ISSN: 1692-7257 - Volume 2 – Number 48 - 2026

RCTA
Revista Colombiana de Tecnologías de Avanzada
UNIPAMPLONA

Assessing the robustness of information theory descriptors in cardiac signal classification

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

PhD. Lely Adriana Luengas Contreras[1], **PhD. Esperanza Camargo Casallas**[1], **PhD. Jonathan Roberto Torres Castillo**[1]

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

Correspondence: laluengasc@udistrital.edu.co

Received: february 23, 2026. **Revised:** june 09, 2026. **Accepted:** july 24, 2026. **Published:** july 28, 2026.

How to cite: L. A. Luengas Contreras, E. Camargo Casallas, and J. R. Torres Castillo, "Assessing the robustness of information theory descriptors in cardiac signal classification", *Revista Colombiana de Tecnologías de Avanzada (RCTA)*, vol. 2, no. 48, pp. 181–189, jul. 2026, doi: 10.24054/rcta.v2i48.4593

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.

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.

1. INTRODUCTION

In recent years, the use of IT and its descriptors—such as entropy, complexity, and dependency measures (mutual information, channel capacity, etc.)—in the analysis of biomedical signals including EEG, ECG, PPG, intracranial pressure, and medical imaging has become well established due to the advantages it offers by not relying on data type or statistical distribution, but instead by analyzing the amount of information contained in the dataset and in each of its variables [1]-[6].

Reviews focused on IT and developed between 2015 and 2024 have systematized its use in physiological signals, evidencing a significant expansion in the measurement of diverse biological parameters and their domains of application in biomedical engineering [2], [3], [7]-[10]. The use of IT quantifiers (entropy and the complexity–entropy plane) in electrical biosignals has enabled the characterization of disorder, organization, and structure in various pathologies, for example identifying gaps particularly in digestive disorders [2].

In biological systems, Shannon theory has been employed to quantify information transmission and channel capacity in signaling pathways and to infer molecular interaction networks, highlighting issues related to sample size and high dimensionality, which are common challenges in biomedicine [11]. In the EEG domain, the combination of complexity quantifiers with advanced signal processing techniques and machine learning has been used for neurological diagnosis, brain–computer interfaces, and clinical monitoring, covering applications ranging from stroke classification, wakefulness/sleep states, and lie detection to emotion recognition. In these applications, concepts associated with IT are fundamental and have demonstrated high-quality results [12]-[15].

In the ECG analysis, for heartbeat and arrhythmia classification, spectral entropy computed in the time–frequency domain has been used in conjunction with convolutional neural networks, achieving effective discrimination among different heartbeat types [16], [17]. Similarly, wavelet entropy or "wavelet packet entropy" has been combined with random forest algorithms for arrhythmia classification, achieving superior performance compared to other traditional approaches and demonstrating the usefulness of entropy-based measures as robust descriptors of ECG morphology and variability [17].

In the detection of coronary artery disease, it has been shown that entropy analysis applied to ECG signals from healthy subjects and patients with coronary artery disease yield acceptable results. Complementarily, hierarchical and multiscale entropy approaches allow the evaluation of the nonlinear and multifrequency characteristics of the ECG by integrating low- and high-frequency information and providing discrimination rates between normal recordings and heart failure when combined with probabilistic neural networks [18]. Recent reviews on multiscale entropy in physiological signals support these findings and highlight the role of multiscale ECG complexity as a sensitive marker of pathological cardiovascular states. Therefore, quantifiers derived from IT can constitute a central axis in ECG processing [16]-[20].

On the other hand, mutual information and transfer entropy measures have been employed to study information flow among the twelve standard ECG leads, identifying which leads contribute greater informational content and how information is shared among them. This has direct implications for the optimal selection of leads and for the design of diagnostic systems with fewer channels, such as Holter monitors or portable devices [21]-[25].

Along the same lines, redundancy metrics based on mutual information have been proposed to quantify informational overlap among leads and to optimize the input of convolutional neural networks, maximizing relevant information and reducing redundancy without degrading diagnostic performance [22].

Previous work has presented ECG signal descriptors based on information theory, including entropy and mutual information; but their integration with a discretization strategy designed to maximize information from a system, without assuming a normal distribution and enabling normal and abnormal segment separation, has been relatively under-explored.

This work hypothesizes that discretization using Local Interval Length (LIL) preserves relevant diagnostic information and enables the identification of dependencies between ECG signal segments and their type. The contribution of this research lies in conducting a quantitative evaluation that integrates LIL categorization, Shannon entropy, mutual information, and a classification model, while also examining the effect of zero-padding on resulting descriptors, specifically within the context of ECG signal methodology validation.

1.1. Information Theory

Every system can be described as an information system using a mathematical expression. In nonlinear physical systems (eg biological systems), linear correlation is a blunt instrument. Biological system *X* exists in a state space with a probability distribution *p(x)*. Its average uncertainty, or the potential for information transmission, is defined by Shannon Entropy (*H*) (1) [26].

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

Where $P(x_i)$ is the probability of a specific outcome. When b=2, the unit is bits.

The relationship between variables in a system is similar to the transfer of information on a communication channel. Every channel has a limit, dictated by physics and noise. The Shannon-Hartley theorem defines the maximum rate (*C*) at which you can transmit error-free data over a noisy channel (2) [26].

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

Where *B*: Bandwidth (Hertz); *S/N*: Signal-to-noise ratio
Mutual Information *I(X;Y)* is the reduction in uncertainty of *X* due to the knowledge of *Y* (3) [26].

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

In a perfect system, *I(X;Y)* equals the total entropy of the source. In reality, noise leaks *H(X|Y)*, which represents the equivocation or the uncertainty remaining about the input after seeing the output (4) [26].

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

2. MATERIALS AND METHODS

The binary subset of the ECG Heartbeat Categorization Dataset, derived from the PTB Diagnostic ECG Database and publicly available [27], was used for the analysis. The dataset contains 14,552 ECG heartbeat segments: 4,046 labeled as normal and 10,506 as abnormal. Each row includes 187 amplitude time values and one additional diagnostic label; therefore, the analyzed matrix contains 2,721,224 amplitude values (14,552 × 187). The 187 positions do not correspond to a selection of subjects or a sampling of observations, but rather to each segment's fixed length. At a sampling frequency of 125 Hz, this length is approximately equivalent to 1.496 s. All available segments from both categories were included, and no random sampling was performed. The diagnostic label was used as the output variable and was not included in the analysis of the 187 amplitudes. The counts should not be interpreted as numbers of independent patients because the preprocessed files are organized by segments.

Data was extracted from CSV files and organized for analysis. The methodology proposed by Luengas (2016) [26] was used as a basis, the general sequence of which is presented in Figure 1. Initially, an exploratory data analysis was performed to characterize the behavior of the signals. Measures of central tendency and dispersion were calculated to describe the variability between segments and compare the normal and abnormal categories.

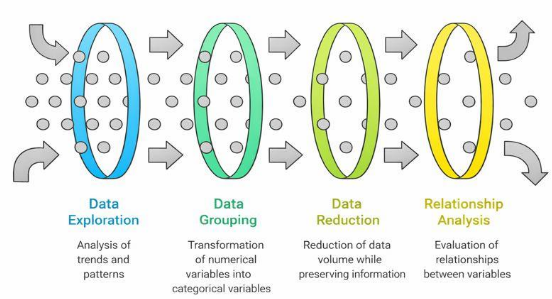

Figure. 1. *Methodology scheme for ECG signal analysis.*
Source: *own elaboration*

Once the data were statistically analyzed, the data grouping stage was performed. Since the foundations of Information Theory (IT) require the handling of discrete variables and considering that ECG amplitude measurements are continuous

RCTA
Revista Colombiana de Tecnologías de Avanzada
UNIPAMPLONA

(numerical) in nature, the discretization or binning process known as Least Information Loss (LIL) was implemented. LIL calculates the information conveyed by the variables to evaluate the feasible limits for the categories; the selected intervals frame the data that best describes the system's behavior with the least increase in entropy. The intervals or bins used in the informational analysis are obtained in this way.

The amount of categorical data was reduced to reduce the computational load and obtain a simplified representation of the system. The procedure sought to preserve the positions with the most information regarding the output. For this purpose, an algorithm was used to calculate the transmitted information and the signal-to-noise ratio (SNR). Data with an SNR of >70% were discarded because they contain little information and their correspondence with the input is minimal [15].

Finally, using the free version of Powerhouse®, the relationship between the entry time positions and the exit diagnostic category was analyzed. Two clusters were established, corresponding to the two categories of the dataset: normal and abnormal. The procedure integrated the transmitted information, the software-reported signal-to-noise ratio, and the verification of non-collinearity between variables. The correspondence between cluster assignments and diagnostic labels was evaluated using a confusion matrix.

To interpret the confusion matrix, category 1 (abnormal) was considered positive. The columns represent the expected categories and the rows the obtained assignment; percentages are normalized by expected category. Sensitivity, specificity, precision, overall accuracy, balanced accuracy, and F1 were calculated. As external clustering indicators, purity, the Rand Index (RI), the Adjusted Rand Index (ARI), and normalized mutual information (NMI) were estimated from the contingency table. Because the figure reports percentages rounded to two decimal places, the reconstructed counts and derived metrics are presented as approximate values.

3. RESULTS

3.1. Data Exploration

The temporal and statistical variability of ECG segments classified as normal or abnormal was analyzed. Data, sampled at 125 Hz and normalized in amplitude, were subjected to non-parametric analysis to identify potentially discriminative statistical descriptors. Figure 2 shows some of the analyzed signals.

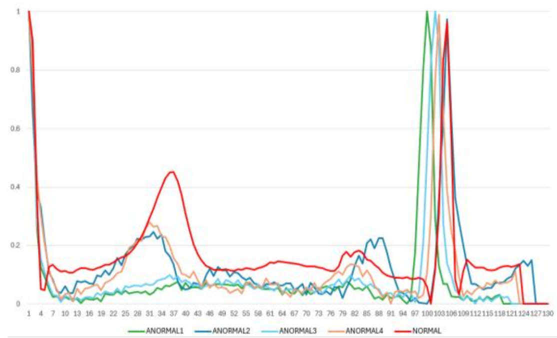

Figure. 2. *Sample of the analyzed signals.*
Source: *own elaboration*

3.1.1. Distribution and Normality

Analysis of the amplitude distribution across the 187 temporal positions of the segments established the analytical framework. The D'Agostino–Pearson test confirmed that the data do not follow a normal distribution ($p < 0.001$).

The Mann–Whitney U test was used to compare amplitudes between categories, which showed statistically significant differences in more than 96% of the temporal positions. The Kolmogorov–Smirnov test also revealed differences between the distributions. The concentration of zero values at the end positions is related to the zero-padding used to fill in the shorter segments up to 187 positions; therefore, bias may be introduced into the statistical and informational measurements.

3.1.2. Measures of Central Tendency

Given the lack of normality, the median and root mean square (RMS) values were used as estimators of central tendency and energy, respectively. Differences in amplitude, local energy, and median were observed between categories, indicating that ECG signals describe morphological variations in segments without establishing a clinical diagnosis.

3.1.3. Variability between the segments

The analysis reveals that normal signals exhibit a highly predictable morphology with a reduced Interquartile Range (IQR) during the QRS complex. In contrast, the pathological group shows significantly greater dispersion ($p < 0.01$, Mann–Whitney U test), suggesting that the represented pathologies do not converge to a single morphology but instead increase the entropy of the system.

3.1.4. Energy Descriptors (RMS)

The RMS analysis revealed that subjects with cardiac abnormalities had higher values compared to subjects without abnormalities. Although this finding may be linked to abnormalities, it is not attributable to a specific clinical condition without additional diagnostic information.

3.1.5. Summary of Statistical Analysis

The statistical analysis revealed differences between the normal and abnormal categories. The abnormal category showed a lower median peak amplitude (0.92 ± 0.18 vs. 1.00 ± 0.04; $p < 0.05$) and a higher RMS (0.3128 vs. 0.2741; $p < 0.01$). It also exhibited lower kurtosis (9.8 vs. 14.2; $p < 0.001$) and a larger interquartile range (> 0.28 vs. < 0.15; $p < 0.001$). These results indicate greater morphological and energetic variability in abnormal segments. Table 1 summarizes the statistical data.

***Table 1:** Summary of Statistical Analysis*

Statistical Parameter	Normal Group	Abnormal Group	Sig (p)
Amplitude Peak (Median)	1.00 ± 0.04	0.92 ± 0.18	< 0.05
Mean Power (RMS)	0.2741	0.3128	< 0.01
Average Kurtosis	14.2	9.8	< 0.001
Segment Stability (IQR)	Low (< 0.15)	High (> 0.28)	< 0.001

***Source:** own elaboration.*

3.2. Data Clustering and Reduction

Discretization using LIL aimed to preserve relevant amplitude information regarding the output category and reduce distortion introduced during quantization. The transformation generated specific partitions for time positions and allowed the selection of a reduced representation. A 56% reduction in time positions was achieved, leaving 44% of the inputs with the most transmitted information according to the implemented criteria.

3.3. Relationship Analysis

The data were analyzed using IT to obtain the infometric map shown in Figure 3. Powerhouse® reported a 44% relationship between the information from the selected inputs and the output. This percentage demonstrates a partial informational association between the ECG temporal positions and the diagnostic category, but it does not constitute a direct measure of accuracy, nor is there a universal threshold that allows it to be classified as high or low in isolation. The existence of a relationship indicates that the system can be modeled because the variables are not independent [28].

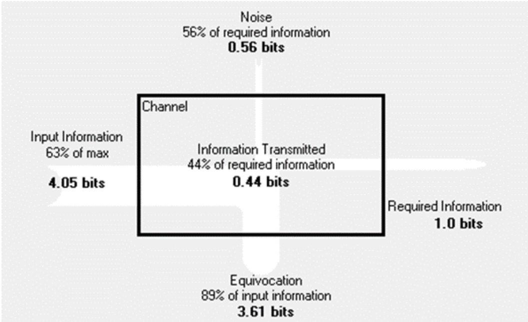

***Figure. 3.** Infometric map of the dataset. The relationship between inputs and output is 44%.*
***Source:** own elaboration*

In the software representation, 44% corresponds to 0.44 bits of transmitted information, and the remaining percentage corresponds to untransmitted information or errors within that scheme. This value's practical utility should be analyzed along with the confusion matrix and external clustering indicators.

As indicated, the infometric map supports the existence of an interdependence between the input and output variables, but it does not show the possibility of creating a high-performance predictive model.

3.4. Clustering

To validate the robustness of the (IT)-based analysis, a classification model using clustering techniques was implemented, considering the segregation and comparison of two distinct cohorts: one composed of normal ECG records and another representative of pathological cardiac morphologies. Powerhouse® software was used.

Unlike conventional approaches based on Euclidean distances, the evaluation of intra-cluster similarity (cohesion) and inter-cluster similarity (separation) measures was based on transmitted mutual information, according to the previous analysis. This criterion makes it possible to quantify the statistical dependence between input vectors (ECG samples) and the output classification of the dataset. Thus, cluster formation responds not only to the geometric proximity of signals, but also to informational coherence and to the reduction of diagnostic uncertainty inherent to the system.

Powerhouse® calculated the mutual information, untransmitted information, and error and represented the clusters using the density map shown in Figure 4. Cluster 1 had a higher concentration of abnormal segments, whereas cluster 2 had a higher concentration of normal segments.

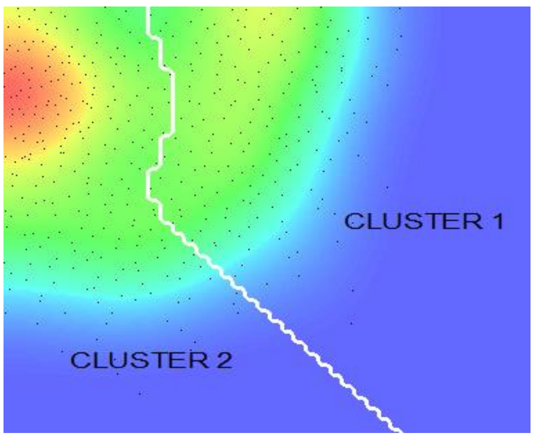

Figure. 4. *Density map.*
Source: *own elaboration*

Figure 5 shows the distribution of categories in the initial set and each cluster. The representation shows a partial separation between normal and abnormal segments; therefore, the results suggest discriminative capacity, but do not imply perfect separation or clinical validation of the procedure.

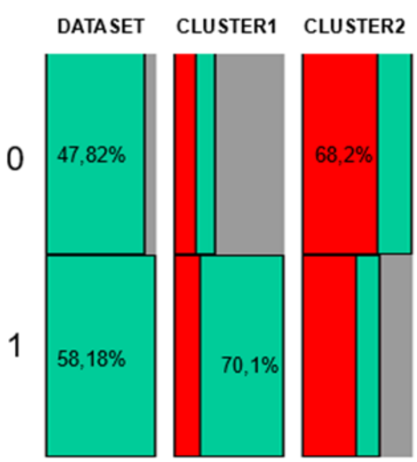

Figure. 5. *Data distribution in each cluster.*
Source: *own elaboration*

Figure 6 shows the normalized confusion matrix. Considering the category 1 as abnormal, a sensitivity of 94.38% and a specificity of 70.53% were obtained. From the class sizes and rounded percentages, an overall accuracy of 87.75%, precision of 89.27%, balanced accuracy of 82.46%, and F1 of 91.75% were estimated. The approximate external clustering metrics obtained were a purity of 0.877, a Rand Index of 0.785, an Adjusted Rand Index of 0.545, and an NMI of 0.391. These values were derived from the reconstructed contingency table and should be confirmed with the original counts if they are available.

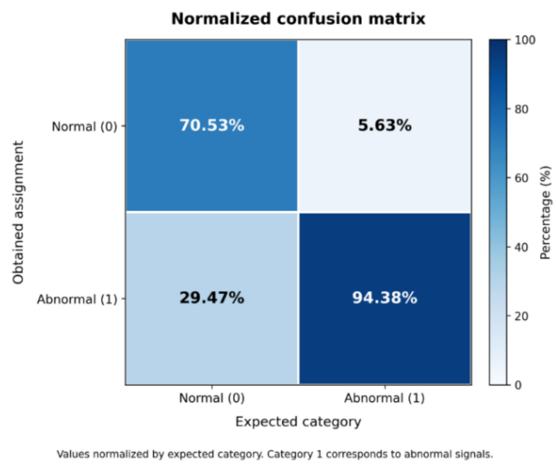

Figure 6. *Distribution of the data in each cluster*
Source: *own elaboration*

3.5. Discussion and limitations

Measures of entropy and mutual information have been used in previous research as descriptors along with machine learning classifiers [16]–[20]. In contrast, the present work estimated the information transmission to investigate the amount of information contained after the data discretization. Hence, 44% of the information transmitted shows a relationship between inputs and output, but not an estimate of a computational model's performance. The sensitivity, specificity, and external metrics of the cluster model must be analyzed.

4. CONCLUSIONS

There exists a statistically significant difference between normal and abnormal ECGs, both in terms of amplitude and distribution shape, confirmed by non-parametric tests (Mann–Whitney U and Kolmogorov–Smirnov). These results validate the use of automatic classification techniques and justify the extraction of temporal and morphological features for cardiac anomaly detection.

Most of the time-series samples are significantly different, as demonstrated by the statistical analysis. This suggests that cardiac pathology affects not only an isolated point but also the overall structure of the ECG signal.

To group ranges of values into categories or intervals and thus reduce variability (eliminate some of the fine structure) in a dataset, LIL was used for discretization. This technique allowed for the transformation of numerical data into categorical data while preserving the transmitted information. A transmission rate of 44% is reported, indicating a relationship between inputs and outputs, but this is not a measure of model performance.

The performance of model with the confusion matrix shows balanced accuracy for a model that allows discrimination between data points, but it should be validated with another dataset.

Zero-padding particularly influences the last positions of the segments, affecting statistical measures and those associated with the time index (TI); therefore, the results related to these positions should be evaluated with caution.

Owing to the non-parametric nature of the data, any classification model (such as neural networks or support vector machines) should be trained considering that the arithmetic mean is not a reliable descriptor; the use of the median and the RMS is suggested.

The results generally support the use of TI-related descriptors for ECG discrimination. However, independent validation is required before establishing its usefulness as a predictive model or extrapolating the findings to other clinical contexts.

Author Contributions: All authors have read and approved the final version of the manuscript.

Funding: Product derived from the research project titled "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)". CODE: PPEE-CV SUE DC 02 2023. Project funded by Sistema Universitario Estatal (SUE)

Data Availability: The preprocessed dataset used in this study corresponds to the binary subset of the ECG Heartbeat Categorization Dataset, derived from the PTB Diagnostic ECG Database and publicly available on Kaggle [27]. The analysis was performed on heartbeat segments and not patient-identified records.

REFERENCES

[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