

Article

Cardiovascular Diseases Diagnosis Using an ECG Multi-Band Non-Linear Machine Learning Framework Analysis

Pedro Ribeiro , Joana Sá, Daniela Paiva  and Pedro Miguel Rodrigues * 

CBQF—Centro de Biotecnologia e Química Fina, Laboratório Associado, Escola Superior de Biotecnologia, Universidade Católica Portuguesa, Rua de Diogo Botelho 1327, 4169-005 Porto, Portugal; s-pmsbribeiro@ucp.pt (P.R.); joanasa8@gmail.com (J.S.); danielaseabrapaiva@outlook.com (D.P.)

* Correspondence: pmrodrigues@ucp.pt

Abstract: Background: cardiovascular diseases (CVDs), which encompass heart and blood vessel issues, stand as the leading cause of global mortality for many people. Methods: the present study intends to perform discrimination between seven well-known CVDs (bundle branch block, cardiomyopathy, myocarditis, myocardial hypertrophy, myocardial infarction, valvular heart disease, and dysrhythmia) and one healthy control group, respectively, by feeding a set of machine learning (ML) models with 10 non-linear features extracted every 1 s from electrocardiography (ECG) lead signals of a well-known ECG database (PTB diagnostic ECG database) using multi-band analysis performed by discrete wavelet transform (DWT). The ML models were trained and tested using a leave-one-out cross-validation approach, assessing the individual and combined capabilities of features, per each lead or combined, to distinguish between pairs of study groups and for conducting a comprehensive all vs. all analysis. Results: the *Accuracy* discrimination results ranged between 73% and 100%, the *Recall* between 68% and 100%, and the *AUC* between 0.42 and 1. Conclusions: the results suggest that our method is a good tool for distinguishing CVDs, offering significant advantages over other studies that used the same dataset, including a multi-class comparison group (all vs. all), a wider range of binary comparisons, and the use of classical non-linear analysis under ECG multi-band analysis performed by DWT.



Citation: Ribeiro, P.; Sá, J.; Paiva, D.; Rodrigues, P.M. Cardiovascular Diseases Diagnosis Using an ECG Multi-Band Non-Linear Machine Learning Framework Analysis. *Bioengineering* **2024**, *11*, 58. <https://doi.org/10.3390/bioengineering11010058>

Academic Editors: Dante Mantini and Mark L. Weiss

Received: 7 November 2023

Revised: 13 December 2023

Accepted: 5 January 2024

Published: 7 January 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: ECG signals; cardiovascular diseases; machine learning models; discrete wavelet transform; non-linear analysis; discrimination

1. Introduction

Heart and blood vessel problems, known as cardiovascular diseases (CVDs), are the main reason why many people die around the world [1]. According to the World Health Organization, 32% of global mortality is attributed to cardiovascular diseases, with the most prevalent being arrhythmias, cardiac arrests, and heart failure. It is estimated that CVDs take about 17.9 million lives every year [2]. Focusing on cardiac pathology and considering how much work the heart constantly does, it is amazing that it functions so well for a long time for many people. However, it can also experience problems and stop working properly due to risk factors like cholesterol, high blood pressure, cigarette smoking, diabetes mellitus, and adiposity [3].

Heart disease is a term for health issues that affect the heart's function and condition. There are different types of heart disease, including: (1) Cardiomyopathy: heart muscle structural and functional abnormality without underlying coronary issues [4]; (2) Endocarditis: infection and inflammation of heart valves and inner lining [5]; (3) Myocarditis: inflammation of middle heart wall layer affecting blood pumping [6]; (4) Pericarditis: inflammation of the thin sac surrounding the heart [7]; (5) Coronary artery disease: cholesterol-filled plaques blocking heart arteries [8]; (6) Heart attack: sudden blockage of blood flow to heart muscle [9]; (7) Heart failure: symptoms include breathlessness,

ankle swelling, and fatigue [10]; (8) Heart rhythm disorders (arrhythmias): irregular heartbeats [11]; (9) Sudden cardiac arrest: sudden stoppage of heartbeat [12]; (10) Heart valve disorders: issues with valves controlling blood flow [13]; (11) Congenital heart disease: heart abnormalities present from birth [14].

The beginning of the diagnosis of heart disease involves evaluating the patient's medical history and conducting a physical examination. Afterwards, laboratory tests and/or additional non-invasive and invasive diagnostic exams can be performed [2]. Natriuretic peptides are the most common laboratory tests used to diagnose heart diseases. They can help identify individuals at higher risk of sudden cardiac death in the general population or patients with coronary artery disease [11]. However, several other non-invasive and invasive tests can be performed: (1) Electrocardiogram (ECG) and ambulatory monitoring: the 12-lead ECG is a key diagnostic test for cardiovascular diseases, assessing risk, and identifying arrhythmias [11]. Choose monitoring time based on symptom frequency. Holter for daily arrhythmias, patient-activated ECG for less frequent events, and ILRs for serious cases [11,15]; (2) Stress tests: monitor the heart during treadmill/bike exercise to assess response and detect exercise-related disorders like arrhythmias, ventricular tachycardia, coronary artery disease, and long QT syndrome [11,16]. Exercise tests aid in diagnosing long QT syndrome by measuring the QTc interval after 4 min of exercise [16]; (3) Imaging tests: essential for assessing heart function and detecting problems like cardiomyopathies [17]. Negative results may indicate primary electrical diseases [4]; (4) Electrophysiological study: exam to diagnose and guide treatment, involving measuring cardiac intervals, controlling electrical stimulation, and mapping heart structures. Effectiveness varies based on heart condition, presence of spontaneous ventricular tachycardia, medication use, and stimulation mode [18]; (5) Provocative diagnostic tests: use sodium channel blockers, adenosine, or epinephrine to detect syndromes. Use acetylcholine or ergonovine to assess coronary spasm as the cause of ventricular fibrillation [19]; (6) Genetic testing: next-gen sequencing made genetic testing accessible. Comprehensive gene panels reveal variations causing or modifying features in syndromes like Brugada, long QT, and hypertrophic and dilated cardiomyopathy [20]; (7) Cardiac catheterisation: catheter inserted into a blood vessel, and guided to the heart with X-ray images and dye to check for blockages [11].

In recent years, there has been a notable surge in computational power, driven by advanced hardware, parallel computing, cloud resources, and increased data accessibility. These developments have significantly enhanced the applications of machine learning (ML) in the diagnosis of CVDs [21]. The role of ML in CVDs is pivotal, as it harnesses data from medical tests to improve diagnostics and management, reducing human error, improving efficiency, and enhancing patient outcomes [22]. It contributes to early disease detection, precise risk assessment, advanced image analysis, predictive modelling, tailored treatment plans, remote patient monitoring, and expedited drug discovery [23]. However, ongoing research is essential to enhance this critical field further and save lives. Thus, for this study, our method will focus on ML-based ECG signal analysis approaches for discriminating CVDs, and thus we present in Table 1 the state of the art of this topic. The heart, operating as a non-linear system, manifests its electrical activity through the ECG signal [24]. The inherent non-linearity underscores the inadequacy of traditional linear analyses and standard clinical features in comprehensively capturing the intricate dynamics of the ECG signal. This complexity is further underscored by the challenges posed to deep learning tools, as their extraction of features may lack explainable understanding. Consequently, a deeper comprehension of how these tools reach and compute features becomes imperative for a more robust interpretation of ECG signals. Unlike prevailing state-of-the-art methods for the topic (Table 1), which have typically abstained from incorporating non-linear feature extraction in their methodology, our study aims to explore a non-linear approach to ECG analysis supported by classical ML tools. By doing so, we want to seek a more comprehensive understanding that embraces the inherent complexities of the heart's electrical behaviour for improving CVDs diagnosis. For that, we defined three objectives for this study:

Table 1. State-of-the-art literature report on CVDs detection with information about the database, the comparison groups, the features extracted, used classifiers, limitations, and *Accuracy*.

Ref	Year	Database	Comparison Group (Number of Participants)	Feature Extracted	Classifier	Limitations	Validation	Accuracy
[25]	2004	University Hospital in Lund database	Normal (1119) vs. Myocardial infarction (1119)	Hermite decomposition	ANN	Exclusive assessment of myocardial infarction. Lack of diversity of CVDs.	Cross-validation	94%
[26]	2016	PTB diagnostic ECG database	Normal (49) vs. Cardiomyopathy (14)	ECG PR, QT, RR and QRS intervals	Feed-forward back-propagation Neural Network	Small and unbalanced dataset. Exclusive assessment of cardiomyopathy. Lack of diversity of CVDs.	Cross-validation	95.2%
[27]	2018	PTB diagnostic ECG database	Normal (25) vs. Myocardial infarction (36)	Feature extracted from DNN	DNN (InceptionV3)	Small database for hold-on. Lack of diversity of CVDs.	Hold-on	99.64%
[28]	2018	PTB diagnostic ECG datasets	Normal (52) vs. Myocardial infarction (148)	Features extracted from CNN	CNN	Small and unbalanced database for hold-on. It is impossible to know what features were extracted due to the nature of deep learning algorithms. Lack of diversity of CVDs.	Hold-on	95.9%
[29]	2019	PTB diagnostic ECG database	Healthy (52) vs. Myocardial infarction (148)	Feature extracted from CNN	CNN	Small and unbalanced dataset. There was no discrimination of diseases outside of the MI class. Lack of diversity of CVDs.	Cross-validation	99.78%
[30]	2019	Cleveland heart disease database	Healthy (150) vs. Heart Disease (147)	Patient clinical information	SVM	Small dataset. There was no discrimination of diseases outside of the Heart Disease class. Lack of diversity of CVDs.	Hold-on	92.22%
[31]	2019	Cleveland heart disease database	Healthy (150) vs. Heart Disease (147)	Patient clinical information	DNN	Small dataset. There was no discrimination of diseases outside of the Heart Disease class. Lack of diversity of CVDs.	Hold-on	93.33%
[32]	2020	Heart sound dataset	Normal (400) vs. Mitral valve prolapse (400) vs. Mitral stenosis (400) vs. Mitral regurgitation (400) vs. Aortic stenosis (400)	Feature extracted from CNN	CNN	Low variety of classes. There is a high risk of over-fitting because of the augmentation technique used.	Cross-validation	98.6%
[33]	2020	PTB diagnostic ECG datasets	Normal (313) vs. Abnormal (318)	Features extracted from ANN	Ensemble	Use class weights when training with artificial neural networks to solve the class unbalance problem.	Hold-on	94.14%
[34]	2020	PTB diagnostic ECG database	No Myocardial infarction (141) vs. Myocardial infarction (148)	Feature extracted from CNN	CNN	Small dataset. Just MI different types of discrimination. Small database.	Cross-validation	81%
[35]	2021	Ch. Pervaiz Elahi Institute of Cardiology Multan Dataset	Normal (3408) vs. Abnormal (2796) vs. Myocardial infarction (2880) vs. Previous history of Myocardial infarction (2064)	Feature extracted from SSD MobileNetV2 (CNN)	SSD MobileNetV2 (CNN)	Just MI different types of discrimination. It is impossible to know what features were extracted due to the nature of deep learning algorithms.	Hold-on	98.33%
[36]	2021	PTB diagnostic ECG database	Healthy (52) vs. Abnormal (216)	Domain features and disease-specific features	XGBoost	Small and unbalanced dataset. There is no disease discrimination, just normal vs. abnormal. Lack of diversity of CVDs.	Cross-validation	98.23%
[37]	2021	Cleveland dataset	Normal (170) vs. Abnormal (140)	Patient clinical information	Decision tree and Random forest combined	Small database for hold-on. There is no disease discrimination, just normal vs. abnormal. Lack of diversity of CVDs.	Hold-on	88.00%
[38]	2021	Cleveland HD dataset	Healthy (135) vs. Heart disease (135)	Feature extracted from MAFW	CNN model with the MAFW	Small dataset. A limited number of classifiers were used. High computational cost and time complexity. Runtime is not considered as an evaluation criterion. Lack of diversity of CVDs.	Cross-validation	90.1%

Table 1. Cont.

Ref	Year	Database	Comparison Group (Number of Participants)	Feature Extracted	Classifier	Limitations	Validation	Accuracy
[39]	2022	Cleveland, Hungary, Switzerland, and Long Beach V datasets	Healthy (500) vs. Heart disease (550)	Patient clinical information	Extreme gradient boosting	No disease discrimination, just normal vs. abnormal. Lack of diversity of CVDs.	Cross-validation	100%
[40]	2022	UC Irvine Machine Learning Repository CVD datasets	Healthy (500) vs. Heart disease (550)	Patient clinical information	CNN and BiLSTM hybrid	No disease discrimination, just normal vs. abnormal. Lack of diversity of CVDs.	Hold-on	94.51%
[41]	2022	ECG dataset of Cardiac and COVID-19 Patients and ECG dataset of Cardiac Patients	Normal (284) vs. Abnormal (233) vs. MI (239) vs. Previous history of MI (102)	Feature extracted from MobileNet V2 (CNN)	MobileNet V2 (CNN)	Small database for hold-on, lack of a truly independent test group. Did not consider optimisation techniques.	Hold-on	95.18%
[42]	2022	PTB-XL dataset	Normal (1608) vs. Abnormal (1357)	Feature extracted from DNN	XGBoost	Lack of diversity of CVDs. Unbalanced dataset.	Hold-on	78.65%
Present Work	2023	PTB diagnostic ECG database	VHD vs. M; VHD vs. MI; VHD vs. MH; VHD vs. HC; VHD vs. Dis; VHD vs. CardMyo; VHD vs. BBB; M vs. MI; M vs. MH; M vs. HC; M vs. Dis; M vs. CardMyo; M vs. BBB; MI vs. MH; MI vs. HC; MI vs. Dis; MI vs. CardMyo; MI vs. BBB; MH vs. HC; MH vs. Dis; MH vs. CardMyo; MH vs. BBB; HC vs. Dis; HC vs. CardMyo and HC vs. BBB	Approximate Entropy, Logarithmic Entropy, Shannon Entropy, Correlation Dimension, Detrended Fluctuation Analysis, Energy, Higuchi Fractal Dimension, Hurst Exponent, Katz Fractal Dimension and Lyapunov Exponent	19 ML Classifiers	Small data sample for some classes and unbalanced dataset.	Cross-validation	73–100%

- Introduce the utilisation of 10 non-linear features (entropies—approximate, logarithmic, and Shannon, correlation dimension, detrended fluctuation analysis, energy, Higuchi fractal dimension, Hurst exponent, Katz fractal dimension, and Lyapunov exponent) extracted under discrete wavelet transform multi-band ECG signal analysis for characterising CVDs.
- Enhance the evaluation of distinguishing between various CVDs by accessing and comparing the individual and combined power of non-linear features.
- Evaluate the discriminatory performance of these features by inputting them into a comprehensive set of ML models.

The fulfilment of the goals will provide insights into the predictive power of these non-linear features, both independently and synergistically, contributing to a comprehensive understanding of their impact on CVD discrimination.

Finally, the paper is divided into five major sections in terms of structure. In Section 2, the applied methodology, including the database, signal processing, and feature extraction, is explained. The study results are indicated in Section 3 and discussed in Section 4. Finally, Section 5 draws the study conclusions.

2. Methodology

This proposed methodology, illustrated in Figure 1, is split into 4 main parts:

- Data collection/pre-processing and artifacts removal;
- Feature extraction;
- Data compressor;
- Machine Learning classification and statistical analysis.

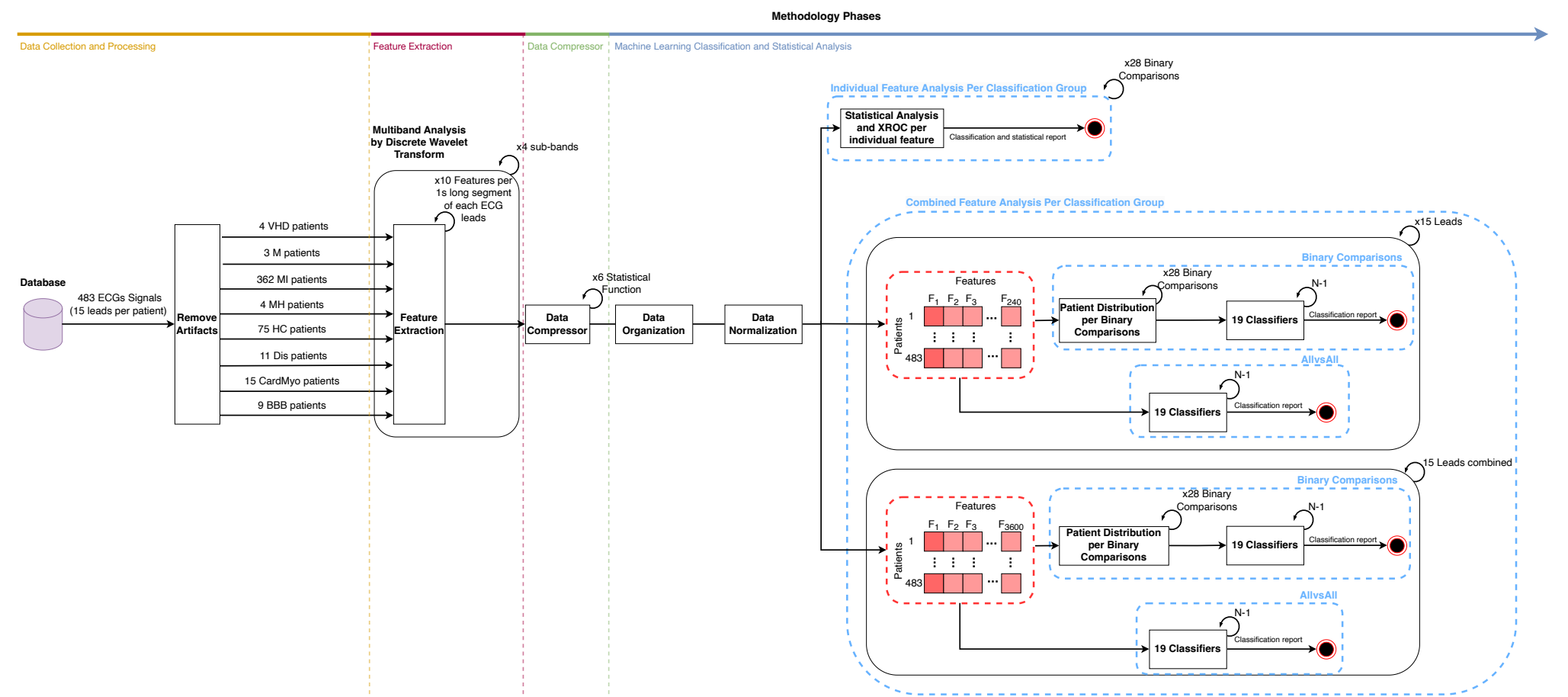


Figure 1. Workflow diagram.

2.1. Experimental Setup

This study involved the use of two distinct programming languages: MATLAB and Python. MATLAB (version R2022a) was employed to eliminate noise from the ECG signal, extract non-linear features from the ECG data, and compress and structure the data for classification purposes. Python (version 3.9.12) was utilised to develop and implement various ML models and generate a discrimination report based on the obtained results. The choice between MATLAB and Python programming languages is driven by optimisation needs: MATLAB is particularly proficient in signal processing and feature extraction with highly optimised toolboxes, whereas Python takes the lead in optimising ML models.

This research was conducted using a MacBook Pro 14 equipped with an M1 Pro chip featuring an 8-core CPU, a 14-core GPU, and 16 GB of RAM.

2.2. Database Characterisation

The PTB diagnostic ECG database [43] comprises data from seven distinct cardiovascular disease groups as well as a healthy control group. This dataset consists of a total of 512 ECG records, each containing 12 conventional leads (I, II, III, α Vr, α VL, α Vf, V1, V2, V3, V4, V5, and V6), along with 3 ECG Frank leads (Vx, Vy, and Vz). The ECG data have been digitised at a sampling frequency of 1000 Hz.

Each lead contains ECG signal samples of 10 s and has been recorded by an electroencephalograph with: (1) input voltage: ± 16 mV; (2) input resistance: $100\ \Omega$ (DC); (3) bandwidth: 0–1 kHz; and (4) noise voltage: 10 μ V.

Table 2 represents the number of ECGs per diagnostic class present in the database.

Table 2. Number of ECGs per diagnosis class.

Diagnostic Class	Number of ECGs
Bundle branch block (<i>BBB</i>)	17
Cardiomyopathy (<i>CardMyo</i>)	20
Healthy controls (<i>HC</i>)	80
Myocarditis (<i>M</i>)	4
Myocardial hypertrophy (<i>MH</i>)	4
Myocardial infarction (<i>MI</i>)	367
Valvular heart disease (<i>VHD</i>)	6
Dysrhythmia (<i>Dis</i>)	16

2.3. Artifacts Removal

The ECG signals' raw data in the database showed artifacts. To ensure the signal quality, complete signal deletion was performed. In the beginning, the database had 512 records. After the removal stage, the number of available signals in the database for the following tasks was reduced to 483 ECG records. Table 3 represents the number of ECGs per diagnostic class after the removal.

Table 3. Number of ECGs per diagnosis class after signal quality assessment and artifacts removal.

Diagnostic Class	Number of ECGs
Bundle branch block	9
Cardiomyopathy	15
Healthy controls	75
Myocarditis	3
Myocardial hypertrophy	4
Myocardial infarction	362
Valvular heart disease	4
Dysrhythmia	11

2.4. Signal Normalisation

The ECG signals, $x(n)$, were loaded into MATLAB® and normalised according to the following equation [44].

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

where N represents the signal's length. Then its mean value was removed.

Multi-Band Decomposition via Wavelet Transform and Features Extraction

The discrete-time wavelet transform (DWT) is a powerful technique used to analyse discrete-time signals with finite energy. It involves breaking down the signal into a set of basis functions composed of a limited number of prototype sequences and their time-shifted variations. This process, as described in Guido's research in 2022 [45], offers significant advantages for analysing signals in the time–frequency domain. By seamlessly transitioning between the time and frequency domains, it enables the localisation of the source of frequency compounds in time.

To perform the decomposition and subsequent reconstruction, an octave-band critically decimated filter bank is employed. This approach, pioneered by Malvar in 1992 and further developed by Vetterli in 1995 [46,47], provides an effective framework. When considering only the positive frequencies, each sub-band in the transform is confined to a specific range,

$$W_k = \begin{cases} [0, \pi/2^S], & m = 0, \\ [\pi/2^{S-m+1}, \pi/2^{S-m}], & m = 1, 2, \dots, S, \end{cases} \quad (2)$$

where S is the number of levels, $S + 1$ is the number of sub-bands, and π is the normalised angular frequency equivalent to half the sampling rate.

The DWT employs an analysis scale function, denoted as $\tilde{\phi}_1(n)$, and an analysis wavelet function, denoted as $\tilde{\psi}_1(n)$, which are defined as follows:

$$\tilde{\phi}_1(n) = h_{LP}(n) \quad (3)$$

and

$$\tilde{\psi}_1(n) = h_{HP}(n), \quad (4)$$

where $h_{LP}(n)$ and $h_{HP}(n)$ represent the impulse responses of the analysis filters for the half-band low-pass and high-pass components, respectively.

Defining the following recursion formulas

$$\tilde{\phi}_{i+1}(n) = \tilde{\phi}_i(n/2) * \tilde{\phi}_1(n), \quad (5)$$

$$\tilde{\psi}_{i+1}(n) = \tilde{\phi}_i(n) * \tilde{\psi}_1(n/2^i), \quad (6)$$

where the symbol “*” signifies the convolution operation, the analysis filter corresponding to the m th sub-band is expressed as follows:

$$h_m(n) = \begin{cases} \tilde{\phi}_S(n), & m = 0, \\ \tilde{\psi}_{S+1-m}(n), & m = 1, 2, \dots, S. \end{cases} \quad (7)$$

The m th sub-band signal is computed as

$$x_m(n) = \begin{cases} \sum_{k=-\infty}^{\infty} x(k)h_m(2^S n - k), & m = 0, \\ \sum_{k=-\infty}^{\infty} x(k)h_m(2^{S-m+1} n - k), & m = 1, 2, \dots, S. \end{cases} \quad (8)$$

In this research, the DWT was employed to decompose each ECG segment of 1 s length into sub-bands ($x_m(n)$) up to level three ($S = 3$). The applied wavelet was Symlet7, and this wavelet proved to be good for ECG signals analysis until decomposition at level 3 [48,49]. To ensure consistency with the original sampling rate, the sub-band signals, $x_m(n)$, underwent re-sampling using the wavelet interpolation method [50]. After that, 10 non-linear features (check Table 4 for more information) were collected from each signal sub-band of 1 s length from a total of 10 s signal length. Then, the resulting time series per feature and sub-band were compressed over time, respectively, by 6 distinct statistical functions: average (*Avg*), standard deviation (*Std*), 95th percentile (*P95*), variance (*Var*), median (*Med*), and kurtosis (*Kur*) [49]. At the end of the process, the data matrix, comprised of all 10-second time series vectors of features extracted from all sub-bands over time for all patients, underwent normalisation using the z-score method [51].

Table 4. The extracted features with the corresponding equations and definitions.

Feature	Equation	Definition
Approximate Entropy (<i>ApEn</i>)	$ApEn(m, r) = \lim_{N \rightarrow \infty} \Theta^m(r) - \Theta^{m+1}(r),$ <p>Θ is the Heaviside step function and m is the dimension [52].</p>	<i>ApEn</i> evaluates the likelihood that similar patterns within the data will remain similar when additional data points are included. The lower the <i>ApEn</i> value is, the more regular or predictable the data are, whereas a higher <i>ApEn</i> value suggests greater complexity or irregularity.
Correlation Dimension (<i>CorrDim</i>)	$CorrDim = \lim_{M \rightarrow \infty} \frac{2 \sum_{i=1}^{M-k} \sum_{j=i+k}^M \Theta(l X_i - X_j)}{M^2},$ <p>where $\Theta(x)$ is the Heaviside step function, X_i and X_j are the position vectors on attractor, l is the distance under consideration, k is the summation offset, and M is the reconstructed vector numbers from the $x(n)$ [52].</p>	<i>CorrDim</i> is used to measure self-similarity, and higher values of <i>CorrDim</i> means a high degree of complexity and less similarity.
Detrended Fluctuation Analysis (<i>DFA</i>)	$DFA(n) = \sqrt{\frac{\sum_{k=1}^N [y(k) - y_n(k)]^2}{N}},$ <p>where N is the length, $y_n(k)$ is the local trend, and $y(k)$ is defined as</p> $y(k) = \sum_{i=1}^k [x(i) - \bar{x}],$ <p>with $x(i)$ as the inter-beat interval and \bar{x} as its average [53].</p>	<i>DFA</i> is a technique for measuring the power scaling observed through R/S analysis.
Energy (<i>En</i>)	$En = \sum_{n=0}^{N-1} x(n) ^2$	<i>En</i> is the capacity of a system to perform work [54].
Higuchi Fractal Dimension (<i>H</i>)	$H = \frac{\ln(L(k))}{\ln(\frac{1}{k})},$ <p>where k is a number of composed sub-series and $L(k)$ is the averaged curve size.</p>	<i>H</i> estimates the fractal dimension of a time series signal [55].

Table 4. Cont.

Feature	Equation	Definition
Hurst Exponent (EH)	$K_q(\tau) \sim \left(\frac{\tau}{\nu}\right)^{qEH(q)},$ <p>with</p> $K_q(\tau) = \frac{(X(t+\tau) - X(t))^q}{(X(t))^q},$ <p>where q is the order moments of the distribution increments, ν is the time resolution, τ is the incorporation time delay of the attractor, and t is the period of a given time series signal $X(t)$ [56].</p>	EH quantifies how chaotic or unpredictable a time series is.
Katz Fractal Dimension (K)	$K = \frac{\log(n)}{\log(n) + \log\left(\frac{\max_n(\sqrt{(n-1)^2 + (x(n)-x(1))^2})}{\sum_{n=2}^N \sqrt{1+(x(n-1)-x(n))^2}}\right)},$	K estimates the fractal dimensions through a waveform analysis of a time series [56].
Logarithmic Entropy (LogEn)	$LogEn = \sum_{n=1}^N \log_2[x(n) ^2]$	LogEn quantifies the average amount of information (in bits) needed to represent each event in the probability distribution. Higher logarithm entropy values indicate greater unpredictability or randomness in the distribution, while lower values suggest more certainty or order [54].
Lyapunov Exponent (ELay)	$ELay(x_0) = \lim_{n \rightarrow \infty} \frac{\sum_{k=1}^n \ln f'(x_k - 1) }{n},$ <p>where f' is the f derivative [57].</p>	ELay evaluates the system's predictability and sensitivity to change.
Shannon Entropy (ShaEn)	$ShaEn = - \sum_{n=1}^N x(n) ^2 \log_2[x(n) ^2]$	ShaEn is measured in bits when the base-2 logarithm (\log_2) is used. This means that the result provides a quantification of the average number of bits required to represent each outcome in a given probability distribution. Higher entropy values indicate greater uncertainty, unpredictability, or randomness in the distribution, while lower values suggest more order or certainty [54].

2.5. Data Driven Framework Analysis

2.5.1. Individual Feature Power Analysis over Binary Groups

The evaluation of the discriminating power of each feature distribution between pairs of study groups, such as VHD vs. M, VHD vs. MI, VHD vs. MH, VHD vs. HC, VHD vs. Dis, VHD vs. CardMyo, VHD vs. BBB, M vs. MI, M vs. MH, M vs. HC, M vs. Dis, M vs. CardMyo, M vs. BBB, MI vs. MH, MI vs. HC, MI vs. Dis, MI vs. CardMyo, MI vs. BBB, MH vs. HC, MH vs. Dis, MH vs. CardMyo, MH vs. BBB, HC vs. Dis, HC vs. CardMyo, and HC vs. BBB, was conducted using the XROC classifier [58], a binary classifier working within a leave-one-out cross-validation process, and using the Mann–Whitney test. A total of 3600 features, consisting of 10 non-linear features time series compressed over time by (\times) 6 statistical measures over (\times) 4 sub-bands for each one of the (\times) 15 leads per participant, were individually assessed to measure their potential to differentiate between these groups. The methodology variation to perform individual feature assessment for discrimination is signalled in Figure 1. It should be noted that the normality and homoscedasticity of each one of the time series feature vector distributions have been assessed for distinguishing binary classes with the MATLAB function *ktest*, which performs

the Kolmogorov–Smirnov and Levene tests, respectively. The hypothesis of parametric tests was not met, so we applied a non-parametric test, such as the Mann–Whitney test.

2.5.2. Combined Features Power Analysis for Groups Discrimination Using Sci-Learn ML Models

In this case, the model’s performance for discriminating between pairs of study groups and between *All* vs. *All* was evaluated by feeding 19 selected Sci-learn ML models [59], presented in Table 5, with combined features—240 features (10 features extracted from (×) 4 sub-bands and compressed (×) by 6 statistics) for the individual lead case or 3600 features (240 features per lead × 15 leads) per combined leads case, for each group comparison, within a leave-one-out cross-validation procedure. The methodology variation to perform combined feature assessment for discrimination is signalled in Figure 1.

Table 5. 19 Sci-learn ML classifiers configurations.

Classifier	Hyperparameters
AdaBoostClassifier (AdaBoost)	Default parameters
BaggingClassifier (BaggC)	Default parameters
DecisionTreeClassifier (DeTreeC)	max_depth: 5
ExtraTreesClassifier (ExTreeC)	n_estimators: 300
GaussianNB (GauNB)	Default parameters
GaussianProcessClassifier (GauPro)	1.0 × RBF(1.0)
GradientBoostingClassifier (GradBoost)	Default parameters
KNearestNeighborsClassifier (KNN)	Default parameters
LinearDiscriminantAnalysis (LinDis)	Default parameters
LinearSVC (LinSVC)	Default parameters
LogisticRegression (LogReg)	solver: “lbfgs”
LogisticRegressionCV (LogRegCV)	cv: 3
MLPClassifier (MLP)	α: 1, max_iter: 1000
OneVsRestClassifier (OvsR)	random_state: 0
RandomForestClassifier (RF)	max_depth: 5, n_estimators: 300, max_features: 1
SGDClassifier (SGD)	max_iter: 100, tol: 0.001
SGDClassifierMod (SGDCMod)	Default parameters
Support-vector Machines (SVC)	γ: “auto”

2.5.3. Classification Metrics

The model’s performance evaluation was carried out using 9 metrics: *Accuracy*, *Precision*, *Recall*, *F1-Score*, *AUC*, *Kappa*, *MCC*, *CSI*, and *Gmean*.

The *Accuracy* represents the number of corrected classified classes concerning all cases [60] and can be defined as

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100\%, \quad (9)$$

where, a *TP*, *TN*, *FP*, and *FN* are, respectively, the true positives, true negatives, false positives, and false negatives [61].

The *Precision*, also known as a positive predictive value, shows the proportion of well-classified positive cases to the total cases predicted as positive [62]. The *Precision* can be defined as

$$Precision = \frac{TP}{TP + FP} \times 100\%. \quad (10)$$

The *Recall*, defined as

$$Recall = \frac{TP}{TP + FN} \times 100\%, \quad (11)$$

represents the proportion of correctly predicted positive cases concerning the total number of positive cases [62].

The *F1-Score* is the harmonic average between the *Recall* and the *Precision* [63], and the equation is defined as

$$F1\text{-score} = \frac{2 \times Precision \times Recall}{Precision + Recall} \times 100\%. \quad (12)$$

The *Kappa* normalises the *Accuracy* by the possibility of agreement by chance [64] and is defined as

$$Kappa = \frac{2 \times (TP \times TN - FN \times FP)}{(TP + FP) \times (FP + TN) + (TP + FN) \times (FN + TN)}. \quad (13)$$

The *MCC* is useful for uneven data [65]. It varies between 0 and 1, with 0 as the worst scenario and 1 as the best. It is defined as

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP) \times (TP + FN) \times (TN + FP) \times (TN + FN)}}. \quad (14)$$

The *CSI* provides a more nuanced evaluation of a binary classification model's effectiveness by considering both the correct identification of positive instances and the ability to avoid false positives. [66]. The *CSI* equation can be defined as

$$CSI = \frac{TP}{TP + FP + FN}. \quad (15)$$

The *Gmean* is the measure that considers a balance between the performance of all classes. The higher the value is, the lower is the risk of models over-fitting. It is defined as

$$Gmean = \sqrt{Recall \times Specificity}, \quad (16)$$

where *Specificity* is defined as

$$Specificity = \frac{TN}{FN + TN}. \quad (17)$$

The area under the curve (*AUC*) of the receiver operating characteristic curve (*ROC*) is a metric that evaluates how well a model can distinguish between positive and negative classes. It achieves this by comparing the rate of *TP* against the rate of *FP* at different classification thresholds. The value of *AUC* ranges between 0 and 1, with the perfect classifier resulting in a value of 1, while a random classifier has an *AUC* of 0.5. Using *AUC* allows for a single-value measure of a model's performance. This is especially useful for comparing models and assessing performance in scenarios where there is an imbalance between classes [67].

3. Results

Table 6 displays the individual features' discrimination power that yielded the best results for statistical and XROC analysis conducted across all 28 binary comparisons (VHD vs. M, VHD vs. MI, VHD vs. MH, VHD vs. HC, VHD vs. Dis, VHD vs. CardMyo, VHD vs. BBB, M vs. MI, M vs. MH, M vs. HC, M vs. Dis, M vs. CardMyo, M vs. BBB, MI vs. MH, MI vs. HC, MI vs. Dis, MI vs. CardMyo, MI vs. BBB, MH vs. HC, MH vs. Dis, MH vs. CardMyo, MH vs. BBB, HC vs. Dis, HC vs. CardMyo, HC vs. BBB, Dis vs. CardMyo, Dis vs. BBB, and CardMyo vs. BBB), respectively.

Table 6. Statistical and XROC results for individual feature power analysis per binary groups, where N.S. means no significance.

Comparison Group	Feature	Compressor	m th Sub-Band	Lead	p -Value	Recall	Accuracy
VHD vs. M	CorrDim	Avg	2	V4	0.0339	100%	100%
VHD vs. MI	CorrDim	Avg	2	I	0.0486	0%	98.91%
VHD vs. MH	CorrDim	Avg	2	V1	N.S.	75.00%	87.50%
VHD vs. HC	CorrDim	Avg	2	Vy	0.0301	0%	94.94%
VHD vs. Dis	CorrDim	Avg	2	V5	N.S.	100%	80.00%
VHD vs. $CardMyo$	CorrDim	Avg	2	Vy	N.S.	50.00%	78.94%
VHD vs. BBB	CorrDim	Avg	2	V1	0.0308	75.00%	84.62%
M vs. MI	CorrDim	Avg	2	V4	N.S.	0%	99.18%
M vs. MH	CorrDim	Avg	2	III	N.S.	100%	85.71%
M vs. HC	CorrDim	Avg	2	V6	N.S.	0%	96.15%
M vs. Dis	CorrDim	Avg	2	V2	N.S.	33.33%	85.71%
M vs. $CardMyo$	CorrDim	Avg	2	V2	N.S.	66.67%	94.44%
M vs. BBB	CorrDim	Avg	2	V4	0.0126	100%	100%
MI vs. MH	CorrDim	Avg	2	II	N.S.	100%	98.91%
MI vs. HC	CorrDim	Avg	2	V6	0.0017	100%	82.84%
MI vs. Dis	CorrDim	Avg	2	Vy	N.S.	100%	97.05%
MI vs. $CardMyo$	CorrDim	Avg	2	V6	0.0326	100%	96.02%
MI vs. BBB	CorrDim	Avg	2	aVr	0.0027	100%	97.57%
MH vs. HC	CorrDim	Avg	2	II	0.0078	0%	94.94%
MH vs. Dis	CorrDim	Avg	2	Vz	N.S.	50.00%	80.00%
MH vs. $CardMyo$	CorrDim	Avg	2	Vz	N.S.	50.00%	84.21%
MH vs. BBB	CorrDim	Avg	2	V2	N.S.	75.00%	92.31%
HC vs. Dis	CorrDim	Avg	2	V1	N.S.	100%	87.21%
HC vs. $CardMyo$	CorrDim	Avg	2	III	0.0071	94.67%	83.33%
HC vs. BBB	CorrDim	Avg	2	V2	0.0047	100%	89.29%
Dis vs. $CardMyo$	CorrDim	Avg	2	V3	N.S.	54.54%	73.08%
Dis vs. BBB	CorrDim	Avg	2	aVr	0.0167	81.81%	80.00%
$CardMyo$ vs. BBB	CorrDim	Avg	2	aVr	N.S.	73.33%	75.00%

Table 7 shows the number of occasions where a feature distribution is shown to be significant ($p < 0.05$) for separating binary classes.

Table 7. The total number of occasions that a feature was shown to be statistically significant ($p < 0.05$) across all sub-band analyses and leads.

[illegible]

Table 7. Cont.

Comparison Group	ApEn	CorrDim	DFA	En	H	EH	K	LogEn	Elay	ShaEn	Total
MI vs. CardMyo	48	42	48	48	48	48	48	48	48	48	474
MI vs. BBB	48	41	48	48	48	48	48	48	48	48	473
MH vs. HC	24	21	24	24	24	24	24	24	24	24	237
MH vs. Dis	0	0	0	0	0	0	0	0	0	0	0
MH vs. CardMyo	0	0	0	0	0	0	0	0	0	0	0
MH vs. BBB	0	0	0	0	0	0	0	0	0	0	0
HC vs. Dis	0	0	0	0	0	0	0	0	0	0	0
HC vs. CardMyo	48	42	48	48	48	48	48	48	48	48	474
HC vs. BBB	120	104	120	120	120	120	120	120	120	120	1184
Dis vs. CardMyo	0	0	0	0	0	0	0	0	0	0	0
Dis vs. BBB	24	21	24	24	24	24	24	24	24	24	237
CardMyo vs. BBB	0	0	0	0	0	0	0	0	0	0	0
Total	600	523	600	600	600	600	600	600	600	600	

Figure 2 illustrates the violin plots for the comparison groups where there was a significant difference, reported in Table 6.

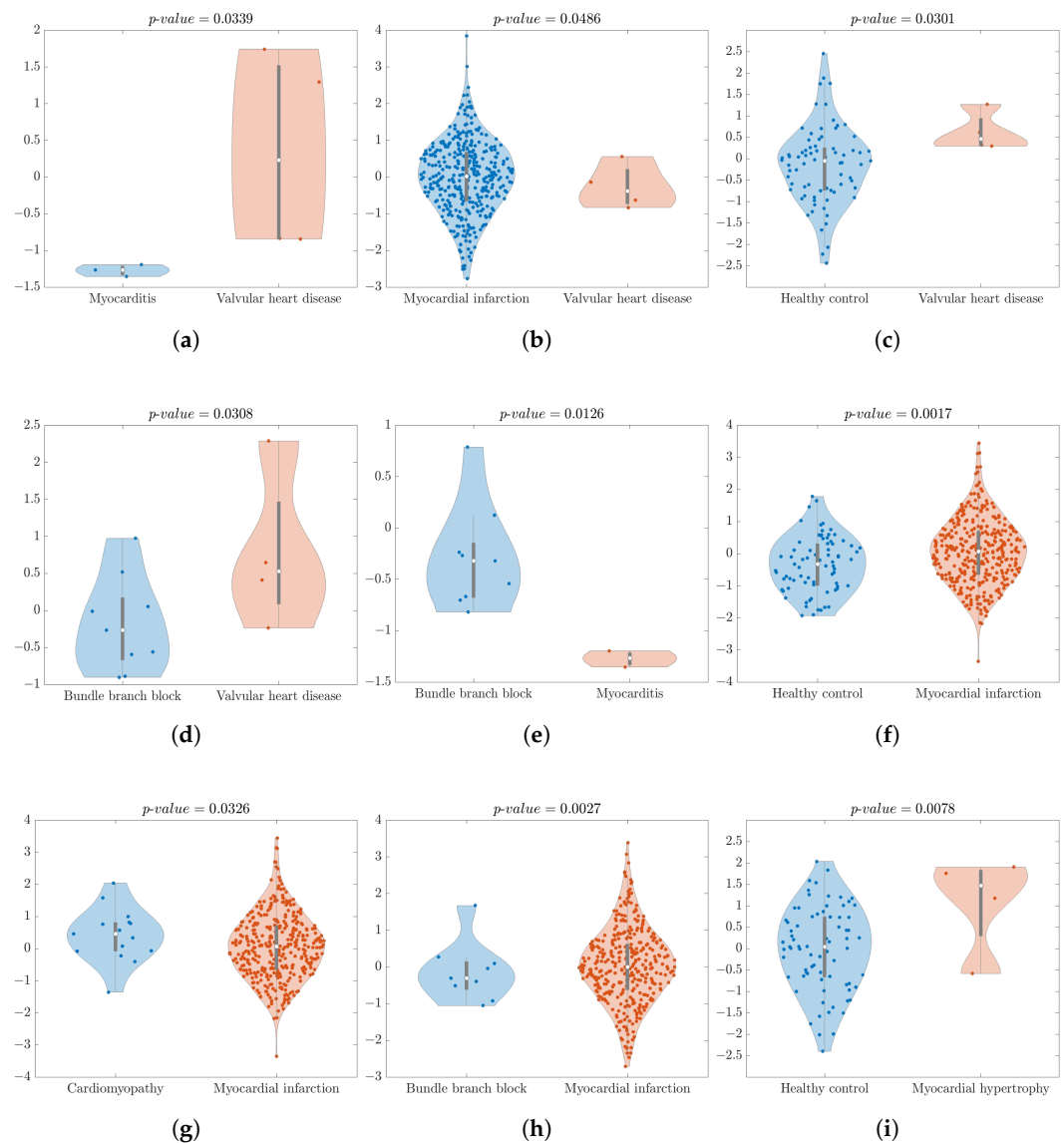


Figure 2. Cont.

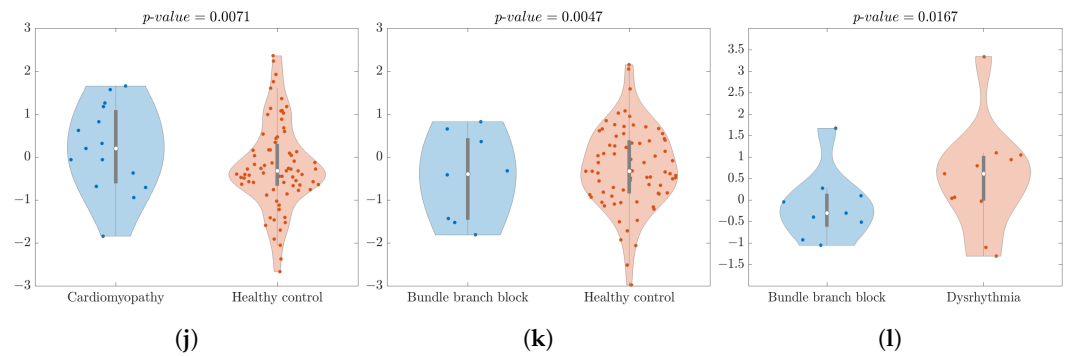


Figure 2. Violin plots of binary group distributions with significant differences—individual feature power analysis for discrimination. (a) VHD vs. M; (b) VHD vs. MI; (c) VHD vs. HC; (d) VHD vs. BBB; (e) M vs. BBB; (f) MI vs. HC; (g) MI vs. CardMyo; (h) MI vs. BBB; (i) MH vs. HC; (j) HC vs. CardMyo; (k) HC vs. BBB; (l) Dis vs. BBB.

The classification results regarding combined feature power analysis performed by ML classifiers can be found as a heatmap in Figure 3.

	Accuracy	Recall	Precision	F1-score	Kappa	MCC	CSI	Gmean	AUC	Lead	Classifier
All vs. All	81.16	72.93	81.16	76.34	0.4018	0.4399	0.6713	0.7417	0.5552	Combined	LogRegCV
VHD vs. M	100.00	100.00	100.00	100.00	1.0	1.0	1.0	1.0	1.0	I	LinSVC
VHD vs. MI	98.91	97.83	98.91	98.36	0.0	0.0	0.0	0.0	0.5	Combined	KNN
VHD vs. MH	75.00	83.33	75.00	73.33	0.5	0.5774	0.6667	0.8165	0.75	V3	AdaBoost
VHD vs. HC	100.00	100.00	100.00	100.00	1.0	1.0	1.0	1.0	1.0	V3	DeTreeC
VHD vs. Dis	100.00	100.00	100.00	100.00	1.0	1.0	1.0	1.0	1.0	V2	SGD
VHD vs. CardMyo	100.00	100.00	100.00	100.00	1.0	1.0	1.0	1.0	1.0	Combined	LinSVC
VHD vs. BBB	92.31	93.85	92.31	92.51	0.8312	0.8433	0.8	0.8944	0.9444	V3	LinSVC
M vs. MI	99.18	98.36	99.18	98.77	0.0	0.0	0.0	0.0	0.5	Combined	LinSVC
M vs. MH	100.00	100.00	100.00	100.00	1.0	1.0	1.0	1.0	1.0	III	SGD
M vs. HC	97.44	97.50	97.44	96.81	0.4902	0.5698	0.3333	0.9869	0.6667	V5	MLP
M vs. Dis	100.00	100.00	100.00	100.00	1.0	1.0	1.0	1.0	1.0	III	AdaBoost
M vs. CardMyo	100.00	100.00	100.00	100.00	1.0	1.0	1.0	1.0	1.0	V2	LinSVC
M vs. BBB	100.00	100.00	100.00	100.00	1.0	1.0	1.0	1.0	1.0	Combined	GradBoost
MI vs. MH	98.36	97.82	98.36	98.09	0.0073	0.0078	0.9836	0.0	0.4751	Combined	GauPro
MI vs. Dis	82.84	68.62	82.84	75.06	0.0	0.0	0.8192	0.0	0.5	Combined	LogRegCV
MI vs. CardMyo	97.32	97.39	97.32	96.22	0.1626	0.2974	0.9731	0.9865	0.5454	V5	AdaBoost
MI vs. BBB	95.23	92.17	95.23	93.67	0.0134	0.0182	0.9523	0.0	0.4272	Combined	GauPro
MI vs. HC	97.84	97.89	97.84	96.99	0.1961	0.3297	0.9784	0.9891	0.5556	V6	AdaBoost
MH vs. HC	96.20	96.35	96.20	95.10	0.3876	0.4903	0.25	0.9806	0.625	III	LogReg
MH vs. Dis	93.33	93.33	93.89	93.00	0.8148	0.8292	0.75	0.9574	0.875	III	DeTreeC
MH vs. CardMyo	78.95	78.95	78.95	78.95	0.3667	0.3667	0.3333	0.6583	0.6833	V1	SGD
MH vs. BBB	100.00	100.00	100.00	100.00	1.0	1.0	1.0	1.0	1.0	Vy	LinSVC
HC vs. Dis	95.35	95.18	95.35	95.15	0.774	0.779	0.9487	0.9243	0.857	Vz	LinSVC
HC vs. CardMyo	94.44	94.34	94.44	94.19	0.7826	0.7894	0.9367	0.9326	0.86	V1	GradBoost
HC vs. BBB	94.05	93.64	94.05	93.51	0.6354	0.6512	0.9367	0.8892	0.7711	avf	GradBoost
Dis vs. CardMyo	69.23	72.59	69.23	69.23	0.3988	0.4182	0.5294	0.7006	0.7091	avr	SGD
Dis vs. BBB	75.00	75.00	75.00	74.81	0.4898	0.4924	0.6429	0.75	0.7424	V2	DeTreeC
CardMyo vs. BBB	95.83	96.25	95.83	95.87	0.913	0.9165	0.9333	0.9487	0.9667	I	AdaBoost

Figure 3. Heatmap classification report regarding combined feature discriminant power analysis—the best Accuracy, Recall, Precision, F1-Score, AUC, Kappa, MCC, CSI, and Gmean results for each comparison group plus the information of lead and ML classifier applied for signal analysis.

The direct comparison between the individual and combined feature power analyses for discrimination is shown in Figure 4.

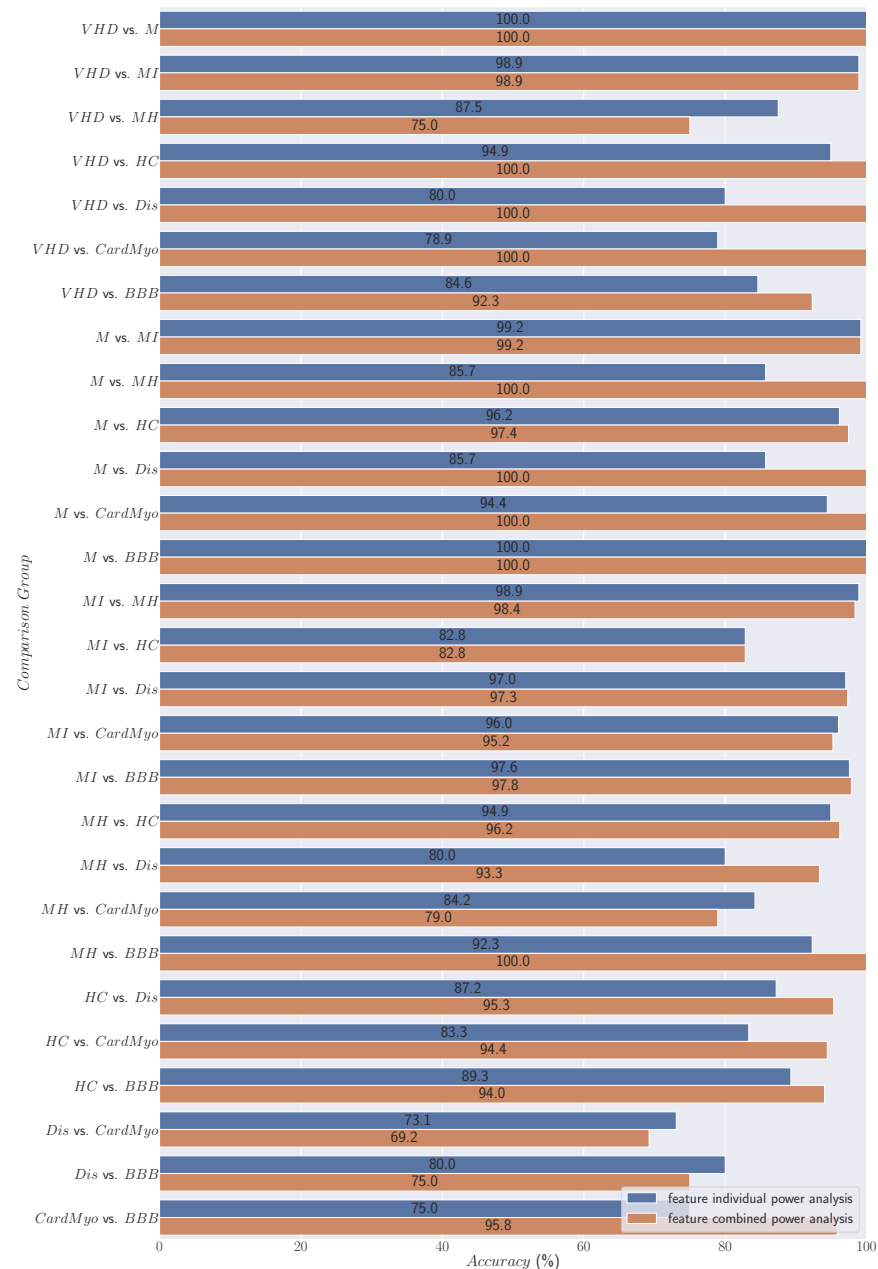


Figure 4. Direct comparison using *Accuracy* between individual and combined feature power analyses for binary groups' discrimination performed by ML models.

4. Discussion

For a more comprehensive discussion, we divided this section into three subsections according to the two variations of analysis employed in this study—individual and combined feature power analyses for discrimination—and compared our results with state-of-the-art results. While acknowledging that, in medicine, *Accuracy* may not fully capture the balance between *Recall* and *Specificity*, our discussion will primarily focus on *Accuracy* for checking the model's performance as it enables a more direct comparison of our results with those achieved by state-of-the-art methods.

4.1. Data Driven Analysis—Individual Feature Power Analysis

From a broader perspective, we can observe in Table 6 that the top-performing feature, compressor, and wavelet sub-band were the *CorrDim*, *Avg*, and 2nd sub-band (DWT

details 2nd level), respectively. Notably, they were present in 100% of the best results for comparison groups, encompassing all 28 binary comparison groups.

In addition, 12 of the 28 binary comparisons were shown to be statistically significant (42.86% of all analyses) and, out of the 15 leads utilised in this study, 8 exhibited at least one analysis with statistically significant differences. The most frequently represented lead in the table was V2, appearing in 16% of the cases (4 out of 28 binary groups).

The classes most frequently represented in binary groups with significant differences were the *HC* and the *BBB* classes. Both classes were present in 5 out of the 12 comparison groups with significant differences (41.67% of the cases). Additionally, the *CardMyo* and *MH* classes were the only ones that did not show significant differences when compared with the *BBB* class, and the *M* and *Dis* classes did not show significant differences when compared with the *HC* class.

Regarding each binary analysis:

- *VHD* vs. *M* analysis yielded a significant *p*-value of 0.0339, and an *Accuracy* and *Recall* of 100%. The feature *CorrDim*, with the compressor *Avg*, and lead V4 extracted from the 2nd sub-band provided excellent results. As shown in Figure 2a, the violin plot easily illustrates a distinct separation between these two classes.
- *VHD* vs. *MI* statistical analysis produced a significant result (*p*-value = 0.0486). The XROC analysis achieved an *Accuracy* of 98.91% and *Recall* of 0% for this binary comparison. The achieved *Recall* of 0% underscores one of the primary limitations of the dataset, namely its imbalance with the XROC, over-adjusting itself too much to the predominant class—*MI*, and it corroborates also the difficulty of splitting groups by the statistical test. In Figure 2b, we can see some outlier values but the highest density of the data is located close to the median.
- *VHD* vs. *MH* statistical analysis revealed a non-significance *p*-value. The XROC metrics—*Accuracy* and *Recall*—demonstrated strong performance for discrimination between groups, with values of 87.50% and 75.00%, respectively.
- *VHD* vs. *HC* group analysis displayed a significant difference, with a *p*-value of 0.0301. It achieved an *Accuracy* of 94.94% and a *Recall* of 0%. Figure 2c indicates some outliers in the *HC* class, but the highest data density is close to the median. Despite the good statistical analysis results, once more the XROC over-adjusts itself too much to the predominant class—*HC*, achieving a *Recall* of 0% for discriminating the class *VHD*. The imbalanced database and the *HC*'s large number of outliers contribute to these results. The XROC employs an averaging method within its genesis, which assigns significant weight to outliers in the final results.
- *VHD* vs. *Dis* analysis resulted in a non-significant *p*-value, an *Accuracy* of 80.00%, and a *Recall* of 100%. Despite being statistically non-significant, the XROC results showed a good performance for discriminating.
- *VHD* vs. *CardMyo* analysis resulted in a non-significant *p*-value. The XROC metrics *Accuracy* and *Recall* were 78.94% and 50.00%, respectively.
- *VHD* vs. *BBB* statistical analysis yielded a *p*-value of 0.0308, and the XROC achieved an *Accuracy* of 84.62% and *Recall* of 75.00%. Figure 2d illustrates a higher density of *BBB*'s data being located close to the median.
- *M* vs. *MI* statistical analysis yielded a *p*-value of non-significance. The *Accuracy* achieved a value of 99.18% and the *Recall* resulted in 0%.
- *M* vs. *MH* statistical analysis exhibited no significant difference in the *p*-value. The *Accuracy* and *Recall* reached 85.71% and 100%, respectively.
- *M* vs. *HC* analysis also showed no significant *p*-value. *Accuracy* and *Recall* demonstrated an interesting performance, achieving values of 96.15% and 0%, respectively. This underscores the difficulty of the XROC classifier in accurately discriminating between unbalanced data sample sizes.
- *M* vs. *Dis* analysis also revealed a non-significant *p*-value, with the *Accuracy* and *Recall* showcasing the values of 85.71% and 33.33%, respectively.

- *M* vs. *CardMyo* analysis resulted in a non-significant *p*-value, an *Accuracy* of 94.44% and a *Recall* of 66.67%. Despite being statistically non-significant, the XROC showed good behaviour for discriminating between classes.
- *M* vs. *BBB* statistical analysis indicated significant differences with a *p*-value of 0.0126. The *Accuracy* and *Recall* reached 100%. In Figure 2e, the violin plot displayed an outlier in the *BBB* class, but the highest data density was slightly below the median. It is worth noting that there was a clear separation between the two classes.
- *MI* vs. *MH* analysis resulted in a non-significant *p*-value. Despite that, the XROC performed well with a discrimination *Accuracy* of 98.91% and a *Recall* of 100%.
- *MI* vs. *HC* statistical analysis yielded a *p*-value of 0.0017. The *Accuracy* and the *Recall* achieved values of 82.84% and 100%, respectively. In Figure 2f, the violin plot exhibited some outliers, but the highest data density was close to the median for both classes.
- *MI* vs. *Dis* analysis was shown to be non-significant. The XROC metrics of *Accuracy* and *Recall* displayed significant performance percentages, with values of 97.05% and 100%, respectively.
- *MI* vs. *CardMyo* statistical analysis revealed a *p*-value of 0.0326 alongside impressive classification metrics, boasting an *Accuracy* of 96.29% and a perfect *Recall* of 100%. In Figure 2g, the violin plot exhibited some outliers, but the highest data density was close to the median for both classes.
- *MI* vs. *BBB* analysis showed statistical significance and the *Accuracy* stood at 97.57%, with a flawless *Recall* of 100%. Figure 2h gives us the opportunity to see some outliers in both classes but the majority of the data were located close to the median.
- *MH* vs. *HC* analysis achieved a significant *p*-value, reaching a statistical analysis value of 0.0078. The *Accuracy* was 94.94% and the *Recall* was 0%, which perfectly illustrates the imbalance of the dataset. Figure 2i shows the *HC* class, with the highest density of data close to the median.
- *MH* vs. *Dis* analysis exhibited non-significant *p*-values, with *Accuracy* rates of 80.00% and 50.00%, respectively.
- *MH* vs. *CardMyo* demonstrated an *Accuracy* of 84.21% and a *Recall* of 50.00%, while the *MH* vs. *BBB* group yielded an *Accuracy* of 92.31% and a *Recall* of 75.00%, with both analyses showing a non-statistical significance. While statistical significance may be elusive, the consistently high *Accuracy* and *Recall* values underscore the potential efficacy of the model in discriminating between different conditions within the studied groups.
- *HC* vs. *Dis* analysis showed non-significant difference. The *Accuracy* and *Recall* displayed great performance, with values of 87.21% and 100%, respectively.
- *HC* vs. *CardMyo* showed a significant difference, with a *p*-value of 0.0071. The *Accuracy* and *Recall* exhibited strong performance, with values of 83.33% and 94.67%, respectively. In Figure 2j, the violin plot displayed some outliers in the *HC* class, but the highest data density was close to the median.
- *HC* vs. *BBB* analysis provided a significant *p*-value of 0.0047 accompanied by an *Accuracy* of 89.29% and an impressive *Recall* of 100%. Figure 2k shows the violin plot with some outliers in the *HC* class, but there was a higher density of data close to the median.
- *Dis* vs. *CardMyo* comparison analysis yielded a non-significant *p*-value, with an *Accuracy* of 73.08% and a *Recall* of 54.54%.
- *Dis* vs. *BBB* comparison analysis showed a *p*-value of 0.0167, achieving an *Accuracy* of 80.00% and a *Recall* of 81.81%. Figure 2l shows a violin plot with a couple of outliers in both classes, but the highest density of data was close to the median.
- *CardMyo* vs. *BBB* analysis provided a non-significant *p*-value, with an *Accuracy* of 75.00% and a *Recall* of 73.33%.

Looking to Table 7, we can see the total number of occasions that a feature was demonstrated to be significant over binary groups and in total. It should be noted that

each originally defined feature generated 360 features per analysis; for more information check Section 2.5.1. While *CorrDim* emerged as a standout performer individually, the results emphasise that the other nine features also demonstrated statistical significance in distinguishing between classes with more moments of appearing to be significant than actually *CorrDim* (523 vs. 600). *MI* vs. *HC* and *HC* vs. *BBB* showed the highest number of results with significant differences, which were 1185 and 1184, respectively. *VHD* vs. *MI*, *VHD* vs. *BBB*, *M* vs. *BBB*, *MH* vs. *HC*, and *Dis* vs. *BBB* were the binary groups with the lowest amount of occasions of significant feature distributions, 237 each.

4.2. Data Driven Analysis—Combined Feature Power Analysis

Figure 3 presents the classification metrics report for the comparison groups provided by 19 Sci-learn ML classifiers with combined features as entries. The heatmap employs a gradient of green shades in its colour scheme, serving to vividly illustrate the method's discrimination capabilities for *Accuracy*, *Recall*, *Precision*, *F1-Score*, *AUC*, *Kappa*, *MCC*, *CSI*, and *Gmean* in each comparative analysis. Lighter shades of green represent lower discriminatory power, while deeper, richer greens signify higher effectiveness. By looking into the results, it can be seen that *VHD* vs. *M*, *VHD* vs. *MI*, *VHD* vs. *HC*, *VHD* vs. *Dis*, *VHD* vs. *CardMyo*, *M* vs. *MH*, *M* vs. *Dis*, *M* vs. *CardMyo*, *M* vs. *BBB*, and *MH* vs. *BBB* obtained 100% on all evaluation metrics. Comparing the individual power discrimination results presented in Table 6, it can be seen that generally the discrimination results have increased, and the ratio of 100% on all evaluated metrics per binary analysis has increased (2/28 to 10/28). Comparing the *Accuracy* results achieved through combined feature power analysis with those obtained through individual feature power analysis (see Figure 4 for a visual representation of the analysis for each binary comparison; this figure provides a clear and concise overview, facilitating an easy assessment of performance differences between the two approaches described in Sections 2.5.1 and 2.5.2), we observe a significant overall improvement across all binary comparisons. Among the 28 comparisons conducted, the results indicate that 17 exhibits enhanced discrimination *Accuracy* when utilising combined features analysis. In contrast, in five instances, the *Accuracy* remained the same as that observed in individual feature power analysis. There are only six cases where we notice a decrease in *Accuracy* compared with individual power analysis (*VHD* vs. *MH*, *MI* vs. *MH*, *MI* vs. *CardMyo*, *MH* vs. *CardMyo*, *Dis* vs. *CardMyo*, and *Dis* vs. *BBB*).

Returning to the analysis of Figure 3, in *All* vs. *All*, an *Accuracy* of 81.16%, *Recall* of 72.93%, *Precision* of 81.16%, 76.34% for the *F1-Score*, *Kappa* of 0.4018, *MCC* of 0.4399, *CSI* of 0.6713, *Gmean* of 0.7417, and *AUC* of 0.5552 were achieved. The leads ensemble combination was the most represented in the table, corresponding to 28% of the total appearances. The classifier with the most frequent appearances was *LinSVC*, representing 24% of the cases. The binary groups *VHD* vs. *MH*, *MI* vs. *HC*, *MH* vs. *CardMyo*, *Dis* vs. *CardMyo*, and *Dis* vs. *BBB*, exhibited *Precision* values below 90%. This challenge in correct classifying can be attributed to the close relationship between *CardMyo* and *MH* or *Dis*. In a clinical context, it is common for patients to present with *CardMyo* alongside either *VHD* or *Dis* [68,69]. This clinical overlap makes accurate differentiation challenging. Understanding and addressing these interconnected conditions are essential for improving classification *Accuracy* in these scenarios. The *MI* vs. *HC* classification, with an 82.84% *Precision*, presents challenges due to the potential presence of acute *MI* within the *HC* class. Additionally, some patients who have recovered from *MI* may be categorised as *HC* [70]. These factors contribute to a slightly lower classification performance of ML models for discrimination within this context.

Moreover, upon assessing various models and their performance metrics, a notable observation is the impact of the imbalanced dataset, particularly evident in comparison groups involving one of either *MI* or *HC* classes. In such instances, we observed a range of *AUC* results from 0.4272 to 0.6667 across all nine comparison groups where at least one of these two classes was present. These findings underscore the substantial challenge of distinguishing between unevenly represented classes. The *Gmean* further highlights the

noteworthy observation that in 71.42% of cases (five out of seven binary comparisons) where the *MI* class is pitted against another class, the *Gmean* metric yields a result of 0. However, in comparisons involving *MI* against *Dis* and *BBB*, a lower risk of over-fitting is evident, with *Gmean* values of 0.9865 and 0.9891, respectively. The *CSI* metric reveals that the preponderance of comparison groups, specifically 17 out of 29, exhibits results surpassing 0.9. This observation underscores a notable challenge in classification, particularly when dealing with classes characterised by higher data abundance. The *MCC* metric highlights a notable trend, with 31% of the comparison groups (9 out of 29 groups) achieving perfect predictions, each obtaining a maximum value of 1. Notably, the class *MI* demonstrates the least favourable outcomes, with its highest *MCC* value capped at 0.3297 when included in a comparison group. The *Kappa* metric reveals a noteworthy pattern, with 31% of the comparison groups (9 out of 29 groups) achieving perfect agreement, each attaining a maximum value of 1. Additionally, 41.37% of the groups surpass a *Kappa* value higher than 0.083.

4.3. Study Results vs. State-of-the-Art Results

When we analyse Table 1, it becomes evident that our results closely match or slightly surpass the achievements of the state of the art, offering valuable insights for enhancing robustness. In particular, when considering the eight state-of-the-art studies that utilised the PTB database, our results are lower in the binary comparisons of *MI* vs. *HC* and *HC* vs. *CardMyo*, with differences of less than 13% and 0.76%, respectively. Furthermore, the present study offers significant advantages over other studies as it includes a multi-class comparison group (*All* vs. *All*), a higher variety of binary comparisons, and the application of classical non-linear analysis under ECG multi-band analysis performed by DWT. These particularities allow a high capacity of differentiation of each class present in the database, a level of detail not typically found in state-of-the-art articles. Moreover, it is imperative to underscore that the developed algorithm relies on ECG signals, presenting distinctive advantages when compared with alternative diagnostic sources such as stress tests, imaging tests, electrophysiological studies, provocative diagnostic tests, genetic tests, and cardiac catheterisation, among others. The affordability, non-invasiveness, widespread use in clinical settings, and user-friendly nature of ECG make it an optimal choice. Its efficacy not only facilitates the easy adoption of our algorithm globally but also addresses the unique needs of patients unable to leave their hospital beds. This highlights the algorithm's versatility and accessibility in diverse healthcare settings.

5. Conclusions

For this research, 10 non-linear features (*En*, *ApEn*, *LogEn*, *ShaEn*, *EH*, *Elay*, *H*, *K*, *CorrDim*, and *DFA*) were extracted from a well-known ECG database (PTB diagnostic ECG database). From the recorded 15 leads per patient (12 conventional leads and 3 Frank leads), each signal lead underwent a 1-second length non-overlapped windowing process over time for extracting a total of 10 non-linear features per window. At the end of the process, each feature time series was compacted by six statistics. The individual power and combined power were accessed from discriminating between different cardiovascular pathologies (*VHD* vs. *M*, *VHD* vs. *MI*, *VHD* vs. *MH*, *VHD* vs. *HC*, *VHD* vs. *Dis*, *VHD* vs. *CardMyo*, *VHD* vs. *BBB*, *M* vs. *MI*, *M* vs. *MH*, *M* vs. *HC*, *M* vs. *Dis*, *M* vs. *CardMyo*, *M* vs. *BBB*, *MI* vs. *MH*, *MI* vs. *HC*, *MI* vs. *Dis*, *MI* vs. *CardMyo*, *MI* vs. *BBB*, *MH* vs. *HC*, *MH* vs. *Dis*, *MH* vs. *CardMyo*, *MH* vs. *BBB*, *HC* vs. *Dis*, *HC* vs. *CardMyo* and *HC* vs. *BBB*), *Dis* vs. *CardMyo*, *Dis* vs. *BBB*), and *CardMyo* vs. *BBB*) and one multi-class comparison (*All* vs. *All*).

The *Accuracy* discrimination results ranged between 81% and 100%. The results demonstrate that the applied method serves as a robust tool for effectively distinguishing cardiovascular diseases (CVDs) through the analysis of ECG signals. The level of detail and discrimination achieved surpasses what is typically observed in state-of-the-art studies using the same dataset. Despite our results indicating a great ability of the proposed method

to diagnose, offering in this way another alternative avenue for medical doctors to arrive at more confident diagnoses, this study had some limitations. (1) The inherently technical nature of utilising unusual standard clinical features extracted from ECG signals may hinder complete interpretability from a clinician's standpoint. This could pose a challenge to its rapid and widespread integration into clinical practice. (2) The high computation time of multi-band analysis for the chosen methodology led us to choose just one wavelet (Symlet7) from tens of wavelets with the level of decomposition set to 3, based on prior work [48,49], as the main wavelet. A more meticulous analysis needs to be carried out in future to choose the wavelet and level of decomposition that adjusts itself better to each CVD activity. (3) The results should be further enhanced by updating them with a larger and more balanced population to ensure a more reliable generalisation and to split data as hold-out for classifying (e.g., 70% for training and 30% for testing) without employing cross-validation methods. Another possible solution would be, in a future work, to reduce the number of cases inside the highest classes to reduce the uneven data distribution. (4) Additional CVDs should be studied and evaluated in future work to enhance the discriminative capabilities of our algorithm (e.g., arrhythmias such as premature atrial contraction, premature ventricular contraction, and atrial fibrillation).

Nevertheless, upon reviewing state-of-the-art works (refer to Table 1), it becomes apparent that many have encountered similar limitations. These constraints predominantly revolved around imbalances in data distribution, as a significant portion of these studies relied on the same database. Additionally, limitations in computational time and resources, and a restricted variety and diversity of CVD classes were commonly shared among these works. This collective set of limitations across the consulted literature underscores the need for addressing data imbalances and expanding the diversity of CVD classes in future research efforts.

Author Contributions: Conceptualization, P.R. and P.M.R.; methodology, P.R. and P.M.R.; validation, P.M.R.; investigation, P.R. and P.M.R.; writing—original, P.R., J.S., D.P. and P.M.R.; writing—review and editing, P.R. and P.M.R.; supervision, P.M.R.; funding acquisition, P.M.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data presented in this study are available on request from the corresponding author (accurately indicate status).

Acknowledgments: This work was supported by National Funds from FCT—Fundação para a Ciência e a Tecnologia through project UIDB/50016/2020.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. American Heart Association. What is Cardiovascular Disease? 2017. Available online: <https://www.heart.org/en/health-topics/consumer-healthcare/what-is-cardiovascular-disease>
2. World Health Organization. Cardiovascular Diseases CVDs. 2021. Available online: [https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-\(cvds\)](https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)) (accessed on 5 October 2023).
3. Visseren, F.L.J.; Mach, F.; Smulders, Y.M.; Carballo, D.; Koskinas, K.C.; Bäck, M.; Benetos, A.; Biffi, A.; Boavida, J.M.; Capodanno, D.; et al. 2021 ESC Guidelines on cardiovascular disease prevention in clinical practice. *Eur. Heart J.* **2021**, *42*, 3227–3337. [CrossRef] [PubMed]
4. Arbelo, E.; Protonotarios, A.; Gimeno, J.; Arbustini, E.; Barriales-Villa, R.; Basso, C.; Bezzina, C.; Biagini, E.; Blom, N.; Boer, R.; et al. 2023 ESC Guidelines for the management of cardiomyopathies. *Eur. Heart J.* **2023**, *44*, 3503–3626. [CrossRef] [PubMed]
5. Delgado, V.; Marsan, N.A.; de Waha, S.; Bonaros, N.; Brida, M.; Burri, H.; Caselli, S.; Doenst, T.; Ederhy, S.; Erba, P.A.; et al. 2023 ESC Guidelines for the management of endocarditis. *Eur. Heart J.* **2023**, *44*, 3948–4042. [CrossRef] [PubMed]
6. Brito, D.; Cardim, N.; Rocha-Lopes, L.; Freitas, A.; Lacerda, A.P.D.; Menezes, M.; Belo, A.; Martins, E.; Peres, M.; Goncalves, L.; et al. P3514Diagnosis and treatment of acute myocarditis in Portugal. Data from the national multicenter registry on myocarditis. *Eur. Heart J.* **2017**, *38*, ehx504.P3514. [CrossRef]
7. Adler, Y.; Charron, P.; Imazio, M.; Badano, L.; Barón-Esquivias, G.; Bogaert, J.; Brucato, A.; Gueret, P.; Klingel, K.; Lionis, C.; et al. 2015 ESC Guidelines for the diagnosis and management of pericardial diseases. *Eur. Heart J.* **2015**, *36*, 2921–2964. [CrossRef]

8. Hamm, C.; Bassand, J.P.; Agewall, S.; Bax, J.; Boersma, E.; Bueno, H.; Caso, P.; Dudek, D.; Gielen, S.; Huber, K.; et al. ESC Guidelines for the management of acute coronary syndromes in patients presenting without persistent ST-segment elevation: The Task Force for the management of acute coronary syndromes (ACS) in patients presenting without persistent ST-segment elevation of the European Society of Cardiology (ESC). *Eur. Heart J.* **2011**, *32*, 2999–3054. [\[CrossRef\]](#)
9. Byrne, R.; Rossello, X.; Coughlan, J.; Barbato, E.; Berry, C.; Chieffo, A.; Claeys, M.; Dan, G.A.; Dweck, M.; Galbraith, M.; et al. 2023 ESC Guidelines for the management of acute coronary syndromes: Developed by the task force on the management of acute coronary syndromes of the European Society of Cardiology (ESC). *Eur. Heart J.* **2023**, *44*, 3720–3826. [\[CrossRef\]](#)
10. McDonagh, T.A.; Metra, M.; Adamo, M.; Gardner, R.S.; Baumbach, A.; Böhm, M.; Burri, H.; Butler, J.; Čelutkienė, J.; Chioncel, O.; et al. 2021 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure. *Eur. Heart J.* **2021**, *42*, 3599–3726. [\[CrossRef\]](#)
11. Zeppenfeld, K.; Tfelt-Hansen, J.; de Riva, M.; Winkel, B.G.; Behr, E.R.; Blom, N.A.; Charron, P.; Corrado, D.; Dagres, N.; de Chillou, C.; et al. 2022 ESC Guidelines for the management of patients with ventricular arrhythmias and the prevention of sudden cardiac death. *Eur. Heart J.* **2022**, *43*, 3997–4126. [\[CrossRef\]](#)
12. Kim, M.; Yoon, M.; Yang, P.; Kim, T.; Uhm, J.; Kim, J.; Pak, H.; Lee, M.; Joung, B. P6422Sex-based disparities in incidence, treatment, and outcomes of sudden cardiac arrest. *Eur. Heart J.* **2017**, *38*, ehx493.P6422. [\[CrossRef\]](#)
13. Authors/Task Force Members.; Vahanian, A.; Alfieri, O.; Andreotti, F.; Antunes, M.J.; Barón-Esquivias, G.; Baumgartner, H.; Borger, M.A.; Carrel, T.P.; Bonis, M.D.; et al. Guidelines on the management of valvular heart disease (version 2012). *Eur. Heart J.* **2012**, *33*, 2451–2496. [\[CrossRef\]](#)
14. Baumgartner, H.; Backer, J.D.; Babu-Narayan, S.V.; Budts, W.; Chessa, M.; Diller, G.P.; lung, B.; Kluin, J.; Lang, I.M.; Meijboom, F.; et al. 2020 ESC Guidelines for the management of adult congenital heart disease. *Eur. Heart J.* **2020**, *42*, 563–645. [\[CrossRef\]](#) [\[PubMed\]](#)
15. Kligfield, P.; Gettes, L.S.; Bailey, J.J.; Childers, R.; Deal, B.J.; Hancock, E.W.; van Herpen, G.; Kors, J.A.; Macfarlane, P.; Mirvis, D.M.; et al. Recommendations for the Standardization and Interpretation of the Electrocardiogram. *Circulation* **2007**, *115*, 1306–1324. [\[CrossRef\]](#) [\[PubMed\]](#)
16. Garner, K.; Pomeroy, W.; Arnold, J. Exercise Stress Testing: Indications and Common Questions. *Am. Acad. Fam. Physicians* **2017**, *96*, 293–299.
17. Maron, B. American College of Cardiology/European Society of Cardiology Clinical Expert Consensus Document on Hypertrophic Cardiomyopathy a Rteport of the American College of Cardiology Foundation Task Force on Clinical Expert Consensus Documents and the European Society of Cardiology Committee for Practice Guidelines. *Eur. Heart J.* **2003**, *24*, 1965–1991. [\[CrossRef\]](#)
18. Perrot, B.; Clozel, J.P.; de la Chaise, A.T.; Cherrier, F.; Faivre, G. Electrophysiological effects of intravenous prostacyclin in man. *Eur. Heart J.* **1984**, *5*, 883–889. [\[CrossRef\]](#)
19. Maseri, A. Safety of provocative tests of coronary artery spasm and prediction of long-term outcome: need for an innovative clinical research strategy. *Eur. Heart J.* **2012**, *34*, 252–254. [\[CrossRef\]](#)
20. Grondin, S.; Davies, B.; Cadrin-Tourigny, J.; Steinberg, C.; Cheung, C.C.; Jorda, P.; Healey, J.S.; Green, M.S.; Sanatani, S.; Alqarawi, W.; et al. Importance of genetic testing in unexplained cardiac arrest. *Eur. Heart J.* **2022**, *43*, 3071–3081. [\[CrossRef\]](#)
21. Azmi, J.; Arif, M.; Nafis, M.T.; Alam, M.A.; Tanweer, S.; Wang, G. A systematic review on machine learning approaches for cardiovascular disease prediction using medical big data. *Med. Eng. Phys.* **2022**, *105*, 103825. [\[CrossRef\]](#)
22. Rodrigues, P.M.; Madeiro, J.P.; Marques, J.A.L. Enhancing Health and Public Health through Machine Learning: Decision Support for Smarter Choices. *Bioengineering* **2023**, *10*, 792. [\[CrossRef\]](#)
23. Krittanawong, C.; Virk, H.U.H.; Bangalore, S.; Wang, Z.; Johnson, K.W.; Pinotti, R.; Zhang, H.; Kaplin, S.; Narasimhan, B.; Kitai, T.; et al. Machine learning prediction in cardiovascular diseases: a meta-analysis. *Sci. Rep.* **2020**, *10*, 16057. [\[CrossRef\]](#) [\[PubMed\]](#)
24. Qu, Z.; Hu, G.; Garfinkel, A.; Weiss, J. Nonlinear and stochastic dynamics in the heart. *Phys. Rep.* **2014**, *543*, 61–162. [\[CrossRef\]](#)
25. Haraldsson, H.; Edenbrandt, L.; Ohlsson, M. Detecting acute myocardial infarction in the 12-lead ECG using Hermite expansions and neural networks. *Artif. Intell. Med.* **2004**, *32*, 127–136. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Begum, R.; Ramesh, M. Detection of cardiomyopathy using support vector machine and artificial neural network. *Int. J. Comput. Appl.* **2016**, *133*, 29–34. [\[CrossRef\]](#)
27. Chowdhury, H.; Sultana, M.; Ghosh, R.; Ahamed, J.; Mahmood, M. AI Assisted Portable ECG for Fast and Patient Specific Diagnosis. In Proceedings of the 2018 International Conference on Computer, Communication, Chemical, Material and Electronic Engineering (IC4ME2), Rajshahi, Bangladesh, 8–9 February 2018. [\[CrossRef\]](#)
28. Kachuee, M.; Fazeli, S.; Sarrafzadeh, M. ECG Heartbeat Classification: A Deep Transferable Representation. In Proceedings of the 2018 IEEE International Conference on Healthcare Informatics (ICHI), New York, NY, USA, 4–7 June 2018. [\[CrossRef\]](#)
29. Baloglu, U.; Talo, M.; Yildirim, O.; Tan, R.; Acharya, U. Classification of myocardial infarction with multi-lead ECG signals and deep CNN. *Pattern Recognit. Lett.* **2019**, *122*, 23–30. [\[CrossRef\]](#)
30. Ali, L.; Niamat, A.; Khan, J.A.; Golilarz, N.A.; Xingzhong, X.; Noor, A.; Nour, R.; Bukhari, S.A.C. An Optimized Stacked Support Vector Machines Based Expert System for the Effective Prediction of Heart Failure. *IEEE Access* **2019**, *7*, 54007–54014. [\[CrossRef\]](#)
31. Ali, L.; Rahman, A.; Khan, A.; Zhou, M.; Javeed, A.; Khan, J.A. An Automated Diagnostic System for Heart Disease Prediction Based on χ^2 Statistical Model and Optimally Configured Deep Neural Network. *IEEE Access* **2019**, *7*, 34938–34945. [\[CrossRef\]](#)

32. Baghel, N.; Dutta, M.K.; Burget, R. Automatic diagnosis of multiple cardiac diseases from PCG signals using convolutional neural network. *Comput. Methods Programs Biomed.* **2020**, *197*, 105750. [\[CrossRef\]](#)
33. Ahamed, A.; Hasan, K.; Monowar, K.; Mashnoor, N.; Hossain, A. ECG Heartbeat Classification Using Ensemble of Efficient Machine Learning Approaches on Imbalanced Datasets. In Proceedings of the 2020 2nd International Conference on Advanced Information and Communication Technology (ICAICT), Dhaka, Bangladesh, 28–29 November 2020. [\[CrossRef\]](#)
34. Makimoto, H.; Höckmann, M.; Lin, T.; Glöckner, D.; Gerguri, S.; Clasen, L.; Schmidt, J.; Assadi-Schmidt, A.; Bejinariu, A.; Müller, P.; et al. Performance of a convolutional neural network derived from an ECG database in recognizing myocardial infarction. *Sci. Rep.* **2020**, *10*, 8445. [\[CrossRef\]](#)
35. Khan, A.; Hussain, M.; Malik, M. Cardiac Disorder Classification by Electrocardiogram Sensing Using Deep Neural Network. *Complexity* **2021**, *2021*, 5512243. [\[CrossRef\]](#)
36. Premanand, S.; Narayanan, S. A Tree Based Machine Learning Approach for PTB Diagnostic Dataset. *J. Phys. Conf. Ser.* **2021**, *2115*, 012042. [\[CrossRef\]](#)
37. Kavitha, M.; Gnaneswar, G.; Dinesh, R.; Sai, Y.; Suraj, R. Heart Disease Prediction using Hybrid machine Learning Model. In Proceedings of the 2021 6th International Conference on Inventive Computation Technologies (ICICT), Coimbatore, India, 20–22 January 2021. [\[CrossRef\]](#)
38. Elhoseny, M.; Mohammed, M.; Mostafa, S.; Abdulkareem, K.; Maashi, M.; Garcia-Zapirain, B.; Mutlag, A.; Maashi, M. A New Multi-Agent Feature Wrapper Machine Learning Approach for Heart Disease Diagnosis. *Comput. Mater. Contin.* **2021**, *67*, 51–71. [\[CrossRef\]](#)
39. Ahmad, G.; Fatima, H.; Ullah, S.; Saidi, A.; Imdadullah. Efficient Medical Diagnosis of Human Heart Diseases Using Machine Learning Techniques with and without GridSearchCV. *IEEE Access* **2022**, *10*, 80151–80173. [\[CrossRef\]](#)
40. Ahmad, S.; Asghar, M.; Alotaibi, F.; Alotaibi, Y. Diagnosis of cardiovascular disease using deep learning technique. *Soft Comput.* **2022**, *27*, 8971–8990. [\[CrossRef\]](#)
41. Mhamdi, L.; Dammak, O.; Cottin, F.; Dhaou, I. Artificial Intelligence for Cardiac Diseases Diagnosis and Prediction Using ECG Images on Embedded Systems. *Biomedicines* **2022**, *10*, 2013. [\[CrossRef\]](#) [\[PubMed\]](#)
42. Karthik, S.; Santhosh, M.; S. Kavitha, M.; Christopher Paul, A. Automated Deep Learning Based Cardiovascular Disease Diagnosis Using ECG Signals. *Comput. Syst. Sci. Eng.* **2022**, *42*, 183–199. [\[CrossRef\]](#)
43. Bousseljot, R.D.; Kreiseler, D.; Schnabel, A. Nutzung der EKG-Signaldatenbank CARDIODAT der PTB über das Internet. *Biomed. Tech. Biomed. Eng.* **1995**, *40*, 317–318. [\[CrossRef\]](#)
44. Rodrigues, P.; Bispo, B.; Garrett, C.; Alves, D.; Teixeira, J.; Freitas, D. Lacsogram: A New EEG Tool to Diagnose Alzheimer's Disease. *IEEE J. Biomed. Health Inform.* **2021**, *25*, 3384–3395. [\[CrossRef\]](#)
45. Guido, R. Wavelets behind the scenes: Practical aspects, insights, and perspectives. *Phys. Rep.* **2022**, *985*, 1–23. [\[CrossRef\]](#)
46. Malvar, H. *Signal Processing with Lapped Transforms*; Artech House: Norwood, MA, USA, 1992.
47. Vetterli, M.; Kovačević, J. *Wavelets and Subband Coding*; Prentice Hall: Englewood Cliffs, NJ, USA, 1995.
48. Chen, C.C.; Tsui, F.R. Comparing different wavelet transforms on removing electrocardiogram baseline wanders and special trends. *BMC Med. Inform. Decis. Mak.* **2020**, *20*, 343. [\[CrossRef\]](#)
49. Ribeiro, P.; Marques, J.A.L.; Pordeus, D.; Zacarias, L.; Leite, C.F.; Sobreira-Neto, M.A.; Peixoto, A.A.; de Oliveira, A.; do Vale Madeiro, J.P.; Rodrigues, P.M. Machine learning-based cardiac activity non-linear analysis for discriminating COVID-19 patients with different degrees of severity. *Biomed. Signal Process. Control.* **2024**, *87*, 105558. [\[CrossRef\]](#)
50. Rioul, O.; Vetterli, M. Wavelets and signal processing. *IEEE Signal Process. Mag.* **1991**, *8*, 14–38. [\[CrossRef\]](#)
51. Peck, R.; Olsen, C.; Devore, J. *Introduction to Statistics and Data Analysis*; Cengage Learning: Boston, MA, USA, 2008; p. 880.
52. Caesarendra, W.; Kosasih, B.; Tieu, K.; Moodie, C. An application of nonlinear feature extraction—A case study for low speed slewing bearing condition monitoring and prognosis. In Proceedings of the 2013 IEEE/ASME International Conference on Advanced Intelligent Mechatronics, Wollongong, NSW, Australia, 9–12 July 2013; pp. 1713–1718. [\[CrossRef\]](#)
53. Hardstone, R.; Poil, S.S.; Schiavone, G.; Jansen, R.; Nikulin, V.; Mansvelder, H.; Linkenkaer-Hansen, K. Detrended Fluctuation Analysis: A Scale-Free View on Neuronal Oscillations. *Front. Physiol.* **2012**, *3*, 450. [\[CrossRef\]](#) [\[PubMed\]](#)
54. Sundararajan, D. *Discrete Wavelet Transform a Signal Processing Approach*, 1st ed.; John Wiley & Sons: Hoboken, NJ, USA, 2015.
55. Silva, M.; Ribeiro, P.; Bispo, B.C.; Rodrigues, P.M. Detecção da Doença de Alzheimer através de Parâmetros Não-Lineares de Sinais de Fala. In Proceedings of the Anais do XLI Simpósio Brasileiro de Telecomunicações e Processamento de Sinais. Sociedade Brasileira de Telecomunicações, São José dos Campos, SP, Brazil, 8–11 October 2023. [\[CrossRef\]](#)
56. Garcia, A.; Garcia, C.; Villasenor-Pineda, L.; Montoya, O. *Biosignal Processing and Classification Using Computational Learning and Intelligence Principles, Algorithms, and Applications*; Academic Press: London, UK, 2022; pp. 59–91.
57. Silva, G.; Batista, P.; Rodrigues, P.M. COVID-19 activity screening by a smart-data-driven multi-band voice analysis. *J. Voice* **2022**, *in press*. [\[CrossRef\]](#) [\[PubMed\]](#)
58. Nakas, C.; Yiannoutsos, C. Ordered multiple-class ROC analysis with continuous measurements. *Stat. Med.* **2004**, *23*, 3437–3449. [\[CrossRef\]](#)
59. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
60. Sammut, C.; Webb, G.I. (Eds.) *Accuracy*. In *Encyclopedia of Machine Learning and Data Mining*; Springer: New York, NY, USA, 2017; p. 8. [\[CrossRef\]](#)

61. Doğan, O. 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, PA, USA, 2019; pp. 108–127. [\[CrossRef\]](#)
62. Ting, K.M. Precision and Recall. In *Encyclopedia of Machine Learning and Data Mining*; Springer: New York, NY, USA, 2017; pp. 990–991. [\[CrossRef\]](#)
63. Goutte, C.; Gaussier, E., A Probabilistic Interpretation of Precision, Recall and F-Score, with Implication for Evaluation. In *Advances in Information Retrieval*; Springer: Berlin/Heidelberg, Germany, 2005; pp. 345–359. [\[CrossRef\]](#)
64. Vieira, S.M.; Kaymak, U.; Sousa, J.M.C. Cohen’s kappa coefficient as a performance measure for feature selection. In Proceedings of the International Conference on Fuzzy Systems, Barcelona, Spain, 18–23 July 2010. [\[CrossRef\]](#)
65. Chicco, D.; Jurman, G. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genom.* **2020**, *21*, 6. [\[CrossRef\]](#) [\[PubMed\]](#)
66. Larner, A. Assessing cognitive screeners with the critical success index. *Prog. Neurol. Psychiatry* **2021**, *25*, 33–37. [\[CrossRef\]](#)
67. Nahm, F. Receiver operating characteristic curve: overview and practical use for clinicians. *Korean J. Anesthesiol.* **2022**, *75*, 25–36. [\[CrossRef\]](#)
68. Spirito, P.; Bellone, P.; Harris, K.M.; Bernabò, P.; Bruzzi, P.; Maron, B.J. Magnitude of Left Ventricular Hypertrophy and Risk of Sudden Death in Hypertrophic Cardiomyopathy. *N. Engl. J. Med.* **2000**, *342*, 1778–1785. [\[CrossRef\]](#)
69. Sossalla, S.; Vollmann, D. Arrhythmia-Induced Cardiomyopathy. *Dtsch. Ärzteblatt Int.* **2018**, *115*, 335. [\[CrossRef\]](#) [\[PubMed\]](#)
70. Sun, B.; Wang, L.; Guo, W.; Chen, S.; Ma, Y.; Wang, D. New treatment methods for myocardial infarction. *Front. Cardiovasc. Med.* **2023**, *10*, 1251669. [\[CrossRef\]](#) [\[PubMed\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.