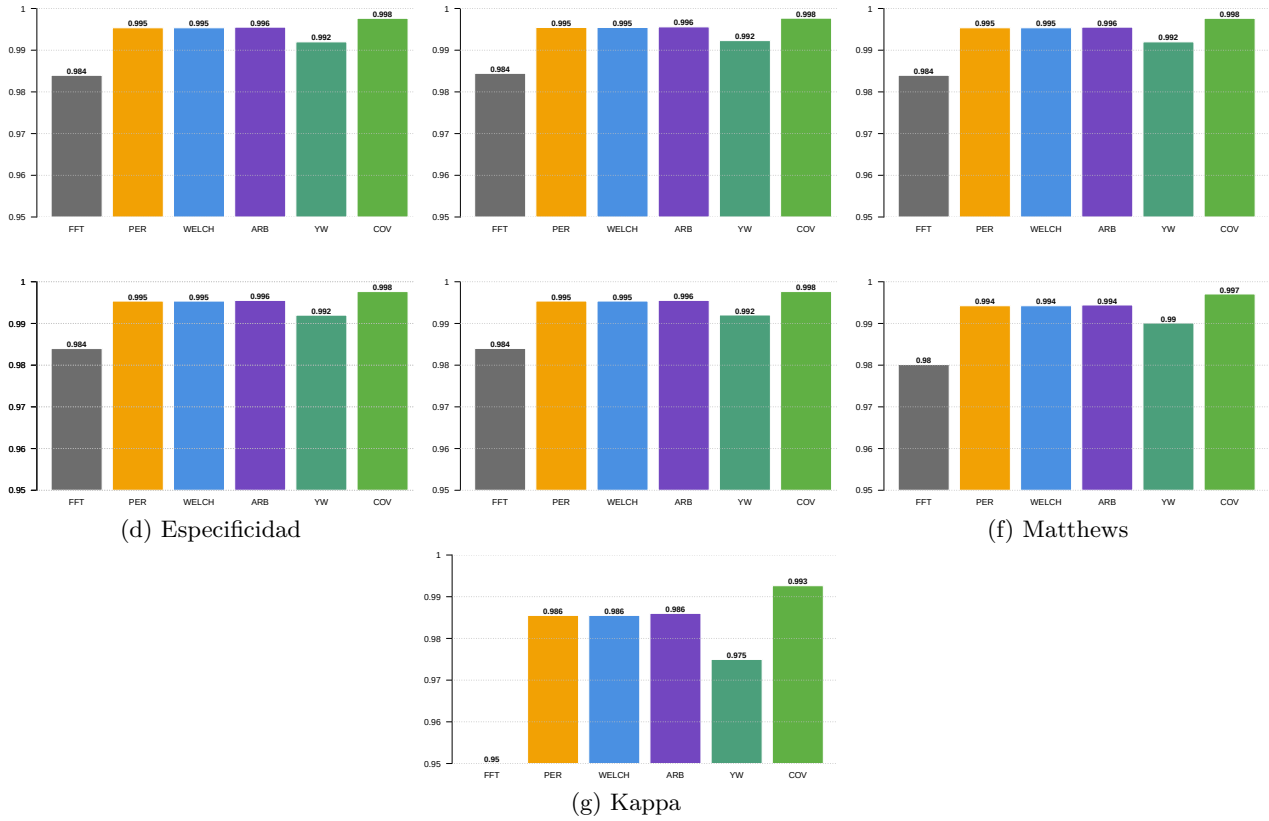


Figura 5.5: Comparación de métricas del experimento 4

En la Figura anterior, se presentan las métricas obtenidas para cada espectro. Los resultados mostraron que el espectro de Covarianza obtuvo el mejor desempeño, seguido por AR Burg, Welch y Periodograma. Las principales observaciones fueron:

- **Valores de métricas altos:** Todos los espectros alcanzaron métricas superiores a 95 %, siendo el mejor de ellos Covariance, lo que confirma la eficacia de las estimaciones espectrales para capturar patrones discriminativos.
- **Mejora notoria:** Incremento de los valores de métricas en un 12 % en comparación con el Experimento 3.
- **Reducción de errores:** Todos los espectros mostraron una disminución significativa en falsos positivos y negativos, siendo Covarianza el más robusto.

En conclusión, el análisis espectral evidenció un incremento sustancial en el rendimiento del modelo gracias a la riqueza de información proporcionada por las características espectrales. Este experimento reafirmó la importancia del análisis espectral en la clasificación de arritmias cardíacas. La utilización del espectro de Covarianza como representación principal demostró ser crucial para maximizar los valores en las métricas de calidad, posicionándolo como una solución efectiva para esta tarea.

5.4.5. Experimento 5: Comparación con otros modelos del estado del arte

Condiciones del Experimento

El Experimento 5 tiene como objetivo comparar el desempeño del modelo propuesto con otros modelos del estado del arte en términos de métricas de calidad.

Datos Utilizados

Para garantizar una comparación justa, por cada modelo evaluado se emplearon los mismos datos y configuración utilizados en el Experimento 4, lo que permite evaluar de manera objetiva la capacidad del modelo propuesto frente a técnicas previamente reportadas en la literatura. Se consideró solo el espectro covariante debido a que este generó los mejores resultados en el experimento anterior. La Tabla 5.7 muestra los valores obtenidos por cada modelo.

Tabla 5.7: Métricas del experimento 5

Modelo	Exactitud	Precisión	F1-score	Sensibilidad	Especificidad	Kappa	Matthews	Inferencia (s)
MLP [49]	0.982	0.982	0.982	0.982	0.994	0.963	0.976	0.070
Transformer [5]	0.990	0.990	0.996	0.990	0.996	0.986	0.973	0.080
BILSTM [50]	0.992	0.992	0.997	0.992	0.997	0.980	0.990	0.006
Propuesto	0.997	0.998	0.997	0.994	0.999	0.992	0.996	0.001

5.4.6. Experimento 6: Evaluación del modelo propuesto con otra base de datos

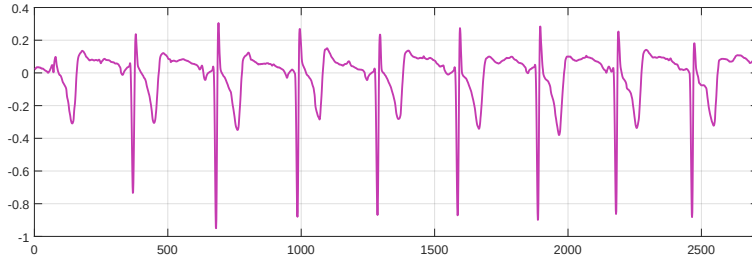
Condiciones del Experimento

El Experimento 6 se diseñó con objetivo de evaluar la capacidad de generalización del modelo propuesto al ser aplicado a una base de datos distinta a la utilizada en los experimentos previos. Para ello, se empleó el mismo enfoque metodológico del Experimento 4, donde el modelo fue entrenado con espectros extraídos de señales ECG, con el fin de comparar el rendimiento obtenido en una base de datos externa. Este análisis permite evaluar la robustez del modelo frente a variaciones en la adquisición y características de las señales ECG, así como validar su desempeño en entornos clínicos distintos.

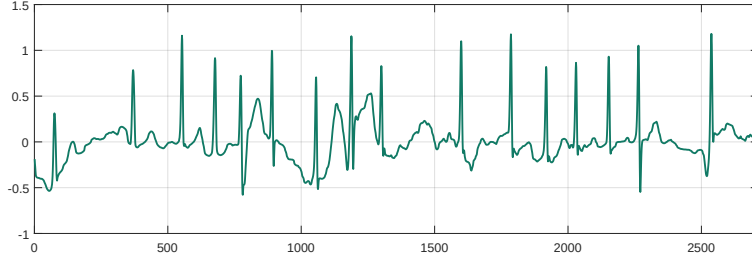
El modelo fue evaluado bajo las mismas condiciones de hiperparámetros utilizados en el Experimento 4, garantizando la consistencia en los resultados. Además, se tomaron en cuenta métricas de calidad previamente descritas para comparar el desempeño del modelo en ambas bases de datos.

Datos Utilizados

Para este experimento, se utilizó la base de datos PhysioNet CinC 2017, la cual contiene muestras de ritmos cardíacos normales (N) y fibrilación auricular (AF). Esta base de datos ha sido ampliamente utilizada en la literatura para la detección automática de arritmias, proporcionando un entorno adecuado para evaluar la capacidad del modelo en la clasificación de señales ECG en un contexto clínico más amplio. La base de datos está compuesta por 738 señales por clase, registradas a una frecuencia de muestreo de 300 Hz y con una longitud de 2714 muestras por señal. Esto garantiza que cada señal proporciona una cantidad suficiente de información para la caracterización del ritmo cardíaco, permitiendo una evaluación robusta del modelo. La Figura 5.6 muestra los diferentes tipos de señales ECG contenidos en la base de datos PhysioNet CinC 2017.



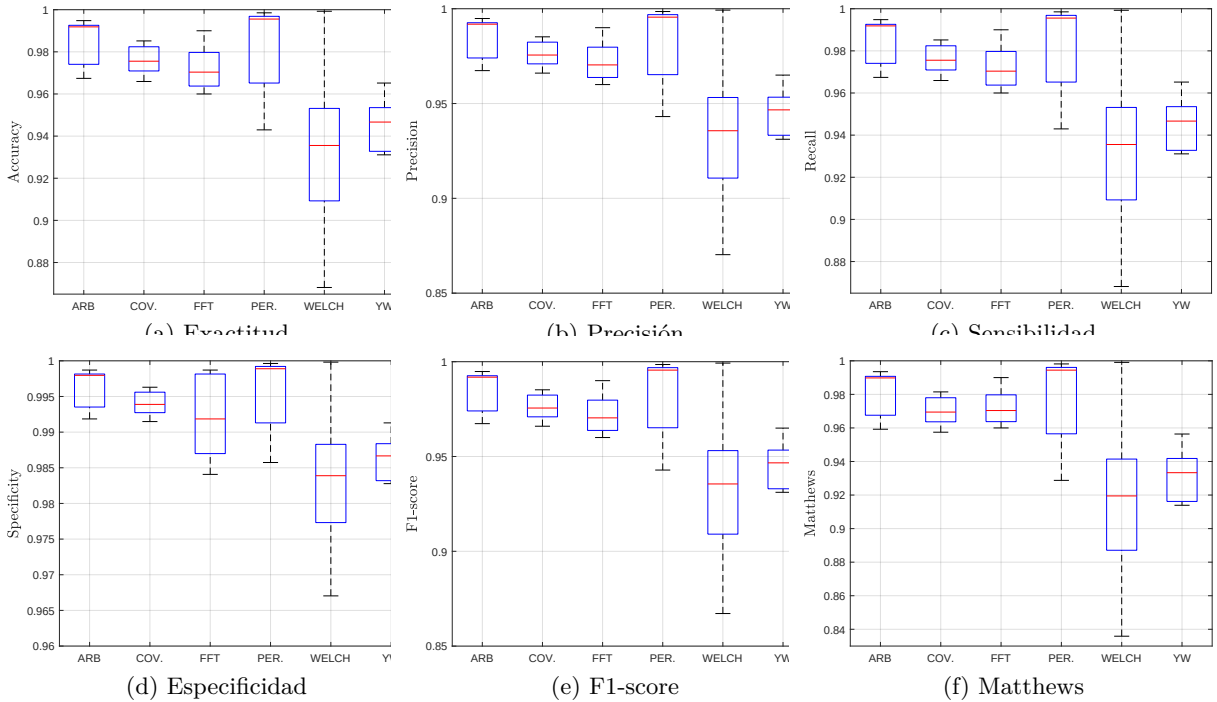
(a) Ritmo Cardíaco Normal



(b) Fibrilación Auricular

Figura 5.6: Muestras de la base de datos Physionet CinC 2017

La partición de los datos se realizó utilizando una técnica de validación cruzada de 10 partes, teniendo un 90 % de los datos para entrenamiento y 10 % para prueba por cada iteración. Esta estrategia permite evaluar la capacidad del modelo para generalizar en una base de datos distinta y verificar si mantiene un rendimiento estable al clasificar señales ECG provenientes de diferentes fuentes. Las señales fueron preprocesadas siguiendo el mismo procedimiento utilizado en los experimentos anteriores: se aplicó un filtro pasa-bajo con frecuencia de corte de 35 Hz para eliminar ruido de alta frecuencia y posteriormente se transformaron al dominio de la frecuencia mediante los seis espectros considerados en el Experimento 4. La Figura 5.7 muestra algunos valores estadísticos de las métricas obtenidas a partir de la validación cruzada.



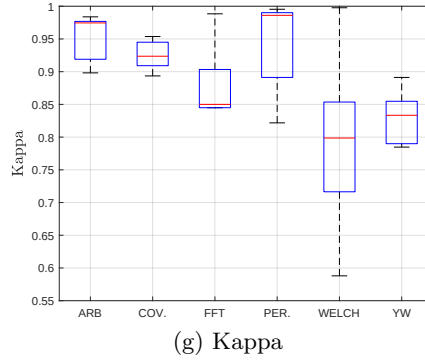


Figura 5.7: Valores estadísticos de validación cruzada para la evaluación del modelo propuesto sobre la base de datos Physionet CinC 2017

La Tabla 5.8 presenta los valores medios obtenidos a partir de la Figura 5.7, tomando como base al espectro Periodograma al considerarse como la mejor opción en este caso. Adicionalmente se incluyeron los resultados obtenidos en el trabajo previo a esta investigación [2] para comparar resultados.

Tabla 5.8: Métricas del experimento 6

Modelo	Exactitud	Precisión	F1-score	Sensibilidad	Especificidad	Kappa	Matthews	Inferencia (s)
Propuesto	0.996	0.998	0.996	0.994	0.998	0.993	0.993	0.009
BiLSTM [2]	0.935	0.928	0.935	0.935	0.935	0.871	0.871	0.35

Los resultados obtenidos reflejan un desempeño notablemente alto, teniendo a Periodograma como la mejor alternativa al mantener métricas de calidad por encima del 95 % en la mayoría de los casos. Lo anterior demuestra que el modelo conserva su capacidad de clasificación efectiva incluso cuando se aplica a datos no vistos previamente. Comparado con el modelo BiLSTM previo a este trabajo [2], el modelo propuesto superó significativamente su desempeño, alcanzando una mejora de más del 6 % en todas las métricas, lo que sugiere una mejor capacidad de clasificación y menor sesgo en la predicción de las clases. La mejora del rendimiento del modelo en esta base de datos sugiere que la combinación de *Autoencoder* profundo con espectros de frecuencia permite capturar de manera más efectiva las características discriminantes de los ritmos cardíacos.

En términos de costo computacional, el modelo propuesto se mantiene competitivo, ya que logra una clasificación eficiente con una arquitectura más simple en comparación con modelos más pesados como BiLSTM. Esto refuerza su aplicabilidad en entornos clínicos donde los recursos computacionales pueden ser limitados. Por otro lado, la validación cruzada permitió evaluar la estabilidad del modelo, mostrando variaciones mínimas en las métricas obtenidas entre las distintas iteraciones, lo que respalda la confiabilidad del enfoque.

En conclusión, los resultados del Experimento 5 confirman que el modelo propuesto es una alternativa viable y eficiente para la clasificación automática de arritmias en distintos conjuntos de datos. Su capacidad de generalización lo posiciona como una solución prometedora para la implementación en entornos clínicos reales, con la posibilidad de ser mejorado mediante ajustes en su arquitectura.

5.4.7. Experimento 7: Entrenamiento del modelo con señales sintéticas y evaluación con señales reales

Condiciones del Experimento

El objetivo de este experimento es evaluar la capacidad de generalización del modelo propuesto cuando es entrenado exclusivamente con señales sintéticas y evaluado con señales reales. Se busca determinar si la utilización de datos generados artificialmente es viable para la clasificación en escenarios donde los datos reales son limitados o difíciles de obtener. Para la evaluación del modelo, se utilizaron los mismos hiperparámetros y métricas utilizados que el experimento 4.

Datos Utilizados

Para este experimento, se utilizó una Red Generativa Antagónica (GAN) para la generación de señales sintéticas. La arquitectura de la GAN utilizada se muestra en la Figura 5.8.

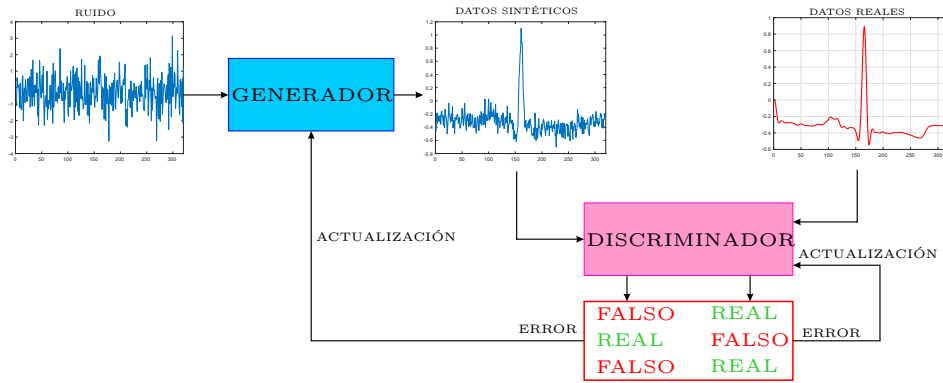
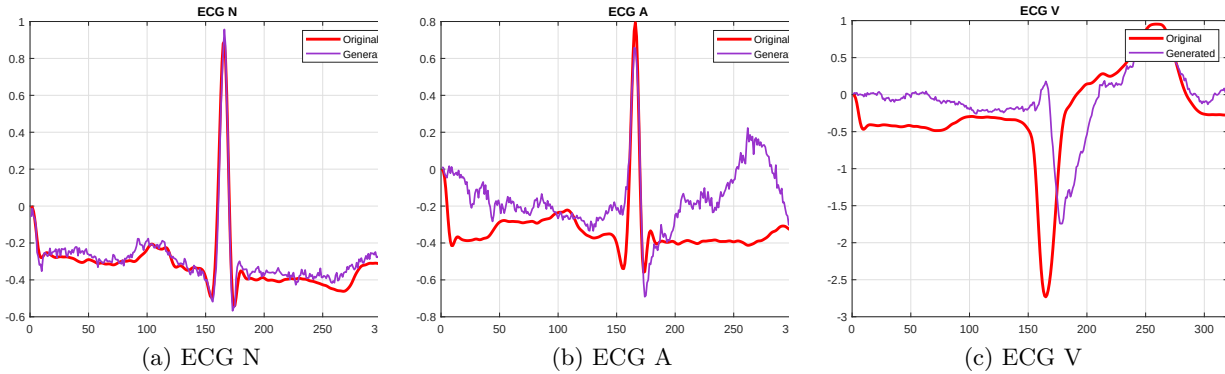


Figura 5.8: Arquitectura de la GAN utilizada

De la base de datos original, la cuál contiene 2546 instancias por clase, se realizó una distribución correspondiente a 30 % (764 instancias por clase) para el entrenamiento de la GAN, la cuál generó 2500 señales sintéticas (500 por clase). El 70 % restante (1782 instancias por clase) se utilizó para la posterior evaluación del *Autoencoder* modificado. La Figura 5.9 muestra un contraste entre las señales sintéticas y originales.



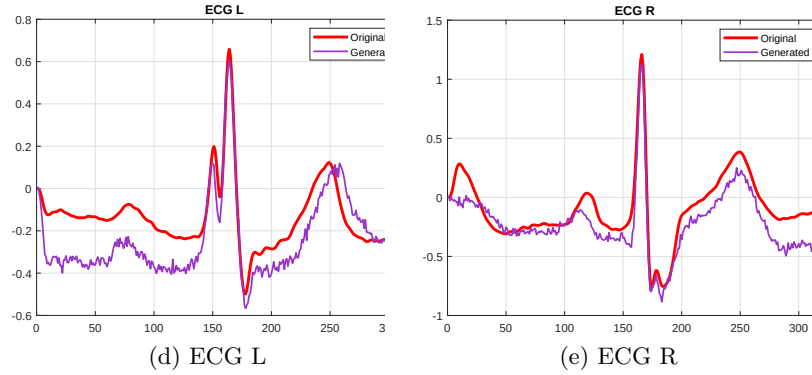


Figura 5.9: Comparación entre señales sintéticas y originales

Posteriormente, debido a que el espectro Covarianza resultó la mejor alternativa en el experimento 5, se procedió a obtener dicho espectro tanto de las señales sintéticas como de las originales. Después, se entrenó el modelo propuesto con las señales sintéticas y se evaluó con las señales originales. Los resultados obtenidos se muestran en la Tabla 5.9.

Tabla 5.9: Métricas del experimento 7

Exactitud	Precisión	F1-score	Sensibilidad	Especificidad	Kappa	Matthews
0.997	0.997	0.997	0.997	0.999	0.993	0.996

A pesar de haber sido entrenado únicamente con señales sintéticas, el modelo logró un desempeño sobresaliente en la clasificación de señales reales. La utilización del espectro Covarianza generó valores de métricas superiores a 95 %. Los resultados obtenidos demuestran la viabilidad de entrenar modelos de clasificación de ECG con señales sintéticas, obteniendo un rendimiento comparable o incluso superior al de modelos entrenados con señales reales. Esto sugiere que la generación de datos sintéticos puede ser una alternativa eficaz en situaciones donde la disponibilidad de datos reales es limitada.

Si bien el modelo demostró una notable capacidad de generalización, la calidad de las señales sintéticas sigue siendo un factor crítico. Investigaciones futuras podrían explorar mejoras en la generación de señales, evaluar su impacto en la robustez del modelo y analizar su aplicabilidad en escenarios clínicos reales.

5.4.8. Experimento 8: Entrenamiento con momentos de tiempo-frecuencia

Condiciones del Experimento

El objetivo de este experimento es evaluar el desempeño del modelo propuesto cuando se entrena con características de tiempo-frecuencia obtenidas a partir de señales ECG, correspondientes a las características utilizadas en [2]. A diferencia de los enfoques anteriores, donde las señales fueron transformadas al dominio de la frecuencia mediante estimaciones espectrales, en este caso se emplean técnicas de análisis tiempo-frecuencia para extraer información dinámica de las señales. Este análisis permite capturar variaciones locales en la distribución de energía de las señales ECG a lo largo del tiempo, proporcionando un conjunto de características complementarias para la clasificación. Para garantizar una comparación justa con los experimentos anteriores, se mantuvieron las configuraciones de hiperparámetros optimizadas en el Experimento 4 y se evaluó el impacto de estas nuevas características en el desempeño del modelo.

Datos Utilizados

Se obtuvieron momentos de tiempo-frecuencia a partir de las señales ECG filtradas, manteniendo la misma distribución para datos de entrenamiento y prueba que el Experimento 4. Las características obtenidas tienen una longitud de 29 características por instancia. La Figura 5.10 muestra representaciones de los momentos tiempo-frecuencia extraídos y la Tabla 5.10 muestra los resultados de este experimento.

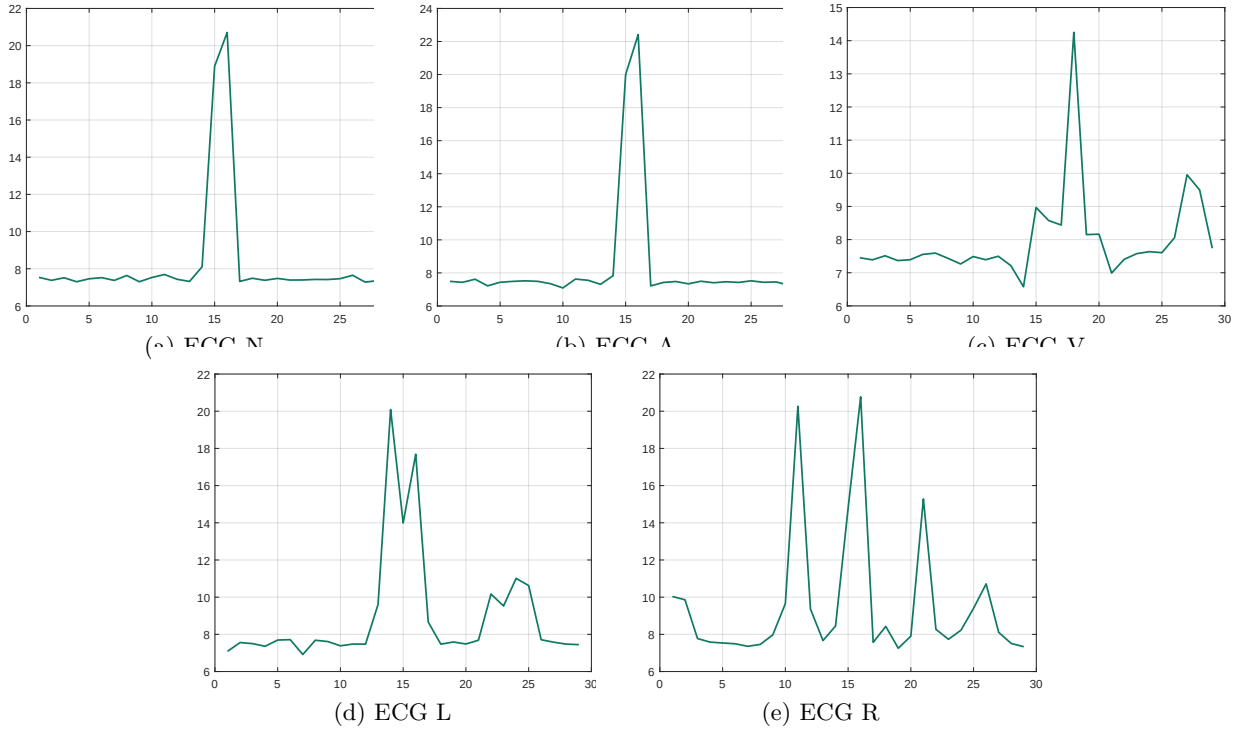


Figura 5.10: Momentos tiempo-frecuencia extraídos a partir de las señales filtradas

Tabla 5.10: Métricas del experimento 8

Exactitud	Precisión	F1-score	Sensibilidad	Especificidad	Kappa	Matthews
0.956	0.956	0.956	0.956	0.989	0.865	0.947

Los resultados obtenidos reflejan que, si bien este enfoque permitió capturar información dinámica de las señales, su desempeño fue ligeramente inferior en comparación con las representaciones espectrales evaluadas en experimentos previos. En general, las métricas de calidad alcanzaron valores superiores al 95 %, lo que demuestra que los momentos tiempo-frecuencia son adecuados para la clasificación de arritmias, superando a los resultados obtenidos en [2]. Sin embargo, en comparación con los espectros utilizados en el Experimento 4, las métricas fueron marginalmente menores, lo que sugiere que la capacidad discriminan de los momentos tiempo-frecuencia no es tan sólida como la proporcionada por las estimaciones espectrales.

Otro aspecto relevante es el impacto en la estabilidad del modelo. A pesar de que los momentos tiempo-frecuencia ofrecen una representación enriquecida de las señales ECG, la variabilidad en los coeficientes extraídos podría haber afectado la robustez del modelo, generando un leve incremento en los errores de clasificación. No obstante, estos resultados resaltan la importancia de

explorar múltiples representaciones de las señales ECG, ya que cada una puede aportar información complementaria para mejorar la clasificación de arritmias en distintos escenarios clínicos.

5.5. Discusión

Los experimentos realizados proporcionaron una evaluación progresiva del modelo propuesto, permitiendo analizar su comportamiento bajo diferentes condiciones. En primera instancia, los experimentos iniciales evidenciaron la influencia del preprocesamiento en la calidad de los resultados, destacando que la eliminación de ruido mediante filtrado mejoró significativamente el rendimiento del modelo en comparación con el uso de señales en crudo. A pesar de esta mejora, el análisis en el dominio del tiempo presentó limitaciones para capturar patrones distintivos entre clases, lo que motivó la exploración de representaciones espectrales.

La transformación a espectros en el Experimento 4 marcó una mejora importante en el desarrollo experimental, ya que permitió extraer características más representativas de las señales ECG. Entre los distintos métodos evaluados, el espectro de Covarianza resultó la mejor alternativa, logrando métricas de calidad superiores al 99 %, lo que confirma su idoneidad para la clasificación de arritmias. En contraste, el resto de espectros mostraron resultados ligeramente inferiores pero competentes, lo que sugiere que la estructura del modelo propuesto se beneficia de representaciones espectrales que enfatizan la densidad espectral de potencia.

El Experimento 5 comparó el rendimiento del modelo con arquitecturas del estado del arte. Los resultados indicaron que el modelo propuesto no solo alcanzó métricas más altas, sino que también presentó un menor costo computacional, lo que lo convierte en una alternativa eficiente para sistemas embebidos o aplicaciones en tiempo real. Además, el análisis del tiempo de inferencia reflejó que la implementación basada en *Autoencoders* con espectros espectrales requiere menos procesamiento en comparación con modelos más complejos de aprendizaje profundo.

Finalmente, el Experimento 6 evaluó la capacidad de generalización del modelo en la base de datos PhysioNet CinC 2017. Los resultados obtenidos validaron la robustez del modelo, evidenciando que mantiene un rendimiento estable incluso cuando se aplicó a datos de distinta procedencia. La validación cruzada realizada permitió reducir posibles sesgos en la evaluación, demostrando que la metodología propuesta es adaptable a diferentes escenarios clínicos. En conjunto, estos hallazgos consolidan la arquitectura propuesta como una solución efectiva para la clasificación automática de arritmias, con ventajas en términos de precisión, eficiencia y aplicabilidad en contextos reales.

6 Conclusiones

A continuación se detalla de forma general los resultados obtenidos de las actividades que se realizaron para el desarrollo de este trabajo. A su vez, se hace una descripción de los productos generados y las aportaciones realizadas.

6.1. Objetivos y alcances logrados

En las tablas 6.1, 6.2 y 6.3 se presenta un resumen de los objetivos, alcances y limitaciones tomados en cuenta en este trabajo.

Tabla 6.1: Objetivos específicos

Objetivo	Actividades
Analizar y comprender las señales ECG, explorando las diversas características que se pueden extraer de ellas.	Se realizó el estudio del marco conceptual y el estado del arte de clasificación ECG y técnicas utilizadas, se redactó un reporte del Estado del Arte.
Profundizar en los conceptos relacionados con la estimación espectral. Además, se abordará la arquitectura de los <i>Autoencoders</i> , con especial atención al concepto de apilamiento.	Se extrajeron 6 diferentes espectros de las señales ECG. Se estudió las arquitecturas de los <i>Autoencoders</i> y las modificaciones posibles que se podían realizar.
Rediseñar la estructura de un <i>Autoencoder</i> mediante la incorporación de una capa totalmente conectada para realizar la tarea de clasificación.	Se diseñó un modelo basado en un <i>Autoencoder</i> para la clasificación multiclase de señales ECG.
Evaluar el rendimiento del modelo desarrollado utilizando métricas mencionadas en el estado del arte, así como aquellas que se desarrollen a lo largo de esta investigación. Este análisis incluirá la consideración de los diferentes espectros previamente mencionados.	Se evaluó el desempeño del modelo tomando en cuenta los diferentes espectros utilizados. Se evaluó y comparó los mejores resultados obtenidos con otros modelos utilizados en el estado del arte.
Desarrollar un sistema computacional que integre todos los módulos del procesamiento de señales ECG y el modelo propuesto.	Se implementaron los módulos de preprocesamiento de señales ECG, extracción de características espectrales y el modelo de clasificación propuesto en un sistema computacional unificado, permitiendo la evaluación y comparación de las señales en tiempo real con los modelos previamente entrenados.

Tabla 6.2: Alcances

Objetivo	Actividades
Se considerarán los espectros <i>AR-Burg, Covariance, Yule-walker, Fast Fourier Transform, Periodogram y Welch.</i>	Se obtuvieron los 6 diferentes espectros de las señales en crudo para evaluar sus diferencias.
Se hará una comparación entre distintas estimaciones espectrales	Después de los entrenamientos correspondientes, se evaluaron y compararon los resultados utilizando métricas de calidad utilizadas en el estado del arte.
Se considerará el dominio espectral de las señales.	Se realizaron 6 diferentes entrenamientos del modelo propuesto, usando cada uno de los espectros extraídos.
Se buscará mantener un costo computacional bajo con el modelo propuesto. Se buscará obtener un rendimiento alto en la clasificación multiclase.	Se realizó una comparación de los resultados entre el modelo propuesto y algunos modelos en el estado del arte, obteniendo los mejores resultados. Se realizó una sintonización de hiperparámetros al modelo propuesto y un preprocesamiento adecuado a los datos para garantizar resultados competentes.

Tabla 6.3: Limitaciones

Objetivo	Actividades
Se utilizarán las bases de datos mencionadas en este documento, así como alguna otra que se encuentre en el desarrollo de esta investigación que contenga señales de ECG de arritmias cardíacas.	Se utilizó la base de datos MIT-BIH, que contiene 5 tipos de ritmos cardíacos.
Se utilizarán únicamente estimaciones espectrales de las señales.	Se consideraron 6 diferentes espectros para evaluar el rendimiento del modelo con cada uno de ellos y comparar resultados.
Se centrará únicamente en el reconocimiento de 4 tipos de arritmias y 1 ritmo cardíaco normal.	Se utilizaron 5 clases de la base de datos MIT-BIH, las cuales son: ritmo cardíaco normal, contracción prematura auricular, contracción prematura ventricular, bloqueo de rama izquierda y bloqueo de rama derecha.

6.2. Productos

Durante el desarrollo de esta investigación, se obtuvieron los siguientes productos:

1. Reporte del estado del arte: documento en el que se presenta un resumen de diversas publicaciones recientes consideradas a partir de los núcleos temáticos contemplados en la investigación como lo es clasificación de arritmias cardíacas, procesamiento digital de señales, métricas y bases de datos.

2. Diseño e implementación de los algoritmos para procesamiento de señales: códigos fuente para la aplicación de filtros pasa-bajos y obtención de diferentes espectros a partir de señales ECG.
3. Diseño e implementación del *Autoencoder* profundo: códigos fuente para la conformación de las capas, sintonización de hiperparámetros, entrenamiento y pruebas del modelo con espectros.
4. Capítulo de libro “*Multiclass classification of ECG signals using a LSTM network trained with spectral features*” indizado en Scopus para el congreso internacional *International Congress of Telematics and Computing* (WITCOM) 2024.
5. Artículo “*Atrial Fibrillation Classification Using a Deep Spectral Autoencoder*” indizado en Scopus y JCR para la revista *Computación y Sistemas* (CyS) 2024.
6. Artículo sometido “*ECG Classification Using a CNN Stacked with an Autoencoder Trained with Synthetic Signals*” indizado en JCR para la revista *Neural Computing and Applications*.
7. Poster titulado “Clasificación de arritmias cardíacas mediante un *Autoencoder* profundo espectral” presentado en el 1er. coloquio de investigación, tecnología y sustentabilidad.
8. Documento de tesis.

6.3. Conclusión

En este trabajo se desarrolló un nuevo enfoque para la clasificación multiclase de señales ECG utilizando un modelo apilado entre un *Autoencoder* y una capa totalmente conectada. La arquitectura propuesta fue diseñada con el objetivo de mantener un bajo costo computacional sin comprometer el rendimiento en tareas de clasificación, lo cual constituye una de las principales aportaciones de esta investigación. Para desarrollar el modelo propuesto se emplearon las técnicas que se detallaron en este documento, desde el filtrado de ruido de las señales usando un filtro pasa-bajos hasta la transformación de las señales al dominio de la frecuencia mediante estimaciones espectrales. Esta última etapa permitió una representación más precisa y discriminante de las características relevantes de las señales ECG, favoreciendo el rendimiento del modelo.

Una de las contribuciones clave de este trabajo es la modificación estructural del *Autoencoder*, mediante la incorporación de una capa completamente conectada, lo que permitió extender su funcionalidad más allá de la reconstrucción, habilitando su uso para clasificación. Esta modificación, junto con la compresión eficiente de características, permitió reducir la dimensionalidad de los datos y facilitar su procesamiento por parte del clasificador, mejorando su desempeño y eficiencia. Asimismo, la integración de estimaciones espectrales como entrada al modelo representa un aporte significativo, ya que estas capturan de manera más efectiva las características discriminativas entre clases, en comparación con las representaciones temporales.

Los resultados obtenidos evidenciaron tanto las fortalezas como las áreas de mejora de la propuesta. En los Experimentos 1 y 2, se evaluó el modelo utilizando señales ECG en crudo y filtradas, observando que el preprocesamiento mediante filtrado permitió un incremento en las métricas de clasificación. No obstante, las señales en el dominio temporal aún mostraron limitaciones para capturar patrones complejos. En el Experimento 3, se implementó una sintonización hiperparamétrica exhaustiva, lo que optimizó la configuración del modelo y mejoró significativamente su capacidad para adaptarse a las características específicas de las señales.

Posteriormente, en el Experimento 4, se realizó la transformación de las señales al dominio de la frecuencia mediante estimaciones espectrales, lo que presentó mejoras significantes. Esta transformación proporcionó una representación más discriminante, destacando patrones espectrales específicos de cada clase de arritmia. Entre las técnicas espectrales evaluadas, el espectro de Covarianza demostró ser el más efectivo, seguido por AR Burg, debido a su capacidad para representar con precisión la densidad espectral de potencia y reducir los errores de predicción. Gracias a este enfoque, el modelo alcanzó valores de métricas superiores al 95 %, consolidándose como la mejor solución dentro de los experimentos realizados hasta ese punto.

En los Experimentos 5 y 6, se llevó a cabo una evaluación más rigurosa del modelo propuesto. En el Experimento 5, se comparó su desempeño con modelos del estado del arte, evidenciando que la solución propuesta alcanzó métricas de clasificación superiores en comparación con enfoques tradicionales, validando su eficacia. Además, en términos de costo computacional, el modelo se mantuvo competitivo, logrando una clasificación eficiente con una arquitectura más simple en comparación con otras técnicas de la literatura. Estos resultados refuerzan su aplicabilidad en entornos clínicos donde los recursos computacionales pueden ser limitados. Finalmente, en el Experimento 6, se evaluó la capacidad de generalización del modelo en una base de datos diferente, utilizando la base PhysioNet CinC 2017. Se observó que el modelo mantuvo un rendimiento estable, alcanzando métricas de calidad por encima del 95 % en la mayoría de los casos, lo que demuestra su robustez y capacidad de adaptación a diferentes bases de datos. Además, la validación cruzada permitió evaluar la estabilidad del modelo, mostrando variaciones mínimas en las métricas obtenidas en distintas iteraciones, lo que respalda la confiabilidad del enfoque.

En conclusión, la evolución de la propuesta reflejó un progreso significativo desde su configuración inicial hasta la integración de representaciones espectrales avanzadas. Cada etapa permitió identificar fortalezas y debilidades clave, desde la importancia del preprocesamiento hasta la optimización hiperparamétrica y la relevancia de los datos en el dominio de la frecuencia. El modelo propuesto demostró ser una arquitectura prometedora para la clasificación automática de arritmias cardíacas, estableciendo una base sólida para su futura implementación en entornos clínicos reales.

Como trabajo futuro, se recomienda evaluar el modelo con otras bases de datos de ECG para reforzar su capacidad de generalización. También se sugiere explorar otras técnicas de procesamiento que optimicen la información de los espectros más efectivos. Además, se considera la implementación en sistemas embebidos de bajo costo, para garantizar su aplicabilidad en entornos clínicos reales y contribuir al diagnóstico temprano de arritmias.

Bibliografía

- [1] Andersen, R. S., Peimankar, A., and Puthusserypady, S. (2019). A deep learning approach for real-time detection of atrial fibrillation. *Expert Systems with Applications*, 115:465–473.
- [2] Aquino, C. J. G. (2022). Clasificación de la fibrilación auricular en señales electrocardiográficas utilizando una red recurrente profunda entrenada con momentos de tiempo-frecuencia. Technical report, Tecnológico Nacional de México, Centro Nacional de Investigación y Desarrollo Tecnológico, Cuernavaca, Morelos. Reporte de Resultados.
- [3] Awad Pérez, K. and Saldarriaga Echeverri, F. (2023). Sistema de detección de fibrilación auricular a partir de procesamiento de señales de ecg e inteligencia artificial.
- [4] Azcona, L. (2009). El electrocardiograma. *López Farré A, Macaya Miguel C, directores. Libro de la salud cardiovascular del Hospital Clínico San Carlos y la fundación BBVA. 1ª ed. Bilbao: Fundación BBVA*, pages 49–56.
- [5] Busia, P., Scrugli, M. A., Jung, V. J. B., Benini, L., and Meloni, P. (2024). A tiny transformer for low-power arrhythmia classification on microcontrollers. *IEEE Transactions on Biomedical Circuits and Systems*.
- [6] Cao, P., Li, X., Mao, K., Lu, F., Ning, G., Fang, L., and Pan, Q. (2020). A novel data augmentation method to enhance deep neural networks for detection of atrial fibrillation. *Biomedical Signal Processing and Control*, 56:101675.
- [7] Cortes Martinez, K. V. et al. (2019). Redes neuronales pulsantes para la detección e interpretación de lesiones en el cerebro humano.
- [8] Costa Cortéz, N. A. et al. (2021). Diagnóstico de la fibrilación auricular mediante el análisis inteligente de registros intracardíacos.
- [9] Daydulo, Y. D., Thamineni, B. L., and Dawud, A. A. (2023). Cardiac arrhythmia detection using deep learning approach and time frequency representation of ecg signals. *BMC Medical Informatics and Decision Making*, 23(1):232.
- [10] Gao, H., Wang, X., Chen, Z., Wu, M., Li, J., and Liu, C. (2023). ECG-CL: A Comprehensive Electrocardiogram Interpretation Method Based on Continual Learning. *arXiv preprint arXiv:2304.04646*.
- [11] Ghil, M. and Taricco, C. (1997). Advanced spectral-analysis methods. In *Past and present variability of the solar-terrestrial system: measurement, data analysis and theoretical models*, pages 137–159. Ios Press.
- [12] Grune, J., Yamazoe, M., and Nahrendorf, M. (2021). Electroimmunology and cardiac arrhythmia. *Nature reviews cardiology*, 18(8):547–564.
- [13] He, R., Liu, Y., Wang, K., Zhao, N., Yuan, Y., Li, Q., and Zhang, H. (2019). Automatic cardiac arrhythmia classification using combination of deep residual network and bidirectional lstm. *IEEE access*, 7:102119–102135.
- [14] He, R., Liu, Y., and Zhang, H. (2020). Study on automatic classification of arrhythmias. *Feature engineering and computational intelligence in ECG monitoring*, pages 113–141.
- [15] Holton, T. (2021). *Digital signal processing: Principles and applications*. Cambridge University

Press.

- [16] Hsieh, C.-H., Li, Y.-S., Hwang, B.-J., and Hsiao, C.-H. (2020). Detection of atrial fibrillation using 1d convolutional neural network. *Sensors*, 20(7):2136.
- [17] Hsu, P.-Y. and Cheng, C.-K. (2020). Arrhythmia classification using deep learning and machine learning with features extracted from waveform-based signal processing. In *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pages 292–295. IEEE.
- [18] Huang, J.-S., Chen, B.-Q., Zeng, N.-Y., Cao, X.-C., and Li, Y. (2020). Accurate classification of ecg arrhythmia using mowpt enhanced fast compression deep learning networks. *Journal of Ambient Intelligence and Humanized Computing*, pages 1–18.
- [19] Jha, C. K. and Kolekar, M. H. (2020). Cardiac arrhythmia classification using tunable q-wavelet transform based features and support vector machine classifier. *Biomedical Signal Processing and Control*, 59:101875.
- [20] Khatibi, T. and Rabinezhadsadatmahaleh, N. (2020). Proposing feature engineering method based on deep learning and k-nns for ecg beat classification and arrhythmia detection. *Physical and Engineering Sciences in Medicine*, 43(1):49–68.
- [21] Kumar, R. and D, C. (2023). A hybrid approach using svm, knn and random forest for ecg classification. *Proceedings of the International Symposium on Electronic System Design*, pages 178–183.
- [22] Lebert, J., Deng, D., Fan, L., Lee, L. C., and Christoph, J. (2023). Deep learning-based prediction of electrical arrhythmia circuits from cardiac motion: An in-silico study. *arXiv preprint arXiv:2305.07822*.
- [23] LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *nature*, 521(7553):436–444.
- [24] León, A. A. S. and Alvarez, J. R. N. (2019). 1d convolutional neural network for detecting ventricular heartbeats. *IEEE Latin America Transactions*, 17(12):1970–1977.
- [25] Luna Alvarez, J. (2020). Conducción autónoma de un vehículo simulado mediante un modelo de red neuronal convolucional recurrente.
- [26] Marple Jr, S. L. (2019). *Digital spectral analysis*. Courier Dover Publications.
- [27] Marzana Peralta, Y. H. et al. (2019). *Clasificación de arritmias cardíacas utilizando análisis de componentes principales y redes neuronales pre-alimentadas en segmentos cortos de electrocardiograma*. PhD thesis.
- [28] Matuz Cruz, M. d. J. (2017). Detección de anomalías en mamografías utilizando la red neuronal convolucional alexnet.
- [29] Nurmaini, S., Tondas, A. E., Darmawahyuni, A., Rachmatullah, M. N., Partan, R. U., Firdaus, F., Tutuko, B., Pratiwi, F., Juliano, A. H., and Khoirani, R. (2020). Robust detection of atrial fibrillation from short-term electrocardiogram using convolutional neural networks. *Future Generation Computer Systems*, 113:304–317.
- [30] Pal, A., Srivastva, R., and Singh, Y. N. (2021). Cardionet: An efficient ecg arrhythmia classification system using transfer learning. *Big Data Research*, 26:100271.
- [31] Pandey, S. K. and Janghel, R. R. (2020). Automatic arrhythmia recognition from electrocar-

- diogram signals using different feature methods with long short-term memory network model. *Signal, Image and Video Processing*, 14(6):1255–1263.
- [32] Petmezas, G., Haris, K., Stefanopoulos, L., Kilintzis, V., Tzavelis, A., Rogers, J. A., Katsagelos, A. K., and Maglaveras, N. (2021). Automated atrial fibrillation detection using a hybrid cnn-lstm network on imbalanced ecg datasets. *Biomedical Signal Processing and Control*, 63:102194.
 - [33] Rahul, J., Sora, M., Sharma, L. D., and Bohat, V. K. (2021). An improved cardiac arrhythmia classification using an rr interval-based approach. *Biocybernetics and Biomedical Engineering*, 41(2):656–666.
 - [34] Ramirez, E., Melin, P., and Prado-Arechiga, G. (2019). Hybrid model based on neural networks, type-1 and type-2 fuzzy systems for 2-lead cardiac arrhythmia classification. *Expert Systems with Applications*, 126:295–307.
 - [35] Rao, K. D. and Swamy, M. N. (2018). *Digital signal processing: Theory and practice*. Springer.
 - [36] Rueda, C., Fernández, I., Larriba, Y., Rodríguez-Collado, A., and Canedo, C. (2021). Compelling new electrocardiographic markers for automatic diagnosis. *arXiv preprint arXiv:2112.12196*.
 - [37] Rueda, C., Rodríguez-Collado, A., Fernández, I., Canedo, C., Ugarte, M. D., and Larriba, Y. (2022). A unique cardiac electrophysiological 3d model. *arXiv preprint arXiv:2202.03938*.
 - [38] Santander-Baños, F., Hernández-Romero, N., Barragán-Vite, I., Karelin, O., Medina-Marín, J., et al. (2022). Clasificación de arritmias cardíacas mediante redes neuronales convolucionales y optimización por enjambre de partículas. *Pädi Boletín Científico de Ciencias Básicas e Ingenierías del ICBI*, 10(Especial2):42–55.
 - [39] Sharma, P., Dinkar, S. K., and Gupta, D. (2021). A novel hybrid deep learning method with cuckoo search algorithm for classification of arrhythmia disease using ecg signals. *Neural computing and Applications*, 33:13123–13143.
 - [40] Subasi, A. (2019a). *Practical guide for biomedical signals analysis using machine learning techniques: A MATLAB based approach*. Academic Press.
 - [41] Subasi, A. (2019b). *Practical guide for biomedical signals analysis using machine learning techniques: A MATLAB based approach*. Academic Press.
 - [42] Subhash, C., Mylarappa, N. K., and Sharan, M. (2024). Optimized ecg classification using wavelet decomposition and cnn’s. *International Advanced Research Journal in Science, Engineering and Technology*, 11(6):137–143.
 - [43] Tahsin, K., Hossain, M. F., and Rahman, M. A. (2023). Atrial fibrillation diagnosis from ecg signal using statistical and time-frequency attributes-based lstm network. *Proceedings of the International Conference on Computer and Information Technology*, pages 225–230.
 - [44] The MIT-BIH Arrhythmia Database (1999). MIT-BIH Arrhythmia Database.
 - [45] Tran, L., Li, Y., Nocera, L., Shahabi, C., and Xiong, L. (2020). Multifusionnet: atrial fibrillation detection with deep neural networks. *AMIA Summits on Translational Science Proceedings*, 2020:654.
 - [46] Villazana, S., Seijas, C., Montilla, G., and Pérez, E. (2021). Agrupamiento de señales eeg con rasgos aprendidos usando autoencoder profundo. *Revista Ingeniería UC*, 28(1):180–192.

- [47] Wang, J. (2020). A deep learning approach for atrial fibrillation signals classification based on convolutional and modified elman neural network. *Future Generation Computer Systems*, 102:670–679.
- [48] Wang, J. and Li, W. (2020). Atrial fibrillation detection and ecg classification based on cnn-bilstm. *arXiv preprint arXiv:2011.06187*.
- [49] Wenbo Wang, Jian Guan, X. C. and Wang, W. (2022). Ms-mlp: Multi-scale sampling mlp for ecg classification. In *Proceedings of the 2022 European Signal Processing Conference (EUSIPCO)*, pages 1288–1292, Belgrade, Serbia. IEEE, IEEE.
- [50] Xiao, Q., Lee, K., Mokhtar, S. A., Ismail, I., Pauzi, A. L. b. M., Zhang, Q., and Lim, P. Y. (2023). Deep learning-based ecg arrhythmia classification: A systematic review. *Applied Sciences*, 13(8):4964.
- [51] Yan, W. (2021). *Computational methods for deep learning*. Springer.
- [52] Zarei, A. and Asl, B. M. (2020). Performance evaluation of the spectral autocorrelation function and autoregressive models for automated sleep apnea detection using single-lead ecg signal. *Computer Methods and Programs in Biomedicine*, 195:105626.

Anexos

A Productos

En esta sección se darán detalles acerca de la producción científica lograda en el periodo de reporte.

- Se publicó el artículo titulado “*Atrial Fibrillation Classification Using a Deep Spectral Autoencoder*” en la revista Computación y Sistemas (CyS). La primera página del artículo se muestra en la Figura 6.1.

ISSN 2007-9737

Atrial Fibrillation Classification Using a Deep Spectral Autoencoder

Enrique Quezada-Próspero¹, Dante Mújica-Vargas^{1,*}, Luis A. Cruz-Próspero¹,
Christian García-Aquino², Ángel A. Rendón-Castro¹

¹ Centro Nacional de Investigación y Desarrollo Tecnológico,
Computer Sciences,
Mexico

² Universidad Politécnica de Tapachula,
Mexico

{m23ce074, dante.mv}@cenidet.tecnm.mx,
{lacr, angel.rendon18ce}@cenidet.edu.mx,
cj.garcia51@gmail.com

Abstract. This study introduces a novel approach for atrial fibrillation classification using a deep autoencoder stacked with a fully connected softmax layer. The model is trained with spectral features extracted from ECG signals through spectral and signal analysis. The primary goal is to enhance existing algorithms in the state-of-the-art by delivering superior results with reduced computational cost. The PhysioNet Challenge 2017 database was used, which contains normal and atrial fibrillation rhythms. The signals were normalized before spectral feature extraction. These features were used to train the autoencoder, which performed additional feature extraction and dimensionality reduction. The resulting features were then used to train the fully connected layer responsible for classification. Performance was evaluated through quality metrics mentioned in the state-of-the-art using cross-validation to ensure robustness. The best median results obtained were: 99.7% Accuracy, 99.8% Precision, 99.5% Recall, 99.8% Specificity, 99.7% F1-score, 99.3% Matthews Correlation, and 99.3% Kappa.

Keywords. Autoencoder, spectral features, ECG signals, signal analysis, feature extraction.

1 Introduction

Cardiac arrhythmias (CAs) are irregularities in the heart's contraction and relaxation cycles. Among

the most common types are Atrial Fibrillation (AF), Ventricular Fibrillation (VF), Atrial Flutter (AFL), Paroxysmal Supraventricular Tachycardias (PSVT), and Atrioventricular Block (AB). According to the World Health Organization (WHO), CAs are a leading cause of death worldwide, with a prevalence affecting approximately 3% of the global population [13].

The automatic classification of cardiac arrhythmias has garnered significant interest due to the severity of this condition. In the field of machine learning, detection techniques have been developed to support specialists.

These techniques are based on convolutional neural networks (CNN) [15], recurrent neural networks (RNN) [16], and support vector machines (SVM) [8], among others. Hybrid architectures combining these algorithms have also been employed to enhance model performance [1].

However, these models require large amounts of data and numerous hidden layers to extract and analyze features, leading to significant processing time and computational cost. This issue becomes problematic for real-world applications that necessitate dedicated components like GPUs to achieve satisfactory results [11]. Training these models is time-consuming and computationally

Figura 6.1: Primera página del Artículo 1

- Se publicó el artículo titulado “*Multiclass classification of ECG signals using a LSTM network trained with spectral features*” en el congreso WITCOM 2024. La constancia de participación se muestra en la Figura 6.2 y la primera página del artículo se muestra en la Figura 6.3.



Figura 6.2: Constancia de participación en el congreso WITCOM 2024



Multiclass Classification of ECG Signals Using a LSTM Network Trained with Spectral Features

Enrique Quezada-Prospero¹ , Dante Mújica-Vargas¹ ,
Luis A. Cruz-Prospero¹, Orvil A. Villarreal-Molina¹ ,
and Christian García-Aquino²

¹ Departamento de Ciencias Computacionales, TecNM/CENIDET, Cuernavaca,
Morelos, Mexico

{m23ce074,dante.mv,m23ce090}@cenidet.tecnm.mx, lacp@cenidet.edu.mx

² Universidad Politécnica de Tapachula, Tapachula, Chiapas, Mexico

Abstract. In this work, an implementation of a deep learning algorithm based on a LSTM network trained with spectral estimates obtained from ECG signals is presented, among which are: AR Burg spectrum, fast fourier transform spectrum, welch spectrum, Yule-Walker spectrum, periodogram spectrum, covariance spectrum. The objective of this work is to demonstrate that spectral estimates can generate good results when training an LSTM, taking into account the processing time, computational cost, the complexity of feature extraction and the value of quality metrics and based on the results, this model can be adapted to other types of approaches. The data used for training corresponds to the MIT-BIH database, which includes normal rhythm, auricular premature contraction, ventricular premature contraction, left bundle branch block and right bundle branch block. Using the signal analysis techniques mentioned above, the data was transformed into sequential data to improve the results, the corresponding features were extracted and a dimensionality reduction was applied. This features were used to train the BiLSTM network and for each train with different features, the results were validated and compared with the quality metrics. The spectral estimate that obtained the best results was FFT with values: 94.6% Accuracy, 94.6% Recall, 94.9% Precision, 94.7% F1-score, 93.4% Matthews and 83.3% Kappa.

Figura 6.3: Primera página del artículo 2

Keywords: LSTM · Spectral Estimates · ECG · Deep Learning ·
Computational Cost · Feature extraction · Dimensionality reduction

1 Introduction

The electrocardiogram (ECG) is an invaluable tool in medical practice, used to record the electrical activity of the heart and detect a wide range of cardiovascular disorders. The ECG has been instrumental in the evaluation and diagnosis of

- Se publicó el artículo titulado “*ECG Classification Using a CNN Stacked with an Autoencoder Trained with Synthetic Signals*” en la revista *Neural Computing and Applications*. La primera página del artículo se muestra en la Figura 6.3.

Neural Computing and Applications

S.I.: 2023 INDIA INTERNATIONAL CONGRESS ON COMPUTATIONAL INTELLIGENCE

ECG classification using a CNN stacked with an autoencoder trained with synthetic signals

Enrique Quezada-Prospero¹ · Dante Mújica-Vargas¹ · Antonio Luna-Álvarez² · Viridiana Vela Rincón¹ · Andrés A. Arenas Muñiz¹

Received: 29 September 2024 / Accepted: 5 April 2025

© The Author(s), under exclusive licence to Springer-Verlag London Ltd., part of Springer Nature 2025

Abstract

The detection of cardiac arrhythmias is crucial for preventing severe cardiovascular diseases. As a result, deep learning algorithms have been developed in the state-of-the-art for the automatic detection of these conditions. While these models have produced promising results, they rely on having sufficient data to generalize well. This presents a challenge in the medical field, where data scarcity is often an issue. To address these challenges, this work proposes a hybrid architecture combining a deep autoencoder with a 1D CNN, trained on synthetic signals and evaluated on real signals. The objective is to demonstrate that synthetic signals can complement or even replace real signals when there are insufficient data. The MIT-BIH arrhythmia database, containing five classes, was used for this process. The dataset was split to train a generative adversarial network for generating synthetic signals, while the remaining data were used to evaluate the model on real signals. Signal preprocessing involved transforming the ECG data into 2D representations using different methods, including similarity maps, Gramian angular fields, Markov transition fields, and co-occurrence matrices (CoMs), each of which was evaluated with the proposed model to observe the performance of each one. The extracted features were compressed by an autoencoder and then passed through a 1D CNN for classification. The best results in real data evaluation were obtained with CoMs, achieving an accuracy of 99.7%, precision of 99.7%, recall of 99.7%, specificity of 99.9%, F1-score of 99.7%, Matthews correlation coefficient of 99.6%, and Cohen's kappa of 99.3%.

Keywords ECG classification · Autoencoder · 2D signal transformations · Generative adversarial network

1 Introduction

Cardiac arrhythmias are one of the leading causes of mortality and morbidity worldwide. These irregular heart rhythms can indicate severe cardiovascular diseases, and their early detection is crucial in preventing serious complications such as stroke and heart failure [1]. Electrocardiography (ECG) is a widely used non-invasive diagnostic tool for monitoring and diagnosing arrhythmias. According to the World Health Organization (WHO), the prevalence of these conditions is about 3% of the world's population [2].

Interpreting ECG signals manually is labor-intensive and error-prone, underscoring the necessity for reliable and efficient automated arrhythmia classification methods. Consequently, various techniques have been developed for this purpose. Traditional methods typically extract morphological features from the P, QRS, and T waves and analyze inter-beat intervals (R-R) [3]. Despite their effectiveness in certain applications, these methods have notable limitations, such as variability in data across different databases and challenges in identifying key points in noisy signals.

Enrique Quezada-Prospero, Dante Mújica-Vargas, Antonio Luna-Álvarez, Viridiana Vela Rincón, Andrés A. Arenas Muñiz have contributed equally to this work.



Neural Computing and Applications

<https://doi.org/10.1007/s00521-025-11284-5>

Published online: 13 May 2025

Springer

Figura 6.4: Primera página del artículo 3

- Se presentó el póster titulado “Clasificación de arritmias cardíacas mediante un *Autoencoder* profundo espectral” en el 1er. coloquio de investigación, tecnología y sustentabilidad. La constancia de presentación se muestra en la Figura 6.5.



Figura 6.5: Comprobante de presentación del poster