

Clasificación de señales ECG

^{1st} Juan Camilo Carvajal Sierra
Universidad de Antioquia
Medellín, Colombia
juanc.carvajal1@udea.edu.co

^{2nd} Ana María Uribe Betancur
Universidad de Antioquia
Medellín, Colombia
ana.uribeb@udea.edu.co

^{3rd} Duvan Vega Cardona
Universidad de Antioquia
Medellín, Colombia
duvan.vegac@udea.edu.co

Abstract—This work addresses the classification of electrocardiographic (ECG) signals to identify four common arrhythmias: Sinus Bradycardia (SB), Atrial Fibrillation (AFIB), Sinus Tachycardia (ST), and Supraventricular Tachycardia (SVT). ECG records were analyzed using a feature set that included temporal, spectral, cepstral (MFCC), and clinical variables. Three neural network models were implemented: a Multilayer Perceptron (MLP), a Long Short-Term Memory (LSTM) network, and a deep neural network with Batch Normalization (DNN+BN). The DNN+BN model achieved the best multiclass performance with 94% accuracy and improved balance across classes. Additionally, a binary classification between AFIB and SB achieved 98% accuracy in all models, highlighting the discriminative power of the extracted features. These findings confirm the relevance of multimodal ECG feature integration and deep learning architectures for reliable arrhythmia classification in clinical contexts.

Index Terms—Electrocardiography (ECG), Arrhythmia Classification, Deep Learning, Multilayer Perceptron (MLP), Long Short-Term Memory (LSTM), Batch Normalization, Signal Processing, MFCC, Biomedical Signal Analysis, Clinical Decision Support.

I. INTRODUCCIÓN

La electrocardiografía (ECG) es una herramienta fundamental en el diagnóstico y monitoreo de enfermedades cardiovasculares, permitiendo la identificación de arritmias a partir de la actividad eléctrica del corazón [1]. Sin embargo, la interpretación manual de trazos ECG puede ser compleja y demandante, especialmente cuando se manejan grandes volúmenes de datos o se presentan señales con ruido, variabilidad morfológica y patrones sutiles [2]. En este contexto, los métodos computacionales basados en aprendizaje automático han cobrado relevancia como herramientas de apoyo diagnóstico, al permitir la clasificación automática de ritmos cardíacos mediante la extracción y análisis de características relevantes [3].

El presente proyecto propone una estrategia integral para la clasificación multicategoría de arritmias cardíacas, utilizando un conjunto de registros ECG preprocesados y anotados, del cual se extrajeron características temporales, espectrales, cepstrales y clínicas. Posteriormente, se entrenaron y evaluaron tres arquitecturas de redes neuronales profundas: perceptrón multicapa (MLP), red LSTM y una red densa con normalización por lotes (DNN+BN), con el fin de comparar su desempeño en tareas de clasificación supervisada [4]. Además, se implementó un escenario de validación binaria entre dos clases con alta distinción fisiológica, como Atrial Fibrillation

(AFIB) y Sinus Bradycardia (SB), para evaluar la sensibilidad de los modelos ante patrones extremos [5].

Este enfoque busca no solo automatizar la clasificación de arritmias, sino también validar la efectividad de técnicas de procesamiento de bioseñales y aprendizaje profundo en el contexto clínico, aportando una base para futuras aplicaciones en dispositivos inteligentes de apoyo diagnóstico.

II. METODOLOGÍA

A. Base de Datos y Selección de Pacientes

Para la realización de este estudio fue empleada la base de datos utilizada previamente, la cual se encontraba constituida por señales electrocardiográficas (ECG) correspondientes a pacientes que presentaban diferentes tipos de arritmias. A partir del conjunto de datos completo, fueron seleccionadas las cuatro clases más representativas: Sinus Bradycardia (SB), Atrial Fibrillation (AFIB), Sinus Tachycardia (ST) y Supraventricular Tachycardia (SVT). Posteriormente, fueron aplicados filtros con el objetivo de conservar exclusivamente los registros correspondientes a pacientes cuyas edades se encontraban comprendidas entre 18 y 90 años, mientras que fueron eliminadas las clases minoritarias como AVNRT, AVRT y SAAWR. El conjunto de datos final quedó conformado por 7616 registros válidos.

B. Extracción de Características

Las señales de ECG fueron obtenidas a partir de la derivación II, empleando una frecuencia de muestreo de 500 Hz. Fue desarrollado un sistema de extracción de características que combinaba atributos temporales, espectrales, cepstrales (MFCC) y clínicos, tal como se describe a continuación:

- **Temporales:** Los picos R fueron identificados mediante la función `find_peaks`, y fueron calculadas métricas de variabilidad tales como SDNN, RMSSD y el rango de intervalos RR.
- **Espectrales:** Fue aplicado el método de Welch para la estimación de la densidad espectral de potencia, a partir del cual fueron extraídas la frecuencia pico, el ancho de banda espectral y la energía contenida en la banda QRS (5–15 Hz).
- **MFCC (Mel Frequency Cepstral Coefficients):** A través del paquete `librosa`, fueron calculados los coeficientes cepstrales en ventanas de 2 segundos, y fueron

extraídas estadísticas del segundo coeficiente (media, desviación estándar y varianza).

- **Clínicas:** Fueron incorporadas variables relevantes disponibles en el archivo Diagnostics.xlsx, tales como edad del paciente, género, frecuencia ventricular y auricular, duración del QRS, intervalo QT corregido, y parámetros del eje cardíaco (RAxis, TAxis), entre otros.

Las señales fueron procesadas de manera automatizada para la generación de un DataFrame consolidado con 24 características por registro, el cual fue exportado para su posterior análisis.

C. Pre-procesamiento de Datos

Las variables categóricas, tales como el género y la clase de arritmia, fueron codificadas de manera numérica. Posteriormente, fueron eliminados los registros que presentaban valores faltantes (NaN) y fueron normalizadas todas las variables numéricas mediante escalado estándar (z-score) con el fin de facilitar el entrenamiento de los modelos.

D. Modelado y Arquitecturas de Clasificación

Fueron diseñados y comparados tres modelos de redes neuronales para abordar la clasificación multiclase de las arritmias:

- **MLP (Multilayer Perceptron):** Red neuronal constituida por dos capas ocultas (128 y 64 neuronas), función de activación ReLU, regularización mediante Dropout y capa de salida softmax.
- **LSTM (Long Short-Term Memory):** Para la simulación de una estructura secuencial, cada señal fue dividida en 100 ventanas consecutivas de 50 muestras, y el modelo fue implementado con dos capas LSTM que incluían Dropout intermedio.
- **DNN con Batch Normalization:** Fue implementada una red profunda constituida por tres capas densas (256, 128, 64 neuronas), normalización por lotes (BatchNormalization) y regularización mediante Dropout.

Todos los modelos fueron entrenados empleando el optimizador Adam, función de pérdida categorical cross-entropy, y fue aplicado class weighting para compensar el desbalance existente entre las clases.

E. Evaluación del Modelo

Los modelos fueron evaluados mediante un conjunto de prueba correspondiente al 20% del total de datos, el cual fue estratificado por clase. Fueron reportadas métricas de rendimiento tales como la exactitud global, F1-score por clase y matrices de confusión. Adicionalmente, fueron analizadas las curvas de entrenamiento (precisión y pérdida por época). De manera complementaria, fue desarrollada una segunda etapa enfocada en la clasificación binaria entre SB y AFIB, debido a que estas clases constituyen representaciones de disfunciones auriculares y del nodo sinusal. En esta fase fueron reutilizadas las tres arquitecturas previamente descritas, las cuales fueron adaptadas al caso binario, manteniendo el mismo enfoque de validación y análisis.

III. RESULTADOS

Fueron entrenadas y evaluadas tres arquitecturas de red neuronal para la clasificación multiclase de arritmias: un perceptrón multicapa (MLP), una red LSTM con entradas simuladas como secuencias, y una red profunda con Batch Normalization. A continuación, son presentados los resultados obtenidos en términos de precisión, F1-score por clase, matrices de confusión y evolución del entrenamiento.

A. Resultados del modelo MLP

Por el modelo MLP básico fue mostrado un buen rendimiento general, siendo alcanzada una exactitud del 94.2% en el conjunto de prueba. El F1-score fue particularmente alto para la clase Sinus Bradycardia (SB) (0.99), mientras que para SVT fue menor, reflejando el desequilibrio de clases. La matriz de confusión es mostrada a continuación, donde puede ser observada una alta tasa de aciertos en todas las clases, con ligeras confusiones entre ST y SVT. El reporte de clasificación correspondiente confirma estos resultados con una media ponderada de F1-score de 0.94.

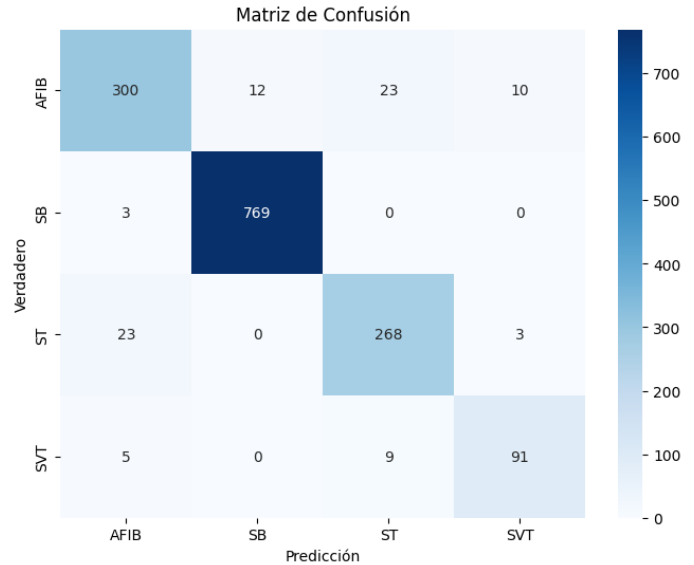


Fig. 1. Matriz de Confusión Modelo 1: MLP Básica

Clase	Precisión	Recall	F1-Score	Support
AFIB	0.88	0.88	0.88	345
SB	0.99	0.98	0.99	772
ST	0.88	0.93	0.90	294
SVT	0.87	0.82	0.84	105
Accuracy			0.93	1516
Macro avg	0.90	0.90	0.90	1516
Weighted avg	0.94	0.93	0.93	1516

TABLE I
REPORTE DE CLASIFICACIÓN DEL MODELO 1

La evolución de la exactitud y pérdida durante las 50 épocas de entrenamiento es presentada en la siguiente figura, donde puede ser observada una buena convergencia sin señales de sobreajuste.

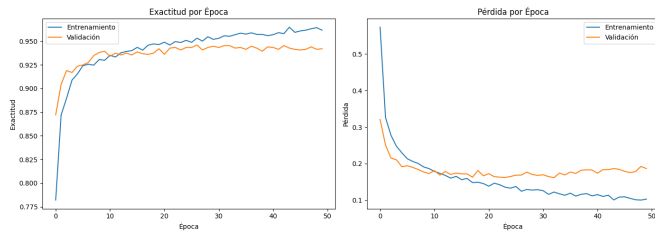


Fig. 2. Gráfica de Exactitud y Pérdida por época Modelo 1

B. Resultados del modelo LSTM

La red LSTM fue entrenada utilizando señales segmentadas en 100 ventanas consecutivas de 50 muestras. Por este modelo fue obtenida una exactitud del 91% en la clasificación de las cuatro arritmias. La clase SVT mostró el F1-score más bajo (0.81), mientras que SB alcanzó nuevamente un desempeño destacado (0.98). En la matriz de confusión es evidenciada una mayor confusión entre SVT y ST. Por el entrenamiento del modelo fue mostrada una mejora sostenida en las primeras 10–15 épocas, momento a partir del cual se estabilizó.

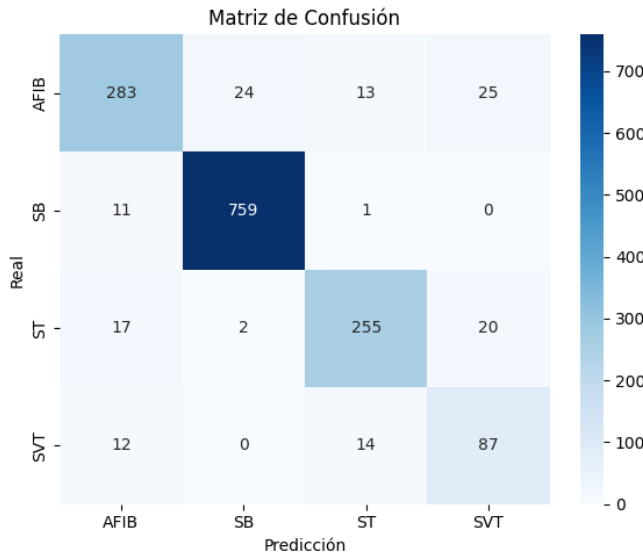


Fig. 3. Matriz de Confusión Modelo 2: Red LSTM

Clase	Precision	Recall	F1-Score	Support
AFIB	0.88	0.90	0.89	345
SB	0.99	0.97	0.98	771
ST	0.94	0.92	0.93	294
SVT	0.77	0.86	0.81	113
Accuracy			0.94	1523
Macro avg	0.89	0.91	0.90	1523
Weighted avg	0.94	0.94	0.94	1523

TABLE II
REPORTE DE CLASIFICACIÓN DEL MODELO 2

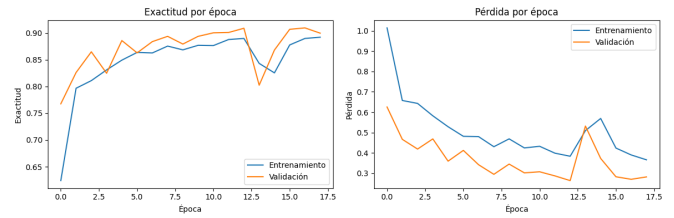


Fig. 4. Gráfica de Exactitud y Pérdida por época Modelo 2

C. Resultados del modelo DNN con Batch Normalization

Por este modelo fue alcanzado el mejor desempeño global con una exactitud de 94% y un F1-score más equilibrado entre clases. En particular, por SVT fue lograda una mejora ($F1 = 0.84$), respecto a las otras arquitecturas. La matriz de confusión y el reporte de clasificación confirman una mejora en la sensibilidad de las clases minoritarias.

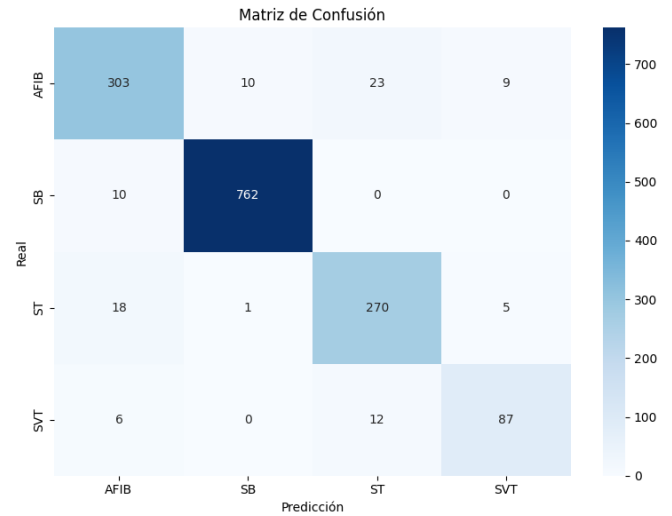


Fig. 5. Matriz de Confusión Modelo 3: DNN con Batch Normalization

Clase	Precision	Recall	F1-Score	Support
AFIB	0.90	0.88	0.89	345
SB	0.99	0.99	0.99	772
ST	0.89	0.92	0.90	294
SVT	0.86	0.83	0.84	105
Accuracy			0.94	1516
Macro avg	0.91	0.90	0.91	1516
Weighted avg	0.94	0.94	0.94	1516

TABLE III
REPORTE DE CLASIFICACIÓN DEL MODELO 3

Por la evolución del entrenamiento es mostrado un patrón estable con precisión creciente y pérdida decreciente, sin signos de sobreajuste durante las 50 épocas.

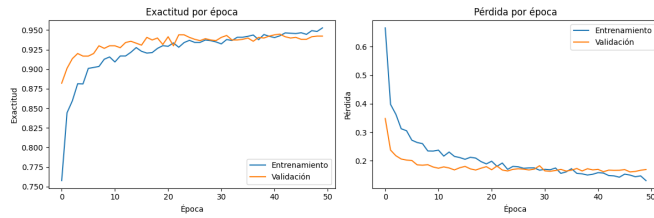


Fig. 6. Gráfica de Exactitud y Pérdida por época Modelo 3

D. Comparación de modelos

La comparación del F1-score por clase entre los tres modelos es presentada en el siguiente diagrama de barras, donde puede ser apreciado que el modelo DNN con Batch Normalization exhibe la mayor consistencia entre las diferentes clases de arritmias evaluadas.

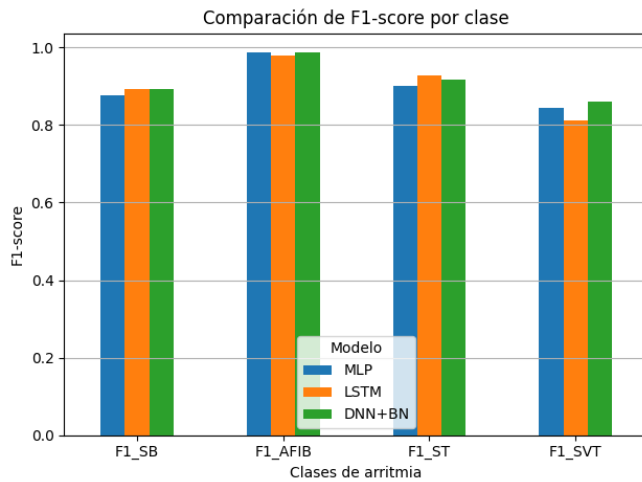


Fig. 7. Diagrama de barras comparativo (F1-Score) entre modelos

La exactitud global por arquitectura es resumida en la siguiente figura, donde es confirmada la ligera superioridad del modelo profundo en comparación con las arquitecturas MLP y LSTM.

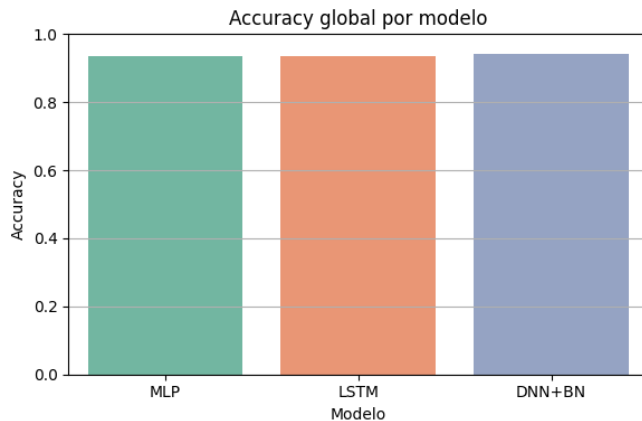


Fig. 8. Diagrama de barras comparativo (Accuracy) entre modelos

E. Clasificación binaria entre AFIB y SB

Con el propósito de explorar el desempeño de los modelos en una tarea más específica, fue implementada una clasificación binaria entre dos tipos de arritmias clínicamente relevantes: fibrilación auricular (AFIB) y bradicardia sinusal (SB). Estas clases fueron seleccionadas debido a que constituyen las más representadas en el conjunto de datos (345 y 772 muestras, respectivamente), además de su relevancia diagnóstica: la AFIB se encuentra asociada a disfunciones auriculares, mientras que la SB refleja alteraciones en el nodo sinusal.

Para esta tarea, fue reutilizado el conjunto de características extraídas previamente (temporales, espectrales, MFCC y clínicas), y fueron aplicadas las mismas etapas de preprocesamiento: codificación de etiquetas, normalización z-score y división estratificada en conjuntos de entrenamiento y prueba (80/20).

Fueron entrenados y comparados tres modelos arquitectónicos:

- **MLP (Perceptrón Multicapa):** red neuronal constituida por dos capas ocultas (128 y 64 neuronas) con función de activación ReLU.
- **DNN con Batch Normalization:** arquitectura más profunda conformada por tres capas densas (256, 128 y 64 neuronas), normalización por lotes y regularización mediante Dropout.
- **LSTM (secuencial simulada):** red recurrente entrenada sobre señales segmentadas en 100 ventanas de 50 muestras, simulando una estructura secuencial temporal.

Por los tres modelos fue alcanzado un desempeño prácticamente idéntico en el conjunto de prueba, siendo obtenida una exactitud del 98% y valores de F1-score de 0.97 para AFIB y 0.99 para SB. En la siguiente tabla son resumidos estos resultados.

Modelo	Exactitud	F1 AFIB	F1 SB	F1 Prom.
MLP	0.98	0.97	0.99	0.98
DNN + BatchNorm	0.98	0.97	0.99	0.98
LSTM (secuencial)	0.98	0.97	0.99	0.98

TABLE IV
COMPARACIÓN DE RENDIMIENTO ENTRE MODELOS EN CLASIFICACIÓN BINARIA

Por las matrices de confusión es confirmada la alta precisión alcanzada, evidenciándose un número mínimo de errores de clasificación y predicciones prácticamente perfectas en ambas clases.

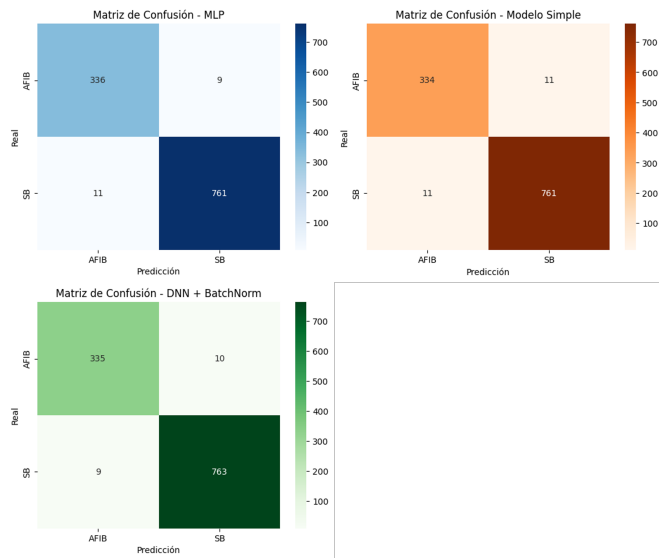


Fig. 9. Matrices de confusión para modelos de clasificación binaria

IV. DISCUSIÓN

Por los resultados obtenidos es evidenciado que las características extraídas de las señales ECG, combinadas con arquitecturas de red neuronal bien diseñadas, permiten una clasificación precisa y robusta de distintos tipos de arritmias cardíacas. En la tarea multiclase, por los tres modelos implementados —MLP, LSTM y DNN con Batch Normalization— fue superado el 91% de exactitud, siendo obtenido el mejor rendimiento global por la red DNN que alcanzó un F1-score de 0.98 para SB y una mejora considerable para SVT, la clase más minoritaria.

Por la red LSTM, pese a simular secuencias a partir de ventanas temporales fijas, fue mostrada una capacidad de generalización comparable a las arquitecturas densas, lo que sugiere que la estructura temporal inherente a las señales ECG puede ser aprovechada incluso con simulaciones aproximadas de secuencialidad. Sin embargo, fueron observados errores frecuentes entre las clases ST y SVT, reflejando su cercanía morfológica en la señal, lo que sugiere la necesidad de que sean incorporadas características adicionales (por ejemplo, morfología de la onda P o análisis de latidos individuales) para mejorar esta discriminación.

En la clasificación binaria entre AFIB y SB, por los tres modelos fue alcanzado un rendimiento muy alto (exactitud del 98%, F1-score > 0.97), con matrices de confusión prácticamente diagonales. Esto sugiere que el conjunto de características seleccionadas es especialmente discriminativo entre estas dos patologías, las cuales presentan patrones electrocardiográficos claramente diferenciados: alta irregularidad en AFIB versus frecuencia baja y regular en SB. Por la arquitectura DNN con Batch Normalization fue mantenido un rendimiento estable y consistente, lo que refuerza su utilidad como alternativa eficiente incluso en tareas con menor complejidad como esta.

Un hallazgo relevante es que las diferencias entre modelos fueron mínimas en la tarea binaria, pero más marcadas en el escenario multiclase, donde por el uso de técnicas de regularización y normalización interna (como Batch Normalization y Dropout) fue favorecida la estabilidad y precisión. Esto sugiere que, aunque las arquitecturas simples como MLP pueden ser suficientes para tareas binarias bien separadas, en entornos multiclase más desafiantes es preferible que sea optado por redes más profundas y con mejores mecanismos de control del aprendizaje.

En la actualidad, algoritmos de ayuda diagnóstica han sido incorporados por diversos equipos comerciales de electrocardiografía para automatizar la interpretación de señales ECG, facilitando de esta manera la detección temprana de arritmias y otras alteraciones cardíacas.

Entre los equipos disponibles, se destaca el electrocardiógrafo clínico Philips PageWriterTC70, el cual incorpora el algoritmo DXL de 16 derivaciones. Este algoritmo es capaz de generar más de 600 interpretaciones automáticas, incluyendo la detección de arritmias, anomalías en el segmento ST, desviaciones en los ejes cardíacos y valores corregidos del intervalo QT. Para ello, se emplea un análisis morfológico de las ondas P, QRS y T, junto con mediciones precisas de intervalos temporales, lo que guarda una estrecha relación con las técnicas de extracción de características temporales abordadas en el curso, como la identificación de picos R y la evaluación de la variabilidad. En Colombia, este dispositivo es distribuido por Philips y se utiliza en instituciones hospitalarias de niveles II y III [6].

Por otra parte, el GE MAC 5500 HD, que incorpora el algoritmo Marquette 12SL, constituye un electrocardiógrafo clínico de alta gama que ofrece interpretación automatizada mediante el análisis del ritmo, eje cardíaco, intervalos (PR, QRS, QT) y segmentos ST-T. Entre sus características técnicas más relevantes, se destaca la aplicación de filtros digitales que permiten la atenuación del ruido de alta frecuencia —como el generado por interferencias musculares o electromagnéticas—, así como la corrección de la línea base para compensar desplazamientos originados por la respiración o el movimiento del paciente. Estos procedimientos se encuentran alineados con los contenidos del curso en cuanto al preprocesamiento de bioseñales, donde se resalta la importancia de eliminar artefactos y estabilizar la señal antes de proceder a la extracción de características, como los picos R o los intervalos RR. Al implementar estas técnicas en tiempo real, el MAC 5500 HD incrementa la precisión de sus algoritmos de detección, facilitando así decisiones clínicas más confiables. En Colombia, este dispositivo es distribuido por GE Healthcare [8].

Finalmente, el CardioSecur Pro, desarrollado por Personal MedSystems GmbH, corresponde a un sistema portátil de monitoreo ECG que emplea cuatro electrodos para reconstruir hasta 22 derivaciones, transmitiendo los datos a una aplicación que realiza el análisis automático en la nube. En este dispositivo se implementan dos técnicas fundamentales del procesamiento de bioseñales abordadas en el curso: el análisis por ventanas temporales y la normalización de la

señal. Mediante la segmentación de la señal ECG en bloques breves, se facilita la extracción eficiente de características dinámicas, tales como los intervalos RR y la morfología de los complejos QRS, lo que permite la detección en tiempo real de arritmias como la fibrilación auricular o la bradicardia. Esta segmentación contribuye a mejorar la respuesta del algoritmo ante eventos transitorios y a reducir la latencia en la interpretación. Asimismo, se aplica una normalización por ventana que estandariza la señal —estableciendo media cero y varianza unitaria—, incrementando así la robustez del modelo frente a variaciones entre pacientes y permitiendo el uso de clasificadores livianos. Ambos componentes, esenciales en tareas de clasificación automática, fortalecen la precisión diagnóstica del dispositivo en entornos ambulatorios [7].

V. CONCLUSIONES

En este trabajo fue presentado un sistema de clasificación de arritmias cardíacas basado en el análisis de señales ECG, utilizando un conjunto de características combinadas (temporales, espectrales, cepstrales y clínicas) y tres arquitecturas de redes neuronales artificiales.

Los principales hallazgos pueden ser resumidos en los siguientes puntos:

- 1) Por las características extraídas fue permitida una representación robusta de la señal ECG, siendo facilitada una clasificación precisa tanto en tareas multiclase como binarias.
- 2) En la tarea multiclase (AFIB, SB, ST, SVT), por la red DNN con Batch Normalization fue mostrado el mejor rendimiento global, siendo alcanzado un F1-score macro de 0.91 y una exactitud del 94%, superando ligeramente al MLP y al modelo LSTM. Por la clase SVT fue presentada la mayor dificultad de clasificación, posiblemente debido a su baja frecuencia relativa y similitud morfológica con ST.
- 3) En la clasificación binaria entre AFIB y SB, por los tres modelos evaluados (MLP, DNN y LSTM) fue alcanzado un F1-score promedio de 0.98, con una exactitud del 98%, siendo demostrada la alta discriminabilidad entre estas dos patologías.
- 4) Aunque por todos los modelos fue mostrado un rendimiento destacado, fue observado que las arquitecturas más profundas y regularizadas (como la DNN con Batch Normalization) ofrecieron ventajas en términos de generalización, especialmente en escenarios multiclase.
- 5) Finalmente, fue comprobado que un enfoque híbrido de características (fisiológicas y espectrales), aplicado sobre una única derivación ECG (derivación II), es suficiente para que sea lograda una clasificación automática eficaz de arritmias, lo cual abre posibilidades para su implementación en herramientas de apoyo diagnóstico portátiles o de bajo costo.

REFERENCES

- [1] J. G. Webster, *Medical Instrumentation: Application and Design*, 4th ed. Hoboken, NJ: Wiley, 2009.
- [2] P. Laguna, R. Jané, and P. Caminal, "Adaptive filtering of ECG baseline wander," in *Proc. IEEE EMBS*, 1992, pp. 508–509.
- [3] A. Hannun et al., "Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network," *Nature Medicine*, vol. 25, pp. 65–69, 2019.
- [4] M. Faust, U. Acharya, H. Adeli, and L. Adeli, "Wavelet-based EEG processing for computer-aided seizure detection and epilepsy diagnosis," *Seizure*, vol. 26, pp. 56–64, 2015.
- [5] G. Clifford et al., "AF Classification from a Short Single Lead ECG Recording: The PhysioNet/Computing in Cardiology Challenge 2017," *Computing in Cardiology*, vol. 44, 2017.
- [6] Philips, "CARDIOGRAFOS PageWriter TC70," Philips Healthcare Colombia. [Online]. Available: <https://www.philips.com.co/healthcare/product/HC860315/cardigrafos-pagewriter-tc70-cardigrafo#features> [Accessed: Jul. 15, 2025].
- [7] GE Healthcare, "Marquette 12SL ECG Analysis Program Physician's Guide*," [Online]. Available: <https://www.gehealthcare.co.uk/-/jssmedia/global/uk/diagnostic-cardiology/2018/06/18/marquette-12sl-algorithm.pdf> [Accessed: Jul. 15, 2025].
- [8] CardioSecur, "Mobile ECG Technology for Professional Use," CardioSecur. [Online]. Available: <https://www.cardiosecur.com/> [Accessed: Jul. 15, 2025].