



EDUCACIÓN

SECRETARÍA DE EDUCACIÓN PÚBLICA



TECNOLÓGICO
NACIONAL DE MÉXICO

Tecnológico Nacional de México

Centro Nacional de Investigación
y Desarrollo Tecnológico

Tesis de Maestría

**Clasificación de la Fibrilación Auricular en Señales
Electrocardiográficas Utilizando una Red Recurrente
Profunda Entrenada con Momentos de Tiempo-
Frecuencia**

presentada por

Ing. Christian Jaime García Aquino

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

Director de tesis

Dr. Dante Mújica Vargas

Cuernavaca, Morelos, México. Enero de 2023.



EDUCACIÓN
SECRETARÍA DE EDUCACIÓN PÚBLICA



TECNOLÓGICO
NACIONAL DE MÉXICO

Centro Nacional de Investigación
y Desarrollo Tecnológico
Departamento de Ciencias Computacionales

Cuernavaca, Mor., 11/enero/2023

OFICIO No. DCC/007/2023

Asunto: Aceptación de documento de tesis
CENIDET-AC-004-M14-OFCIO

CARLOS MANUEL ASTORGA ZARAGOZA
SUBDIRECTOR ACADÉMICO
PRESENTE

Por este conducto, los integrantes de Comité Tutorial de GARCIA AQUINO CHRISTIAN JAIME, con número de control M21CE012, de la Maestría en Ciencias en Ciencias de la Computación, le informamos que hemos revisado el trabajo de tesis de grado titulado **"Clasificación de la fibrilación auricular en señales electrocardiográficas utilizando una red recurrente profunda entrenada con momentos de tiempo-frecuencia"** y hemos encontrado que se han atendido todas las observaciones que se le indicaron, por lo que hemos acordado aceptar el documento de tesis y le solicitamos la autorización de impresión definitiva.

Dante

DANTE MÚJICA VARGAS
Director de tesis

Juan Gabriel

JUAN GABRIEL GONZÁLEZ SERNA
Revisor 1

Nimrod

NIMROD GONZÁLEZ FRANCO
Revisor 2

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

SEP TecNM CENTRO NACIONAL DE INVESTIGACIÓN
Y DESARROLLO TECNOLÓGICO
RECIBIDO
13 ENE 2023
CMA
SUBDIRECCIÓN ACADÉMICA

EDUCACIÓN | TECNOLOGICO NACIONAL DE MEXICO
13 ENE 2023
CENTRO NACIONAL DE INVESTIGACIÓN
Y DESARROLLO TECNOLÓGICO
SERVICIOS ESCOLARES
RECIBIDO



Interior Internado Palmira S/N, Col. Palmira, C. P. 62490, Cuernavaca, Morelos
Tel. 01 (777) 3627770, ext. 3202, e-mail: dcc@cenidet.tecnm.mx | cenidet.tecnm.mx



2023
AÑO DE
Francisco VILLA
EL PROFESOR Y EL SOLDADO

Cuernavaca, Mor.,

No. De Oficio:

Asunto:

13/enero/2023

SAC/016/2023

**Autorización de
impresión de tesis**

**CHRISTIAN JAIME GARCÍA AQUINO
CANDIDATO AL GRADO DE MAESTRO EN CIENCIAS
DE LA COMPUTACIÓN
P R E S E N T E**

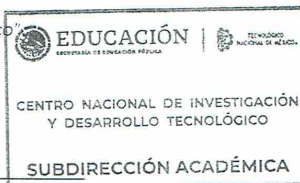
Por este conducto, tengo el agrado de comunicarle que el Comité Tutorial asignado a su trabajo de tesis titulado **"CLASIFICACIÓN DE LA FIBRILACIÓN AURICULAR EN SEÑALES ELECTROCARDIOGRÁFICAS UTILIZANDO UNA RED RECURRENTE PROFUNDA ENTRENADA CON MOMENTOS DE TIEMPO-FRECUENCIA"**, ha informado a esta Subdirección Académica, que están de acuerdo con el trabajo presentado. Por lo anterior, se le autoriza a que proceda con la impresión definitiva de su trabajo de tesis.

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

ATENTAMENTE

Excelencia en Educación Tecnológica®

"Conocimiento y tecnología al servicio de México"



**CARLOS MANUEL ASTORGA ZARAGOZA
SUBDIRECTOR ACADÉMICO**

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

CMAZ/RMA

Dedicatoria

Dedico esta tesis principalmente a mis papás Reyna y Jaime por siempre impulsarme a seguir adelante, a mis hermanas Lourdes, Gisela y Gladys por su apoyo incondicional, a mi sobrina Alejandra por darme motivos a ser mejor cada día.

A todos los reales que siempre han estado conmigo, gracias por soportarme.

Sigo agradeciendo con hechos.

Agradecimientos

Agradezco al Consejo Nacional de Ciencia y Tecnología (CONACYT) por el apoyo brindado durante mis estudios de maestría mediante su sistema de becas de posgrado. Al Centro Nacional de Investigación y Desarrollo Tecnológico (CENIDET), perteneciente al TECNM, por haberme permitido realizar los estudios de una Maestría en Ciencias en sus instalaciones.

Agradezco a mi director de tesis Dr. Dante Mújica Vargas por su apoyo, esfuerzo, dedicación y guía brindada durante mi formación. A el comité revisor, Dr. Gabriel González Serna y al Dr. Nimrod González Franco, por sus aportaciones y esfuerzo para el desarrollo de esta investigación. Al Mtro. Manuel Matuz por su confianza, amistad y dedicación brindada.

Resumen

En este trabajo se propone un método eficiente para la clasificación de la Fibrilación Auricular basado en características en el Dominio de Tiempo-Frecuencia y una Red Neuronal Híbrida Bidireccional. La Metodología de solución se divide en 3 módulos principales: Transformación al dominio de Tiempo-Frecuencia, Extracción de características y reducción de dimensionalidad y entrenamiento de la Red Neuronal Híbrida Bidireccional.

El primer módulo se enfoca a la adquisición de los datos a usar, así como el proceso de transformación de las señales electrocardiográficas del dominio secuencial y caótico al dominio de Tiempo-Frecuencia a partir de la Transformada de Fourier de Tiempo Corto. El segundo módulo es referente a la extracción de características como la frecuencia instantánea y la entropía espectral para reducir dimensionalidad de los datos conservando el dominio de Tiempo-Frecuencia. En el tercer módulo, se detalla la formalización de la Red Neuronal Híbrida Bidireccional, la descripción puntual de cada capa constituyente de la red, así como el procesamiento interno de las características dentro de la misma.

El método propuesto se evaluó con métricas derivadas de la matriz de confusión (*Accuracy*, *Precision*, *Recall*, *F1 – Score*) y otras consideradas de la literatura (*CohenKappa* y *MCC*). Para los experimentos se utilizó la base de datos Physionet CinC 2017, la cuál contiene señales electrocardiográficas con Fibrilación Auricular y señales con ritmos cardíacos normales. Los resultados cualitativos, muestran que el método propuesto resulta eficiente para la clasificación de la Fibrilación Auricular, teniendo un rendimiento mayor a 0.90 en las métricas de evaluación y una reducción del costo computacional comparado con características en crudo de mas del 90%.

Palabras Clave: Fibrilación Auricular, Tiempo-Frecuencia, Red Neuronal Híbrida Bidireccional, Transformada de Fourier de Tiempo Corto.

Abstract

In this work, an efficient method for the classification of Atrial Fibrillation based on characteristics in the Time-Frequency Domain and a Bidirectional Hybrid Neural Network is proposed. The solution Methodology is divided into 3 main modules: Transformation to the Time-Frequency domain, Feature extraction and dimensionality reduction and Bidirectional Hybrid Neural Network training.

The first module focuses on the acquisition of the data to be used, as well as the transformation process of the electrocardiographic signals from the sequential and chaotic domain to the Time-Frequency domain from the Short Time Fourier Transform. The second module refers to the extraction of characteristics such as instantaneous frequency and spectral entropy to reduce the dimensionality of the data while preserving the Time-Frequency domain. In the third module, the formalization of the Bidirectional Hybrid Neural Network is detailed, the specific description of each constituent layer of the network, as well as the internal processing of the characteristics within it.

The proposed method was evaluated with metrics derived from the confusion matrix (*Accuracy, Precision, Recall, F1 – Score*) and others considered from the literature (*CohenKappa* and *MCC*). For the experiments, the Physionet CinC 2017 database was obtained, which contains electrocardiographic signals with Atrial Fibrillation and signals with normal heart rhythms. The qualitative results show that the proposed method is efficient for the classification of Atrial Fibrillation, having a performance greater than 0.90 in the evaluation metrics and a reduction in computational cost compared to raw characteristics of more than 90%.

Keywords: Atrial Fibrillation, Time-Frequency, Bidirectional Hybrid Neural Network, Short Time Fourier Transform.

Índice General

Resumen	vii
Índice de Figuras	xi
Índice de Tablas	xiii
1 Introducción	1
1.1 Planteamiento del Problema	1
1.2 Objetivos	2
1.3 Alcances y Limitaciones	3
1.4 Justificación	4
1.5 Organización de la Tesis	4
2 Marco Teórico	5
2.1 Fibrilación Auricular	5
2.2 Electrocardiograma	6
2.3 Transformada de Fourier de Corto Plazo	6
2.4 Aprendizaje Profundo	7
2.5 Redes Neuronales Recurrentes	8
2.6 Arquitecturas de Redes Neuronales Recurrentes	11
2.7 Discusión	13
3 Estado del Arte	14
3.1 Antecedentes	14
3.2 Trabajos relacionados	15
3.3 Discusión	33
4 Metodología	37
4.1 Transformación al dominio de Tiempo-Frecuencia	38
4.2 Extracción de características y reducción de dimensionalidad	41
4.3 Red Neuronal Híbrida Bidireccional	45
4.4 Discusión	52
5 Experimentación y Resultados	53
5.1 Diseño Experimental	53

5.2	Sintonización de la Red Neuronal Híbrida Bidireccional	56
5.3	Comparación con diferentes tipos de características	62
5.4	Comparación del modelo propuesto con otros modelos neuronales . . .	67
5.5	Discusión	68
6	Conclusiones	69
6.1	Objetivos y alcances logrados	69
6.2	Resultados del trabajo	70
6.3	Conclusiones	71
6.4	Trabajos futuros	72
A	Matrices de confusión	73
B	Producción científica	81
C	Acrónimos	89
	Bibliografía	90

Índice de Figuras

2.1	Sistema eléctrico del corazón con FA (Imagen extraída de NHLBI, 2021).	5
2.2	Diagrama de una señal de ECG normal (Imagen extraída de Vogt, 2019).	6
2.3	La meseta de rendimiento en el aprendizaje automático (Imagen extraída de Buduma and Locascio, 2017).	7
2.4	Estructura básica de una RNN (Vogt, 2019).	9
2.5	Estructura de una célula LSTM (Vogt, 2019).	12
2.6	Estructura de una BiLSTM (Vogt, 2019).	13
3.1	Arquitectura de red.	16
3.2	Arquitectura general de 8CSL.	17
3.3	Estructura del modelo propuesto.	18
3.4	Arquitectura del modelo híbrido propuesto.	19
3.5	Arquitectura de MultiFusionNet.	20
3.6	Arquitectura del modelo de aprendizaje propuesto.	21
3.7	Señal de ECG de un voluntario.	22
3.8	(a) Señal en crudo. (b) Señal con picos R filtrados.	23
3.9	Diagrama de bloques del método propuesto.	23
3.10	Descripción general de la arquitectura del modelo propuesta.	25
3.11	Diferentes técnicas de aumento realizadas en el conjunto de datos.	26
3.12	Procesamiento de lenguaje de ECG.	27
3.13	Modelos usados en el trabajo.	29
3.14	Arquitectura general DenseNet.	30
3.15	El diagrama de bloques del método propuesto para la clasificación de señales de ECG.	31
3.16	Arquitectura para la detección y clasificación de Fibrilación Auricular.	32
4.1	Diagrama de bloques de la metodología.	37
4.2	Señales de la base de datos seleccionada.	38
4.3	Fases de la STFT.	39
4.4	Representación de un espectrograma.	41
4.5	Diagrama de una señal y un espectro hipotético en un instante de tiempo (Extraído de Benjumea,2003).	41
4.6	Representación de la Frecuencia Instantánea.	43
4.7	Representación de la función de Entropía.	44

4.8	Representación de la Entropía Espectral.	45
4.9	Arquitectura de la BiHNN.	45
4.10	Estructura interna de la capa de entrada de secuencias.	46
4.11	Estructura interna de una capa BiLSTM.	47
4.12	Capa totalmente conectada.	48
5.1	Resultados con el optimizador RMSPROP.	57
5.2	Resultados con el optimizador SGDM.	58
5.3	Resultados con el optimizador ADAM.	59
5.4	Resultados con diferentes capas recurrentes.	61
5.5	Resultados con distintas capas BiLSTM.	62
5.6	Resultados con datos en crudo	63
5.7	Resultados con características estadísticas	64
5.8	Descripción de las capas de la red modificada.	65
5.9	Resultados con características de TF	66
5.10	Resultados con distintos tipos de CNN.	68
A.1	Matrices de confusión obtenidas con el optimizador RMSPROP.	73
A.2	Matrices de confusión obtenidas con el optimizador SGDM.	74
A.3	Matrices de confusión obtenidas con el optimizador ADAM.	75
A.4	Matrices de confusión obtenidas con datos en crudo.	76
A.5	Matrices de confusión obtenidas con características estadísticas.	77
A.6	Matrices de confusión obtenidas con características de TF.	78
A.7	Matrices de confusión obtenidas con diferentes capas recurrentes.	79
A.8	Matrices de confusión obtenidas con diferentes cantidad de capas BiLSTM.	79
A.9	Matrices de confusión obtenidas con diferentes redes neuronales convolucionales.	80
B.1	Constancia de autor del artículo publicado en el congreso internacional WITCOM 2021.	81
B.2	Primera página del artículo publicado en el congreso internacional WITCOM 2021.	82
B.3	Constancia de autor del artículo publicado en el congreso COMIA 2022.	83
B.4	Primera página del artículo publicado en el congreso COMIA 2022.	84
B.5	Constancia de autor del artículo publicado en el congreso internacional WITCOM 2022.	85
B.6	Primera página del artículo publicado en el congreso internacional WITCOM 2022.	86
B.7	Constancia de co-autor de artículo publicado en el congreso MCPR 2021.	87
B.8	Primera página del artículo publicado en el congreso MCPR 2021.	88

Índice de Tablas

3.1	Resumen del Estado del Arte	33
5.1	Cantidad de señales por clase.	54
5.2	Señales divididas para entrenamiento y pruebas	55
5.3	Resultados con el optimizador RMSPROP	57
5.4	Resultados con el optimizador SGDM	58
5.5	Resultados con el optimizador ADAM	59
5.6	Resultados con diferentes capas recurrentes	60
5.7	Resultados con diferente cantidad de capas BiLSTM	61
5.8	Resultados de Rendimiento con Datos en crudo.	63
5.9	Resultados con características estadísticas	64
5.10	Resultados del rendimiento con características de Tiempo-Frecuencia. . .	65
5.11	Costo computacional de clasificación	66
5.12	Resultados con distintos tipos de CNN	67
6.1	Objetivos específicos.	69
6.2	Alcances.	70

Introducción

La importancia clínica de las arritmias cardíacas va en aumento dada la alza de incidencia y prevalencia, que se asocian principalmente con el envejecimiento de la población. Entre estas enfermedades cardíacas, se destaca la Fibrilación Auricular (FA) ya que es el tipo de arritmia común en adultos con una tendencia creciente más significativa, siendo esta en diversas ocasiones difícil diagnosticar debido que en algunos casos se presenta un comportamiento paroxístico y/o un cuadro asintomático (Jiménez, 2018). Por ende, existe el interés por realizar la clasificación automática de la FA utilizando técnicas de Inteligencia Artificial, específicamente algoritmos de Aprendizaje Profundo como las Redes Neuronales Profundas, las cuales permiten contribuir en gran medida en el dictámen médico como parte de la práctica clínica.

En la literatura, dichas tareas de clasificación se realizan empleando la información en crudo que se obtiene de las señales de ECG, que en términos prácticos, no son significativas y/o discriminatorias dada la naturaleza caótica por la que está caracterizada, por ende, se generan problemas en los algoritmos de Aprendizaje Profundo como un alto costo computacional en las fases de entrenamiento, además de un rendimiento reducido y limitado en las tareas de clasificación.

En este sentido, el procesamiento digital de señales es una parte fundamental en cuestiones de tareas clasificación en señales electrocardiográficas (ECG), ya que ayuda en gran manera a mitigar los problemas que adolecen los algoritmos de Aprendizaje Profundo a la hora de entrenar y clasificar dichas señales.

Por lo que el objetivo principal es clasificar la Fibrilación Auricular usando una Red Recurrente Profunda, la cuál será entrenada con características significativas conocidas como momentos en el dominio del tiempo y la frecuencia que, como se había mencionado anteriormente, serán características que permitirán mejorar el costo computacional y el rendimiento de clasificación de la red neuronal.

1.1 Planteamiento del Problema

En general las tareas de clasificación de la Fibrilación Auricular se realizan utilizando los datos que son propios de los ECG, y dado que la información en crudo que se obtiene de las señales de ECG no logra ser completamente significativa, ocasiona problemas en los algoritmos computacionales actuales, como un alto costo computacional en las fases de entrenamiento, además de un rendimiento reducido y limitado en las tareas de clasificación.

Para evitar los problemas que ocasiona trabajar con datos en crudo, es necesario realizar el tratamiento de las señales ECG para obtener características de Tiempo-Frecuencia que servirán para entrenar un modelo clasificador mejorando su rendimiento y reduciendo los tiempos de entrenamiento.

Hipótesis: Las Redes Neuronales Recurrentes (RNN) al caracterizarse por trabajar con datos secuenciales y aprender las sucesiones de estados, permiten realizar la predicción de señales futuras cercanas, así como las pasadas. Por lo tanto, trabajar con características de Tiempo-Frecuencia conllevará a mejorar el rendimiento en la clasificación de las señales de ECG y reducir el costo computacional del mismo.

1.1.1 Delimitación del Problema

En este trabajo de investigación se desarrollará la clasificación de Fibrilación Auricular con la cual, se pueda determinar si la señal procesada es un latido normal del corazón o un latido con el ritmo cardiaco alterado causado por la Fibrilación Auricular. Entre las diferentes arquitecturas de Red Neuronal Recurrente que se encuentran en la literatura, se hará uso de un Modelo de Red Neuronal Recurrente de tipo LSTM para realizar la clasificación de forma automática. Dicho modelo permitirá realizar el entrenamiento con características de Tiempo-Frecuencia, y así poder realizar la clasificación con un costo computacional moderado.

1.1.2 Complejidad del Problema

La complejidad del problema radica en los siguientes puntos:

1. Extraer las características de Tiempo-Frecuencia de las señales electrocardiográficas.
2. Modificar el modelo RNN para ser entrenada y evaluada con las características de tiempo-frecuencia.

1.2 Objetivos

1.2.1 Objetivo General

Diseñar, implementar y evaluar un clasificador que permita detectar la Fibrilación Auricular en señales electrocardiográficas, utilizando una Red Recurrente Profunda del tipo LSTM entrenada con momentos de Tiempo-Frecuencia.

1.2.2 Objetivos Específicos

Aunado al objetivo principal se consideran los siguientes objetivos específicos:

- Analizar y comprender el paradigma de las Redes Neuronales Recurrentes Profundas, con mayor énfasis en la arquitectura *Long Short-Term Memory*.
- Analizar y comprender las señales electrocardiográficas.
- Extraer características de Tiempo-Frecuencia a partir de espectrogramas.
- Definir la arquitectura de RNN.
- Entrenar y evaluar la RNN utilizando las series de tiempo en crudo.
- Modificar la arquitectura de RNN considerando los momentos de tiempo-frecuencia extraídos de los espectrogramas.
- Entrenar y evaluar la RNN con funciones de Tiempo-Frecuencia elegidas, y publicar los resultados obtenidos.

1.3 Alcances y Limitaciones

1.3.1 Alcances

- Entrenar con diferentes tipos de datos para observar el desempeño en cada uno de ellos.
- Modificar el modelo RNN para entrenarlo con características de Tiempo-Frecuencia.
- Mantener un costo computacional óptimo con el modelo propuesto.
- Determinar con al menos un 80% en el rendimiento general de clasificación.
- Evaluar los resultados obtenidos con base en las métricas definidas en la literatura.

1.3.2 Limitaciones

- La mayoría de las bases de datos de señales tienen clases desbalanceadas, limitando la cantidad de señales utilizables para clasificar.
- El trabajo se centrará solo en la clasificación de las señales.
- Únicamente se dará un reporte simple de señales clasificadas sin llegar a reemplazar a un diagnóstico médico.

1.4 Justificación

La clasificación de Fibrilación Auricular es un tema de interés general en los últimos años, pues la prevalencia no solo de la Fibrilación Auricular sino de las arritmias en general va en crecimiento con el paso de los años y afectando con mayor frecuencia a personas de la mediana y tercera edad. El trabajo de investigación por su parte, resalta su importancia al contribuir en el dictamen médico para saber el estado del paciente. La detección temprana de la Fibrilación Auricular reduce considerablemente la tasa de prevalencia y de mortalidad, en cuanto a sufrir de un infarto agudo al miocardio se refiere. El empleo de las Redes Neuronales Recurrentes resulta ser una opción viable sobre las Redes de Aprendizaje Profundo clásicas debido al entrenamiento con datos secuenciales, es decir, las salidas temporales toman en cuenta la información de los tiempos $t - 1$ y $t + 1$, permitiendo así la predicción de las salidas en el presente y en el futuro.

1.5 Organización de la Tesis

Este documento está compuesto por seis capítulos, además de anexos y referencias. El capítulo 2 contiene el marco conceptual donde se definen conceptos como Fibrilación Auricular, Electrocardiograma, Aprendizaje Profundo y Redes Neuronales Recurrentes que facilitarán la comprensión de este proyecto de investigación. En el capítulo 3 se presentan los antecedentes y el estado del arte mencionando aquellos trabajos que fueron considerados de gran relevancia para esta investigación. El capítulo 4 se refiere a la metodología que se desarrolló en esta investigación, detallando las principales aportaciones. En el capítulo 5 se detalla la experimentación realizada y los resultados obtenidos. Finalmente, en el capítulo 6 se describen las conclusiones generales de la investigación, los productos desarrollados y se exponen las propuestas para trabajos futuros.

CAPÍTULO 2

Marco Teórico

En este capítulo se detallan diversos conceptos básicos y específicos que son necesarios para la contextualización y desarrollo del tema de tesis propuesto.

2.1 Fibrilación Auricular

La Fibrilación Auricular es uno de los tipos más comunes de arritmias, que son ritmos cardiacos irregulares. La Fibrilación Auricular provoca latidos más rápidos de lo normal debido a que las cámaras superior e inferior del corazón no funcionan de manera sincrona como se puede observar en la Figura 2.1. Cuando esto sucede, las cámaras inferiores no se llenan por completo ni logran bombear la sangre suficiente a los pulmones y al cuerpo, haciéndose notable una sensación de cansancio, mareo, palpitaciones del corazón o un dolor en el pecho (NHLBI, 2021).

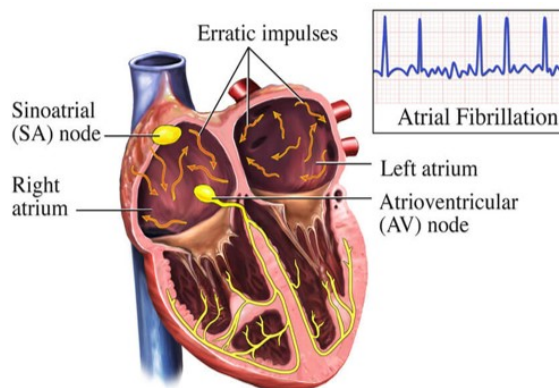


Figura 2.1: Sistema eléctrico del corazón con FA (Imagen extraída de NHLBI, 2021).

La enfermedad es causada por trastornos en la actividad eléctrica de las aurículas desarrollándose contracciones cardiacas poco regulares que perturban el funcionamiento mecánico del corazón, afectando también al sistema cardiovascular, lo cual incrementa los riesgos de sufrir un accidente cerebrovascular, enfermedad de las arterias coronarias y en el peor de los casos la muerte (Vogt, 2019).

2.2 Electrocardiograma

El electrocardiograma (ECG) es una prueba común e indolora (no invasiva) que registra la actividad eléctrica del corazón para la detección de problemas cardíacos basándose en el estudio de la despolarización y repolarización de las fibras cardíacas. Los trastornos son notorios cuando se presentan cambios en las amplitudes e intervalos de las ondas P, Q, R, S y T. El origen de su forma es a partir de la actividad eléctrica comenzando por la despolarización de las fibras musculares auriculares y su recorrido por el nodo auriculoventricular (onda P), la activación de las fibras musculares ventriculares (complejo QRS) y la repolarización de los ventrículos (onda T) (Vogt, 2019). En la Figura 2.2 se puede apreciar el diagrama de un latido cardíaco sano.

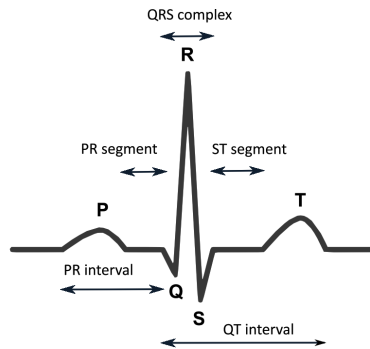


Figura 2.2: Diagrama de una señal de ECG normal (Imagen extraída de Vogt, 2019).

2.3 Transformada de Fourier de Corto Plazo

La Transformada de Fourier de Corto Plazo (del inglés STFT) analiza las señales no estacionarias a través de la Transformada de Fourier (TF). La STFT consiste en dividir la señal en pequeños segmentos de tiempo de tal manera que se pueda suponer que cada segmento de la señal es estacionaria, y así calcular la TF en cada porción de la señal, la cual se toma como una ventana que se desliza a lo largo del eje del tiempo, lo que da como resultado una representación bidimensional de la señal (Smith, 2007).

La representación matemática general de la STFT es:

$$X_m(\omega) = \sum_{n=-\infty}^{\infty} x(n) w(n - mR) e^{-j\omega n} \quad (2.1)$$

donde $x(n)$ es la señal de entrada en el tiempo n dado, $w(n)$ es la función de la longitud de ventana y R es el tamaño del salto en las muestras entre las Transformadas de Fourier a Corto Plazo.

2.4 Aprendizaje Profundo

El Aprendizaje Profundo es un subconjunto del Aprendizaje Automático que se centra en un área de algoritmos inspirados en nuestra comprensión de cómo funciona el cerebro para obtener conocimiento, basándose en la idea de las redes neuronales artificiales y la escala para poder consumir grandes cantidades de datos mediante la profundización de las redes de una manera específica. A través de una red más profunda, un modelo de Aprendizaje Profundo tiene la capacidad de extraer características de datos sin procesar y aprender sobre ellas poco a poco en cada capa, construyendo un conocimiento de los datos de nivel superior.

Una de las características más distintivas del Aprendizaje Profundo, y una que lo hizo bastante popular y práctico, es que escala bien; es decir, cuantos más datos se le proporcionen, mejor funcionará, a diferencia de muchos algoritmos de aprendizaje automático más antiguos que tienen un límite superior para la cantidad de datos que pueden procesar el cual se conoce como la meseta de rendimiento y se puede observar en la Figura 2.3.

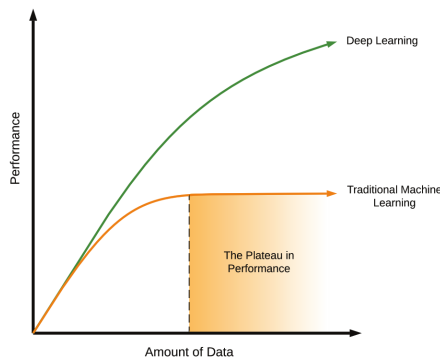


Figura 2.3: La meseta de rendimiento en el aprendizaje automático (Imagen extraída de Buduma and Locascio, 2017).

Entre las redes neuronales que han sido más empleadas en la actualidad se encuentran (Goodfellow et al., 2016):

- **Perceptrón Multicapa:** es un tipo de red neuronal el cual es una versión mejorada del perceptrón simple que consta de un número mayor de neuronas y capas ocultas de estas, obteniendo una capacidad mejorada de clasificación al eliminar la limitante de una clasificación lineal, además de la posibilidad de procesar una cantidad de datos mucho mayor con respecto al perceptrón simple.
- **Redes Neuronales Convolucionales:** son utilizadas principalmente por su enfoque en el procesamiento de imágenes, segmentación, extracción de características, reconocimiento de patrones gráficos, detección y clasificación de objetos.

- Redes Neuronales Recurrentes: son una derivación del perceptrón multicapa, con la diferencia de que las RNN son utilizadas para el análisis de datos que cambian con el tiempo.
- Autocodificadores: este tipo de red neuronal se utiliza para aprender una codificación de un conjunto de datos de entrada, la cual es entrenada para aprender la reconstrucción para intentar generar una salida similar a su entrada.

2.4.1 Ventajas del Aprendizaje Profundo

- Las características se van transformando automáticamente al paso sobre la red neuronal y ajustándose de forma óptima para el resultado deseado (Smebook, 2021).
- El mismo enfoque basado en redes neuronales tiende a emplearse a diversas aplicaciones y tipos de datos diferentes.
- La arquitectura de Aprendizaje Profundo es totalmente flexible, ya que pueden derivarse nuevos y mejorados enfoques con el paso del tiempo para adaptarse a nuevos retos computacionales en el futuro.

2.4.2 Limitaciones del Aprendizaje Profundo

- Las redes de aprendizaje profundo requieren una inmensa cantidad de costo computacional debido a la gran cantidad de parámetros en el modelo.
- No es fácil comprender la salida basada en el mero aprendizaje y requiere clasificadores para hacerlo, lo que hace más complejo el proceso de interpretación.
- Requiere una gran cantidad de datos para entrenar y evitar un sobreajuste.

2.5 Redes Neuronales Recurrentes

Las Redes Neuronales Recurrentes (RNN) fueron introducidas para el análisis de datos que cambian con el tiempo. Las tareas para las que se emplean regularmente son en el reconocimiento de voz, subtítulos de imágenes o la predicción caracteres para la generación de código (Vogt, 2019). Una idea arraigada junto al procesamiento de secuencias es que la salida obtenida de una neurona no solo debería depender de los datos dados en el presente, sino también de la información del pasado. Para permitir a la red memorizar y acceder a los historiales de entrada, se introducen las conexiones recurrentes para predecir el paso de tiempo actual y transferir esa predicción al siguiente paso como una entrada.

Como se observa en la Figura 2.4, el esquema de un modelo RNN tiene la misma estructura que los modelos de Redes Neuronales Artificiales, contando con una capa de entrada, n capas ocultas y una de salida, sin olvidar el parámetro t correspondiente al tiempo, siendo x_{t-1} , x_t y x_{t+1} las entradas del modelo RNN en diferentes instantes de tiempo.

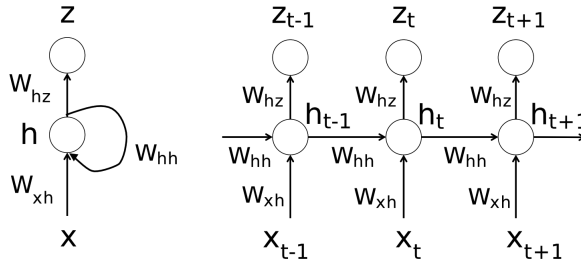


Figura 2.4: Estructura básica de una RNN (Vogt, 2019).

Para la forma más básica de una RNN, se encuentra su función de aprendizaje en la ecuación 2.2 para las capas ocultas y en la ecuación 2.3 para la capa de salida.

$$h[t] = f(W_i^h(x[t] + b_i) + W_h^h(h[t-1] + b_h)) \quad (2.2)$$

$$y[t] = g(W_h^o(h[t] + b_o)) \quad (2.3)$$

donde $f(\cdot)$ es la función de activación, generalmente implementada por una función sigmoideal o tangente hiperbólica. El estado oculto $\mathbf{h}[t]$ transmite el contenido de la memoria en el paso de tiempo t , normalmente se inicializa con un vector de ceros y depende de las entradas pasadas y los estados de la red. La salida $\mathbf{y}[t] \in \mathbb{R}^{N_o}$ se calcula con una transformación $g(\cdot)$, generalmente lineal, en la matriz de los pesos de salida $\mathbf{W}_h^o \in \mathbb{R}^{N_r \times N_o}$ es aplicada la suma del estado actual $\mathbf{h}[t]$ y el sesgo $\mathbf{b}_o \in \mathbb{R}^{N_o}$.

2.5.1 Propagación hacia atrás a través del tiempo

La propagación hacia atrás para las RNN es la misma aplicada para las redes neuronales de retroalimentación, con una diferencia en el gradiente, en cual, la suma se realiza con respecto a la pérdida en cada momento. Para calcular el descenso del gradiente basta con repetir iterativamente dos pasos hasta alcanzar la convergencia:

1. La función de pérdida L_k se evalúa en la RNN configurado con pesos W_k , cuándo un conjunto de entrada X_k se procesa. Considerando que W_k se refiere a los parámetros de red, mientras que el índice k identifica sus valores en la época k .
2. El gradiente $\frac{\partial L_k}{\partial W_k}$ se retropropaga a través de la red para actualizar sus parámetros.

La función de pérdida se define en la ecuación 2.4.

$$L_k = E(X_k, Y_k^*; W_k) + R_\lambda(W_k) \quad (2.4)$$

donde E es una función encargada de evaluar el error de predicción de la red cuando esta se alimenta con entradas en X_k , con respecto a una respuesta esperada Y_k^* . R_λ es una función de regularización que depende de un hiperparámetro λ , que pondera la contribución de la regularización en la pérdida total. Una de las funciones de error para E que se menciona en la literatura es el error cuadrático medio (MSE). El cual se define como:

$$MSE(\Upsilon_k, \Upsilon_k^*) = \frac{1}{|X_k|} \sum_{x \in X_k} (y_x - y_x^*)^2 \quad (2.5)$$

Donde $y_x \in Y_k$ es la salida de la red cuando se procesa la entrada $x_x \in X_k$ y $y_k^* \in \Upsilon_k^*$ es el valor de verdad fundamental que la red debe aprender a reproducir. El término de regularización R_λ introduce un bias que mejora las capacidades de generalización del modelo RNN, al reducir el sobreajuste en los datos de entrenamiento. Los parámetros de regularización mencionados en la literatura son (Bianchi et al., 2019):

- $L1$: $R_\lambda(W_k) = \lambda_1 \|W_k\|_1$
- $L1$: $R_\lambda(W_k) = \lambda_1 \|W_k\|_1$
- *Penalización de red elástica* : $R_\lambda(W_k) = \lambda_1 \|W_k\|_1 + \lambda_2 \|W_k\|_2$
- $\lambda_1 = 0$ y $\lambda_2 = 0$

2.5.2 Desvanecimiento y Explosión del Gradiente

Uno de los principales inconvenientes de las primeras arquitecturas RNN era su capacidad limitada de memoria causada por el problema de desvanecimiento o explosión del gradiente, que se hace notar cuando la información contenida en entradas pasadas debe recuperarse tras un intervalo de tiempo prolongado. Para ilustrar el problema de desaparición del gradiente, se puede considerar la influencia de la función de pérdida L_t que depende de las entradas de la red y de los parámetros W_t cuando su gradiente es retro propagado a través de la red como se muestra en la ecuación 2.6 (Vogt, 2019).

$$\frac{\partial L[t]}{\partial W} = \sum_{\tau} \frac{\partial L[t]}{\partial h[t]} \frac{\partial h[t]}{\partial h[\tau]} \frac{\partial h[\tau]}{\partial W} \quad (2.6)$$

En la ecuación anterior, las derivadas parciales de los estados con respecto a sus valores anteriores se pueden factorizar como:

$$\frac{\partial h[t]}{\partial h[\tau]} = \frac{\partial h[t]}{\partial h[t-1]} \cdots \frac{\partial h[t-1]}{\partial h[\tau]} = ft \dots ft + 1. \quad (2.7)$$

Para garantizar la estabilidad local, la red debe operar en un régimen ordenado, una propiedad garantizada por la condición $|f'_t| < 1$. Sin embargo, en este caso el producto en la ecuación 2.7 converge rápidamente de manera exponencial a 0, cuando $t - \tau$ aumenta. En consecuencia, la sumatoria de la ecuación 2.6 se ve dominada por los términos correspondientes a las dependencias a corto plazo y se produce el efecto del desvanecimiento del gradiente. Como efecto secundario principal, los pesos se actualizan cada vez menos a medida que fluye el gradiente hacia atrás a través de las capas de la red. Por otro parte, el fenómeno de la explosión del gradiente se manifiesta cuando $|f'_t| > 1$ y la red se vuelve localmente inestable.

2.6 Arquitecturas de Redes Neuronales Recurrentes

2.6.1 Vanilla Recurrent Neural Network

Vanilla RNN es el modelo clásico y la base de las redes recurrentes. Fue propuesto en (Elman, 1995) inicialmente para aplicaciones de Procesamiento del Lenguaje Natural. La estructura de esta red neuronal es la de una RNN como se aprecia en la Figura 2.4 y sus ecuaciones de las capas ocultas y de salida son descritas en las ecuaciones 2.2 y 2.3.

2.6.2 Long Short-Term Memory (LSTM)

Dado que el problema de “explosión del gradiente” puede obstaculizar gravemente el proceso de aprendizaje, se propuso LSTM para el control mejorado del flujo del gradiente. Por esta razón, la formulación de la celda LSTM se amplía con la definición de unidades de puerta y un estado de memoria. En la Figura 2.5 se observa la estructura de una unidad LSTM de este tipo, incluidas las puertas de entrada, salida, candidato y olvido, las cuales permiten el almacenamiento persistente de secuencias en memoria.

Para cada muestra X_t de la secuencia de entrada, la estructura de la célula LSTM se visualiza en la Figura 2.5 y las 5 puertas de una célula LSTM se expresan en las siguientes ecuaciones (Goodfellow et al., 2016):

$$\text{Input gate : } i_t = \sigma(W_{xi}x_t + b_{xi} + W_{hi}h_{t-1} + b_{hi}) \quad (2.8)$$

$$\text{Forget gate : } f_t = \sigma(W_{xf}x_t + b_{xf} + W_{hf}h_{t-1} + b_{hf}) \quad (2.9)$$

$$\text{Output gate : } o_t = \sigma(W_{xo}x_t + b_{xo} + W_{ho}h_{t-1} + b_{ho}) \quad (2.10)$$

$$\text{Candidate gate : } g_t = \sigma(W_{xg}x_t + b_{xg} + W_{hg}h_{t-1} + b_{hg}) \quad (2.11)$$

$$\text{Hidden state : } h_t = o_t \tanh(c_t) \quad (2.12)$$

donde h_{t-1} y c_{t-1} son el estado oculto y el estado de la celda del paso de tiempo anterior (o los estados iniciales en el paso del tiempo $t = 0$). Las puertas aplican la función sigmoïdal σ :

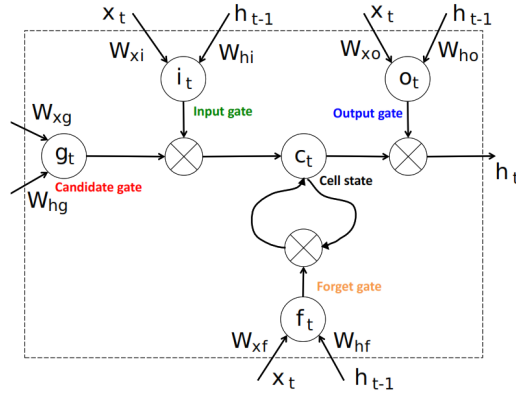


Figura 2.5: Estructura de una célula LSTM (Vogt, 2019).

$$\sigma(x) = \frac{1}{1 + \exp(-x)} \quad (2.13)$$

para calcular valores escalares en el rango de $[0, 1]$ permitiendo dejar pasar toda información por las puertas que corresponda a un valor de 1 o no dejar pasar ninguna información con un valor de 0.

Mientras que el estado de la celda c_t se considera una memoria a largo plazo, el estado oculto h_t más bien representa la memoria de trabajo que se enfoca en información inmediatamente útil de la memoria a largo plazo. Por lo tanto, se afirma que el estado oculto es una versión definida del estado de la celda (ya que resulta de una multiplicación del estado de la celda con la puerta de salida, lo que limita la información que se transmite). Durante el proceso de entrenamiento, el LSTM adapta los pesos W que están asociados con todas las puertas para saber qué información recordar, actualizar y prestar atención.

2.6.3 BiLSTM

Las redes bidireccionales de memorias a corto y largo plazo son beneficiosas en configuraciones donde la salida en un paso de tiempo dado no solo depende del pasado sino también de la información futura. La BiLSTM consiste básicamente en entrenar dos capas LSTM independientes, una LSTM “hacia adelante” que procesa la secuencia de entradas con una orden de tiempo regular $(0, \dots, N - 1)$ y una LSTM “hacia atrás” con un orden de tiempo invertido $(N - 1, \dots, 0)$. Para combinar la información obtenida de ambas direcciones, los estados ocultos hacia adelante y hacia atrás se concatenan en cada paso del tiempo (Vogt, 2019). La estructura de una BiLSTM se observa en la Figura 2.6.

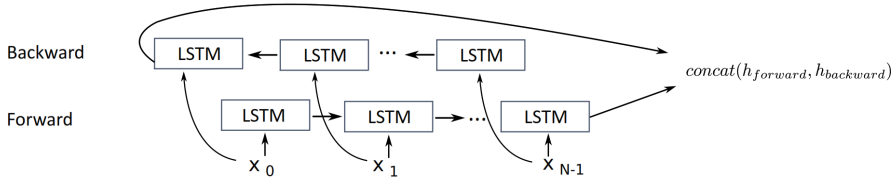


Figura 2.6: Estructura de una BiLSTM (Vogt, 2019).

2.6.4 Redes neuronales de unidades recurrentes cerradas

Las GRU son una variante LSTM simplificada que utiliza una formulación combinada de la celda y el estado oculto, así como una formulación combinada de la puerta de olvido y la puerta de entrada en una única puerta de actualización. Para cada elemento de la secuencia de entrada, la GRU calcula las siguientes funciones (Vogt, 2019):

$$\text{Reset gate} : r_t = \sigma(W_{ir}x_t + b_{ir} + W_{hr}h_{t-1} + b_{hr}) \quad (2.14)$$

$$\text{Update gate} : z_t = \sigma(W_{iz}x_t + b_{iz} + W_{hz}h_{t-1} + b_{hz}) \quad (2.15)$$

$$\text{Update gate} : z_t = \sigma(W_{iz}x_t + b_{iz} + W_{hz}h_{t-1} + b_{hz}) \quad (2.16)$$

$$\text{Hidden state} : h_t = (1 - z_t)n_t + z_th_{t-1} \quad (2.17)$$

2.7 Discusión

Este capítulo presentó un panorama general sobre los temas relacionados con el trabajo de investigación que se desarrolla en este tema de tesis, abarcando los conceptos básicos de la Fibrilación Auricular, Caracterización del Electrocardiograma, Espectrogramas y Redes Neuronales Recurrentes. La información presentada en este capítulo es requerida para el entendimiento fundamental de la transformación del dominio de la información, la extracción de características, reducción de dimensionalidad de los datos e implementación del modelo de clasificación que será útil en capítulos posteriores.

CAPÍTULO 3

Estado del Arte

En este capítulo se presentan brevemente los trabajos más relevantes en relación a la clasificación de la Fibrilación auricular, considerando técnicas de procesamiento de señales, modelos neuronales empleados, métricas y bases de datos utilizadas.

3.1 Antecedentes

Los trabajos mostrados a continuación fueron desarrollados en el Centro Nacional de Investigación y Desarrollo Tecnológico (CENIDET) y abordan temas relacionados con este trabajo de investigación.

Conducción Autónoma de un Vehículo Simulado mediante un Modelo de Red Neuronal Convolutiva Recurrente (Luna, 2020)

En este trabajo se propuso un modelo de Red Neuronal Híbrido compuesto por capas de Convulsión y Recurrentes para realizar el control de un vehículo simulado en el Sistema Operativo Robótico. El modelo propuesto deriva de un modelo secuencial llamado *Chauffeur* utilizado para evaluar la predicción correcta de la dirección de un vehículo en condiciones climáticas no favorables. El modelo propuesto cambia el bloque convolutivo del *Chauffeur* por un bloque de convulsión distribuida en el tiempo, que toma como entrada secuencias de imágenes en lugar de una sola.

Los resultados mostraron que el modelo propuesto aportó al vehículo una mejora media en la autonomía de 11.4% con respecto a los métodos de comparación en escenarios libres, y que el uso del canal de profundidad mejoró en 8.9% la autonomía en escenarios con obstáculos.

Segmentación no Paramétrica de Tejidos Cerebrales Mediante una Arquitectura Paralela de Redes Neuronales Convolutivas (Morales, 2018)

En este trabajo se desarrolló e implementó un método de segmentación no paramétrica de tejidos cerebrales, basada en una arquitectura paralela de Redes Neuronales Convolutivas (CNN). El modelo U-Net fue utilizado para la adaptación de una arquitectura paralela. Los repositorios de imágenes utilizados fueron BrainWeb para entrenamiento y BraTS 2017 para validación y pruebas. A los estudios de resonancia magnética utilizados para entrenamiento se les aplicó un preprocesamiento para

la supresión de tejidos no blandos, seguido de una transformación a formato TIF (Tagged ImageFile Format) para su uso en los framework Keras y TensorFlow.

La metodología obtuvo un buen rendimiento de segmentación con respecto a las métricas: Jaccard, Coeficiente de similitud Dice (Dice Similarity Coefficient) y Área bajo la curva (AUC). En promedio, los resultados obtenidos con el repositorio BrainWeb fueron de 0.9653 y en cuanto al repositorio BraTS 2017 se obtuvo un 0.9692.

Sintonización de una red totalmente conectada para segmentación de dos clases de objetos en imágenes (Suárez, 2018)

En este trabajo se realizó la sintonización de una Red Neuronal Convolutiva en el entorno de trabajo Caffe, para la segmentación de dos clases (persona y ave), en imágenes del repositorio BSDS500. Se utilizaron los modelos FCN-Alexnet y FCN-8s para segmentar imágenes a color sin ruido, pero con sombras, texturas y colores similares, iluminación no uniforme y oclusión. La sintonización de los hiperparámetros en Caffe fueron de 200 iteraciones, tasa de aprendizaje de $1e-12$ y 86 imágenes de entrenamiento. El rendimiento de la segmentación fue de un 3.43% con respecto al modelo FCN-Alexnet y un 1.23 % con respecto al modelo FCN-8s.

Detección de anomalías en mamografías utilizando la Red Neuronal Convolutiva AlexNet (Matuz, 2017)

En este trabajo se presentó una metodología para la detección de anomalías en mamografías, mediante una Red Neuronal Convolutiva. La metodología propuesta fue constituida en dos partes, en la primera se realizó el preprocesamiento de las imágenes y la segunda parte consistió en el diseño y caracterización de la red convolutiva, basada en la arquitectura AlexNet para la detección de anomalías en mamografías. Esta implementación fue probada con la base de datos MIAS (322 mamografías) y una muestra de DDSM con (250 imágenes), fue realizada una clasificación cambiando los parámetros de clasificación al modelo de red neuronal, formando cinco clasificadores de mamografías como sanas o con algún padecimiento. El modelo de clasificación con mejores resultados tenía los siguientes parámetros: imágenes a color con dimensión de píxeles, en formato PNG, con 50 épocas y una tasa de aprendizaje de 0.001 la cual obtuvo una exactitud de 92%, sensibilidad del 88% y especificidad del 96%.

3.2 Trabajos relacionados

La investigación relacionada con la clasificación de la Fibrilación Auricular ha crecido considerablemente durante los últimos años. Por ende, la literatura sobre este tema es extensa. En esta sección se revisaron los trabajos más relevantes encontrados en la literatura relacionados al tema de investigación, como implementación de Redes Neuronales Profundas que se aplican a las tareas de clasificación de la Fibrilación

Auricular, así mismo se mencionan trabajos relacionados a técnicas procesamiento de señales electrocardiográficas y métricas para cuantificar el rendimiento de los métodos propuestos y bases de datos utilizadas.

Enfoque de aprendizaje profundo para la detección del aleteo y fibrilación auricular altamente específicos basado en intervalos RR (Ivanovic et al., 2019)

En este trabajo se propuso una arquitectura de red neuronal profunda, que combinó redes neuronales convolucionales y recurrentes, para extraer características de la secuencia de intervalos RR (distancia entre dos ondas R sucesivas). La red de aprendizaje profundo propuesta constaba de tres partes diferentes. El primero utilizó una CNN para la extracción de características locales, mientras que el segundo utilizó una estructura de RNN llamada LSTM para codificar los patrones secuenciales extraídos por CNN. Finalmente, la salida de la capa LSTM se alimentó a la capa densa con la función de activación softmax para su clasificación. La arquitectura de red se detalla en la Figura 3.1.

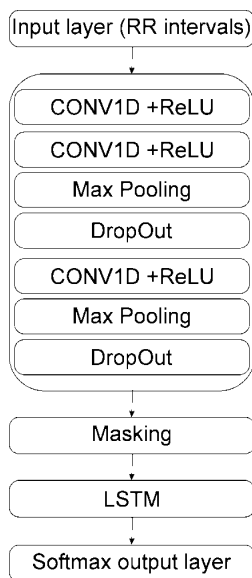


Figura 3.1: Arquitectura de red.

Los resultados obtenidos en la prueba de rendimiento del modelo propuesto fueron de 89.67% en Precisión, 94.20% en Sensibilidad y 93.13% en Especificidad en cuanto a la clasificación de fibrilación auricular se refiere. Estos resultados demostraron que se pudo distinguir con gran confianza la actividad cardíaca saludable de las relacionadas con las demás enfermedades mencionadas.

Detección automática de fibrilación auricular basada en CNN-lstm y conexión de acceso directo (Ping et al., 2020)

En este artículo, se propuso un modelo de aprendizaje profundo híbrido, el cual constó en la combinación de una CNN de 8 capas con una conexión de acceso directo y una LSTM de 1 capa, a este modelo híbrido se le conoce como 8CSL, para las tareas de clasificación de ECG. En comparación con las RNN y las redes neuronales de convolución de múltiples escalas (MCNN), 8CSL extrae características con habilidad, además de no ser susceptible a la dependencia a largo plazo entre los datos. En particular, 8CSL incluye ocho conexiones de acceso directo que pueden mejorar la velocidad de transmisión y procesamiento de datos como resultado de las conexiones de acceso directo. En la Figura 3.2, se puede apreciar la arquitectura general del modelo propuesto (8CSL).

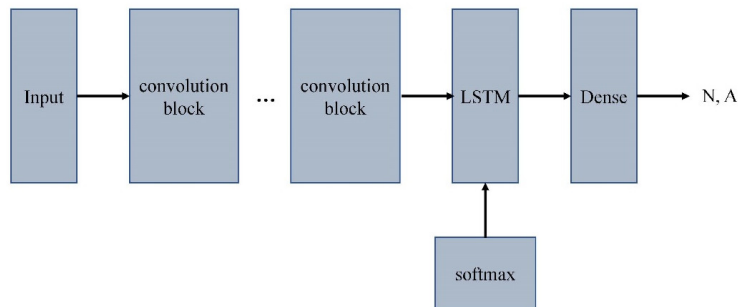


Figura 3.2: Arquitectura general de 8CSL.

En comparación con otros modelos, 8CSL tiene el mejor rendimiento en el equipo de prueba dado que 8CSL tiene un modelo de red más profundo y más formas de evitar un ajuste excesivo. Y en comparación con el mismo modelo, los tres modelos funcionan mejor en la longitud del segmento es de 10 s.

Una memoria a corto y largo plazo apilado para la predicción de la fibrilación auricular basada en ECG multivariados (Sun et al., 2020)

En este trabajo se desarrolló un modelo RNN compuesto de capas LSTM apiladas una sobre otra para la detección de fibrilación auricular. Comenzaron resaltando y a su vez justificando el uso de las RNN y sus variantes con respecto a los modelos CNN propuestos en los últimos años para la tarea de detección de FA dado que estos últimos son buenos en cuanto al reconocimiento de patrones en gráficos. Sin embargo, en el trabajo demostró que las CNN no pueden lidiar con las características de tiempo acompañado con la desaparición gradual de series temporales.

Posteriormente, mencionaron el modelo de capas apiladas de LSTM debido a que una RNN convencional a pesar de su eficacia en el modelado de las series de tiempo,

tiene el problema de desaparición del gradiente y el descenso del gradiente cuando se realiza la propagación hacia atrás. Aunado a los problemas antes mencionados, el hacer uso de una red LSTM tradicional presenta una limitación en la extracción de características. La estructura detallada del modelo propuesto se puede apreciar en la Figura 3.3.

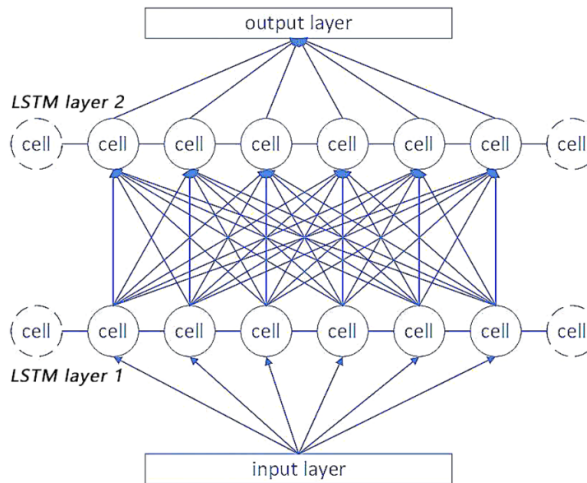


Figura 3.3: Estructura del modelo propuesto.

En los resultados obtenidos, se destacó que el modelo propuesto obtuvo mejores resultados en comparación con la red LSTM de una sola capa y la RNN convencional, justificando las deficiencias de las demás arquitecturas que mencionan al comienzo del artículo. Los porcentajes son de 92% para la precisión, sensibilidad y el puntaje F1 respectivamente. También realizaron el análisis de rendimiento del modelo propuesto con diferente número de capas, obteniendo el mayor rendimiento con 2 capas apiladas y con esto demostraron que un número pequeño de capas, el modelo carece de diversidad y cuando el número de capas es grande, ocasiona un fenómeno de ajuste insuficiente en los pesos. Por lo tanto, para evitar el ajuste insuficiente sugirieron 2 enfoques: encontrar el número óptimo de capas en proporción al tamaño de los datos o aumentar el tamaño de los datos y los tiempos de iteración.

Detección automatizada de fibrilación auricular utilizando una red híbrida CNN-LSTM en conjuntos de datos de ECG desequilibrados (Petmezas et al., 2021)

En esta publicación se propuso un modelo de red neuronal híbrido que combina CNN y LSTM para clasificar diferentes tipos de arritmias cardiovasculares. La arquitectura de la red se muestra en la Figura 3.4.

Lo interesante de este artículo fue el preprocesamiento que le fue aplicado a las señales de ECG, ya que implementaron un filtro de paso alto de Butterworth de

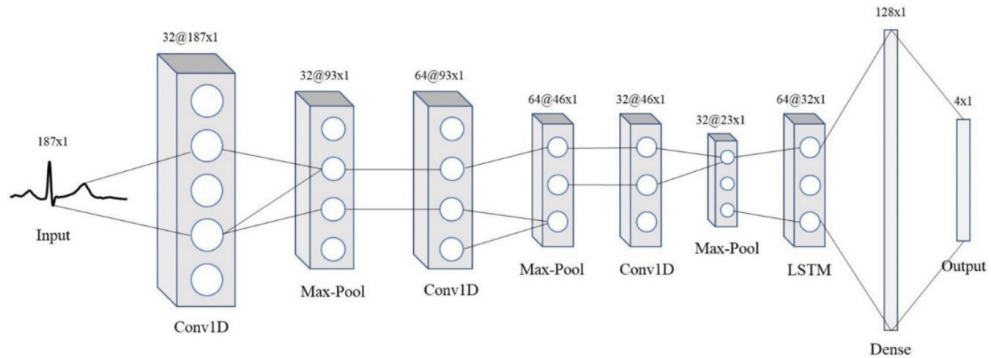


Figura 3.4: Arquitectura del modelo híbrido propuesto.

séptimo orden con una frecuencia de corte de 0.5 Hz para la eliminación del ruido de fluctuación de fase de la línea base, además de un filtro de paso bajo de sexto orden para eliminar la interferencia eléctrica y otros componentes de alta frecuencia del ruido. Para el ruido blanco gaussiano, implementaron la transformada discreta de ondículas de tres niveles usando una ondícula de la familia Daubechies y un umbral suave.

Posteriormente al preprocesamiento, las señales fueron sometidas a la red neuronal en donde la tarea principal de la red neuronal convolucional fue extraer características, mismas que fueron suministradas como entrada para la parte de la LSTM, generando características más discriminativas y logrando paralelamente una reducción de dimensionalidad. El modelo fue entrenado con la base de datos de Fibrilación Auricular del MIT-BIH, obteniendo una sensibilidad del 97.87% y una especificidad del 99.29% a través de una validación cruzada de diez veces.

MultiFusionNet: Detección de fibrilación auricular con redes neuronales profundas (Tran et al., 2020)

En este trabajo se propuso MultiFusionNet, una red de aprendizaje profundo que utilizó un método de fusión multiplicativo enfocado en la combinación de dos Redes Neuronales Profundas entrenadas con diferentes bases de conocimiento, es decir, los datos sin procesar y las características extraídas relevantes. Este enfoque que propusieron los autores, utilizó las características hechas manualmente para mejorar el modelo al momento de analizar datos sin procesar. La estructura de la red propuesta se muestra en la Figura 3.5.

El conjunto de características extraídas pasó a través de una subred totalmente conectada, ya que esa subred al tener más capas se aproximó a una función más complicada de los datos de entrada. Por otra parte, los datos sin procesar pasaron a través de otra subred especialmente para datos sin procesar para extraer características latentes y capturar la dependencia temporal. Posteriormente, las salidas de ambas subredes se

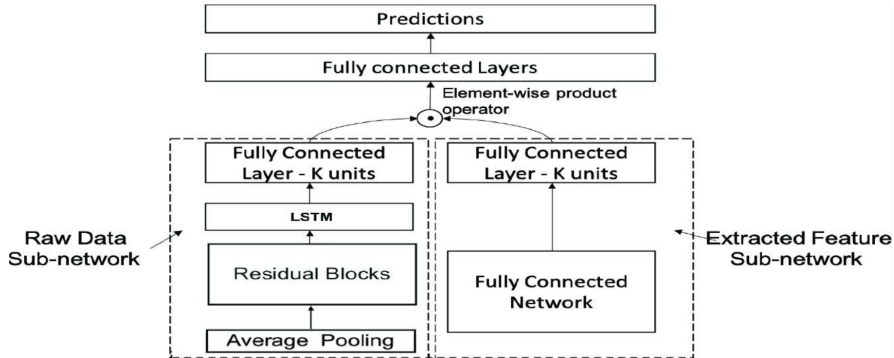


Figura 3.5: Arquitectura de MultiFusionNet.

fusionaron mediante un operador de fusión multiplicativo, obteniendo como salida un vector de acceso rápido $y = [y_0, y_1, y_2, y_3]$ con $y_i = 1$ en relación a la etiqueta predicha $l = L_i$, $L = [RitmoNormal, FibrilacinAuricular, Otros, SealRuidosa]$. Los resultados obtenidos durante el entrenamiento con 6000 señales de ECG fueron de 80% de F-1 Score y .85 en la curva característica de operación y el area sobre la curva respectivamente.

Detección automatizada de fibrilación auricular utilizando una red de memoria a corto plazo con señales de intervalo RR (Faust et al., 2018)

En este trabajo se propuso un sistema de aprendizaje profundo para detectar latidos con fibrilación auricular en señales de frecuencia cardiaca. Los experimentos se realizaron con la base de datos de fibrilación auricular del MIT-BIH. Estas grabaciones también contienen anotaciones de pulso y anotaciones de ritmo realizadas manualmente por médicos expertos. Además, los picos R están etiquetados y la secuencia del intervalo RR se extrajo basándose en estas etiquetas. Las secuencias de intervalos RR se han dividido en secuencias superpuestas de 100 latidos. 2 para cada trazo de FC.

Utilizaron la agrupación máxima global en una dimensión entre las capas LSTM bidireccionales y las capas completamente conectadas para comprimir las características de las secuencias de salida producidas por las capas LSTM bidireccionales. Efectivamente, las capas LSTM bidireccionales actuaron para aprender y extraer las características de las secuencias de datos de FC de entrada, antes de pasar estas características al modelo superior completamente conectado para clasificar si hay signos de FA presentes o no. El modelo propuesto en este artículo se implementó utilizando Keras y Tensor flow. La arquitectura del modelo de aprendizaje propuesto se puede apreciar en la Figura 3.6.

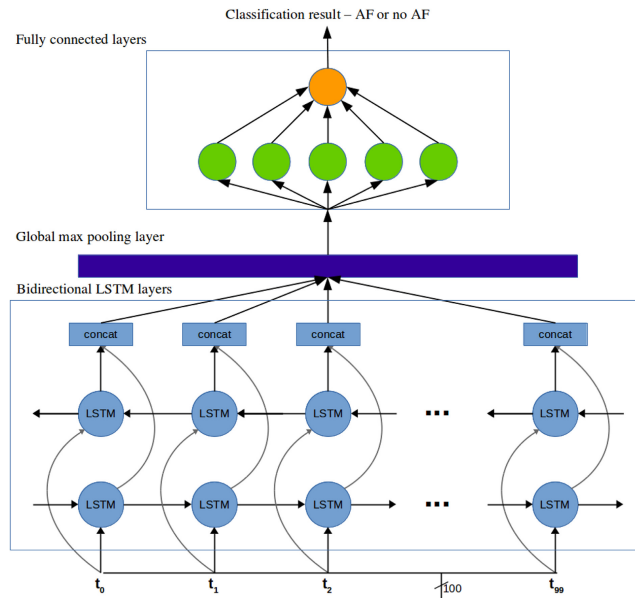


Figura 3.6: Arquitectura del modelo de aprendizaje propuesto.

Clasificación de la fibrilación auricular basada en redes MLP extrayendo los parámetros de Jitter y Shimmer (Borghi et al., 2021)

En este trabajo, se propuso un sistema basado en la red neuronal artificial MLP para diferenciar entre Fibrilación Auricular y no Fibrilación Auricular mediante la segmentación de características del ECG. Utilizando la base de datos de fibrilación auricular PhysioNet del MIT-BIH, se realizaron experimentos de reconocimiento de Fibrilación Auricular basados en la segmentación y extracción de características de la señal de ECG en intervalos de 30 picos R consecutivos.

El modelo de aprendizaje, red neuronal artificial MLP, se sometió a algunos conjuntos de entrada compuestos por cuatro combinaciones de características y demostró un ligero aumento en la precisión a medida que se introdujeron más características. El mejor resultado (80,67% de precisión en el conjunto de prueba) se logró mediante la arquitectura con 3 capas ocultas, 165 neuronas en cada capa, función de activación sigmoidea tangente, uso de todas las características en el conjunto de entrada y considerando segmentos de AF cuando el 10% de los intervalos RR están etiquetados como AF.

Detección de fibrilación auricular a partir de señales de ECG cortas mediante un modelo híbrido de aprendizaje profundo (Wu et al., 2019)

En este artículo se propuso un modelo híbrido CNN-LSTM que utiliza la señal de ECG corta del conjunto de datos PhysioNet / CinC Challenges 2017 para explorar y evaluar el rendimiento relativo de cuatro algoritmos de minería de datos y tres arquitecturas de aprendizaje profundo, CNN, LSTM y CNN-LSTM. Los resultados mostraron que todas las arquitecturas de aprendizaje profundo, excepto LSTM, se desempeñaron mucho mejor que los algoritmos de aprendizaje automático sin necesidad de una extracción de características complicada.

CNN-LSTM tuvo el mejor desempeño, logrando 97.08% de precisión, 95.52% de sensibilidad, 98.57% de especificidad, 98.46% de precisión, 0.99 AUC (Área bajo la curva ROC) valor y 0.97 puntuación F1. Con un diseño adecuado de la configuración, el aprendizaje profundo puede ser efectivo para la detección automática de AF, mientras que los métodos de minería de datos requieren conocimiento del dominio y un extenso proceso de extracción y selección de características para obtener resultados satisfactorios.

Diseño y desarrollo del modelo LSTM-RNN para la predicción de intervalos RR en señales de ECG (Dhananjay et al., 2021)

El trabajo propuesto tuvo como objetivo predecir el intervalo RR (RRI) durante 10 segundos en señales de ECG de ritmo sinusal normal (NSR) mediante el diseño y desarrollo de un modelo basado en una LSTM-RNN. La señal de ECG considerada en este trabajo fue la base de datos del hospital de la Universidad de Chapman y Shaoxing People's que tiene señales de ECG estándar de 12 derivaciones. La base de datos tiene 10646 ECG's de diferentes pacientes y todas las señales muestreadas a 500 Hz. La duración del registro de ECG disponible para cada paciente en la base de datos es de 10 segundos. El foco principal fue la señal de ECG de derivación II para el trabajo propuesto. La Figura 3.7 representa la entrada de un ECG de voluntarios de ritmo sinusal normal (NSR) descargada de la base de datos.

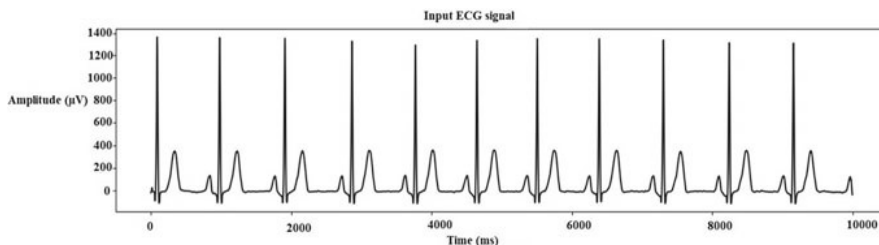


Figura 3.7: Señal de ECG de un voluntario.

La señal de ECG descargada se sometió a una técnica de procesamiento de señales para eliminar el ruido mediante el desarrollo de una ventana de 10 puntos. Después

de promediar la señal, se realizó la detección del pico R en la señal de ECG para calcular el RRI. La representación de señales de ECG de base de datos mostradas en amplitud y dominio de tiempo en microvoltios (μV) y milisegundos (ms).

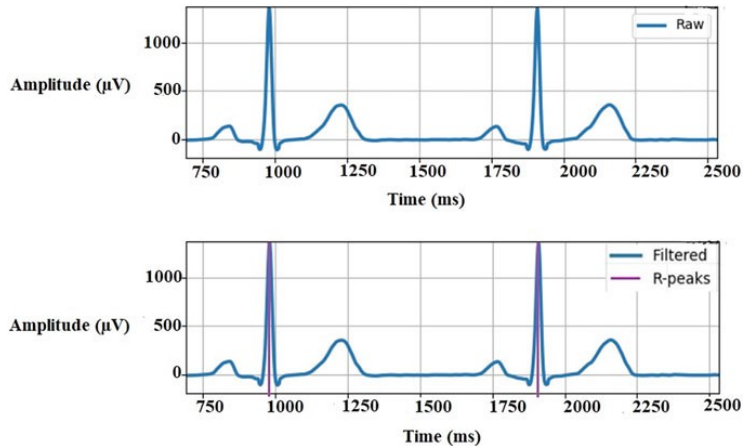


Figura 3.8: (a) Señal en crudo. (b) Señal con picos R filtrados.

Detección de fibrilación auricular a partir de datos de fotopletimografía de muñeca mediante redes neuronales artificiales (Yousefi et al., 2018)

El objetivo de este artículo fue evaluar un método de detección de FA basado en Redes Neuronales Artificiales a partir de datos de intervalo latido a latido derivados de PPG utilizados para fines de detección o monitorización primaria.

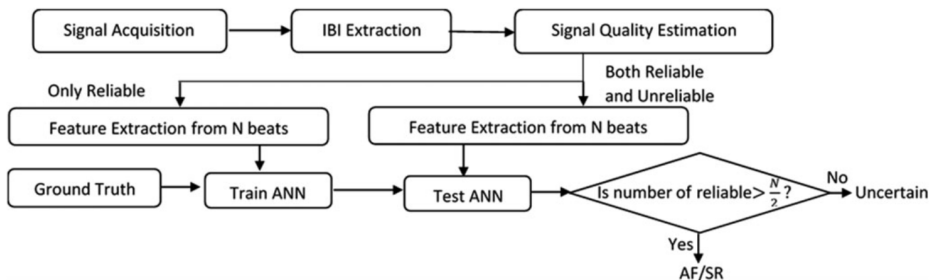


Figura 3.9: Diagrama de bloques del método propuesto.

El clasificador propuesto fue capaz de distinguir entre FA y ritmos sinusales (SR). En total, 30 pacientes (15 con FA, 15 con RS, edad media 71,5 años) con múltiples comorbilidades fueron monitorizados durante el tratamiento postoperatorio de rutina. La monitorización incluyó un ECG estándar y un monitor PPG de pulsera con fuentes de luz verde e infrarroja. Las características de entrada de la ANN se basaron en la información obtenida de las secuencias de intervalo entre latidos (IBI) de 30 pulsos PPG consecutivos.

Por lo tanto, en el método propuesto, la confiabilidad del IBI se evaluó automáticamente considerando los siguientes aspectos: la cantidad de incertidumbre debida a latidos no confiables fue del 15,42%. La sensibilidad y especificidad alcanzadas de detección de FA para secuencias de 30 latidos fueron $99.20 \pm 1.3\%$ y $99.54 \pm 0.64\%$, respectivamente. Con base en estos resultados, el algoritmo ANN demostró un rendimiento excelente en el reconocimiento de AF de SR utilizando datos PPG de muñeca.

Un dispositivo de ECG móvil de una derivación incorporado en un programa de monitoreo remoto de arritmias basado en síntomas (Selder et al., 2019)

El objetivo de este estudio fue evaluar el beneficio del dispositivo AliveCor Kardia Mobile (KM) y su algoritmo en la detección de arritmias cardíacas en una cohorte del mundo real de pacientes ambulatorios. Se incluyeron todos los ECG registrados en el programa Hartwatch Arrhythmia (HA) entre enero de 2017 y marzo de 2018. La clasificación por el algoritmo KM se comparó con la del equipo de Hartwacht dirigido por un cardiólogo.

Se recibieron 5.982 KM ECG de 233 pacientes. El algoritmo KM categorizó el 59% como SR, el 22% como posible FA, el 17% como no clasificado y el 2% como ilegible. Según el equipo de Hartwacht, 498 (8%) ECG no eran interpretables. El valor predictivo negativo para la detección de FA fue del 98%. Sin embargo, el valor predictivo positivo, así como la detección de otras arritmias, fue deficiente. En el 81% de los ECG no clasificados, el equipo de Hartwacht pudo proporcionar un diagnóstico. Menos del 10% de los ECG fueron ininterpretables. Sin embargo, el rendimiento actual del algoritmo KM hizo que el dispositivo fuese inadecuado como aplicación independiente, lo que respaldó la necesidad de un análisis de ECG manual en HA y programas similares.

Red neuronal profunda de fusión multiescala para la detección de arritmias multiclase (Wang et al., 2020)

En este artículo, se formuló una nueva arquitectura de red neuronal convolucional de fusión profunda de múltiples escalas (DMSFNet) de extremo a extremo para la detección de arritmias. El enfoque propuesto pudo capturar de manera efectiva patrones anormales de enfermedades y suprimir la interferencia de ruido mediante la extracción de características de múltiples escalas y la complementariedad de información de escalas de las señales. La arquitectura se visualiza en la Figura 3.10.

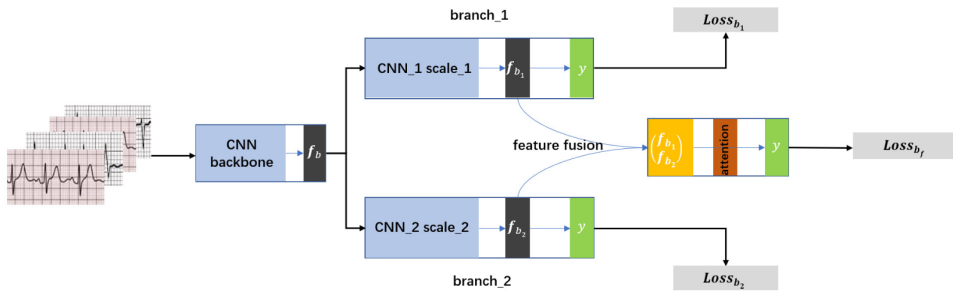


Figura 3.10: Descripción general de la arquitectura del modelo propuesta.

En este trabajo, se probó DMSFNet en dos conjuntos de datos abiertos (CPSC 2018 y PhysioNet / CinC 2017) y generó el rendimiento de vanguardia en ellos. Entre ellos, CPSC 2018 es un conjunto de datos de ECG de 12 derivaciones y CinC 2017 es un conjunto de datos de una sola derivación. Para estos dos conjuntos de datos, lograron la puntuación F1 de 82,8% y 84,1% respectivamente. Los resultados demostraron que DMSFNet de extremo a extremo tuvo un rendimiento sobresaliente para la extracción de características de una amplia gama de arritmias distintas y una elegante capacidad de generalización para manejar eficazmente las señales de ECG con diferentes derivaciones.

ECGDeepNET: un enfoque de aprendizaje profundo para clasificar los latidos del ECG (Mahmud et al., 2019)

En este artículo, se propuso un método para detectar anomalías en los latidos del corazón de manera eficiente. La estructura propuesta fue bastante liviana y requirió menor capacidad de cálculo y memoria. Además, para reducir el desequilibrio de clases y aumentar la precisión, realizaron un procesamiento en los datos y realizaron un aumento de información a las clases con números más bajos con 6 operaciones diferentes. Las cuales se pueden observar en la Figura 3.11.

Para la clasificación de arritmias, lograron una precisión promedio de 97.3%, 98.9% con puntaje F1 de 97.21%, 99.2% y especificidad de 99.3%, 98.95% para la base de datos de arritmias MIT BIH y la base de datos de ECG de diagnóstico PTB respectivamente.

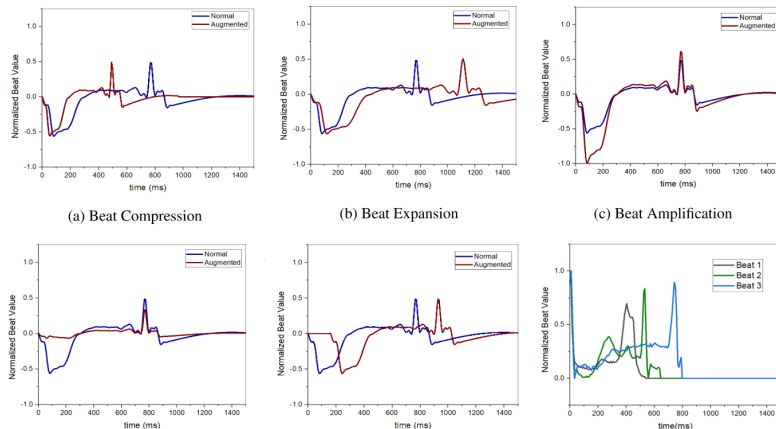


Figura 3.11: Diferentes técnicas de aumento realizadas en el conjunto de datos.

Detección de arritmias cardíacas a partir de ECG con redes neuronales convolucionales recurrentes (Zaen et al., 2020)

En este trabajo se propusieron tres arquitecturas de redes neuronales para detectar ritmos anormales a partir de señales de ECG de una sola derivación. Estas arquitecturas combinaron capas convolucionales para extraer características de alto nivel pertinentes para la detección de arritmias de ventanas deslizantes y capas recurrentes para agregar estas características sobre señales de duración variable.

Se usaron las redes neuronales en el conjunto de datos utilizado para el desafío de la Computación en Cardiología 2017 y un conjunto de datos construido al unir tres bases de datos disponibles en PhysioNet. El primer conjunto incluye 8528 señales de ECG de una sola derivación registradas con un dispositivo AlivCor. El segundo conjunto como se mencionó anteriormente, se creó mediante la combinación de tres bases de datos de PhysioNet: la base de datos de fibrilación auricular MIT-BIH, la base de datos de arritmias MIT-BIH y la base de datos de fibrilación auricular a largo plazo.

La base de datos de fibrilación auricular a largo plazo incluye 84 registros de ECG de dos derivaciones muestreados a 128 Hz. Estos registros se obtuvieron de sujetos con fibrilación auricular paroxística o sostenida y sus duraciones variaron, pero típicamente fueron entre 24 y 25 horas. Las arquitecturas lograron una precisión del 86,23% en el primer conjunto de datos, similar a las entradas ganadoras del desafío, y una Precisión del 92,02% en el segundo conjunto de datos.

Procesamiento del lenguaje de ECG (ELP): una nueva técnica para analizar señales de ECG (Mousavi et al., 2021)

En este trabajo, el objetivo fue proponer una técnica de análisis de ECG novedosa, el procesamiento del lenguaje de ECG (PLECG), que se centró en capacitar a las computadoras para que comprendan las señales de ECG. Se evaluó el método propuesto en dos tareas, incluida la clasificación de los latidos cardíacos y la detección de fibrilación auricular en las señales del ECG.

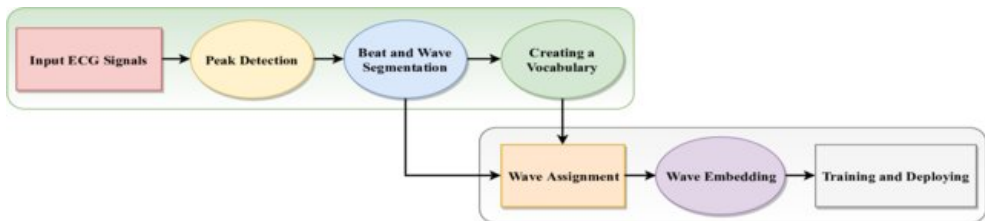


Figura 3.12: Procesamiento de lenguaje de ECG.

Evaluaron el método propuesto en dos tareas, incluida la clasificación de los latidos cardíacos y la detección de fibrilación auricular en las señales del ECG. Los resultados experimentales en tres bases de datos (es decir, las bases de datos MIT-BIH, MIT-BIH AFIB y PhysioNet Challenge 2017 AFIB Dataset de PhysioNet) demostraron que el método propuesto es una idea general que se puede aplicar a una variedad de aplicaciones biomédicas y es capaz de lograr un rendimiento notable.

Modelo híbrido de aprendizaje profundo CNN-LSTM y técnica de conjunto para la detección automática de infarto de miocardio utilizando grandes datos de ECG (Rai and Chatterjee, 2021)

En este trabajo se propuso un sistema de detección automatizado de MI que utilizó señales de electrocardiograma (ECG) mediante una red neuronal convolucional (CNN), CNN híbrida: red de memoria a largo plazo a corto plazo (LSTM) y una técnica de conjunto para elegir el modelo de rendimiento óptimo. Se utilizaron 123,998 latidos de ECG obtenidos de la Base de datos de diagnóstico de PTB (PTBDB) y la Base de datos de arritmias MIT-BIH (MITDB) para desarrollar el modelo.

El experimento se realizó en dos etapas: (i) usando conjuntos de datos originales y no balanceados y (ii) usando un conjunto de datos balanceados, obtenidos a partir de la técnica de muestreo de datos de minorías sintéticas (SMOTE). Se obtuvo la máxima precisión de clasificación de 99,82%, 99,88% y 99,89% utilizando CNN, CNN-LSTM híbrido y técnicas de conjunto, respectivamente.

Enfocándose en las métricas para cuantificar el rendimiento del modelo, en este trabajo se utilizó la matriz de confusión (CM), la precisión y la curva del historial de pérdidas y hemos calculado las diversas métricas de evaluación del rendimiento para verificar el resultado del rendimiento del modelo. Un CM es una tabla que resume

el resultado de cualquier algoritmo de clasificación por cualquier clasificador, es un gráfico de clasificación real Vs clasificación predicha.

Usar solo la precisión de la clasificación para autenticar el modelo del clasificador puede ser engañoso o puede no dar la evaluación correcta, especialmente en el caso de un número desigual de observaciones. Por lo tanto, el CM no solo pudo proporcionar una mejor idea del rendimiento de cualquier clasificador, sino que también proporcionó los muchos tipos de métricas de evaluación de rendimiento y parámetros. De la tabla CM, se pudo obtener fácilmente la muestra clasificada con precisión y clasificada incorrectamente.

En esta investigación se utilizó *Recall*, *Especificity*, *Precision*, *Accuracy*, *F1 – Score* y *Classificationerrorrate* y el Índice de Youden para validar el rendimiento del clasificador los cuales se describen a continuación.

$$Precision = \frac{TP}{TP + FP} \times 100 \quad (3.1)$$

$$Recall = \frac{TP}{TP + FN} \times 100 \quad (3.2)$$

$$Especificity = \frac{TN}{TN + FP} \times 100 \quad (3.3)$$

$$Classificationerrorrate = \frac{FP + FN}{TN + TP + FN + FP} \times 100 \quad (3.4)$$

$$Accuracy = \frac{TN + TP}{TN + TP + FN + FP} \times 100 \quad (3.5)$$

$$F1 - Score = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (3.6)$$

$$IndicedeYouden = (Recall + Especificity) - 100 \quad (3.7)$$

Detección de fibrilación auricular utilizando una descomposición multiescala mejorada en una Red neuronal convolucional residual mejorada (Cao et al., 2019)

En esta investigación, se propuso una red neuronal convolucional residual mejorada de descomposición de múltiples escalas. Se aplicó a un gran conjunto de datos de ECG de una sola derivación proporcionado por PhysioNet / CinC Challenge 2017. En el método propuesto, segmentó el registro de ECG original en muestras cortas de 9 segundos. Luego, utilizando la descomposición de la trama de ondícula derivada, las muestras cortas segmentadas se descompusieron y reconstituyeron en muestras de subseñal de diferentes escalas. Se entrenaron redes neuronales convolucionales

residuales de muestreo descendente rápido (FDResNets) con el conjunto de datos de señal corta original y el conjunto de datos reconstruido de cada escala.

La técnica de aprendizaje por transferencia se aplicó para acoplar las tres FDResNets con buen rendimiento en una red neuronal convolucional residual mejorada de descomposición de múltiples escalas (MSResNet). Ambas redes se muestran en la Figura 3.13.

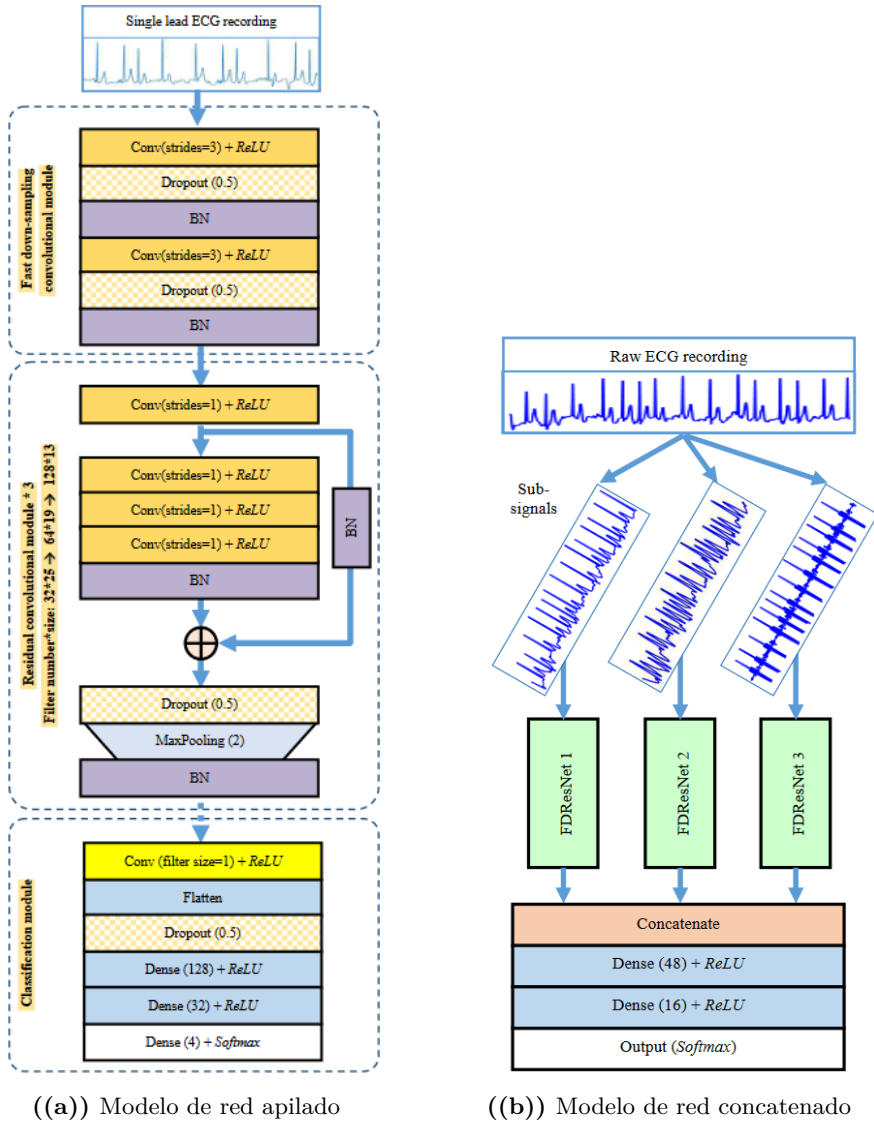


Figura 3.13: Modelos usados en el trabajo.

El método propuesto se validó en el conjunto de datos públicos de ECG de derivación única corta del PhysioNet / CinC Challenge 2017 con una precisión de prueba del 92,1% y una puntuación de F1 del 89,9%. Los principales hallazgos de este estudio se pueden resumir de la siguiente manera: La red neuronal convolucional residual profunda unidimensional pudo aprender características de clasificación efectivas a partir de la forma de onda del ECG en el dominio del tiempo. La convolución de paso amplio pudo mejorar la estabilidad del modelo al tiempo que mejora en gran medida la velocidad de aprendizaje del modelo. La ruta de cortocircuito residual fue fundamental para la capacidad de aprendizaje del modelo. En el dominio de la frecuencia, la información de características de la señal de ECG se concentró principalmente en la banda de baja frecuencia. Con la trama de paquetes de ondas derivadas, la señal de ECG se pudo descomponer en sub-señales de diferentes escalas con invariancia de traducción.

Aplicación de redes neuronales densas para la detección de fibrilación auricular y clasificación de conjunto de características de ECG aumentado (Kraševa et al., 2021)

En esta investigación, se optimizó un clasificador de RNA y se realizó la clasificación de 137 características importantes de ECG mejoradas que se calcularon a partir de representaciones de señales de ECG de tiempo y frecuencia de tiras cortas de una sola derivación disponibles en la base de datos Physionet / CinC Challenge 2017. Este enfoque se basó en la búsqueda de cuadrícula de hiperparámetros de capas NN densamente conectadas.

La reducción de características estuvo vinculada a la aplicación efectiva de un método integral para el cálculo del mapa de características basado en los pesos de las neuronas activadas a través de la ruta total desde la entrada hasta la salida específica en DenseNet.

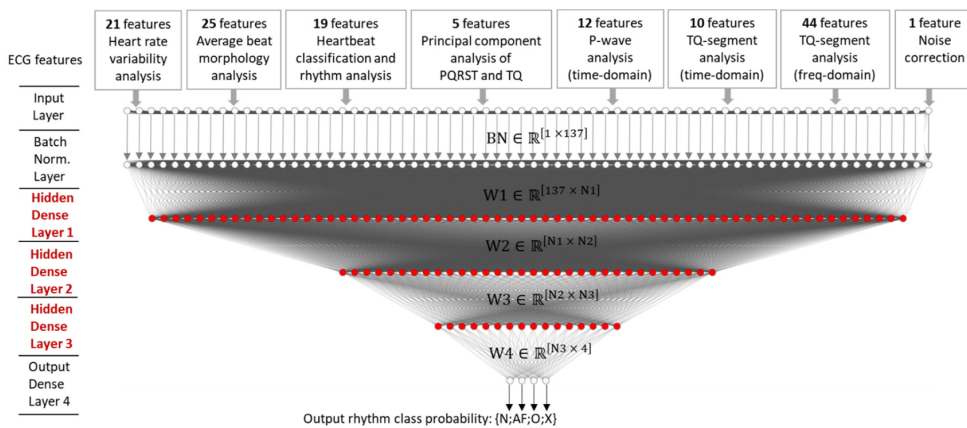


Figura 3.14: Arquitectura general DenseNet.

Las importantes aportaciones de este estudio se resumieron en dos aspectos: la optimización y la interpretación del proceso de toma de decisiones en redes neuronales artificiales, que comúnmente se tratan como clasificadores de caja negra. En primer lugar, el enfoque sistemático de búsqueda de cuadrícula de un clasificador de red neuronal denso llevó a obtener la configuración óptima que funciona bien con características de diagnóstico de 32 a 137 ECG, siendo capaz de clasificar el ritmo incluso con conjuntos muy limitados.

Aprendizaje profundo para la detección del tipo de arritmia de latido de ECG de una sola derivación mediante una representación de espectrograma de iris (Alqudah and Alqudah, 2021)

Este artículo presentó una metodología para detectar hasta 17 clases de arritmias cardíacas basadas en el análisis de señales de electrocardiografía (ECG) por latidos utilizando un espectrograma de iris. El análisis automático de cada latido del ECG permitió detectar anomalías. El objetivo principal de este artículo fue desarrollar un aprendizaje profundo rápido y un enfoque eficiente para clasificar las arritmias cardíacas. Se basó en analizar un solo latido del ECG y calcular el espectrograma del iris. Luego, el espectrograma del iris se alimentó a una red neuronal convolucional. El método propuesto se puede apreciar en la siguiente Figura.

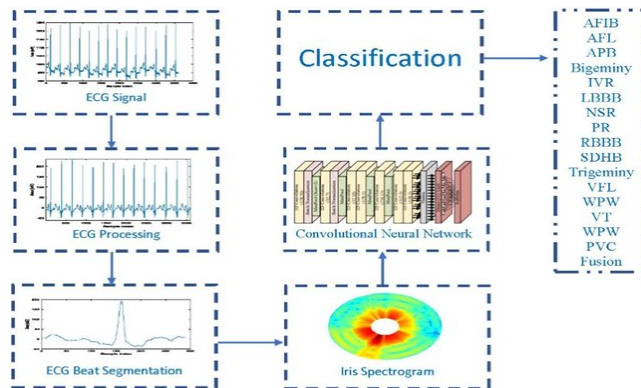


Figura 3.15: El diagrama de bloques del método propuesto para la clasificación de señales de ECG.

El método propuesto demostró ser eficiente, simple y rápido, lo que lo hace factible para la clasificación en tiempo real. Los resultados mostraron que la metodología propuesta tuvo una precisión de reconocimiento general de $99.13\% \pm 0.25$, $98.223\% \pm 0.85$ y $97.494\% \pm 1.26$ para 13, 15 y 17 clases de arritmias, respectivamente. El entrenamiento/prueba se realizó utilizando una validación cruzada diez veces mayor en comparación con los estudios existentes, demostrando ser prometedor y superando a otros enfoques.

Reconocimiento de episodios de AF mediante funciones optimizadas de frecuencia de tiempo y SVM sensible al costo (Hamil et al., 2021)

En este trabajo se presentó y validó experimentalmente un nuevo enfoque eficiente para la detección y clasificación de Fibrilación Auricular utilizando múltiples señales de electrocardiograma (ECG). Este trabajo consistió en aplicar la transformada de Stockwell (ST) con kernel de soporte compacto (ST-CSK) para el análisis de tiempo-frecuencia de ECG. La matriz de características se empleó como entrada de las máquinas de vectores de soporte (SVM) que trabajan en modo binario y asimétrico con una opción de rechazo incorporada.

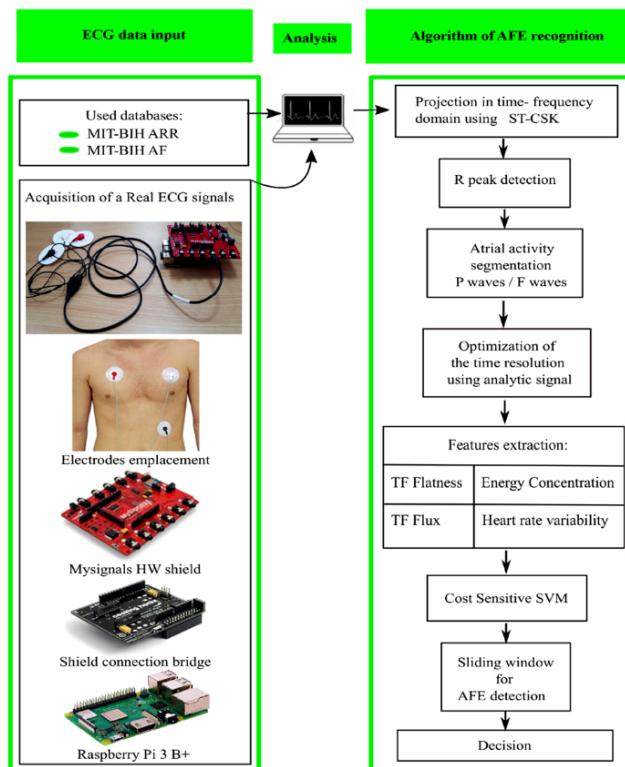


Figura 3.16: Arquitectura para la detección y clasificación de Fibrilación Auricular.

El algoritmo propuesto se entrenó y probó utilizando diferentes fuentes de ECG, en específico, dos bases de datos proporcionadas por PhysionNet (arritmia MIT-BIH, Fibrilación Auricular MIT-BIH). El método utilizado alcanzó el 98,46% y el 97,81% de Recall y Especificity, respectivamente.

3.3 Discusión

Los trabajos que se mencionan en el estado del arte contemplan varios enfoques novedosos referentes a la clasificación de la fibrilación auricular que son relevantes para el presente trabajo. A continuación, se muestra en la Tabla 3.1 información resumida de estos trabajos considerando el objetivo, algoritmos/técnicas y la utilidad para esta investigación.

Tabla 3.1: Resumen del Estado del Arte

Trabajo	Objetivo	Algoritmos/Técnicas	Utilidad
Enfoque de aprendizaje profundo para la detección del aleteo y fibrilación auricular altamente específicos basado en intervalos RR (Ivanovic et al., 2019).	Detectar el aleteo y la Fibrilación Auricular usando intervalos RR de las señales electrocardiográficas.	CNN-LSTM	Modelo utilizado (LSTM) y métricas empleadas (Precisión, Sensibilidad y Especificidad).
Detección automática de fibrilación auricular basada en CNN-lstm y conexión de acceso directo (Ping et al., 2020).	Detección de Fibrilación Auricular usando una red híbrida.	Red Neuronal Híbrida 8CSL	Comparación con otros modelos y métricas para la evaluación del rendimiento.
Una memoria a corto y largo plazo apilado para la predicción de la fibrilación auricular basada en ECG multivariados (Sun et al., 2020).	Clasificar la Fibrilación Auricular usando una red neuronal con capas apiladas de LSTM.	2 Layer-LSTM	Modelo utilizado (LSTM) y comparación de resultados con diferente número de capas apiladas.
Detección automatizada de fibrilación auricular utilizando una red híbrida CNN-LSTM en conjuntos de datos de ECG desequilibrados (Petmezas et al., 2021).	Detectar automáticamente la Fibrilación Auricular usando bases de datos con clases desbalanceadas.	Modelo empleado: CNN-LSTM. Técnicas de preprocesamiento: Filtro de paso alto de Butterworth, Transformada discreta de ondículas.	Técnicas de preprocesamiento de ECGs y red neuronal empleada (LSTM).
MultiFusionNet: Detección de fibrilación auricular con redes neuronales profundas (Tran et al., 2020).	Detectar la Fibrilación Auricular combinando dos redes neuronales profundas entrenadas con diferentes datos de entrada.	RNA x LSTM.	Fusión de los datos de salida de ambas redes neuronales.

Resumen de Artículos

Trabajo	Objetivo	Algoritmos/Técnicas	Utilidad
Detección automatizada de fibrilación auricular utilizando una red de memoria a corto plazo con señales de intervalo RR (Faust et al., 2018).	Detección de fibrilación auricular usando intervalos RR divididos en secuencias superpuestas	Red Neuronal Recurrente Bidireccional de Memorias a largo y corto plazo.	Modelo utilizado (Bi-LSTM) y base de datos utilizada (MIT-BIH AFDB).
Clasificación de la fibrilación auricular basada en redes MLP extrayendo los parámetros de Jitter y Shimmer (Borghi et al., 2021)	Sistema basado en un MLP para diferenciar entre FA y no FA mediante la segmentación de características del ECG	Extracción de parámetro Jitter y Shimmer	Técnicas de procesamiento de señales
Detección de fibrilación auricular a partir de señales de ECG cortas mediante un modelo híbrido de aprendizaje profundo (Wu et al., 2019)	Detectar Fibrilación Auricular a partir de señales de ECG mediante un modelo híbrido	En este trabajo, se destacó el conjunto de datos CinC Challenge 2017 de Physionet para su exploración y posteriormente evaluar un modelo propuesto	Conjunto de datos utilizado
Diseño y desarrollo del modelo LSTM-RNN para la predicción de intervalos RR en señales de ECG (Dhananjay et al., 2021)	Predecir intervalos RR en señales de ECG	Detección y filtrado de los picos R de las señales ECG	Técnicas de procesamiento de señales
Detección de fibrilación auricular a partir de datos de fotopleletismografía de muñeca mediante redes neuronales artificiales (Yousefi et al., 2018)	Evaluar el método propuesto para la detección de Fibrilación Auricular	Se usaron las métricas AUC, Sensibilidad y Especificidad	Métricas de evaluación de la calidad de clasificación

Resumen de Artículos - Continuación

Trabajo	Objetivo	Algoritmos/Técnicas	Utilidad
Un dispositivo de ECG móvil de una derivación incorporado en un programa de monitoreo remoto de arritmias basado en síntomas (Selder et al.,2019)	Evaluar el algoritmo para la detección de arritmias cardiacas.	Se usaron las métricas Exactitud, Sensibilidad y Especificidad	Métricas de evaluación de la calidad de clasificación
Red neuronal profunda de fusión multiescala para la detección de arritmias multi-clase (Wang et al.,2020)	Formular una nueva arquitectura de Red Neuronal Convolutiva de fusión profunda de múltiples escala de extremo a extremo	DMSFNet	Modelo de red neuronal para la clasificación y métricas para la evaluación del modelo
ECGDeepNET: un enfoque de aprendizaje profundo para clasificar los latidos del ECG (Mahmud et al.,2019)	Clasificar anomalías en ECG con clases desbalanceadas	ECGDeepNET como modelo de clasificación en donde se empleó la Precisión, F1-Score, Especificidad para medir el rendimiento del modelo	Modelo de red neuronal y métricas de evaluación empleadas
Detección de arritmias cardiacas a partir de ECG con redes neuronales convolucionales recurrentes (Zaen et al.,2020)	Detectar arritmias cardiacas a partir de ECG de una sola derivación utilizando RCNN	Conjunto de datos CinC Challenge 2017	Bases de datos de ECG
Procesamiento del lenguaje de ECG (ELP): una nueva técnica para analizar señales de ECG (Mousavi et al.,2021)	Análisis de señales mediante el procesamiento del lenguaje de ECG	Las bases de datos que se emplearon para el análisis y su posterior evaluación con modelos de aprendizaje profundo son: MIT-BIH, MIT-BIH AFIB y CinC Challenge 2017	Bases de datos de ECG

Resumen de Artículos - Continuación

Trabajo	Objetivo	Algoritmos/Técnicas	Utilidad
Modelo híbrido de aprendizaje profundo CNN-LSTM y técnica de conjunto para la detección automática de infarto de miocardio utilizando grandes datos de ECG (Rai and Chatterjee,2021)	Detectar automáticamente el infarto de miocardio	Las métricas que se utilizaron para cuantificar el rendimiento del modelo propuesto son: Precisión, Sensibilidad, Especificidad, Tasa de error de clasificación, Exactitud, Puntuación F1 y el Índice Youden	Métricas para la evaluación de modelos de red
Detección de fibrilación auricular utilizando una descomposición multiescala mejorada en una Red neuronal convolucional residual mejorada (Cao et al.,2019)	Detectar la Fibrilación Auricular usando descomposición multi escala en una red residual	El modelo de red utilizado fue FDResNet y el conjunto de datos fue CinC Challenge 2017	Modelo de red neuronal para la clasificación y bases de datos
Aplicación de redes neuronales densas para la detección de fibrilación auricular y clasificación de conjunto de características de ECG aumentado (Krasteva et al.,2021)	Clasificación de la Fibrilación Auricular mediante 137 características importantes	La base de datos empleada en el trabajo fue CinC Challenge 2017	Base de datos de ECG
Aprendizaje profundo para la detección del tipo de arritmia de latido de ECG de una sola derivación mediante una representación de espectrograma de iris (Alqudah and Alqudah,2021)	Detección de 17 clases diferente de arritmias cardíacas	CNN	Modelo de red neuronal para la clasificación
Reconocimiento de episodios de AF mediante funciones optimizadas de frecuencia de tiempo y SVM sensible al costo (Hamil et al., 2021)	Detectar y clasificar la Fibrilación Auricular	Transformada de Stockwell	Técnica de procesamiento de señales ECG

CAPÍTULO 4

Metodología

En este capítulo se aborda el tema principal de esta investigación y se presenta el enfoque propuesto para la clasificación de señales electrocardiográficas utilizando características de Tiempo-Frecuencia. El objetivo principal es realizar una transformación de la señales en crudo del dominio caótico al dominio de Tiempo-Frecuencia para posteriormente, realizar una extracción de características conservando el dominio de Tiempo-Frecuencia y hacer una reducción de dimensionalidad para entrenar la Red Neuronal Recurrente Profunda. La metodología propuesta se plantea para la clasificación de la Fibrilación Auricular utilizando momentos de Tiempo-Frecuencia, dada su importancia actual en el campo de la investigación médica, pudiéndose apreciar gráficamente en la Figura 4.1.

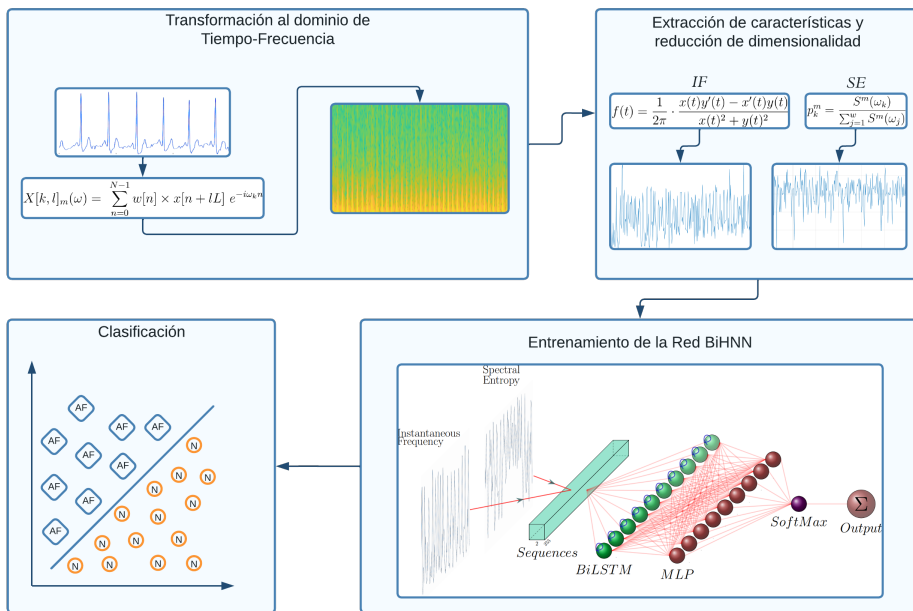


Figura 4.1: Diagrama de bloques de la metodología

De acuerdo con el esquema final propuesto visualizado anteriormente, para una mejor organización, la metodología de solución se desglosará en 3 módulos principales. El primer módulo se enfoca a la adquisición de los datos a usar, así como el proceso de transformación al dominio de tiempo-frecuencia. El segundo módulo es referente a la extracción de características y una reducción de dimensionalidad de las mismas, conservando el dominio de Tiempo-Frecuencia. En el tercer módulo, se detalla la formalización de la Red Neuronal Híbrida Bidireccional, la descripción puntual de cada capa constituyente de la red, así como el procesamiento interno de la información dentro de la misma. El módulo de clasificación se detalla en el capítulo 5 junto con la experimentación y resultados.

4.1 Transformación al dominio de Tiempo-Frecuencia

4.1.1 Conjunto de datos

Para los fines de experimentación se emplearon los ECG de la base de datos 2017 Challenge de PhysioNet (Goldberger et al.,2000), constituidos en un conjunto de ECG's muestreadas a 300 Hz. El uso de esta base de datos se debe a la existencia dos clases que son de interés para las tareas de clasificación, siendo estas: Normal (N), AFib (A), aunado a que los ECG son de una sola derivación y la duración en segundos de una señal individual en su mayoría ronda los 30 segundos. Pudiéndose observar una muestra de señal ECG de cada clase mencionada en la Figura 4.2.

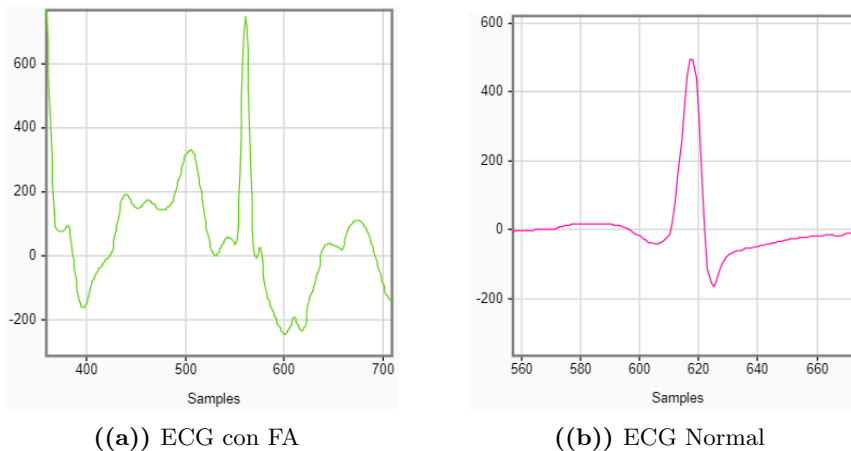


Figura 4.2: Señales de la base de datos seleccionada.

4.1.2 Generación de espectrogramas

Como se había mencionado, la información en crudo de los ECG son de naturaleza caótica, por lo tanto, es necesario realizar una transformación al dominio de tiempo-frecuencia. En este sentido, se emplea la Transformada de Fourier de Tiempo Corto (STFT), el cual es un método comúnmente utilizado para el estudio de señales ECG, que se encarga de mapear la señal en una función bidimensional que se encuentra en el dominio del tiempo y la frecuencia. Para obtener información del ECG, se aplica STFT a la señal, dividiéndola en diferentes segmentos de tiempo y según superposiciones, el proceso es comúnmente conocido como ventaneo y el proceso para generar un espectrograma aplicando la STFT se puede apreciar gráficamente en la Figura 4.3.

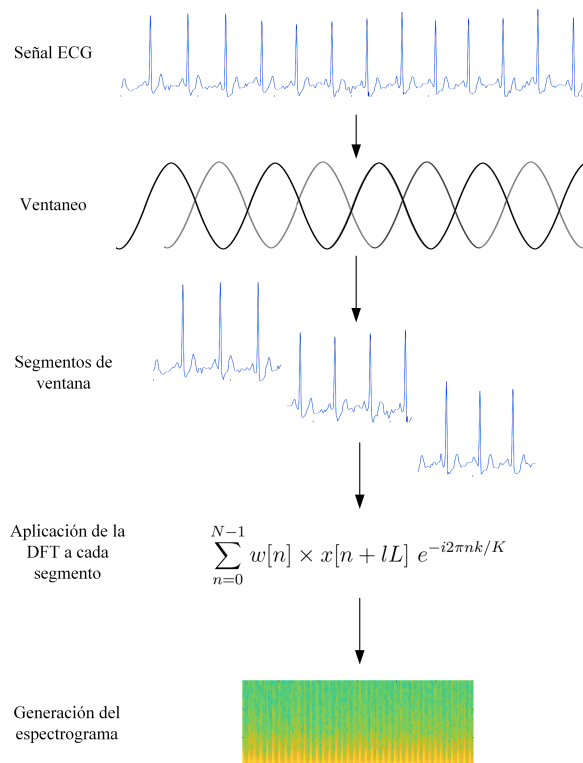


Figura 4.3: Fases de la STFT.

En concordancia con la función de ventana de tiempo w_n , el gráfico espectral resultante (espectrograma) se clasifica como de banda estrecha o de banda ancha. Cuando el intervalo de las ventanas de tiempo sea corto, la STFT será de banda ancha; durante un período de tiempo más largo, se producirá un espectro de banda estrecha. Los segmentos de señal (tramas) se pueden representar mediante:

$$x_l[n] = w_n \times x[n + lL], \quad 0 \leq n \leq N - 1 \quad (4.1)$$

donde N es el tamaño de la ventana, l es el índice de la trama y L es el tamaño del desplazamiento, n es el índice del instante correspondiente al comienzo del desplazamiento de la ventana. En la expresión $x[n + lL]$ representa una posición determinada sobre la señal y \times representa una modulación entre dos términos. Finalmente, la Transformada Discreta de Fourier es aplicada a cada trama de la señal de la siguiente manera (Smith, 2007):

$$X[k, l] = \sum_{n=0}^{N-1} x_l[n] e^{-i2\pi nk/K} = \sum_{n=0}^{N-1} w[n] \times x[n + lL] e^{-i2\pi nk/K} \quad (4.2)$$

donde K es la longitud de la DFT y k es un índice de frecuencia o índice bin. La STFT $X[k, l]$ representa el comportamiento de un segmento de la señal a través del tiempo lL y el número de bins k para una frecuencia de muestreo de F_s . Estos índices discretos corresponden a la longitud temporal de la muestra, que se denota por lL/F_s y la frecuencia es dada por kF_s/K y para poder simplificar la notación, se escribe la siguiente frecuencia radial de:

$$\omega_k = 2\pi k/K \quad (4.3)$$

Por lo tanto, la expresión de la STFT se reescribe como:

$$X[k, l]_m(\omega) = \sum_{n=0}^{N-1} w[n] \times x[n + lL] e^{-i\omega_k n} \quad (4.4)$$

Aplicando la ventana a lo largo del tiempo, la STFT resulta en una representación espectral de un segmento de tiempo de la señal de entrada; mostrando $X[k, l]$ como una función de la frecuencia k para cada valor del índice de tiempo l , $e^{-i\omega_k n}$ es el factor de fase para definir la distribución conjunta y así poder analizar el comportamiento de la señal a una frecuencia ω particular, la STFT corresponde a una serie de espectros ubicados en el tiempo. Expresándose como la serie al cuadrado de las STFT de dichos espectros como se observa en la expresión 4.5.

$$|X(k, l)|^2 \quad (4.5)$$

Para la construcción del espectrograma se muestra la potencia de cada espectro segmento por segmento, representando las magnitudes una al lado de la otra como una imagen con un mapa de colores dependiente de la magnitud.

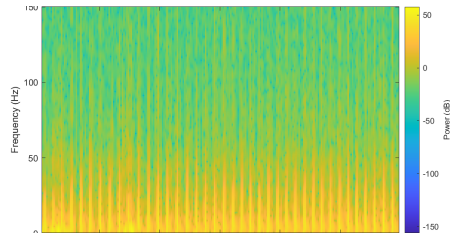


Figura 4.4: Representación de un espectrograma.

4.2 Extracción de características y reducción de dimensionalidad

Una vez generados los espectrogramas, ahora el problema a abordar es que las redes neuronales recurrentes solo pueden trabajar con datos en una dimensión. En este sentido, fue necesario extraer características específicas que se encontraran en el dominio de Tiempo-Frecuencia y así poder realizar una reducción de dimensionalidad de los datos, por ende, se realizó la extracción de momentos de Tiempo-Frecuencia conocidos como Frecuencia Instantánea y Entropía Espectral, que se detallan a continuación.

4.2.1 Frecuencia Instantánea

La Frecuencia Instantánea (FI) de una señal se considera como un parámetro variable en el tiempo que se relaciona con la media de las frecuencias en la señal a medida que evoluciona. Desde la perspectiva probabilística, (Flandrin, 1998) y (Barnes, 1993) muestran que la frecuencia instantánea es el primer momento o media de la densidad espectral de la señal en un instante de tiempo, como se puede observar en la Figura 4.5.

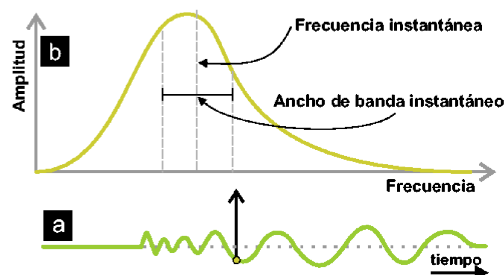


Figura 4.5: Diagrama de una señal y un espectro hipotético en un instante de tiempo (Extraído de Benjumea,2003).

En la Figura 4.5a se observa la señal en el dominio del tiempo, en la Figura 4.5b el espectro de la señal se resalta y pudiéndose observar la frecuencia instantánea como el valor medio del espectro y el ancho de banda instantáneo como la desviación estándar de los datos con respecto a la media. Tomando en cuenta que una señal compleja como el espectrograma puede ser expresada como $\hat{s}(t) = s(t) + s^*(t)$, siendo $s(t)$ la magnitud de la señal y $s^*(t)$ la fase de la misma, se puede realizar una mejor representación para esta investigación, por lo cuál la representación de la señal compleja se reescribe como:

$$\hat{s}(t) = A(t) + \exp^{\theta(t)} \quad (4.6)$$

Inferiendo que si los términos reales de una señal compleja son conocidos, entonces se puede encontrar la magnitud (tiempo) $A(t) = \sqrt{s^2(t) + s^{*2}(t)} = |S(t)|$ y la fase instantánea $\theta(t) = \arctan\left(\frac{s^*(t)}{s(t)}\right)$ de la señal. Por otra parte, la variación de la fase en el tiempo está relacionada con la frecuencia angular de la señal y es denominada frecuencia angular instantánea (Taner et al., 1979). Expresado de otra forma, la frecuencia angular instantánea es la derivada con respecto al tiempo de fase, expresado como:

$$\frac{d}{dt}\theta(t) = \omega(t) \quad (4.7)$$

En el caso de esta investigación se necesita saber la frecuencia (cantidad de ciclos por unidad de tiempo), en lugar de la frecuencia angular (radianes por unidad de tiempo), por lo que se emplea la relación entre frecuencia y frecuencia angular $f = \omega/2\pi$ con lo cuál se define la frecuencia instantánea de un espectrograma mediante la expresión (Boashash, 1992):

$$f(t) = \frac{1}{2\pi} \frac{d}{dt}\theta(t) \quad (4.8)$$

donde $\theta(t)$ es la fase instantánea en el tiempo. Una ambigüedad inherente a la fase instantánea hace que la expresión 4.8 no sea práctica para calcular la frecuencia instantánea ya que solo se calculan los valores principales de la fase, lo que provoca discontinuidades de fase de 2π . En cambio, la frecuencia instantánea se calcula mediante otra ecuación, reescribiendo la expresión 4.9 como:

$$f(t) = \frac{1}{2\pi} \cdot \frac{x(t)y'(t) - x'(t)y(t)}{x(t)^2 + y(t)^2} \quad (4.9)$$

donde $x(t)$ en el tiempo dado, $y(t)$ la transformación de Hilbert o el rastro imaginario, denotando una diferenciación en el numerador con respecto al tiempo. Aplicando la frecuencia instantánea a un espectrograma de una señal de la base de datos se puede observar la frecuencia instantánea de un ECG en la Figura 4.6.

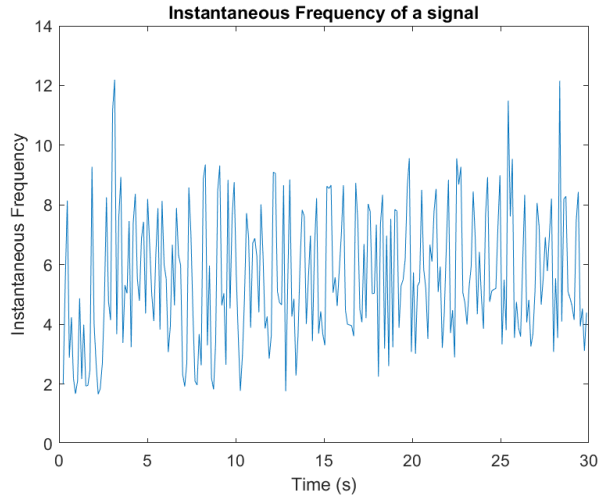


Figura 4.6: Representación de la Frecuencia Instantánea.

4.2.2 Entropía Espectral

La Entropía Espectral (SE) de una señal es una medida de su distribución de potencia espectral. El concepto se basa en la entropía de Shannon, o entropía de la información, que, en la teoría de la información, es una medida de la incertidumbre promedio, calculada a partir de la probabilidad de que ocurra cada evento. Para lograr la entropía, es necesario utilizar la teoría de la probabilidad, cuya función principal es gestionar una gran cantidad de eventos o eventos aleatorios individuales, pero que pueden predecirse globalmente en condiciones probabilísticas. Un sistema general consiste en un conjunto de eventos que ocurren uno tras otro, es decir, solo uno de ellos puede ocurrir en cada iteración. El conjunto que compone todo este sistema puede contener n eventos $S = E_1, E_2, E_3, \dots, E_n$ en donde cada uno de estos tiene una probabilidad de ocurrencia $p_1, p_2, p_3, \dots, p_n$ respectivamente, siempre y cuando:

$$\sum_{i=0}^n p_i = 1 \quad \text{donde} \quad p_i \geq 0 \quad (4.10)$$

Al tener los eventos con sus respectivas probabilidades de ocurrencia se tiene un esquema finito el cual es descrito como un estado de incertidumbre

$$S = \begin{pmatrix} E_1 & E_2 & E_3 & \dots & E_n \\ p_1 & p_2 & p_3 & \dots & p_n \end{pmatrix} \quad (4.11)$$

Observando también, la incertidumbre máxima de un sistema con dos eventos en la Figura 4.7.

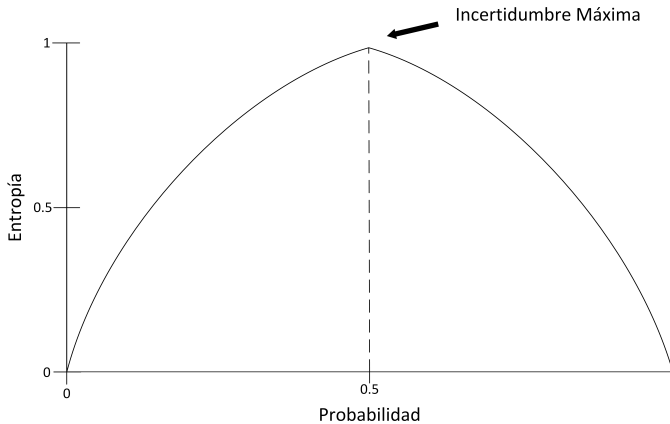


Figura 4.7: Representación de la función de Entropía.

Adecuando el concepto a esta investigación, la Entropía Espectral trata la distribución de potencia normalizada de la señal en el dominio de la frecuencia como una distribución de probabilidad, y calcula la incertidumbre o entropía del espectro (Pan et al., 2008). Por lo tanto, conociendo el espectrograma de potencia de frecuencia de tiempo, se obtiene la Entropía Espectral en el tiempo mediante:

$$SE(m) = - \sum_{k=1}^w p_k^m \log_2 p_k^m \quad (4.12)$$

donde nuevamente m es el índice de la ventana, w es la duración de las ventanas y p_k^m es la Función Masa de Probabilidad (FMP) que corresponde a la m -ésima ventana, obtenida a partir de la densidad espectral de potencia o también conocido como el espectrograma ($S^m(\omega_k)$)

$$p_k^m = \frac{S^m(\omega_k)}{\sum_{j=1}^w S^m(\omega_j)} \quad (4.13)$$

De esta forma, la función p_k^m se encarga de verificar la condición necesario de las FMPs: $\sum_k p_k^m = 1$. Con la normalización realizada en (4.13) se obtiene una medida que se independiza de la potencia total de la señal, y por lo tanto la SE es independiente de la amplitud, observándose finalmente la representación gráfica la Entropía Espectral a partir de un espectro de potencia (espectrograma) en la Figura 4.8.

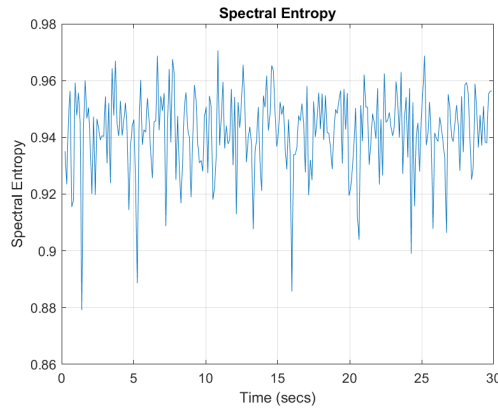


Figura 4.8: Representación de la Entropía Espectral.

4.3 Red Neuronal Híbrida Bidireccional

En cuanto a la red empleada para este trabajo se encuentra dada por una Red Neuronal Híbrida Bidireccional (BiHNN), conformada inicialmente por una capa de entrada de secuencias, una capa recurrente LSTM bidireccional (BiLSTM) propuesta por (Schuster and Paliwal, 1997) formada por dos capas ocultas LSTM apiladas en direcciones opuestas (Hochreiter and Schmidhuber, 1997) para aprender las dependencias bidireccionales a largo plazo entre los pasos temporales, una tercer capa totalmente conectada para de obtener las probabilidades de pertenecer a las diferentes clases, una capa Softmax para representar las probabilidades mediante una distribución categórica y por último una capa de clasificación. El diagrama de la arquitectura se puede observar en la Figura 4.9.

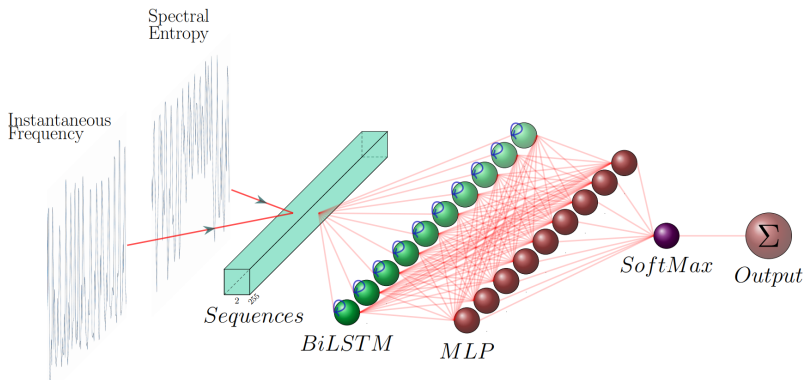


Figura 4.9: Arquitectura de la BiHNN.

De acuerdo a la figura anterior, la arquitectura esta compuesta inicialmente de una capa de entrada de secuencias para procesar las características extraídas, con una dimensionalidad de [2,255] haciendo referencia a la frecuencia instantánea y la entropía espectral.

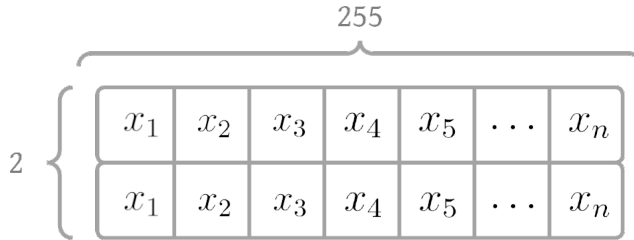


Figura 4.10: Estructura interna de la capa de entrada de secuencias.

La información presente en la capa de entrada secuencias , se introduce en la capa BiLSTM, la cuál consta de 2 dos capas ocultas LSTM con 100 neuronas cada una (Hochreiter and Schmidhuber, 1997). Se emplea este tipo de neuronas, dado que a diferencia de las RNN convencionales (Elman, 2001), las LSTM son un tipo de neuronas recurrentes especialmente para el aprendizaje de dependencias distantes entre los valores de datos secuenciales (Pattanayak, 2017).

Al igual que otros tipos de neuronas recurrentes, esta fue propuesta para resolver un problema en específico, la cual consta en evitar el problema del desvanecimiento del gradiente provocado por la propagación hacia atrás del error a través de las capas ocultas en la etapa de entrenamiento descrito en la expresión 4.14:

$$\frac{\partial L[t]}{\partial W} = \sum_{\tau} \frac{\partial L[t]}{\partial h[t]} \frac{\partial h[t]}{\partial h[\tau]} \frac{\partial h[\tau]}{\partial W} \quad (4.14)$$

donde t representa un tiempo de entrenamiento, τ el último tiempo, h las capas ocultas y L la función de error. Al acumularse el error a través de t , el error tiende al infinito o en el peor de los casos a 0, mermando el rendimiento de aprendizaje de la red hacia nuevos patrones. Por lo tanto, para subsanar este problema, se hace uso de las neuronas LSTM en donde la estructura central o más bien la inteligencia de esta neurona reside en las 5 compuertas por las que está compuesta, realizando una función específica mostradas en la Figura 2.5 y descritas en las expresiones 2.8-2.12.

Principalmente se tienen 3 compuertas para controlar un vector de activación de celdas de memoria. La primera compuerta es la de olvido, la cuál determina cuanta información del estado de la celda en el tiempo anterior c_{t-1} se debe conservar hasta el estado actual c_t de la celda de estado; la segunda compuerta es la de entrada, misma que se encarga de delimitar la entrada x_t de la red que se guardará en la celda de estado actual; la tercera compuerta determina cuanta información de la celda de estado c_t se debe transmitir hacia al valor h_t de la salida actual de la neurona.

Las tres compuertas funcionan por si solas como una capa totalmente conectada, o de otra manera se puede interpretar como una red de redes, teniendo como entradas un vector y como salida se tienen numeros reales en $[0,1]$.

$$\text{Input gate : } i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + b_i) \quad (4.15)$$

$$\text{Forget gate : } f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + b_f) \quad (4.16)$$

$$\text{Output gate : } o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + b_o) \quad (4.17)$$

$$\text{Cell state : } c_t = (f_t \times c_{t-1}) + (i_t \times \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c)) \quad (4.18)$$

$$\text{Cell output : } h_t = o_t \tanh(c_t) \quad (4.19)$$

donde $\sigma(\cdot)$ es la función sigmoide $\sigma(x) = \frac{1}{1 + e^{-x}}$, x_t indica el vector del t-ésimo valor de la señal electrocardiográfica y h_t indica el estado oculto, los términos W y b , respectivamente, representan las matrices de peso y los vectores de sesgo de las distintas compuertas de la neurona LSTM (p.ej., W_{xo} representa la matriz de pesos de la compuerta de salida y b_i representa el vector de sesgos de la compuerta de entrada). Por lo que de manera sintetizada, estas compuertas funguen como limitantes durante la actualización de la memoria de la neurona LSTM mediante la recurrencia, evitando así, el sobre aprendizaje y el problema originado por el metodo de la propagación hacia atrás a través del tiempo mencionado en la expresión 2.7.

Dado que una sola capa LSTM solo puede memorizar los tiempos pasados que imposibilita memorizar los tiempos futuros, para superar esta deficiencia característica, (Schuster and Paliwal, 1997) propusieron las redes neuronales recurrentes bidireccionales (BiRNN) para poder combinar dos capas LSTM ocultas separadas en direcciones opuestas pero siempre apuntando a la misma salida. La estructura interna de una capa BiLSTM se puede apreciar en la Figura 4.11.

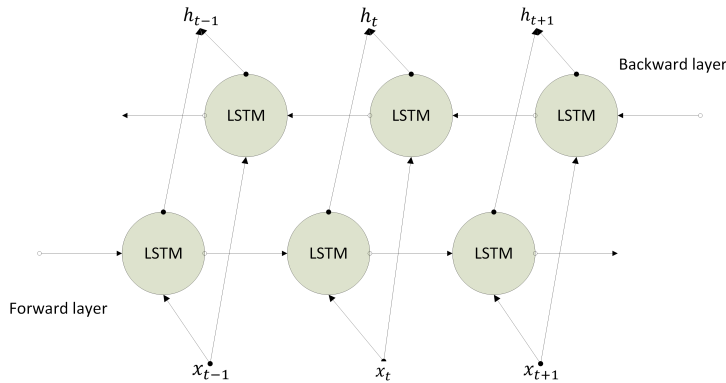


Figura 4.11: Estructura interna de una capa BiLSTM.

Con esta estructura, la capa de salida puede hacer uso de la información relacionada tanto del tiempo anterior como del futuro. En otras palabras, una BiLSTM calcula la secuencia de entradas $x = (x_1, x_2, \dots, x_n)$ desde la dirección opuesta hacia una secuencia oculta hacia adelante $\vec{h} = (\vec{h}_1, \vec{h}_2, \dots, \vec{h}_n)$ y una secuencia oculta hacia atrás $\overleftarrow{h} = (\overleftarrow{h}_1, \overleftarrow{h}_2, \dots, \overleftarrow{h}_n)$. El vector codificado de salidas h_t se genera a partir de la concatenación de las salidas finales hacia adelante y hacia atrás, $h_t = [\vec{h}_t, \overleftarrow{h}_t]$. Expresándose lo anterior mencionado mediante las siguientes ecuaciones.

$$\vec{h}_t = \sigma(W_{x\vec{h}}x_t + W_{\vec{h}\vec{h}}\vec{h}_{t-1} + b_{\vec{h}}) \quad (4.20)$$

$$\overleftarrow{h}_t = \sigma(W_{x\overleftarrow{h}}x_t + W_{\overleftarrow{h}\overleftarrow{h}}\overleftarrow{h}_{t+1} + b_{\overleftarrow{h}}) \quad (4.21)$$

$$h_t = W_{y\vec{h}}\vec{h}_t + W_{y\overleftarrow{h}}\overleftarrow{h}_t + b_h \quad (4.22)$$

donde $h = (h_1, h_2, \dots, h_t, \dots, h_n)$ es la secuencia de salida de la capa BiLSTM. La salida de la capa BiLSTM sirve como datos de entrada para una capa totalmente conectada con 100 neuronas de tipo perceptrón.

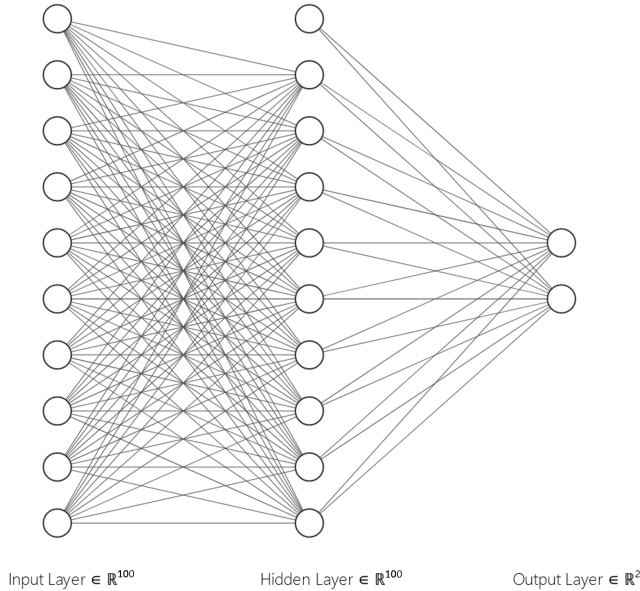


Figura 4.12: Capa totalmente conectada.

Las salidas de las neuronas de la capa oculta se ocultan mediante las sumatorias de las entradas, esto siendo ponderados por el peso menos el sesgo de la respectiva capa, escribiendo la siguiente expresión (Caterini and Chang, 2018):

$$\zeta_i^{(1)} = \sum_{i,j=1}^n W_{ij}^{(1)} x_i^{(0)} - b_i^{(1)} \quad (4.23)$$

Posteriormente, las salidas de las neuronas se convierten en entradas para la siguiente capa, no sin antes procesarla mediante una función de activación.

$$y_i = x_i^{(1)} = \psi(\zeta_i^{(1)}) = \psi\left(\sum_{i,j=1}^n W_{ij}^{(1)} x_i^{(0)} - b_i^{(1)}\right) \quad (4.24)$$

Para efectos de practicidad, se puede representar la expresión 4.23 en su forma matricial, reescribiéndose de la siguiente manera:

$$\zeta^i = \varpi^{(1)T} \chi^{(0)} - b^{(1)} \quad (4.25)$$

Donde ϖ^{jT} son los pesos de todas las entradas, χ^j y b^j son las entradas y los sesgos de la capa totalmente conectada respectivamente. Finalmente, se calculan las salidas de las neuronas de la capa de salida.

$$\zeta_i^{(2)} = \sum_{i,j=1}^n W_{ij}^2 x_i^1 - b_i^2 \quad (4.26)$$

Concatenando todas las entradas y los pesos de la capa anterior en una notación matricial.

$$y_i = x_i^{(2)} = \psi(\zeta_i^{(2)}) = \psi(\varpi^{(2)T} \chi^{(1)} - b^{(2)}) \quad (4.27)$$

Por ende, la expresión general que constituye la capa totalmente conectada se escribe como:

$$y_i = \psi(\varpi^{(2)T} \psi(\varpi^{(1)T} \chi^{(0)} - b^{(1)}) - b^{(2)}) \quad (4.28)$$

Con una función de activación $\psi(\cdot)$ regularmente de tipo sigmoideal de acuerdo a la literatura para generar salidas acotadas en $[0, 1]$, esta función de activación es la que regularmente se emplea, sin embargo, se puede recurrir a otros tipos de funciones dependiendo del criterio de cada investigador. La salida de la capa totalmente conectada pasa a través de una función softmax para transformar (normalizar) dichas salidas a una representación de distribución probabilística, de tal manera que la sumatoria de todas las probabilidades de las salidas sea 1 (Svensén and Bishop, 2007). Definiéndose en la siguiente expresión.

$$f(y_i) = \frac{\exp(y_i)}{\sum_{k=1}^K \exp(y_i)} \quad (4.29)$$

Finalmente se tiene una capa de clasificación para determinar si la salida obtenida pertenece a la clase de señales con Fibrilación Auricular o señales con latidos normales.

4.3.1 Algoritmo de optimización del error

El modelo en cuestión es entrenado con el método *Adaptive Moment Estimation* (ADAM) (Kingma and Ba, 2014) para la propagación del error. En el algoritmo 1 se encuentra de manera sintetizada los procedimientos que se efectúan para el ajuste de los parámetros de la red. Donde g_t representa el gradiente ∇f , el cuál es calculado a partir de la función de costo y el momentum (línea 3 del Algoritmo 1), α es la tasa de aprendizaje recomendada con un valor de 0.01, β_1 y β_2 son las tasas de decaimiento exponencial para las estimaciones del momentum y se recomiendan los valores de 0.9 y 0.999 respectivamente y el hiperparámetro ϵ es recomendado en 10^{-8}

Algorithm 1 Algoritmo de propagación del error ADAM.

Require: δ_j
Ensure: δ_k

- 1: **while** δ_k not converged **do**
- 2: $k \leftarrow k + 1$
- 3: $g_t \leftarrow \nabla_{\delta} f_t(\delta_{k-1})$
- 4: $m_k \leftarrow \beta_1 \cdot m_{k-1} + (1 - \beta_1) \cdot g_k$
- 5: $v_k \leftarrow \beta_2 \cdot v_{k-1} + (1 - \beta_2) \cdot g_k^2$
- 6: $\hat{m}_k \leftarrow \frac{m_k}{1 - \beta_1^k}$
- 7: $\hat{v}_k \leftarrow \frac{v_k}{1 - \beta_2^k}$
- 8: $\delta_k \leftarrow \frac{\delta_{k-1} - \alpha \cdot \hat{m}_k}{(\sqrt{\hat{v}_k} + \epsilon)}$
- 9: **end while**
- 10: **end while**
- 11: **return** δ_k

δ_k se hace referencia a los parámetros entrenables de toda la red que se ajustan a través de $\delta_k \leftarrow \delta_{k-1} + \nabla \delta$, donde $\nabla \delta$ refiere al incremento o decremento de los pesos obtenidos en la línea 8 del algoritmo 1. Por último, el índice k se refiere a todas las iteraciones de entrenamiento e independiente de los tiempos t de la capa recurrente distribuida en el tiempo.

4.3.2 Propagación hacia atrás a través del tiempo

Al igual que en las redes neuronales existentes en la literatura, las redes neuronales recurrentes también actualizan sus parámetros mediante propagación hacia atrás al encontrar el gradiente del error (pérdida) con respecto a sus pesos. En el caso de la Red Neuronal Híbrida propuesta no es la excepción, ya que en este tipo de Red Neuronal, al utilizar celdas LSTM en la capa birireccional, hace que cada nodo en la red tenga un paso de tiempo.

Por ende, para actualizar los parámetros de la red se utiliza la Propagación Hacia Atrás a Través del Tiempo (del inglés BPTT). Aquí, usando BPTT, se quiere averiguar cuánto afectan las unidades ocultas y la salida al error total, así como también cuánto afecta el cambio de pesos (W_y, W_h, W_x) a la salida. W_x , es constante en toda la red, por lo que se debe retroceder hasta el paso de tiempo inicial para actualizarlo (Dawani, 2020).

Cuando se propaga hacia atrás una red recurrente, se aplica la regla de la cadena. Lo que dificulta el entrenamiento de la red recurrente, debido a que la función de pérdida no solo depende de la activación de la capa de salida sino también de la activación de la capa oculta actual y su efecto en la capa oculta en el siguiente paso de tiempo.

En la siguiente ecuación, podemos ver cómo la retropropagación en una red recurrente. Como se puede observar, primero se encuentra la pérdida de entropía cruzada, por lo que el error total es el siguiente:

$$\frac{\partial E}{\partial W_x} = \sum_{i=1}^{\tau} \frac{\partial E_t}{\partial W_x} \quad (4.30)$$

Expandiendo la ecuación anterior usando la regla de la cadena con respecto a las pérdidas y capas ocultas de la siguiente manera:

$$\frac{\partial E_t}{\partial W_x} = \sum_{k=1}^t \frac{\partial E_t}{\partial y_t} \frac{\partial y_t}{\partial h_t} \frac{\partial h_t}{\partial h_k} \frac{\partial h_k}{\partial W_x} \quad (4.31)$$

Centrándose en el tercer término del lado derecho, se amplía como:

$$\frac{\partial h_t}{\partial h_j} = \prod_{j=k+1}^t \frac{\partial h_j}{\partial h_{j-1}} \quad (4.32)$$

Tomando en cuenta que cada derivada parcial es una matriz jacobiana, ahora se pueden combinar las ecuaciones anteriores para obtener una visión holística de cómo calcular el error, que se ve de la siguiente manera:

$$\frac{\partial E_t}{\partial W_x} = \sum_{k=1}^t \frac{\partial E_t}{\partial y_t} \frac{\partial y_t}{\partial h_t} \left(\prod_{j=k+1}^t \frac{\partial h_j}{\partial h_{j-1}} \right) \frac{\partial h_k}{\partial W_x} \quad (4.33)$$

Teniendo en cuenta que para obtener el estado oculto de una compuerta LSTM h_t se calcula usando la siguiente ecuación:

$$h_t = f(W_x x_t + W_h h_{t-1} + b) \quad (4.34)$$

Entonces, se calcula el gradiente de h_t como:

$$\frac{\partial h_t}{\partial h_{t-1}} = W_h^\top \text{diag}(f(W_x x_t + W_h h_{t-1} + b)) \quad (4.35)$$

Dado que las neuronas ocultas también toman x_t como entrada, también debemos tomar la derivada con respecto a W_x . Podemos hacer esto de la siguiente manera:

$$\frac{\partial E_t}{\partial W_x} = \frac{\partial E_t}{\partial y_t} \frac{\partial y_t}{\partial h_t} \frac{\partial h_t}{\partial W_x} \quad (4.36)$$

Cabe recalcar que las unidades ocultas toman dos entradas. Entonces, se retropropaga un paso de tiempo usando lo anterior expresado y entonces se puede reescribir como:

$$\frac{\partial E_t}{\partial W_x} = \frac{\partial E_t}{\partial y_t} \frac{\partial y_t}{\partial h_t} \frac{\partial h_t}{\partial W_x} + \frac{\partial E_t}{\partial y_t} \frac{\partial y_t}{\partial h_t} \frac{\partial h_t}{\partial h_{t-1}} \frac{\partial h_{t-1}}{\partial W_x} \quad (4.37)$$

Usando esto, ahora se pueden sumar todos los gradientes anteriores hasta el presente, de la siguiente manera:

$$\frac{E_t}{\partial W_x} = \frac{\partial E_t}{\partial y_t} \frac{\partial y_t}{\partial h_t} \frac{\partial h_t}{\partial h_{t-1}} \frac{\partial h_{t-1}}{\partial W_x} \quad (4.38)$$

Ahora bien, la expresión anterior solo denota la BPTT de una sola capa recurrente, en el caso de este trabajo de investigación lo que se precisa es obtener la BPTT tanto para la capa hacia adelante como para la capa hacia atrás, por lo que haría falta hacer una adecuación a la expresión 3.38, quedando la BPTT para las capas internas de la capa BiLSTM de la siguiente manera:

$$\frac{\vec{E}_t}{\partial W_x} = \frac{\partial \vec{E}_t}{\partial y_t} \frac{\partial y_t}{\partial \vec{h}_t} \frac{\partial \vec{h}_t}{\partial \vec{h}_{t-1}} \frac{\partial \vec{h}_{t-1}}{\partial W_x} \quad (4.39)$$

$$\frac{\overleftarrow{E}_t}{\partial W_x} = \frac{\partial \overleftarrow{E}_t}{\partial y_t} \frac{\partial y_t}{\partial \overleftarrow{h}_t} \frac{\partial \overleftarrow{h}_t}{\partial \overleftarrow{h}_{t+1}} \frac{\partial \overleftarrow{h}_{t+1}}{\partial W_x} \quad (4.40)$$

4.4 Discusión

Hasta este punto se abordó el diseño y desarrollo de la metodología para la clasificación de la Fibrilación Auricular que fue propuesto en tres módulos principales. Como resultado se obtuvo un método funcional que se evalúa cuantitativamente mediante métricas derivadas de la matriz de confusión encontradas en el estado del arte que se adaptaron a los experimentos descritos en el siguiente capítulo.

CAPÍTULO 5

Experimentación y Resultados

En este capítulo se detallan los experimentos diseñados para la evaluación cuantitativa de la metodología desarrollada, así como las métricas empleadas y comparaciones con diferentes enfoques que se encuentran en la literatura, además de los resultados correspondientes por cada experimentación.

5.1 Diseño Experimental

Con los datos procesados y la arquitectura de red definida, se realizó una serie de experimentaciones con las señales electrocardiográficas en donde se utilizaron los datos de entrenamiento. Los objetivos de la serie de experimentaciones se detalla en los siguientes puntos:

- Sintonizar los hiperparámetros de la red neuronal.
- Sintonizar la arquitectura híbrida con distintas capas recurrentes.
- Sintonizar la arquitectura híbrida con distinto número de capas BiLSTM.
- Comparar las características de TF con otro tipos.
- Comparar la arquitectura híbrida frente a otros modelos neuronales.

5.1.1 Entorno de desarrollo

Para la implementación del sistema propuesto se precisó el uso de un equipo de cómputo de alto rendimiento, debido a que en el mismo equipo se realizó tanto el entrenamiento como tareas de clasificación de la red neuronal. Cabe mencionar que es indispensable contar con una Unidad de Procesamiento Gráfico (GPU) por la potencia del cómputo en paralelo de la misma al procesar grandes cantidades de datos. En cuanto al software y librerías utilizadas son estrictas debido a la compatibilidad en entre sí. Por lo tanto, ya que en los experimentos realizados el uso medio de los recursos fue elevado, se considera como requisitos:

Arquitectura de hardware:

- Procesador Intel Core i7-10870H @ 2.20 GHz.
- 16 GB de memoria RAM.
- Disco de estado sólido de 512 GB.
- GPU GeForce RTX 3060 Laptop con 3840 núcleos CUDA y 6 GB de memoria VRAM.

Arquitectura de software:

- Sistema Operativo Xubuntu 20.04 LTS.
- MATLAB R2021a.
- Librería NVIDIA CUDA 11.
- Librería Scikit-Learn de Python.

5.1.2 Datos de entrenamiento y prueba

Se examinó el contenido de la base de datos propuesta para conocer la cantidad de señales presentes en cada clase, observando las cantidades en la siguiente tabla.

Tabla 5.1: Cantidad de señales por clase.

Etiqueta	Cantidad
Normal	4937
Fibrilación Auricular	718

Teniendo en cuenta las señales del conjunto de datos se realizó la división en un conjunto de datos para entrenamiento designando un 90% y el resto en un conjunto de pruebas. Tanto al conjunto de entrenamiento como al de pruebas, se le aplicó un aumento de datos para poder normalizar la cantidad de datos en ambas clases, y como se puede observar en la Tabla 5.2 se obtuvieron 4438 instancias de entrenamiento así como 490 instancias de prueba.

Tabla 5.2: Señales divididas para entrenamiento y pruebas

Etiqueta	Datos/Entrenamiento	Datos/Prueba
Normal	4438	490
Fibrilación Auricular	4438	490

5.1.3 Métricas

Para la evaluación de los resultados de esta investigación, se contemplan provisionalmente tres aspectos más utilizadas en estado del arte, siendo *Accuracy*, *Precision*, *Recall* y *F1 – Score* las métricas que se derivan de la matriz de confusión (Chen et al., 2021), detallándose de la siguiente manera:

La matriz de confusión es una tabla resumida utilizada para evaluar el rendimiento de un modelo de clasificación. La ponderación de predicciones correctas e incorrectas son resumidas con los valores de conteo y se separan por clase como se aprecia en la siguiente expresión:

$$ConfMtrx = \begin{pmatrix} TP & FN \\ FP & TN \end{pmatrix} \quad (5.1)$$

Donde *TP* son los resultados predichos correctamente en la clase positiva, *TN* son los resultados predichos correctamente en la clase negativa, *FP* son los resultados predichos incorrectamente en la clase positiva cuando en realidad es negativa y *FN* son los resultados predichos incorrectamente en la clase negativa cuando en realidad es positiva.

Accuracy calcula básicamente la exactitud de clasificación y se obtiene dividiendo la cantidad de predicciones correctas sobre la cantidad total de predicciones como se expresa a continuación:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (5.2)$$

Precision es la relación existente entre las instancias relevantes recuperadas y la cantidad total de instancias recuperadas y se calcula con la siguiente expresión:

$$Precision = \frac{TP}{TP + FP} \quad (5.3)$$

Recall es la tasa de aciertos o la tasa positiva real, la cual indica la cantidad total de instancias se han predicho correctamente y se tiene la siguiente expresión:

$$Recall = \frac{TP}{TP + FN} \quad (5.4)$$

F1 – Score es una medida de precisión y robustez en una prueba, calculando la media armónica de *Precision* y *Recall* como se expresa a continuación:

$$F1\ Score = \frac{2 \times (precision \times recall)}{precision + recall} \quad (5.5)$$

Por otra parte, se realizó una búsqueda en la literatura sobre otras métricas para evaluar el método propuesto y se consideraron dos métricas más, siendo el primero el puntaje Kappa de Cohen (CKS) (SAMUIC, 2021), el cual es medida de concordancia que se basa en comparar la concordancia observada en un conjunto de datos tras realizar una tarea de clasificación y viene dada por:

$$\kappa = \frac{\xi - \varkappa}{1 - \varkappa} \quad (5.6)$$

Donde ξ es la proporción de la concordancia observada y \varkappa es la proporción de la concordancia esperada. Si κ es cero, significa que la concordancia observada coincide con la que ocurriría por puro azar, si el resultado fuera 1, se trataría de una concordancia perfecta.

La última métrica considerada es el Coeficiente de Correlación de Mathews (MCC), el cual es un método de matriz de contingencia para escalar el coeficiente de correlación entre los valores reales y predichos (Chicco and Jurman, 2020). Esta métrica está pensada a manera de complemento alternativo a la métrica F1-Score, ya que MCC no se ve afectada cuando los resultados de clasificación se realizan con conjuntos de datos desequilibrados, definiéndose como:

$$MCC = \frac{(TP \times TN) - (FP \times FN)}{\sqrt{(TP + FP) \times (TP + FN) \times (TN + FP) \times (TN + FN)}} \quad (5.7)$$

5.2 Sintonización de la Red Neuronal Híbrida Bidireccional

5.2.1 Experimentación con distintas sintonizaciones de hiperparámetros

Como parte del proceso de entrenamiento de una red neuronal profunda, es indispensable realizar el proceso de sintonización de los hiperparámetros. En consecuencia, se realizó una primer experimentación con distintos hiperparámetros para encontrar la mejor configuración de clasificación. Dentro de los hiperparámetros se consideraron algoritmos de optimización del error de tipo adaptativos, dentro de los más conocidos son *Root Mean Square Propagation* (RMSPROP), *Stochastic Gradient Descent with momentum* (SGDM) y *Adaptive Moment Estimation* (ADAM). Aunado a lo anterior, se consideraron tasas de aprendizaje de 0.1, 0.01 y 0.001, y épocas de 10, 20 y 30. Los resultados de la experimentación se pueden observar en las tablas 5.3, 5.4 y 5.5.

Tabla 5.3: Resultados con el optimizador RMSPROP

Configuración	Accuracy	Precision	Recall	F1-Score	CKS	MCC
RMSPROP-10 0.001	0.7551	0.974	0.6781	0.7987	0.51	0.5659
RMSPROP-10 0.01	0.7214	0.7714	0.7013	0.7347	0.443	0.4451
RMSPROP-10 0.1	0.5122	0.0571	0.6364	0.1049	0.024	0.0591
RMSPROP-20 0.001	0.7388	0.7143	0.7511	0.7322	0.478	0.4781
RMSPROP-20 0.01	0.6133	0.9286	0.5695	0.7060	0.227	0.2919
RMSPROP-20 0.1	0.5153	0.9429	0.5083	0.6605	0.031	0.059
RMSPROP-30 0.001	0.7724	0.8429	0.7388	0.7874	0.545	0.5504
RMSPROP-30 0.01	0.9173	0.9143	0.9199	0.9171	0.835	0.8347
RMSPROP-30 0.1	0.5	0	0	0	0	0

Para un mejor análisis de los resultados, se generó un gráfico de barras con los resultados previos para observar las diferencias presentadas entre las distintas configuraciones de hiperparámetros.

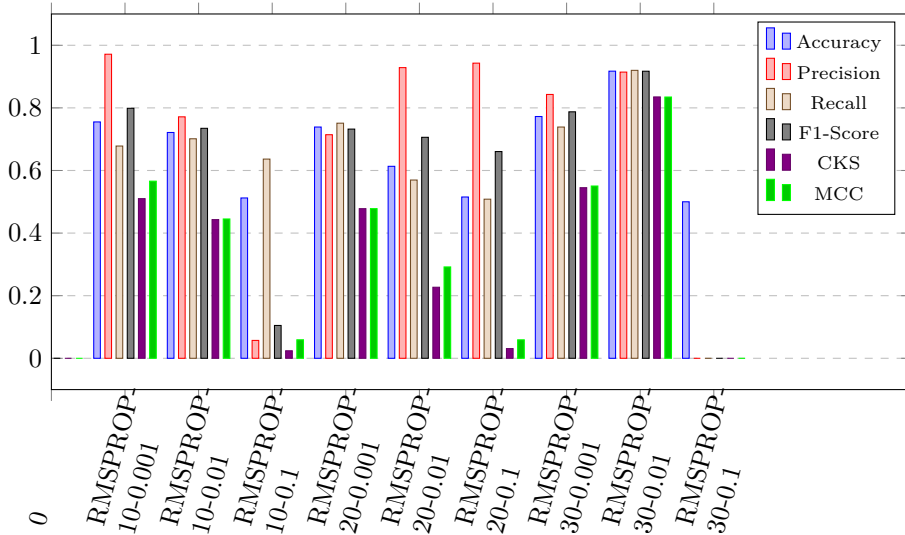
**Figura 5.1:** Resultados con el optimizador RMSPROP.

Tabla 5.4: Resultados con el optimizador SGDM

Configuración	Accuracy	Precision	Recall	F1-Score	CKS	MCC
SGDM-10 0.001	0.5286	0.4143	0.5370	0.4677	0.057	0.0587
SGDM-10 0.01	0.6796	0.6	0.7136	0.6519	0.359	0.3638
SGDM-10 0.1	0.6786	0.6143	0.7049	0.6565	0.357	0.3601
SGDM-20 0.001	0.5673	0.4867	0.5805	0.5289	0.135	0.1365
SGDM-20 0.01	0.7071	0.7571	0.6883	0.7211	0.414	0.4164
SGDM-20 0.1	0.8347	0.8	0.8596	0.8288	0.669	0.6710
SGDM-30 0.001	0.6306	0.5714	0.6481	0.6074	0.261	0.2631
SGDM-30 0.01	0.7010	0.6571	0.7204	0.6873	0.402	0.4036
SGDM-30 0.1	0.9173	0.9429	0.8941	0.9194	0.835	0.8358

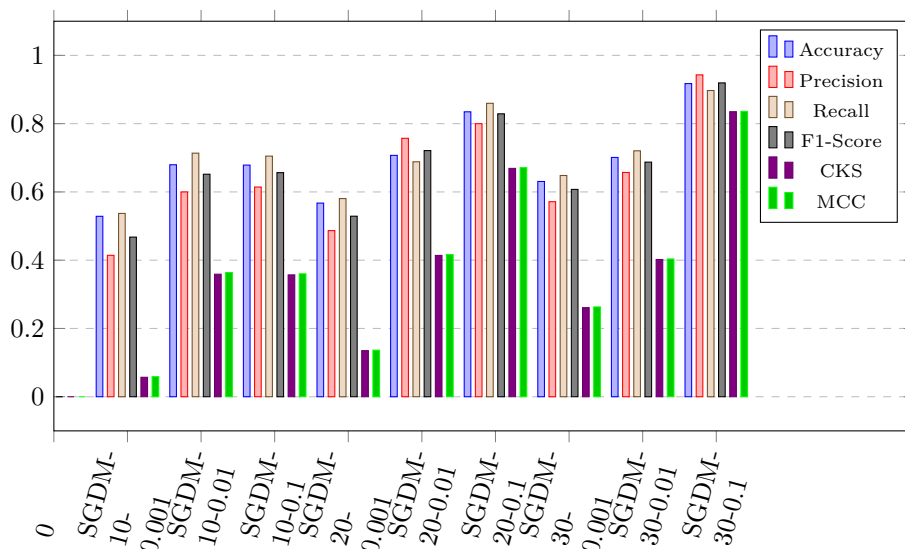
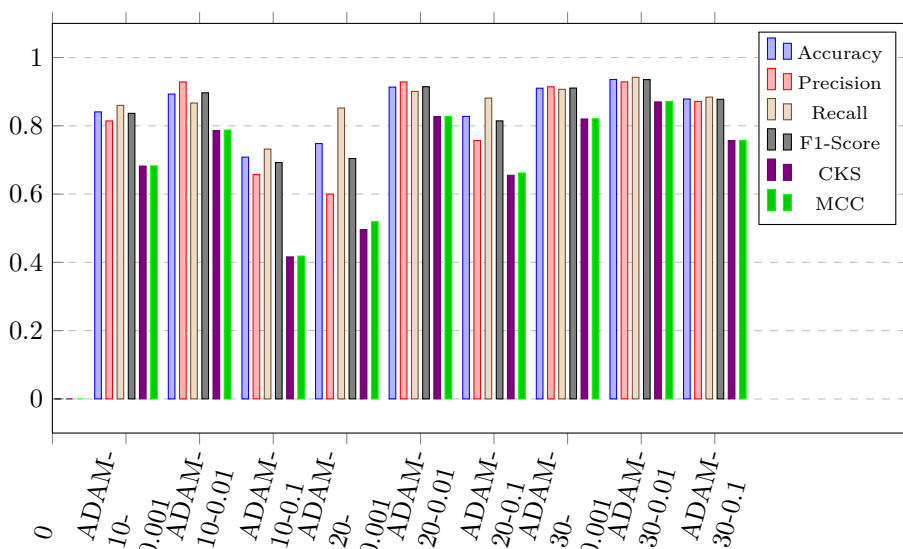
**Figura 5.2:** Resultados con el optimizador SGDM.

Tabla 5.5: Resultados con el optimizador ADAM

Configuración	Accuracy	Precision	Recall	F1-Score	CKS	MCC
SGDM-10 0.001	0.8408	0.8143	0.8599	0.8365	0.682	0.6826
SGDM-10 0.01	0.8929	0.9286	0.8667	0.8966	0.786	0.7877
SGDM-10 0.1	0.7082	0.6571	0.7318	0.6925	0.416	0.4185
SGDM-20 0.001	0.7480	0.6	0.8522	0.7042	0.496	0.5192
SGDM-20 0.01	0.9133	0.9286	0.9010	0.9146	0.827	0.8269
SGDM-20 0.1	0.8276	0.7571	0.8812	0.8145	0.655	0.6617
SGDM-30 0.001	0.9102	0.9143	0.9069	0.9106	0.82	0.8204
SGDM-30 0.01	0.9357	0.9286	0.9420	0.9353	0.87	0.8715
SGDM-30 0.1	0.8786	0.8714	0.8841	0.8777	0.757	0.7572

**Figura 5.3:** Resultados con el optimizador ADAM.

De acuerdo con los resultados presentados en las Tablas 5.3-5.5 y las Figuras 5.1-5.3, se pudo observar que el optimizador que presentó un comportamiento más inestable fue el de RMSPROP, mientras que el más estable en sus resultados fue el algoritmo ADAM. De la misma manera se pudo apreciar que los mejores resultados se obtuvieron con 30 épocas de entrenamiento y una tasa de aprendizaje de 0.01, al menos con el algoritmo de RMSPROP y ADAM.

Por lo tanto, los hiperparámetros para el entrenamiento de la red neuronal híbrida bidireccional se consideraron de acuerdo en las observaciones antes mencionadas y también en recomendaciones de diversos autores encontrados en la literatura, en este sentido, es importante mencionar que la sintonización que se empleó para el resto de los experimentos es la siguiente:

- Optimizador ADAM.
- Tasa de aprendizaje $\alpha = 0.01$.
- Mini lotes de 150 instancias por iteración.
- 30 épocas de entrenamiento.

5.2.2 Sintonización con diferentes capas recurrentes

En la literatura, existen distintas capas recurrentes además de la capa recurrente original (Vanilla RNN) que se diseñaron con el fin de mitigar las deficiencias encontradas en su funcionamiento. Dentro de las más conocidas en la literatura se encuentran la *Long-short Time Memory* (LSTM), *Gated Recurrent Unit* (GRU), *2 Stacked LSTM* (2SLSTM) y *Bidirectional LSTM* (BiLSTM). En este sentido, esta experimentación se enfocó en comparar la capa recurrente BiLSTM frente a las demás mencionadas para observar el comportamiento de clasificación de la FA.

Tabla 5.6: Resultados con diferentes capas recurrentes

Capa Recurrente	Accuracy	Precision	Recall	F1-Score	CKS	MCC
LSTM	0.8959	0.8857	0.9042	0.8948	0.792	0.7920
GRU	0.8612	0.9143	0.8266	0.8682	0.722	0.7266
2SLSTM	0.9010	0.9429	0.8701	0.9050	0.802	0.8049
BiLSTM	0.9357	0.9286	0.9420	0.9353	0.871	0.8715

De acuerdo a los resultados presentes en la Tabla 5.6 e ilustrados en la Figura 5.4, se puede notar que la capa recurrente con el menor rendimiento es la GRU y esto se debe a que este tipo de capa tienen mejor desempeño de clasificación en conjuntos de datos más pequeños dadas sus características de carecer de una puerta de salida y tener menos parámetros a ser entrenables. Posteriormente está la capa 2SLSTM se observa una similitud de rendimiento con la capa LSTM, demostrando que profundizar la red neuronal con múltiples capas LSTM no siempre generará mejores resultados e inclusive llega a incrementar el costo computacional de la red. En cuanto a la capa BiLSTM que se empleó en este trabajo de investigación, demostró tener el mejor rendimiento de clasificación frente a las demás capas recurrentes, esto se debe en parte a que la capa en cuestión tiene dos capas LSTM ocultas que no se encuentran apiladas como la capa 2SLSTM, sino que trabajan par a par haciendo la recurrencia en sentidos opuestos a manera de aprender las dependencias temporales tanto en $t - 1$ como en $t + 1$.

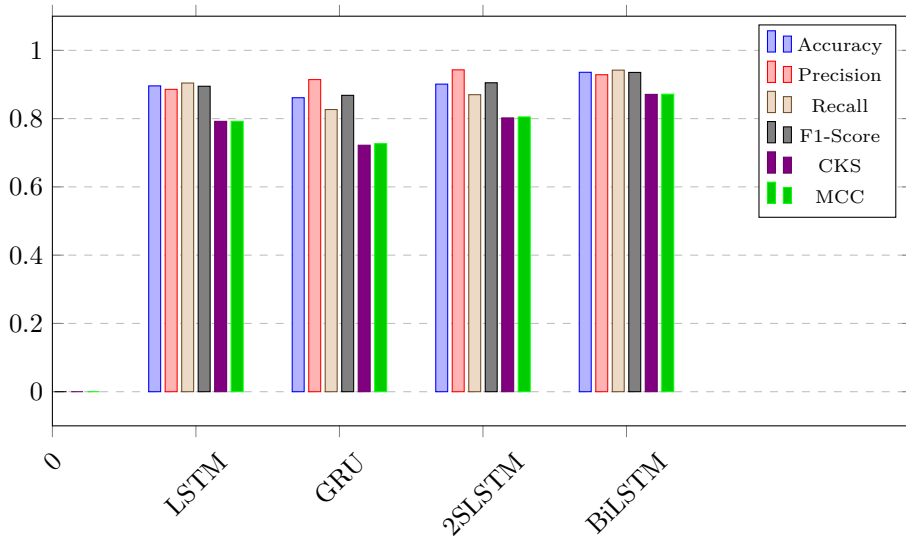


Figura 5.4: Resultados con diferentes capas recurrentes.

5.2.3 Sintonización con diferente cantidad de capas BiLSTM

Trabajar con diferente cantidad de capas permite a las redes neuronales puede mejorar en ciertos casos la capacidad de abstracción y la cantidad de características que genera a costa de un aumento en el costo computacional. Es por ello que se realizó una experimentación con distintos números de capas BiLSTM apiladas para conocer la capacidad de clasificación con una sola capa BiLSTM frente a 2, 3 y 4 capas.

Tabla 5.7: Resultados con diferente cantidad de capas BiLSTM

Capa Recurrente	Accuracy	Precision	Recall	F1-Score	CKS	MCC
4 capas	0.8633	0.9286	0.8213	0.8716	0.727	0.7328
3 capas	0.9235	0.9143	0.9314	0.9228	0.847	0.8471
2 capas	0.9071	0.9	0.913	0.9065	0.814	0.8144
1 capa	0.9357	0.9286	0.9420	0.9353	0.871	0.8715

Los resultados obtenidos que se aprecian en la Tabla 5.7 e ilustrados en la Figura 5.5, descartan que emplear demasiadas capas para profundizar y abstraer la red no asegura un mejor rendimiento e inclusive puede llegar a mermar la capacidad de clasificación de la red y aumentar el costo computacional del mismo. Por otra parte, se encontró una similitud en el rendimiento al momento de utilizar 1, 2 y 3 capas, siendo el primero el que tuvo los mejores porcentajes de clasificación, reafirmando la poca utilidad de profundizar una red neuronal y el aumento del costo computacional.

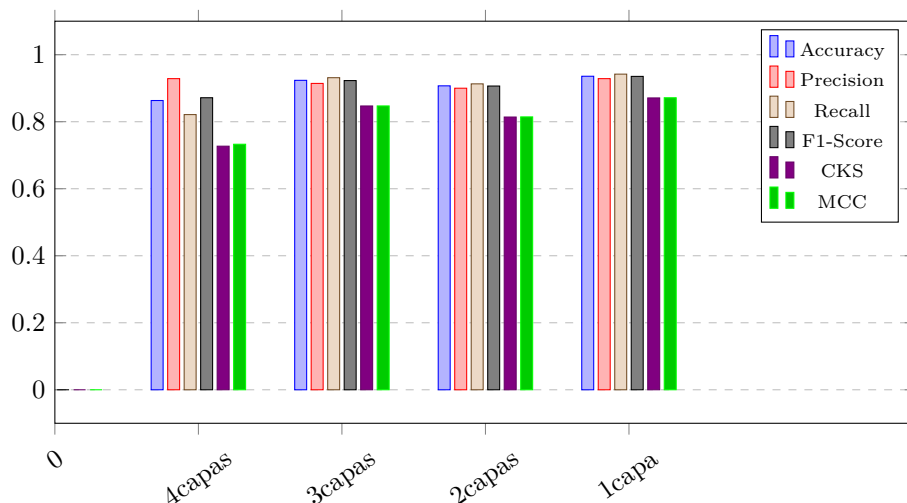


Figura 5.5: Resultados con distintas capas BiLSTM.

5.3 Comparación con diferentes tipos de características

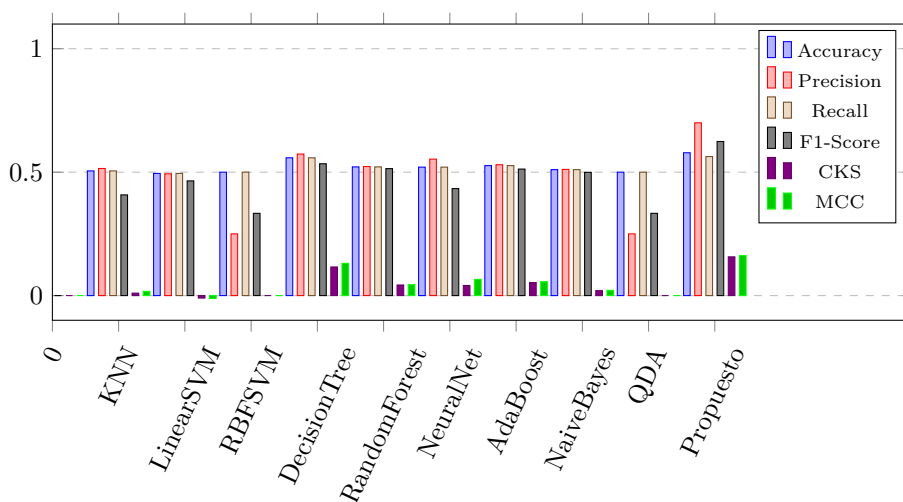
5.3.1 Datos en crudo

En esta experimentación se realizó utilizando las series temporales (ECG's) sin aplicar ningún tipo de preprocesamiento y/o extracción de características específicas. Los objetivos de esta experimentación en enfocan en observar el comportamiento de clasificación del la red neuronal con datos en crudo, además, se incluyeron en la experimentación algoritmos clásicos de aprendizaje automático como KNN, Máquinas de soporte vectorial lineal, Máquinas de soporte vectorial de base radial, Árboles de decisión, Bosques aleatorios, Red Neuronal, AdaBoost, Naive Bayes y el Análisis discriminante cuadrático, todos estos a manera de comparación frente al enfoque utilizado en este trabajo de investigación. Se podrán observar la matrices de confusión generadas por cada método/algoritmo utilizado en la experimentación (Anexo A) y posteriormente, en la Tabla 5.8 e ilustrado en la Figura 5.6 los porcentajes obtenidos por cada métrica.

El resumen cuantitativo para cada método y métrica empleada en la experimentación se presenta en la Figura 5.6 considerando las matrices de confusión (Figura A.4). De los resultados obtenidos, se destaca que el modelo híbrido tuvo el mejor rendimiento de clasificación, en contraste con los algoritmos clásicos que se puede observar en algunas métricas obtuvieron un desempeño por debajo de 0.30, demostrando que los algoritmos no alcanzaron una convergencia en el entrenamiento.

Tabla 5.8: Resultados de Rendimiento con Datos en crudo.

Algoritmo/Método	Accuracy	Precision	Recall	F1-Score	CKS	MCC
Nearest Neighbors	0.5051	0.5149	0.5051	0.4079	0.0	0.0
Linear SVM	0.4949	0.4934	0.4949	0.4646	0.0	0.0
RBF SVM	0.5	0.25	0.5	0.3333	0.0	0.0
Decision Tree	0.5582	0.5734	0.5582	0.534	0.0	0.0
Random Forest	0.5214	0.5228	0.5214	0.5143	0.0	0.0
Neural Net	0.5204	0.5529	0.5204	0.4334	0.0	0.0
AdaBoost	0.5265	0.53	0.5265	0.5125	0.0	0.0
Naive Bayes	0.5102	0.5112	0.5102	0.4994	0.0	0.0
QDA	0.5	0.25	0.5	0.3333	0.0	0.0
Propuesto	0.5786	0.7000	0.5632	0.6242	0.0	0.0

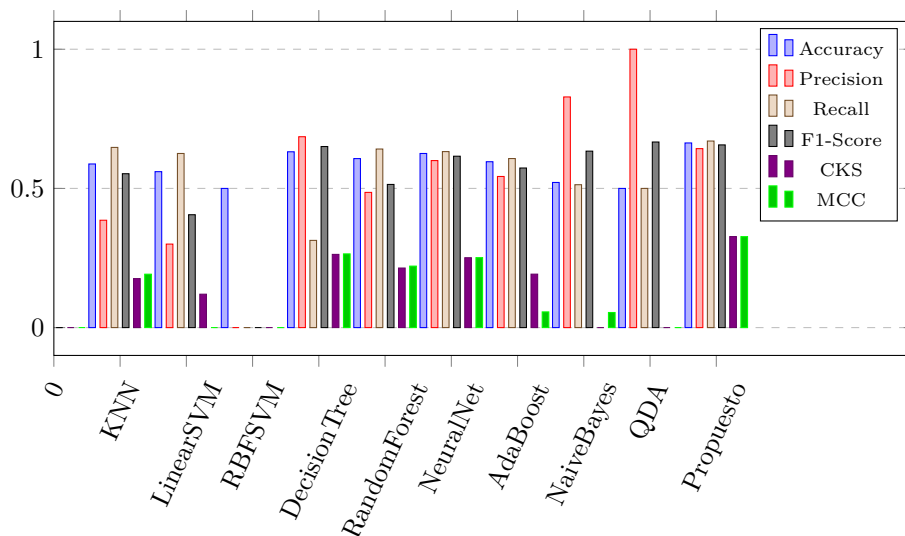
**Figura 5.6:** Resultados con datos en crudo

5.3.2 Características estadísticas

Esta experimentación se enfocó en conocer el comportamiento de la red neuronal frente a los algoritmos clásicos con respecto a características extraídas de manera clásica conocidas como características estadísticas, en específico, la media, desviación estándar, mediana, asimetría, curtosis, raíz media cuadrática media y la proporción. En la Tabla 5.9 se podrán observar una gráfica resumida por medio de métricas derivadas de la matriz de confusión (Figura A.5) el rendimiento de clasificación obtenido por cada método/algoritmo evaluado y visualizado de manera gráfica en la Figura 5.7.

Tabla 5.9: Resultados con características estadísticas

Algoritmo/Método	Accuracy	Precision	Recall	F1-Score	CKS	MCC
Nearest Neighbors	0.5878	0.3857	0.6473	0.5528	0.176	0.1919
Linear SVM	0.5602	0.3	0.6255	0.4055	0.120	0.0
RBF SVM	0.5	0	0	0	0.0	0.0
Decision Tree	0.6316	0.6857	0.3133	0.6505	0.263	0.2648
Random Forest	0.6071	0.4857	0.6415	0.5143	0.214	0.2209
Neural Net	0.6255	0.6	0.6323	0.6157	0.251	0.2513
AdaBoost	0.5959	0.5429	0.6073	0.5733	0.192	0.0564
Naive Bayes	0.5214	0.8286	0.5133	0.6339	0.0	0.0543
QDA	0.5	1	0.5	0.6667	0.0	0.0
Propuesto	0.6633	0.6429	0.6702	0.6563	0.327	0.3268

**Figura 5.7:** Resultados con características estadísticas

Del resumen cuantitativo se pueden observar los puntajes obtenidos, pudiéndose observar un ligero aumento en el rendimiento de algoritmos como Random Forest, Neural Net y la red neuronal híbrida. A pesar del ligero aumento en el rendimiento de clasificación, estas características aún no logran ser del todo significativas y discriminatorias dado que en el momento de extraer dichas características, los ECG se encontraban en el dominio caótico, por ende, era de esperarse que los algoritmos no logran clasificar de manera satisfactoria el conjunto de prueba.

5.3.3 Características de Tiempo-Frecuencia

Para este caso, de acuerdo a las características detalladas en la sección 3.1, se trabajaron como datos de entrada la entropía espectral y la frecuencia instantánea. Por ende, se trabajó con una matriz de secuencias para albergar ambas características, teniendo que modificar la dimensionalidad en la capa de entrada de secuencias, quedando la nueva red modificada de la siguiente manera (Figura 5.8).

1	''	Sequence Input	Sequence input with 2 dimensions
2	''	BiLSTM	BiLSTM with 100 hidden units
3	''	Fully Connected	2 fully connected layer
4	''	Softmax	softmax
5	''	Classification Output	crossentropyex

Figura 5.8: Descripción de las capas de la red modificada.

En este sentido, para esta experimentación se utilizaron las características de tiempo-frecuencia propuestas en este trabajo de investigación para observar el comportamiento y demostrar un aumento en el rendimiento de clasificación de los algoritmos. De nueva cuenta se generó una gráfica resumiendo a partir de métricas derivadas de la matriz de confusión el rendimiento de clasificación obtenido de la experimentación.

Tabla 5.10: Resultados del rendimiento con características de Tiempo-Frecuencia.

Algoritmo/Método	Accuracy	Precision	Recall	F1-Score	CKS	MCC
Nearest Neighbors	0.8199	0.8199	0.8199	0.8199	1.0	1.0
Linear SVM	0.574	0.5794	0.574	0.5665	0.0	0.0
RBF SVM	0.5	0.25	0.5	0.3333	0.0	0.0
Decision Tree	0.6357	0.636	0.6357	0.6355	0.0	0.0
Random Forest	0.6168	0.6236	0.6168	0.6115	0.0	0.0
Neural Net	0.6724	0.757	0.6724	0.6431	0.0	0.0
AdaBoost	0.6393	0.6425	0.6393	0.6372	0.0	0.0
Naive Bayes	0.6061	0.6093	0.6061	0.6033	0.0	0.0
QDA	0.5378	0.7067	0.5378	0.419	0.0	0.0
Propuesto	0.9357	0.9286	0.9420	0.9353	1.0	1.0

De acuerdo a lo presentado en la Figura 5.9, en los resultados obtenidos se destaca que el modelo híbrido tuvo el mejor desempeño de clasificación, superando las métricas utilizadas el 0.90 en el rendimiento de clasificación, a diferencia del resto de los algoritmos donde se puede ver que tuvieron un rendimiento por debajo de 0.80, debido a que las características extraídas están en el dominio del tiempo y la frecuencia por los cuales no fueron diseñados dichos algoritmos. En adición a lo anterior, también se puede constatar que los algoritmos clásicos tienen un problema en particular que es conocido como la meseta de rendimiento, el cual consiste en que a mayor carga de datos para entrenar y evaluar dichos algoritmos, el rendimiento se trunca.

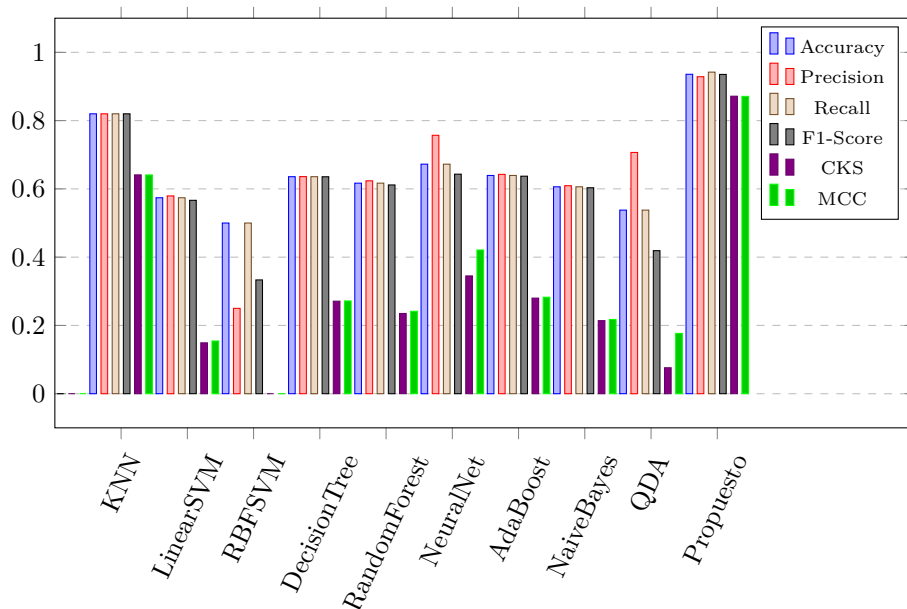


Figura 5.9: Resultados con características de TF

5.3.3.1 Costo computacional

En adición de las experimentaciones con las diferentes características y otro de los objetivos de esta investigación, se calculó el costo computacional de cada algoritmo al momento de clasificar el conjunto de pruebas.

Tabla 5.11: Costo computacional de clasificación

Algoritmo/Método	Con datos en crudo (en Segundos)	Con características estadísticas (en Segundos)	Con características de tiempo-frecuencia (en Segundos)
Nearest Neighbors	9.96	0.03	1.95
Linear SVM	13.99	0.26	5.27
RBF SVM	64.28	0.30	8.14
Decision Tree	0.08	0.002	0.008
Random Forest	0.08	0.003	0.009
Neural Net	0.39	0.008	0.02
AdaBoost	1.17	0.01	0.07
Naive Bayes	0.18	0.002	0.01
QDA	9.73	0.005	0.06
Propuesto	6.41	0.06	0.35

Conforme a lo observado en la Tabla 5.11, se midió el costo computacional de los algoritmos y arquitecturas consideradas. La experimentación se realizó clasificando el conjunto de datos de prueba sin procesar y el conjunto de datos de prueba con características de tiempo-frecuencia. De acuerdo a los resultados obtenidos, se observó que la arquitectura propuesta presentaba un menor costo computacional en comparación con los algoritmos clásicos de aprendizaje automático. A primera vista se tiene la impresión de tener un menor costo computacional en el caso del Arbol de Decisión, sin embargo, el rendimiento observado en la Tabla 5.10 descarta la posibilidad del correcto funcionamiento en dicho algoritmo al momento de clasificar la Fibrilación Auricular.

5.4 Comparación del modelo propuesto con otros modelos neuronales

Como parte final de las experimentaciones, se realizó una comparación de la red neuronal híbrida con distintas redes neuronales convolucionales como SqueezeNet, DarkNet19 y AlexNet. Lo que caracteriza a estas redes es que a diferencia de otras arquitecturas convolucionales, estas contienen menos parámetros entrenables lo que se traduce a redes convolucionales ligeras. Ahora bien, el conjunto de datos de TF es compatible solo para la red neuronal híbrida debido a que son datos en una dimensión, por lo tanto, para entrenar y clasificar la FA con las redes convolucionales, lo que se hizo fue generar un nuevo dataset de imágenes espectrográficas a partir de las mismas señales de la base de datos.

Tabla 5.12: Resultados con distintos tipos de CNN

Capa Recurrente	Accuracy	Precision	Recall	F1-Score	CKS	MCC
SqueezeNet	0.5	1	0.5	0.6667	0	0
DarkNet19	0.8959	0.8776	0.91	0.8940	0.792	0.7924
AlexNet	0.5	1	0.5	0.6667	0	0
Propuesto	0.9357	0.9286	0.9420	0.9353	0.871	0.8715

El resumen cuantitativo que se muestra en la Tabla 5.12 y visualizado gráficamente la Figura 5.10 denota en primera instancia el bajo rendimiento de clasificación en las redes SqueezeNet y AlexNet debido a la poca generalización que se logró durante el periodo de entrenamiento. Lo anterior demuestra que a diferencia de las redes neuronales recurrentes, las redes neuronales convolucionales necesitan de una alta capacidad de abstracción dependiendo del contexto en el que se está aplicando para poder generalizar de mejor manera la información. Por otra parte, la red neuronal DarkNet19 obtuvo un rendimiento superior a las redes mencionadas anteriormente y un desempeño similar a la red neuronal híbrida, la razón por la que DarkNet19 tuvo mejor desempeño se debe a una mayor profundidad en sus capas, teniendo un

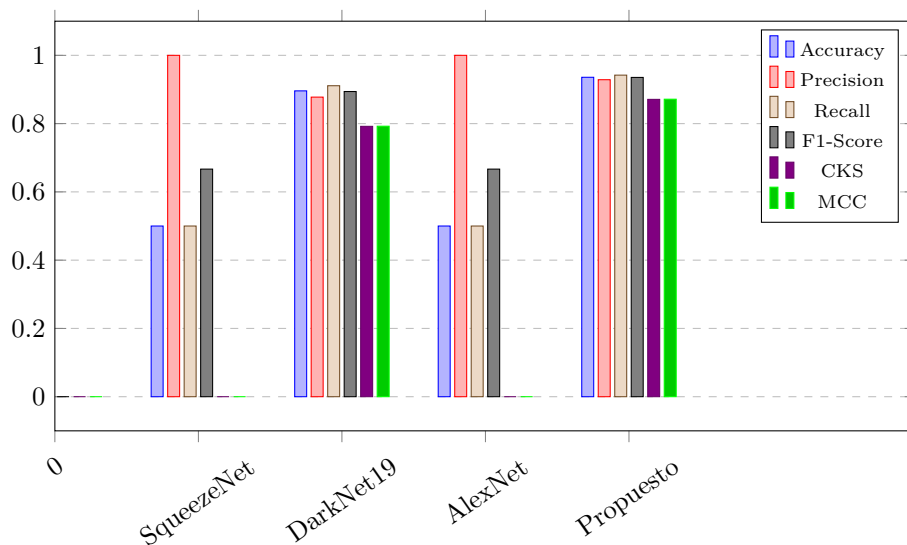


Figura 5.10: Resultados con distintos tipos de CNN.

total de 19 capas convolucionales frente a las 5 convolucionales y 3 capas densas de SqueezeNet y AlexNet. Sin embargo, la red neuronal propuesta tuvo un desempeño superior en la mayoría de las métricas con las que se evaluó.

5.5 Discusión

En este capítulo se detallaron los experimentos para evaluar cuantitativamente la metodología desarrollada a través de múltiples experimentos. Los cuales, fueron diseñados para encontrar la mejor sintonización de hiperparámetros de la Red Neuronal Híbrida Bidireccional, demostrar el aumento en el rendimiento de clasificación y reducción del costo computacional con los momentos de Tiempo-Frecuencia y un mejor desempeño frente a otros modelos neuronales. Como resultado se obtuvo un incremento de al menos 30% al emplear características de Tiempo-Frecuencia y una reducción considerable del costo computacional en la etapa de entrenamiento y clasificación respectivamente.

Conclusiones

En este capítulo se realiza un análisis de las conclusiones a las que se llegaron después de la evaluación de la metodología propuesta y de estudiar los resultados obtenidos. También se realiza un análisis de trabajos futuros.

6.1 Objetivos y alcances logrados

En la Tabla 6.1, se muestran las actividades realizadas en relación a los objetivos y en la Tabla 6.2 los alcances planteados para el desarrollo de la tesis.

Tabla 6.1: Objetivos específicos.

Objetivo	Actividades
Analizar y comprender el paradigma de las Redes Neuronales Recurrentes Profundas, con mayor énfasis en la arquitectura Long Short-Term Memory.	Se realizó el estudio del marco conceptual y el estado del arte de las Redes Neuronales Recurrentes, se redactó un reporte del Estado del Arte.
Analizar y comprender las señales electrocardiográficas.	Se realizó el estudio del marco conceptual de las señales electrocardiográficas.
Extraer características de tiempo-frecuencia a partir de espectrogramas.	Se extrajeron características como la Frecuencia Instantánea que se encuentra en el dominio del tiempo y la Entropía Espectral que se encuentra en el dominio de la frecuencia.
Definir la arquitectura de RNN.	La arquitectura que se definió en esta investigación fue una Red Híbrida Bidireccional .
Entrenar y evaluar la RNN utilizando las series de tiempo en crudo.	Se realizó el entrenamiento y evaluación de la red híbrida con datos en crudo para conocer el comportamiento de las señales caóticas en crudo.
Modificar la arquitectura de RNN considerando los momentos de tiempo-frecuencia extraídos de los espectrogramas.	Se modificó la dimensionalidad de la capa de entradas secuenciales a manera de procesar simultáneamente las características de tiempo-frecuencia.
Entrenar y evaluar la RNN con funciones de tiempo-frecuencia elegidas, y publicar los resultados obtenidos.	Con la red híbrida modificada, se realizó el entrenamiento y prueba de dichas características, y se escribió un artículo científico con los resultados obtenidos.

Tabla 6.2: Alcances.

Alcances	Actividades
Entrenar con diferentes tipos de datos para observar el desempeño en cada uno de ellos.	Se evaluó la red híbrida con datos en crudo, características estadísticas y las características de tiempo-frecuencia.
Modificar el modelo RNN para entrenarlo con características de tiempo-frecuencia.	La capa de entradas secuenciales se modificó para procesar 2 características simultáneamente.
Mantener un costo computacional óptimo con el modelo propuesto.	El modelo propuesto logró reducir al menos un 90% del costo computacional empleando características de tiempo-frecuencia, con respecto a los datos en crudo.
Determinar con al menos un 80 % en el rendimiento general de clasificación.	El modelo propuesto logró en la mayoría de sus métricas, obtener un rendimiento superior al 90%.
Evaluar los resultados obtenidos con base en las métricas definidas en la literatura.	Se emplearon métricas de evaluación cuantitativa como Accuracy, Precision, Recall, F1-Score, y otras encontradas en la literatura como el puntaje Kappa de cohen y la Matriz de correlación de Mathews.

6.2 Resultados del trabajo

6.2.1 Productos

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

1. Reporte del estado del arte: documento en el que se presenta un resumen de diversas publicaciones recientes consideradas a partir de los núcleos temáticos contemplados en la investigación como lo es redes neuronales recurrentes, procesamiento digital de señales, métricas y bases de datos.
2. Extracción de características de tiempo-frecuencia a partir de señales electrocardiográficas.
3. Diseño e implementación de la Red Neuronal Híbrida Bidireccional: códigos fuente para la conformación de las capas, sintonización de hiperparámetros, entrenamiento y pruebas de la Red con características de tiempo-frecuencia.
4. Artículo indizado en Scopus para el congreso internacional WITCOM 2021 (Figura B.1 y B.2).
5. Artículo para el congreso COMIA 2022 (Figura B.3 y B.4).
6. Artículo indizado en Scopus para el congreso internacional WITCOM 2022 (Figura B.5 y B.6).

7. Artículo para el congreso mexicano de reconocimiento de patrones (MCPR 2021) (Figura B.7 y B.8).

6.2.2 Aportaciones

Como aportaciones al tema de clasificación de la fibrilación auricular y al marco teórico se tiene

- Características de Tiempo-Frecuencia extraídas a partir de los espectrogramas.
- Red Neuronal Híbrida Bidireccional para la clasificación de los momentos de Tiempo-Frecuencia.

6.3 Conclusiones

En esta tesis se presentó un nuevo enfoque para la clasificación de la Fibrilación Auricular utilizando momentos de Tiempo-Frecuencia, detallando formalmente el método propuesto que parte desde la transformación al dominio de tiempo-frecuencia de las señales electrocardiográficas, la extracción de características y reducción de dimensionalidad de información, hasta las definiciones matemáticas de las capas que conforman la red neuronal híbrida bidireccional.

En primer instancia, las experimentaciones realizadas en la sección 5.2, demostraron que la mejor configuración de los hiperparámetros para poder entrenar la red neuronal híbrida fue con el Optimizador ADAM, una tasa de aprendizaje $\alpha = 0.01$, un total de 30 épocas y una sola capa BiLSTM.

Por otro lado, se evidenció que el correcto tratamiento de las señales mejora el proceso de entrenamiento de los algoritmos de clasificación utilizados durante la investigación, sugiriendo además que la Arquitectura Neural Híbrida propuesta obtuvo el mejor desempeño para la clasificación de arritmias, ya que aunque algoritmos como k -Vecinos más cercanos, los Árboles de Decisión, a pesar de caracterizarse como algoritmos de clasificación multiclase, no logran superar el desempeño de la Arquitectura Neural Híbrida propuesta al momento de analizar señales de ECG, debido a la mencionada meseta de rendimiento que reduce el desempeño de estos.

De igual manera, se comprobó que la red neuronal híbrida al compararse con distintas enfoques de capas recurrentes, distintas capas BiLSTM y con las redes neuronales convolucionales, obtuvo el mejor desempeño en la clasificación de la FA; asegurando ser la mejor red neuronal para clasificar las características de TF. Por otro lado, la implementación de equipos de cómputo de alto rendimiento es de gran ayuda en cuanto al procesamiento de grandes cantidades de señales electrocardiográficas, como se ejemplificó en los experimentos realizados.

6.4 Trabajos futuros

Como trabajo futuro derivado de esta investigación, se contempla la implementación de la metodología propuesta en un sistema embebido de bajo consumo con el fin de concebir un sistema portátil y remoto, además de mejorar los tiempos de procesamiento, de tal forma que el sistema pueda trabajar lo más cerca posible de tiempo real. En cuanto al tema de conjunto de datos, se considera generar un conjunto de datos propio de señales sanas y con fibrilación auricular para mejorar el entrenamiento y conseguir que la red neuronal híbrida pueda generalizar aún mejor la información, lo que se traduciría en un mejor desempeño de clasificación.

APÉNDICE A

Matrices de confusión

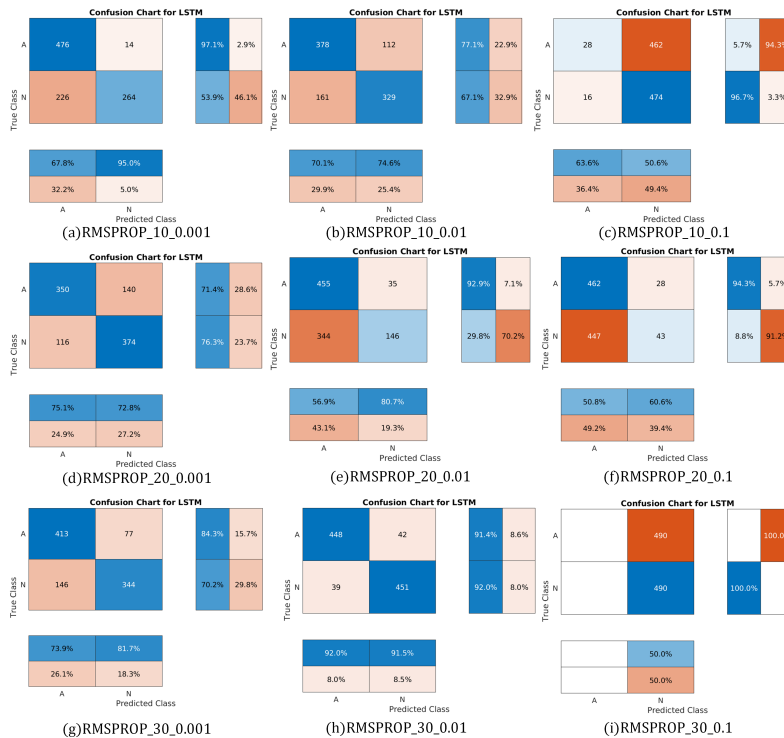


Figura A.1: Matrices de confusión obtenidas con el optimizador RMSPROP.



Figura A.2: Matrices de confusión obtenidas con el optimizador SGDM.

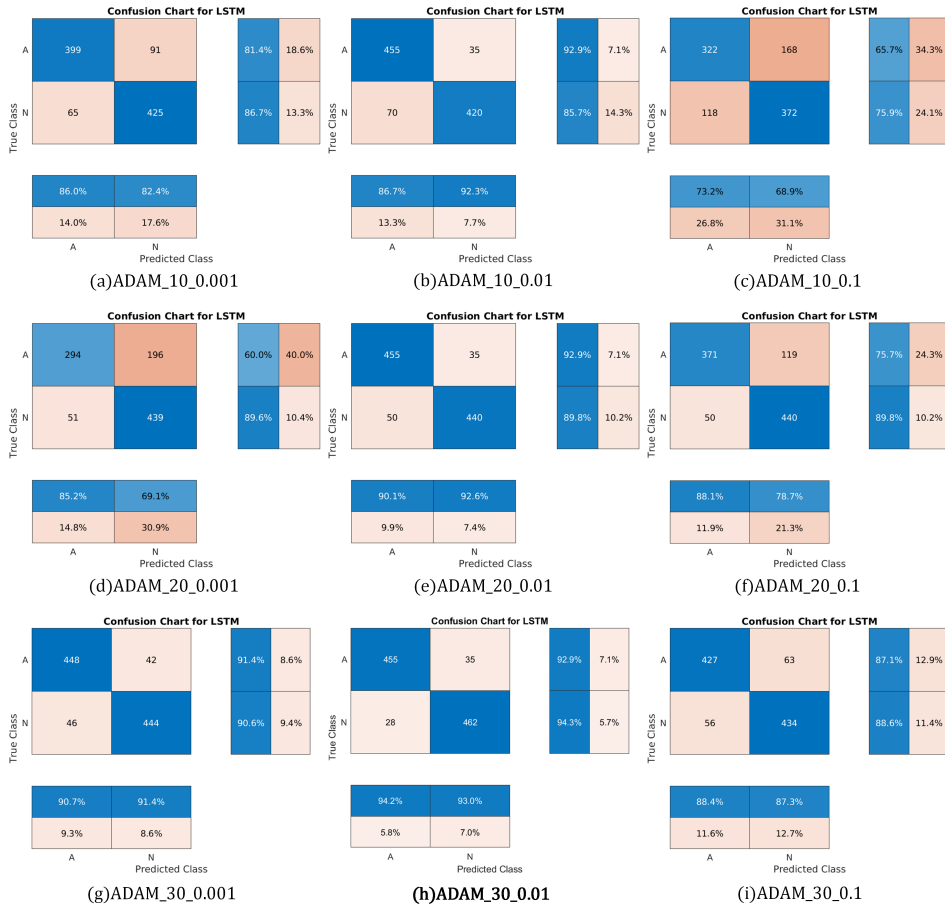


Figura A.3: Matrices de confusión obtenidas con el optimizador ADAM.

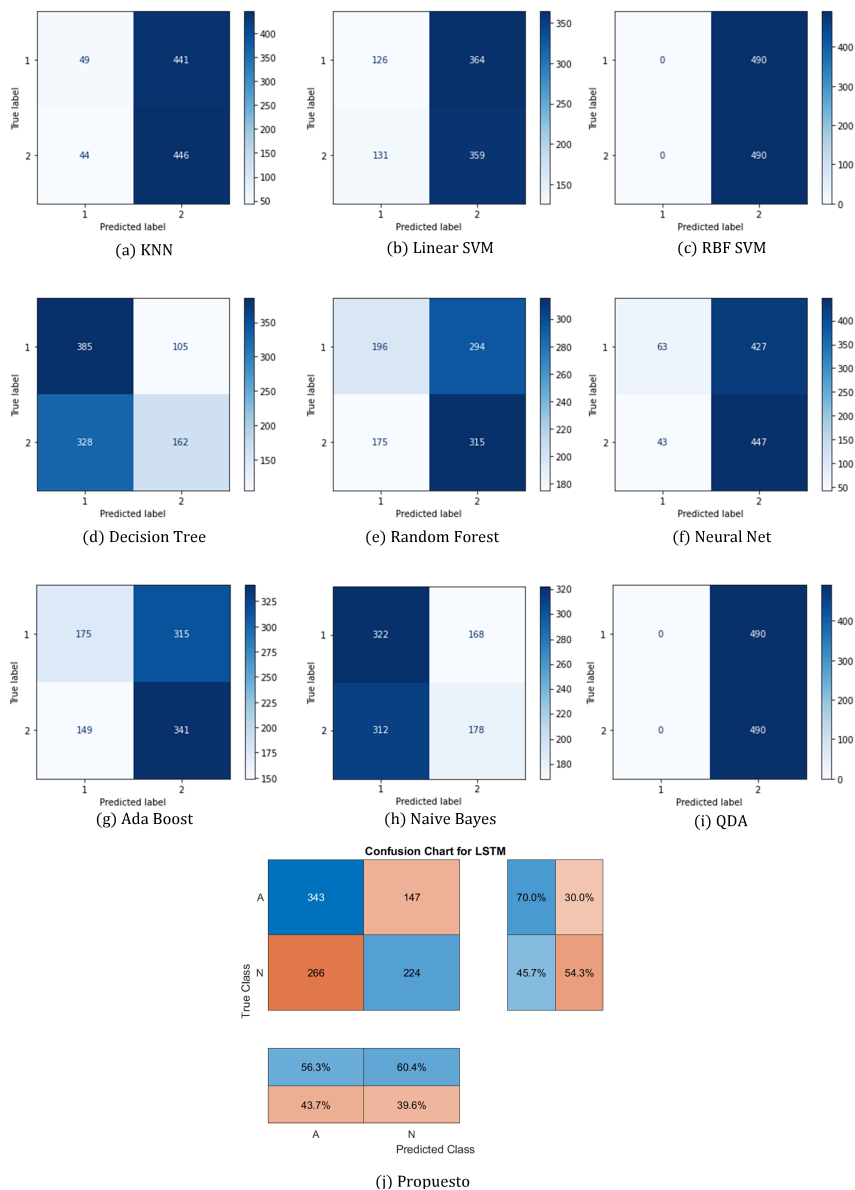


Figura A.4: Matrices de confusión obtenidas con datos en crudo.

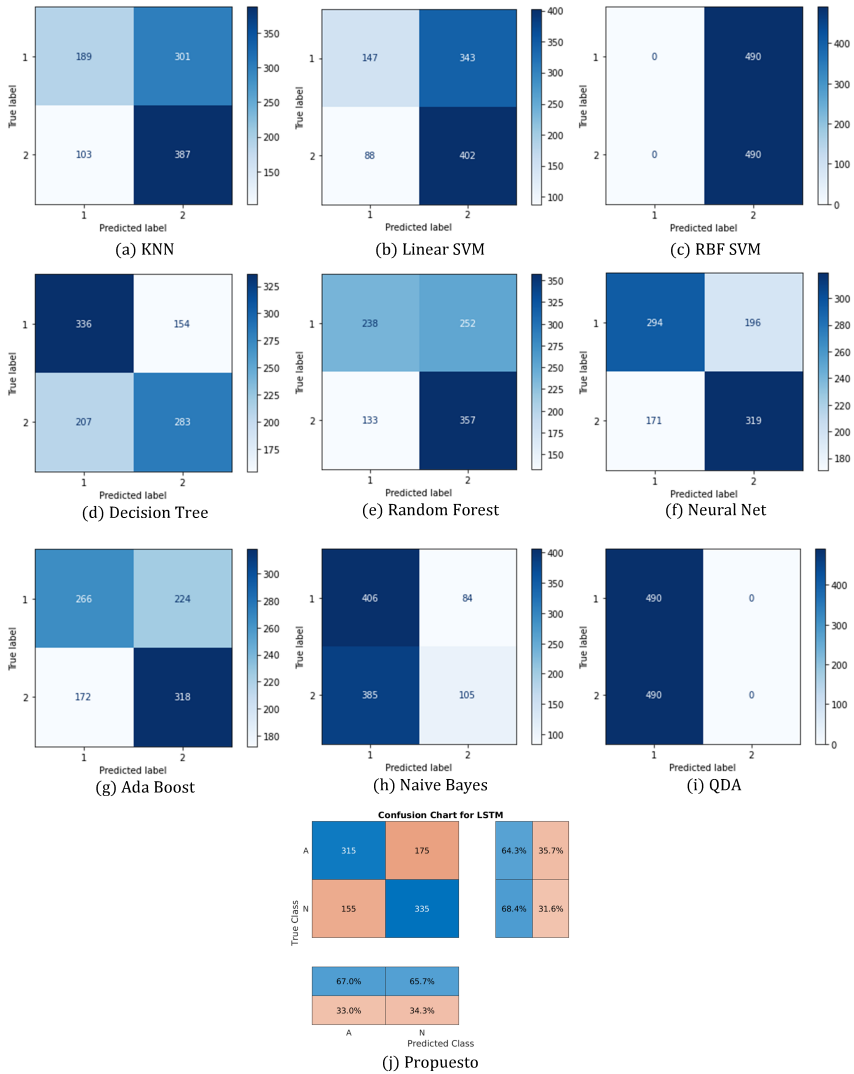


Figura A.5: Matrices de confusión obtenidas con características estadísticas.

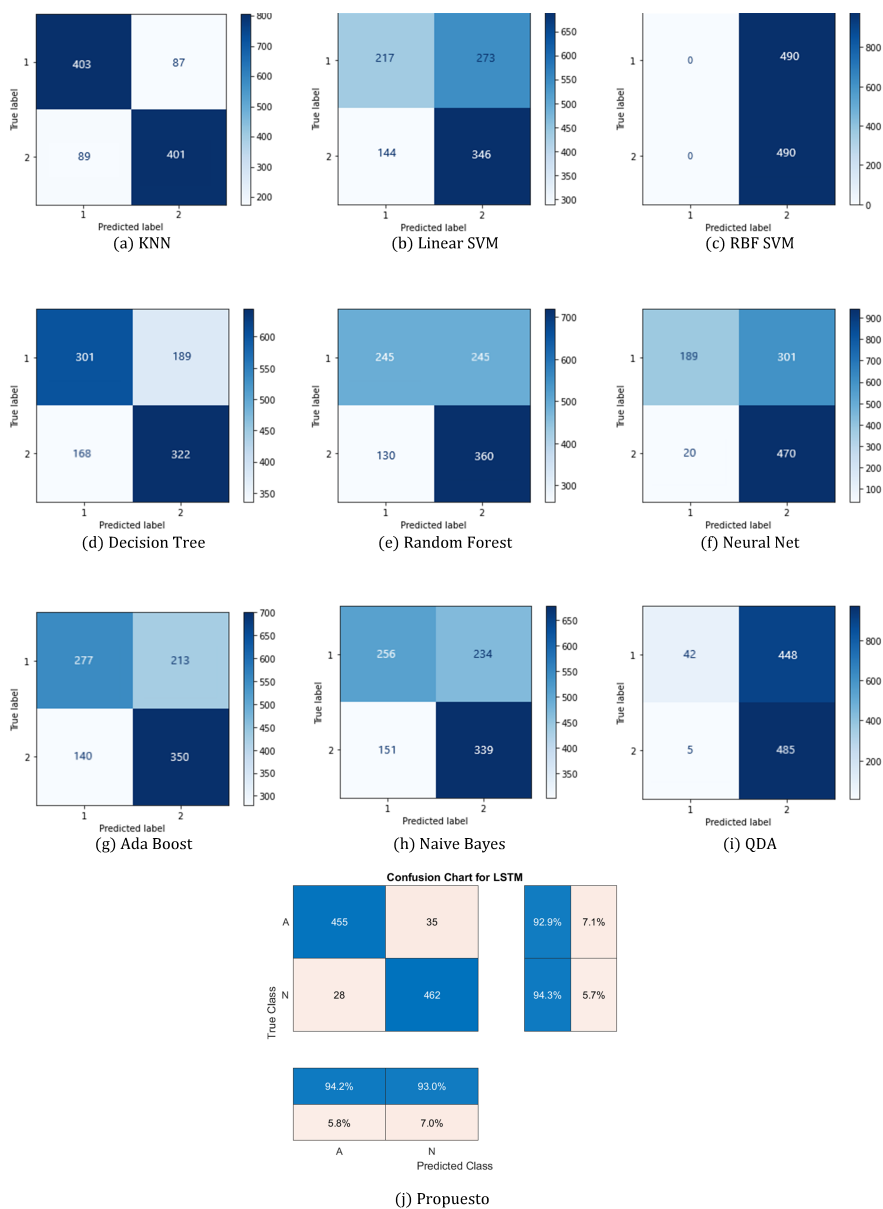


Figura A.6: Matrices de confusión obtenidas con características de TF.

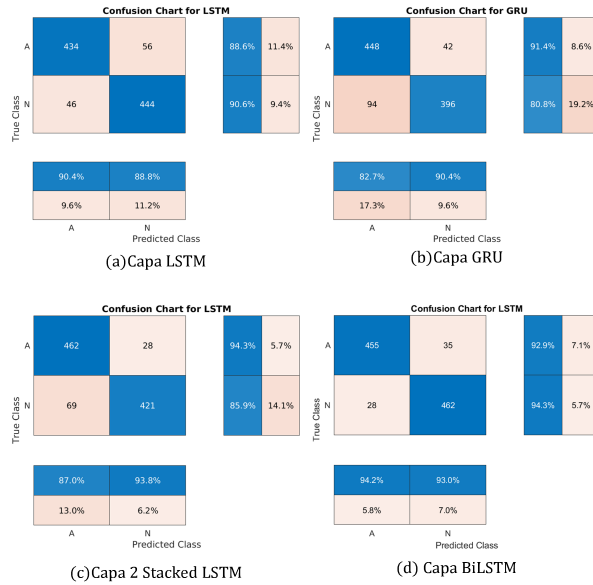


Figura A.7: Matrices de confusión obtenidas con diferentes capas recurrentes.

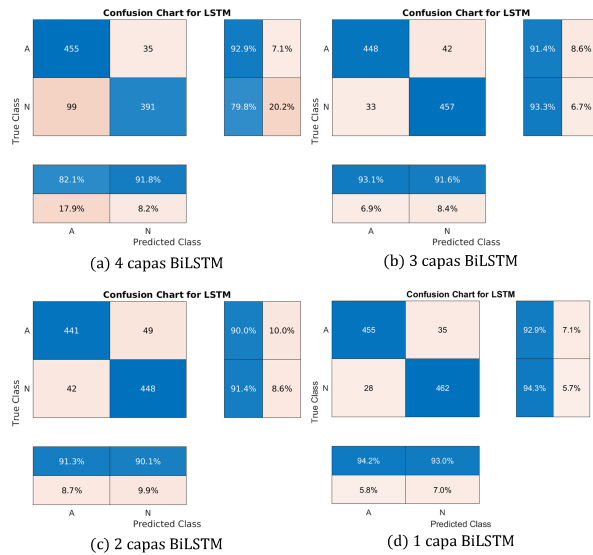


Figura A.8: Matrices de confusión obtenidas con diferentes cantidad de capas BiLSTM.

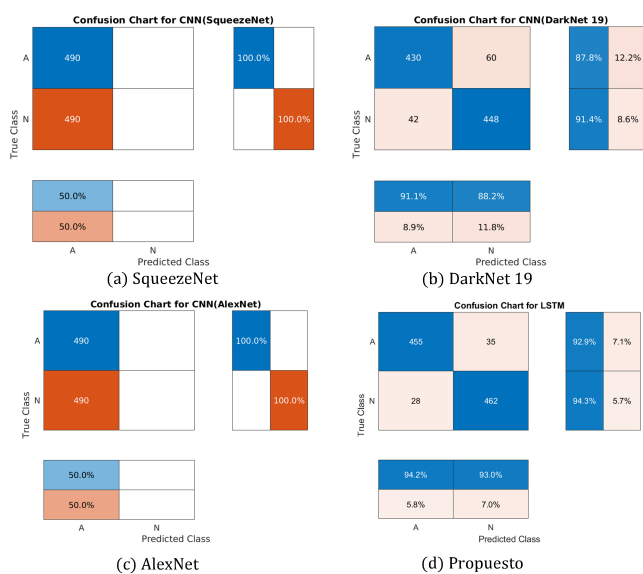


Figura A.9: Matrices de confusión obtenidas con diferentes redes neuronales convolucionales.

APÉNDICE **B**

Producción científica

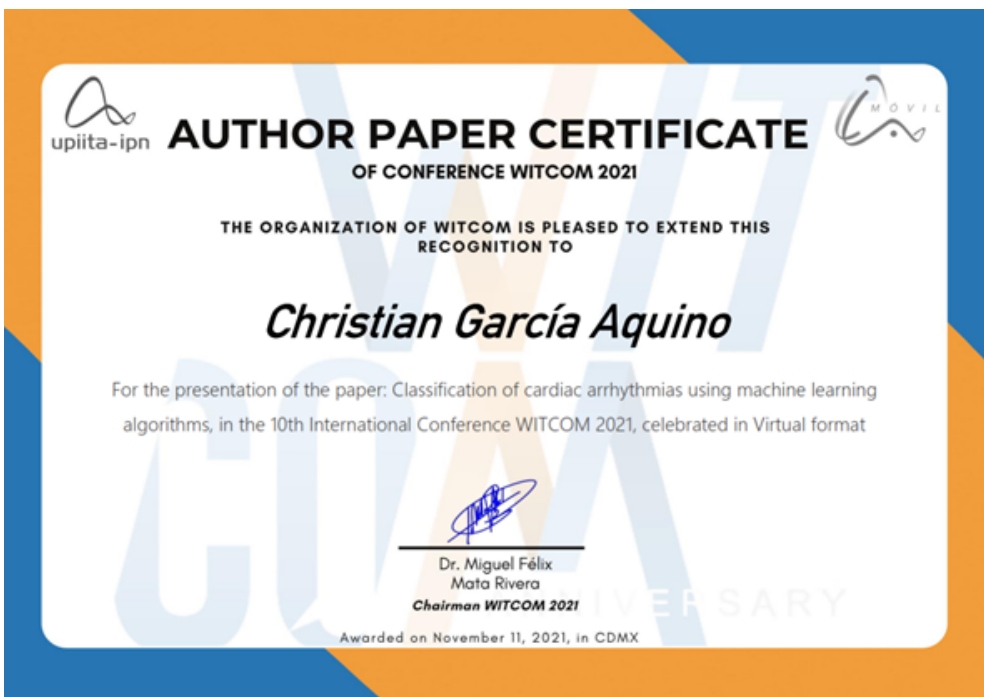


Figura B.1: Constancia de autor del artículo publicado en el congreso internacional WITCOM 2021.



Classification of Cardiac Arrhythmias Using Machine Learning Algorithms

Christian García-Aquino¹ , Dante Mújica-Vargas¹ ,
and Manuel Matuz-Cruz² 

¹ Tecnológico Nacional de México, CENIDET, Cuernavaca, Mexico
m21ce012@cenidet.tecnm.mx

² Tecnológico Nacional de México, Campus Tapachula, Mexico City, Mexico

Abstract. This article proposes a method to classify cardiac arrhythmias using feature extraction and dimensionality reduction techniques. The experiment was carried out with the QRS complexes of the electrocardiographic signals, which are part of the Physionet MIT-BIH arrhythmia database. The machine learning algorithms used to perform the classification were k Nearest Neighbors, Linear SVM, RBF SVM, Decision Tree, Random Forest, Neural Net, AdaBoost, Naive Bayes and QDA. To measure the efficiency and quality of the proposed method, the metrics of Accuracy, Precision, Recall, F1 Score, as well as the Cohen Kappa score and the Mathews correlation coefficient were used. The results obtained show a better classification performance in the neural network with 96.33%, 96.33%, 96.33%, 96.32%, 1.0 and 1.0 of the aforementioned metrics.

Keywords: Arrhythmia · ECG · Feature extraction · Dimensionality reduction · QRS complex

1 Introduction

In recent years, cardiovascular diseases have been the leading cause of deaths around the world. Among them, cardiac arrhythmias are very common conditions that require novel approaches in terms of early detection and effective treatments, and thus prevent heart problems. Cardiac arrhythmia, on the other hand, is a series of disturbances in which the heart rate varies between excessively fast, excessively slow or irregular [1].

Electrocardiography records the electrical activity of the heart, which is the most used technique in the diagnosis of patients during clinical practice [2]. The signals present in the electrocardiogram (ECG) present morphological and temporal characteristics of multiple components, which provide an improvement in the diagnosis and treatment planning processes, as well as in the control and prediction of cardiovascular diseases such as: arrhythmias, acute coronary syndromes, among other congenital or genetically unknown cardiovascular diseases, as studied in [3], [4] and [5].

© Springer Nature Switzerland AG 2021
M. F. Mata-Rivera and R. Zagal-Flores (Eds.): WITCOM 2021, CCIS 1430, pp. 174–185, 2021.
https://doi.org/10.1007/978-3-030-89586-0_14

Figura B.2: Primera página del artículo publicado en el congreso internacional WITCOM 2021.



Figura B.3: Constancia de autor del artículo publicado en el congreso COMIA 2022.

ISSN 1870-4069

Implementación de modelos de red neuronal profundos en GPUs embebidas para la clasificación de secuencias de fibrilación auricular

Christian García-Aquino¹, Dante Mújica-Vargas¹,
Juan Gabriel González-Serna¹, Mammel Matuz-Cruz²

¹ Tecnológico Nacional de México,
Centro Nacional de Investigación y Desarrollo Tecnológico,
México

² Tecnológico Nacional de México,
México

m21ce012@cenidet.tecnm.mx

Resumen. En este artículo se implementaron arquitecturas basadas en aprendizaje profundo en tarjetas embebidas. La experimentación se realizó en una tarjeta embebida Nvidia Jetson Nano y se emplearon las señales ECG de la base de datos CinC Challenge 2017 de PhysioNet, de las cuales, se extrajeron características de Tiempo-Frecuencia para el entrenamiento de los modelos. Los modelos de aprendizaje profundo que se emplearon para las tareas de clasificación fue una CNN SqueezeNet y una RNN del tipo BiLSTM. Para cuantificar el rendimiento de clasificación de los modelos, se obtuvo la exactitud y el error promedio por cada clase. En donde se sugiere que el modelo convolucional tiene un menor costo computacional, sin embargo, en tareas de clasificación se sugiere que el modelo recurrente tuvo un mayor rendimiento y una mejor propagación del error frente al modelo convolucional.

Palabras clave: Jetson nano, physionet, squeezeNet, BiLSTM, tiempo-frecuencia.

Implementation of Deep Neural Network Models on Embedded GPUs for Classification of Atrial Fibrillation Sequences

Abstract. In this article, architectures based on deep learning in embedded cards were implemented. The experimentation was carried out on an embedded Nvidia Jetson Nano card and ECG signals from the PhysioNet CinC Challenge 2017 database were used, from which Time-Frequency characteristics were extracted for model training.

pp. 129–142; rec. 2022-04-06; acc. 2022-05-11 129 *Research in Computing Science* 151(5), 2022






Figura B.4: Primera página del artículo publicado en el congreso COMIA 2022.



Figura B.5: Constancia de autor del artículo publicado en el congreso internacional WITCOM 2022.



Implementation of Time-Frequency Moments for the Classification of Atrial Fibrillation Sequences Through a Bidirectional Long-Short Term Memory Network

Christian García-Aquino¹ , Dante Mújica-Vargas¹ ,
Manuel Matuz-Cruz² , Nimrod Gonzalez-Franco¹ ,
and Gabriel González-Serna¹ 

¹ Tecnológico Nacional de México, CENIDET, Cuernavaca, Mexico
{m21ce012,dante_mv}@cenidet.tecnm.mx

² Tecnológico Nacional de México, Campus Tapachula, Tapachula, Mexico

Abstract. This article proposes a method to classify atrial fibrillation signals using time-frequency characteristics through a BiLSTM network. The experiment was performed with the ECG signals, which are part of the PhysioNet CinC 2017 database. In addition to the BiLSTM network, machine learning algorithms such as k Nearest Neighbors, Linear SVM, RBF SVM, Decision Tree, Random Forest, Neural Net, AdaBoost, Naive Bayes and QDA were used for the classification experiments. To measure the efficiency and quality of the proposed method, the Accuracy, Precision, Recall, F1 Score metrics were used, as well as the Cohen Kappa score and the Mathews correlation coefficient. The results obtained show a better classification performance in the BiLSTM Network with 93.57%, 92.86%, 94.20%, 93.53%, 1.0 and 1.0 of the mentioned metrics.

Keywords: Atrial fibrillation · Feature extraction · BiLSTM · Time-frequency · ECG · PhysioNet

1 Introduction

Abnormalities of the circulatory system are the most common cardiac disorders addressed with Electrocardiography (ECG). The importance of these studies is motivated by the prevalence, which is relatively high considering that around 3% of the world population suffers from some anomaly. Such is the case that studies have been carried out focused on the classification of Atrial Fibrillation (AF), this being one of the most common types of arrhythmias [1].

In this sense, the use of deep learning has become popular for its classification, abstraction and, above all, learning capabilities that make it attractive in a wide range of approaches that can hardly be solved with conventional computing equipment [2, 3]. Consequently, works are emerging that emphasize proposing

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2022
M. F. Mata-Rivera et al. (Eds.): WITCOM 2022, CCIS 1659, pp. 201–214, 2022.
https://doi.org/10.1007/978-3-031-18082-8_13

Figura B.6: Primera página del artículo publicado en el congreso internacional WITCOM 2022.

Adicionalmente, se colaboró en un trabajo de investigación titulado “Multi-core Median Redescending M-Estimator for Impulsive Denoising in Color Images”, participación que fue de suma importancia para conocer el proceso de producción científica en congresos de alto prestigio indexados en Scopus como lo es el “Mexican Conference on Pattern Recognition” bajo el auspicio de la Asociación Mexicana de Visión por Computadora, Neurocomputación y Robótica (MACVNR), miembro de la Asociación Internacional para el Reconocimiento de Patrones (IAPR).




The Academic Division of Engineering

Hereby certifies that

Christian Garcia Aquino

*Attended the 13th Mexican Conference on Pattern Recognition, MCPR 2021,
which was held virtually at the Instituto Tecnológico Autónomo de México,
and presented the conference*

***“Multi-core Median Redescending M-Estimator
for Impulsive Denoising in Color Images”***

July 23-25, Mexico City, Mexico.



*Edgar Roman-Ramirez ITAM
General Chairs*



*Ángel Fernando Kuri Morales, ITAM
General Chairs*

Figura B.7: Constancia de co-autor de artículo publicado en el congreso MCPR 2021.



Multi-core Median Redescending M-Estimator for Impulsive Denoising in Color Images

Dante Mújica-Vargas^{1(✉)}, Arturo Rendón-Castro¹, Manuel Matuz-Cruz²,
and Christian García-Aquino¹

¹ Tecnológico Nacional de México-CENIDET, Cuernavaca-Morelos, Mexico
dante.mv@cenidet.tecnm.mx

² Tecnológico Nacional de México-ITTapachula, Tapachula-Chiapas, Mexico

Abstract. In this paper, to reduce impulsive noise in color images we propose an extension of the Median Redescending M-Estimator. For that purpose, a multitasking approach was developed such as a multi-core processing in order to reduce in parallel the noise on R, G and B color channels. With this paradigm, an acceleration up to three times can be guaranteed compared to the sequential paradigm, while having the ability to reduce corrupted data up to densities of 80% of fixed-value and 40% of random-value impulsive noises, guaranteeing the preservation of edges. The effectiveness of our proposal is verified by quantitative and qualitative results.

1 Introduction

The acquisition or transmission of digitized images through digital communication channels inherently brings the incorporation of noise. Usually this corrupted information is modeled as an additive, multiplicative or impulsive noise. In this research paper, we concentrate on the impulsive denoising on color images. The state-of-the-art suggests that this type of noise has been commonly modeled such as salt and pepper impulse or random impulse. Two aspects that must be satisfied regardless of the type of noise are the suppression of atypical information while avoiding destroying the fine details of the images. To ensure better performance in the high-level processing tasks of digital image processing or computer vision, images need to be noise-free with a high visual quality.

In the literature interesting methods to filter impulse noise have been introduced; more specifically: in [1], an Adaptive Fuzzy Filter based on Histogram Estimation (AFHE) was proposed; in which, the size of the processing window was adapted based on local noise densities using fuzzy based criterion. In [2], an Adaptive Interpolation-based Impulse Noise Removal (AIBINR) algorithm was proposed, with the merit of developing parametric self-tuning. A Modified Cascaded Filter (MCF) was submitted in [3], this consisted of a cascading connection of Decision-based Median Filters and Un-symmetric Trimmed-Mean

© Springer Nature Switzerland AG 2021
E. Roman-Rangel et al. (Eds.): MCPR 2021, LNCS 12725, pp. 261–271, 2021.
https://doi.org/10.1007/978-3-030-77004-4_25

Figura B.8: Primera página del artículo publicado en el congreso MCPR 2021.

Acrónimos

FA	Fibrilación Auricular
TF	Transformada de Fourier
BiHNN	<i>Bidirectional Hybrid Neural Network</i>
STFT	<i>Short-time Fourier Transform</i>
ECG	Electrocardiograma
DL	<i>Deep Learning</i>
RNN	Redes Neuronales Recurrentes
RNA	Redes Neuronales Artificiales
LSTM	<i>Long Short-time Memory</i>
BiLSTM	<i>Bidirectional Long Short-time Memory</i>
CENIDET	Centro Nacional de Investigación y Desarrollo Tecnológico
CinC	<i>Chanllenge in Computing</i>
SVM	<i>Support Vector Machines</i>
ADAM	<i>Adaptive Moment Estimation</i>
BPTT	<i>Back Propagation Trough Time</i>

Bibliografía

- Benjumea M. Vibraciones causadas por actividad humana: caracterización, efectos y manejo en la ingeniería civil, Diciembre 2003.
- Jiménez S. Clasificación automática de registros ecg para la detección de fibrilación auricular y otros ritmos cardiacos, Julio 2018.
- Alqudah A. M. and Alqudah A. Deep learning for single-lead ECG beat arrhythmia-type detection using novel iris spectrogram representation. *Soft Computing*, November 2021. doi: 10.1007/s00500-021-06555-x. URL <https://doi.org/10.1007/s00500-021-06555-x>.
- Caterini A. L. and Chang D. E. *Deep Neural Networks in a Mathematical Framework*. Springer International Publishing, 2018. doi: 10.1007/978-3-319-75304-1. URL <https://doi.org/10.1007/978-3-319-75304-1>.
- Goldberger A. L., Amaral L. A. N., Glass L., Hausdorff J. M., Ivanov P. C., Mark R. G., Mietus J. E., Moody G. B., Peng C., and Stanley H. E. PhysioBank, PhysioToolkit, and PhysioNet. *Circulation*, 101(23), June 2000. doi: 10.1161/01.cir.101.23.e215. URL <https://doi.org/10.1161/01.cir.101.23.e215>.
- Barnes A. E. Instantaneous spectral bandwidth and dominant frequency with applications to seismic reflection data, 1993.
- Boashash B. Estimating and interpreting the instantaneous frequency of a signal. i. fundamentals. *Proceedings of the IEEE*, 80(4):520–538, April 1992. doi: 10.1109/5.135376. URL <https://doi.org/10.1109/5.135376>.
- Dhananjay B., Venkatesh N. P., Bhardwaj A., and Sivaraman J. Design and development of LSTM-RNN model for the prediction of RR intervals in ECG signals. In *Advances in Intelligent Systems and Computing*, pages 133–141. Springer Singapore, August 2021. doi: 10.1007/978-981-16-2123-9_10. URL https://doi.org/10.1007/978-981-16-2123-9_10.
- Chicco D. and Jurman G. The advantages of the matthews correlation coefficient (mcc) over f1 score and accuracy in binary classification evaluation. *BMC Genomics*, 21, 01 2020. doi: 10.1186/s12864-019-6413-7.

- Kingma D. P. and Ba J. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.
- Suárez D. G. Sintonización de una red totalmente conectada para segmentación de dos clases de objetos en imágenes, Mayo 2018.
- Bianchi F.M., Maiorino E., Kampffmeyer M.C., A. Rizzi, and Jenssen R. *Recurrent Neural Networks for Short-Term Load Forecasting: An Overview and Comparative Analysis*. SpringerBriefs in Computer Science. Springer International Publishing, 2017. ISBN 9783319703381. URL <https://books.google.com.mx/books?id=wu09DwAAQBAJ>.
- Petmezas G., Haris K., Stefanopoulos L., Kilintzis V., Tzavelis A., Rogers J. A., Katsaggelos A. K., and Maglaveras N. Automated atrial fibrillation detection using a hybrid cnn-lstm network on imbalanced ecg datasets. *Biomedical Signal Processing and Control*, 63:102194, 2021. ISSN 1746-8094. doi: <https://doi.org/10.1016/j.bspc.2020.102194>. URL <https://www.sciencedirect.com/science/article/pii/S1746809420303323>.
- Chen H., Li C., Li X., Hu W., Li Y., Liu W., Sun C., Yao Y., and Grzegorzec M. Gashis-transformer: A multi-scale visual transformer approach for gastric histopathology image classification. *ArXiv*, abs/2104.14528, 2021a.
- Hamil H., Zidelmal Z., Salah M., Sakhi S., Kaibou R., and Ould D. AF episodes recognition using optimized time-frequency features and cost-sensitive SVM. *Physical and Engineering Sciences in Medicine*, 44(3):613–624, June 2021b. doi: 10.1007/s13246-021-01005-2. URL <https://doi.org/10.1007/s13246-021-01005-2>.
- Rai H. M. and Chatterjee K. Hybrid CNN-LSTM deep learning model and ensemble technique for automatic detection of myocardial infarction using big ECG data. *Applied Intelligence*, August 2021. doi: 10.1007/s10489-021-02696-6. URL <https://doi.org/10.1007/s10489-021-02696-6>.
- Goodfellow I., Bengio Y., and Courville A. *Deep Learning*. MIT Press, 2016. URL <http://www.deeplearningbook.org>.
- Luna J. Conducción autónoma de un vehículo simulado mediante un modelo de red neuronal convolucional recurrente, Enero 2020a.
- Smith J. O. *Mathematics of the Discrete Fourier Transform (DFT)*. W3K Publishing, 2007. ISBN 978-0-9745607-4-8. URL <http://www.w3k.org/books/>.
- Selder J. L., Breukel L., Blok S., van Rossum A. C., Tulevski I. I., and Allaart C. P. Correction to: A mobile one-lead ECG device incorporated in a symptom-driven remote arrhythmia monitoring program. the first 5, 982 hartwacht ECGs. *Netherlands Heart Journal*, 27(3):165–165, January 2019. doi: 10.1007/s12471-019-1225-6. URL <https://doi.org/10.1007/s12471-019-1225-6>.

- Zaen J. V., Delgado-Gonzalo R., Ferrario D., and Lemay M. Cardiac arrhythmia detection from ECG with convolutional recurrent neural networks. In *Biomedical Engineering Systems and Technologies*, pages 311–327. Springer International Publishing, 2020. doi: 10.1007/978-3-030-46970-2_15. URL https://doi.org/10.1007/978-3-030-46970-2_15.
- Elman J. L. Language as a dynamical system. In Tim van Gelder and Robert Port, editors, *Mind as Motion: Explorations in the Dynamics of Cognition*, pages 195–223. MIT Press, 1995.
- Dawani J. *Hands-On Mathematics for Deep Learning: Build a Solid Mathematical Foundation for Training Efficient Deep Neural Networks*. Packt Publishing, 2020b. ISBN 9781838647292. URL <https://books.google.com.mx/books?id=Uj11zQEACAAJ>.
- Cai L., Zhou S., Yan X., and Yuan R. A stacked BiLSTM neural network based on coattention mechanism for question answering. *Computational Intelligence and Neuroscience*, 2019:1–12, August 2019. doi: 10.1155/2019/9543490. URL <https://doi.org/10.1155/2019/9543490>.
- Sun L., Wang Y., He J., Li H., Peng D., and Wang Y. A stacked lstm for atrial fibrillation prediction based on multivariate ecgs. *Health Information Science and Systems*, 8, 04 2020a. doi: 10.1007/s13755-020-00103-x.
- Tran L., Li Y., Nocera L., Shahabi C., and Xiong L. Multifusionnet: Atrial fibrillation detection with deep neural networks. *AMIA Joint Summits on Translational Science proceedings. AMIA Joint Summits on Translational Science*, 2020:654–663, 05 2020b.
- Segismundo M. J., Villanueva M. M. F., Zaldua A., and Lee-Ramos C. M. Spectrogram analysis of electrocardiogram with normal sinus rhythm, arrhythmia and atrial fibrillation. In *TENCON 2012 IEEE Region 10 Conference*, pages 1–5, 2012. doi: 10.1109/TENCON.2012.6412329.
- Ivanovic M. D., Atanasoski V., Shvilkin A., Hadzievski L., and Maluckov A. Deep learning approach for highly specific atrial fibrillation and flutter detection based on rr intervals. In *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 1780–1783, 2019. doi: 10.1109/EMBC.2019.8856806.
- Matuz M. J. Detección de anormalidades en mamografías utilizando la red neuronal convolucional alexnet, Enero 2017.
- Morales M. X. Segmentación no paramétrica de tejidos cerebrales mediante una arquitectura paralela de redes neuronales convolucionales, Octubre 2018.
- Schuster M. and Paliwal K. K. Bidirectional recurrent neural networks. *IEEE transactions on Signal Processing*, 45(11):2673–2681, 1997.

- Svensén M. and Bishop C. M. Pattern recognition and machine learning, 2007.
- Taner M. T., Koehler F., and Sheriff R. Complex seismic trace analysis. *Geophysics*, 44(6):1041–1063, 1979.
- NHLBI NIH. Atrial fibrillation. URL (Accedido el 10-06-2021) <https://www.nhlbi.nih.gov/health-topics/atrial-fibrillation>, 2021.
- Buduma N. and Locascio N. *Fundamentals of deep learning: Designing next-generation machine intelligence algorithms.* ” O’Reilly Media, Inc.”, 2017.
- Vogt N. Cnns, lstms, and attention networks for pathology detection in medical data, 12 2019.
- Faust O., Shenfield A., Kareem M., San T. R., H. Fujita, and Acharya U. R. Automated detection of atrial fibrillation using long short-term memory network with rr interval signals. *Computers in Biology and Medicine*, 102:327–335, 2018. ISSN 0010-4825. doi: <https://doi.org/10.1016/j.compbiomed.2018.07.001>. URL <https://www.sciencedirect.com/science/article/pii/S0010482518301847>.
- Borghi P. H., Borges R. C., and J. P. Teixeira. Atrial fibrillation classification based on MLP networks by extracting jitter and shimmer parameters. *Procedia Computer Science*, 181:931–939, 2021. doi: 10.1016/j.procs.2021.01.249.
- Flandrin P. Time-frequency/time-scale analysis, 1998.
- Wang R., Fan J., and Li Y. Deep multi-scale fusion neural network for multi-class arrhythmia detection. *IEEE Journal of Biomedical and Health Informatics*, 24(9): 2461–2472, September 2020. doi: 10.1109/jbhi.2020.2981526. URL <https://doi.org/10.1109/jbhi.2020.2981526>.
- Mousavi S., Afghah F., Khadem F., and Acharya U. R. ECG language processing (ELP): A new technique to analyze ECG signals. *Computer Methods and Programs in Biomedicine*, 202:105959, April 2021. doi: 10.1016/j.cmpb.2021.105959. URL <https://doi.org/10.1016/j.cmpb.2021.105959>.
- Pattanayak S. *Pro Deep Learning with TensorFlow*. Apress, 2017. doi: 10.1007/978-1-4842-3096-1. URL <https://doi.org/10.1007/978-1-4842-3096-1>.
- SAMUIC. Kappa de cohen. URL (Accedido el 07-05-2021) <https://www.samiuc.es/estadisticas-variables-binarias/medidas-de-concordancia/kappa-de-cohen/>, 2018.
- Smebook. The advantages and disadvantages of deep learning. URL (Accedido el 27-02-2021) <https://smebook.eu/knowledge-base/deep-learning/the-advantages-and-disadvantages-of-deep-learning/>, 2021.
- Hochreiter S. and Schmidhuber J. Long short-term memory. *Neural computation*, 9 (8):1735–1780, 1997.

- Mahmud T., Hossain A. R., and S. A. Fattah. ECGDeepNET: A deep learning approach for classifying ECG beats. In *2019 7th International Conference on Robot Intelligence Technology and Applications (RiTA)*. IEEE, November 2019. doi: 10.1109/ritapp.2019.8932850. URL <https://doi.org/10.1109/ritapp.2019.8932850>.
- Krasteva V., Christov I., Naydenov S., Stoyanov T., and Jekova I. Application of dense neural networks for detection of atrial fibrillation and ranking of augmented ECG feature set. *Sensors*, 21(20):6848, October 2021. doi: 10.3390/s21206848. URL <https://doi.org/10.3390/s21206848>.
- Cao X., Yao B., and Chen B. Atrial fibrillation detection using an improved multi-scale decomposition enhanced residual convolutional neural network. *IEEE Access*, 7:89152–89161, 2019a. doi: 10.1109/access.2019.2926749. URL <https://doi.org/10.1109/access.2019.2926749>.
- Wu X., Sui Z., Chu C., and Huang G. Detection of atrial fibrillation from short ECG signals using a hybrid deep learning model. In *Smart Health*, pages 269–282. Springer International Publishing, 2019b. doi: 10.1007/978-3-030-34482-5_24.
- Ping Y., Chen C., Wu Lu, and Shu M. Automatic atrial fibrillation detection based on deep learning model with shortcut connection. In *2020 IEEE 5th Information Technology and Mechatronics Engineering Conference (ITOEC)*, pages 1075–1079, 2020. doi: 10.1109/ITOEC49072.2020.9141689.
- Pan Y. N., Chen J., and Li X. L. Spectral entropy: A complementary index for rolling element bearing performance degradation assessment. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, 223(5):1223–1231, December 2008. doi: 10.1243/09544062jmes1224. URL <https://doi.org/10.1243/09544062jmes1224>.
- Yousefi Z. R., Parak J., Tarniceriu A., Harju J., Yli-Hankala A., Korhonen I., and Vehkaoja A. Atrial fibrillation detection from wrist photoplethysmography data using artificial neural networks. In *IFMBE Proceedings*, pages 399–404. Springer Singapore, May 2018. doi: 10.1007/978-981-10-9038-7_75. URL https://doi.org/10.1007/978-981-10-9038-7_75.



**TECNOLÓGICO
NACIONAL DE MÉXICO**

cenidet
Centro Nacional de Investigación
y Desarrollo Tecnológico