ISSN: 1692-7257 - Volumen 2 – Número 48 - 2026

ORIGINAL

RCTA
Revista Colombiana de Tecnologías de Avanzada
UNIPAMPLONA

Evaluación de la robustez de los descriptores de teoría de la información en la clasificación de señales cardíacas

Assessing the robustness of information theory descriptors in cardiac signal classification

Lely Adriana Luengas Contreras[1]
Esperanza Camargo Casallas[1]
Jonathan Roberto Torres Castillo[1]

[1] **Universidad Distrital Francisco José de Caldas, Bogotá, Colombia.**

Recibido: 23 febrero 2026. Revisado: 09 junio 2026. Aceptado: 24 julio 2026. Publicado: 28 julio 2026.
Autor de correspondencia: Lely Adriana Luengas Contreras

Cómo citar: L. A. Luengas Contreras, E. Camargo Casallas, and J. R. Torres Castillo, "Evaluación de la robustez de los descriptores de teoría de la información en la clasificación de señales cardíacas", *Revista Colombiana de Tecnologías de Avanzada (RCTA)*, vol. 2, no. 48, jul. 2026, art. no. 4593, doi: 10.24054/rcta.v2i48.4593

Esta obra está bajo una licencia internacional Creative Commons Atribución-NoComercial 4.0.

Resumen:

La clasificación automatizada de señales electrocardiográficas (ECG) puede constituirse en una herramienta de apoyo para la identificación de alteraciones cardiovasculares. No obstante, la naturaleza no lineal y no estacionaria de estas señales puede limitar el uso de métodos paramétricos convencionales. Este estudio evalúa descriptores de Teoría de la Información (TI) como alternativa no paramétrica para caracterizar segmentos de latidos cardíacos. Se analizaron 14 552 segmentos del ECG Heartbeat Categorization Dataset, cada uno representado por 187 posiciones temporales, y se cuantificó la interdependencia entre las categorías diagnósticas y las amplitudes mediante entropía e información mutua. Además, se implementó un procedimiento de agrupamiento basado en TI. La matriz de confusión para el modelo de clasificación mostró una sensibilidad de 94,38 % y una especificidad de 70,53 %; la exactitud balanceada fue 82,46 %. La pureza de los clústeres fue cercana a 0,877. Los resultados muestran una separación parcial entre segmentos normales y anormales y deben interpretarse considerando el desbalance de clases, el preprocesamiento y el efecto del zero-padding.

Palabras clave: entropía de Shannon, análisis de complejidad, procesamiento de señales de ECG, información mutua, menor pérdida de información, teoría de la información.

Abstract:

Automated classification of electrocardiographic (ECG) signals can be a useful tool for identifying cardiovascular abnormalities. However, the nonlinear and nonstationary nature of these signals may limit the use of conventional parametric methods. This study evaluates Information Theory (IT) descriptors as a nonparametric alternative for characterizing heartbeat segments. A total of 14,552 segments from the ECG Heartbeat Categorization Dataset were analyzed, each represented by 187 temporal positions, and the interdependence between diagnostic categories and amplitudes was quantified using entropy and mutual information. An IT based clustering procedure was also implemented. For the abnormal category, the confusion matrix showed a sensitivity of 94.38% and a specificity of 70.53%; balanced accuracy was 82.46%. Cluster purity was close to 0.877. The results indicate partial separation between normal and abnormal segments and should be interpreted considering class imbalance, preprocessing, and the effect of zero-padding.

Keywords: Shannon entropy, complexity analysis, ECG signal processing, mutual information, least information lost, information theory.

INTRODUCCIÓN

En los últimos años, el uso de teoría de la información (TI) y sus descriptores tales como medidas de entropía, complejidad y dependencia (información mutua, capacidad de canal, etc.) en el análisis de señales biomédicas en EEG, ECG, PPG, presión intracraneal e imágenes médicas, se ha consolidado por las ventajas que ofrece al no depender del tipo de datos y su distribución estadística, sino que analiza la cantidad de información que el conjunto de datos y cada una de las variables posee [1]-[6]. Revisiones centradas en TI y desarrolladas entre 2015 y 2024 han sistematizado su uso en señales fisiológicas, evidenciando una expansión significativa en la toma de medidas de diversos parámetros biológicos y de sus dominios de aplicación en ingeniería biomédica[2],[3], [7]-[10]. El uso de cuantificadores de teoría de la información (entropía y plano complejidad–entropía) en bioseñales eléctricas ha permitido caracterizar desorden, organización y estructura en diversas patologías, identificando vacíos especialmente en trastornos digestivos[2].

En sistemas biológicos, se ha empleado la teoría de Shannon para cuantificar transmisión de información y capacidad de canal en vías de señalización y para inferir redes de interacción molecular, destacando problemas de tamaño muestral y alta dimensionalidad, problemas habituales en biomedicina[11].

En el dominio de EEG, la combinación de cuantificadores de complejidad con técnicas avanzadas de procesado de señal y aprendizaje automático se ha utilizado para diagnóstico neurológico, interfaces cerebro–computador y monitorización clínica, abarcando desde la clasificación de ictus, estados de vigilia/sueño y detección de mentiras hasta el reconocimiento de emociones, siendo los conceptos asociados a la TI fundamentales en estas aplicaciones, mostrando resultados de alta calidad [12]-[15].

En cuanto al ECG, para la clasificación de latidos y arritmias, la entropía espectral calculada en el dominio tiempo–frecuencia se ha utilizado junto con redes neuronales convolucionales, logrando una discriminación eficaz entre diferentes tipos de latidos[16],[17]. De forma análoga, la entropía de wavelet o “wavelet packet entropy” se ha combinado con algoritmos de tipo random forest para la clasificación de arritmias, alcanzando resultados superiores a otros enfoques tradicionales y demostrando la utilidad de medidas entrópicas como descriptores robustos de la morfología y variabilidad del ECG[17]. En la detección de enfermedad coronaria, se ha mostrado que el análisis de la entropía aplicada a ECG de sujetos sanos y pacientes con enfermedad de arterias coronarias presenta resultados aceptables. De manera complementaria, enfoques de entropía jerárquica y multiescala permiten evaluar las características no lineales y multifrecuencia del ECG, integrando información de bajas y altas frecuencias y proporcionando tasas de discriminación entre registros normales e insuficiencia cardiaca cuando se fusionan con redes neuronales probabilísticas[18]. Revisiones recientes sobre entropía multiescala en señales fisiológicas respaldan estos hallazgos y subrayan el papel de la complejidad multiescala del ECG como marcador sensible de estados patológicos cardiovasculares[18], de allí que los cuantificadores derivados de la teoría de la información pueden constituir un eje central en el procesamiento de ECG [16]-[20].

Por otro lado, medidas de información mutua y entropía de transferencia se han empleado para estudiar el flujo de información entre las doce derivaciones estándar, identificando qué derivaciones aportan mayor contenido informativo y cómo se comparte la información entre ellas, lo que tiene implicaciones directas en la selección óptima de derivaciones y en el diseño de sistemas de diagnóstico con menos canales, como monitores Holter o dispositivos portátiles [21]-[25]. En esta misma línea, se han propuesto métricas de redundancia basadas en información mutua para cuantificar el solapamiento informativo entre derivaciones y optimizar la entrada de redes neuronales convolucionales, maximizando la información relevante y reduciendo la redundancia sin degradar el rendimiento diagnóstico[22].

Los descriptores de señales ECG haciendo uso de TI, tales como la entropía y la información mutua, se han mostrado en trabajos previos; no obstante, su integración con una discretización orientada a maximizar la cantidad de información de un sistema de forma que no se asuma distribución normal, pero que permita la disgregación de segmentos normales y anormales, ha sido poco explorada. Se plantea como hipótesis que la discretización mediante LIL conserva parte de la información diagnóstica relevante y permite identificar dependencia entre los segmentos de valores medidos de ECG y el tipo de señal. En este escenario, el aporte del trabajo reside en llevar a cabo una evaluación cuantitativa que emplee la teoría de la información de Shannon para integrar categorización LIL, entropía de Shannon, información mutua y generación de un modelo clasificatorio, así como examinar la incidencia del zero-padding a los descriptores obtenidos. Por tanto, la contribución se centra en una validación metodológica aplicada a señales ECG.

Teoría de la Información

Un sistema independiente de su naturaleza puede ser descrito como un sistema de información usando expresiones matemáticas. En sistemas físicos no lineales (por ejemplo, sistemas biológicos) la correlación lineal es una herramienta poco precisa. Un sistema biológico X existe en el espacio de los estados con una distribución de probabilidad p(x). Su incertidumbre promedio, o el potencial de transmisión de información, se define por la Entropía de Shannon (H) (1)[26].

$$H(X) = -\sum_{i=1}^{n} P(x_i)\log_b P(x_i) \quad (1)$$

Donde $P(x_i)$ es la probabilidad de un evento específico. Cuando b=2, la unidad es el bit.

La relación entre variables en un sistema es similar a la transferencia de información en un canal de comunicación. Cada canal tiene un límite condicionado por la parte física y el ruido. El Teorema de Shannon define la máxima rata (C) a la cual puede transmitirse datos sin error en un canal con ruido (2)[26].

$$C = B\log_2\left(1 + \frac{S}{N}\right) \quad (2)$$

Donde B es el ancho de banda (Hertz); S/N es la relación señal ruido.

La reducción de incertidumbre de X debido al conocimiento de Y se denomina Información Mutua I(X;Y) (3)[26].

$$I(X;Y) = H(X) - H(X|Y) \quad (3)$$

En un sistema ideal, I(X;Y) es igual a la entropía total de la fuente. En la parte práctica, el ruido H(X|Y) representa la equivocación o incertidumbre de la entrada al conocer la salida (4)[26].

$$H(X|Y) = -\sum_{x} p(x|y)\log_2 p(x|y) \quad (4)$$

MÉTODOS

Para el análisis se utilizó el subconjunto binario del ECG Heartbeat Categorization Dataset, derivado del PTB Diagnostic ECG Database y disponible públicamente[27]. El conjunto contiene 14 552 segmentos de latidos ECG: 4 046 etiquetados como normales y 10 506 como anormales. Cada fila incluye 187 valores temporales de amplitud y una etiqueta diagnóstica adicional; por tanto, la matriz analizada contiene 2 721 224 valores de amplitud (14 552 × 187). Las 187 posiciones no corresponden a una selección de sujetos ni a un muestreo de observaciones, sino a la longitud fija de cada segmento. A una frecuencia de muestreo de 125 Hz, esta longitud equivale aproximadamente a 1,496 s. Se incluyeron todos los segmentos disponibles de las dos categorías y no se realizó muestreo aleatorio. La etiqueta diagnóstica se empleó como variable de salida y no se incluyó entre las 187 amplitudes analizadas. Debido a que los archivos preprocesados se encuentran organizados por segmentos, los conteos no deben interpretarse como números de pacientes independientes.

Los datos fueron extraídos de archivos CSV y organizados para su análisis. Se tomó como base la metodología propuesta porLuengas (2016)[26], cuya secuencia general se presenta en la Figura 1. Inicialmente se efectuó un análisis exploratorio de datos para caracterizar el comportamiento de las señales. Se calcularon medidas de tendencia central y dispersión con el fin de describir la variabilidad entre segmentos y comparar las categorías normal y anormal.

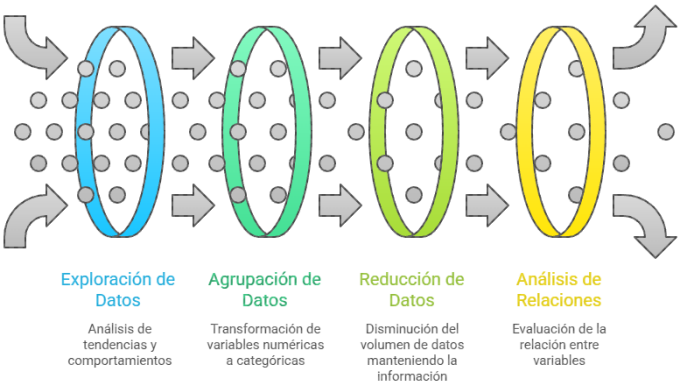

Figura 1. *Esquema de la metodología para el análisis de señales ECG.*
Fuente: elaboración propia.

Una vez examinados estadísticamente los datos, se realizó la etapa de agrupación de datos. Dado que los fundamentos de la Teoría de la Información (TI) requieren el manejo de variables discretas, y considerando que las mediciones de amplitud del ECG son de naturaleza continua (numéricas), se implementó el proceso de discretización o agrupación (binning) denominado Menor Pérdida de Información (Least Information Loss, LIL). LIL realiza el cálculo de la información transmitida por las variables para evaluar límites factibles de las categorías; los intervalos seleccionados enmarcan los datos que mejor describen el comportamiento del sistema con el menor incremento de entropía. De esta manera se obtienen los intervalos o bins empleados en el análisis informacional.

La cantidad de los datos categóricos se redujeron para reducir la carga computacional y obtener una representación simplificada del sistema. El procedimiento buscó preservar las posiciones con mayor información respecto de la salida. Para ello se empleó un algoritmo de cálculo de la información trasmitida y la relación señal/ruido (SNR), aquellos datos con una SNR mayor al 70% se descartaron, ya que la información que poseen es baja y su correspondencia con la entrada es mínima[15].

Finalmente, la relación entre las posiciones temporales de entrada y la categoría diagnóstica de salida se analizó mediante la versión gratuita de Powerhouse®. Se fijaron dos agrupaciones, en correspondencia con las dos categorías del conjunto: normal y anormal. El procedimiento integró la información transmitida, el indicador de señal/ruido reportado por el software y la verificación de no colinealidad entre variables. La correspondencia entre la asignación de los clústeres y las etiquetas diagnósticas se evaluó mediante una matriz de confusión.

Para interpretar la matriz de confusión se consideró la categoría 1 (anormal) como positiva. Las columnas representan las categorías esperadas y las filas la asignación obtenida; los porcentajes se encuentran normalizados por categoría esperada. Se calcularon sensibilidad, especificidad, precisión, exactitud global, exactitud balanceada y F1. Como indicadores externos de agrupamiento se estimaron la pureza, el Rand Index (RI), el Adjusted Rand Index (ARI) y la información mutua normalizada (NMI) a partir de la tabla de contingencia. Debido a que la figura reporta porcentajes redondeados a dos decimales, los conteos reconstruidos y las métricas derivadas se presentan como valores aproximados.

Disponibilidad de los datos

El conjunto preprocesado utilizado en este estudio corresponde al subconjunto binario del ECG Heartbeat Categorization Dataset, derivado del PTB Diagnostic ECG Database y disponible públicamente en Kaggle [27]. El análisis se realizó sobre segmentos de latidos y no sobre registros identificados a nivel de paciente.

RESULTADOS

Exploración de los datos

Se analizó la variabilidad temporal y estadística de los segmentos ECG clasificados como normales o anormales. Los datos, muestreados a 125 Hz y normalizados en amplitud, fueron sometidos a un análisis no paramétrico para identificar descriptores estadísticos potencialmente discriminativos. La Figura 2 muestra algunas de las señales analizadas.

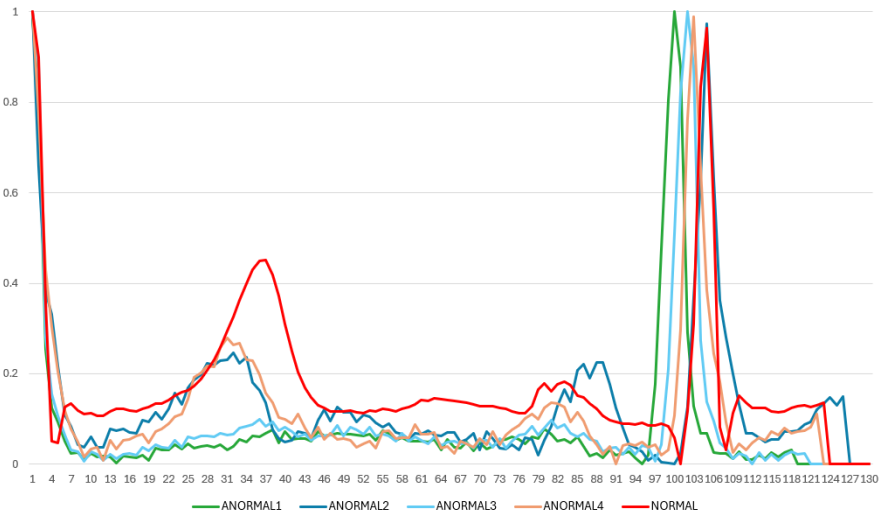

Figura 2. *Muestra de las señales analizadas.*
Fuente: elaboración propia.

Distribución y Normalidad

El análisis de la distribución de amplitudes en las 187 posiciones temporales de los segmentos permitió establecer el marco analítico. La prueba de D'Agostino–Pearson confirmó que los datos no siguen una distribución normal ($p < 0{,}001$).

La comparación de las amplitudes entre las categorías se realizó mediante la prueba U de Mann–Whitney y mostró diferencias estadísticamente significativas en más del 96 % de las posiciones temporales. La prueba de Kolmogorov–Smirnov también indicó diferencias entre las distribuciones. La concentración de valores iguales a cero en las posiciones finales se relaciona con el zero-padding utilizado para completar los segmentos de menor longitud hasta 187 posiciones;

por ello, estas posiciones pueden introducir sesgo en las medidas estadísticas e informacionales.

Medidas de tendencia

Dada la ausencia de normalidad, se emplearon la mediana y el valor eficaz (RMS) como estimadores de tendencia central y energía, respectivamente. Se observaron diferencias en amplitud, energía local y mediana entre las categorías, lo cual indica que las señales ECG describen variaciones morfológicas en los segmentos, sin establecer por sí solos una deducción de diagnosis clínica.

Variabilidad entre segmentos

El análisis revela que las señales normales exhiben una morfología altamente predecible con un Rango intercuartil (RIC) reducido. En contraste, el grupo con patologías muestra una dispersión significativamente mayor ($p < 0.01$, prueba de Mann-Whitney U), lo que sugiere que las patologías representadas no convergen en una única morfología, sino que aumentan la entropía del sistema.

Descriptores de Energía (RMS)

El análisis RMS reveló que los sujetos con alteraciones cardíacas muestran valores más altos en comparación con los sujetos sin alteraciones. Este hallazgo puede vincularse a anomalías, pero no es atribuible a una condición clínica concreta sin contar con información diagnóstica adicional.

Resumen del Análisis Estadístico

El análisis estadístico reveló diferencias entre las categorías normal y anormal. La categoría anormal mostró una mediana de amplitud pico menor ($0,92 \pm 0,18$ frente a $1,00 \pm 0,04$; $p < 0,05$) y un RMS mayor (0,3128 frente a 0,2741; $p < 0,01$). También presentó una curtosis menor (9,8 frente a 14,2; $p < 0,001$) y un rango intercuartílico mayor ($> 0,28$ frente a $< 0,15$; $p < 0,001$). Estos resultados describen una mayor variabilidad morfológica y energética en los segmentos anormales. La Tabla 1 resume los datos estadísticos.

Tabla 1: Resumen de las medidas estadísticas

Parámetro Estadístico	Grupo Normal	Grupo Anormal	Sig (p)
Mediana	1.00 ± 0.04	0.92 ± 0.18	< 0.05
RMS	0.2741	0.3128	< 0.01
Curtosis	14.2	9.8	< 0.001
RIC	Bajo (< 0.15)	Alto (> 0.28)	< 0.001

Fuente: elaboración propia

Agrupación y Reducción de los datos

La discretización mediante LIL buscó conservar la información relevante de las amplitudes respecto de la categoría de salida y reducir la distorsión introducida durante la cuantización. La transformación generó particiones específicas para las posiciones temporales y permitió seleccionar una representación reducida. Se obtuvo una reducción del 56 % de las posiciones temporales, por lo que permaneció el 44 % de las entradas con mayor información transmitida según el criterio implementado.

Análisis de Relaciones

Los datos se analizaron mediante TI para obtener el mapa infométrico de la Figura 3. Powerhouse® reportó una relación del 44 % entre la información de las entradas seleccionadas y la salida. Este porcentaje evidencia una asociación informativa parcial entre las posiciones temporales del ECG y la categoría diagnóstica, pero no constituye una medida directa de exactitud ni existe un umbral universal que permita calificarlo de manera aislada como alto o bajo. La existencia de una relación indica que el sistema puede ser modelado ya que las variables no son independientes[28].

En la representación del software, el 44 % equivale a 0,44 bits de información transmitida y el porcentaje restante corresponde a información no transmitida o equivocación dentro de ese esquema. La utilidad práctica del valor se debe analizar en conjunto con la matriz de confusión y los indicadores externos de agrupamiento.

Según lo indicado, el mapa infométrico respalda la existencia de una interdependencia entre las variables de entrada y la variable de salida, pero por sí mismo no muestra la posibilidad de crear un modelo predictivo de alto rendimiento.

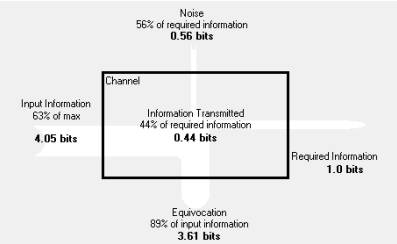

Figura 3. *Mapa infométrico del dataset. La relación entre las entradas y la salida es del 44%.*
Fuente: elaboración propia

Agrupamiento

Con el propósito de validar el desempeño del análisis basado en la Teoría de la Información (TI), se implementó un modelo de clasificación con la técnica de agrupamiento (clustering), donde se contempló la segregación y comparación de dos cohortes distintivas: un grupo compuesto por registros ECG normales y otro representativo de morfologías cardíacas patológicas. Se utilizó el software Powerhouse®.

En distinción a las orientaciones tradicionales que se basan en distancias euclidianas, para las medidas de similitud intraclúster (cohesión) e interclúster (separación) se tomó como fundamento la información mutua transmitida, considerando el análisis previo; de esta forma se cuantifica la dependencia entre los vectores de entrada (muestras de ECG) y la clasificación de salida del conjunto de datos; así, la consolidación de los clústeres responde tanto a la proximidad geométrica de las señales como a la coherencia informativa y la reducción de la incertidumbre diagnóstica inherente al sistema.

Powerhouse® calculó la información mutua, la información no transmitida y la equivocación, y representó las agrupaciones mediante el mapa de densidad de la Figura 4. El clúster 1 presentó una mayor concentración de segmentos anormales y el clúster 2 una mayor concentración de segmentos normales.

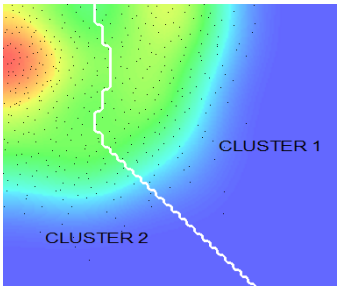

Figura 4. *Mapa de densidad.*
Fuente: elaboración propia

La Figura 5 presenta la distribución de las categorías en el conjunto inicial y en cada clúster. La representación muestra una separación parcial entre segmentos normales y anormales; por tanto, los resultados sugieren capacidad discriminativa, pero no implican una separación perfecta ni una validación clínica del procedimiento.

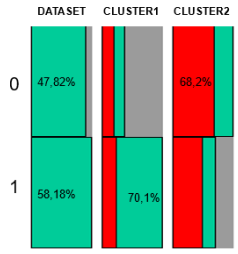

Figura 5. *Distribución de los datos en cada clúster*
Fuente: elaboración propia

La Figura 6 presenta la matriz de confusión normalizada. Considerando la categoría 1 como anormal, se obtuvo una sensibilidad de 94,38 % y una especificidad de 70,53 %. A partir de los tamaños de clase y de los porcentajes redondeados se estimaron una exactitud global de 87,75 %, una precisión de 89,27 %, una exactitud balanceada de 82,46 % y un F1 de 91,75 %. Como métricas externas de agrupamiento se obtuvieron, de manera aproximada, una pureza de 0,877, un Rand Index de 0,785, un Adjusted Rand Index de 0,545 y una NMI de 0,391. Estos valores se derivaron de la tabla de contingencia reconstruida y

deben confirmarse con los conteos originales si estos se encuentran disponibles.

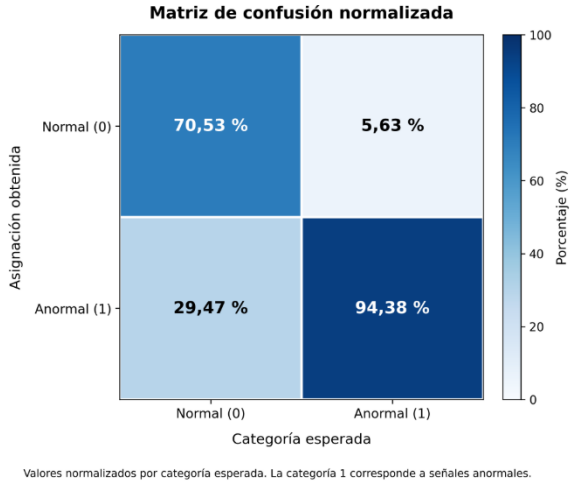

Figura 6. *Distribución de los datos en cada clúster*
Fuente: elaboración propia

Discusión y limitaciones

Medidas de entropía e información mutua han sido usadas en investigaciones previas como descriptores junto con clasificadores de aprendizaje automático [16]–[20]. A diferencia de esto, el presente trabajo estimó la transmisión de información para indagar la cantidad de información contenida luego de la discretización de los datos. De allí que, el 44% de información transmitida muestra una relación entre entradas y salida, pero no una estimación del rendimiento de un modelo computacional, para esto debe analizarse la sensibilidad, la especificidad y las métricas externas del modelo de clúster.

Con el fin de tener un formato específico para los datos se empleó zero-padding, de esta forma se agregan ceros en posiciones faltantes de valores; sin embargo, puede llegar a presentarse modificación en la entropía, la información mutua y las medidas de tendencia.

Dentro de las limitantes de la investigación se tiene el uso de un único conjunto de datos preprocesados, el análisis a nivel de segmentos de señales ECG y la información limitada sobre los algoritmos empelados en Powerhouse®, así como su configuración interna. Se necesita validación adicional con otros datos, sin que estén procesados y fortalecer las métricas de desempeño del modelo antes de generalizar los resultados.

CONCLUSIONES

Existe una diferencia estadísticamente significativa entre ECG normales y anormales, tanto en términos de amplitud como de forma de la distribución, confirmada por pruebas no paramétricas (Mann–Whitney U y Kolmogorov–Smirnov). Estos resultados validan el uso de técnicas de clasificación automática y justifican la extracción de características temporales y morfológicas para detección de anomalías cardíacas.

La mayoría de las muestras temporales son significativamente distintas, lo cual se demostró con el análisis estadístico, esto apunta a que la patología cardíaca no afecta solo un punto aislado y modifica la estructura global de la señal ECG.

Para agrupar rangos de valores en categorías o intervalos y así reducir la variabilidad (eliminar parte de la estructura fina) en un conjunto de datos se realizó la discretización mediante LIL, esta técnica permitió transformar datos numéricos en categóricos, preservando la información transmitida. Se reporta un 44% de información transmitida, aludiendo relación entre entradas y salida, pero no es una medida de desempeño de un modelo.

El desempeño del modelo con la matriz de confusión muestra una exactitud balanceada para un modelo que permite discriminación entre datos, pero que debe ser validado con otro conjunto de datos.

El zero-padding influencia especialmente las últimas posiciones de los segmentos, influyendo en las medidas estadísticas y las asociadas a la TI; de allí que, los resultados vinculados a esas posiciones deben ser evaluados con precaución.

Debido a la naturaleza no paramétrica de los datos, cualquier modelo de clasificación (como redes neuronales o máquinas de soporte vectorial) debe ser entrenado considerando que la media aritmética no es un descriptor fiable, se sugiere el uso de la Mediana y el RMS.

En general, los resultados avalan el uso de descriptores asociados a la TI para la discriminación de ECG. No obstante, se requiere validación independiente antes de establecer su utilidad como base de un modelo predictivo o extrapolar los hallazgos a otros contextos clínicos.

REFERENCIAS

[1] K. Balasubramanian, S. Rajendran, and S. Pati, "Complexity Measures in Biomedical Signal Analysis: A Clinically-Grounded Survey Across EEG, ECG, Intracranial Pressure, and Photoplethysmogram Modalities," *IEEE Access*, vol. 13, pp. 155285–155304, 2025. doi: 10.1109/access.2025.3603848

[2] L. Acosta and M. Pacheco, "Entropy and Statistical Complexity in Bioelectrical Signals: A Literature Review," *Signals*, vol. 6, no. 1, p. 7, 2025. doi: 10.3390/signals6010007

[3] J. Li, G. Feng, C. Ling, X. Ren, X. Liu, S. Zhang, L. Wang, Y. Chen, X. Zeng, and R. Chen, "A Resource-Efficient Multi-Entropy Fusion Method and Its Application for EEG-Based Emotion Recognition," *Entropy*, vol. 27, no. 1, p. 96, 2025. doi: 10.3390/e27010096

[4] C. Wu, B. Yao, X. Zhang, T. Li, J. Wang, and J. Pu, "The Application of Entropy in Motor Imagery Paradigms of Brain–Computer Interfaces," *Brain Sci.*, vol. 15, no. 2, p. 168, 2025. doi: 10.3390/brainsci15020168

[5] D. Platakis and G. Manis, "Review of Recent (2015–2024) Popular Entropy Definitions Applied to Physiological Signals," *Entropy*, vol. 27, no. 9, p. 983, 2025. doi: 10.3390/e27090983

[6] L. A. Luengas and L. F. Wanumen Silva, "Modelos computacionales en la posturografía," *Tecnura*, vol. 26, no. 73, pp. 30–48, 2022. doi: 10.14483/22487638.18060.

[7] Zandbagleh, S. Sanei, L. Penalba-Sánchez, P. Rodrigues, M. Crook-Rumsey, and H. Azami, "Intra- and Inter-Regional Complexity in Multi-Channel Awake EEG Through Multivariate Multiscale Dispersion Entropy for Assessing Sleep Quality and Aging," *Biosensors*, vol. 15, no. 4, p. 240, 2025. doi: 10.3390/bios15040240

[8] L. Sparacino, Y. Antonacci, G. Mijatović, and L. Faes, "Measuring hierarchically-organized interactions in dynamic networks through spectral entropy rates: Theory, estimation, and illustrative application to physiological networks," *Neurocomputing*, vol. 630, p. 129675, 2025. doi: 10.1016/j.neucom.2025.129675

[9] D. Sengupta, P. Gupta, and A. Biswas, "An efficient similarity metric for 3D medical image registration," *Multimedia Tools Appl.*, vol. 83, pp. 87987–88017, 2024. doi: 10.1007/s11042-024-18710-1

[10] M. Zbili and S. Rama, "A Quick and Easy Way to Estimate Entropy and Mutual Information for Neuroscience," *Front. Neuroinform.*, vol. 15, p. 596443, 2021. doi: 10.3389/fninf.2021.596443

[11] S. Uda, "Application of information theory in systems biology," *Biophys. Rev.*, vol. 12, pp. 377–384, 2020. doi: 10.1007/s12551-020-00665-w

[12] N. Amer and S. Belhaouari, "EEG Signal Processing for Medical Diagnosis, Healthcare, and Monitoring," *IEEE Access*, vol. 11, pp. 143116–143142, 2023. doi: 10.1109/access.2023.3341419

[13] T. Wang, W. Jia, F. Li, X. Liu, X. Zhang, and F. Hu, "Multi-dimensional feature extraction of EEG signal and its application in stroke classification," *Sci. Rep.*, vol. 15, 2025. doi: 10.1038/s41598-025-04756-0

[14] F. Von Wegner, M. Wiemers, G. Hermann, I. Tödt, E. Tagliazucchi, and H. Laufs, "Complexity Measures for EEG Microstate Sequences: Concepts and Algorithms," *Brain Topogr.*, vol. 37, pp. 296–311, 2023. doi: 10.1007/s10548-023-01006-2

[15] Z. Zhang, "Research on Mutual Information Analysis in EEG Based on Different Entropy Measures," in Proc. IEEE 5th Int. Conf. Information Systems and Computer Aided Education (ICISCAE), 2022, pp. 548–553. doi: 10.1109/iciscae55891.2022.9927610.

[16] Asgharzadeh-Bonab, M. Amirani, and A. Mehri, "Spectral entropy and deep convolutional neural network for ECG beat classification," *Biocybern. Biomed. Eng.*, vol. 40, pp. 691–700, 2020. doi: 10.1016/j.bbe.2020.02.004

[17] T. Li and M. Zhou, "ECG Classification Using Wavelet Packet Entropy and Random Forests," *Entropy*, vol. 18, no. 8, p. 285, 2016. doi: 10.3390/e18080285

[18] H. Wu, "Multiscale entropy with electrocardiograph, electromyography, electroencephalography, and photoplethysmography signals," *Biomed. Signal Process. Control*, vol. 93, p. 106124, 2024. doi: 10.1016/j.bspc.2024.106124

[19] M. Coq, A. Rey, O. Rosso, R. Armentano, and W. Legnani, "Detection of cardiac arrhythmia patterns in ECG through H × C plane," *Chaos*, vol. 32, no. 12, p. 123118, 2022. doi: 10.1063/5.0118717

[20] S. Aulia and S. Hadiyoso, "Abnormal ECG Classification using Empirical Mode Decomposition and Entropy," *Jurnal Rekayasa Elektrika*, vol. 17, no. 3, 2021. doi: 10.17529/jre.v17i3.22070

[21] L. Xie, Z. Li, Y. Zhou, Y. He, and J. Zhu, "Computational Diagnostic Techniques for Electrocardiogram Signal Analysis," *Sensors*, vol. 20, no. 21, p. 6318, 2020. doi: 10.3390/s20216318

[22] E. Ramírez, S. Ruipérez-Campillo, R. Casado-Arroyo, J. Merino, J. Vogt, F. Castells, and J. Millet, "The art of selecting the ECG input in neural networks to classify heart diseases," *Front. Physiol.*, vol. 15, 2024. doi: 10.3389/fphys.2024.1452829

[23] J. Sahores, V. Bank, M. Meijome, I. Ziccardi, and W. Legnani, "Study of Information Flow Among ECG Leads," in IEEE Biennial Congress of Argentina (ARGENCON), 2024, pp. 1–8. doi: 10.1109/argencon62399.2024.10735814

[24] B. Andayeshgar, F. Abdali-Mohammadi, M. Sepahvand, A. Almasi, and N. Salari, "Arrhythmia detection by the graph convolution network and a proposed structure for communication between cardiac leads," *BMC Med. Res. Methodol.*, vol. 24, 2024. doi: 10.1186/s12874-024-02223-4

[25] H. Zhang, Z. Bian, X. Guo, X. Li, S. Yin, and D. Cui, "Measuring phase-amplitude coupling using dispersion fuzzy mutual information," *Comput. Methods Programs Biomed.*, vol. 272, p. 109075, 2025. doi: 10.1016/j.cmpb.2025.109075

[26] L. A. Luengas, "Modelo de alineación estática para prótesis transtibiales", PhD Thesis, Facultad de Ingeniería, Pontificia Universidad Javeriana, 2016. https://repository.javeriana.edu.co/items/4297eb56-d54f-401c-bb6c-8cb1b7c029a1

[27] M. Kachuee, S. Fazeli, and M. Sarrafzadeh, "ECG Heartbeat Classification: A Deep Transferable Representation," 2018 IEEE International Conference on Healthcare Informatics (ICHI), New York, NY, USA, 2018, pp. 443–444. doi: 10.1109/ICHI.2018.00092.

[28] Pyle, D. (1999). Data preparation for data mining. In Morgan Kaufmann Publishers, Inc. Morgan Kaufmann Publishers, Inc. https://doi.org/10.1080/713827180

FINANCIACIÓN

Producto derivado del proyecto de investigación titulado "Desarrollo de una aplicación para la detección temprana de fibrilación auricular a partir de técnicas de procesamiento digital de señales aplicadas a datos de fotopletismografía (PPG)". Proyecto financiado por el Sistema Universitario Estatal (SUE), bajo el código de proyecto PPEE-CV SUE DC 02 2023.

CONFLICTOS DE INTERÉS

No existen

CONTRIBUCIÓN DE AUTORÍA

Todos los autores han leído y aprobado la versión final del manuscrito.