



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA

**DSIC**  
DEPARTAMENT DE SISTEMES  
INFORMÀTICS I COMPUTACIÓ

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Dpto. de Sistemas Informáticos y Computación

Generación automática de señales de ECG en los  
asistentes cognitivos para mejorar la detección de  
problemas cardíacos

Trabajo Fin de Máster

Máster Universitario en Inteligencia Artificial, Reconocimiento de  
Formas e Imagen Digital

AUTOR/A: Girón Arevalo, María Fernanda

Tutor/a: Julian Inglada, Vicente Javier

Cotutor/a: Rincón Arango, Jaime Andrés

CURSO ACADÉMICO: 2022/2023



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA

Universitat Politècnica  
de València

**Departamento de Sistemas Informáticos y  
Computación**



Máster Universitario en Inteligencia Artificial, Reconocimiento  
de Formas e Imagen Digital

Trabajo Fin de Máster

**Generación automática de señales de ECG  
en asistentes cognitivos para mejorar la  
detección de problemas cardíacos**

Autor(a): María Fernanda Girón Arevalo  
Director(a): Vicente Javier Julián Inglada  
Cotutor(a): Jaime Andrés Rincón Arango

Valencia, Julio 2023

Este Trabajo Fin de Máster se ha depositado en el Departamento de Sistemas Informáticos y Computación de la Universitat Politècnica de València para su defensa.

*Trabajo Fin de Máster*

*Máster Universitario en Inteligencia Artificial, Reconocimiento de Formas e Imagen Digital*

*Título: Generación automática de señales de ECG en asistentes cognitivos para mejorar la detección de problemas cardíacos*

*Julio 2023*

*Autor(a):* María Fernanda Girón Arevalo

*Director(a):* Vicente Javier Julián Inglada

Departamento de Sistemas Informáticos y Computación  
Universitat Politècnica de València

*Co-director(a):* Jaime Andrés Rincón Arango

# Resum

Les malalties cardiovasculars són una de les principals causes de mortalitat a tot el món, amb un impacte significatiu en la població d'edat avançada i en països de baixos i mitjans ingressos. En particular, la fibril·lació auricular, una forma comuna d'arrítmia cardíaca, presenta desafiaments en el seu diagnòstic degut a la seua naturalesa asimptomàtica o intermitent.

La detecció precoç i el tractament adequat són fonamentals per a reduir la taxa de mortalitat causada per aquestes malalties. Considerant aquest context, els dispositius portàtils han guanyat interès com a eines de monitoratge en entorns clínics i investigatius. No obstant això, encara existeixen desafiaments en el desenvolupament de dispositius portàtils que integren mètodes automàtics per a la detecció de malalties cardíques a partir de senyals ECG adquirides de manera autònoma, especialment en el cas d'afeccions cardíques menys comunes.

El present treball es centra en el desenvolupament d'un model de classificació de senyals ECG basat en tècniques d'aprenentatge profund, amb la finalitat de ser implementat en un dispositiu portàtil que funcione com a assistent cognitiu. A més, es proposa la generació de senyals ECG sintètiques com a estratègia per a abordar el desequilibri de classes en els conjunts de dades disponibles i millorar el rendiment del model de classificació. Per a aconseguir aquests objectius, es duran a terme diferents etapes, que inclouen la revisió exhaustiva de la literatura existent, el disseny i avaluació de models de classificació de senyals, el desenvolupament d'algorismes eficients d'adquisició i classificació de senyals a implementar en el maquinari proposat, la construcció d'un model generatiu de senyals ECG i la millora del rendiment dels models de classificació mitjançant l'ús de conjunts de dades equilibrades.

**Paraules clau: arrítmies, senyals ECG, generació sintètica, model de classificació, GAN, assistents cognitius.**



# Resumen

Las enfermedades cardiovasculares son una de las principales causas de mortalidad en todo el mundo, con un impacto significativo en la población de edad avanzada y en países de bajos y medianos ingresos. En particular, la fibrilación auricular, una forma común de arritmia cardíaca, presenta desafíos en su diagnóstico debido a su naturaleza asintomática o intermitente.

La detección temprana y el tratamiento adecuado son fundamentales para reducir la tasa de mortalidad causada por estas enfermedades. Considerando este contexto, los dispositivos portátiles han ganado interés como herramientas de monitoreo en entornos clínicos e investigativos. Sin embargo, todavía existen desafíos en el desarrollo de dispositivos portátiles que integren métodos automáticos para la detección de enfermedades cardíacas a partir de señales ECG adquiridas de forma autónoma, especialmente en el caso de afecciones cardíacas menos comunes.

El presente trabajo se centra en el desarrollo de un modelo de clasificación de señales ECG basado en técnicas de aprendizaje profundo, con la finalidad de ser implementado en un dispositivo portátil que funcione como asistente cognitivo. Además, se propone la generación de señales ECG sintéticas como estrategia para abordar el desequilibrio de clases en los conjuntos de datos disponibles y mejorar el rendimiento del modelo de clasificación. Para lograr estos objetivos, se llevarán a cabo diferentes etapas, que incluyen la revisión exhaustiva de la literatura existente, el diseño y evaluación de modelos de clasificación de señales, el desarrollo de algoritmos eficientes de adquisición y clasificación de señales a implementar en el hardware propuesto, la construcción de un modelo generativo de señales ECG y la mejora del rendimiento de los modelos de clasificación mediante el uso de conjuntos de datos equilibrados.

**Palabras clave:** arritmias, señales ECG, generación sintética, modelo de clasificación, GAN, asistentes cognitivos.



# Abstract

Cardiovascular diseases are one of the main causes of mortality worldwide, with a significant impact on the elderly population and in low- and middle-income countries. In particular, atrial fibrillation, which is a common form of arrhythmia, presents challenges in its diagnosis due to its asymptomatic or intermittent nature.

Early detection and appropriate treatment are critical to reduce the mortality rate caused by these diseases. Considering this context, portable devices have gained interest as monitoring tools in clinical and research settings. However, there are still challenges in the development of portable devices that integrate automatic methods for the detection of cardiac diseases from autonomously acquired ECG signals, especially in the case of less common cardiac conditions.

This work focuses on the development of an ECG signal classification model based on deep learning techniques, with the aim of being implemented in a portable device that performs as a cognitive assistant. In addition, the generation of synthetic ECG signals is proposed as a strategy to address the imbalance of classes in the available datasets and to improve the performance of the classification model. To achieve these objectives, different stages will be carried out, including the exhaustive review of the existing literature, the design and evaluation of signal classification models, the development of efficient signal acquisition and classification algorithms to be implemented in the proposed hardware, the construction of a generative model of ECG signals, and the improvement of the performance of the classification models by using balanced datasets.

**Keywords: arrhythmias, ECG signals, synthetic generation, classification model, GAN, cognitive assistants.**





# Tabla de contenidos

<b>1. Introducción</b>	<b>1</b>
1.1. Motivación . . . . .	1
1.2. Objetivos . . . . .	2
1.3. Estructura de la memoria . . . . .	3
<b>2. Estado del Arte</b>	<b>5</b>
2.1. Electrocardiograma y señal ECG . . . . .	5
2.2. Clasificación señales electrocardiográficas (ECG) . . . . .	8
2.3. Técnicas empleadas para la generación sintética de secuencias temporales	9
2.3.1. Autoencoders (AE) . . . . .	9
2.3.2. Variational Autoencoders (VAE) . . . . .	10
2.3.3. Redes generativas adversarias (GAN) . . . . .	11
<b>3. Preprocesamiento de señales ECG</b>	<b>15</b>
3.1. Base de datos de señales ECG . . . . .	15
3.2. Preprocesamiento . . . . .	17
<b>4. Clasificación señales ECG</b>	<b>21</b>
4.1. Descripción y resultados del primer modelo de clasificación propuesto .	21
4.2. Descripción y resultados del segundo modelo de clasificación propuesto	23
<b>5. Adquisición y análisis de señales ECG</b>	<b>27</b>
5.1. Descripción del hardware . . . . .	27
5.1.1. Cardio chip BMD101 . . . . .	27
5.1.2. M5Stick-C Plus . . . . .	27
5.2. Algoritmo para lectura de datos . . . . .	28
5.3. Descripción del modelo de clasificación . . . . .	29
5.4. Algoritmo para clasificación de datos. . . . .	31
5.5. Resultados de clasificación en el dispositivo . . . . .	32
<b>6. Generación de señales ECG</b>	<b>35</b>
6.1. Desarrollo red GAN . . . . .	35
6.1.1. Descripción del modelo generador . . . . .	35
6.1.2. Descripción del modelo discriminador . . . . .	36
6.2. Resultados primera experimentación . . . . .	37
6.3. Preprocesamiento de datos . . . . .	40
6.4. Resultados segunda experimentación. . . . .	41
<b>7. Resultados de clasificación con señales ECG generadas sintéticamente</b>	<b>45</b>

7.1. Base de datos PTB Diagnostic . . . . .	45
7.2. Base de datos del Computing in Cardiology Challenge 2017 . . . . .	47
<b>8. Conclusiones</b>	<b>51</b>
<b>Bibliografía</b>	<b>56</b>

# Capítulo 1

## Introducción

### 1.1. Motivación

Según la Organización Mundial de la Salud (OMS) las enfermedades cardiovasculares son la principal causa de mortalidad y discapacidad en el mundo. Provocando la muerte de alrededor de 17,7 millones de personas al año, representando un 31 % de todas las muertes registradas en el mundo. Estas enfermedades tienden a afectar en su mayoría a la población de edad avanzada o con trastornos de obesidad, sin embargo, se estima que una de cada 10 personas de 30 a 70 años muere a causa de enfermedades cardiovasculares en el mundo. Sobre todo en países de bajos y medianos ingresos, que representan el 80% de estas muertes, ya que no cuentan con los recursos económicos y humanos para hacer frente a los retos que implican estas enfermedades. Se ha demostrado que la detección precoz y el tratamiento temprano, puede contribuir a la reducción del índice de mortalidad causado por estas enfermedades [1],[2].

La fibrilación auricular es un tipo de alteración del ritmo cardíaco que destaca por ser el tipo de arritmia más común en adultos con un índice de morbilidad significativo. Además, se asocia con mayor riesgo de desarrollar complicaciones graves, como accidentes cardiovasculares. La detección de la fibrilación auricular plantea un desafío debido a su naturaleza asintomática o intermitente, lo que requiere un monitoreo periódico de la frecuencia cardíaca para lograr un diagnóstico preciso de esta afección [3].

En vista de esta información, es evidente la necesidad de innovar y desarrollar tecnologías que contribuyan al diagnóstico temprano de estas enfermedades. Ante esta necesidad, los dispositivos portátiles han ganado un considerable interés como herramientas de monitoreo tanto en entornos de investigación como en aplicaciones clínicas. Sin embargo, aún existe un reto y un amplio margen de desarrollo en lo que respecta a dispositivos portátiles que integren métodos automáticos para detectar enfermedades cardíacas a partir de señales ECG adquiridas de forma autónoma. Sobre todo, cuando se consideran afecciones cardíacas poco comunes [4].

Durante los últimos años, se han propuesto y aplicado varios métodos para la clasificación de arritmias a partir de señales de ECG, desde enfoques basados en aprendizaje automático, en donde paso a paso se realiza el procesamiento, extracción de características y clasificación de los datos. Hasta técnicas basadas en aprendiza-

je profundo, en donde la extracción de características y la clasificación se realizan simultáneamente. En esta categoría se pueden mencionar arquitecturas como las redes neuronales convolucionales (CNN), las redes *Long short-term memory* (LSTM), autoencoders (AE), entre otras. Sin embargo, estos métodos de clasificación no toman en cuenta el desequilibrio de clases en los datos usados para entrenamiento, lo que generalmente conlleva a un modelo cuya predicción está sesgada hacia la clase mayoritaria [5].

El problema de desequilibrio de clases es un tema cada vez más relevante, ya que los conjuntos de datos desequilibrados entre clases existen en muchos contextos del mundo real. Y en algunos de estos contextos, resulta de mayor importancia la clasificación precisa en las clases minoritarias. Como es el caso de la clasificación de arritmias en señales ECG, en donde, es sumamente común encontrar conjuntos de datos en donde algunas afecciones cardíacas, como la fibrilación auricular o ciertos tipos de arritmias, son menos frecuentes en comparación con el ritmo sinusal normal, por lo que es crucial lograr una mayor sensibilidad en la detección de las clases minoritarias [6].

Este trabajo se enfocará en explorar técnicas de aprendizaje profundo, para el desarrollo de un modelo de clasificación de arritmias que pueda facilitar la detección de enfermedades cardíacas y técnicas de generación de señales sintéticas con el fin de abordar el problema de desequilibrio de clases en bases de datos de señales ECG y mejorar el rendimiento del modelo de clasificación. Se espera que este trabajo final de máster contribuya al avance en el campo de la detección y clasificación de arritmias, brindando una herramienta eficaz y precisa para el diagnóstico temprano de enfermedades cardíacas.

## 1.2. Objetivos

El objetivo principal del presente Trabajo Final de Máster consiste en desarrollar un modelo de generación de señales ECG que aborde el desafío que presenta trabajar con conjuntos de datos altamente desbalanceados y en consecuencia, permita diseñar un modelo de clasificación con resultados menos sesgados, que incremente la fiabilidad y efectividad en la detección de problemas cardíacos en asistentes cognitivos.

Este objetivo principal se subdivide en los siguientes objetivos específicos:

1. Revisión bibliográfica de trabajos relacionados con la clasificación de señales ECG y de estudios referentes al diseño de modelos generativos para la generación sintética de datos.
2. Diseñar y evaluar el rendimiento de modelos de clasificación de registros de ECG de una derivación, utilizando diferentes bases de datos de señales ECG.
3. Investigar y evaluar opciones de componentes (microcontroladores, microchips, pantallas, etc) que formarán parte del hardware utilizado para la adquisición y análisis de señales ECG.
4. Desarrollar un algoritmo de adquisición de señales y un algoritmo de clasificación de señales que sean eficientes y puedan ser integrados en un microcontrolador.

5. Construir e implementar un modelo generativo que genere señales sintéticas que capturen las características morfológicas presentes en señales ECG reales.
6. Mejorar el rendimiento de los modelos de clasificación, utilizando conjuntos de datos equilibrados como muestras de señales ECG generadas sintéticamente.

### 1.3. Estructura de la memoria

El presente trabajo se estructura en una serie de capítulos que abordan diferentes aspectos relacionados con el análisis y clasificación de señales ECG. El primer capítulo de este trabajo es la introducción, en donde se presenta la motivación principal que dio origen a esta investigación y los objetivos que se espera cumplir al culminarla. En el capítulo 2, se realiza un análisis del estado del arte, explorando acerca de los fundamentos del ECG y su relevancia en el diagnóstico y monitoreo de enfermedades cardíacas. Además, se analizan estudios relevantes sobre técnicas y modelos utilizados en la clasificación de señales ECG y en la generación sintética de datos. A continuación, en el capítulo 3 se describen las bases de datos utilizadas en este trabajo y se detalla el procesamiento realizado para preparar los datos para su posterior análisis y clasificación.

En el capítulo 4 se presentan dos modelos de clasificación empleando redes neuronales convolucionales (CNN) y redes residuales (ResNet), respectivamente, y se analizan los resultados obtenidos con ambos modelos. Posteriormente, en el capítulo 5, se describe el hardware utilizado y el algoritmo implementado para la adquisición y clasificación de datos, resaltando las consideraciones y limitaciones que pueden existir en este proceso.

En el capítulo 6 se propone el uso de redes generativas adversarias (GAN) para la generación de señales ECG sintéticas con el objetivo de abordar el desbalanceo de clases en las bases de datos y mejorar el rendimiento de los modelos de clasificación. Seguidamente, en el capítulo 7, se evalúa el rendimiento de los modelos de clasificación utilizando conjuntos de datos balanceados, que combina señales ECG reales y sintéticas. Por último, en el capítulo 8, se discuten las conclusiones derivadas de los resultados obtenidos en este trabajo.



## Capítulo 2

# Estado del Arte

En este capítulo se presenta un análisis del estado del arte en el campo del análisis de señales electrocardiográficas (ECG). Se abordan aspectos fundamentales relacionados con el electrocardiograma y la morfología de la señal ECG, destacando su importancia en el diagnóstico y monitoreo de enfermedades cardíacas. Además, se revisan estudios relevantes, en los cuales se exploran diferentes técnicas y modelos utilizados para el procesamiento y la clasificación de señales ECG, así como para la generación sintética de señales o series temporales. El objetivo principal es proveer una visión general del estado actual de la investigación en este ámbito y sentar las bases para el desarrollo de nuevas soluciones.

### 2.1. Electrocardiograma y señal ECG

El análisis de señales electrocardiográficas (ECG) desempeña un papel fundamental en el campo de la medicina y la ingeniería biomédica y ha sido objeto de muchos estudios e investigaciones, debido a la importancia de estas señales en el diagnóstico y monitoreo de enfermedades cardíacas.

Un electrocardiograma es una representación gráfica de los impulsos eléctricos que se generan durante cada ciclo cardíaco, esta información se captura de manera no invasiva, utilizando dos electrodos o más que se colocan sobre la piel del paciente, en puntos estratégicos del tórax, para registrar la actividad eléctrica del corazón en forma de ondas [7].

Generalmente, un electrocardiograma se compone de 12 derivaciones, que permiten observar la actividad eléctrica del corazón desde 12 perspectivas diferentes [8]. Estas derivaciones están compuestas por:

- **Derivaciones precordiales:** Proporcionan una vista frontal del corazón y registran la actividad eléctrica en diferentes regiones del ventrículo izquierdo y derecho. Estas derivaciones se dividen en V1, V2, V3, V4, V5 y V6.
- **Derivaciones unipolares:** Estas derivaciones son aVR, aVL y aVF. Y se encargan de medir el potencial absoluto en ciertas extremidades, tomando como referencia un punto en común.
- **Derivaciones bipolares:** Registran la diferencia de potencial eléctrico entre dos extremidades del cuerpo, se dividen en DI, DII, DIII.



## 2.1. Electrocardiograma y señal ECG

El uso de 12 derivaciones en un ECG facilita el diagnóstico de enfermedades cardíacas, ya que al proporcionar una visión global de la actividad eléctrica del corazón desde varios ángulos, permite detectar anomalías que afectan a una parte concreta del corazón que pueden no ser advertidas desde una derivación, pero sí desde otra. En la figura 2.1 se observa la distribución de los electrodos conocida como triángulo de Einthoven, utilizada para capturar las derivaciones bipolares [9].

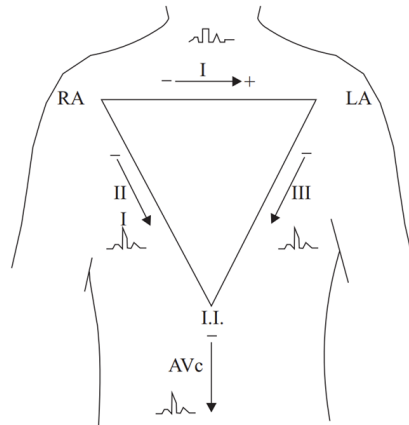


Figura 2.1: Triángulo de Einthoven [9].

Las ondas representativas de las señales ECG se originan debido a los cambios en el potencial eléctrico generados por la despolarización y la repolarización de las células cardíacas. Las distancias entre deflexiones u ondas son denominadas segmentos o intervalos. El trazo normal de un ciclo cardíaco consta de una onda P, un complejo QRS, una onda T, tal como se observa en la figura 2.2. En algunos casos incluso puede ser visible una pequeña onda U.

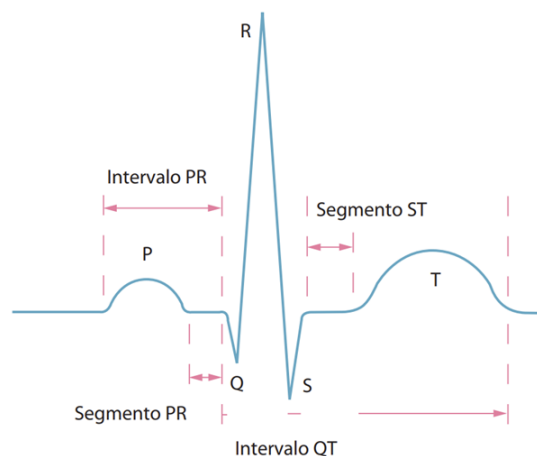


Figura 2.2: ECG y sus formas de onda más representativas [7].

A continuación, se presenta una descripción más a detalle de la ondas, intervalos, complejos y segmentos que se considera deben estar presentes en el trazo ECG de un individuo normal.

- **Onda P:** Representa la despolarización auricular.

- **Onda T:** Representa la repolarización de los ventrículos.
- **Complejo QRS:** Representa la despolarización de los ventrículos.
- **Segmento ST:** Es la parte del trazado que se extiende desde el final del complejo QRS hasta el inicio de la onda T. Representa el periodo de tiempo en que los ventrículos permanecen en estado activado y puede comenzar la repolarización ventricular.
- **Intervalo PR:** Corresponde al intervalo de tiempo que abarca desde el inicio de la onda P hasta el inicio del complejo QRS. Representa el tiempo de conducción eléctrica desde el inicio de activación de las aurículas hasta la activación de los ventrículos.
- **Intervalo QT:** Representa el intervalo de tiempo que va desde el comienzo del complejo QRS hasta el final de la onda T. Y refleja la duración de la sístole eléctrica, que representa el período en el que los ventrículos se contraen y se relajan.
- **Intervalo RR:** Representa la duración de un latido cardíaco completo, es decir, el tiempo que tarda el corazón en completar un ciclo de contracción y relajación. En el trazado ECG corresponde al intervalo de tiempo entre la onda R de un complejo QRS y la onda R del siguiente complejo QRS [10].

Las señales ECG tienen un espectro frecuencial que abarca desde los 0 Hz hasta los 100 Hz, con amplitudes que oscilan entre 0 y 5 mV [11]. Generalmente, las señales ECG adquiridas deben pasar por una etapa de preprocesamiento antes de ser utilizadas para realizar cualquier tipo de diagnóstico clínico. Esto se debe a la presencia de elementos ajenos no deseados en las señales, causados por diferentes fuentes de ruido externas, como el ruido generado por el contacto electrodo-piel, el ruido generado por aparatos electrónicos utilizados en el proceso de adquisición, el ruido ocasionado por las señales eléctricas de los músculos (EMG), entre otros factores [12].

En cuanto a sistemas comerciales electrocardiógrafos, existen diversas alternativas para la adquisición y acondicionamiento de señales ECG. Sin embargo, hay características y capacidades fundamentales que debe disponer el sistema para garantizar la precisión y calidad de los datos recopilados. En general, estos sistemas de adquisición utilizan amplificadores de instrumentación en las etapas iniciales para amplificar la señal captada, además de combinaciones de filtros analógicos, como filtros pasa bajas y pasa altas, para eliminar el ruido no deseado y resaltar las características relevantes del electrocardiograma. Otro componente muy importante en estos sistemas es el microcontrolador. En primer lugar este debe contar con una capacidad de procesamiento de señal de alto rendimiento. Además, debe incorporar un convertidor analógico-digital de alta resolución para capturar con precisión las señales analógicas. Y contar con la capacidad de transmitir señales a otros dispositivos [13].

En conclusión, el electrocardiograma es una herramienta clave para el registro y análisis de la actividad eléctrica del corazón, proporcionando información valiosa sobre el ritmo cardíaco y la función cardíaca. Además, el análisis adecuado de las ondas y componentes presentes en una señal ECG, desempeña un papel crucial en el diagnóstico temprano de diferentes afecciones cardíacas, como arritmias, enfermedades coronarias y alteraciones en la conducción eléctrica del corazón, contribuyendo así a mejorar la salud cardiovascular de los pacientes.

## 2.2. Clasificación señales electrocardiográficas (ECG)

Una señal electrocardiográfica (ECG) es una representación gráfica de la actividad eléctrica del corazón. Estas señales se obtienen principalmente mediante la colocación de electrodos en la piel del paciente, que capturan las corrientes eléctricas que se generan cuando el corazón se contrae y se relaja.

Las señales ECG son una herramienta clave en la evaluación y seguimiento de la salud cardíaca, ya que mediante su análisis se puede obtener la siguiente información relevante. Primero, con el tiempo que tarda la onda eléctrica en pasar a través del sistema de conducción eléctrica del corazón, se puede determinar si la actividad eléctrica es regular o irregular, rápida o lenta. En segundo lugar, midiendo la fuerza de la actividad eléctrica, se puede determinar si partes del corazón son demasiado grandes o están sobrecargadas.

Cualquier trastorno en la actividad eléctrica de las células nerviosas del corazón se ve reflejado en las señales ECG. Por lo que, son de suma importancia para la detección y diagnóstico de arritmias cardíacas. El análisis repetitivo de señales de este tipo es una tarea altamente compleja y que requiere un considerable esfuerzo temporal y técnico. Para superar esta limitación, se han desarrollado tecnologías de diagnóstico asistido por ordenador, que permiten analizar los datos del ECG de manera automatizada y precisa. A continuación se describen algunas de las técnicas de aprendizaje automático y aprendizaje profundo que se han implementado para la detección de arritmias, a partir de clasificación de señales ECG.

En un estudio realizado por [14], se empleó la transformada de *wavelet* discreta para identificar el complejo QRS de la señal, descomposición empírica en modos para reducción del ruido y máquinas de vectores de soporte (SVM) para una tarea de clasificación de cinco tipos de arritmia. En [15], se emplea la técnica *high-order spectra* para extracción de características, la técnica *Principal Component Analysis* (PCA) para reducir la dimensionalidad del conjunto y un modelo basado en SVM para clasificación de cinco diferentes tipos de arritmia, con un valor de precisión igual a 99.28%. En [16], se implementó un método de clasificación jerárquica basado en la técnica *weighted extreme gradient boosting* (XGBoost). En donde, se extrajo un gran número de características a señales previamente procesadas, luego se redujo esta cantidad utilizando una técnica de eliminación recursiva y por último se implementó un clasificador jerárquico para seis diferentes categorías, con el cual se obtuvo un 92.1% de precisión.

Se han llevado a cabo varios estudios que se centran en la base de datos de arritmias MIT-BIH. Entre ellos, destaca el trabajo de Kumar et al. [17], en donde se propone un clasificador basado en la técnica de *random forest* y la transformada de coseno discreta para identificar los intervalos R-R de las señales ECG y utilizarlos como características relevantes en el proceso de clasificación. Otro estudio notable es el de Park et al. [18], en donde, inicialmente emplean el método Pan-Tompkins para identificar las ondas P y el complejo QRS de las señales y posteriormente emplean un clasificador KNN para detectar arritmias en 17 tipos diferentes de latidos cardiacos.

Los enfoques de aprendizaje automático tradicionales implican la extracción y filtrado manual de características. Por el contrario, las técnicas de aprendizaje profundo, en particular las redes neuronales convolucionales (CNN), tienen la capacidad de aprender automáticamente características a partir de datos con un mínimo o nulo procesa-

miento y sin necesidad de extraer características de forma manual. Como resultado, se ha visto un aumento en el uso de técnicas de aprendizaje profundo para la clasificación de arritmias cardíacas. Como es el caso de Acharya et al. [19, 20, 21, 22], en donde se realizó un estudio exhaustivo sobre la clasificación de arritmias mediante técnicas de aprendizaje profundo. Inicialmente, en [21] se propone un clasificador automatizado para la detección de arritmias. Más adelante, en [20], se implementa una arquitectura de red neuronal convolucional (CNN) para predecir el infarto de miocardio con una precisión igual a 95.22%. Mientras que, en [22] se propone un modelo CNN automatizado para la categorización de arritmias ventriculares desfibrilables y no desfibrilables, que logra una precisión del 93.18%.

Por otra parte, en el estudio referenciado en [23], se utilizó una red tipo ResNet para la clasificación automática de señales ECG. Cada dato de entrada constaba de dos latidos, mejorando la capacidad del modelo para aprender las características relevantes. Además, en este trabajo se abordó el problema del desequilibrio de datos, empleando una técnica de pérdida ponderada por *batches*, que eliminó la necesidad de realizar un complejo proceso de *data augmentation*. Similar a [23] en [24], se propuso un nuevo marco para el análisis de señales ECG que permite representar la señal de manera transferible entre distintas tareas. Para ello, se empleó una arquitectura con redes residuales que es capaz de aprender tales representaciones y adaptarse a la tarea de detección de arritmias cardíacas.

### 2.3. Técnicas empleadas para la generación sintética de secuencias temporales

Uno de los desafíos que se presenta al trabajar con bases de datos de señales ECG es el desequilibrio de clases, es decir, la distribución desigual de las muestras en diferentes categorías. Este desequilibrio se origina debido a varios factores. Siendo el factor principal el hecho de que ciertas afecciones cardíacas son menos comunes que otras, lo que se refleja en una menor cantidad de muestras disponibles para estas clases.

Este desequilibrio de clases, representa un desafío en el desarrollo de modelos de clasificación de señales ECG. Ya que el modelo no recibirá suficiente información de las clases subrepresentadas, lo que se reflejara en un modelo sesgado hacia la clase mayoritaria con menos capacidad de capturar patrones y características de las clases minoritarias. Por lo tanto, abordar el desafío del desbalance de clases en bases de datos de señales ECG es crucial para lograr un análisis confiable y asegurar una detección precisa de problemas cardíacos.

En esta sección se resumen algunas de las técnicas empleadas para la generación de secuencias temporales sintéticas, con el objetivo de mitigar el sesgo hacia la clase mayoritaria.

#### 2.3.1. Autoencoders (AE)

Los Autoencoders (AE) son un tipo de redes neuronales artificiales que se utilizan en el aprendizaje no supervisado, y han sido aplicados en la generación de datos sintéticos. Su arquitectura se compone de múltiples redes neuronales conectadas: un codificador y un decodificador. La red codificadora procesa la secuencia de entrada

### **2.3. Técnicas empleadas para la generación sintética de secuencias temporales**

y la comprime a una representación latente de menor dimensión. Mientras que, el decodificador se encarga de reconstruir la secuencia de entrada a partir de la representación aprendida por el decodificador [26].

Este tipo de redes se ha empleado en estudios como el mencionado en [25], en donde se propone un modelo AE con extrapolador (ExtraMAE) y un entrenamiento autosupervisado, para la generación de series temporales.

Al generar datos sintéticos con autoencoders, es importante tener en cuenta la calidad y diversidad de los datos de entrenamiento. Un autoencoder bien entrenado puede capturar las características importantes de los datos de entrada y generar datos sintéticos que sean similares a los datos reales. Sin embargo, es necesario tener cuidado de no generar datos que simplemente reproduzcan los patrones existentes en los datos de entrenamiento sin añadir variabilidad o explorar nuevas instancias

#### **2.3.2. Variational Autoencoders (VAE)**

Los Variational Autoencoders (VAE) son modelos generativos que mezclan las redes neuronales con distribuciones de probabilidad, con el fin de generar datos sintéticos que siguen los mismos patrones que el conjunto de datos empleado para su entrenamiento. Los VAE poseen similitudes en su arquitectura con los autoencoders al estar compuestos por dos componentes principales: un codificador y un decodificador. Sin embargo, se diferencian por que el autoencoder variacional proporciona una forma estadística de describir las muestras del conjunto de datos en el espacio latente, lo que lo hace un modelo más apto para la generación de datos sintéticos en comparación a los AE [29].

En un VAE, el codificador tiene como función mapear los datos de entrada a un espacio latente de menor dimensión, que generalmente es representado por una distribución gaussiana que captura las características principales de los datos de entrada. Mientras que, el decodificador toma como entrada la salida mapeada por el codificador para generar datos sintéticos que se asemejen lo más posible a la entrada original. Los resultados se propagan desde la red neuronal en forma de una función de pérdida, que mide la similitud entre un objeto decodificado y el objeto que se codifica [28].

La aplicación de autoencoders variacionales ha tenido una amplia atención, sobre todo en el campo de la generación de imágenes sintéticas, debido a su capacidad de capturar características visuales y generar muestras de alta calidad. Como es el caso de las investigaciones realizadas en [27], en donde se emplea un VAE para la generación sintética de bocetos a partir de dibujos secuenciales. O el estudio realizado en [28], en donde proponen superar el problema de desequilibrio de clases entre imágenes de resonancias magnéticas de sujetos con enfermedad de Parkinson y pacientes sanos, utilizando como modelo generativo un VAE que genere imágenes sintéticas de resonancias magnéticas y así mejora la detección de la enfermedad de Parkinson. Y aunque la generación de imágenes ha sido el enfoque más común de los VAE, estos modelos también se han adaptado a diferentes tipos de datos. Como en el estudio realizado en [29], en donde se propone una novedosa arquitectura que emplea autoencoders variacionales para la generación de datos de series temporales multivariantes. Y la investigación realizada en [30], en donde reportan resultados inconsistentes al utilizar autoencoders variacionales para generar ondas sinusoida-

les. En la investigación mencionada anteriormente, se determinó que los resultados obtenidos con VAEs eran claramente inferiores a los obtenidos con otros modelos generadores, en este caso específico las Recurrent Conditional GAN (RCGAN).

### 2.3.3. Redes generativas adversarias (GAN)

Las redes generativas adversarias (por sus siglas en inglés GAN) son un tipo de redes neuronales utilizadas en el campo de la inteligencia artificial para la generación de datos sintéticos que se asemejen a los datos reales de entrenamiento. Las redes GAN se componen de dos partes principales: el generador  $G$  y el discriminador  $D$ .

El generador consiste en una red neuronal que toma como entrada un ruido aleatorio  $z$ , que es forzado por el generador para producir una salida de datos sintéticos que modelen la distribución de los datos reales. Por otra parte, el discriminador es una red neuronal que toma como entrada tanto datos reales como datos generados sintéticamente, y está entrenada para clasificar correctamente si estos datos son reales o falsos.

El proceso de entrenamiento de una red GAN se basa en la competencia entre el generador y el discriminador. En donde el objetivo del generador es maximizar la tasa de fallos del discriminador, generando muestras cada vez más semejantes a las reales. Mientras que, el discriminador pretende minimizar esta tasa de fallos, al ser entrenado para diferenciar con la mayor precisión posible entre las muestras generadas sintéticamente y las reales. Este comportamiento se puede describir por la función 2.1, en donde  $D_x$  representa la probabilidad de que  $x$  pertenezca a los datos reales [33].

$$\min_G \max_D V(G, D) = E_{x \sim p_{\text{data}}(x)}[\log D(x)] + E_{z \sim p_z(z)}[\log(1 - D(G(z)))] \quad (2.1)$$

A lo largo del tiempo, se han realizado muchas investigaciones relacionadas al problema del desequilibrio de datos, desde implementaciones de técnicas de remuestreo de datos, hasta combinaciones de métodos de sobremuestreo con técnicas de aprendizaje profundo semisupervisado [34, 36]. Sin embargo, estas técnicas de aumento de datos que se basan principalmente en remuestrear las señales originales, demostraron tener mejoras limitadas. Otros acercamientos, como el implementado en [37], se basan en la creación de modelos dinámicos para generar señales ECG sintéticas. En resumen, estos modelos constan de tres ecuaciones diferenciales ordinarias acopladas y un usuario debe detallar las características específicas de las señales que se desea generar, lo que implica que se requiere de expertos con conocimiento especializado en el dominio para la generación de datos sintéticos. Por estas razones y con el fin de eliminar estas dependencias, se introduce la aplicación de redes generativas adversarias (GAN) como un nuevo acercamiento para solucionar el problema de aprendizaje con datos desequilibrados.

Las redes GAN fueron implementadas inicialmente para la generación de muestras sintéticas de imágenes y en el transcurso de los años se ha hecho hincapié al desarrollo de GAN enfocadas en dicha tarea, lo que ha generado que su aplicación tenga un gran éxito especialmente en el campo de la visión por computador. A pesar de ello, no se ha explorado profundamente la contribución que estas técnicas en la generación de series temporales, especialmente, series temporales fisiológicas que en el

### 2.3. Técnicas empleadas para la generación sintética de secuencias temporales

campo de la medicina pueden ser de suma utilidad para el desarrollo de sistemas de clasificación y diagnóstico de enfermedades más eficaces.

A continuación, se describen algunos de los trabajos más destacados en donde se implementa el uso de redes generativas adversarias (GAN) para la generación de señales temporales. En el estudio realizado por Delaney et al. [38], se experimentó con una serie de arquitecturas de redes GAN para la generación de señales sintéticas. En este estudio, se implementaron como métricas de evaluación los valores de *Maximum Mean Discrepancy* (MMD) y *Dynamic Time Warping* (DTW). En [39], se implementa una variante del modelo GAN conocida como ACGAN, que incorpora una red auxiliar clasificadora para generar latidos sintéticos representativos de señales ECG. En el caso de [40], se generaron señales ECG sintéticas con una red GAN con arquitectura basada en LSTMs bidireccionales en el generador y en el discriminador. Y como métrica de evaluación el valor MMD. El uso de redes LSTM (*Long Short Term Memory*) se plantea con el fin de tratar las limitaciones que presentan las redes neuronales convencionales, ya que las redes LSTM utilizan una unidad de memoria adicional que les permite manejar dependencias a largo plazo, lo que contribuye a prevenir el problema de desvanecimiento del gradiente [46]. Similar a este estudio, en [41], se propuso el uso de aprendizaje profundo para generar señales electrocardiográficas sintéticas utilizando una arquitectura que combina redes LSTM bidireccionales con redes neuronales convolucionales (CNN). Por otra parte, en [42], se propone la implementación de un modelo GAN basado en atención, con el objetivo de generar señales ECG que contribuyan a la anonimización de datos médicos. Empleando como métricas de evaluación valores como DTW, *Root Mean Square Error* (RMSE), *Frechet distance* (FD), entre otros.

Cabe destacar que, como se observa en los trabajos mencionados anteriormente, no existe un consenso en cómo evaluar los datos generados utilizando redes GAN. Sin embargo, se suele medir la distorsión entre señales reales y señales sintéticas utilizando métricas comunes como lo son:

- **Root Mean Square Error (RMSE):** Es una métrica utilizada comúnmente para evaluar la precisión de modelos de regresión o pronóstico, que calcula la diferencia entre los valores observados y los valores predichos por un modelo o estimación mediante la ecuación 2.2 [43].

$$RMSE = \sqrt{\frac{1}{T} \sum_{t=1}^T (x_t - x'_t)^2} \quad (2.2)$$

- **Dynamic Time Warping (DTW):** Es una métrica empleada para estimar la similitud entre dos secuencias de datos temporales. Esta técnica deforma las series a lo largo del eje temporal, con el fin de encontrar una alineación óptima entre los puntos de tiempo de las series y medir la distancia entre ellas, mientras menor sea esta distancia más similares serán las secuencias comparadas [44]. Al emplear DTW el costo acumulado se calcula de la siguiente manera:

$$D_{i,j} = f(x_i, y_j) + \min \{D_{i,j-1}, D_{i-1,j}, D_{i-1,j-1}\}$$

Para,  $i = 1, \dots, N$  y  $j = 1, \dots, M$  en donde  $N$  y  $M$  son las longitudes de las series  $x$  e  $y$ , respectivamente. Y por lo general,  $f(x_i, y_j) = (x_i - y_j)^2$  [45].

## Estado del Arte

---

Esta métrica es comúnmente empleada en aplicaciones de reconocimiento de patrones, reconocimiento de voz, procesamiento de señales, análisis de señales biomédicas, análisis de señales musicales y otras comparaciones de series temporales en general.

El estado del arte en el campo del análisis de señales ECG ha demostrado la importancia de estas señales en el campo de la medicina y la ingeniería biomédica, ya que desempeñan un papel crucial en el diagnóstico y monitoreo de enfermedades cardíacas. Además, ha proporcionado una visión general sobre las técnicas y enfoques utilizados en la clasificación y generación de señales ECG. A pesar de que el uso de autoencoders variacionales ha demostrado resultados prometedores en la generación de datos sintéticos, estudios como los realizados por Desai et al. [29] y Esteban et al. [30], han demostrado que esta técnica puede llegar a presentar resultados inconsistentes en la generación de datos de series temporales. En consecuencia, se concluyó que las redes generativas adversarias (GAN) son la metodología más adecuada para abordar esta aplicación específica. Por lo tanto, en este trabajo se dará especial énfasis al uso de las redes GAN como enfoque principal para la generación de datos.





## Capítulo 3

# Preprocesamiento de señales ECG

En este capítulo se presenta una descripción detallada de las bases de datos de señales ECG empleadas para el desarrollo de este trabajo. Se explica la importancia de estos conjuntos de datos en el campo de la clasificación de arritmias cardíacas y se detalla la distribución de clases en cada uno de ellos. Además, se describe el procesamiento realizado en las señales ECG con el fin de extraer latidos representativos y preparar los datos para el entrenamiento del modelo de clasificación.

### 3.1. Base de datos de señales ECG

Para llevar a cabo este trabajo, se requirió una gran cantidad de registros de señales ECG etiquetadas, por lo que se trabajó con tres bases de datos diferentes. La base de datos del *Massachusetts Institute of Technology-Beth Israel Hospital (MIT-BIH Arrhythmia)*<sup>1</sup>, la base de datos *PTB Diagnostic ECG*<sup>2</sup> y la proporcionada por *PhysioNet* para el *Computing in Cardiology Challenge 2017*<sup>3</sup>.

La base de datos *MIT-BIH Arrhythmia* consta de registros de ECG de 47 sujetos, las señales están grabadas a una frecuencia de muestreo de 360Hz con una resolución de 5  $\mu$ V/bit y han sido anotadas por al menos dos cardiólogos. En total, cuenta con 109446 muestras, divididas en quince tipos de latidos que de acuerdo con la Asociación para el Avance de la Instrumentación Médica (AAMI), se pueden agrupar en cinco categorías diferentes. Esta base de datos es ampliamente utilizada en el campo de la clasificación de arritmias cardíacas y ha sido objeto de numerosos estudios para evaluar el rendimiento de diferentes técnicas de clasificación de señales ECG.

En el cuadro 3.1 se detalla la distribución de clases según la Asociación para el Avance de la Instrumentación Médica (AAMI), para la base de datos *MIT-BIH Arrhythmia*. Y en la figura 3.1 se observa gráficamente la distribución de los datos de entrenamiento según su categoría.

---

<sup>1</sup><https://www.physionet.org/content/mitdb/1.0.0/>

<sup>2</sup><https://www.physionet.org/content/ptbdb/1.0.0/>

<sup>3</sup><https://www.physionet.org/content/challenge-2017/1.0.0/>

### 3.1. Base de datos de señales ECG

N	S	V	F	Q
Normal	Atrial premature	Premature ventricular contraction	Fusion of ventricular and normal	Paced
Left bundle branch block	Aberrant atrial premature	Ventricular escape		Fusion of paced and normal
Right bundle branch block	Nodal (junctional) premature			
Atrial escape	Supra-ventricular premature			
Nodal (junctional) escape				Unclassifiable

Cuadro 3.1: Distribución de clases según la AAMI de las señales ECG pertenecientes a la base de datos *MIT-BIH Arrhythmia* [35].

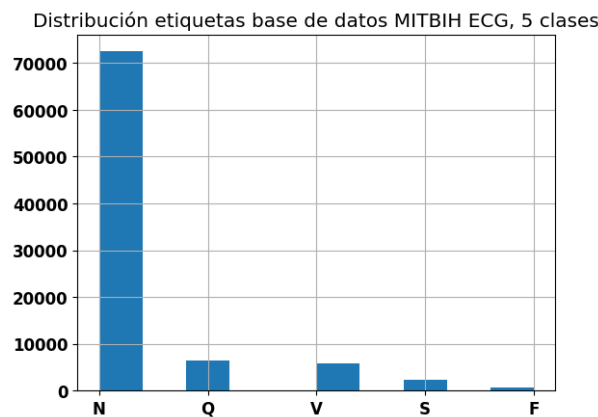


Figura 3.1: Distribución de clases de la base de datos de MITBIH ECG.

La base de datos *PTB Diagnostic ECG* consta de registros de ECG de 290 sujetos, con edades entre 17 y 87 años. Cada grabación contiene señales ECG de 12 derivaciones muestreadas a 10KHz, sin embargo, en este trabajo se implementa únicamente la información de la II derivación. La señales contenidas en esta base de datos se encuentran divididas en nueve categorías, una de control sano y ocho de enfermedades cardiacas diferentes. En este estudio se trabajó con 14552 muestras, pertenecientes a dos categorías: control sano y control anormal (infarto de miocardio).

En la figura 3.2 se observa gráficamente la distribución de los datos según su categoría.

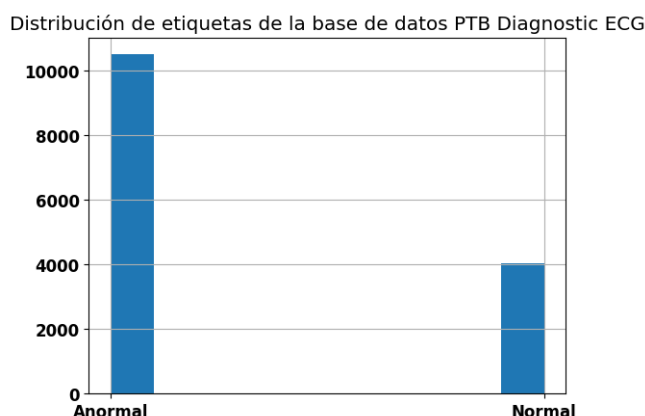


Figura 3.2: Distribución de etiquetas de la base de datos PTB Diagnostic ECG.

Por último, la base de datos de *PhysioNet* para el *Computing in Cardiology Challenge 2017*, consta de 8,528 grabaciones de señales ECG de una sola derivación con duraciones entre 9 y 60 segundos. Las grabaciones fueron obtenidas con el dispositivo *AliveCor*, muestreadas a 300Hz y sometidas a un proceso de filtrado. Las señales se dividen en cuatro categorías: ritmo sinusal normal, fibrilación auricular, ritmo alternativo y señal con ruido.

Con estos datos se realizaron dos aproximaciones, para la primera aproximación se utilizaron únicamente dos categorías: ritmo sinusal normal y fibrilación auricular. Mientras que para la segunda se emplearon las cuatro categorías disponibles. Cabe mencionar que estas señales pasaron por una etapa de preprocesamiento, que se detallara en la siguiente sección, con el objetivo de extraer latidos representativos de la señal ECG de cada grabación disponible.

En el cuadro 3.2 se detalla la distribución de la base de datos según su categoría.

Etiqueta	Categoría	No. de grabaciones	Porcentaje
N	ritmo sinusal normal	5076	60 %
A	fibrilación auricular	758	9 %
O	ritmo alternativo (otros)	2415	28 %
~	señal con ruído	279	3 %

Cuadro 3.2: Distribución de etiquetas de la base de datos de Physionet.

### 3.2. Preprocesamiento

Dado que la base de datos de *PhysioNet* para el *Computing in Cardiology Challenge 2017* cuenta con grabaciones de señales ECG con duraciones variables, fue necesario emplear una etapa de preprocesamiento, con el fin de preparar los datos antes de ser empleados para el entrenamiento del modelo de clasificación.

El objetivo principal en esta etapa consistió en segmentar latidos del corazón representativos de las señales ECG. Para ello, se llevaron a cabo los siguientes pasos:

1. El primer paso consistió en reducir el tamaño de todas las grabaciones de señales ECG a ventanas con una duración de 9 segundos, lo cual corresponde al

tamaño de la grabación con menor duración.

2. A continuación, se identificaron los picos R de las señales ECG. Para ello, se utilizó el *Heart Rate Analysis Toolkit* de Phyton. La librería `HeartPy` de Phyton, es una herramienta diseñada específicamente para el análisis de la frecuencia cardíaca en señales ECG regulares o muy ruidosas, que cuenta con funciones capaces de detectar los picos R presentes en dichas señales [47].
3. Como siguiente paso, para cada pico R encontrado se extrajo un segmento de señal de tamaño  $T$ . Con el fin de contar con al menos un pico R dentro del segmento extraído, el tamaño de dicho segmento se calculó de la siguiente manera:

$$T = \left\lfloor \frac{1}{2} (R_i + R_{i+1}) \right\rfloor - \left\lfloor \frac{1}{2} (R_{i-1} + R_i) \right\rfloor + 1$$

En donde,  $R_i$  representa la posición del pico R actual,  $R_{i-1}$  la posición del pico R anterior y  $R_{i+1}$  la posición del pico R posterior.

4. Seguidamente, se normalizaron las señales segmentadas para que los valores de amplitud se encontraran en un rango entre 0 y 1.
5. Por último, dado que la longitud  $T$  de cada segmento extraído era diferente, se definió un tamaño de señal  $D$  fijo. Si los segmentos contaban con un tamaño inferior a  $D$  se les realizó *zero padding* y en caso contrario, se redujo la señal a un tamaño  $D$ . Para los experimentos realizados en este trabajo  $D = 200$ .

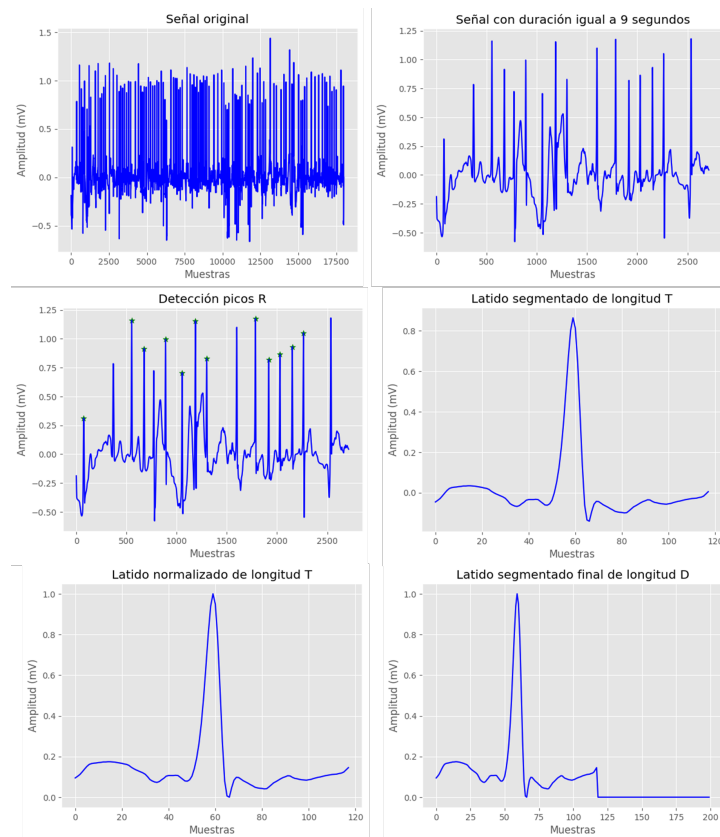


Figura 3.3: Pasos realizados en la etapa de preprocesamiento de señales ECG.

## Preprocesamiento de señales ECG

Después de esta etapa de procesamiento el número de muestras disponibles incrementó considerablemente, pasando de un total de 8,528 grabaciones de duraciones variables a un total de 67,471 latidos segmentados en ventanas de 200 muestras. En la figura 3.4 se muestra gráficamente la distribución de etiquetas de la base de datos de Physionet, después de aplicada la etapa de preprocesamiento.

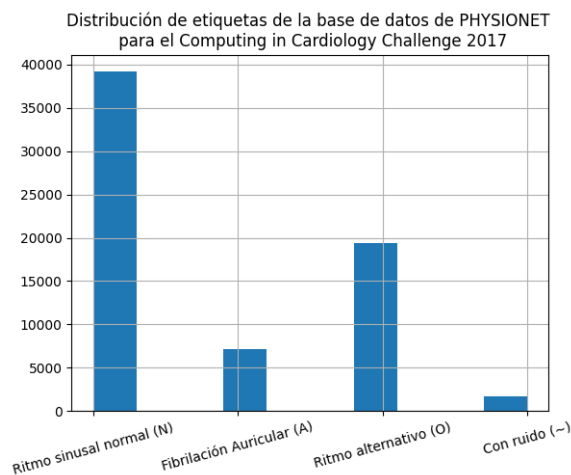


Figura 3.4: Distribución de etiquetas de la base de datos de Physionet para el *Computing in Cardiology Challenge 2017*, después de aplicar preprocesamiento.

Cabe mencionar que, las bases de datos *MIT-BIH Arrhythmia* y *PTB Diagnostic ECG* que se emplearon para este trabajo, ya contaban con un preprocesamiento anterior, por lo que se prescindió de esta etapa. Las señales contenidas en estas bases de datos fueron sometidas previamente a las siguientes técnicas con el fin de limpiar y preparar las señales para entrenamiento.

- Las señales fueron remuestreadas de 360Hz y 10KHz a 125Hz.
- Las grabaciones fueron modificadas dejando únicamente 10 segundos de duración.
- Se normalizaron las señales.
- Se identificaron los picos R en las señales.
- Se determinó la distancia entre cada pico R y se calculó la mediana T.
- Se segmentaron series de 187 muestras, tomando 1.2T puntos sucesivos iniciando en cada pico R. Si la longitud de la señal segmentada (1.2T) era menor a 187 se aplicó zero-padding.

Como se ha expuesto en el presente capítulo, las bases de datos *MIT-BIH Arrhythmia*, *PTB Diagnostic ECG* y la otorgada por Physionet para el *Computing in Cardiology Challenge 2017*, proporcionan una gran cantidad de registros de señales ECG etiquetadas, que son fundamentales para el desarrollo y evaluación de modelos de clasificación de arritmias cardíacas. No obstante, se ha observado que un tema recurrente en estos conjuntos de datos es el desbalanceo de clases, lo cual como se ha mencionado en capítulos anteriores puede llegar a representar un problema en el proceso de clasificación. En el siguiente capítulo, se llevarán a cabo una serie de pruebas para determinar si este desbalanceo de clases efectivamente representa un

### **3.2. Preprocesamiento**

---

desafío significativo en la tarea de clasificación. En base a los resultados obtenidos se podrá comprender mejor la naturaleza de este desafío y explorar en los capítulos subsiguientes estrategias adecuadas para abordarlo de manera efectiva.

## Capítulo 4

# Clasificación señales ECG

Basados en el estado del arte relacionado a clasificación de señales electrocardiográficas, se planteó experimentar con dos modelos de clasificación utilizando la biblioteca TensorFlow. El primer modelo consistió en un modelo de redes neuronales convolucionales (CNN), mientras que el segundo modelo cuenta con una arquitectura más compleja al emplear redes residuales (ResNet). A continuación, se describe a detalle cada modelo y los resultados de clasificación obtenidos.

### 4.1. Descripción y resultados del primer modelo de clasificación propuesto

Inicialmente se planteó el uso de un modelo de clasificación con redes convolucionales para llevar a cabo la tarea de clasificación. La arquitectura del modelo consta de una capa de entrada tipo convolucional 1D, seguida por una capa MaxPooling1D, con un tamaño de ventana igual a 5. Una capa Dropout con una tasa igual 0.1, que contribuye a prevenir el sobre ajuste. Luego una capa Flatten, seguida de dos capas densas con 64 y 32 neuronas respectivamente y con función de activación ReLu. Y por último una capa de densa de salida con función de activación softmax y número de neuronas igual al número de clases.

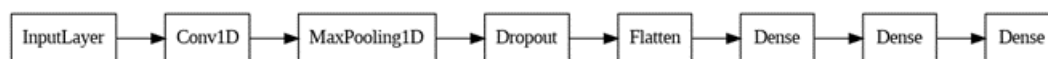


Figura 4.1: Arquitectura del primer modelo propuesto (CNN).

Como función de pérdida se utilizó *binary crossentropy* en tareas de clasificación binaria y *sparse categorical crossentropy* en tareas de clasificación multiclase. En el proceso de entrenamiento del modelo, se empleó un optimizador Adam (*Adaptive moment estimation*) con valor inicial de *learning rate* igual a 0.001. Para cada prueba el modelo se entrenó durante 50 *epochs* y los datos se dividieron en 80 % para entrenamiento, 20 % para validación y 20 % para prueba.

Los resultados de clasificación obtenidos utilizando este modelo para cada una de las bases de datos propuestas, se pueden observar en las figuras 4.2, 4.3 y 4.4.



#### 4.1. Descripción y resultados del primer modelo de clasificación propuesto

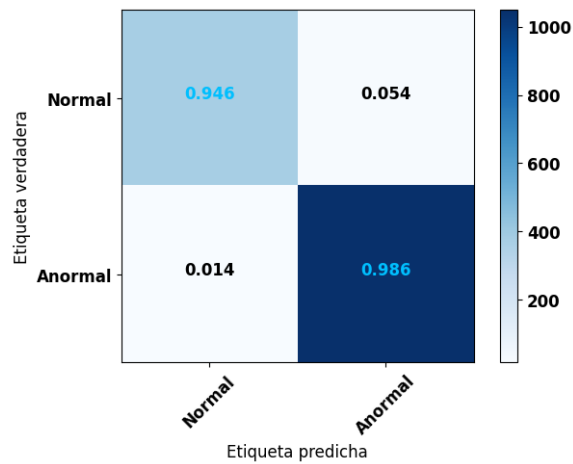


Figura 4.2: Matriz de confusión con los resultados de clasificación de señales para la base de datos PTB Diagnostic ECG.

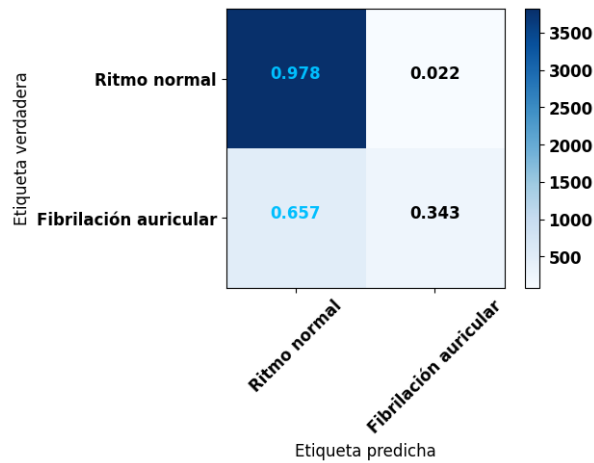


Figura 4.3: Matriz de confusión con los resultados de clasificación de señales para la base de datos de Physionet.

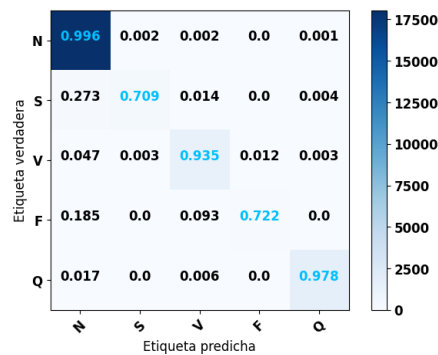


Figura 4.4: Matriz de confusión con los resultados de clasificación de señales para la base de datos MITBIH ECG.

## 4.2. Descripción y resultados del segundo modelo de clasificación propuesto

Con el objetivo de mejorar los resultados obtenidos durante la primera experimentación y partiendo de la propuesta realizada en [24]. En donde proponen el uso de redes residuales con el fin de mejorar la capacidad de aprendizaje de la red, ya que este tipo de arquitectura permite entrenar redes más profundas y capturar características más detalladas y discriminativas en comparación con las arquitecturas convolucionales tradicionales.

La arquitectura del segundo modelo propuesto consta de una capa convolucional 1D de entrada que recibe los latidos específicos extraídos de las señales ECG. Seguida por cinco capas residuales, cada una formada por dos capas convolucionales 1D, dos funciones de activación ReLu, una conexión residual entre las dos capas de convolución y por último una capa MaxPooling1D.

Después de las cinco capas residuales, se implementa una capa Flatten que transforma el tensor de entrada a una dimensión, dos capas totalmente conectadas con 32 y 16 neuronas respectivamente y con función de activación ReLu. Por último, la salida se procesa a través de una capa densa final con función de activación tipo sigmoide. Todas las capas convolucionales 1D mencionadas anteriormente, poseen 32 filtros con un tamaño de kernel igual a 5.

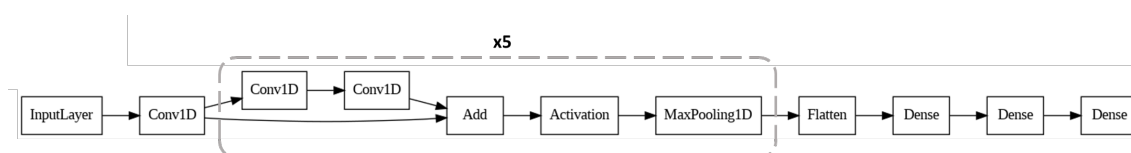


Figura 4.5: Arquitectura del segundo modelo propuesto (ResNet).

Similar a la experimentación anterior, en el proceso de entrenamiento del modelo, se empleó un optimizador Adam con valor inicial de *learning rate* igual a 0.001. Para cada prueba el modelo se entrenó durante 50 *epochs* y los datos se dividieron en 80 % para entrenamiento, 20 % para validación y 20 % para prueba.

Los resultados de clasificación obtenidos utilizando este modelo para cada una de las bases de datos propuestas, se pueden observar en las figuras 4.6, 4.7 y 4.8. Demostrando que el uso de redes residuales, contribuyó a la mejora del desempeño del modelo de clasificación.

## 4.2. Descripción y resultados del segundo modelo de clasificación propuesto

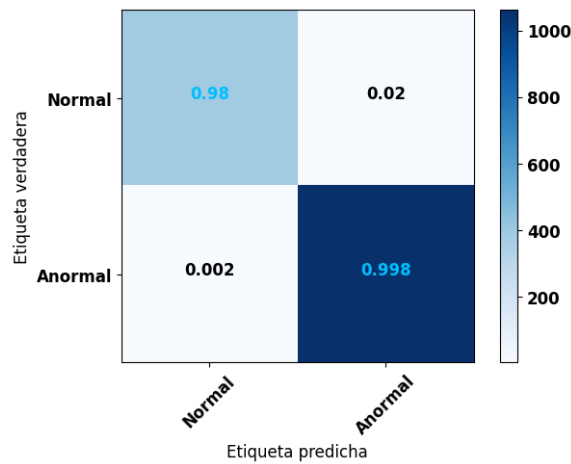


Figura 4.6: Matriz de confusión con los resultados de clasificación de señales para la base de datos PTB Diagnostic ECG.

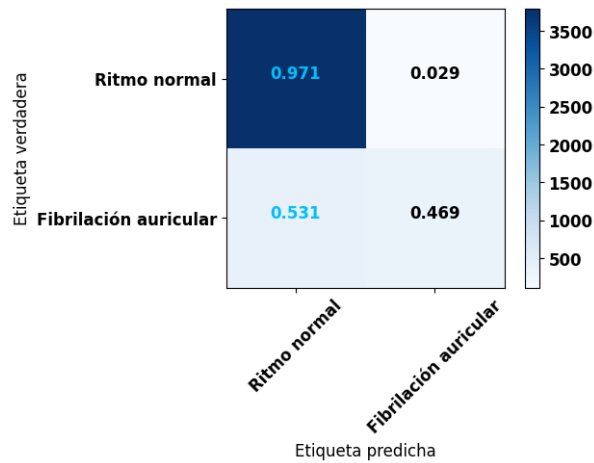


Figura 4.7: Matriz de confusión con los resultados de clasificación de señales para la base de datos de Physionet.

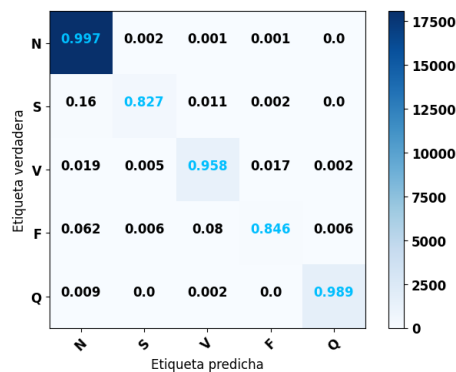


Figura 4.8: Matriz de confusión con los resultados de clasificación de señales para la base de datos MITBIH ECG.

Como se detalló en la sección 3.1 las bases de datos disponibles para este tipo de aplicaciones generalmente se encuentran desbalanceadas en términos de la distribu-

ción de clases. Este desequilibrio por lo regular ocasiona que los modelos aprendan a clasificar correctamente solo las clases mayoritarias, ignorando las clases con menor cantidad de información, obteniendo como resultado modelos de clasificación sesgados y poco efectivos.

Como es el caso de los experimentos realizados en esta sección, en donde los modelos se ven menos expuestos a las clases minoritarias, dado que se entrenaron con más información de las clases mayoritarias y una cantidad muy reducida de las clases minoritarias. Como se observa en los resultados de la experimentación 1 (figuras 4.2, 4.3 y 4.4) y en los resultados de la experimentación 2 (figuras 4.6, 4.7 y 4.8) en donde, a pesar que el uso de redes residuales haya contribuido a mejorar la capacidad de aprendizaje de la red, se puede observar que los modelos son menos precisos al clasificar las clases minoritaria en comparación a las clases mayoritarias. Sobre todo, en casos donde la diferencia entre la cantidad de datos disponibles entre clases es sumamente significativa, como ocurre con la base de datos de *Physionet* después de haber aplicado una etapa de preprocesamiento para extraer latidos representativos de cada grabación.

En conclusión, tomando como base a los resultados presentados en este capítulo, en donde se exploraron dos modelos de clasificación de señales electrocardiográficas, un modelo empleando redes neuronales convolucionales (CNN) y un modelo con arquitectura de redes residuales (ResNet). Se observó que el uso de redes residuales mejora el desempeño en comparación con las arquitecturas CNN tradicionales. Sin embargo, es evidente que trabajar con bases de datos desbalanceadas en cuanto a la distribución de clases, afecta la precisión en la clasificación de las clases minoritarias. Este sesgo en la clasificación representa un problema significativo en la detección temprana y precisa de enfermedades cardíacas menos comunes y resalta la importancia de considerar en los próximos capítulos la implementación de modelos generativos que sinteticen datos de las clases minoritarias para aumentar su participación en el conjunto de datos. Y de esta manera poder entrenar modelos de clasificación más equilibrados y eficientes.



## Capítulo 5

# Adquisición y análisis de señales ECG

En este capítulo se presenta una descripción del hardware utilizado en el proceso de adquisición y análisis de señales ECG. Además, se describe el algoritmo implementado para la lectura y clasificación de datos. Así como el modelo de clasificación utilizado y los pasos que se llevaron a cabo para embeber dicho modelo en el sistema, tomando siempre en consideración las limitaciones del dispositivo.

### 5.1. Descripción del hardware

#### 5.1.1. Cardio chip BMD101

El BMD101 es un SoC (system-on-chip) de alto rendimiento especializado en la detección de bioseñales, diseñado por NeuroSky, una empresa especializada en la producción de circuitos integrados y aplicaciones relacionadas con la monitorización cardíaca.

El BMD101 integra un amplificador de instrumentación de bajo ruido, que permite capturar señales débiles con mayor sensibilidad y precisión. Filtros analógicos integrados, que contribuyen a eliminar ruido e interferencias no deseadas. Y convertidores analógico-digitales ADC de 16 bits, que convierten señales ECG analógicas a digitales con alta resolución. Además, el dispositivo está diseñado para detectar bioseñales del cuerpo humano de forma automática lo que facilita el proceso de adquisición y análisis de señales ECG [31].

El BMD101 funciona de forma que mediante electrodos conectados a las entradas SEP y SEN, se reciben señales analógicas, que se transforman a señales digitales que posteriormente pueden ser transmitidas mediante las salidas TX y RX, utilizando protocolo de comunicación serial.

#### 5.1.2. M5Stick-C Plus

El M5Stick-C Plus es un dispositivo compacto, versátil, portátil y de fácil uso desarrollado por la empresa M5Stack, como una solución de desarrollo de IoT (Internet de las cosas). El dispositivo incorpora un microcontrolador ESP32 de doble núcleo

## 5.2. Algoritmo para lectura de datos

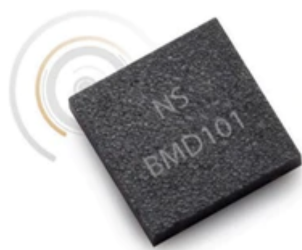


Figura 5.1: Cardio chip BMD101 [31].

que proporciona capacidades de procesamiento avanzadas. Cuenta con una pantalla a color TFT de 0.96 pulgadas, una batería de 120 mAh, conector de carga USB-C y capacidad de conectividad inalámbrica por medio de Wi-Fi y Bluetooth [32].

Además, cuenta con recursos de hardware extra que pueden ser útiles en distintas aplicaciones como, micrófono, sensores de luz infrarroja, buzzer, botones, acelerómetro, entre otros.



Figura 5.2: Dispositivo M5Stick-C Plus [32].

En este trabajo, se utilizaron ambos dispositivos (chip BMD101 y M5Stick-C Plus) para desarrollar un dispositivo capaz de adquirir señales electrocardiográficas. Para ello, el cardio chip BMD101 cuenta con electrodos conectados a las entradas, que se encargan de recibir la señal ECG analógica para luego convertirla a una señal digital que se envía por medio de protocolo serial al M5Stick-C Plus, para su posterior análisis. Una vez que la señal es recibida por el dispositivo M5Stick-C Plus, se procede a su representación gráfica en la pantalla.

## 5.2. Algoritmo para lectura de datos

El algoritmo para la lectura de datos de señales ECG se implementó utilizando la plataforma Arduino, dada su flexibilidad, simplicidad y compatibilidad con una amplia gama de microcontroladores.

Al desarrollar este algoritmo primero se realizó la importación de las librerías necesarias: `EloquentTinyML.h`, `eloquent_tinyml/tensorflow.h` y `M5StickCPlus.h`. Seguido de la declaración de entradas, la inicialización de variables y dispositivos y la inicializa-

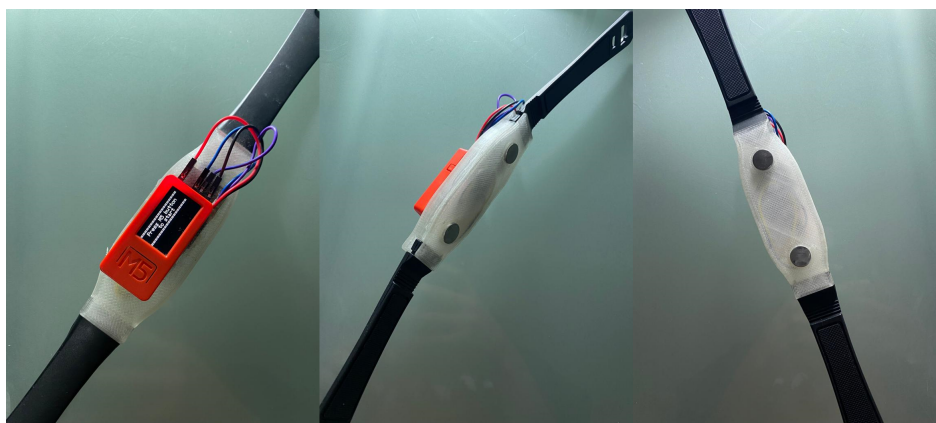


Figura 5.3: Dispositivo utilizado para adquirir y clasificar señales ECG.

ción de la comunicación serial, estableciendo un baudaje igual a 115200 baudios/s. Posteriormente, se definieron y crearon una serie de tareas empleando el sistema operativo FreeRTOS.

Más adelante, se indica en la pantalla al usuario que la adquisición de señales iniciará al presionar el botón A del dispositivo. Por lo que, se procede a realizar una evaluación constante del estado del botón A del M5Stick-C Plus. Al momento de detectar que se presionó y soltó este botón, se procede a realizar la lectura de datos del dispositivo BMD101 a través del puerto serial, cada vez que haya datos disponibles en el puerto. Los bytes recibidos se almacena en un búfer y a medida que son recibidos se verifica la integridad del paquete, si se detecta un paquete completo se pasa a la etapa de análisis de los datos.

En esta etapa, se evalúa la calidad de la señal, si la calidad es buena se procesan los bytes recibidos para obtener el valor bruto de amplitud de la señal. Este dato se representa visualmente en pantalla y se almacena constantemente en un búfer de tamaño definido, que se limpia constantemente al alcanzar su límite (187 muestras).

Una vez se ha llenado este búfer, es decir, se tiene una señal ECG de buena calidad con la cantidad de muestras establecida, se procede a la etapa de clasificación de dicha señal. Este proceso se explica detalladamente en la subsección 5.4.

### 5.3. Descripción del modelo de clasificación

El ESP32 incorporado en el dispositivo M5Stick-C Plus es un microcontrolador de recursos limitados en comparación a plataformas más potentes como un computador o servidor. El ESP32 tiene capacidad de almacenamiento reducida, memoria RAM relativamente pequeña y una capacidad de procesamiento limitada. Por lo que, es inasequible que pueda manejar modelos de clasificación complejos y pesados que requieren de una cantidad significativa de memoria para ejecutarse eficientemente.

Tiny Machine Learning (TinyML) es un campo de estudio emergente que se enfoca en desarrollar modelos de aprendizaje automático para dispositivos de baja potencia y recursos limitados, como el microcontrolador ESP32. La aplicación de esta técnica implica el uso de modelos ligeros y optimizados que se ajustan específicamente a las restricciones impuestas por el hardware. Por los que, generalmente estos modelos



### 5.3. Descripción del modelo de clasificación

son diseñados para ejecutar tareas de clasificación o predicción específicas con un porcentaje de precisión aceptable, pero sin exigir una cantidad elevada de recursos.

Dado estas restricciones, el modelo previamente explorado en la sección 4.2, no cumple con las condiciones necesarias para ser embebido en el microcontrolador. Ya que al ser muy complejo y pesado, puede conducir a un rendimiento deficiente, carencia de respuesta o un agotamiento de los recursos disponibles, provocando bloqueos o reinicios del sistema.

Por este motivo, se planteó un nuevo modelo más simple y mucho más ligero. Este nuevo modelo consistió en cuatro capas densas secuenciales, las primeras tres capas con función de activación tipo ReLU. La primera capa, cuenta con 15 neuronas y recibe como entrada un vector de longitud igual al tamaño de muestras en la señal (187), con los valores de amplitud de la señal ECG. La segunda capa densa cuenta con 10 neuronas y la tercera capa con 5 neuronas. En el caso de la última capa, la configuración depende si la tarea de clasificación es multiclase o binaria. En el caso de una clasificación binaria, se tiene una sola neurona y se emplea una función de activación tipo sigmoide. Mientras que, para tareas de clasificación multiclase, la última capa densa cuenta con 5 neuronas y función de activación softmax.

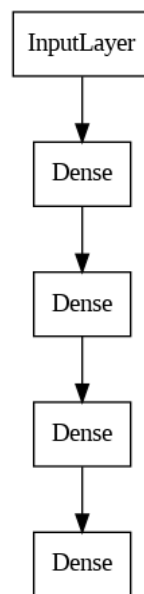


Figura 5.4: Arquitectura del modelo simplificado para implementar con TinyML.

Como función de pérdida se utilizó *binary crossentropy* en tareas de clasificación binaria y *sparse categorical crossentropy* en tareas de clasificación multiclase. En el proceso de entrenamiento del modelo, se empleó un optimizador Adam con valor de *learning rate* igual a 0.001. Para cada prueba el modelo se entrenó durante 125 *epochs*, dividiendo los datos en 80 % para entrenamiento, 20 % para validación y 20 % para prueba.

Es importante resaltar que aunque este presenta una precisión inferior en comparación con el modelo propuesto inicialmente, cuya arquitectura es más grande y compleja. Sin embargo, con este modelo nos aseguramos de cumplir con las restricciones de recursos en términos de capacidad de cálculo y memoria del hardware.

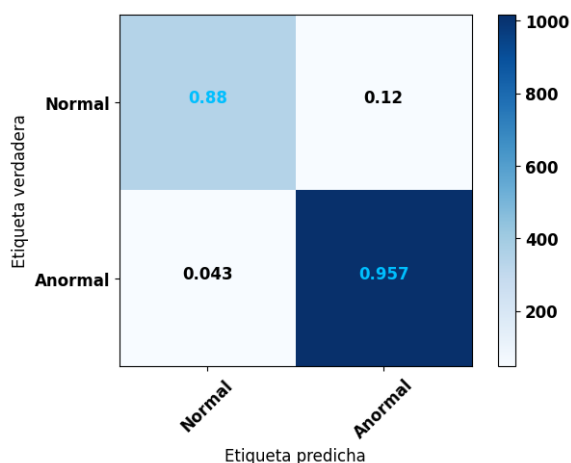


Figura 5.5: Matriz de confusión con los resultados de clasificación de señales, para la base de datos PTB Diagnostic ECG, utilizando un modelo compacto.

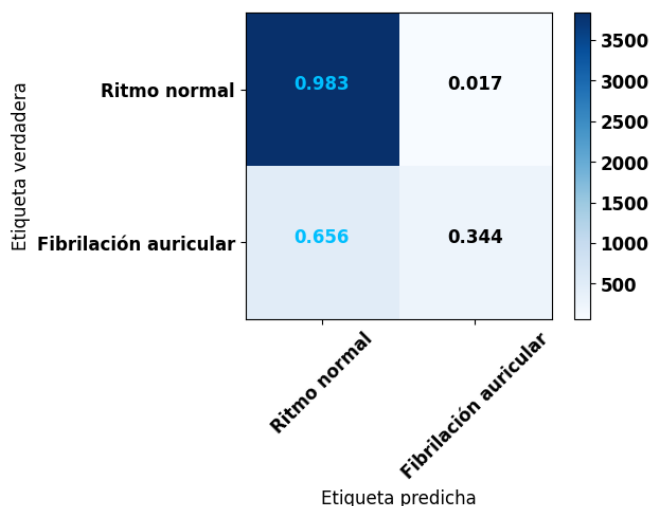


Figura 5.6: Matriz de confusión con los resultados de clasificación de señales, para la base de datos de Physionet, utilizando un modelo compacto.

### 5.4. Algoritmo para clasificación de datos.

Con el fin de lograr la clasificación de señales directamente desde el dispositivo, fue necesario embeber el modelo previamente entrenado en el microcontrolador ESP32. A continuación, se expone detalladamente el proceso empleado para la integración de dicho modelo en el microcontrolador y la consecuente clasificación realizada desde el dispositivo.

Inicialmente, fue imperativo exportar el modelo entrenado a un archivo con formato .h compatible con TinyML, con el objetivo de obtener una representación reducida del modelo que conserva solo la información esencial para que el dispositivo pueda realizar inferencias respecto a la clasificación. Para llevar a cabo este proceso se utilizó la librería `tinymolgen`, diseñada especialmente para generar código optimizado en formatos compatibles con microcontroladores, a partir de modelos entrenados con TensorFlow.

## 5.5. Resultados de clasificación en el dispositivo

Una vez se exportó el modelo a formato .h, este se anexó al código de Arduino. Para realizar el proceso de carga y ejecución del modelo previamente entrenado, se utilizó la librería `EloquentTinyML` desarrollada por la plataforma Arduino. Al implementar esta librería es necesario definir el número de entradas y salidas del modelo y el tamaño del área del tensor que se reserva para el modelo. Además, se debe crear una instancia de `TensorFlow` con los parámetros previamente definidos.

En este caso, el modelo entrenado en 5.5, tiene como entrada un vector con longitud igual a 187 y realiza una tarea de clasificación binaria, por lo que el número de entradas y salidas se definió como se muestra a continuación. En el caso del tamaño del área del sensor, este varía según el tamaño del modelo a implementar, por lo que el valor adecuado se encontró de forma iterativa y experimental.

```
1 #define N_INPUTS 187
2 #define N_OUTPUTS 2
3 #define TENSOR_ARENA_SIZE 5 * 1024
4
5 Eloquent::TinyML::TensorFlow::TensorFlow<N_INPUTS, N_OUTPUTS, TENSOR_ARENA_SIZE> tf;
6
7 float input[1][187];
8 float output_data = 0;
```

Posteriormente, se inicializó el modelo y se procedió a crear una tarea dedicada específicamente a realizar el proceso de clasificación. Esta tarea se activa únicamente después de adquirir una señal ECG de buena calidad del tamaño especificado (187 muestras). Su principal objetivo es efectuar una transformación de los datos almacenados en el búfer que contiene la señal ECG, a fin de adecuarlos al formato de entrada requerido por el modelo, para luego realizar la clasificación de la señal y en función al resultado obtenido mostrar en pantalla si el latido cardiaco es considerado normal o anormal. Este proceso de adquisición y clasificación se repite constantemente mientras se detecte, mediante los electrodos conectados al dispositivo, una señal con calidad suficiente para ser medida por el cardio chip BMD101.

Cabe recalcar la importancia de simplificar la arquitectura del modelo entrenado. Ya que si al exportar el modelo a formato .h se genera un archivo demasiado pesado, al momento de embeberlo en el microcontrolador se puede superar la capacidad de almacenamiento y la memoria disponible del dispositivo, generando que la carga y ejecución del modelo sean ineficientes o incluso imposibles de efectuar.

## 5.5. Resultados de clasificación en el dispositivo

Con el fin de evaluar el desempeño de la arquitectura del modelo propuesto en 5.4, una vez embebido en el dispositivo, se llevaron a cabo dos pruebas. La primera prueba se realizó evaluando el modelo entrenado previamente con las señales contenidas en la base de datos *PTB Diagnostic ECG*, resultados con datos de prueba en la figura 5.5. Mientras que, para la segunda prueba se entrenó el modelo con las señales de la base de datos de *Physionet*, resultados con datos de prueba en la figura 5.6.

El comportamiento observado al adquirir y clasificar señales directamente con el dispositivo, coincidió con el que se espera al observar los resultados de las matrices de confusión de las figuras 5.5 y 5.6. En la figura 5.5 se pudo observar que el modelo presenta menor precisión al momento de clasificar latidos cardíacos normales. Mien-

tras que, en 5.6 sucede el caso contrario, en donde la precisión disminuye al clasificar latidos cardíacos anormales.

La razón de este comportamiento, como se mencionó en secciones anteriores, radica en la falta de representación equitativa de ambas clases en el conjunto de datos de entrenamiento, como se observa en la figura 3.2 y en el cuadro 3.2. Dicha desigualdad en la distribución de clases puede afectar negativamente el rendimiento del modelo, ya que su capacidad para aprender y generalizar de manera efectiva puede estar sesgada hacia la clase mayoritaria.

Este comportamiento se reflejó claramente en los experimentos realizados. En el primer experimento, se observó un notable número de falsos positivos en la clasificación de las señales por parte del dispositivo para la clase anormal. En contraste, en el segundo experimento, se encontró una significativa cantidad de falsos positivos en la clasificación para la clase normal.

Los resultados observados reiteran la importancia de disponer de bases de datos equilibradas al llevar a cabo el entrenamiento y validación del modelo de clasificación, particularmente, en aplicaciones de naturaleza médica, en donde un modelo de clasificación sesgado podría tener consecuencias significativas en la detección precisa de latidos cardíacos normales o latidos cardíacos que presentan algún tipo de arritmia.

Para abordar el problema de desequilibrio de clases, se propone el uso de redes generativas adversarias (GAN) para generar muestras sintéticas de señales ECG pertenecientes a las clases minoritarias. De esta manera, se plantea equilibrar la distribución de clases, proporcionar al modelo una mayor cantidad de datos de entrenamiento y así mismo mejorar su rendimiento, ya que será más preciso al clasificar todas las clases por igual.



## Capítulo 6

# Generación de señales ECG

Como se evidenció en las secciones anteriores el problema principal de entrenar modelos de clasificación con bases de datos desbalanceadas en términos de la distribución de clases, radica en que se generan como resultado modelos de clasificación sesgados y poco efectivos.

Dado que las redes GAN han demostrado ser una herramienta valiosa en el campo del aprendizaje automático, especialmente en la generación de contenido sintético realista. Se propuso como solución a este problema, el desarrollo de un modelo basado en redes GAN que genere señales ECG sintéticas pertenecientes a las clases minoritarias, con el fin de lograr un equilibrio en la distribución de clases y mejorar el desempeño del modelo de clasificación.

### 6.1. Desarrollo red GAN

El modelo de la red GAN propuesto sigue la estructura de una GAN regular, formada por dos componentes principales: el generador y el discriminador. A continuación se da una descripción de las arquitecturas empleadas para cada uno de los componentes de la red.

#### 6.1.1. Descripción del modelo generador

El modelo de generador consta de dos capas LSTM bidireccionales (BiLSTM), que permiten que el modelo capture dependencias en ambas direcciones en los datos de entrada, ambas capas BiLSTM constan de 64 unidades ocultas. A continuación, se cuenta con una capa Lambda, que se encarga de sumar las salidas del estado oculto final de las capas LSTM con el objetivo de combinar la información contextual de ambas direcciones. Y por último, se tiene una capa densa con función de activación lineal y una dimensión de salida igual al número de muestras de salida establecido por el usuario, en este caso igual a 187.

Dado que el objetivo del generador es producir latidos cardíacos sintéticos que sean morfológicamente similares a las señales con que se alimenta la red, se aplicó la función de pérdida *mean absolute error*, cuyo objetivo es minimizar la diferencia absoluta promedio entre las predicciones del modelo y los valores reales del conjunto de datos.

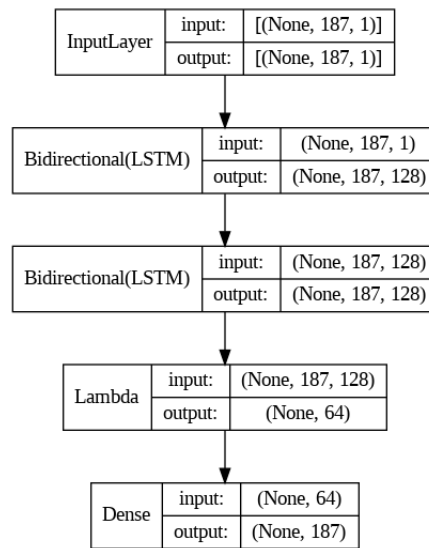


Figura 6.1: Arquitectura modelo generador.

### 6.1.2. Descripción del modelo discriminador

El modelo del discriminador consiste en una red neuronal (CNN) diseñada con el objetivo de clasificar y distinguir entre muestras de señales generadas y reales. El modelo consta de dos capas es 1D, con 32 y 64 filtros respectivamente, cada una seguida por una función de activación LeakyReLU. Seguido de este par de capas, se cuenta con una capa MaxPooling1D. Esta misma arquitectura se repite, agregando otro par de capas es 1D, con 128 y 256 filtros respectivamente y funciones de activación LeakyReLU. Por último, se tiene una capa densa totalmente conectada con una neurona y función de activación sigmoide.

En el caso del discriminador, se empleó la función de pérdida *binary\_crossentropy* al tratarse de un problema de clasificación binaria, en donde el objetivo consistía en predecir si las señales de entrada pertenecían al conjunto de datos reales o al conjunto de datos generados sintéticamente.

Para cada experimento realizado la red GAN se entrenó durante 10,000 *epochs*. En cada *epoch* la red se entrenó en *batches* de 256 grabaciones, extraídas del conjunto completo de datos de entrenamiento. Se utilizó un optimizador Adam, por su eficiencia computacional, con un valor de *learning rate* igual a 0.001.

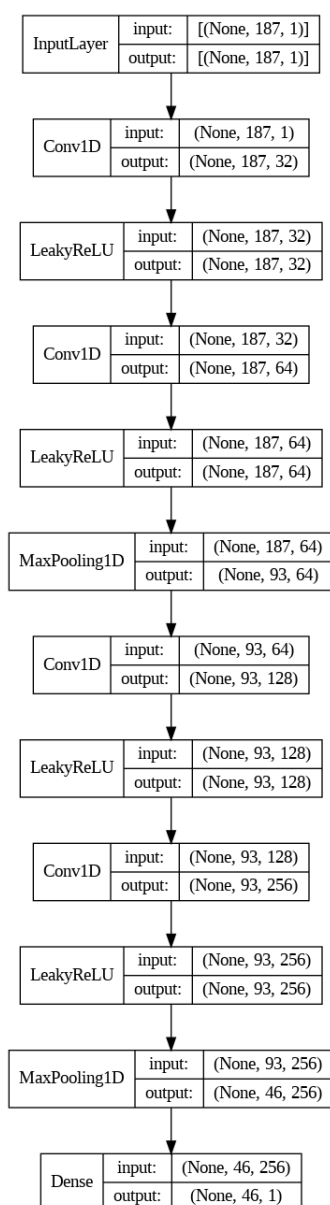


Figura 6.2: Arquitectura modelo discriminador.

## 6.2. Resultados primera experimentación

Como experimento inicial, se trabajó con el conjunto de datos *PTB Diagnostic*, esta base de datos cuenta con señales clasificadas en dos categorías: control sano, que es equivalente a ritmo sinusal normal y ritmo anormal. Con una distribución desigual entre categorías, siendo ritmo anormal la clase con mayor cantidad de muestras. Para este experimento se entrenaron dos redes GAN, una para generar señales ECG representativas de la clase minoritaria (ritmo sinusal normal) y otra para generar señales ECG correspondientes a la clase anormal.

En la figura 6.3 se observan los resultados obtenidos después de entrenar una red GAN durante 10,000 *epochs*, utilizando como datos de entrada las señales clasificadas en la categoría “ritmo sinusal normal” de la base de datos *PTB Diagnostic*.



## 6.2. Resultados primera experimentación

En esta figura se observa claramente que las muestras generadas presentan gran similitud morfológica con la distribución de las muestras originales.

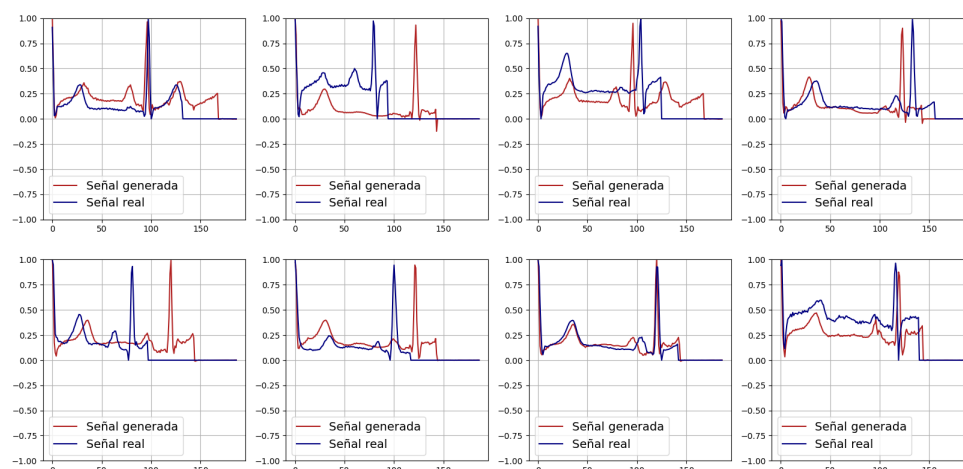


Figura 6.3: Ejemplo de los resultados obtenidos del entrenamiento de la red GAN, para generar latidos representativos de señales cardíacas normales, tomando como datos de entrada señales de la base de datos *PTB Diagnostic*.

Aunque visualmente las muestras generadas por la red GAN presentan resultados satisfactorios, se consideró necesario realizar una evaluación cuantitativa para obtener una medida objetiva de la similitud entre los datos generados sintéticamente y los datos reales. Si bien es cierto que existe una falta de consenso sobre la forma adecuada de evaluar los datos generados utilizando redes GAN, para este trabajo se ha decidido emplear las métricas DTW (Dynamic Time Warping) y RMSE (Root Mean Square Error).

A pesar de que estas métricas no son las únicas opciones disponibles, se decidió trabajar con ellas debido su amplia aceptación y utilización en diversos campos para evaluar la similitud entre señales o series temporales. Particularmente, la métrica DWT es útil al tratar con variaciones en la escala temporal, desplazamientos en el tiempo o deformaciones en los datos. Mientras que, la métrica RMSE se centra principalmente en cuantificar la discrepancia entre los valores reales y los valores generados.

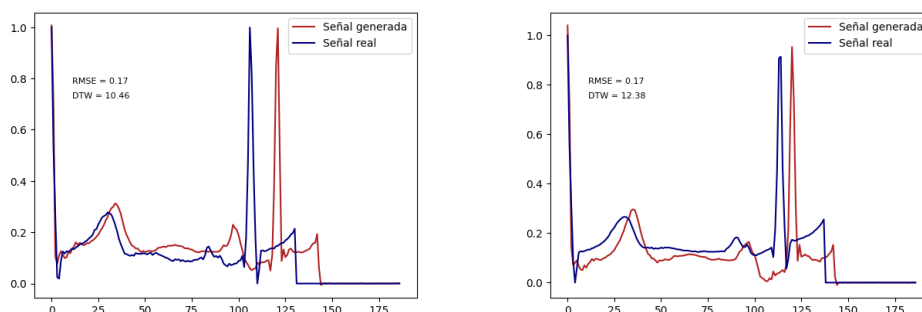


Figura 6.4: Ejemplos de señales ECG pertenecientes a la clase control sano (normal) y su correspondiente valor RSME y DWT.

En las figuras 6.4 y 6.5 se presentan ejemplos de las métricas DTW y RMSE obtenidas al evaluar las señales generadas sintéticamente con la arquitectura de red GAN

## Generación de señales ECG

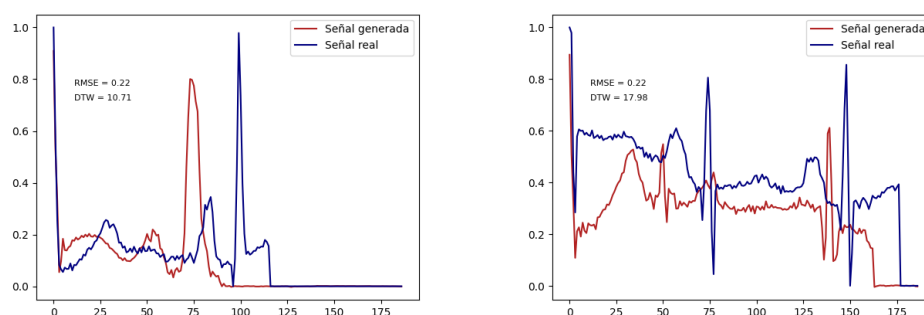


Figura 6.5: Ejemplos de señales ECG pertenecientes a la clase control anormal y su correspondiente valor RSME y DWT.

planteada (señales rojas), con las señales reales pertenecientes a la base de datos (señales azules). Los resultados obtenidos para la métrica DTW indican una gran similitud en la forma y el patrón temporal de las señales, mientras que, los resultados de RMSE obtenidos reflejan una buena correspondencia entre las magnitudes de ambas señales. Estos resultados contribuyen a confirmar que las señales generadas por la GAN logran capturar las características fundamentales de las señales ECG reales, tanto en términos de su forma como de sus valores numéricos.

Es importante destacar la relevancia de la cantidad de *epochs* empleados para el entrenamiento de la red GAN. Ya que teniendo presente que las señales contenidas en las bases de datos cuentan con un procesamiento previo en que se aplicó la técnica zero-padding. Y sabiendo que el generador aprende a generar muestras sintéticas tratando de replicar las características de las señales reales presentes en los datos de entrenamiento. El uso de la técnica zero-padding puede afectar negativamente el rendimiento de la red y dificultar la capacidad de la misma de generar señales ECG de calidad, si la red se entrena durante una cantidad de *epochs* reducida. Como es el caso de los resultados de la figura 6.6, en donde después de 1,500 *epochs* de entrenamiento se obtienen como resultado líneas rectas de amplitud cero.

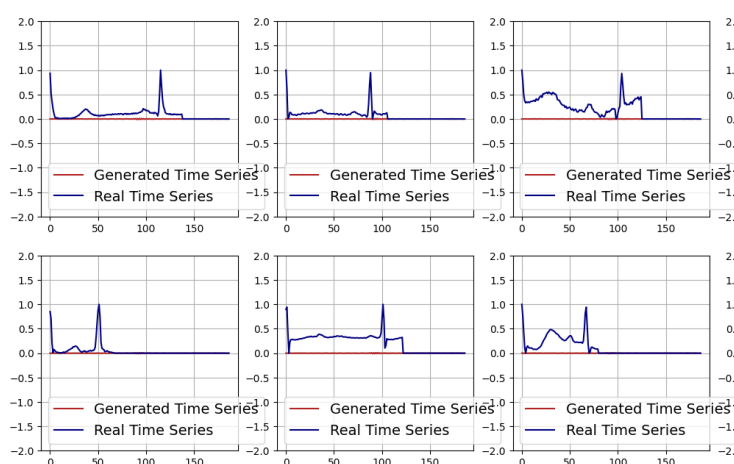


Figura 6.6: Ejemplo de los resultados obtenidos del entrenamiento de la red GAN durante 1,500 *epochs*, para generar latidos representativos de señales cardíacas normales, tomando como datos de entrada señales de la base de datos *MIT-BIH Arrhythmia*.

Aunque no se puede garantizar que el comportamiento observado en la figura 6.6,

se reproducirá de manera consistente en todo tipo de señales ECG a las cuales se les aplicó previamente la técnica zero-padding. Se decidió abordar este desafío antes de proceder con el siguiente experimento, realizando modificaciones a la etapa de preprocesamiento descrita previamente en la sección 3.2. Estas modificaciones se aplicaron exclusivamente al conjunto de datos proporcionados para el desafío *Computing in Cardiology Challenge 2017* de *PhysioNet*, esta decisión se basó en el hecho de que estos datos no habían sido sometidos a ningún tipo de preprocesamiento previo, a diferencia de los conjuntos de datos de *MIT-BIH* y *PTB Diagnostic*. En la siguiente sección, se detallan los pasos a realizar en esta nueva etapa de preprocesamiento.

### 6.3. Preprocesamiento de datos

Similar a lo establecido en la sección 3.2, el propósito de esta etapa se centró en la adecuación de los datos de entrada antes de su utilización en el entrenamiento de la red GAN. En este caso, con la restricción de no aplicar la técnica de zero-padding con el objetivo de igualar las longitudes de las señales. Para cumplir con dicho objetivo, se llevaron a cabo los siguientes pasos:

1. El primer paso consistió en reducir el tamaño de todas las grabaciones de señales ECG a ventanas de 9 segundos.
2. Seguidamente, se identificaron los picos R de las señales ECG utilizando el *Heart Rate Analysis Toolkit* de Phyton.
3. A continuación, por cada pico R encontrado, se extrajo un segmento de señal de longitud D, buscando dejar el pico R como punto central. Para este trabajo D = 200.
4. Por último, se normalizaron las señales segmentadas, para que sus valores de amplitud se encontraran en un rango entre 0 y 1.

Tras este proceso de procesamiento, se obtuvo un incremento significativo en el número de muestras disponibles, alcanzando un total de 67,471 latidos segmentados en ventanas de 200 muestras. En la tabla 6.1 se da una visión más detallada de la distribución de etiquetas en la base de datos de Physionet después de aplicar esta etapa de preprocesamiento.

Etiqueta	Categoría	No. de muestras	Porcentaje
N	ritmo sinusal normal	39187	58%
A	fibrilación auricular	7118	10.5%
O	ritmo alternativo (otros)	19441	29%
~	señal con ruído	1723	2.5%

Cuadro 6.1: Distribución de clases después de aplicar preprocesamiento para la base de datos de Physionet, 4 clases.

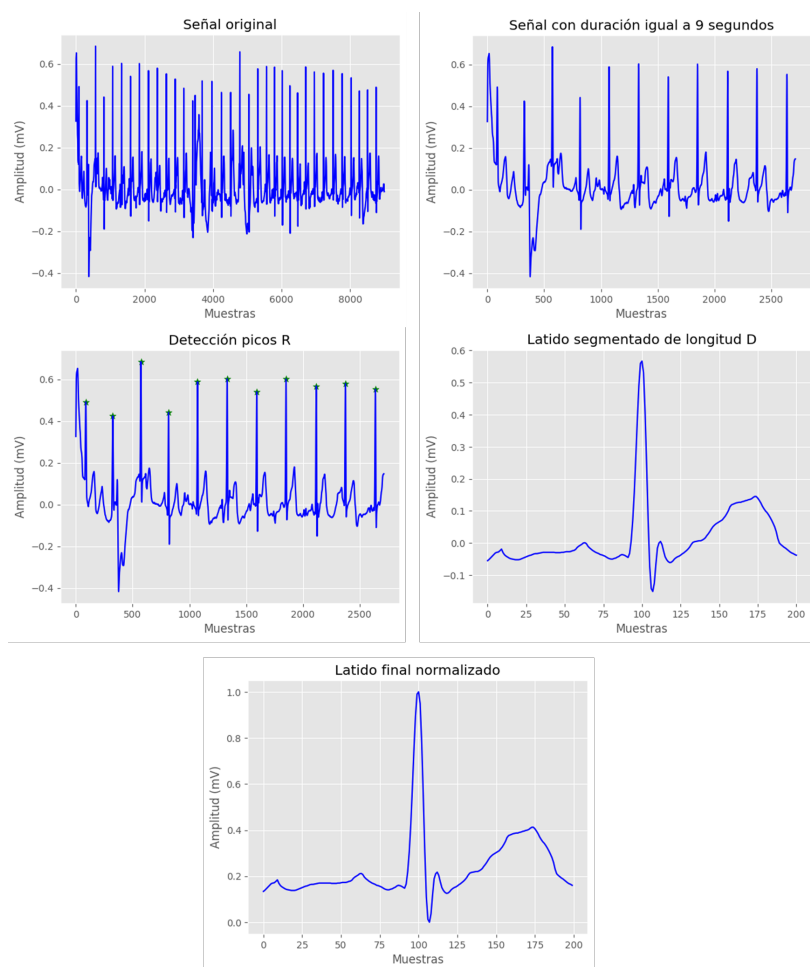


Figura 6.7: Pasos realizados en la etapa de preprocesamiento de señales.

### 6.4. Resultados segunda experimentación.

Para este experimento se empleó el conjunto de señales de la base de datos de Physionet para el *Computing in Cardiology Challenge 2017*. Este conjunto de señales ECG se encuentra dividido en cuatro categorías: ritmo sinusal normal, fibrilación auricular, ritmos alternativos y señales con ruido. Dado que la distribución de estas categorías es desigual, siendo el ritmo sinusal normal la clase mayoritaria, representando el 58% de los datos disponibles. Se planteó como objetivo emplear la red GAN planteada para generar señales ECG sintéticas correspondientes a cada una de las categorías, con el propósito de construir un conjunto de datos balanceado, en el que todas las categorías estén representadas de forma equitativa.

En la figura 6.8 se observan algunos ejemplos de señales generadas sintéticamente con la red GAN, después de ser entrenada durante 8,000 epochs. En la figura se observan los resultados para las cuatro clases contenidas en la base de datos.

## 6.4. Resultados segunda experimentación.

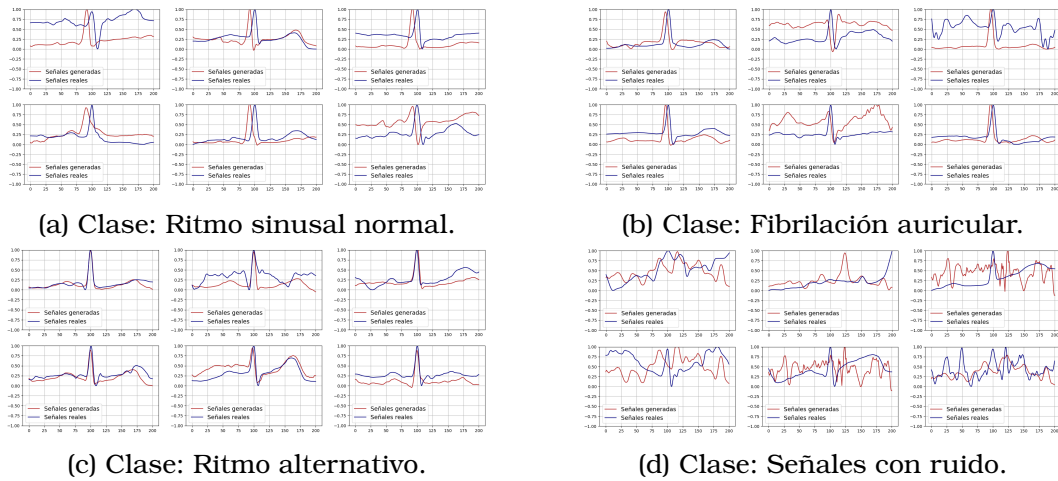


Figura 6.8: Ejemplos de señales ECG reales vs generadas, para cada clase contenida en la base de datos de Physionet.

Similar al experimento realizado previamente, de emplearon las métricas DTW y RMSE para evaluar y cuantificar la similitud entre las señales generadas y las señales reales. Los resultados de esta evaluación se presentan en la figura 6.9, en donde se demuestra tanto a nivel visual como cuantitativo la presencia de patrones y características morfológicas similares a las señales reales en las señales generadas con la red GAN.

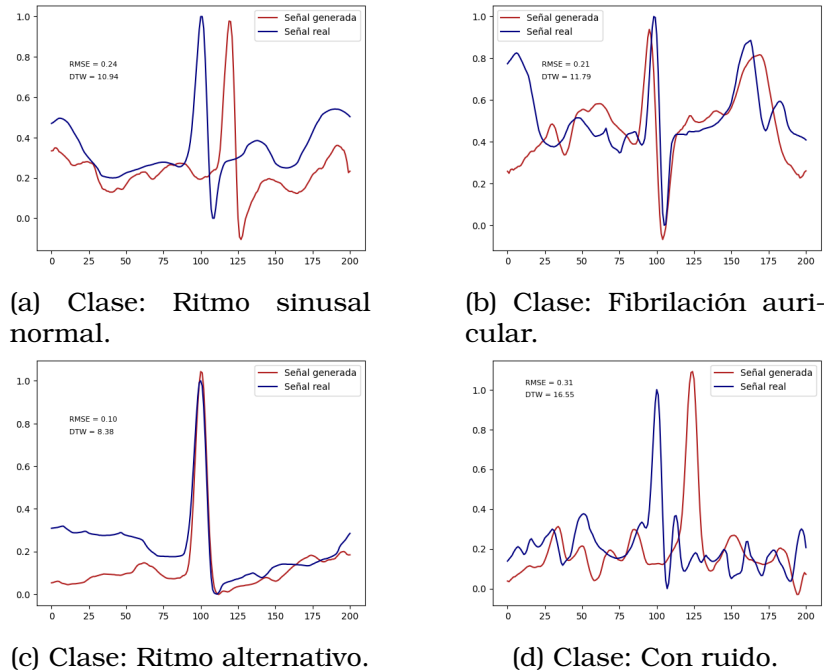


Figura 6.9: Ejemplos de señales ECG y su correspondiente valor RSME y DWT.

Es relevante destacar que, incluso después de aplicar el proceso de preprocesamiento, los segmentos de señales extraídos de la base de datos presentan una notable diversidad entre sí e incluso se encuentran señales que exhiben ciertas irregularidades, a diferencia de las señales contenidas en la base de datos de PTB Diagnostic.

## **Generación de señales ECG**

---

Estas variaciones intrínsecas entre las señales reales de una misma clase se evidencian claramente en la figura 6.8, y pueden representar un desafío al momento de entrenar la red GAN. Por lo tanto, como recomendación para trabajos futuros, se sugiere llevar a cabo un segundo proceso de filtrado, en donde se seleccionen visualmente aquellas señales que presenten los patrones y características morfológicas definidas por expertos para cada clase y se entrene únicamente con esta información el modelo de la red.

En base a los resultados observados en este capítulo, podemos concluir que la arquitectura de red GAN planteada es capaz de generar señales sintéticas que capturan de manera precisa las características morfológicas de las señales ECG reales utilizadas para el entrenamiento del modelo. Lo que demuestra que el uso de redes GAN es una solución efectiva para abordar el desafío del desequilibrio de clases en las bases de datos de señales ECG y mejorar el desempeño de los modelos de clasificación y detección de enfermedades cardíacas.



## Capítulo 7

# Resultados de clasificación con señales ECG generadas sintéticamente

En este capítulo se presentan los resultados de la clasificación de señales, utilizando conjuntos de datos balanceados que combinan señales ECG reales y generadas sintéticamente con la red GAN propuesta en el capítulo anterior. Con el fin de comparar el rendimiento entre los modelos entrenados con estos nuevos datos y los modelos entrenados con bases de datos desequilibradas. Los resultados mostrados se dividirán en dos partes, en la primera se mostraran los resultados obtenidos al trabajar con la base de datos *PTB Diagnostic*, realizando una clasificación entre señales ECG normales y anormales. Mientras que, en la segunda parte se utilizó la base de datos otorgada por Physionet para el *Computing in Cardiology Challenge 2017*.

### 7.1. Base de datos PTB Diagnostic

Como se describió en la sección 3.1, la base de datos *PTB Diagnostic* utilizada en este trabajo cuenta con 14,552 muestras, divididas en dos categorías: control sano y control anormal. Con un desequilibrio que favorece a la clase control anormal, que abarca el 72 % de los datos que forman el conjunto.

Después de comprobar que la arquitectura de red GAN propuesta logra generar señales ECG que capturan los patrones y característica morfológicas de las señales utilizadas en el entrenamiento de la red. Como se observó para el caso de este conjunto de datos específico en las figuras 6.4 y 6.5. Se utilizaron estos modelos previamente entrenados para construir un conjunto de datos balanceado que combina tanto señales ECG reales como sintéticas y se evaluó de nuevo el rendimiento de los modelos de clasificación propuestos.

El nuevo conjunto de datos balanceados utilizado para el entrenamiento del clasificador consistió en 21,012 muestras, con una distribución equitativa del 50 % para cada clase. La distribución específica entre muestras reales y sintéticas se detalla en el cuadro 7.1. Se mantuvieron los mismos parámetros de entrenamiento utilizados en los capítulos 4 y 5. En los que, respectivamente, se empleó un modelo ResNet entrenado durante 50 *epochs* y un modelo de arquitectura más compacta entrenado



## 7.1. Base de datos PTB Diagnostic

durante 125 *epochs*. Además, el conjunto de datos de prueba utilizado para evaluar el rendimiento fue el mismo conjunto que se empleó para evaluar los modelos entrenados con datos desbalanceados.

Clase	No. de muestras reales	No. muestras sintéticas	Total
Normal	4,046	6,460	10,506
Anormal	8,506	2,000	10,506

Cuadro 7.1: Distribución conjunto de datos balanceado, base de datos PTB Diagnostic.

En las figuras 7.1 y 7.2, se presentan los resultados obtenidos al emplear un modelo ResNet y un modelo compacto, respectivamente. Se puede observar que, aunque ambos modelos logran altos porcentajes de acierto en la clasificación tanto con el conjunto de datos desbalanceado como con el balanceado, existe una mejora en el rendimiento al emplear el conjunto de datos balanceado con muestras sintéticas. Sobre todo, en el porcentaje de muestras clasificadas como verdaderos positivos de la clase minoritaria, ritmo normal, en donde se observa una mejora del 1 % en el modelo ResNet y de 4.1 % en el modelo compacto. Este comportamiento coincide con lo que se esperaba obtener al balancear el conjunto de datos utilizando datos generados sintéticamente.

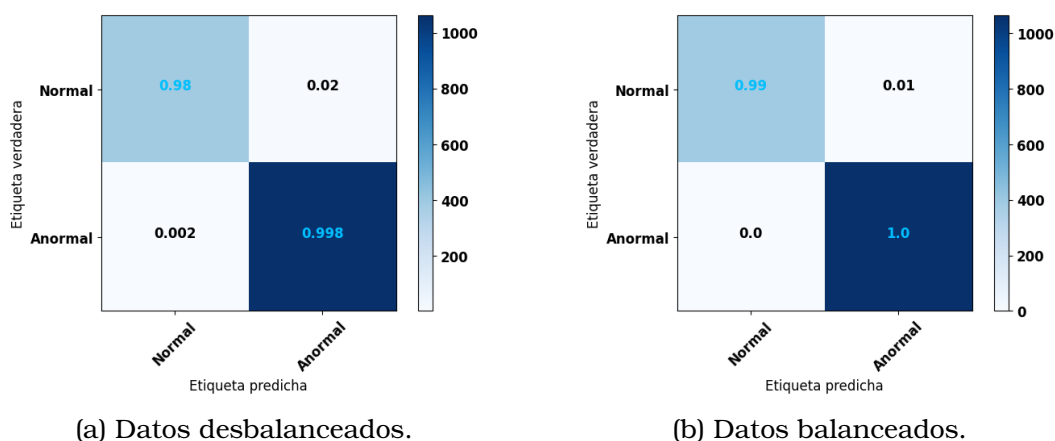


Figura 7.1: Resultados de clasificación de señales ECG, para la base de datos PTB Diagnostic ECG, utilizando un modelo ResNet.

## Resultados de clasificación con señales ECG generadas sintéticamente

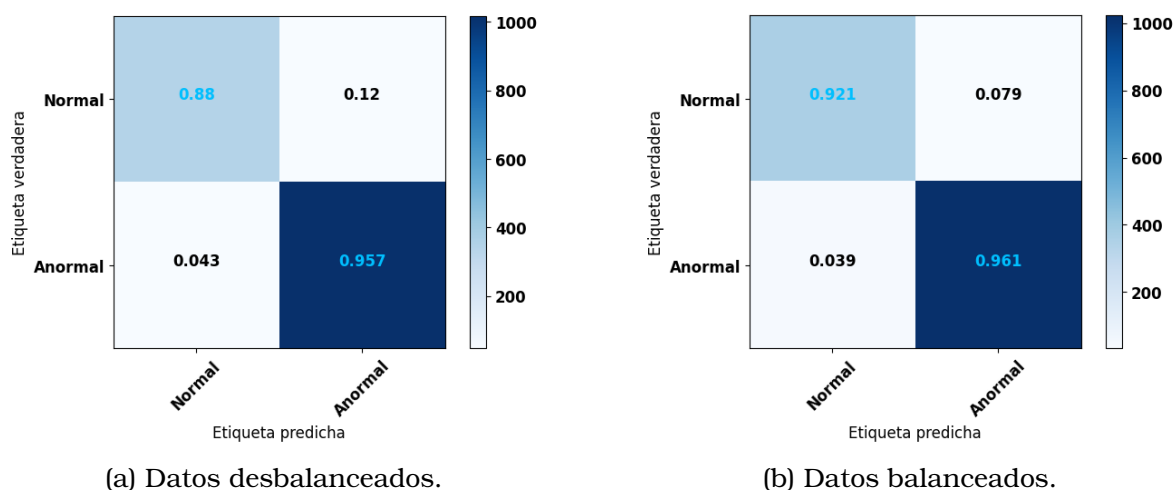


Figura 7.2: Resultados de clasificación de señales ECG, para la base de datos PTB Diagnostic ECG, utilizando un modelo más compacto.

## 7.2. Base de datos del Computing in Cardiology Challenge 2017

En esta segunda experimentación se trabajó con la base de datos brindada por Physionet para el *Computing in Cardiology Challenge 2017* que cuenta con información de 4 tipos diferentes de señales ECG. Después de aplicar la etapa de preprocesamiento indicada en 6.3 a este conjunto de datos, se obtiene un total de 67,471 muestras disponibles, desequilibradas de tal forma que el 58% de los datos corresponde a señales de ritmo sinusal normal y el 42% restante se divide entre la tres clases sobrantes.

Siguiendo la línea del experimento anterior, se utilizó el conjunto de modelos de red GAN previamente entrenados y cuyos resultados se presentan en la sección 6.4, para construir un conjunto de datos en el que las 4 clases se encontraran representadas de forma equitativa. Este nuevo conjunto de datos se compone de 30,000 muestras por clase, con la distribución entre muestras reales y sintéticas detallada en el cuadro 7.2.

Clase	Muestras reales	Muestras sintéticas	Total
N	20,000	10,000	30,000
A	7,118	22,882	30,000
O	19,000	11,000	30,000
~	1,723	28,277	30,000

Cuadro 7.2: Distribución conjunto de datos balanceado, base de datos del Cardiology Challenge 2017.

Cabe mencionar que para este conjunto de datos se realizaron dos diferentes experimentos. Para el primer experimento se tomaron en consideración únicamente las clases ritmo sinusal normal y fibrilación auricular. Mientras que para el segundo experimento se consideraron las cuatro clases disponibles.

Como parte del primer experimento se trabajó con los modelos propuestos en los capítulos 4 y 5. Para el modelo ResNet propuesto en la sección 4.2, se realizó un entre-

## 7.2. Base de datos del Computing in Cardiology Challenge 2017

namiento durante 50 *epochs*. Mientras que el modelo más compacto propuesto en la sección 5.3 se entrenó durante 125 *epochs*. Para ambos modelos se mantuvieron los mismo parámetros de entrenamiento tanto para el conjunto de datos desequilibrado como para el conjunto de datos equilibrado. Además, se utilizó el mismo conjunto de datos de prueba, el cual consiste exclusivamente en muestras reales, para evaluar el desempeño de los modelos. Los resultados obtenidos de esta evaluación se presentan en las figuras 7.3 y 7.4.

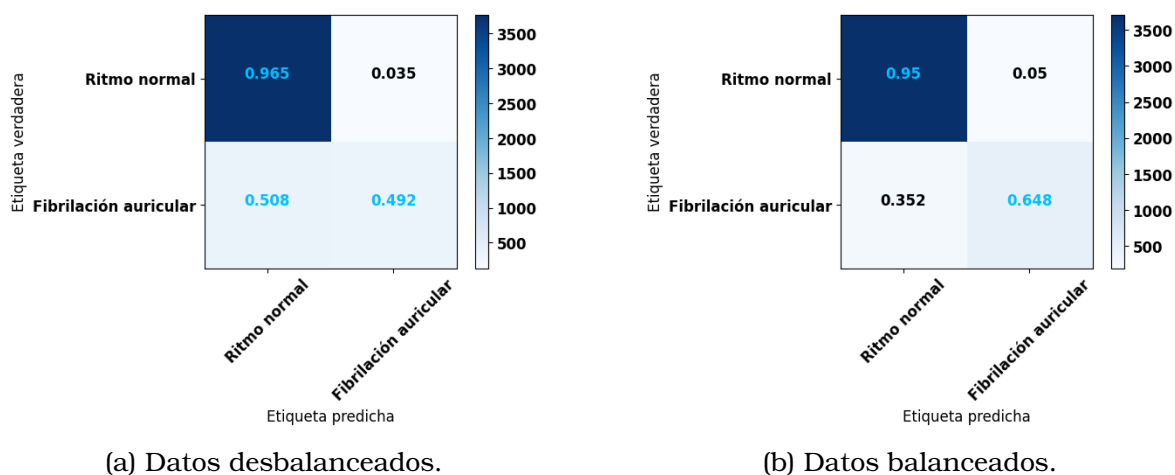


Figura 7.3: Resultados de clasificación (2 clases) de señales ECG, para la base de datos de Physionet Computing in Cardiology Challenge 2017, utilizando un modelo ResNet.

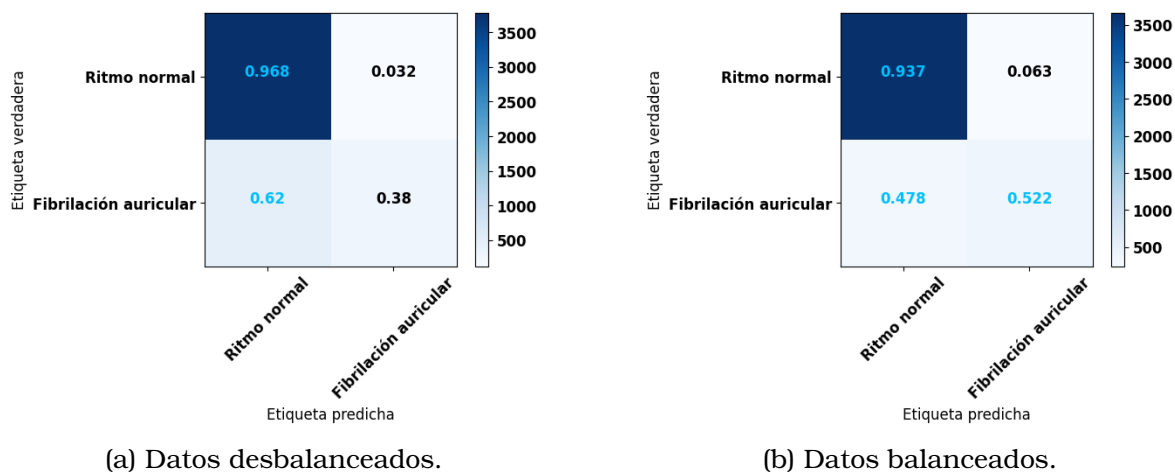


Figura 7.4: Resultados de clasificación (2 clases) de señales ECG, para la base de datos de Physionet Computing in Cardiology Challenge 2017, utilizando un modelo más compacto.

Por otra parte, para el segundo experimento se implementó únicamente el modelo ResNet propuesto en la sección 4.2, el cual fue entrenado durante 50 *epochs* para realizar una clasificación multiclase. En la figura 7.5 se presentan las matrices de confusión con los resultados de evaluación del modelo tanto para el conjunto de datos equilibrado como para el conjunto de datos desequilibrado. Posteriormente, se analizarán en detalle los resultados obtenidos en ambos experimentos.

## Resultados de clasificación con señales ECG generadas sintéticamente

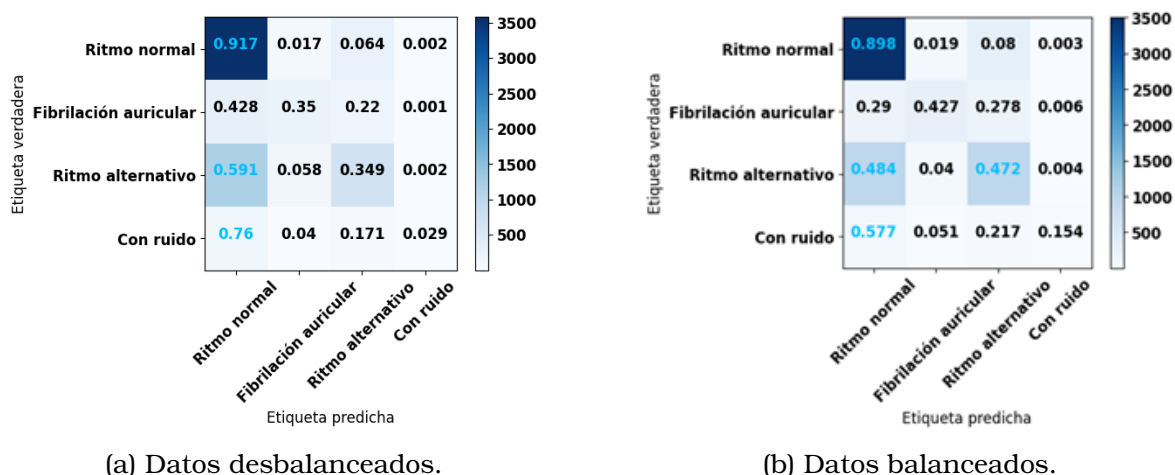


Figura 7.5: Resultados de clasificación (4 clases) de señales ECG, para la base de datos de Physionet Computing in Cardiology Challenge 2017, utilizando un modelo ResNet.

Como se observa en las matrices de confusión de las figuras 7.3a, 7.4a y 7.5a, al trabajar con un conjunto de entrenamiento altamente desbalanceado era predecible que se obtendría un rendimiento bajo en la clasificación de las clases minoritarias (A, O y ~), en comparación a la clase mayoritaria (N). En contraste, al examinar las matrices de confusión en las figuras 7.3b, 7.4b y 7.5b, que corresponden a conjuntos de entrenamiento balanceados, se observa una mejora considerable en la clasificación de las clases minoritarias. Para proporcionar una visión más precisa de la mejora obtenida, los cuadros 7.3 y 7.4 presentan los porcentajes de mejora en la clasificación de cada clase al utilizar conjuntos de datos equilibrados con muestras sintéticas generadas con redes GAN.

Modelo	N	A
ResNet	-1.5 %	21.9 %
Embebido	-3.1 %	14.2 %

Cuadro 7.3: Porcentajes de mejora en los modelos de clasificación, considerando dos clases del conjunto de datos brindado por Physionet para el Computing in Cardiology Challenge 2017.

Modelo	N	A	O	~
ResNet	-1.9 %	7.7 %	12.3 %	12.5 %

Cuadro 7.4: Porcentajes de mejora en el modelo de clasificación, considerando cuatro clases del conjunto de datos brindado por Physionet para el Computing in Cardiology Challenge 2017.

Un aspecto destacado en los resultados presentados en los cuadros 7.3 y 7.4 es la mejora en el rendimiento de las clases minoritarias, que lograron clasificar correctamente un mayor número de muestras como verdaderos positivos. Sin embargo, se observó una ligera disminución en la precisión de clasificación de la clase mayoritaria. A pesar de esto, se ha alcanzado el objetivo principal de este trabajo, que consistía en mejorar la capacidad de clasificación en las clases minoritarias.

## **7.2. Base de datos del Computing in Cardiology Challenge 2017**

---

Los resultados presentados en este capítulo destacan la importancia de abordar el desequilibrio de clases en el conjunto de entrenamiento para lograr una clasificación más precisa y equitativa de todas las clases. Además, respaldan la afirmación de que la arquitectura de red GAN propuesta tiene la capacidad de generar señales sintéticas que contribuyen a la mejora de modelos de clasificación de señales ECG, enfocados en la detección de enfermedades cardíacas.

## Capítulo 8

# Conclusiones

En el presente capítulo se presentan las principales conclusiones alcanzadas a partir de los resultados y análisis realizados en este trabajo final de máster.

En primer lugar, se destaca la importancia de la detección y clasificación precisa de enfermedades cardíacas y cardiovasculares debido a su alta incidencia y riesgo para la salud pública, siendo la principal causa de mortalidad y morbilidad en el mundo. También se destaca el desafío que presenta la detección de la fibrilación auricular debido a su naturaleza asintomática o intermitente, siendo de vital importancia el diagnóstico temprano y preciso de esta afección para prevenir complicaciones graves.

Durante los últimos años, se han propuesto y aplicado diversos métodos de clasificación de arritmias basados en aprendizaje automático y aprendizaje profundo. Sin embargo, uno de los desafíos en la clasificación de arritmias es el desequilibrio de clases en los conjuntos de datos, donde ciertas afecciones cardíacas son menos frecuentes en comparación con el ritmo sinusal normal. Lo que generalmente lleva a una predicción sesgada hacia la clase mayoritaria y una menor sensibilidad en la detección de las clases minoritarias.

En este trabajo, se ha abordado el desafío del desequilibrio de clases mediante el uso de técnicas de generación de señales sintéticas basadas en redes generativas adversarias (GAN). Esto mediante el desarrollo de un modelo cuya arquitectura se compone de un generador formado principalmente por redes BiLSTM y un discriminador formado por redes CNN. En cuanto al modelo de clasificación de señales, se realizaron dos propuestas iniciales. La primera propuesta consistió en un modelo cuya arquitectura empleaba redes CNN simples, mientras que, la segunda propuesta incluía el uso de bloques de redes residuales. Los resultados que se observaron al evaluar dichos modelos, concluyeron que el modelo ResNet presenta un mejor rendimiento en el proceso de clasificación, para todas las bases de datos utilizadas en las pruebas. No obstante, como parte de este trabajo consistía en desarrollar un modelo de clasificación capaz de ser embebido en un dispositivo portátil, era indispensable tomar en cuenta las limitaciones impuestas por el hardware, lo que llevó a una tercera propuesta.

El último modelo de clasificación propuesto consistió en una versión más compacta, compuesto únicamente por cuatro capas densas secuenciales. A pesar de que los resultados de clasificación obtenidos con este modelo fueron inferiores en comparación con el modelo ResNet, es importante destacar que este modelo más compacto

---

logró cumplir con las restricciones impuestas por el hardware, particularmente en términos de memoria.

Por otra parte, el modelo de red GAN implementado para la generación de señales sintéticas demostró ser capaz de reproducir de manera precisa los patrones y características morfológicas presentes en las señales ECG reales utilizadas durante el entrenamiento. Además, mostró ser una estrategia exitosa para abordar el desequilibrio de clases en las bases de datos y mejorar el rendimiento de los modelos de clasificación, obteniendo porcentajes de mejora de hasta el 21.9% en la clasificación de clases minoritarias, lo cual es altamente prometedor. Los códigos implementados en este trabajo se encuentran disponibles en el siguiente repositorio: [enlace al repositorio](#).

Considerando las limitaciones actuales y los resultados obtenidos, se sugiere como trabajo futuro explorar la utilización de microcontroladores con mayor capacidad de procesamiento y memoria. Estos dispositivos permitirían superar las limitaciones presentes en este trabajo y manejar modelos de clasificación mucho más complejos y exigentes, que requieren una cantidad significativa de recursos para ejecutarse eficientemente. Al eliminar las restricciones actuales impuestas por el hardware, se abriría la posibilidad de investigar y evaluar otras arquitecturas de modelos de clasificación más avanzadas, con el objetivo de mejorar aún más los porcentajes de acierto obtenidos en este estudio.

Además, este trabajo sienta las bases para continuar trabajando en el desarrollo de un dispositivo con un nivel mayor de madurez con vistas a una posible comercialización. Dicho dispositivo tendría la funcionalidad de un asistente cognitivo especializado en el monitoreo y detección de diferentes arritmias cardíacas. El objetivo sería contribuir a la detección temprana de estas arritmias y a prevenir su evolución hacia enfermedades cardiovasculares más graves, en especial en áreas rurales o con acceso limitado a instalaciones y servicios médicos.

# Bibliografía

- [1] Organización Mundial de la Salud Enfermedades cardiovasculares. (2017).  
Extraído de: [https://www.who.int/es/news-room/fact-sheets/detail/cardiovascular-diseases-\(cvds\)](https://www.who.int/es/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds))
- [2] Infocoponline Mapa mundial de las enfermedades cardiovasculares de la Federación Mundial del Corazón. (Consejo General de la Psicología de España,2017),  
[http://www.infocop.es/view\\_article.asp?id=6908](http://www.infocop.es/view_article.asp?id=6908)
- [3] Chow, G., Marine, J. & Fleg, J. Epidemiology of Arrhythmias and Conduction Disorders in Older Adults. *Clinics In Geriatric Medicine*. 28, 539-553 (2012).
- [4] Kumari, P., Mahew, L. & Syal, P. Increasing trend of wearables and multimodal interface for human activity monitoring: A review. *Biosensors And Bioelectronics*. 90 pp. 298-307 (2017)
- [5] Fan, W., Si, Y., Yang, W. & Sun, M. Imbalanced ECG data classification using a novel model based on active training subset selection and modified broad learning system. *Measurement*. 198 pp. 111412 (2022),<https://www.sciencedirect.com/science/article/pii/S0263224122006443>
- [6] Han, H., Wang, W. & Mao, B. Borderline-SMOTE: A New Over-Sampling Method in Imbalanced Data Sets Learning. *Advances In Intelligent Computing, ICIC 2005, Lecture Notes in Computer Science, vol 3644*. (2005).
- [7] López Farré, A. & Macaya Miguel, C. Libro de la salud cardiovascular del Hospital Clínico San Carlos y la Fundación BBVA. (2009).
- [8] Romero, J. Análisis de señales electrocardiográficas usando técnicas de procesamiento digital. Universitat Oberta de Catalunya,(2015), <http://hdl.handle.net/10609/40186>
- [9] Kusumoto, F. ECG Interpretation: From Pathophysiology to Clinical Application. (Springer,2009).
- [10] Vera, O., Duque Cardona, E. & Rivera Piedrahita, J. Extracción de características de la señal electrocardiográfica mediante software de análisis matemático. *Scientia Et Technica*. (2006), <https://www.redalyc.org/articulo.oa?id=84911639012>
- [11] Proakis, J. Tratamiento digital de señales: Principios, algoritmos y aplicaciones. 3a ed. Ed. Prentice Hall. (1997).
- [12] Cuesta, D. Revisión de Métodos y Algoritmos para el Tratamiento de Señales Electrocardiográficas. (1999).



- 
- [13] Bistel Esquivel, R. & Fajardo Márquez, A. Diseño de un Sistema de Adquisición y Procesamiento de la Señal de ECG basado en Instrumentación Virtual. *Ingeniería Electrónica, Automática Y Comunicaciones*. 36, 17-30 (2015).
- [14] Sahoo, S., Kanungo, B., Behera, S. & Sabut, S. Multiresolution wavelet transform based feature extraction and ECG classification to detect cardiac abnormalities. *Measurement*. 108, 55-66. (2017).
- [15] Pławiak, P. Novel methodology of cardiac health recognition based on ECG signals and evolutionary-neural system. *Expert Systems With Applications*. 92, 334-349. (2018).
- [16] Shi, H., Wang, H., Huang, Y., Zhao, L., Qin, C. & Liu, C. A hierarchical method based on weighted extreme gradient boosting in ECG heartbeat classification. *Computers In Biology And Medicine*. 171, 1-10. (2019).
- [17] Kumar, R.G.; Kumaraswamy, Y. Investigating cardiac arrhythmia in ECG using random forest classification. *Int. J. Comput. Appl.* 37, 31–34. (2012).
- [18] Park, J.; Lee, K.; Kang, K. Arrhythmia detection from heartbeat using k-nearest neighbor classifier. In Proceedings of the 2013 IEEE International Conference on Bioinformatics and Biomedicine, Shanghai, China, 18–21 December 2013; pp. 15–22.
- [19] Acharya, U., Oh, S., Hagiwara, Y., Tan, J., Adam, M., Gertych, A. & Tan, R. A deep convolutional neural network model to classify heartbeats. *Computers In Biology And Medicine*. 89 pp. 389-396 (2017).
- [20] Acharya, U., Fujita, H., Oh, S., Hagiwara, Y., Tan, J. & Adam, M. Application of deep convolutional neural network for automated detection of myocardial infarction using ecg signals. *Information Sciences*. 415 pp. 190-198 (2017).
- [21] Acharya, U., Fujita, H., Lih, O., Hagiwara, Y., Tan, J. & Adam, M. Automated detection of arrhythmias using different intervals of tachycardia ECG segments with convolutional neural network. *Information Sciences*. 405 pp. 81-90 (2017).
- [22] Acharya, U., Fujita, H., Oh, S., Raghavendra, U., Tan, J., Adam, M., Gertych, A. & Hagiwara, Y. Automated identification of shockable and non-shockable life-threatening ventricular arrhythmias using convolutional neural network. *Future Generation Computer Systems*. 79 pp. 952-959 (2018).
- [23] Sellami, A. & Hwang, H. A robust deep convolutional neural network with batch-weighted loss for heartbeat classification. *Expert Systems With Applications*. 122 pp. 75-84 (2019).
- [24] Kachuee, M., Fazeli, S. & Sarrafzadeh, M. ECG Heartbeat Classification: A Deep Transferable Representation. *2018 IEEE International Conference On Healthcare Informatics (ICHI)*. pp. 443-444 (2018).
- [25] Zha, M. Time Series Generation with Masked Autoencoder. *CoRR*. abs/2201.07006 (2022), <https://arxiv.org/abs/2201.07006>
- [26] Baytas, I., Xiao, C., Zhang, X., Wang, F., Jain, A. & Zhou, J. Patient Subtyping Via Time-Aware LSTM Networks. *Proceedings Of The International Conference On Knowledge Discovery And Data Mining*. (2017).

- [27] Ha, D. & Eck, D. A Neural Representation of Sketch Drawings. *Proceedings Of The International Conference On Learning Representations*. (2018).
- [28] Madan, Y., Veetil, I., V, S., EA, G. & KP, S. Synthetic Data Augmentation of MRI using Generative Variational Autoencoder for Parkinson's Disease Detection. *Evolution In Computational Intelligence*. 267 (2022), [https://doi.org/10.1007/978-981-16-6616-2\\_16](https://doi.org/10.1007/978-981-16-6616-2_16)
- [29] Desai, A., Freeman, C., Wang, Z. & Beaver, I. TimeVAE: A Variational Auto-Encoder for Multivariate Time Series Generation. *CoRR*. abs/2111.08095 (2021), <https://arxiv.org/abs/2111.08095>
- [30] Esteban, C., Hyland, S. & Rätsch, G. Real-valued (medical) time series generation with recurrent conditional GANs. *ArXiv Preprint ArXiv:1706.02633*. (2017).
- [31] Neurosky. ECG-BMD101. Recuperado de <https://store.neurosky.com/products/ecg-bmd101>.
- [32] M5Stack. M5StickC Plus datasheet. Extraído de: <https://cdn.competec.ch/documents2/8/5/9/185624958/185624958.pdf>.
- [33] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. & Bengio, Y. Generative Adversarial Nets. *Advances In Neural Information Processing Systems* 27. pp. 2672-2680 (2014), <http://papers.nips.cc/paper/5423-generative-adversarial-nets.pdf>
- [34] Kandala, N. & Dhuli, R. Classification of imbalanced ECG beats using resampling techniques and AdaBoost ensemble classifier. *Biomedical Signal Processing And Control*. 41 pp. 242-254 (2018).
- [35] Niu, L., Chen, C., Liu, H., Zhou, S. & Shu, M. A Deep-Learning Approach to ECG Classification Based on Adversarial Domain Adaptation. *Healthcare*. (2020).
- [36] Ukil, A., Bandyopadhyay, S., Puri, C., Singh, R. & Pal, A. Class augmented semi-supervised learning for practical clinical analytics on physiological signals. *Proc. Mach. Learn. Health (ML4H) Workshop NeurIPS*. pp. 1-5 (2018).
- [37] McSharry, P., Clifford, G., Tarassenko, L. & Smith, L. A dynamical model for generating synthetic electrocardiogram signals. *IEEE Transactions On Biomedical Engineering*. 50, 289-294 (2003).
- [38] Delaney, A., Brophy, E. & Ward, T. Synthesis of realistic ECG using generative adversarial networks. *ArXiv Preprint ArXiv:1909.09150*. (2019).
- [39] Wang, P., Hou, B., Shao, S. & Yan, R. ECG arrhythmias detection using auxiliary classifier generative adversarial network and residual network. *IEEE Access*. 7 pp. 100910-100922 (2019).
- [40] Esteban, C., Hyland, S. & Ratsch, G. Real-valued (medical) time series generation with recurrent conditional GANs. *ArXiv Preprint ArXiv:1706.02633*. (2017).
- [41] Zhu, F., Ye, F., Fu, Y., Liu, Q. & Shen, B. Electrocardiogram generation with a bidirectional LSTM-CNN generative adversarial network. *Scientific Reports*. 9, 1-11 (2019).

- 
- [42] Dasgupta, S., Das, S. & Bhattacharya, U. CardioGAN: An Attention-based Generative Adversarial Network for Generation of Electrocardiograms. *2020 25th International Conference On Pattern Recognition (ICPR)*. pp. 3193-3200 (2021).
- [43] AlMahamdy, M. & Riley, H. Performance study of different denoising methods for ECG signals. *Procedia Computer Science*. 37 pp. 325-332 (2014).
- [44] Cuturi, M. & Blondel, M. Soft-DTW: a differentiable loss function for time-series. *ICML*. pp. 894-903 (2017).
- [45] Serra, J. & Arcos, J. An empirical evaluation of similarity measures for time series classification. *Knowledge-Based Systems*. 67 pp. 305-314 (2014).
- [46] Zheng, Z., Chen, Z., Hu, F., Zhu, J., Tang, Q. & Liang, Y. An automatic diagnosis of arrhythmias using a combination of CNN and LSTM technology. *Electronics*. 9 pp. 121 (2020).
- [47] Gent, P., Farah, H., Nes, N. & Arem, B. Analysing Noisy Driver Physiology Real-Time Using Off-the-Shelf Sensors: Heart Rate Analysis Software from the Taking the Fast Lane Project. *Journal Of Open Research Software*. 7, 32 (2019), <http://doi.org/10.5334/jors.241>