



Machine learning-based stratification of chagas heart failure severity using ECG power spectral biomarkers

Pedro Ribeiro¹ · João Alexandre Lobo Marques² · Maria Inês Barbosa¹ · Roberto C. Pedrosa³ · João Paulo do Vale Madeiro⁴ · Pedro Miguel Rodrigues¹

Received: 28 November 2025 / Accepted: 3 April 2026
© The Author(s) 2026

Abstract

Purpose This study presents a machine learning methodology to automatically classify heart failure severity in Chagas disease (CD) patients using non-invasive 24-hour ECG-Holter signals.

Methods Following American Heart Association (AHA) guidelines, the cohort was stratified into three Left Ventricular Ejection Fraction (LVEF)-based severity groups: Normal (LVEF ≥ 0.50 , $n=197$), Moderate ($0.40 \leq \text{LVEF} < 0.50$, $n=106$), and Severe (LVEF < 0.40 , $n=77$), totaling $N=380$ patients. From short 10-second ECG segments, we extracted eleven spectral features derived from the power spectral density (PSD). Class imbalance was addressed through oversampling applied to the training folds. All classifiers were evaluated over 50 random stratified train-test splits (80/20) across three pairwise tasks (Normal vs. Moderate, Normal vs. Severe, Moderate vs. Severe).

Results Analysis revealed a consistent leftward shift in PSD, with increased low-frequency power in more severe cases, consistent with morphological ECG changes including P-wave attenuation, QRS alterations, and ST-segment shifts. Using this spectral biomarker, the best models achieved mean *AUC/PR-AUC* values of 0.79/0.76 for Normal vs. Severe and 0.83/0.85 for Moderate vs. Severe across 50 random states. The Normal vs. Moderate task showed moderate separability (*AUC* = 0.75, *PR-AUC* = 0.72).

Conclusion These findings highlight the potential of power spectral ECG analysis as a low-cost, fully automated tool for risk stratification in CD. The methodology shows promise for improving triage and clinical decision-making in resource-limited settings where CD remains highly prevalent.

Keywords Chagas disease · Left ventricular ejection fraction · Power spectral density · Machine learning · Heart failure severity · Discrimination

✉ Pedro Miguel Rodrigues
pmrodrigues@ucp.pt

Pedro Ribeiro
s-pmsbribeiro@ucp.pt

João Alexandre Lobo Marques
alexandre.lobo@usj.edu.mo

Maria Inês Barbosa
mibarbosa@ucp.pt

Roberto C. Pedrosa
coury@hucff.ufjf.br

João Paulo do Vale Madeiro
jpaulo.vale@dc.ufc.br

¹ Universidade Católica Portuguesa, CBQF – Centro de Biotecnologia e Química Fina – Laboratório Associado, Escola Superior de Biotecnologia, Rua de Diogo Botelho 1327, 4169-005 Porto, Portugal

² University of Saint Joseph, Laboratory of Applied Neurosciences, Macao 999078, China

³ Edson Saad Heart Institute, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil

⁴ Federal University of Ceará, Department of Computing, Fortaleza, Ceará, Brazil

1 Introduction

Chagas disease (CD), caused by the parasite *Trypanosoma cruzi*, was first described in 1909 by Carlos Chagas. It is more prevalent in the Americas, especially in Central and South America [1, 2].

CD can manifest with both acute and chronic symptoms. The acute phase may present with fever, subcutaneous oedema, malaise, and enlargement of the spleen, liver, and lymph nodes. Additionally, some patients exhibit electrocardiogram (ECG) abnormalities such as sinus tachycardia, atrioventricular block, and T-wave changes [3].

In the chronic phase, the disease ranges from asymptomatic to highly symptomatic, with indicators grouped into cardiac, digestive, and cardiogastrointestinal categories. Cardiac manifestations are the most frequent and consequential, affecting about 20–30% of chronic cases and causing ECG abnormalities such as right bundle branch block, left anterior fascicular block, ventricular premature beats, ST–T changes, abnormal Q waves, and low QRS voltage [3].

Diagnosis of acute CD is typically performed via microscopic detection, aiming to identify trypomastigotes in blood samples. For the chronic phase, confirmation requires at least two serological tests to detect antibodies against *Trypanosoma cruzi* [3]. Treatment during the acute phase is primarily pharmacological, with medications such as Benznidazole and Nifurtimox targeting the parasite [4]. In the chronic phase, Benznidazole remains the only available treatment in South American countries like Brazil, though no fully effective therapy exists for this stage [5].

ECG signals represent the heart's electrical activity [6], and the characteristic changes induced by CD [7] make them a promising source of early diagnostic information [6]. Modern machine learning (ML) algorithms, originating from work in cybernetics and computer science in the 1950s [8, 9], can automatically learn patterns from such data. In this context, multi-band non-linear analysis (MLBNA) decomposes ECG signals into frequency bands and extracts non-linear features [10], providing rich ECG descriptors for ML models. Leveraging these features can accelerate detection and enhance early diagnosis and treatment of cardiovascular diseases [11].

The novelty of our approach lies in analyzing broadband spectral and power spectral ECG features extracted from 24-hour Holter data, rather than relying solely on classical Heart Rate Variability (HRV) spectral indices. Traditional HRV measures capture only narrow-band autonomic modulation and therefore overlook broader alterations in the ECG frequency structure that accompany disease progression [12]. By employing a richer and physiologically grounded representation of the signal, our framework enables the characterization of a shift in spectral power from higher to lower frequencies - a pattern consistent with

conduction slowing and myocardial remodeling in Chagas cardiomyopathy [13]. Such broadband and power spectral descriptors remain underexplored in the literature, despite their potential to reveal early electro-physiological signatures of cardiac dysfunction.

Building on this rationale, we extract a comprehensive set of eleven Power Spectral Density (PSD)-based features, including broadband spectral metrics and power ratios across the 1–25 Hz range. These descriptors are then used to train and validate ML models designed to predict Left Ventricular Ejection Fraction (LVEF) severity, a key marker of CD-induced heart failure. This work therefore presents a novel ML-based framework that leverages richer ECG-derived information to improve the non-invasive stratification of CD-related cardiac dysfunction.

2 Literature review

To establish a comprehensive overview of the state-of-the-art, a systematic literature review was conducted from August to October 2025, with Google Scholar serving as the primary search engine. This review focused on scholarly articles from prominent publishers (i.e., Elsevier, MDPI, Springer Nature, Wiley, SAGE, Taylor & Francis, and PLOS), as well as conference proceedings from SBIS and CinC, and publications from the Al-Furat Al-Awsat Technical University. The search covered the period from 2018 to the present to ensure the inclusion of recent and relevant advances in the field. A set of specific keywords guided the search: “Chagas Disease classification,” “Chagas disease Deep learning,” “Chagas disease Ejection fraction classification,” “Chagas Disease ECG screening,” “Left ventricular ejection fraction machine learning,” and “Cardiovascular diseases LVEF classification.” Table 1 summarizes recent state-of-the-art approaches (2019–2024) for heart failure diagnosis from CD or other cardiac conditions based on LVEF severity. The review includes 12 key LVEF-focused studies, 4 of which specifically address CD, illustrating substantial progress in ML for this domain. Methods range from conventional algorithms (e.g., Random Forest) to deep learning architectures (CNNs, ANNs, ResNets). Reported accuracies span 74.60%–92.41%, with 9 studies achieving an Area Under the Receiver Operating Characteristic Curve (AUC) > 0.80. The best performance, by Alkhodari et al. [14], reached 92.41% accuracy and AUC = 0.975 using a deep learning classifier on clinical features.

Analysis of Table 1 shows substantial variation in dataset sizes and validation strategies. Sample sizes range from 33 subjects in Ferreira et al. [17] to over 147000 in Vaid et al. [21]. This disparity is critical, as 9 of the 12 studies explicitly report limitations due to small or unbalanced

Table 1 Summary of state-of-the-art studies on diagnosing heart failure cardiac conditions deivered from CD or other diseases. Each entry details the database, comparison groups, features, classifier, limitations, and accuracy

Ref	Year	Features Extracted	Comparison Group (number of samples)	Classifier	Limitation	Validation	Acc	AUC
LVEF derived from CD								
Brito, B. et. al. [15]	2021	Features extracted from neural network	LVSD (93) vs. No-LVSD (1211)	AI-ECG	Unbalanced dataset.	Hold-out	89%	0.874
Silva, L. et. al. [16]	2021	28 HRV features	Preserved (38) vs. Reduced (15) vs. Borderline (10)	Random Forest	Small and unbalanced dataset.	Cross-validation	74.60%	-
Ferreira, G. et. al. [17]	2023	Features extracted from ANN	not preserved LVEF (17) vs. preserved LVEF (16)	ANN	Small test set.	Hold-out	79%	-
Madeiro, J. et. al. [18]	2023	Statistical measures related to waveform amplitudes and durations, wavelet decomposition coefficients, heart rate variability parameters and non-linear analysis	LVEF (19) vs. non-LVEF (36)	GBC	Unbalanced dataset.	Hold-out	-	0.75
LVEF derived from other diseases								
Attia, Z. et. al. [19]	2019	Features extracted from CNN	ALVD (4131) vs. non-ALVD (48739)	CNN	Unbalanced dataset.	Hold-out	85.7%	0.93
Kwon, J. et. al. [20]	2019	Sex, rhythm, age, weight, heart rate and QT interval	HFrEF (5245) vs. no-HFrEF (5245)	DEHF	They included patients with multiple ECGs; this could slightly bias survival analysis since multiple entries per individual.	Hold-out	-	0.889
Alkhodari, M. et. al. [14]	2021	Age, BMI, Diabetes, Anti-arrhythmics, Diuretics, VT, prior MI	HFpEF (129) vs. HFnEF (92) vs. HFrEF (82)	CNN	Small dataset.	Cross-validation	92.41%	0.975
Vaid, A. et. al. [21]	2022	Features extracted from CNN	LVEF (17840) vs. non-LVEF (129201)	CNN	Unbalanced dataset.	Hold-out	-	0.94
Choi, J. et. al. [22]	2022	features extracted from CNN	HFrEF (600) vs. No-HFrEF (691)	DeepECG-HFrEF	They included patients with multiple ECGs; this could slightly bias survival analysis since multiple entries per individual.	Hold-out	-	0.844
Akerman, A. et. al. [23]	2023	Features extracted from 3D-CNN	HFpEF (646) vs. non-HFpEF (638)	3D-CNN	Possibility that some control participants had subclinical diseases. Complete matching for age was not possible where patients with HFpEF were older.	Hold-out	-	0.95

Table 1 (continued)

Ref	Year	Features Extracted	Comparison Group (number of samples)	Classifier	Limitation	Validation	Acc	AUC
Decoodt, P. et. al. [24]	2024	Features extracted from AutoML	rEF (208) vs. mEF (64) vs. pEF (143)	AutoML	Unbalanced dataset.	Hold-out	75.9%	0.856
Younis, S. et. al. [25]	2024	TP (ms), QT (ms), ST-T (ms), QRS (ms), Mean-Entropy and Mean-Instant requency	HFmEF (92) vs. HFpEF (129) vs. HfrEF (82)	Decision Trees	Small dataset.	Cross-validation	91.2%	0.98

datasets, highlighting a persistent challenge for generalizable models. Validation methods also vary: 9 studies used hold-out validation, whereas only 3 employed cross-validation. This predominance of hold-out validation, especially in smaller cohorts, suggests that future work should adopt more rigorous validation protocols to improve reliability and reproducibility.

3 Methodology and materials

Our methodology (Fig. 1) comprised ECG acquisition, automated artifact detection, normalization, and feature extraction. After extracting the eleven *PSD*-based features, the dataset was evaluated across 50 random train–test splits (80/20). The entire preprocessing pipeline, including

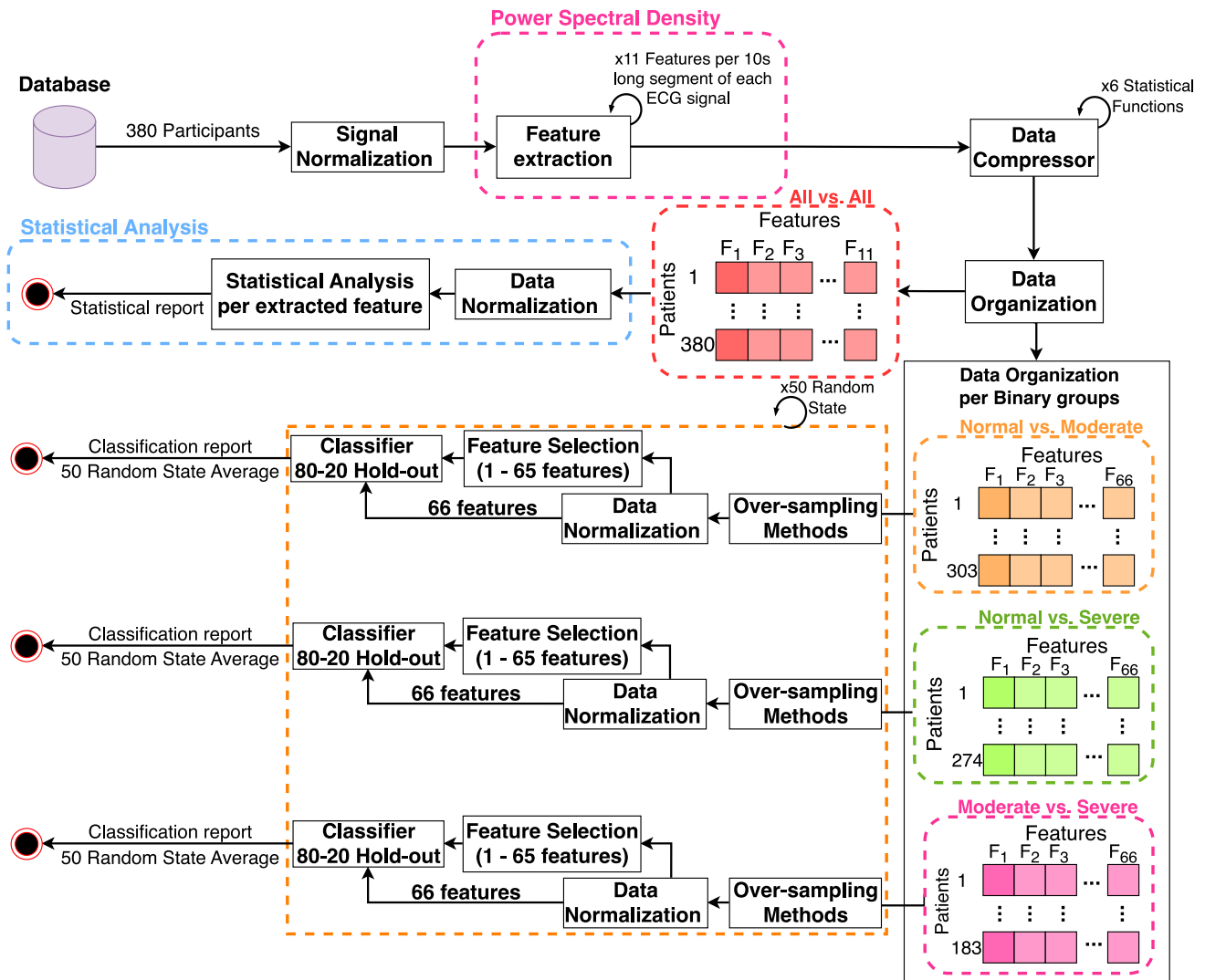


Fig. 1 Methodology workflow diagram

automated oversampling, feature selection, and classifier training, was fitted exclusively on the training split for each iteration, ensuring no information leakage. Five machine learning classifiers were then evaluated on the independent test sets using metrics such as *AUC*, Precision–Recall *AUC* (*PR-AUC*), *Accuracy*, *Recall*, and *Precision*.

3.1 The dataset

3.1.1 Population characterization

The data originate from a cross-sectional, descriptive study involving 380 patients (56% female) treated for Chagas Heart Disease (ChHD) at Hospital Universitário Clementino Fraga Filho (HUCFF) and the Federal University of Rio de Janeiro (UFRJ) between 1992 and 2016. All patients were managed according to the protocols of the Brazilian Ministry of Health [26].

From the initial cohort of 556 individuals, strict exclusion criteria were applied to ensure that the final study population represented a clinically homogeneous group with *confirmed* Chagas cardiomyopathy. A total of 82 patients were excluded due to malignancy, psychiatric disorders, severe hepatic disease (cirrhosis), advanced renal failure (dialysis or renal transplantation), or other forms of non-cardiac organ failure. Sixteen patients were excluded due to loss to follow-up, 17 were removed because they had received anti-*T. cruzi* treatment (based on disease duration, absence of severe cardiomyopathy, and patient preference), and 7 were excluded due to having undergone isolated cardiac resynchronization therapy. After these exclusions, 434 eligible patients were invited to participate in a structured clinical follow-up programme. At enrollment, additional exclusion criteria were applied: atrial fibrillation, anomalies in ECG analogue-to-digital conversion, or the presence of a ventricular-paced rhythm. This resulted in a final cohort of 380 patients. Importantly, all patients included in the final dataset had a definitive diagnosis of Chagas cardiomyopathy, established by at least two positive serological tests for *T. cruzi* infection together with compatible electrocardiographic abnormalities, following established diagnostic criteria. No patients with other etiologies of cardiomyopathy or alternative cardiac pathologies were included.

3.1.2 Ethical approval

Data acquisition and usage were approved by the institutional ethics committee (approval no. 45360915.1.1001.5262) of HUCFF. In accordance with the Declaration of Helsinki and Brazilian research ethics standards, the requirement for written informed consent was waived.

3.1.3 Data acquisition and target variable

Each patient record included a 24-hour ECG Holter (Lead II) acquisition sampled at 128Hz. The LVEF, calculated using the Teicholz method, served as the target variable for classification. All ECG Holter recordings included in this study were obtained exclusively from patients with confirmed chronic Chagas cardiomyopathy, as ensured by at least two positive serological tests for *T. cruzi* antibodies together with characteristic ECG abnormalities. This guaranteed that no signals from non-Chagas etiologies were inadvertently incorporated into the dataset. The 24-hour Holter recordings and the echocardiographic examinations used to estimate LVEF were performed with a median interval of 6.3 months (interquartile range: 3.2–8.9 months). Both assessments were systematically conducted between 8:00 a.m. and 12:00 p.m., a time window characterized by heightened sympathetic activity [27, 28], thereby maximizing autonomic modulation of cardiac electrophysiology. Harmonizing the timing of these examinations minimized circadian physiological variability and strengthened the temporal alignment between ECG-derived features and the corresponding LVEF measurements. During Holter monitoring, patients documented symptoms, activity levels, and relevant daily events to contextualize physiologic fluctuations observed in the raw signal. Thus, all computational analyses performed on present work relied exclusively on the previously mentioned 4-hour segment, selected to optimize signal quality and minimize artifacts. All remaining portions of the 24-hour ECG were excluded to ensure methodological uniformity across participants.

Echocardiographic measurements followed the recommendations of the American Society of Echocardiography. Although LVEF was calculated using the Teicholz M-mode method, an additional subjective visual assessment by experienced cardiologists was incorporated, given that regional wall-motion abnormalities and segmental asynergia, common in chronic Chagas cardiomyopathy, may impair formula-based LVEF estimation accuracy. Visual assessment has been shown to outperform automated estimation under such heterogeneous contraction patterns [29]. All echocardiographic studies were recorded and independently reviewed by two observers to ensure interpretative consistency. Likewise, Holter recordings were analyzed twice by the same observer (at acquisition and during dataset preparation) to ensure consistent annotation, artifact detection, and signal-quality verification across the dataset.

Following the American Heart Association (AHA) guidelines [30–33], the dataset was stratified into three heart-failure severity groups based on LVEF: Normal ($LVEF \geq 0.50$, 197 samples), Moderate ($0.40 \leq LVEF < 0.50$, 106 samples), and Severe ($LVEF < 0.40$, 77 samples). These categories were consistently applied throughout the study,

Table 2 Demographic and clinical characteristics of the study population (Mean \pm SD), presented for the full cohort and stratified by LVEF-based classes. F = female; M = male

	General		Normal		Moderate		Severe	
Sex	220 F	160 M	110 F	87 M	62 F	44 M	49 F	28 M
Age	44.55 \pm 9.30	43.59 \pm 9.86	44.21 \pm 9.14	43.34 \pm 9.67	45.43 \pm 9.19	40.20 \pm 9.88	44.18 \pm 9.91	46.71 \pm 9.20
BMI	29.15 \pm 13.45	27.80 \pm 2.86	29.75 \pm 19.02	28.18 \pm 3.08	28.81 \pm 2.58	27.51 \pm 2.56	28.24 \pm 2.31	27.07 \pm 2.45
LVEF	0.52 \pm 0.16		0.66 \pm 0.07		0.44 \pm 0.02		0.31 \pm 0.07	

and a detailed characterization of the cohort is provided in Table 2. ANOVA analyses revealed that only LVEF differed significantly across the three groups ($p < 0.05$), with all pairwise post-hoc comparisons confirming robust between-group differences, whereas Sex, BMI, and Age showed no statistically significant variation across groups.

3.2 Experimental setup

All computational tasks were executed on a MacBook Pro (14-inch, 2021) equipped with an Apple M1 Pro system-on-a-chip (8-core CPU, 14-core GPU) and 16 GB of unified memory. The software tool-chain consisted of Matlab (R2023b, The MathWorks, Inc.) for spectral feature extraction, signal preprocessing, and data organization, while Python (3.9.12, Python Software Foundation) was used for the statistical analysis and reporting & evaluation of ML model performance.

3.3 Signal filtering

Each ECG signal was filtered using a 16th-order elliptic band-pass filter with a pass-band of 1–25 Hz (steepness = 0.85; stop-band attenuation = 60 dB). This upper cutoff frequency was selected after re-examining the spectral content of our recordings, which confirmed that the relevant physiological power of the ECG is concentrated below 25 Hz. Defining the filter in this range ensures that all meaningful ECG components are preserved while higher-frequency noise is effectively suppressed. This choice is further supported by recent findings in the literature. Ádám et al. (2025) [34] demonstrated that an upper frequency boundary around 20–25 Hz

captures all important ECG signal characteristics, with no relevant information being lost. This evidence reinforces the adoption of a 25 Hz upper cutoff in our processing pipeline.

3.4 Artifact removal

To ensure signal quality, a custom artifact removal algorithm was developed, as outlined in Algorithm on Table B1 at Supplementary Material B to remove artifact of the same nature as the one in the Fig. 2. To establish a robust baseline for artifact detection, the median value of each previously filtered signals is computed. The core of the algorithm involves identifying all local peaks in the signal and comparing their amplitudes against dynamic thresholds derived from the median. Any peak with an amplitude exceeding 50 times the median (indicating a large spike artifact) or falling below 0.015 times the median (suggesting noise or a baseline distortion) is classified as an artifact and their located zone (the time-series between the previously peak and the identified peak as an artifact) is removed from the signal. The peaks are processed in descending order of amplitude to ensure that the largest artifacts are removed first.

After removing artifacts from the ECG recordings, we analyzed the remaining signal segments for each subject. Rather than reconstructing a continuous signal, which could introduce artificial information or bias, we performed all analyses directly on these segments using a sliding-window approach. To ensure the robustness of the procedure, we discarded all segments shorter than 10 s, retaining only those with a duration exceeding 10 s for subsequent analysis. As a direct consequence of this quality-driven segmentation, the total amount of analyzable data inherently varies across

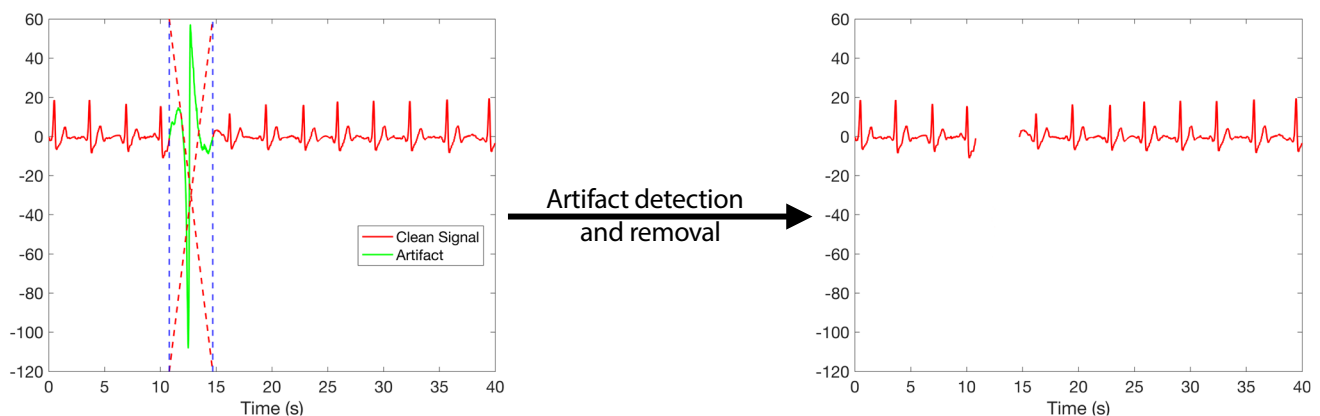


Fig. 2 40-seconds of ECG signal's example for artifact detection and removal

subjects, depending on the number and duration of artifact-free segments preserved. After preprocessing, the mean SNR of the recordings was 12 dB, which falls within the typical range reported for long-term ambulatory ECG monitoring (5–15 dB) [35]. This confirms that the cleaned signals preserved sufficient quality for reliable spectral analysis.

3.5 Signal normalization

After artifact removal, the ECG filtered signals, $x(n)$, were normalized according to the Root Mean Square (RMS) normalization formula:

$$x(n) = \frac{x(n)}{\sqrt{\frac{\sum_{n=0}^{N-1} x^2(n)}{N}}}, \tag{1}$$

where N is the ECG length. Then, the mean value was removed from the entire signal.

3.6 ECG time-frequency analysis through power spectral density and feature extraction,

Spectral analysis is a fundamental technique for interpreting ECG signals, typically involving the decomposition of the time-domain signal into its constituent frequency components [36]. For stationary signals, this transformation is commonly achieved using the Fourier Transform (FT), which extracts spectral information by projecting the signal onto a basis of sinusoidal functions.

Given an ECG signal $x(t)$ sampled at a frequency $f_s = 1/T_s$, where T_s is the sampling period, the resulting discrete-time signal is $x[n] = x(nT_s)$ for $n = 0, \dots, N - 1$, where N is the total number of samples. The spectrum of this signal, $X[k]$, can be obtained using the Discrete-Time Fourier Transform (DTFT). The DTFT assumes that $x[n]$ is periodic with a fundamental period N , corresponding to a fundamental frequency of $\Omega_0 = 2\pi/N$. The discrete frequency variable k then corresponds to harmonics of this fundamental frequency, such that $\Omega = k\Omega_0$.

The forward and inverse DTFT are defined by the transform pair:

$$X[k] = \sum_{n=0}^{N-1} x[n] \cdot e^{-jk\Omega_0 n}, \quad k = 0, \dots, N - 1 \tag{2}$$

$$x[n] = \frac{1}{N} \sum_{k=0}^{N-1} X[k] \cdot e^{jk\Omega_0 n}, \quad n = 0, \dots, N - 1 \tag{3}$$

where $X[k]$ are the complex Fourier coefficients. This transform pair provides a complete representation of the signal in

the frequency domain, allowing for lossless recovery of the time-domain signal from its spectral components and vice-versa [37, 38].

The distribution of signal power across different frequencies is a key characteristic, formally described by the *PSD*. For a discrete-time signal, a common method for estimating the *PSD* is via the Wiener-Khinchin theorem, which states that the *PSD* is the Fourier transform of the signal’s autocorrelation function [39]. The unbiased autocorrelation estimate, $R_{xx}[u]$, for a lag u is given by:

$$R_{xx}[u] = \begin{cases} \frac{1}{N} \sum_{n=0}^{N-m-1} x[n] \cdot x^*[n+u], & \text{if } u \geq 0 \\ R_{xx}^*[-u], & \text{if } u < 0 \end{cases} \tag{4}$$

where x^* denotes the complex conjugate (for real signals, $x^* = x$). The *PSD* is then estimated as the DTFT of this autocorrelation sequence:

$$PSD[k] = \frac{1}{N} \sum_{u=0}^{N-1} R_{xx}[u] \cdot e^{-j\frac{2\pi k u}{N}}, \quad k = 0, \dots, N - 1 \tag{5}$$

To facilitate the computation of spectral features, the *PSD* was normalized. Specifically, each *PSD* was divided by its total power (the integral across all frequencies), thereby scaling the total spectral content to unity (1).

$$PSD_n[k] = \frac{PSD[k]}{\sum_{k=0}^{N_T-1} PSD[k]} \tag{6}$$

$$PSD_n(f) = PSD_n \left[\frac{k * f_s}{N_T} \right] \tag{7}$$

where N_T is the length of the PSD_n and f is the frequency bins.

Figure 3 shows the *PSD* for the 3 classes, while Fig. C1 at Supplementary Material C includes the relative power

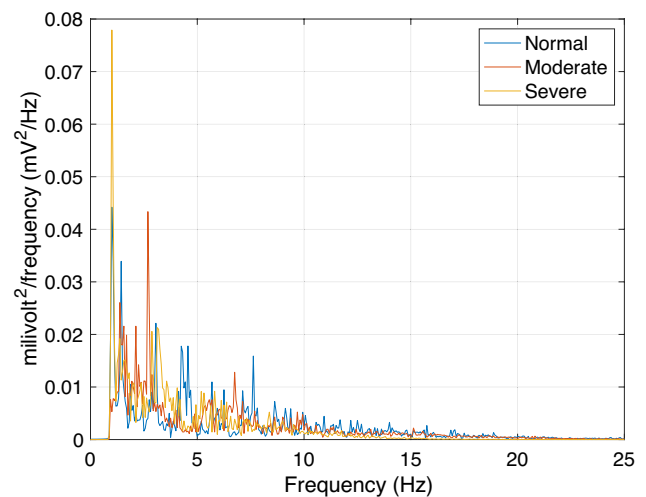


Fig. 3 *PSD* plot for the 3 classes

percentages for each plot, A noticeable trend in the latter is the leftward shift of the power spectra, indicating that spectral power concentrates at lower frequencies as the disease develops.

3.7 Feature extraction

With the spectral “identification” observed in Fig. C1 in Supplementary Material C, and because this phenomenon is fundamentally expressed as changes in *relative* spectral power rather than absolute amplitude, ratio-based and power-derived metrics provide a sensitive and physiologically meaningful way to capture these shifts.

In the first stage of a two-stage feature extraction process, we followed the standard duration of ECG recordings in clinical practice [40] and applied a 10-second, non-overlapping sliding window to the clean segments to compute their *PSD*. From each *PSD* we extracted eleven *PSD*-based features, summarized in Table 3.

To compute these related power features (relative power and ratios), the 1–25 Hz physiological band was divided into four equal-width sub-bands. This uniform segmentation provides a balanced and unbiased partitioning of the spectrum, ensuring that the *PSD* is represented consistently across its entire physiologically relevant range. Importantly,

equal-width sub-bands are particularly well suited to characterizing the progressive left-shift in spectral power observed with advancing disease: as energy becomes increasingly concentrated at lower frequencies, the symmetric partition allows this redistribution to be quantified without imposing a priori assumptions about which regions should dominate. This yields descriptors that are mathematically stable and physiologically interpretable, capturing disease-related changes in the structure of the ECG spectrum.

For each of the eleven time series, we computed six summary statistics (stats): the mean, median, standard deviation (SD), variance (Var), 95th percentile (P95), and kurtosis (Kur) [41]. These summary measures are hereafter denoted by the configuration `feature_stat`. This resulted in a final set of 66 features (11 primary features \times 6 statistical metrics) for each patient record, totaling 380 patients.

3.8 Feature normalization

Prior to model training, all 66 features were standardized using z-score normalization, where each value x_i - where $1 \leq i \leq 66$ - is transformed to \tilde{x}_i according to:

$$\tilde{x}_i = \frac{x_i - \mu}{\sigma} \quad (8)$$

Table 3 Features extracted, corresponding equations, and short description

Feature	Equation	Description
Median Frequency (<i>MF</i>)	$MF = f_i$ such that $\frac{\sum_{1 \leq f \leq f_i} PSD(f)}{\sum_{1 \leq f \leq 25} PSD(f)} = 0.5$	Frequency at which 50% of the PSD within 1–25 Hz is accumulated.
Relative Power of low Frequencies (<i>RP1</i>)	$RP1 = \sum_{1Hz}^{6.25Hz} PSD_n(f)$	Relative power in the low-frequency band (1–6.25 Hz).
Relative Power of intermediate Frequencies (<i>RP2</i>)	$RP2 = \sum_{6.25Hz}^{12.5Hz} PSD_n(f)$	Relative power in the intermediate band (6.25–12.5 Hz).
Relative Power of intermediate high Frequencies (<i>RP3</i>)	$RP3 = \sum_{12.5Hz}^{18.75Hz} PSD_n(f)$	Relative power in the mid-high band (12.5–18.75 Hz).
Relative Power of high Frequencies (<i>RP4</i>)	$RP4 = \sum_{18.75Hz}^{25Hz} PSD_n(f)$	Relative power in the high-frequency band (18.75–25 Hz).
Ratio between <i>RP1</i> and <i>RP2</i> (<i>R₁</i>)	$R_1 = \frac{RP1}{RP2}$	Ratio of low- to intermediate-frequency power. Higher values indicate dominance of low-frequency activity.
Ratio between <i>RP1</i> and <i>RP3</i> (<i>R₂</i>)	$R_2 = \frac{RP1}{RP3}$	Ratio of low- to mid-high-frequency power.
Ratio between <i>RP1</i> and <i>RP4</i> (<i>R₃</i>)	$R_3 = \frac{RP1}{RP4}$	Ratio of low- to high-frequency power.
Ratio between <i>RP2</i> and <i>RP3</i> (<i>R₄</i>)	$R_4 = \frac{RP2}{RP3}$	Ratio of intermediate to mid-high frequency power.
Ratio between <i>RP2</i> and <i>RP4</i> (<i>R₅</i>)	$R_5 = \frac{RP2}{RP4}$	Ratio of intermediate to high-frequency power.
Ratio between <i>RP3</i> and <i>RP4</i> (<i>R₆</i>)	$R_6 = \frac{RP3}{RP4}$	Ratio of mid-high to high-frequency power.

To prevent data leakage and ensure analytical validity, the normalization parameters (mean (μ) and standard deviation (σ)) were computed from the training data and then applied to both the training set and the test set (data splitting explained in the following point Section 3.10). For pairwise comparisons, feature normalization used a reference population comprising only the two groups being analyzed, whereas for all-versus-all comparisons, a pooled reference population from all groups was used.

3.9 Statistical analyses

To assess differences among the Normal, Moderate, and Severe groups, a statistical analysis was conducted on the mean values of the eleven primary features. The assumption of normality for each feature was evaluated using the Kolmogorov–Smirnov test, and homogeneity of variances across the three groups was assessed using Levene’s test. Once these assumptions were confirmed, a one-way Analysis of Variance (ANOVA) was performed to test for significant differences among the three groups, with statistical significance defined as $p < 0.05$; post-hoc pairwise comparisons were also performed between the groups. To control for multiple comparisons across these pairwise tests, the Bonferroni correction was applied.

3.10 Combined features power to discriminate between groups via ML approaches

To evaluate pairwise classification performance across study groups, we employed five ML models from scikit-learn [42], detailed in Table 4, using a repeating stratified 80/20 split across 50 different random seeds, effectively emulating the robustness of repeated cross-validation, as the model is evaluated on 50 independent test sets.

A comprehensive preprocessing and training pipeline was fitted exclusively on the training data, ensuring a robust

Table 4 Configurations of the five Scikit-learn classifiers used in this study

Classifier (Scikit-learn class)	Hyperparameters
AdaBoost (AdaBoostClassifier)	(n_estimators = 50, learning_rate = 1.0, algorithm = “SAMME.R”) + default
Decision Tree (DecisionTreeClassifier)	(max_depth = 5, criterion = “gini”, splitter = “best”, min_samples_split = 2) + default
K-Nearest Neighbors (KNeighborsClassifier)	(n_neighbors = 5, weights = “uniform”, algorithm = “auto”) + default
Linear SVM (LinearSVC)	(penalty = “l2”, loss = “squared_hinge”, dual = True) + default
RBF SVM (SVC)	(γ = “auto”, kernel = “rbf”, C = 1.0, probability = False) + default

evaluation fully insulated from data leakage. This entire procedure was executed independently for each random seed, and the pipeline consisted of three sequential stages:

1. **Class Balancing:** OS techniques (specifically, RandomOverSampler - ROS and ADASYN) were applied to the features to mitigate the effects of class imbalance in the training data.
2. **Feature Normalization:** Following class balancing, we applied z-score normalization to all features’ time-series, as previously described in the Methods Section 3.8.
3. **Feature Selection:** With the features normalized, the next step was to identify the optimal feature subset for each classification task, we employed an ANOVA F-test. Features were ranked by their F-value, which measures the ratio of between-group to within-group variance and indicates discriminatory power [43]. This ranking was used to evaluate subsets of increasing size, from 1 to 66 features, to find the optimal set.
4. **Model Training:** Finally, each classifier was trained on the transformed (feature-selected and balanced) training data.

The fully trained model was then evaluated on an independent test sets across 50 random-state runs, using data that was never involved in feature selection or oversampling procedures. All reported performance metrics correspond to the average across these 50 runs, ensuring a stable and robust estimate of model performance.

3.10.1 Classification evaluation metrics

Model performance was evaluated using a suite of standard classification metrics, with the *AUC* as the primary indicator [44]. To better contextualize the model’s clinical applicability, particularly in scenarios where sensitivity-driven triage may be relevant, we also added the *PR-AUC*. The *AUC* and *PR-AUC* summarize a model’s ability to discriminate between classes across all thresholds, making it effective for comparing models, particularly on imbalanced datasets. For a more granular analysis, we also computed Accuracy (*Acc*), Recall (*Rec*), Specificity (*Spec*), Precision (*Prec*), Negative Predictive Value (*NPV*), and the F1-Score (*F1*) [45] (more information at Supplementary Material A).

4 Results and discussion

This section details the outcomes of our study, beginning with the statistical analysis of the extracted features and culminating in the performance of the ML classifiers.

4.1 Separability between groups, a multiclass and pairwise evaluation via statistical analysis

To investigate the statistical separability between classes, we employed an ANOVA across all 3 groups, followed by pairwise t-tests. The overall ANOVA identified 15 feature distributions with statistically significant differences (Supplementary Material C - Fig. C2), with the most significant result ($p=0.004$) associated with the R_1 feature using the *Kur* data compressor.

Drilling down into pairwise comparisons revealed where these differences were most pronounced. The distinction between *Normal* and *Severe* classes was the strongest, yielding 15 significant feature distributions, with a top p -value of 0.0005 from the R_1 feature (*Kur* compressor) (Figure in Fig. C4 in Supplementary Material C). The *Moderate* vs. *Severe* separation was also robust, with 11 significant distributions and a best p -value of 0.0043, from the $RP4$ feature (*95Prc* compressor) (Supplementary Material C - Fig. C5). In stark contrast, the *Normal* vs. *Moderate* comparison did not produce distributions with significant differences (Fig. C3 in Supplementary Material C). This confirms that while the classes are separable overall, the distinction between the *Normal* and *Moderate* groups is the most subtle and challenging.

4.2 Spectral power shift evaluation

Table 5 summarizes the mean values of the features across the three patient classes. As set on point 3.9., a one-way ANOVA was performed to assess spectral-power differences between classes by testing the extracted features stratified by patient class based on their mean values, revealing that five of the eleven primary features (45.45%) differed significantly across groups.

Measures indexing spectral location and power showed the strongest associations with disease severity. In particular, the mean relative power in the lower frequency band ($RP1_Mean$) increased with severity, rising from 0.28 ± 0.18 (Normal) and 0.27 ± 0.17 (Moderate) to 0.35 ± 0.19 (Severe), reflecting a significant shift of spectral power toward lower frequencies ($p = 0.022$). Similarly, the mean relative power in the intermediate frequency band ($RP2_Mean$) increased from 0.22 ± 0.09 (Normal) and 0.23 ± 0.08 (Moderate) to 0.26 ± 0.09 (Severe), also indicating a significant redistribution of spectral power toward lower frequencies ($p = 0.0376$). In contrast, the mean relative power in the high-frequency band ($RP4_Mean$) decreased with severity, from 0.12 ± 0.05 (Normal) and 0.13 ± 0.05 (Moderate) to 0.10 ± 0.06 (Severe), further supporting a significant shift of spectral power toward lower frequencies ($p = 0.0191$). Finally, it should be emphasized that, although a slight trend toward lower-frequency displacement shown by some features is observable between the Normal and Moderate groups, these differences did not reach statistical significance (Bonferroni-corrected $p \geq 0.05$), indicating that the spectral shift becomes substantial only from Moderate to Severe. This leftward spectral shift is corroborated by the *PSD* plots (Fig. C2 in Supplementary Material C) and interestingly, while our study is in a different field, these findings are coincident with previous reports of spectral slowing in other domains such as neuro-degenerative conditions [46–52]. We examined the ECG morphologies to understand the leftward shift in power observed in the spectral features. As illustrated in Fig. C7 in Supplementary Material C, the class-specific alterations visible in the ECG traces help explain this behavior: (i) in the Moderate class, reductions in P-wave amplitude, occasional QRS-complex changes and mild ST-segment elevation reduce high-frequency content; (ii) in the Severe class, the absence of a clear P wave and a

Table 5 Mean \pm SD distribution of primary features stratified by patient class. Group differences were evaluated using one-way ANOVA with Bonferroni-adjusted post-hoc tests, and independent-samples t-tests when multiple-comparison differences were significant. *N.S.* indicates $p \geq 0.05$. Features correspond to MF, $RP1$ – $RP4$, and their ratios R_1 – R_6

Feature	p -value	Mean \pm SD (Normal, 197 samples)	Mean \pm SD (Moderate, 106 samples)	Mean \pm SD (Severe, 77 samples)	p -value Normal vs. Moderate	p -value Normal vs. Severe	p -value Moderate vs. Severe
MF_Mean	N.S.	9.93 ± 3.60	9.99 ± 3.26	8.81 ± 3.08	N.S.	0.017	0.014
RP1_Mean	0.022	0.28 ± 0.18	0.27 ± 0.17	0.35 ± 0.19	N.S.	0.0047	0.0053
RP2_Mean	0.0376	0.22 ± 0.09	0.23 ± 0.08	0.26 ± 0.09	N.S.	0.0036	0.037
RP3_Mean	N.S.	0.15 ± 0.06	0.16 ± 0.06	0.17 ± 0.06	N.S.	0.039	N.S.
RP4_Mean	0.0191	0.12 ± 0.05	0.13 ± 0.05	0.10 ± 0.06	N.S.	0.0046	0.0051
R_1_Mean	N.S.	1.44 ± 1.14	1.41 ± 1.05	1.61 ± 1.34	N.S.	N.S.	N.S.
R_2_Mean	N.S.	2.72 ± 3.32	2.51 ± 2.85	3.64 ± 6.58	N.S.	N.S.	N.S.
R_3_Mean	N.S.	5.40 ± 9.26	7.80 ± 26.63	38.01 ± 241.6	N.S.	N.S.	N.S.
R_4_Mean	N.S.	1.82 ± 1.31	1.91 ± 1.70	1.97 ± 1.45	N.S.	N.S.	N.S.
R_5_Mean	0.0329	3.12 ± 4.62	3.95 ± 9.70	8.73 ± 27.97	N.S.	0.0068	N.S.
R_6_Mean	0.0051	1.68 ± 1.60	1.79 ± 1.78	2.9 ± 4.71	N.S.	0.0015	0.027

markedly elevated ST segment further attenuate higher-frequency components; and (iii) these progressive morphological degradations smooth the waveform and diminish abrupt transitions, shifting energy toward lower frequencies and reinforcing the leftward spectral shift.

4.3 Classification performance

Following the feature analysis, we evaluated the performance of 5 classifiers on the 3 binary classification tasks. The criteria to select the best results per comparison group was based on the best *PR-AUC*. The optimal results for each task are summarized in Table 6.

The classification task distinguishing Moderate from Severe achieved the highest performance, with the KNN classifier reaching an *AUC* of 0.8348, a *PR-AUC* of 0.853, and an *Acc* of 83.78%. The Normal vs. Severe comparison also showed strong discriminative ability, with the SVC model achieving an *AUC* of 0.7925, a *PR-AUC* of 0.7559, and an *Acc* of 83.64%, although interestingly these values remain slightly lower than those obtained for Moderate vs. Severe.

In contrast, the Normal vs. Moderate classification was the most challenging task, consistent with the statistical analysis. The best-performing model (Decision Tree Classifier) achieved an *AUC* of 0.7458 and a *PR-AUC* of 0.724, indicating moderate separation between the two classes. This outcome is not unexpected, as the statistical analysis showed that Normal and Moderate differ only slightly, although some features exhibit a mild trend toward lower-frequency displacement. Nevertheless, these tendencies are insufficient to generate a clear statistical separation, which explains why the classification performance, while still showing some discriminative relevance, remains lower than in the other pairwise comparisons.

Another key finding emerged from the model optimization, the RandomOverSampler was essential for achieving top performance in all three scenarios (100%), highlighting the critical need to address class imbalance in this dataset.

Figure C6 in Supplementary Material C provides a broader view of performance, showing the distribution of *PR-AUC* scores across all classifiers and feature sets. The boxplots reveal three important trends: (1) performance differences among the top classifiers are relatively small,

(2) the outcomes are consistent across models, offering no indication of selective reporting or model-specific artifacts, and (3) classifier choice has a comparatively smaller influence on overall performance than feature selection or data balancing strategies. Collectively, these observations strengthen the robustness and reliability of our findings.

4.4 Comparison with the state-of-the-art

To contextualize our findings, we benchmarked our model against the state-of-the-art, as summarized in Table 1. Our analysis reveals prevalent methodological trends, including a majority of studies employing binary classifications (66.67%) and hold-out validation (75%). Furthermore, common challenges include small sample sizes (33%) and unbalanced datasets (50%). To mitigate these issues, our work directly addresses these limitations through over-sampling techniques.

Among the literature on heart failure assessed from CD with LVEF-based stratification, the studies by Ferreira et al. [17], Brito et al. [15], and Madeiro et al. [18] constitute the most directly comparable benchmarks. Compared with Ferreira et al., our model achieved a 4.64% higher accuracy in the Normal vs. Severe classification task, although it exhibited a 2% lower accuracy in the Normal vs. Moderate task. Relative to Brito et al., our approach yielded slightly lower accuracy and *AUC*; however, it should be noted that in our study we applied a different stratification scheme for the Normal and Moderate classes, considering only LVEF values above 0.4 for this separation. Most notably, our model substantially outperformed the method proposed by Madeiro et al., achieving improvements of 4.25% and 8% in *AUC* for the Normal vs. Severe and Moderate vs. Severe classifications, respectively.

When compared to studies of heart failure from other etiologies that employ similar LVEF-based class structures [20–23], our model demonstrates competitive performance. Although Kwon et al. [20], Vaid et al. [21], and Akerman et al. [23] reported higher *AUC* values—exceeding ours by 5.42% to 9.65% in classifications that included a “Severe” category—our model achieved a slightly lower *AUCs* relative to Choi et al. [22], with 0.92% and 5.15% less in comparable settings. Furthermore, our results offer an important point of reference for models targeting heart

Table 6 Optimal mean classification results for each pairwise comparison across 50 random states, indicating the best-performing model and its associated metrics

Pairwise	OS	Classifier	<i>Acc</i>	<i>Rec</i>	<i>Prec</i>	<i>Spec</i>	<i>F1</i>	<i>NPV</i>	<i>AUC</i>	<i>PR-AUC</i>
Normal vs. Moderate	ROS	DecisionTreeClassifier	77.05	82.50	82.50	66.67	82.50	66.67	0.7458	0.724
Normal vs. Severe	ROS	SVC	83.64	89.74	87.50	68.75	88.61	73.33	0.7925	0.7559
Moderate vs. Severe	ROS	KNeighborsClassifier	83.78	85.71	85.71	81.25	85.71	81.25	0.8348	0.853

Acc = Accuracy; *AUC* = Area Under the Curve; *F1* = *F1*-score; *NPV* = Negative Predictive Value; OS = oversampling; *Prec* = Precision; *Rec* = Recall; *Spec* = Specificity; *PR - AUC* = Precision–Recall *AUC*

failure associated with CD. While our *AUC* is 11.81% below the top benchmark for this etiology, this difference must be interpreted in context: our model was trained and evaluated on a more finely stratified and balanced dataset, providing what is arguably a more robust and reliable benchmark for future research in the field.

Critically, none of the 12 papers benchmarked in Table 1 reported a similar power shift phenomenon found on present study, underscoring the novelty of our findings.

5 Conclusion

We developed and validated a ML model that successfully classifies heart failure severity in CD patients using features from short ECG segments. Our results demonstrated both a statistically significant power spectral “shift” corresponding to disease severity and high classification accuracies (up to 83.78%). The primary clinical value of this approach lies in its potential as a low-cost, non-invasive tool for rapid bedside risk stratification, enabling better triage decisions in resource-limited environments [53].

To ensure model explainability, we deliberately prioritized a classical ML framework with explicit feature extraction over deep learning. This interpretable setup showed that all PSD-derived metrics were consistently modulated by the leftward shift in the power spectrum. This spectral slowing, in turn, aligns with broader morphological changes in the ECG waveform, such as reduced P-wave amplitude, alterations in QRS morphology, or shifts in ST-segment levels, providing a physiologically coherent basis for the model’s predictions.

Moving forward, research should focus on validating these findings on a larger dataset and integrating standard clinical variables to improve both generalizability and clinical acceptance. Another interesting direction would be introducing power-law *PSD*, as it showed to be useful to characterize CD [54] and compared to the results obtained in present study. Ultimately, this work provides a strong foundation for an accessible decision-support tool that can significantly aid in the management of heart failure severity induced by CD.

6 Supplementary information

The online version contains supplementary material available at doi: (available after being accepted - included in submission - appendices.pdf);

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11517-026-03573-5>.

Acknowledgements This work was supported by National Funds from FCT - Fundação para a Ciência e a Tecnologia through project UID/50016/2025. João Paulo Madeiro thanks for the financial support of the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), grants n. 420576/2023-1, 404683/2024-0, 200528/2025-4.

Author Contributions **Pedro Ribeiro:** Conceptualization, Investigation, Writing-original draft, Writing - review and editing; **Maria Inês Barbosa, Roberto C. Pedrosa, and João Paulo Madeiro:** Validation and Writing-review; **João Alexandre Lobo Marques and Pedro Miguel Rodrigues:** Conceptualization, Validation, Investigation, Writing - original draft, Writing - review and editing, Supervision and Funding acquisition.

Funding Open access funding provided by FCT|FCCN (b-on).

Data Availability The datasets generated during and/or analysed during the current study are available from the corresponding author upon reasonable request.

Declarations

Ethics approval and consent to participate The research was approved by the Ethics Committee of HUCFF with the number: 5360915.1.1001.5262. All patients gave their consent prior.

Consent for publication Not applicable.

Competing interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Jannin J, Villa L (2007) An overview of chagas disease treatment. *Mem Inst Oswaldo Cruz* 102(suppl 1):95–98. <https://doi.org/10.1590/s0074-02762007005000106>
- Nunes MCP, Dones W, Morillo CA, Encina JJ, Ribeiro AL (2013) Chagas disease. *J Am Coll Cardiol* 62(9):767–776. <https://doi.org/10.1016/j.jacc.2013.05.046>
- Rassi A, Rassi A, Marin-Neto JA (2010) Chagas disease. *The Lancet* 375(9723):1388–1402. [https://doi.org/10.1016/s0140-6736\(10\)60061-x](https://doi.org/10.1016/s0140-6736(10)60061-x)
- Schofield CJ, Jannin J, Salvatella R (2006) The future of chagas disease control. *Trends Parasitol* 22(12):583–588. <https://doi.org/10.1016/j.pt.2006.09.011>

5. Ribeiro V, Dias N, Paiva T, Hagström-Bex L, Nitz N, Pratesi R, Hecht M (2020) Current trends in the pharmacological management of chagas disease. *Int J Parasitol Drugs Drug Resist* 12:7–17. <https://doi.org/10.1016/j.ijpddr.2019.11.004>
6. Gacek A (2011) An introduction to ECG signal processing and analysis. Springer, London, pp 21–46. https://doi.org/10.1007/978-0-85729-868-3_2
7. Marin-Neto JA, Simões MV, Sarabanda AVL (1999) Chagas' heart disease. *Arq Bras Cardiol* 72(3). <https://doi.org/10.1590/s0066-782x1999000300001>
8. Fradkov AL (2020) Early history of machine learning. *IFAC-PapersOnLine* 53(2):1385–1390. <https://doi.org/10.1016/j.ifacol.2020.12.1888>
9. Badillo S, Banfai B, Birzele F, Davydov II, Hutchinson L, Kam-Thong T et al (2020) An introduction to machine learning. *Clin Pharmacol Therapeutics* 107(4):871–885. <https://doi.org/10.1002/cpt.1796>
10. Yang H, Bukkapatnam ST, Komanduri R (2007) Nonlinear adaptive wavelet analysis of electrocardiogram signals. *Phys Rev E* 76(2):026214. <https://doi.org/10.1103/physreve.76.026214>
11. Behera SK, Bhattacharya D, Aithal N, Sinha N (2024) Non-linear analysis based ECG classification of cardiovascular disorders. <https://doi.org/10.48550/ARXIV.2408.01542>
12. Hayano J, Yuda E (2019) Pitfalls of assessment of autonomic function by heart rate variability. *J Physiol Anthropol* 38(1). <https://doi.org/10.1186/s40101-019-0193-2>
13. Nunes MCP, Beaton A, Acquatella H, Bern C, Bolger AF, Echeverría LE, Dutra WO, Gascon J, Morillo CA, Oliveira-Filho J, Ribeiro ALP, Marin-Neto JA, American Heart Association Rheumatic Fever E, Cardiovascular KDC, Nursing S, Council S (2018) Chagas cardiomyopathy: An update of current clinical knowledge and management: A scientific statement from the American heart association. *Circulation* 138(12):169–209. <https://doi.org/10.1161/CIR.0000000000000599>
14. Alkhodari M, Jelinek HF, Karlas A, Soulaïdopoulos S, Arsenos P, Doundoulakis I et al (2021) Deep learning predicts heart failure with preserved, mid-range, and reduced left ventricular ejection fraction from patient clinical profiles. *Front Cardiovascular Med* 8. <https://doi.org/10.3389/fcvm.2021.755968>
15. Brito BODF, Attia ZI, Martins LNA, Perel P, Nunes MCP et al (2021) Left ventricular systolic dysfunction predicted by artificial intelligence using the electrocardiogram in chagas disease patients—the sami-trop cohort. *PLoS Negl Trop Dis* 15(12):0009974. <https://doi.org/10.1371/journal.pntd.0009974>
16. Silva LEV, Moreira HT, Bernardo MMM, Schmidt A, Romano MMD, Salgado HC et al (2021) Prediction of echocardiographic parameters in chagas disease using heart rate variability and machine learning. *Biomed Signal Process Control* 67:102513. <https://doi.org/10.1016/j.bspc.2021.102513>
17. Ferreira G, Rigo Jr L, C Pedrosa R, Paulo Madeiro J (2023) Prediction of left ventricular ejection fraction using an eeg-based lstm model in chagas disease patients. In: 2023 Computing in Cardiology Conference (CinC). CinC2023. Computing in Cardiology, Atlanta, GA, USA. <https://doi.org/10.22489/cinc.2023.379>
18. Paulo Madeiro J, Rigo Jr L, Coury Pedrosa R (2023) Feature extraction strategies for predicting reduced left ventricular ejection fraction in chagas disease patients. In: 2023 Computing in cardiology conference (CinC). CinC2023. Computing in Cardiology, Atlanta, GA, USA. <https://doi.org/10.22489/cinc.2023.333>
19. Attia ZI, Kapa S, Lopez-Jimenez F, McKie PM, Ladewig DJ, Satam G et al (2019) Screening for cardiac contractile dysfunction using an artificial intelligence-enabled electrocardiogram. *Nat Med* 25(1):70–74. <https://doi.org/10.1038/s41591-018-0240-2>
20. Kwon J-M, Kim K-H, Jeon K-H, Kim HM, Kim MJ, Lim S-M et al (2019) Development and validation of deep-learning algorithm for electrocardiography-based heart failure identification. *Korean Circul J* 49(7):629. <https://doi.org/10.4070/kcj.2018.0446>
21. Vaid A, Johnson KW, Badgeley MA, Somani SS, Bicak M, Landi I et al (2022) Using deep-learning algorithms to simultaneously identify right and left ventricular dysfunction from the electrocardiogram. *JACC Cardiovasc Imaging* 15(3):395–410. <https://doi.org/10.1016/j.jcmg.2021.08.004>
22. Choi J, Lee S, Chang M, Lee Y, Oh GC, Lee H-Y (2022) Deep learning of eeg waveforms for diagnosis of heart failure with a reduced left ventricular ejection fraction. *Sci Rep* 12(1). <https://doi.org/10.1038/s41598-022-18640-8>
23. Akerman AP, Porumb M, Scott CG, Beqiri A, Chartsias A, Ryu AJ et al (2023) Automated echocardiographic detection of heart failure with preserved ejection fraction using artificial intelligence. *JACC Adv* 2(6):100452. <https://doi.org/10.1016/j.jacadv.2023.100452>
24. Decoodt P, Sierra-Sosa D, Anghel L, Cuminetti G, De Keyzer E, Morissens M (2024) Transfer learning video classification of preserved, mid-range, and reduced left ventricular ejection fraction in echocardiography. *Diagnostics* 14(13):1439. <https://doi.org/10.3390/diagnostics14131439>
25. Al Younis SM, Hadjileontiadis LJ, Khandoker AH, Stefanini C, Soulaïdopoulos S, Arsenos P et al (2024) Prediction of heart failure patients with distinct left ventricular ejection fraction levels using circadian eeg features and machine learning. *PLoS ONE* 19(5):0302639. <https://doi.org/10.1371/journal.pone.0302639>
26. Dias JCP, Ramos AN Jr, Gontijo ED, Luquetti A, Shikanai-Yasuda MA et al (2016) 2nd Brazilian consensus on chagas disease, 2015. *Rev Soc Bras Med Trop* 49(suppl 1):3–60. <https://doi.org/10.1590/0037-8682-0505-2016>
27. Kelters IR, Koop Y, Young ME, Daiber A, Laake LW (2025) Circadian rhythms in cardiovascular disease. *Eur Heart J* 46(36):3532–3545. <https://doi.org/10.1093/eurheartj/ehaf367>
28. Boudreau P, Yeh WH, Dumont GA, Boivin DB (2012) A circadian rhythm in heart rate variability contributes to the increased cardiac sympathovagal response to awakening in the morning. *Chronobiol Int* 29(6):757–768. <https://doi.org/10.3109/07420528.2012.674592>
29. Amico AF, Lichtenberg GS, Reisner SA, Stone CK, Schwartz RG, Meltzer RS (1989) Superiority of visual versus computerized echocardiographic estimation of radionuclide left ventricular ejection fraction. *Am Heart J* 118(6):1259–1265. [https://doi.org/10.1016/0002-8703\(89\)90018-5](https://doi.org/10.1016/0002-8703(89)90018-5)
30. Heidenreich PA, Bozkurt B, Aguilar D, Allen LA, Byun JJ, Colvin MM et al (2022) 2022 aha/acc/hfssa guideline for the management of heart failure: A report of the American college of cardiology/American heart association joint committee on clinical practice guidelines. *Circulation* 145(18). <https://doi.org/10.1161/cir.0000000000001063>
31. Behnouth AH, Khalaji A, Naderi N, Ashraf H, Haehling S (2022) Acc/aha/hfssa 2022 and esc 2021 guidelines on heart failure comparison. *ESC Heart Failure* 10(3):1531–1544. <https://doi.org/10.1002/ehf2.14255>
32. Golla MSG, Shams P (2024) Heart failure with preserved ejection fraction (hfpef). In: StatPearls [Internet]. StatPearls Publishing, Treasure island, Florida, US
33. Santulli G, Wang X, Mone P (2022) Updated acc/aha/hfssa 2022 guidelines on heart failure: what is new? from epidemiology to clinical management. *Eur Heart J Cardiovas Pharmacother* 8(5):23–24. <https://doi.org/10.1093/ehjcvp/pvac029>
34. Ádám N, Val'ko D, Balogh Z, Madoš B, Hurtuk J (2025) Comparative evaluation of filtration techniques for eeg signal denoising with emphasis on stationary wavelet transform. *Sci Rep* 15(1). <https://doi.org/10.1038/s41598-025-26476-1>
35. Rahman S, Karmakar C, Natgunanathan I, Yearwood J, Palaniswami M (2022) Robustness of electrocardiogram signal quality indices. *J R Soc Interface* 19(189). <https://doi.org/10.1098/rsif.2022.0012>

36. Wang R, Wang J, Yu H, Wei X, Yang C, Deng B (2014) Power spectral density and coherence analysis of Alzheimer's eeg. *Cogn Neurodyn* 9(3):291–304. <https://doi.org/10.1007/s11571-014-9325-x>
37. Richards MA (2013) The discrete-time fourier transform and discrete fourier transform of windowed stationary white noise. *Tech. Rep*, Georgia Institute of Technology
38. Sundararajan DD, Sundararajan D (2021) Discrete-time fourier transform. *Digital Signal Process An Introduction* 105–153
39. Youngworth RN, Gallagher BB, Stamper BL (2005) An overview of power spectral density (psd) calculations. *Optical Manuf Test VI* 5869:206–216
40. Padeletti M, Bagliani G, De Ponti R, Leonelli FM, Locati ET (2019) Surface electrocardiogram recording. *Cardiac Electrophysiology. Clinics* 11(2):189–201. <https://doi.org/10.1016/j.ccep.2019.01.004>
41. Ribeiro P, Marques JAL, Pordeus D, Zacarias L, Leite CF, Sobreira-Neto MA et al (2024) Machine learning-based cardiac activity non-linear analysis for discriminating covid-19 patients with different degrees of severity. *Biomed Signal Process Control* 87:105558. <https://doi.org/10.1016/j.bspc.2023.105558>
42. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O et al (2011) Scikit-learn: Machine learning in python. *J Mach Learn Res* 12(Oct):2825–2830
43. Guyon I, Elisseeff A (2003) An introduction to variable and feature selection. *J Mach Learn Res* 3(Mar):1157–1182
44. Nahm F (2022) Receiver operating characteristic curve: overview and practical use for clinicians. *Korean J Anesthesiol* 75(1):25–36
45. Doğan O (2019) Data linkage methods for big data management in industry 4.0. In: *Optimizing Big Data Management and Industrial Systems With Intelligent Techniques*. IGI Global, Hershey, pp 108–127. <https://doi.org/10.4018/978-1-5225-5137-9.ch005>
46. Rodrigues PM, Bispo BC, Garrett C, Alves D, Teixeira JP, Freitas D (2021) Lacsogram: A new eeg tool to diagnose Alzheimer's disease. *IEEE J Biomed Health Inform* 25(9):3384–3395. <https://doi.org/10.1109/jbhi.2021.3069789>
47. Rodrigues PM, Freitas DR, Teixeira JP, Alves D, Garrett C (2021) Electroencephalogram signal analysis in Alzheimer's disease early detection. *IGI Global, PA, USA*, pp 224–244. <https://doi.org/10.4018/978-1-7998-3441-0.ch013>
48. Cunha RM, Silva G, Alves M, Bispo BC, Alves D, Garrett C et al (2022) Eeg wavelet packet power spectrum tool for checking Alzheimer's disease progression. *Int J Biomed Eng Technol* 40(3):289. <https://doi.org/10.1504/ijbet.2022.126497>
49. Araújo T, Teixeira JP, Rodrigues PM (2022) Smart-data-driven system for Alzheimer disease detection through electroencephalographic signals. *Bioengineering* 9(4):141. <https://doi.org/10.3390/bioengineering9040141>
50. Manuel AR, Ribeiro P, Silva G, Rodrigues PM, Nunes MVS (2024) Exploring the relationship between caide dementia risk and eeg signal activity in a healthy population. *Brain Sci* 14(11):1120. <https://doi.org/10.3390/brainsci14111120>
51. Jeong J (2004) Eeg dynamics in patients with Alzheimer's disease. *Clin Neurophysiol* 115(7):1490–1505. <https://doi.org/10.1016/j.clinph.2004.01.001>
52. Rodrigues PM, Freitas D, Teixeira JP, Bispo B, Alves D, Garrett C (2018) Electroencephalogram hybrid method for Alzheimer early detection. *Proc Comput Sci* 138:209–214. <https://doi.org/10.1016/j.procs.2018.10.030>
53. Popescu I-M, Margan M-M, Anghel M, Mocanu A, Laitin SMD, Margan R et al (2023) Developing prediction models for covid-19 outcomes: A valuable tool for resource-limited hospitals. *Int J General Med* 16:3053–3065. <https://doi.org/10.2147/ijgm.s419206>
54. Ribeiro ALP, Lombardi F, Sousa MR, Lins Barros MV, Porta A, Costa Val Barros V, Gomes MED, Santana Machado F, Otávio Costa Rocha M (2002) Power-law behavior of heart rate variability in chagas' disease. *Am J Cardiol* 89(4):414–418. [https://doi.org/10.1016/s0002-9149\(01\)02263-9](https://doi.org/10.1016/s0002-9149(01)02263-9)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Pedro Ribeiro earned his B.Sc. in Bioengineering and M.Sc. in Biomedical Engineering from Universidade Católica Portuguesa, where he is now a Ph.D. student. His research focuses on developing an AI-based prediction tool to assess COVID-19 severity and other cardiovascular diseases using electrocardiographic signals.

João Alexandre Lobo Marques is the Vice-Rector and Full Professor at the University of Saint Joseph in Macau, China, and holds a PhD in Teleinformatics Engineering. His research develops systems for analyzing physiological signals from the brain and heart and their educational applications.

Maria Inês Barbosa holds a Ph.D. in Biomedical Engineering and is an Invited Assistant Professor at the Faculty of Biotechnology - Universidade Católica Portuguesa. Her research focuses on data processing and artificial intelligence, with an emphasis on neurodegenerative diseases and pattern recognition in biomedical signals.

Roberto C. Pedrosa holds a PhD in Medicine (Cardiology) from the Federal University of Rio de Janeiro and is a Permanent Professor in the Graduate Program at ICES-UFRJ. He has experience in the field of Medicine, with an emphasis on Cardiology, working mainly in Chagas Heart Disease and Dysautonomia.

João Paulo do Vale Madeiro is an Assistant Professor in the Department of Computing of the Federal University of Ceará. He holds a Ph.D. in Teleinformatics Engineering. His expertise includes signal and image processing, medical diagnosis support systems, and automatic feature extraction from electrocardiographic signals.

Pedro Miguel Rodrigues holds a Ph.D. in Biomedical Engineering and is the Vice-Dean and Assistant Professor at the Faculty of Biotechnology - Universidade Católica Portuguesa. His research areas focus on digital signal & image processing and artificial intelligence applied to diagnosing neurodegenerative, cardiac, and speech diseases.