

DETECCIÓN DE ARRITMIAS EN ECG EMPLEANDO MÉTODOS DE APRENDIZAJE PROFUNDO

ECG ARRHYTHMIA DETECTION EMPLOYING DEEP LEARNING METHODS

Joshua Gutiérrez Ojeda

Instituto Politécnico Nacional, México
Jgutierrez1300@alumno.ipn.mx

Volodymyr Ponomaryov

Instituto Politécnico Nacional, México
vponomar@ipn.mx

José Agustín Almaraz Damián

Instituto Politécnico Nacional, México
jalmaraz1401@alumno.ipn.mx

Rogelio Reyes Reyes

Instituto Politécnico Nacional, México
rreyesre@ipn.mx

Clara Cruz Ramos

Instituto Politécnico Nacional, México
ccruzra@ipn.mx

Recepción: 24/noviembre/2023

Aceptación: 26/abril/2024

Resumen

Las enfermedades cardiovasculares son la principal causa de muerte en todo el mundo [OMS, 2021]. En este trabajo se presenta un sistema de Detección Asistida por Computadora (CADe) diseñado para identificar y clasificar anomalías en señales electrocardiográficas (ECG). Este sistema consta de dos etapas principales. (i) Se procesa la señal ECG, la cual se convierte en imágenes 2D utilizando la técnica de Recurrence Plot (RP), a partir de dividir la señal en segmentos de dos segundos. (ii) Las imágenes generadas se clasificaron utilizando la arquitectura ResNET-18 y el Módulo de Atención Convolutiva (CBAM). La base de datos utilizada MIT-BIH, está compuesta por 16 clases de anomalías, empleando el estándar de la Asociación para el Avance de Instrumentación Médica (AAMI) se procede a agrupar las 16 anomalías en cinco categorías de relevancia médica. El sistema diseñado

obtiene un rendimiento destacado, con una exactitud del 94.2%, precisión del 93.25%, recall del 93.24%, F1-Score del 93.24%.

Palabras Clave: Electrocardiograma (ECG), Módulo de Atención Convolutacional (CBAM), Recurrence Plot (RP), ResNet-18, Random Under Sampling (RUS).

Abstract

Cardiovascular diseases are the leading cause of death worldwide [OMS, 2021]. In the following work, a Computer-Aided Detection (CADe) system is presented, designed to identify, and classify anomalies in electrocardiographic (ECG) signals. This system consists of two main stages. (i) The ECG signal is processed, which is converted into 2D images using the Recurrence Plot (RP) technique, by dividing the signal into two-second segments. (ii) The generated images were classified using the ResNET-18 architecture and the Convolutional Block Attention Module (CBAM). The MIT-BIH database, composed of 16 classes of anomalies, was used, and following the standard of the Association for the Advancement of Medical Instrumentation (AAMI), the 16 anomalies were grouped into just five categories of medical relevance. The designed system achieves outstanding performance, with an accuracy of 94.2%, precision of 93.25%, recall of 93.24%, and F1-Score of 93.24%.

Keywords: *Convolutional Block Attention Module (CBAM), Electrocardiogram (ECG), Random Under Sampling (RUS), Recurrence Plot (RP), ResNet-18.*

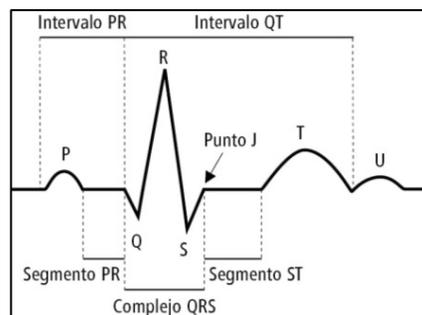
1. Introducción

La Organización Mundial de la Salud (OMS) estima que las enfermedades cardiovasculares son la principal causa de muerte en todo el mundo, se calcula que cada año mueren 17.9 millones de personas y se espera que en el 2030 esta cifra aumente a 23.6 millones de personas [OMS, 2021]. En México, se estima que 220 mil personas fallecieron por alguna enfermedad cardiovascular y de ellas 177 mil fueron por infarto al miocardio [secretaría de Salud, 2023].

El corazón es un órgano humano capaz de generar sus propios impulsos eléctricos, los cuales son formados en el nodo sinoatrial (SA) [Vélez, 2013], en ocasiones estos

impulsos eléctricos se forman o propagan de una manera incorrecta, dando lugar a arritmias, que suelen ser lentas (Bradicardias) o rápidas (Taquicardia) [Vélez, 2013]. Para observar los eventos anteriores se requiere el uso de un electrocardiógrafo, el cual entrega como resultado un electrocardiograma (ECG), que tiene como propósito el registro gráfico de la actividad eléctrica del corazón y se considera la herramienta más utilizada para diagnosticar problemas cardiovasculares debido a su naturaleza no invasiva e indolora.

Dependiendo de las formas de ondas que se registren en el ECG, los médicos o especialistas podrán realizar diferentes diagnósticos. Para [Pan, 1985], es importante extraer características de la señal ECG, las cuales son: duración, localización, amplitud de los complejos y cinco ondas representativas de la señal (P, Q, R, S y T), y en algunos casos también contempla la onda U, como se observa en la figura 1.



Fuente: elaboración propia.

Figura 1 ECG normal y sus ondas.

En la literatura, existen diversos métodos propuestos que realizan la detección y clasificación de las anomalías presentes en la señal ECG que utilizan una señal unidimensional (1D), como es el caso del autor [Jana, 2022], quienes utilizan una combinación de los clasificadores *Support Vector Machine (SVM)* y *Decision Tree*. Mientras que el autor [Han, 2022], emplean un método llamado *Bayesian Belief Network*. Diversos autores incursionan en métodos que emplean imágenes, los cuales se describen a continuación. En [Ahmad, 2021], proponen dos métodos denominados Fusión Multimodal de Imágenes (MIF) y Fusión de Características Multimodales (MTF), donde la señal ECG se convierte en una imagen empleando

las técnicas *Gramian Angular Field (GAF)*, *Recurrence Plot (RP)*, y *Markov Transition Field (MTF)*, se utiliza el conjunto de datos de arritmias del MIT-BIH, las cuales alimentan a los modelos CNN diseñados por los autores, con el fin de ser entrenadas desde cero, para la detección de cinco clases de arritmias.

En [Mathunjwa, 2022], realizan una segmentación de corta duración (dos segundos) de la señal ECG, posteriormente, una transformación de 1D a 2D generando imágenes empleando la técnica RP en la señal ECG. En la etapa de clasificación se contemplan cinco enfermedades cardíacas pertenecientes al conjunto de datos de arritmias del MIT-BIH, empleando dos arquitecturas CNN.

En [Niken, 2022], se propone un método para caracterizar ritmos cardíacos anormales al utilizar la técnica RP. Además, se presenta una arquitectura CNN para la clasificación de arritmias en el ECG que fue evaluada utilizando el conjunto de datos de arritmias del MIT-BIH, y para abordar el problema del sobreajuste (*overfitting*) del modelo, los autores emplean una fase de afinamiento (*Fine tuning*); sin embargo, no especifican qué pesos se usan o si entrenan desde cero.

Los métodos existentes presentan diversos inconvenientes, entre ellos el desbalance de clases que se encuentra en los conjuntos de datos debido a las diferencias significativas de muestras entre las clases mayoritarias y minoritarias.

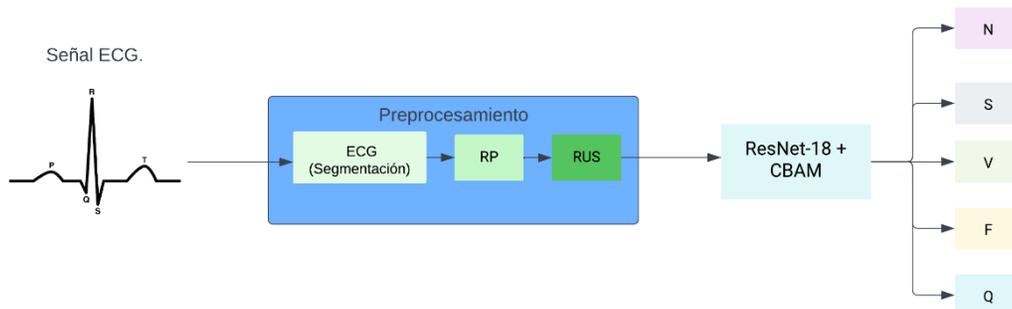
Las principales aportaciones del sistema presentado como *ECG-Recurrence-plot-Atención-Clasificación (ECG-RPAC)* en la clasificación de arritmias se resumen en:

- Se propone el uso de un método para resolver el problema del desbalance entre clases, evitando clasificaciones erróneas.
- Se emplea la red ligera ResNet-18 entrenada desde cero, además de emplear el módulo de atención *CBAM*.
- La validación experimental del rendimiento del sistema *ECG-RPAC* muestra ser competitivo frente a los mejores sistemas existentes, destacando los resultados sobresalientes en términos de *F1-Score*, *IBA*, *Precisión* y *Recall*.

2. Métodos

El diagrama a bloques del sistema propuesto se presenta en la figura 2. Las señales ECG son obtenidas a través de [PhysioNet, 1980]. Para este trabajo se

utilizó la base de datos del MIT-BIH, que está formada por 16 clases de anomalías, empleando el estándar de la Asociación para el Avance de Instrumentación Médica (AAMI) se procede a agrupar las 16 anomalías en cinco categorías de relevancia médica, las cuales son: Normal (N), ectópico ventricular (V), ectópico supraventricular (S), fusión (F) y desconocida (Q), estos tipo de arritmias están detalladas en la tabla 1.



Fuente: elaboración propia.

Figura 2 Diagrama a bloques del sistema propuesto.

Tabla 1 Estándar AAMI y sus categorías con sus anotaciones.

Categorías	Anotaciones
N	Latido normal
	Latido con bloqueo de la rama izquierda/derecha
	Latido de escape auricular
	Latido de escape nodal
S	Latido auricular prematuro
	Latido auricular prematuro aberrante
	Latido nodal prematuro
	Latido supraventricular prematuro
V	Latido ventricular prematuro
	Latido de escape ventricular
F	Fusión de latido ventricular y normal
Q	Latido estimulado (Marcapasos)
	Fusión de latido estimulado y normal
	Ruido
	Latido no clasificable

Fuente: AAMI, 1998.

En la primera etapa se propuso la segmentación de la señal ECG la cual se realiza cada dos segundos, posteriormente se transforma dicha señal a una imagen mediante la técnica de RP. Debido a la cantidad de muestras que se tiene por cada clase, existe un problema considerable con el desbalance de las clases, para mitigarlo, se propone emplear el método de *Random Under Sampling (RUS)*

[Johnson, 2019], en la cual se eliminan muestras aleatoriamente, dejando cada clase con un número equitativo de muestras.

Finalmente, como clasificador se utiliza la arquitectura de red neuronal convolucional ResNet-18 con el módulo de atención denominado como *Convolution Block Attention Module* (CBAM), dicha red se entrena desde cero. En la siguiente sección se aborda con detalle cada una de las partes anteriormente descritas.

Segmentación

Las señales que se encuentran en la base de datos de arritmias del MIT-BIH, contienen 48 extractos con una duración de 1.30 h, estas fueron adquiridas de pacientes ambulatorios y se obtuvieron dos derivaciones por cada paciente, las cuales son utilizadas para el monitoreo a pacientes en cuidados intensivos, debido a que la onda P representa el inicio de un nuevo ciclo cardíaco en el ECG.

Como se ha indicado se elige una segmentación cada dos segundos como [Mathunjwa, 2022], permitiendo un análisis continuo de la señal durante un periodo corto de tiempo. La base de datos tiene principalmente cinco clases en total, para la clase N, se cuenta con 90,630 señales, para la clase S, se tiene 2,780 señales, para la clase V, se tiene 7,236 señales, para la clase F, se cuenta con 803 señales y para la clase Q, se cuenta con 11,147, teniendo un total de 110,000 anotaciones

Recurrence plot

La técnica de *Recurrence plot* (RP), es propuesta por [Eckmann, 1995], teniendo como principal objetivo, visualizar la recurrencia de estados \vec{x}_i en un espacio de fases. La técnica de Eckmann permite investigar la recurrencia en un estado en el tiempo i y en un tiempo diferente j . Un RP puede definirse a partir de la ecuación 1.

$$R(\varepsilon) = \theta(\varepsilon - \|\vec{x}_i - \vec{x}_j\|), \quad i, \quad j = 1, 2, 3, \dots, N \quad (1)$$

Donde N es el número de estados considerados x_i , ε_i es la distancia umbral, $\|\cdot\|$ representa la norma euclidiana y θ es la función Heaviside que esta formulada por la ecuación 2.

$$\theta(Z) = \begin{cases} 0, & \text{si } z < 0 \\ 1, & \text{otro caso} \end{cases} \quad (2)$$

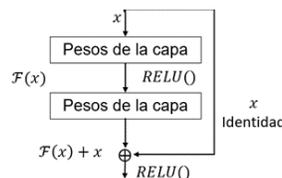
La ecuación 2 es considerada binaria, debido a ε , para este trabajo se considera eliminar ε , como en el trabajo [Faria,2016]. La nueva matriz está definida por la ecuación 3.

$$R(i, j) = \|\vec{x}_i - x_j\|, \quad i, \quad j = 1, 2, 3, \dots, \quad N \quad (3)$$

A partir de dicha técnica las señales ECG 1D se han transformado en imágenes 2D, con un tamaño de 170×170 pixeles. Al ser una imagen de un solo canal, se procede a replicar 3 veces el contenido de esta para generar la información necesaria para la creación de una imagen en el espacio RGB y finalmente concatenar dicha información y que dicha imagen sea utilizada en la siguiente etapa.

Arquitectura CNN

La arquitectura de redes neuronales convolucionales empleada es ResNet-18, propuesta por [He, 2016], en donde se introduce el bloque residual, como se muestra la figura 3.



Fuente: He, 2016.

Figura 3 Aprendizaje residual.

El enfoque propuesto por el autor permite conectar características de capas anteriores con características de capas posteriores mediante conexiones de salto. Esto ayuda a abordar el problema del desvanecimiento del gradiente y permite aumentar la profundidad de la red sin necesidad de agregar más capas.

En la figura 3, x es la entrada de la capa anterior, $F(x)$ es la transformación que se realiza en la entrada original para obtener la salida del bloque residual y $F(x) + x$ es la suma de la función transformada y la entrada original, esta suma se utiliza para obtener la salida del bloque residual.

Las señales que se encuentran en la base de datos de arritmias del MIT-BIH, contienen 48 extractos con una duración de 1.30 h, estas fueron adquiridas de

pacientes ambulatorios y se obtuvieron dos derivaciones por cada paciente, las cuales son utilizadas para el monitoreo a pacientes en cuidados intensivos, debido a que la onda P representa el inicio de un nuevo ciclo cardíaco en el ECG.

Como se ha indicado se elige una segmentación cada dos segundos como [Mathunjwa, 2022], permitiendo un análisis continuo de la señal durante un periodo corto de tiempo.

Módulo de atención

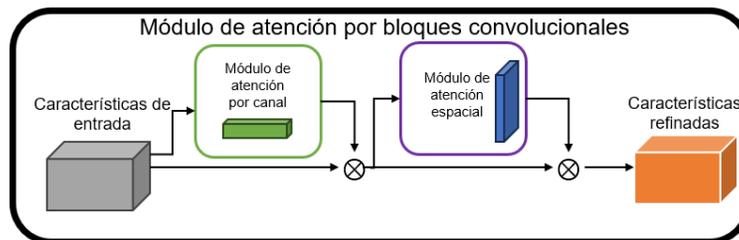
El módulo de atención denominado *Convolutional Block Attention Module* (CBAM), propuesto por [Woo, 2018], tiene como objetivo aprender la importancia relativa de cada mapa de características que se obtiene a partir de una capa convolucional determinada de la arquitectura presentada.

Se emplean operaciones de *pooling global* para capturar información global y se produce un vector de atención que resalta los canales más relevantes. Como se aprecia en la figura 4, se construye a partir de dos submódulos secuenciales (canal y espacial), los cuales se obtienen a partir de las ecuaciones 4 y 5.

$$F' = M_c(F) \otimes F \quad (4)$$

$$F'' = M_s(F') \otimes F' \quad (5)$$

Donde \otimes es la operación de producto punto tensorial o multiplicación de Hadamard, F es un mapa intermedio de características de dimensión $R^{C \times H \times W}$, M_c es el mapa de atención de un solo canal con dimensión $R^{C \times 1 \times 1}$, M_s es el mapa de atención espacial de dimensión $R^{1 \times H \times W}$ y F'' es el mapa de características refinado después de aplicar la atención de canal y la atención espacial.



Fuente: Woo, 2018.

Figura 4 Ejemplo de CBAM.

3. Resultados

Entorno de desarrollo

El método propuesto fue desarrollado y probado en un sistema computacional con las siguientes características: CPU AMD EPYC ®7532, 16 GB de RAM, GPU GeForce® 3090 con 24 GB de RAM, en un sistema operativo Linux de 64-bits, Python versión 3.10, con el uso de las librerías: Keras 2.9.0, Scikit-learn e imbalanced-learn.

Métricas de calidad

Para validar el rendimiento del sistema propuesto, se emplearon las métricas de evaluación conocidas en el estado del arte como: *Accuracy*, *Precision*, *Recall*, *F1-Score*, *Receiver Operating Characteristic curve (ROC)*, *Geometry Mean e Index of Balanced Accuracy (IBA)*, donde se definen como verdadero positivo (*tp*), verdadero negativo (*tn*), falso positivo (*fp*) y falso negativo (*fn*) y están dados por:

- *Accuracy* mide las clasificaciones adecuadas sobre el total de elementos evaluados, ecuación 6.

$$ACC = \frac{tp + tn}{tp + tn + fp + fn} \quad (6)$$

- *Precisión* caracteriza el número de elementos que se clasifican correctamente entre todos los elementos positivos a evaluar, ecuación 7.

$$PRE = \frac{tp}{tp + fp} \quad (7)$$

- *Recall* es la proporción de casos positivos reales que se predicen correctamente. Mide la cobertura de casos positivos reales y refleja los casos predichos correctos, ecuación 8.

$$REC = \frac{tp}{tp + fn} \quad (8)$$

- *F1-score* se define como la media armónica entre la precisión y la recuperación, que está dada por la ecuación 9.

$$F1 - Score = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (9)$$

La *curva Receiver Operating Characteristic (ROC)* es un gráfico bidimensional que ilustra el rendimiento de un clasificador [Yang, 2017] y el área bajo la curva (AUC) resume el rendimiento del clasificador en una única métrica cuantitativa, esta se utiliza para determinar si el clasificador tiene un rendimiento superior.

En el estudio [García, 2009], se propone un método para verificar la precisión para conjuntos de datos desbalanceados, llamado Índice de Precisión Equilibrada (*IBA*). La idea de esta medida es favorecer moderadamente los modelos de clasificación con un mayor índice de predicción de la clase minoritaria y sin subestimar la relevancia de la clase mayoritaria.

Las medidas utilizadas para caracterizar él *IBA* son:

- Geometría Media intenta maximizar la precisión para cada clase, al mismo tiempo, regula esas precisiones para que estén equilibradas, ecuación 10.

$$G_{mean} = \sqrt{Sensitivity \times Specificity} \quad (10)$$

- Dominancia tiene como objetivo cuantificar la relación de prevalencia entre las clases mayoritarias y minoritarias, es utilizada para analizar el comportamiento de un clasificador binario, ecuación 11.

$$Dominance = Sensitivity - Specificity \quad (11)$$

- *IBA* es una métrica de rendimiento en clasificación, sensible para dominios desequilibrados, ecuación 12.

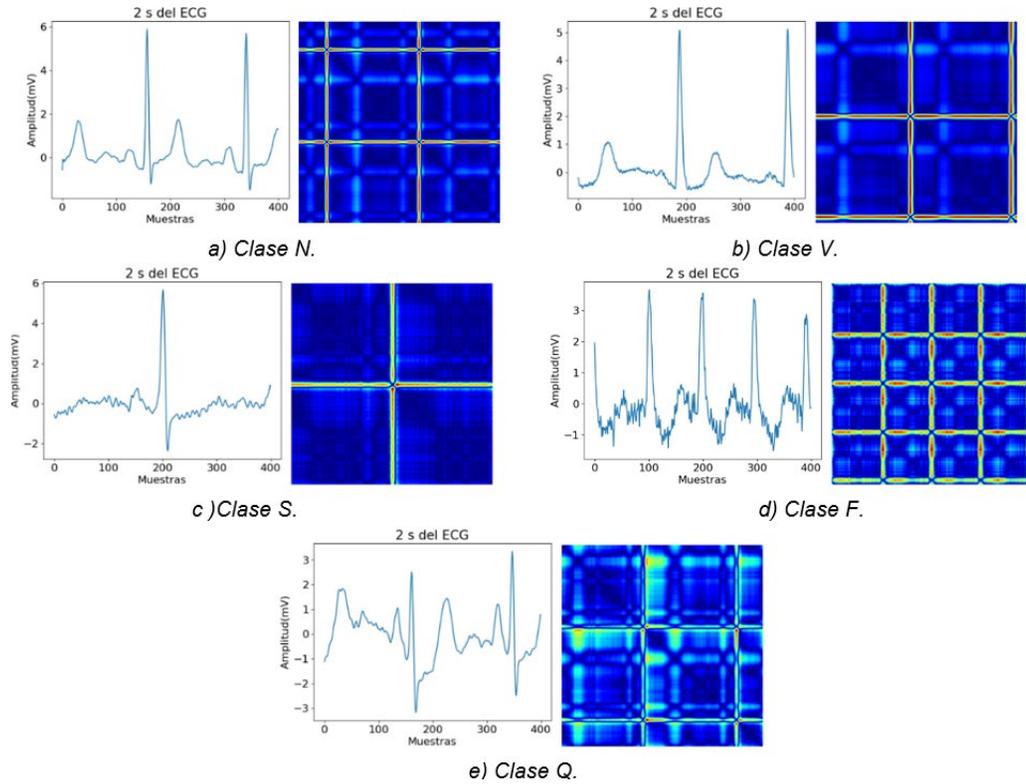
$$IBA = (1 - Dominance) \times G_{mean}^2 \quad (12)$$

Esquema de entrenamiento

La arquitectura ResNet, tiene como entrada imágenes en espacio RGB con un tamaño de $224 \times 224 \times 3$. En este caso las imágenes a ingresar son generadas por el método RP, a las imágenes se les aplica un *upsampling* mediante la técnica de interpolación bicúbica debido a que el tamaño es menor al requerido por la arquitectura. Las cuales se presentan en la figura 5, en a) Clase N, b) Clase V, c) Clase S, d) Clase F y e) Clase Q.

La red es inicializada con los pesos basados con la estrategia *He-Normal*, la cual emplea $N(\mu, \sigma^2)$, con $\mu = 0, \sigma = \sqrt{2/fan}$, donde *fan* es el número de entradas de

la capa anterior. El sistema propuesto es entrenado desde cero (scratch), con un *Learning rate* inicial $1e^{-2}$, con *momentum* de 0.8, un tamaño de lote de 4, con un total de 100 épocas, utilizando un estimador gradiente estocástico descendente (SGD) y la función de pérdida multiclase *CrossEntropy*.



Fuente: elaboración propia.

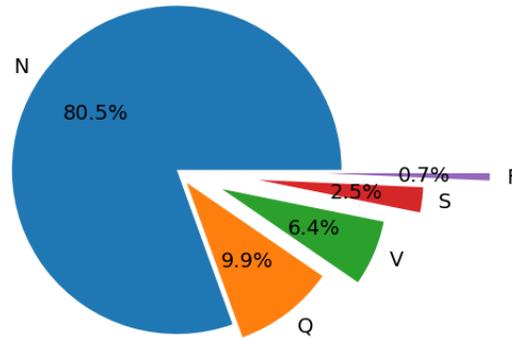
Figura 5 Segmentación y transformación del ECG.

Como optimización se utilizó la técnica de *Step decay*, que permite minimizar *Learning rate*, de acuerdo con una cantidad de épocas definida, se calcula empleando la ecuación 13. Donde lr es el Learning rate actual, lr_0 es el learning rate inicial, $drop$ es el factor de disminución de la tasa de aprendizaje, $floor$ representa el redondeo hacia $-\infty$ de $epoch/epochs_{drop}$, $epoch$ representa el número de época actual y $epochs_{drop}$ representa la época definida en la que el *Learning rate* disminuirá. Los valores utilizados para el presente trabajo son: $lr_0 = 1e^{-2}$, $drop = 0.25$, $epoch_{drop} = 15$.

$$lr = lr_0 * drop^{floor\left(\frac{epoch}{epochs_{drop}}\right)} \quad (13)$$

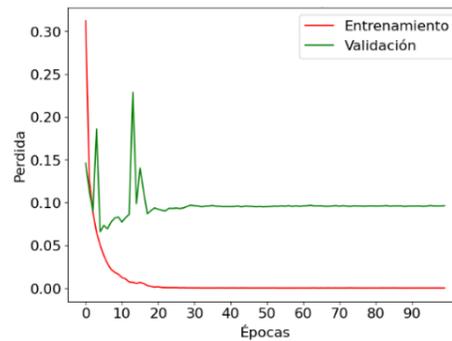
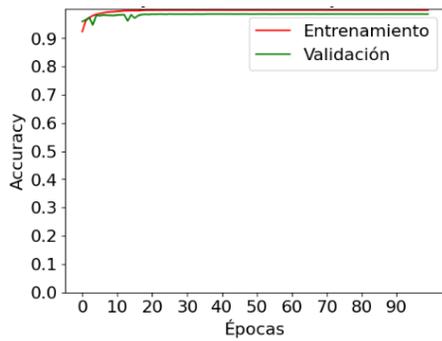
Entrenamiento con la base de datos completa

Como se observa en la figura 6, las clases están en desbalance por lo que cuentan con diferentes cantidades de muestras, por ejemplo, la clase N, contiene el 80.5% de todas las muestras, las clases S y F contienen 2.5 y 0.7% de todas las muestras respectivamente. Cuando es entrenado el método propuesto con todas las muestras de la base de datos se obtienen los resultados, figura 7.

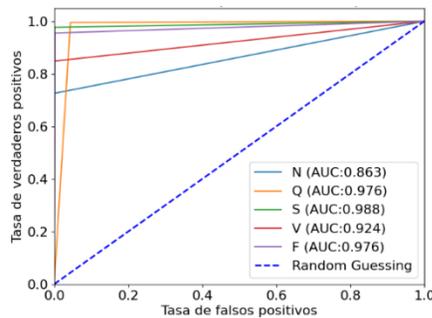


Fuente: elaboración propia.

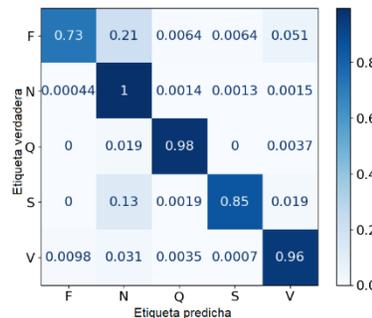
Figura 6 Distribución de la base de datos.



a) Accuracy del entrenamiento y validación. b) Pérdida de entrenamiento y validación.



c) Curvas ROC-AUC para cada clase.



d) Matriz de confusión.

Fuente: elaboración propia.

Figura 7 Resultados del entrenamiento con la base de datos del MIT-BIH.

Se visualizan las diferentes curvas de aprendizaje, en a) el *accuracy* del entrenamiento y validación, b) la pérdida de entrenamiento y validación, c) las diferentes curvas ROC-AUC para cada clase y d) la matriz de confusión. Para las gráficas de *accuracy* de entrenamiento como de validación, se nota que desde las primeras épocas el sistema alcanza un valor elevado de *accuracy*, en comparación de las gráficas de pérdida en las cuales se observa que aproximadamente en la época 30, el sistema es capaz de reducir su tasa de aprendizaje de forma gradual, indicando que el sistema está teniendo un sobreajuste (*overfitting*).

En la tabla 2, se presentan a mayor detalle, las métricas obtenidas, en donde se observa, que existen clases como la N, que tiene un desempeño sobresaliente, con respecto a otras clases; sin embargo, al evaluar la métrica *IBA*, esta disminuye debido a que la clase no tiene la cantidad de suficientes en comparación a las distintas clases evaluadas.

Tabla 2 Resultados con la base de datos en desbalance.

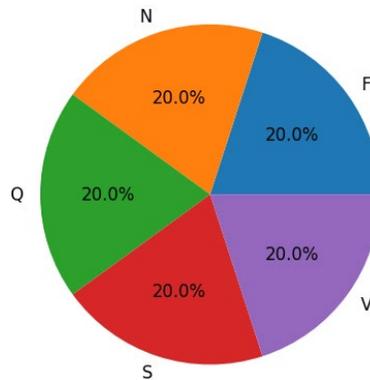
Clase	ACC	PRE	REC	F1-Score	AUC	G-Mean	IBA
N	99.1	99.6	99.6	99.5	87.4	86.5	72.9
Q	98.58	98.4	98.3	98.4	98.5	74.9	85.6
Clase	ACC	PRE	REC	F1-Score	AUC	G-Mean	IBA
S	85.6	93.7	90.6	92.21	99.1	78.4	87.4
V	95.25	80.5	97.3	97.1	95.2	86.2	98.3
F	81.2	88.1	74.80	80.9	98.6	80.1	88.2
Promedio	91.94	92.06	92.66	93.62	95.74	81.22	84.69

Fuente: elaboración propia.

Entrenamiento con la base de datos en equilibrio

Para el sistema propuesto se opta por emplear el método de *Random Under Sampling (RUS)* [Johnson, 2019], en el cual, disminuye las muestras por cada clase aleatoriamente, hasta que cada clase tenga la misma cantidad de muestras que la clase que cuenta con el menor número de muestras.

La técnica RUS reduce a 4,015 muestras el conjunto de datos completo, que son 803 muestras para cada clase, basándose en la clase minoritaria F, contenida en la base de datos de arritmias del MIT-BIH. La distribución del conjunto de datos equilibrado se muestra en la figura 8.



Fuente: elaboración propia.

Figura 8 Distribución de las clases, después de utilizar la técnica RUS.

En la tabla 3 se muestran los resultados obtenidos del sistema propuesto con el nuevo conjunto de datos generado con el método RUS y su comparación con los sistemas del estado del arte.

Tabla 3 Rendimiento del sistema ECG-RPAC contra el estado del arte.

Autor del Sistema	Preprocesamiento	ACC	PRE	REC	F1-Score	AUC	G-mean	IBA
Mathunjwa	RP	98.21	-	-	96.65	-	94.8	86.44
Ahmad	GAF	94.2	74.2	91	81.7	-	-	-
Ahmad	RP	96.3	80	90	84.7	-	-	-
Ahmad	MTF	94	72	86	78.3	-	-	-
Ahmad	MIF	98.6	93	92	82.4	-	-	-
Niken	RP	95.8	75.8	74.6	74.6	-	-	-
Propuesto	RP	94.2	93.25	93.24	93.24	95.7	96.5	92.7

Fuente: elaboración propia.

En la figura 9, se encuentran las diferentes curvas de aprendizaje, en a) se muestra el *accuracy* del entrenamiento y validación, b) se muestra la pérdida de entrenamiento y validación, c) se muestra las diferentes curvas *ROC-AUC* para cada clase y d) la matriz de confusión, después de emplear la técnica RUS.

Grad-CAM

Finalmente, se muestra el mapa de activación de clase ponderada por el gradiente (Grad-CAM), el cual utiliza los gradientes del entrenamiento, con el objetivo de producir un mapa de localización que resalte las regiones importantes, donde el sistema pone énfasis al predecir a que clase corresponden.

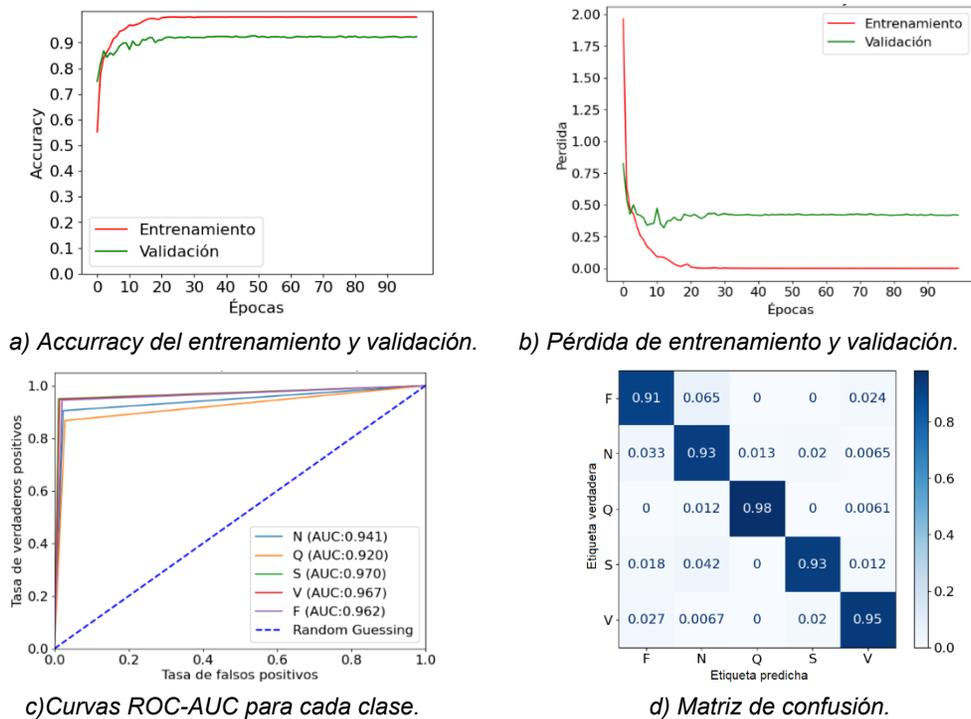
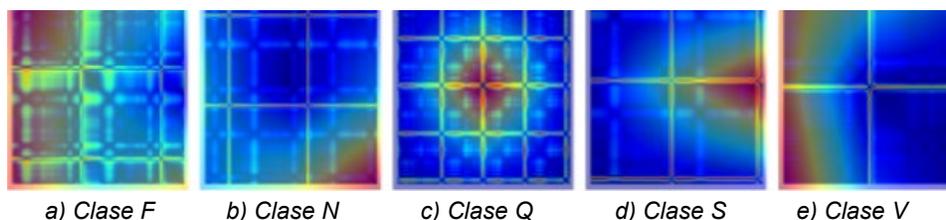


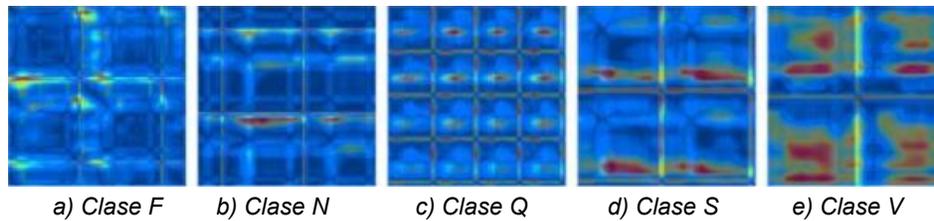
Figura 9 Resultados del entrenamiento con la base de datos del MIT-BIH.

Como se puede observar en las figuras 10 y 11, se muestran las imágenes generadas por la técnica Grad-CAM, donde se tiene para la clase F, la clase N, la clase Q, la clase S y la Clase V, respectivamente. Además, en la figura 10 se muestran las regiones resaltadas para la arquitectura ResNet-18 sin módulo de atención, por lo que se observa que la arquitectura selecciona solo una región de la imagen, como es el caso de las clases Q y V. En cuanto a la figura 11, se observan los resultados de la técnica Grad-CAM de la arquitectura ResNet-18 en combinación con el módulo de atención CBAM, se aprecia que la arquitectura se enfoca en partes más específicas para las clases mencionadas anteriormente.



Fuente: elaboración propia.

Figura 10 Resultados de Grad-CAM con la arquitectura ResNet-18 (Vanilla).



Fuente: elaboración propia.

Figura 11 Resultados de Grad-CAM con la arquitectura ResNet-18 + CBAM.

4. Discusión

Los resultados presentados en la tabla 3 demuestran que el sistema ECG-RPAC obtiene resultados sobresalientes en comparación con el estado del arte. Confirmando el desempeño del método propuesto.

Varios métodos del estado del arte utilizan diferentes técnicas para clasificar las Arritmias. Por ejemplo, en [Ahmad, 2021], el autor sugirió emplear la técnica de sobre muestreo denominada SMOTE para resolver el problema del desequilibrio de datos. Sin embargo, esto es un problema, debido que los datos sintéticos obtenidos podrían generar señales de ECG con patrones diferentes a los de la naturaleza fisiológica y causar un sesgo en el sistema.

Adicionalmente, el autor propuso la transformación 1D a 2D con varias técnicas de preprocesamiento (GAF, MTF y RP), independientes o combinadas, obteniendo una precisión del 98.6% con MIF; sin embargo, reincide el problema que el sistema se entrena sobre un conjunto de datos desequilibrado, lo que puede inducir a un sesgo a favor de la clase mayoritaria.

En [Mathunjwa, 2022], se propuso un sistema con dos etapas de clasificación empleando arquitecturas ResNet-18 y ResNet-50, su sistema no solucionó el problema del desbalance de datos, los cuales se pueden observar en los valores de los criterios G-Means e IBA (véase la tabla 3), donde pueden conllevar a una clasificación incorrecta.

Finalmente, en [Niken, 2022], propone un sistema donde emplea la técnica de RP; sin embargo, su sistema propuesto no resuelve el problema del conjunto de datos desbalanceado, y las métricas de calidad demuestran que su rendimiento no es competitivo en comparación con el método propuesto en este estudio (Tabla 3).

5. Conclusiones

En este estudio, se diseñó el sistema ECG-RPAC para la clasificación de cinco clases representativas de la base de datos de arritmias del MIT-BIH, siguiendo el estándar médico AAMI. El sistema resuelve el problema del desequilibrio en los datos sin utilizar generadores de muestras sintéticas las cuales, podrían generar muestras incorrectas y entregar una clasificación deficiente. ECG-RPAC puede ser entrenado con un volumen reducido de datos y logra resultados de clasificación favorables en comparación con los sistemas más avanzados del estado del arte, en términos de criterios comunes: un accuracy global del 94.2%, una precisión del 93.25%, un recall del 93.24%, un F1-Score del 93.2% y un AUC del 95.7%.

Otras métricas utilizadas para evaluar la calidad, como el F1-Score, el AUC, el G-Mean e IBA, justifican un mejor rendimiento en comparación con los métodos que no consideran el problema del desequilibrio de clases. El sistema propuesto puede ser útil para médicos sin experiencia, debido a que puede identificar morfologías poco frecuentes. Además, el sistema CAD diseñado puede ofrecer una segunda opinión en la clasificación debido a la generación de mapas de calor que explican cómo es que el sistema visualiza los patrones de las imágenes generadas por la técnica de RP a partir de las señales ECG.

6. Bibliografía y Referencias.

- [1] AAMI, (1998). Standard Cardiac Arrhythmia Detection. IEEE, 16th Indian Council International Conference, No. 2, Vol. 5, pp.1-4.
- [2] Ahmad, Z., Tabassum, A., Guan, L., & Khan, N. M., (2021). ECG Heartbeat c.
- [3] Classification Using Multimodal Fusion, (2022). IEEE Access, No. 24, Vol. 9, pp. 100615-100626.
- [4] Eckmann, J. P., Kamphorst, S. O., & Ruelle, D., (1995). Recurrence Plots of Dynamical Systems. World Scientific Series on Nonlinear Science, Series A, No. 16, Vol. 16, pp. 441-446.
- [5] Faria, F. A., Almeida, J., Alberton, B., Morellato, L. P. C., & Torres, R. D. S. (2016), Fusion of Time Series Representations for Plant Recognition in Phenology Studies. Pattern Recognition Letters, No. 55, Vol. 83, pp. 205-214.

- [6] García, V., Mollineda, R. A., & Sánchez, J. S., (2009). Index of Balanced Accuracy: A Performance Measure for Skewed Class Distributions. In Iberian conference on pattern recognition and image analysis, No. 2321, Vol. 5524, pp. 441-448.
- [7] Han, J., Sun, G., Song, X., Zhao, J., Zhang, J., & Mao, Y., (2022). Detecting ECG Abnormalities Using an Ensemble Framework Enhanced by Bayesian Belief Network. *Biomedical Signal Processing and Control*, No. 225, Vol. 72, pp. 103320.
- [8] He, K., Zhang, X., Ren, S., & Sun, J., (2016). Identity Mappings in Deep Residual Networks. In *Computer Vision—ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, Proceedings, Part IV*, No. 54, Vol. 14, pp. 630-645.
- [9] Jana, G. C., Agrawal, A., Pattnaik, P. K., & Sain, M., (2022). DWT-EMD Feature Level Fusion-Based Approach Over Multi and Single Channel EEG Signals for Seizure Detection. *Diagnostics*, No. 21, Vol. 12, pp. 324-334.
- [10] Johnson, J. M., & Khoshgoftaar, T. M., (2019). Survey on Deep Learning with Class Imbalance. *Journal of Big Data*, No. 1, Vol. 6, pp. 1-54.
- [11] Mathunjwa, B. M., Lin, Y. T., Lin, C. H., Abbod, M. F., Sadrawi, M., & Shieh, J. S., (2022). ECG Recurrence Plot-Based Arrhythmia Classification Using Two-Dimensional Deep Residual CNN features. *Sensors*, No. 4, Vol. 22, pp. 1660.
- [12] Niken, Martono P., Nishiguchi, T., & Ohwada, H., (2022). ECG Signal Classification Using Recurrence Plot-Based Approach and Deep Learning for Arrhythmia Prediction. In *Asian Conference on Intelligent Information and Database Systems*, No. 12, Vol. 3, pp. 327-335.
- [13] Organización Mundial de la Salud (OMS), (2021). The Burden of Cardiovascular Disease. No. 23, Vol. 46, pp. 54-86.
- [14] PhysioNet, (1980). The impact of the MIT-BIH Arrhythmia Database. *IEEE Eng in Med and Biol*, No. 3, Vol. 20, pp: 45-50.
- [15] Pan, J., & Tompkins, W. J., (1985). A Real-Time QRS Detection Algorithm. *IEEE transactions on biomedical engineering*. No. 42, Vol. 3, pp. 230-236.

- [16] Secretaria de salud, (2022). Cada año, 220 mil Personas Fallecen Debido a Enfermedades del Corazón. Ciudad de México, México.
- [17] Vélez Rodríguez, (2013). Pautas de la Electrocardiografía (2^a ed.), MARBÁN Libros, Cataluña, España.
- [18] Woo, S., Park, J., Lee, J. Y., & Kweon, I. S., (2018). Cbam: Convolutional block attention module. Proceedings of the European conference on computer vision (ECCV), No. 17, Vol. 39, pp. 3-19.
- [19] Yang, S., Orellana-Barrios, M., & Nugent, K., (2018). Research Methods: Clinical studies based on routine laboratory tests. The Southwest Respiratory and Critical Care Chronicles, No. 23, Vol. 6, pp. 33-39.