Int. j. inf. tecnol.
https://doi.org/10.1007/s41870-023-01641-9

ORIGINAL RESEARCH

Optimized multi-stage sifting approach for ECG arrhythmia classification with shallow machine learning models

Piyush Mahajan[1] · Amit Kaul[1]

Received: 26 July 2023 / Accepted: 26 October 2023

Abstract Early diagnosis of illness is critical for timely initiation of treatment and ultimately curing of the patient. This is especially important for diseases in which fatality rate is high like heart ailments. The approach suggested in this work utilizes the efficacy of different feature sets and shallow machine learning approaches for detection of various classes of arrhythmia. In this work a multi-stage sifting approach for arrhythmia detection has been suggested. For classification of arrhythmia, eleven number of shallow machine learning models have been studied. The arrhythmia detection performance was compared using various metrics such as accuracy, precision, recall (sensitivity), specificity, and F1-Scores (F-measure). To further improve the classification models, optimized weighting was applied on top three performing classifier models. Among the four different optimizers evaluated, the Whale Optimization Algorithm (WOA) and Particle Swarm Optimizer (PSO) emerged as the top performers. The proposed method exhibited substantial improvements compared to existing models, with an average increase of 3.1% in overall accuracy. Additionally, all other parameters such as precision, recall, sensitivity, and F1-Scores showed an average improvement of around 8%.

Keywords ECG · Arrhythmia · Machine Learning (ML) · Support Vector Machine (SVM) · K Nearest Neighbour (KNN)

✉ Piyush Mahajan
piyush.mahajan@nith.ac.in

Amit Kaul
amitkaul@nith.ac.in

[1] Electrical Engineering Department, NIT Hamirpur, Hamirpur, HP 177005, India

1 Introduction

Globally, it has been reported by various agencies that Cardiovascular Diseases (CVDs) result in annual fatality of around 18 million people around the world [1]. A large percentage of these deaths take place in developing nations like India, which account for around one fifth of these fatalities. Through a network of diverse biological signals, our body continually attempts to forewarn us of potential prospective illnesses. Electrocardiogram (ECG) is the most significant signal for determining a person's heart condition. ECG is the measurement of the heart's electrical activity. A Normal Sinus Rhythm (NSR) ECG wave with typical time domain features is depicted in Fig. 1. The signal voltage ranges between 0.5 and 5mV. Due to low voltage levels, the ECG signal is vulnerable to even minute disturbances. The frequency components of an individual's ECG signal span from 0.05 to 100 Hz. ECG is recorded with electrodes placed at various locations on the patient's limbs and chest, resulting in a variety of lead combinations [2]. An ECG records the heart's electrical activity and provides information about rhythm and conditions. It includes characteristic waves (deviations from mean electrical current), intervals (time between ECG events), segments (distance between ECG points at baseline), and complexes (group of waves). The most commonly measured intervals are PR, QRS, QT, and RR as shown in Fig. 1. The main components of an ECG are P wave (atrial depolarization), QRS complex (ventricular depolarization), and T wave (ventricular repolarization). The U wave that may follow the T wave, indicates Purkinje fiber repolarization [3]. ECG signals include a wealth of critical information that may be indicative of cardiovascular diseases. The rhythm of one's heartbeats is one example of this characteristic. Arrhythmia is the medical name for a cardiac rhythm disorder. In many cases, it is a reliable predictor of

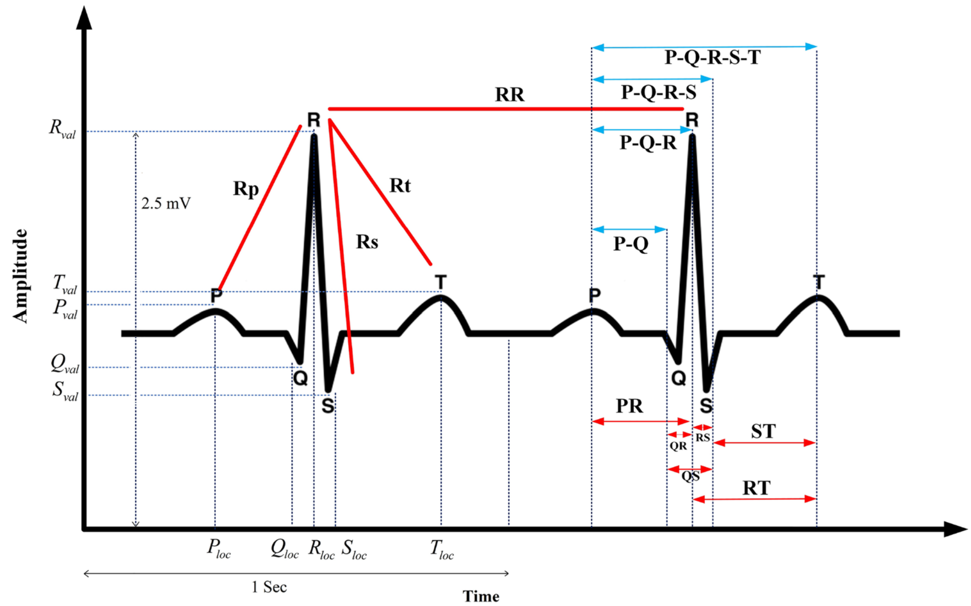

Fig. 1 Typical ECG signal along with time domain features

heart disease. It frequently indicates the health of the heart. It can potentially result in cardiac failure in some cases. The mortality rate in case of CVDs can be decreased if the abnormalities can be detected and disease is diagnosed at an early stage. Cardiologists can correctly interpret an ECG waveform in order to diagnose diseases. However, there is acute shortage of cardiologists in India, with only 1 cardiologist available for every 2,55,000 individuals [4]. Moreover, in countries like Japan which has large geriatric population, there is requirement for continuous monitoring of health state of individuals. This issue can be solved by exploiting the advancements made in domain of Artificial Intelligence (AI) and Information Technology (IT). Automatic tools for diagnosis of cardiovascular diseases can be developed which can assist the physician and can also provide an early warning mechanism.

Following the amassing of datasets, the first step is to preprocess them and eliminate any unwanted noise. The feature extraction and classifiers are used for abnormality detection. A number of researchers have developed automated procedures for arrhythmia detection through ECG signal via Machine Learning (ML) techniques. Some of these contributions have been discussed in the next section.

1.1 Related work

The recent progress in AI has enabled automatic arrhythmia detection using various ML techniques.

Several studies have focused on classifying input ECG signals into normal and abnormal categories. Celin and Vasanth [5] used ML techniques such as Support Vector Machine (SVM), AdaBoost, Artificial Neural Network (ANN), and Naive Bayes Classifier, while Singh and Kaul [6] achieved considerable classification accuracy using multiclass SVM. Subramanian and Prakash [7] employed an SVM classifier on pre-processed ECG data to reduce noise and achieved an overall accuracy of 91%. Sireesha et al. [8] used ML methods such as Decision Tree (DT), Gaussian Naive Bayes, and SVM to classify ECG signals as normal or abnormal, achieving an accuracy of 98.2% with the use of K-fold cross-validation.

Other studies have focused on detecting specific types of arrhythmias. Rao and Martis [9] used several ML techniques to detect Atrial Fibrillation (AFIB) from ECG signals, achieving a highest accuracy of 85.1% using Decision Tree classifier. Jothiramalingam et al. [10] focused on Left Ventricular Hypertrophy (LVH) and used SVM, KNN, and an Ensemble of Bagged trees classifiers to compare findings with a neural network classifier. Kumari et al. [11] used Discrete Wavelet Transform (DWT) to create over one hundred and ninety features for an SVM classifier and achieved an accuracy of 95.92% in classifying NSR, Congestive Heart Failure (CHF), and cardiac arrhythmia categories. Anupuram Pradeepkumar and Amit Kaul [12] presented an ensemble of diverse features to improve the overall accuracy of ML models.

Other researchers have proposed novel methods to improve the performance of ML models. Pham et al. [13] created a Computer-Aided Decision Support System (CDSS) capable of recognizing three types of arrhythmias: AFIB, Atrial Flutter, and Ventricular Fibrillation, achieving an accuracy of 98% using the Random Forest (RF) classifier. Aziz et al. [14] investigated the effect of utilising Fractional Fourier Transform (FrFT) and Two-Event Related Moving Average (TERMA) for feature extraction and achieved an accuracy of 92.2% using SVM. Nurmaini et al. [15]

proposed a Deep Neural Network (DNN) structure using Principal Component Analysis (PCA) for feature selection and showed that DNNs are 2.3% more sensitive than SVM.

Recently, several studies employed deep learning approaches to improve the accuracy of their models. Jiang et al. [16] developed a Multi-Modal Neural Network (MMNN) using Denoising Auto-Encoding (DAE) and Convolutional Neural Network (CNN) for feature extraction, achieving a highest accuracy of 97.3%. Sraritih et al. [17] created a Computer-Assisted Diagnostics System (CADS) using SVM, RF, KNN, and an ensemble of these, achieving an accuracy of 83%. Harrane and Belkhiri [18] utilized CNN and Long Short-Term Memory (LSTM) network for the categorization of ECG arrhythmia, achieving a classification accuracy of 99.2%. Marsa Gholamian and colleagues [19] suggested a method based on Modified Local Binary Pattern (MLBP). LBP operators were utilised to extract the features. Researchers are adopting CNN in order to alleviate the need of feature selection. Saad Irfan[20] and associates introduced a unique architecture for Deep Learning (DL) that combines many networks to build a single robust model. Mohammed Hammad [21] and his team of experts refined DNN by employing a powerful collection of features. The optimal feature sets are recommended to be calculated using a genetic algorithm (GA) procedure. Additionally, in recent research, Goswami et al. [22] proposed an ensemble-based classification model for ECG signal classification. This model achieved an accuracy of 99.98% in classifying arrhythmia datasets. Furthermore, a study by Sengupta et al. [23] presented a novel approach to detecting bradycardia from ECG signals, achieving an accuracy of around 95%. Moreover, Rao et al. [24] developed a method for the detection of AFIB based on Stockwell transformation and CNN. The proposed approach achieved an overall accuracy (Acc) of 99.54% and improved AFIB detection. In addition, Gupta and Avasthi [25] explored person identification using ECG and LSTM, achieving results superior to state-of-the-art methods. Sharma and Sunkaria [26] successfully detected and delineated the U-wave in ECG signals, achieving high accuracy and precision using both the RF algorithm and KNN techniques, making significant advancements in U-wave detection.

These innovative approaches demonstrate the wide range of research efforts in the field of ECG signal analysis and its applications in healthcare.

1.2 Motivation and contribution

This research paper addresses important research gaps in automatic arrhythmia detection. First, it introduces a new multi-level classification method that includes hierarchical classification and gradual refinement of the classification results. This approach increases classification accuracy and fills a gap in the limited study of multi-stage methods. Second, the research examines the effectiveness of ensemble classifiers and overcomes the limitations of using a single classifier. By combining multiple classifiers, the proposed method improves accuracy. In addition, the study uses optimization techniques such as WOA and PSO to further improve performance. The optimization process reduces the number of false negatives (FN), which is essential in medical applications, and improves the accuracy, precision, recall, sensitivity and F1-Score of the ensemble classifiers. Overall, this research contributes to advanced arrhythmia detection by addressing research gaps, improving classification accuracy, and effectively addressing the problem of FN results, thereby providing valuable insights for medical applications.

While there have been various studies on automatic classification of arrhythmia using ECG signals, most of these studies have focused on using a single classifier or a single feature extraction method. In contrast, the suggested method in this work employs multiple classifiers and evaluates the performance of various feature extraction methods. This multi-stage sifting approach can potentially improve the accuracy of arrhythmia classification.

1.3 Paper organization

The rest of the paper is organized as follows: Sect. 2 outlines the methodology, covering the datasets used (2.1), data preprocessing and feature extraction techniques (2.2), the optimizers employed (2.3), and the proposed approach (2.4). Section 3 presents the results and includes a discussion of the findings, while section 4 concludes the paper by summarizing the contributions and outlining potential avenues for future research.

2 Methodology adopted

In this work, investigations are carried out to enhance the arrhythmia detection performance by using combination of features and gradually sieving out different classes through different ML models. It is essential to train the models with appropriate data to aid clinicians and researchers in the precise detection of various CVDs. The identification of arrhythmia can assist in the diagnosis of a wide range of abnormalities. The suggested approach can identify five different forms of arrhythmia. Figure 2 displays the methodology flowchart, illustrating the systematic approach employed in this study. All the components outlined in the flowchart will be elaborated upon in the upcoming subsections, providing a detailed insight into the methodology's structure and decision-making process.

Following is a brief description of the physionet ECG datasets that were used for this study.

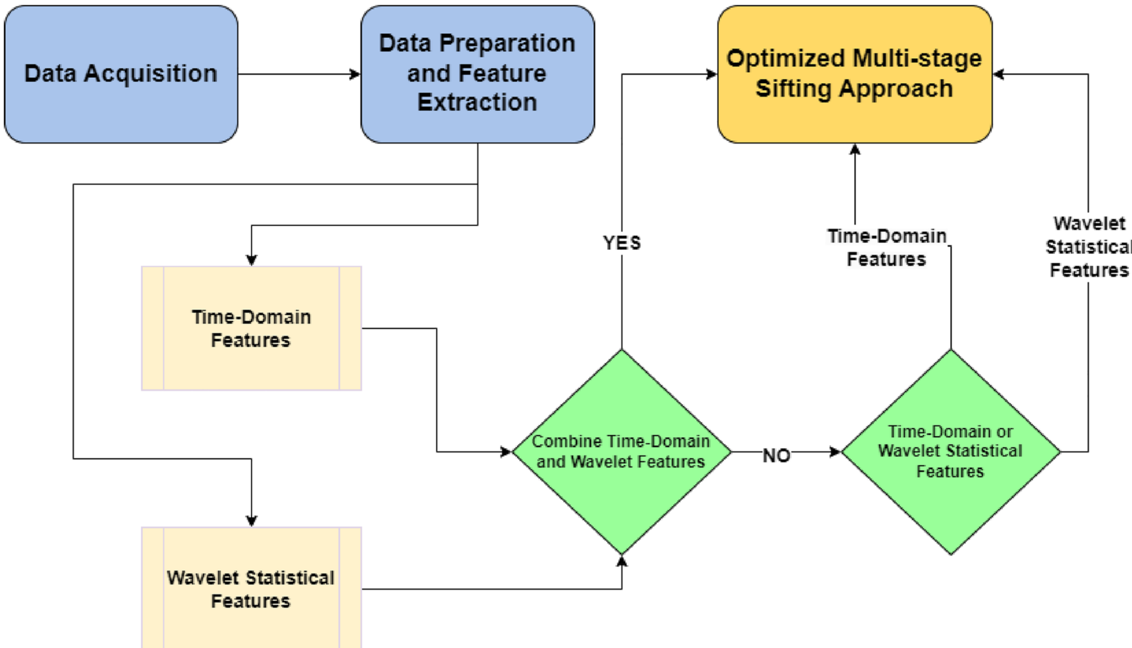

Fig. 2 Methodology flowchart for the proposed multistage approach for ECG arrhythmia classification

2.1 Datasets used

The performance of the suggested approach was evaluated on six PhysioNet datasets. These signals were divided into 10-s-long waveforms. A summary of the datasets used is provided in Table 1. ECG waveforms of NSR and various

Table 1 ECG waveform datasets and their properties

S.No	Name	No. of records	Sampling frequency	No. of waveforms used
1	The MIT-BIH Normal Sinus Rhythm Database (Normal)	18 recordings with a length of up to 12 h are available.	128 Hz	441
2	The MIT-BIH Atrial Fibrillation Database (AFIB)	25 recordings with a maximum duration of 10 hours	250 Hz	274
3	The BIDMC Congestive Heart Failure Database (CHF)	15 recordings with a maximum duration of 20 hours	250 Hz	178
4	The MIT-BIH Malignant Ventricular Ectopy Database (VE)	22 half-hour-long recordings	250 Hz	319
5	The MIT-BIH Supraventricular Arrhythmia Database (SUP)	Each of the 78 recordings has a maximum duration of 12 h.	128 Hz	307
6	Creighton University Ventricular Tachyarrhythmia Database (CUVT)	35 eight-minute ECG recordings	250 Hz	272

arrhythmia classes, such as AFIB, CHF, Malignant Ventricular Ectopy (VE), Supraventricular Arrhythmia (SUP), and Ventricular Tachyarrhythmia (CUVT), are depicted in Fig. 3. Each type of arrhythmia is illustrated in accordance with its respective image. AFIB is characterized by chaotic and rapid firing of action potentials in the pulmonary veins or atrium, resulting in an increased atrial rate of 400–600 beats per minute. CHF results from inadequate pumping of blood by the heart, leading to fluid accumulation in the lungs and other parts of the body. VE or premature ventricular contractions (PVCs) form clusters such as bigeminy, trigeminy, couplets, and triplets. SUP affects the upper chambers of the heart and is characterized by a rapid or irregular heartbeat, which may occur intermittently. CUVT is a potentially life-threatening condition, and identifying the exact QRS morphology can be challenging based on the origin of the arrhythmia.

2.2 Data preprocessing and feature extraction

This research aims to improve the accuracy with which various arrhythmia can be identified. The FN rate must be lowered to acceptable levels, since it is morally indefensible to falsely label a sick person as healthy. While the PhysioNet's

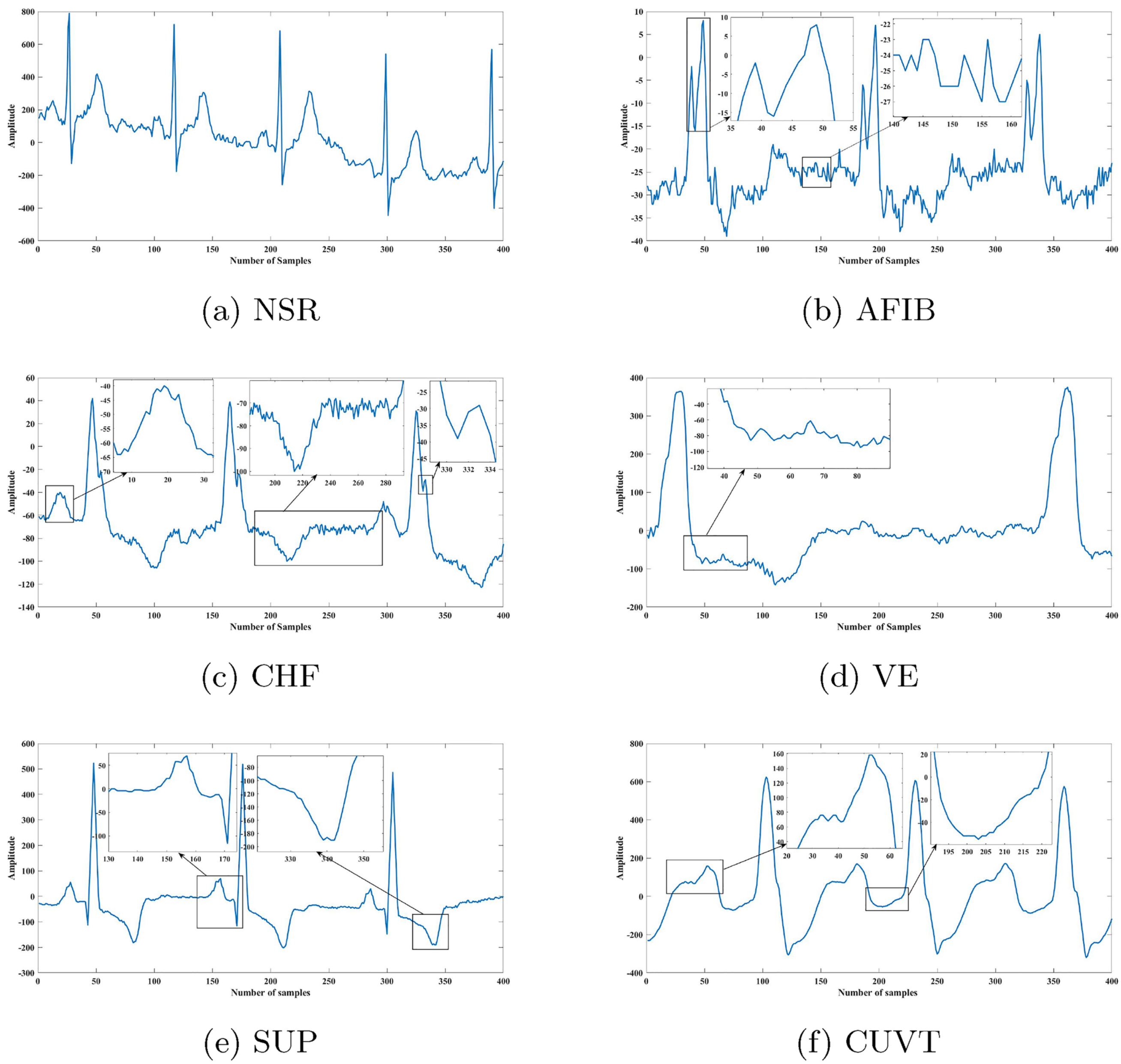

Fig. 3 Different ECG waveforms in time domain

ECG waveforms are somewhat filtered, considerable baseline drift was still present in several cases. Both baseline drift and powerline interference were present in the input ECG waveforms. Low-frequency noise between 0.5 and 0.6 Hz is what experts call baseline wander. It was filtered out using a high-pass digital filter with a cut-off frequency of 0.5 to 0.6 Hz. A digital notch filter with a 50 Hz cut-off frequency was used to get rid of powerline interference (50 Hz noise from mains supply). The pre-processed data was used to extract time-domain and wavelet statistical features.

2.2.1 *Time-domain features*

Time-Domain characteristics represent the time, amplitude variations, and angles between each heartbeat's observed locations [27]. Three peaks (P, R, and T), two valleys (Q and S), and six onsets and offsets make up these points. Heart illness causes modifications to the ECG waveform. These modifications may involve alterations in shape, amplitudes, duration, or angle. Therefore, 50 time-domain characteristics were derived to detect these alterations which are listed in Table 2 and displayed in Fig. 1. The QRS complex is the most crucial and effective component, while P and T are regarded the least reliable and suspicious components. Noise can affect P waves, but they usually do not affect T waves, which are dynamic in their location and always change with the heart rate.

2.2.2 *Wavelet statistical features*

High classification accuracy in an ML model relies on utilizing properly extracted relevant features from the input data. Information that is common knowledge in one domain may not be available in another. Due to the non-stationary nature of ECG waveforms, the Time-Frequency Representation (TFR) approach may provide useful insights on the cause of a certain ailment. As a result, the Wavelet transform was used to yield coefficients and a frequency vector. Though experiments were performed with different wavelets, Symlet 4 was finally selected because it mimics an ECG signal and so produced the most accurate findings. Several statistical functions, including mean absolute value (MAV), average power (AVP), standard deviation (SD), variance (Var), autocorrelation (Xcorr), Kurtosis, skewness, signal-to-noise ratio (SNR), peak-to-rms ratio, total harmonic distortion (THD), maximum frequency, minimum frequency were employed to

Table 2 Descriptions of time-domain features

S. no	Feature	Description	S. no	Feature	Description
1	R_loc	Location of R peaks	26	ST	Distance between S valley and T peak
2	P_loc	Location of P peaks	27	STavg	Average of ST
3	Q_loc	Location of Q valleys	28	RQavg	Average of RQ
4	S_loc	Location of S valleys	29	RS	Distance between R peak and S valley
5	T_loc	Location of T peaks	30	RSavg	Average of RS
6	R_val	Amplitude of R peaks	31	RT	Distance between R and T peaks
7	P_val	Amplitude of P peaks	32	RTavg	Average of RT
8	Q_val	Amplitude of Q valleys	33	QS	Distance between Q and S valleys
9	S_val	Amplitude of S valleys	34	QSavg	Average of QS
10	T_val	Amplitude of T peaks	35	Rp	Length of RP segment
11	$Rloc_avg$	Average of R_loc	36	Rq	Length of RQ segment
12	$Ploc_avg$	Average of P_loc	37	Rs	Length of RS segment
13	$Qloc_avg$	Average of Q_loc	38	Rt	Length of RT segment
14	$Sloc_avg$	Average of S_loc	39	P-Q	Length of PQ segment
15	$Tloc_avg$	Average of T_loc	40	P-Q-R	Length of PQR segment
16	$Rval_avg$	Average of R_val	41	P-Q-R-S	Length of PQRS segment
17	$Pval_avg$	Average of P_val	42	P-Q-R-S-T	Length of PQRST segment
18	$Qval_avg$	Average of Q_val	43	Rpavg	Average of Rp
19	$Sval_avg$	Average of S_val	44	Rqavg	Average of Rq
20	$Tval_avg$	Average of T_val	45	Rsavg	Average of Rs
21	RR	Distance between R peaks	46	Rtavg	Average of Rt
22	RRavg	Average of RR	47	P-Qavg	Average of P-Q
23	RP	Distance between R and P peaks	48	P-Q-Ravg	Average of P-Q-R
24	RPavg	Average of RP	49	P-Q-R-Savg	Average of P-Q-R-S
25	RQ	Distance between R peak and Q valley	50	P-Q-R-S-Tavg	Average of P-Q-R-S-T

extract the characteristics from wavelet coefficients. MAV is obtained using Eq. 1.

$$MAV = \frac{1}{s}\sum_{i=1}^{s}|x_i| \tag{1}$$

Similarly the expressions for calculating AVP, Var and SD are given in Eqs. 2, 3 and 4 respectively. Here x_i is the i th sample of ECG wavelet coefficients obtained from a segment and μ is the mean of the coefficients.

$$AVP = \frac{1}{s}\sum_{i=1}^{s}|x_i|^2 \tag{2}$$

$$VAR = \frac{1}{s-1}\sum_{i=1}^{s}(x_i-\mu)^2 \tag{3}$$

$$SD = \sqrt{\frac{1}{s-1}\sum_{i=1}^{s}\left(x_i-\mu\right)^2} \tag{4}$$

THD is represented by Eq. 5, where V_n is the RMS value of the n th harmonic voltage and V_1 is the RMS value of the fundamental component.

$$THD = \frac{\sqrt{V_1^2+V_2^2+V_3^2\cdots V_n^2}}{V_1} \tag{5}$$

Matlab 2022 has been used to compute a total of 20 wavelet and statistical features listed in table 3.

2.3 Optimizers used

Optimizers, also known as optimization algorithms (OA) or search algorithms, are computational approaches that find the optimum solution to a given problem. These issues frequently entail determining the best settings for parameters or variables that minimise or maximise an objective function.

Optimizers' major goal is to automate the process of finding the optimal solution, reducing the amount of human work and time required for manual exploration of the solution space. Optimizers are widely utilised in a wide range of fields, including engineering, ML, operations research, economics, and many more where complicated problems must be addressed effectively. A brief review of four optimizers used here is given below:

- **Particle Swarm Optimization:** PSO is a population-based optimization method inspired by bird flocking or fish schooling behaviour [28]. It keeps a population of particles that explore the search space by modifying their placements based on their personal experience and the best global solution found thus far.
- **Genetic Algorithm:** GA is a population-based optimization method inspired by the natural selection process [29]. It entails keeping a population of candidate solutions (individuals) and applying selection, crossover, and mutation operators iteratively to develop new solutions. The approach is similar to the idea of survival of the fittest.
- **Whale Optimization Algorithm:** WOA is a population-based optimization algorithm inspired by humpback whale social behaviour. It employs a series of mathematical equations to replicate whale hunting behaviour. The program keeps track of a population of search agents (whales) and changes their positions using a mix of exploration and exploitation tactics [30].
- **Grey Wolf Optimization (GWO):** GWO is a population-based optimization algorithm inspired by grey wolves' social structure and hunting behaviour. It updates the positions of search agents by simulating the wolf leadership hierarchy. To arrive at the best solution, the algo-

Table 3 Descriptions of wavelet statistical features

S. No	Feature	Description	S. no	Feature	Description
1	f_max	Maximum frequency present in the signal	11	SD_2	Standard deviation of input signal
2	f_min	Minimum frequency present in the signal	12	Xcorr	Auto-correlation of input signal
3	f_mean	Mean of frequency vector	13	Kurtosis	Kurtosis of input signal
4	MAV_1	Mean Absolute value of frequency vector	14	Skewness	Skewness of input signal
5	MAV_2	Mean Absolute value of input signal	15	SNR	Signal-to-noise ratio of input signal
6	AVP_1	Average power of frequency vector	16	P2rms	Peak-to-rms ratio of input signal
7	AVP_2	Average power of input signal	17	THD	Total harmonic distortion
8	VAR_1	Variance of frequency vector	18	MAX	Maximum amplitude of input signal
9	VAR_2	Variance of input signal	19	MIN	Minimum amplitude of input signal
10	SD_1	Standard deviation of frequency vector	20	MEAN	Average amplitude of input signal

rithm goes through exploration and exploitation phases [31].

2.4 Optimized multi-stage sifting approach

As stated earlier, reducing the percentage of FN is a crucial aspect of this research. Features and classifiers are two important components of pattern recognition problems like arrhythmia detection. A multistage sifting mechanism has been adopted in this work. This approach utilizes multiple features and different classifiers in conjunction to enhance the arrhythmia classification performance. The basic philosophy of the approach is to sieve out classes one at a time at every stage. The classifiers are trained to function as 'binary classifiers' with one class pitted against all others. The basic structure of the multistage approach is given in figure 4. The main steps involved in the implementation of multistage sifting approach are listed below:-

1. Train a set of classifiers in multi-class configuration.
2. Arrange the accuracy of all 'N' classes obtained from various classifiers in descending order.
3. Train different classifiers in binary classification mode such that class 'C_1'is one with highest accuracy achieved in step 2 and all other 'N-1' classes clubbed as 'C_2'.
4. Sieve out class 'C_1' and retrain the classifiers again by now taking the class having second highest accuracy in step 2 as 'C_1' and remaining 'N-2' classes as 'C_2'.
5. Repeat step 4, till last two classes are left.

Once the classifiers are trained, they are used in a stage-wise manner in order to improve the classification. To further

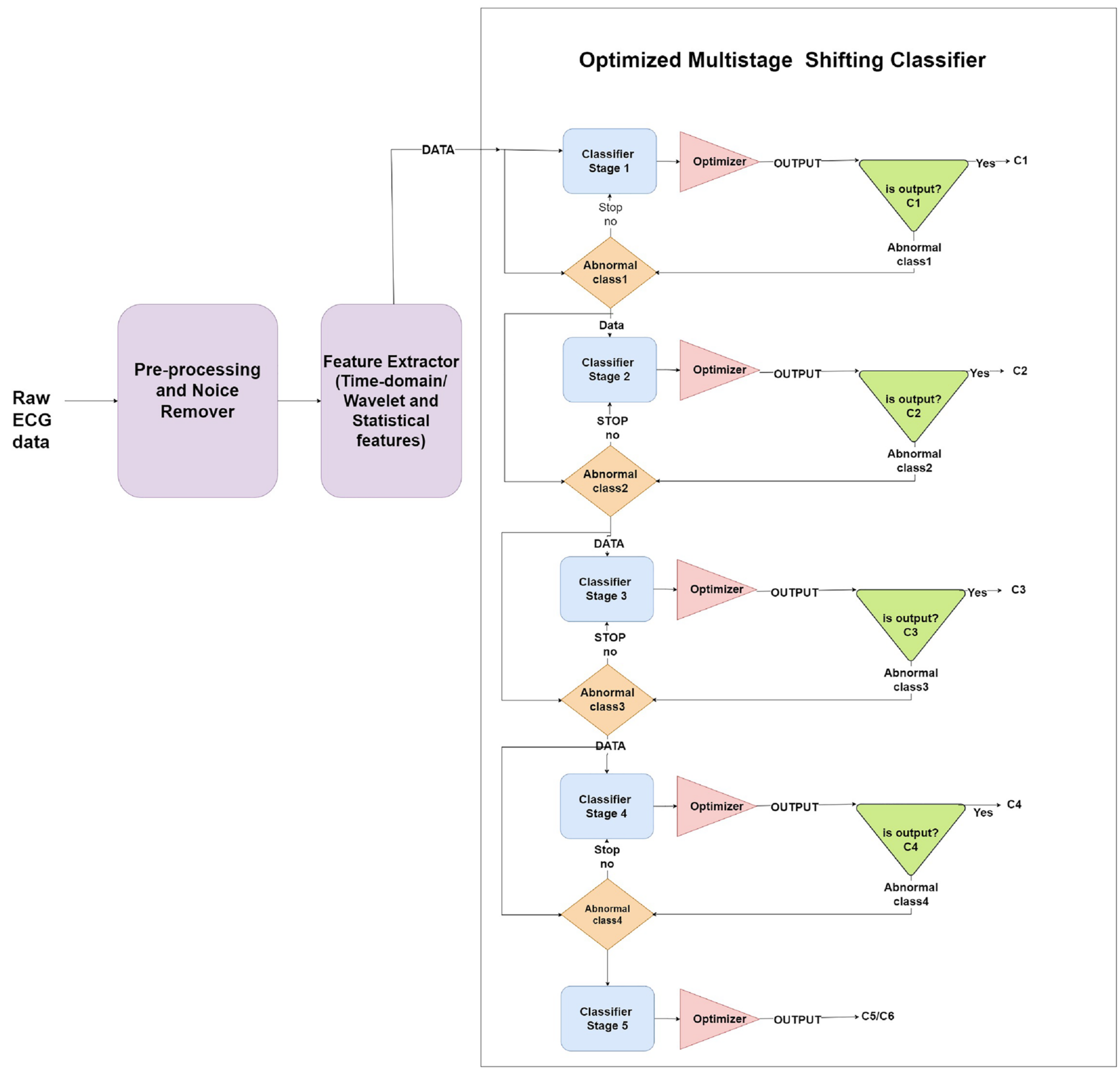

Fig. 4 The proposed multistage approach for ECG arrhythmia classification

enhance the performance, the fusion of classifier is done using four above-mentioned optimization tools. An OF is necessary for optimizers in order to change weights. The objective function used in our approach has been described below.

- **Objective Function:** The OF attempts to reduce the number of FN in the classification system. It is calculated as the weighted total of each model's FN. Each model's weights show its importance in reducing FN. The OF guides the optimisation algorithms to discover an appropriate combination of weights that minimises FN and increases the classification system's performance by optimising these weights. The objective function used can be written mathematically as Eq. 6. Here 'f(X)' is the fitness function representing the objective function to be minimized, 'w_i' is the weight and 'FN_i' is the number of false negatives for the'i_{th}' model.

$$f(X) = \sum w_i * FN_i \tag{6}$$

The findings after implementing the proposed approach are discussed in depth in the upcoming section under "Results and Discussions".

3 Results and discussion

Before the proposed multi-stage sifting process was executed, classifier models were trained to categorise input data into six classes: Normal, AFIB, CHF, CUVT, SUP, and VE. The outcomes of which are shown in Tables 4 and 5. F1-Scores and accuracy indicate how well a model classifies a specific class. This is required for the proposed sifting process to function properly. When the tables are examined closely, it is clear that the precision, recall, and F-measure values for certain misclassified classes, particularly CHF and VE, are relatively low. This shows that the models have difficulty correctly classifying instances belonging to these classes. Normal and SUP classes were found to be the most accurately classified. While F1 scores and accuracy metrics provide an overall evaluation of a model's performance, it is critical to recognise the limitations and weaknesses detected for certain classes. In this situation, the low precision, recall, and F-measure values for the CHF and VE classes show that additional improvement in identifying these diseases is required. The proposed sifting process aims to address these limitations and enhance the classification performance for all classes, including CHF and VE.

The Multi-Stage Sifting Process begins with the classification of input data into normal and abnormal categories. All classes, including AFIB, CHF, CUVT, SUP, and VE, are grouped as abnormal. Many classifiers are capable of obtaining perfect accuracy. As a result, data may now be reliably identified as normal or abnormal. Having isolated the normal class, the trained model is now solely exposed to ECG data indicative of pathological arrhythmia. SUP classification was deemed the most accurate of all the feature sets since it received the highest F1-Score.

In the following stage of the Multi-Stage Sifting Method, abnormal data is separated into two groups: SUP and abnormal. The classification models were fed this data, and their results were compared. Multiple classifiers accomplished perfect accuracy on all types of feature vectors. Therefore, SUP can now be removed from the aberrant data. The second phase involved training models to divide anomalous data into AFIB, CHF, CUVT, and VE. On many occasions, the class CUVT achieved the highest F1-Scores. Consequently, this served as the foundation for the next step.

After analysing the aforementioned results, CUVT is now removed from the abnormal data using the same method as in case of SUP. This information is then given to classifier models, which classify it into the CUVT and abnormal categories. All forms of feature vectors were observed to achieve a high level of classification accuracy. Consequently, by repeating the preceding processes, CUVT has been eliminated from the anomalous data. Again, models were trained to categorise aberrant input data into three categories, namely AFIB, CHF, and VE.

Again, after studying the aforementioned data, it was discovered that the majority of classifiers classified AFIB with the highest F1-Score. Thus, AFIB and abnormal data classes are created from the remaining abnormal data, and then fed into the classification models. As with the prior models, anomalous data is fed in order to categorise it as CHF and VE. After executing the models for both feature sets, it was determined that both sets outperform one another in a variety of circumstances. For certain models and classes, time-domain features performed exceptionally well, as did wavelet statistical features in other instances. These outcomes from a multi-stage sifting procedure are shown in Table 6 and Table 7. The best results were obtained by combining both feature sets. Further investigation showed that the most frequently misclassified classes, CHF and VE, benefited significantly from the combined feature set in terms of precision, recall, and F-measure. The observed improvements in precision, recall, and F-measure for the misclassified classes, namely CHF and VE, can be attributed to a significant decrease in FN. The decrease in FN indicates that the classifier successfully identified more positive instances that were previously missed. As a result, the precision for the CHF class increased by an average of 26.127%, the recall improved by 17.09%, and the F-measure showed a significant gain of 27.99%. Similarly, for the VE class, precision increased by 27.21% and recall improved by

Table 4 Classifier performance on data containing all classes for three distinct feature sets (Part A)

S.No.	Classifier	Parameters	Time Domain Features						Wavelet Statistical Features						All Features Combined					
			AFIB	CHF	CUVT	SUP	VE	Normal	AFIB	CHF	CUVT	SUP	VE	Normal	AFIB	CHF	CUVT	SUP	VE	Normal
1	**Fine Tree**	Accuracy	92.67	94.97	97.77	98.02	93.49	98.02	93.33	94.98	97.28	100.00	92.84	100.00	94.69	96.27	98.26	100.00	94.52	100.00
		Precision	82.07	68.92	89.43	92.21	70.23	97.07	84.40	66.67	86.99	100.00	67.18	100.00	86.76	69.77	91.20	100.00	79.46	100.00
		Recall	86.55	57.30	88.71	92.21	69.70	97.51	86.55	62.92	86.29	100.00	66.67	100.00	90.55	75.95	91.94	100.00	67.42	100.00
		Specificity	94.47	97.96	98.81	98.87	96.40	98.32	95.32	97.51	98.53	100.00	96.03	100.00	95.91	97.69	98.98	100.00	97.86	100.00
		F-measure	84.25	62.58	89.07	92.21	69.96	97.29	85.46	64.74	86.64	100.00	66.92	100.00	88.61	72.73	91.57	100.00	72.95	100.00
2	**Medium Tree**	Accuracy	85.26	92.51	97.69	98.02	89.46	98.02	92.02	94.49	97.20	100.00	92.10	100.00	94.32	95.80	98.27	100.00	93.83	100.00
		Precision	65.79	48.65	89.34	92.21	51.69	97.07	78.90	62.79	87.50	100.00	66.98	100.00	84.11	72.09	91.87	100.00	76.15	100.00
		Recall	72.73	40.45	87.90	92.21	46.21	97.51	88.36	60.67	84.68	100.00	53.79	100.00	92.36	69.66	91.13	100.00	62.88	100.00
		Specificity	88.94	96.63	98.81	98.87	94.74	98.32	93.09	97.16	98.63	100.00	96.77	100.00	94.89	97.87	99.08	100.00	97.60	100.00
		F-measure	69.08	44.17	88.62	92.21	48.80	97.29	83.36	61.71	86.07	100.00	59.66	100.00	88.04	70.86	91.50	100.00	68.88	100.00
3	**Naive Bayes**	Accuracy	90.45	93.42	96.13	98.27	90.04	98.44	92.12	95.41	96.23	99.84	92.21	99.75	92.51	95.39	96.21	99.75	92.92	99.67
		Precision	79.34	54.46	74.84	90.80	55.79	99.07	81.23	65.42	77.46	99.35	72.84	99.55	84.59	68.13	74.07	99.35	73.00	99.32
		Recall	78.18	61.80	93.55	96.10	40.15	96.60	85.30	78.65	88.71	99.35	44.70	99.77	81.82	69.66	96.77	98.70	55.30	99.77
		Specificity	94.04	95.91	96.43	98.59	96.12	99.48	94.15	96.73	97.08	99.91	97.98	99.74	95.64	97.42	96.15	99.91	97.51	99.61
		F-measure	78.75	57.89	83.15	93.38	46.70	97.82	83.22	71.43	82.71	99.35	55.40	99.66	83.18	68.89	83.92	99.02	62.93	99.55
4	**Linear SVM**	Accuracy	89.88	93.09	96.63	99.34	91.36	99.42	93.09	95.31	96.46	100.00	91.52	99.92	92.26	94.65	97.53	100.00	91.36	100.00
		Precision	72.49	54.10	80.29	98.03	66.27	98.87	78.85	67.78	80.92	100.00	71.01	100.00	79.48	65.38	82.64	100.00	64.84	100.00
		Recall	89.09	37.08	88.71	96.75	41.67	99.55	94.91	68.54	85.48	100.00	37.12	99.77	88.73	57.30	95.97	100.00	44.70	100.00
		Specificity	90.11	97.51	97.53	99.72	97.41	99.35	92.55	97.42	97.71	100.00	98.15	100.00	93.30	97.60	97.71	100.00	97.05	100.00
		F-measure	79.93	44.00	84.29	97.39	51.16	99.21	86.14	68.16	83.14	100.00	48.76	99.89	83.85	61.08	88.81	100.00	52.91	100.00
5	**Quadratic SVM**	Accuracy	93.66	95.47	98.02	99.59	94.57	99.59	95.56	96.54	97.53	100.00	94.49	99.92	96.54	97.70	98.77	100.00	95.97	100.00
		Precision	83.00	72.97	88.46	98.69	78.45	99.32	89.89	75.27	86.72	100.00	76.42	100.00	92.06	84.27	90.98	100.00	84.30	100.00
		Recall	90.55	60.67	92.74	98.05	68.94	99.55	90.55	78.65	89.52	100.00	71.21	99.77	92.73	84.27	97.58	100.00	77.27	100.00
		Specificity	94.57	98.22	98.63	99.81	97.69	99.61	97.02	97.96	98.44	100.00	97.32	100.00	97.66	98.76	98.90	100.00	98.25	100.00
		F-measure	86.61	66.26	90.55	98.37	73.39	99.43	90.22	76.92	88.10	100.00	73.73	99.89	92.39	84.27	94.16	100.00	80.63	100.00
6	**Fine KNN**	Accuracy	94.98	97.12	97.78	95.47	95.14	95.64	96.13	96.87	97.78	100.00	95.47	99.75	95.97	97.20	98.44	99.92	95.80	99.67
		Precision	89.05	78.13	90.08	81.53	78.29	94.29	92.22	74.76	90.08	100.00	79.84	100.00	90.94	79.57	94.12	99.35	80.45	99.77
		Recall	88.73	84.27	87.90	83.12	76.52	93.65	90.55	86.52	87.90	100.00	78.03	99.32	91.27	83.15	90.32	100.00	81.06	99.32
		Specificity	96.81	98.13	98.90	97.27	97.41	96.77	97.77	97.69	98.90	100.00	97.60	100.00	97.34	98.31	99.36	99.91	97.60	99.87
		F-measure	88.89	81.08	88.98	82.32	77.39	93.97	91.38	80.21	88.98	100.00	78.93	99.66	91.11	81.32	92.18	99.68	80.75	99.55

Table 5 Classifier performance on data containing all classes for three distinct feature sets (Part B)

S.No.	Classifier	Parameters	Time Domain Features						Wavelet Statistical Features						All Features Combined					
			AFIB	CHF	CUVT	SUP	VE	Normal	AFIB	CHF	CUVT	SUP	VE	Normal	AFIB	CHF	CUVT	SUP	VE	Normal
1	**Weighted KNN**	Accuracy	93.58	95.80	97.20	95.56	94.81	95.56	95.39	96.63	97.28	100.00	94.16	99.75	96.05	98.02	98.11	100.00	96.30	100.00
		Precision	81.47	73.17	89.47	80.49	80.53	95.10	90.11	74.49	85.27	100.00	74.80	100.00	88.74	88.24	89.76	100.00	87.83	100.00
		Recall	92.73	67.42	82.26	85.71	68.94	92.52	89.45	82.02	88.71	100.00	69.70	99.32	94.55	84.27	91.94	100.00	76.52	100.00
		Specificity	93.83	98.05	98.90	96.98	97.97	97.29	97.13	97.78	98.26	100.00	97.14	100.00	96.49	99.11	98.81	100.00	98.71	100.00
		F-measure	86.73	70.18	85.71	83.02	74.29	93.79	89.78	78.07	86.96	100.00	72.16	99.66	91.55	86.21	90.84	100.00	81.78	100.00
2	**Ensemble Bagged Trees**	Accuracy	96.13	97.78	98.35	98.93	97.37	99.09	96.30	96.21	98.19	100.00	95.97	100.00	97.78	98.02	98.93	100.00	97.53	100.00
		Precision	88.78	88.75	90.00	95.48	93.86	98.64	88.85	77.92	89.84	100.00	84.87	100.00	93.97	90.12	91.11	100.00	91.80	100.00
		Recall	94.91	79.78	94.35	96.10	81.06	98.87	95.64	67.42	92.74	100.00	76.52	100.00	96.36	82.02	99.19	100.00	84.85	100.00
		Specificity	96.49	99.20	98.81	99.34	99.35	99.22	96.49	98.49	98.81	100.00	98.34	100.00	98.19	99.29	98.90	100.00	99.08	100.00
		F-measure	91.74	84.02	92.13	95.79	86.99	98.75	92.12	72.29	91.27	100.00	80.48	100.00	95.15	85.88	94.98	100.00	88.19	100.00
3	**Narrow Neural Network**	Accuracy	93.33	94.32	97.45	97.70	92.92	97.78	94.65	96.21	97.04	99.84	94.49	99.67	94.05	96.03	98.35	99.83	92.98	99.59
		Precision	85.93	61.11	88.43	93.75	66.43	96.00	87.23	74.71	84.38	100.00	75.59	99.77	88.01	71.58	91.94	100.00	67.41	99.31
		Recall	84.36	61.80	86.29	87.66	70.45	97.96	89.45	73.03	87.10	98.70	72.73	99.32	85.45	76.40	91.94	98.70	68.94	99.54
		Specificity	95.96	96.89	98.72	99.15	95.66	97.67	96.17	98.05	98.17	100.00	97.14	99.87	96.58	97.59	99.08	100.00	95.92	99.61
		F-measure	85.14	61.45	87.35	90.60	68.38	96.97	88.33	73.86	85.71	99.35	74.13	99.55	86.72	73.91	91.94	99.35	68.16	99.43
4	**Wide Neural Network**	Accuracy	95.80	96.21	96.87	98.11	94.81	98.27	94.73	96.63	97.04	99.75	94.07	99.67	96.38	96.71	98.27	100.00	95.31	99.84
		Precision	91.79	75.29	85.83	92.26	73.47	97.73	89.51	75.00	83.85	100.00	73.08	99.55	92.62	75.79	90.55	100.00	79.07	100.00
		Recall	89.45	71.91	83.06	92.86	81.82	97.51	86.91	80.90	87.90	98.05	71.97	99.55	91.27	80.90	92.74	100.00	77.27	99.55
		Specificity	97.66	98.13	98.44	98.87	96.40	98.71	97.02	97.87	98.08	100.00	96.77	99.74	97.87	97.96	98.90	100.00	97.51	100.00
		F-measure	90.61	73.56	84.43	92.56	77.42	97.62	88.19	77.84	85.83	99.02	72.52	99.55	91.94	78.26	91.63	100.00	78.16	99.77
5	**Trilayered Neural Network**	Accuracy	91.69	93.99	97.04	97.61	93.50	97.86	94.16	95.64	97.53	99.92	93.09	99.75	94.65	94.73	97.70	99.84	93.17	99.67
		Precision	78.62	60.81	84.38	92.52	73.04	96.42	87.23	70.45	89.17	99.35	67.14	100.00	86.97	64.04	87.50	100.00	69.92	99.77
		Recall	86.91	50.56	87.10	88.31	63.64	97.73	86.91	69.66	86.29	100.00	71.21	99.32	89.82	64.04	90.32	98.70	65.15	99.32
		Specificity	93.09	97.42	98.17	98.96	97.14	97.93	96.28	97.69	98.81	99.91	95.75	100.00	96.06	97.16	98.53	100.00	96.58	99.87
		F-measure	82.56	55.21	85.71	90.37	68.02	97.07	87.07	70.06	87.70	99.68	69.12	99.66	88.37	64.04	88.89	99.35	67.45	99.55

Table 6 Classifier performance on data containing all classes for three distinct feature sets after applying Multi-Stage Sifting Process (Part A)

S.No.	Classifier	Parameters	Time Domain Features						Wavelet Statistical Features						All Features Combined					
			AFIB	CHF	CUVT	SUP	VE	Normal	AFIB	CHF	CUVT	SUP	VE	Normal	AFIB	CHF	CUVT	SUP	VE	Normal
1	**Fine Tree**	Accuracy	83.00	83.70	95.30	100.00	83.70	97.86	84.00	90.40	94.80	100.00	90.40	100.00	85.40	92.70	94.50	100.00	92.70	100.00
		Precision	83.00	78.40	91.30	100.00	87.50	97.20	86.80	86.90	86.50	100.00	93.00	100.00	84.10	91.00	90.10	100.00	93.90	100.00
		Recall	86.90	82.00	84.60	100.00	84.80	96.80	84.00	89.80	87.90	100.00	90.90	100.00	90.90	91.00	81.40	100.00	93.90	100.00
		Specificity	78.20	84.80	97.90	100.00	82.00	98.40	84.10	90.90	96.50	100.00	89.80	100.00	78.70	93.90	97.70	100.00	91.00	100.00
		F-measure	85.00	80.20	87.80	100.00	86.10	97.00	85.30	88.30	87.20	100.00	91.90	100.00	87.40	91.00	85.50	100.00	93.90	100.00
2	**Medium Tree**	Accuracy	80.40	83.70	95.30	100.00	83.70	97.86	82.80	90.40	94.80	100.00	90.40	100.00	85.80	92.70	94.50	100.00	92.70	100.00
		Precision	80.20	78.40	91.30	100.00	87.50	97.20	85.10	86.90	86.50	100.00	93.00	100.00	85.40	91.00	90.10	100.00	93.90	100.00
		Recall	85.80	82.00	84.60	100.00	84.80	96.80	83.60	89.80	87.90	100.00	90.90	100.00	89.80	91.00	81.40	100.00	93.90	100.00
		Specificity	73.70	84.80	97.90	100.00	82.00	98.40	81.90	90.90	96.50	100.00	89.80	100.00	80.90	93.90	97.70	100.00	91.00	100.00
		F-measure	82.90	80.20	87.80	100.00	86.10	97.00	84.40	88.30	87.20	100.00	91.90	100.00	87.50	91.00	85.50	100.00	93.90	100.00
3	**Naive Bayes**	Accuracy	76.60	76.00	92.90	99.30	76.00	95.20	81.60	89.50	89.50	100.00	89.50	100.00	81.60	88.20	91.10	100.00	88.20	100.00
		Precision	76.70	64.70	74.30	100.00	89.80	91.50	80.60	83.00	67.20	100.00	95.00	100.00	82.30	81.80	70.10	100.00	93.40	100.00
		Recall	82.90	88.70	98.30	96.00	67.40	95.60	88.00	93.20	92.70	100.00	87.10	100.00	85.00	91.00	96.70	100.00	86.30	100.00
		Specificity	68.70	67.40	91.50	100.00	88.70	94.90	73.70	87.10	88.70	100.00	93.20	100.00	77.30	86.30	89.70	100.00	91.00	100.00
		F-measure	79.70	74.80	84.70	98.30	77.00	93.50	84.10	87.80	77.90	100.00	90.90	100.00	83.70	86.10	81.30	100.00	89.70	100.00
4	**Linear SVM**	Accuracy	76.20	75.50	92.40	99.80	75.50	99.40	79.40	84.60	92.20	100.00	84.60	100.00	81.20	79.60	93.30	100.00	79.60	100.00
		Precision	77.70	69.60	79.30	99.30	79.50	98.60	82.10	75.70	78.70	100.00	92.90	100.00	82.20	76.80	79.00	100.00	81.20	100.00
		Recall	80.00	69.60	83.80	100.00	79.50	99.70	80.30	91.00	83.80	100.00	80.30	100.00	84.30	70.70	91.10	100.00	85.60	100.00
		Specificity	71.40	79.50	94.50	99.80	69.60	99.20	78.20	80.30	94.30	100.00	91.00	100.00	77.30	85.60	93.90	100.00	70.70	100.00
		F-measure	78.80	69.60	98.50	99.60	79.50	99.20	81.20	82.60	81.20	100.00	86.10	100.00	83.30	73.60	84.60	100.00	83.30	100.00
5	**Quadratic SVM**	Accuracy	84.40	91.80	96.10	99.80	91.80	99.20	89.30	92.70	94.60	100.00	92.70	100.00	92.10	90.40	96.20	100.00	90.40	100.00
		Precision	86.30	89.80	89.00	99.30	93.10	98.40	89.60	90.10	83.20	100.00	94.60	100.00	93.30	89.50	87.90	100.00	91.10	100.00
		Recall	85.40	89.80	91.90	100.00	93.10	99.50	91.20	92.10	91.90	100.00	93.10	100.00	92.30	86.50	94.30	100.00	93.10	100.00
		Specificity	83.20	93.10	97.10	99.80	89.80	99.00	86.80	93.10	95.30	100.00	92.10	100.00	91.80	93.10	96.70	100.00	86.50	100.00
		F-measure	85.90	89.80	90.40	99.60	93.10	98.90	90.40	91.10	87.30	100.00	93.80	100.00	92.80	88.00	91.00	100.00	92.10	100.00
6	***Fine KNN***	*Accuracy*	*89.70*	*93.20*	*97.40*	*99.80*	*93.20*	*97.20*	*93.30*	*96.80*	*97.50*	*100.00*	*96.80*	*100.00*	*95.90*	*97.20*	98.20	*100.00*	*97.20*	*100.00*
		Precision	*92.10*	*92.00*	*95.00*	*99.30*	*93.90*	*97.60*	*94.10*	*97.70*	*93.60*	*100.00*	*95.50*	*100.00*	*96.70*	*95.60*	*95.20*	*100.00*	*98.40*	*100.00*
		Recall	*89.00*	*91.00*	*91.90*	*100.00*	*94.60*	*94.50*	*93.80*	*96.90*	*94.30*	*100.00*	*96.60*	*99.84*	*96.00*	*97.70*	*95.90*	*100.00*	*96.90*	*99.81*
		Specificity	*90.40*	*94.60*	*98.70*	*99.80*	*91.00*	*98.70*	*92.70*	*96.60*	*98.30*	*100.00*	*96.90*	*100.00*	*95.90*	*96.90*	*98.70*	*100.00*	*97.70*	*100.00*
		F-measure	*90.50*	*91.50*	*93.40*	*99.60*	*94.30*	*96.00*	*93.90*	*97.30*	*93.90*	*100.00*	*96.00*	*100.00*	*96.30*	*96.60*	*95.50*	*100.00*	*97.70*	*100.00*

The italic values signify the best performing models

Table 7 Classifier performance on data containing all classes for three distinct feature sets after applying Multi-Stage Sifting Process (Part B)

S.No.	Classifier	Parameters	Time Domain Features						Wavelet Statistical Features						All Features Combined					
			AFIB	CHF	CUVT	SUP	VE	Normal	AFIB	CHF	CUVT	SUP	VE	Normal	AFIB	CHF	CUVT	SUP	VE	Normal
1	***Weighted KNN***	*Accuracy*	*90.10*	*90.90*	*95.40*	*99.60*	*90.90*	*97.60*	*93.90*	*94.50*	*96.90*	*100.00*	*94.50*	*100.00*	*93.70*	*95.40*	*97.20*	*100.00*	*95.40*	*100.00*
		Precision	*88.40*	*87.90*	*93.60*	*99.30*	*93.00*	*97.60*	*93.50*	*89.60*	*91.30*	*100.00*	*98.30*	*100.00*	*91.70*	*91.50*	*92.80*	*100.00*	*98.40*	*100.00*
		Recall	*94.50*	*89.80*	*83.00*	*98.70*	*91.60*	*95.60*	*95.60*	*97.70*	*93.50*	*100.00*	*92.40*	*99.69*	*97.40*	*97.70*	*93.50*	*100.00*	*93.90*	*100.00*
		Specificity	*84.60*	*91.60*	*98.50*	*99.80*	*89.80*	*98.70*	*91.80*	*92.40*	*97.70*	*100.00*	*97.70*	*100.00*	*89.10*	*93.90*	*98.10*	*100.00*	*97.70*	*100.00*
		F-measure	*91.30*	*88.80*	*88.00*	*99.00*	*92.30*	*96.60*	*94.60*	*93.50*	*92.40*	*100.00*	*95.30*	*100.00*	*94.50*	*94.50*	*93.10*	*100.00*	*96.10*	*100.00*
2	***Ensemble Bagged Trees***	*Accuracy*	*94.10*	*91.40*	*98.30*	*100.00*	*91.40*	*98.80*	*92.10*	*94.10*	*96.90*	*100.00*	*94.10*	*100.00*	*96.10*	*96.80*	*97.70*	*100.00*	*96.80*	*100.00*
		Precision	*95.50*	*91.60*	*93.80*	*100.00*	*91.20*	*98.60*	*92.40*	*90.40*	*90.60*	*100.00*	*96.80*	*100.00*	*96.70*	*94.50*	*94.30*	*100.00*	*98.40*	*100.00*
		Recall	*93.80*	*86.50*	*98.30*	*100.00*	*94.60*	*98.10*	*93.40*	*95.50*	*94.30*	*100.00*	*93.10*	*100.00*	*96.30*	*97.70*	*94.30*	*100.00*	*96.20*	*100.00*
		Specificity	*94.50*	*94.60*	*98.30*	*100.00*	*86.50*	*99.20*	*90.40*	*93.10*	*97.50*	*100.00*	*95.50*	*100.00*	*95.90*	*96.20*	*98.50*	*100.00*	*97.70*	*100.00*
		F-measure	*94.60*	*89.00*	*96.00*	*100.00*	*92.90*	*98.40*	*92.90*	*92.80*	*92.40*	*100.00*	*94.90*	*100.00*	*96.50*	*96.10*	*94.30*	*100.00*	*97.30*	*100.00*
3	**Narrow Neural Network**	Accuracy	86.20	87.70	95.30	99.70	87.70	98.00	88.90	88.60	95.90	99.62	88.60	100.00	90.70	86.40	95.60	100.00	86.40	99.29
		Precision	88.40	86.90	88.60	98.70	88.30	97.00	89.20	84.70	90.20	99.21	91.40	100.00	90.90	86.40	87.50	99.05	86.40	99.79
		Recall	86.50	82.00	87.90	100.00	91.60	97.50	90.90	87.60	89.50	97.15	89.30	99.78	91.20	78.60	91.10	99.82	91.60	100.00
		Specificity	85.90	91.60	97.10	99.60	82.00	98.30	86.40	89.30	97.50	100.00	87.60	100.00	90.00	91.60	96.70	100.00	78.60	99.19
		F-measure	87.50	84.30	88.20	99.30	89.90	97.20	90.00	86.10	89.80	98.97	90.40	100.00	91.60	82.30	89.30	99.12	88.90	99.13
4	**Wide Neural Network**	Accuracy	89.50	90.00	96.60	99.80	90.00	98.20	88.30	88.60	95.40	100.00	88.60	100.00	92.30	89.50	95.90	100.00	89.50	100.00
		Precision	89.60	89.40	92.50	99.30	90.40	97.50	89.40	85.50	89.30	100.00	90.80	100.00	93.00	90.20	87.20	100.00	89.20	100.00
		Recall	91.60	85.30	90.30	100.00	93.10	97.70	89.40	86.50	87.90	99.12	90.10	100.00	93.00	83.10	93.50	99.42	93.90	100.00
		Specificity	86.80	93.10	98.10	99.80	85.30	98.50	86.80	90.10	97.30	100.00	86.50	100.00	91.40	93.90	96.50	100.00	83.10	100.00
		F-measure	90.60	87.30	91.40	99.60	91.70	97.60	89.40	86.00	88.60	100.00	90.40	100.00	93.00	86.50	90.20	100.00	91.50	100.00
5	**Trilayered Neural Network**	Accuracy	87.20	85.50	95.80	99.70	85.50	98.60	88.30	92.70	95.10	100.00	92.70	100.00	90.90	86.80	96.10	100.00	86.80	100.00
		Precision	89.20	85.10	92.20	98.70	85.70	98.10	89.40	91.90	88.50	100.00	93.20	100.00	92.20	87.50	88.40	99.98	86.50	100.00
		Recall	87.60	77.50	86.20	100.00	90.90	97.90	89.40	89.80	87.00	98.80	94.60	99.18	91.20	78.60	92.70	100.00	92.40	100.00
		Specificity	86.80	90.90	98.10	99.60	77.50	98.90	86.40	94.60	97.10	100.00	89.80	100.00	90.40	92.40	96.90	100.00	78.60	100.00
		F-measure	88.40	81.10	89.10	99.30	88.20	98.00	89.40	90.90	87.80	99.24	93.90	100.00	91.70	82.80	90.50	100.00	89.30	100.00

The italic values signify the best performing models

37.13%. These notable improvements in the evaluation metrics suggest that the classifier's ability to correctly identify positive instances has been enhanced, primarily by reducing the occurrences of FN. The reduction in FN played a crucial role in achieving higher precision, recall, and overall classification performance, highlighting the importance of addressing FN to improve the performance of the classifier. The proposed multistage sifting approach improves the F1-Scores obtained by the majority of misclassified classes. F1-Score is often considered a better measurement than accuracy for classification models because accuracy can be misleading in certain situations, especially when dealing with imbalanced datasets. Accuracy is defined as the number of correctly classified instances divided by the total number of instances in the dataset. While this metric is useful in many cases, it can be misleading when the dataset has a class imbalance, which occurs when one class has significantly more or fewer examples than the other. In such cases, a model that always predicts the majority class will have a high accuracy, but may not be useful in practice as it fails to correctly identify the minority class. The F1-Score, on the other hand, is a weighted average of precision and recall. Although there were considerably fewer false positives (FP), accuracy suffered as a result; therefore, optimizers were utilised for optimised weightage at the conclusion of each stage to increase accuracy. Therefore, the suggested Multistage Sifting method was carried out again on three best performing classifiers highlighted in Tables 6 and 7. These are Fine KNN, Weighted KNN and Ensemble Bagged Trees classifiers. It was noted that each stage's overall accuracy improved as well. The OF used by optimizers to weight the models according to the number of FN. PSO, GA, WOA, and GOA were the four optimizers used, and the results are shown in the Table 8. PSO and WOA were found to perform better than GA and GWA by obtaining nearly 100 percent accuracy at all stages. These findings underscore the effectiveness of the proposed strategy in enhancing the accuracy of arrhythmia classification models. They also highlight the potential to contribute to the development of more precise and reliable diagnostic tools for detecting arrhythmias in clinical settings. Additionally, the Multistage Sifting method introduced in this study showcases competitive performance compared to the techniques outlined in Table 9, which represent established approaches by various researchers in the field along with their results.

4 Conclusion and future scope

The proposed multi-stage sifting approach has shown promising results in improving the classification performance for various arrhythmia classes. By combining time-domain and wavelet statistical features, the classifier achieved notable improvements in precision, recall, and

Table 8 Results obtained after applying PSO, GAO,WOA and GWO techniques

S.No.	Optimization Technique	Parameters	All Features Combined					
			AFIB	CHF	CUVT	SUP	VE	Normal
1	**Particle Swarm Optimization (PSO)**	Accuracy	99.19	99.09	100.00	100.00	99.09	100.00
		Precision	100.00	100.00	100.00	100.00	98.50	100.00
		Recall	98.54	97.75	100.00	100.00	100.00	100.00
		Specificity	100.00	100.00	100.00	100.00	97.75	100.00
		F-measure	99.26	98.86	100.00	100.00	99.24	100.00
2	**Genetic Algorithm Optimization (GAO)**	Accuracy	94.35	96.83	98.39	100.00	96.83	94.96
		Precision	94.27	94.57	95.97	100.00	98.45	95.29
		Recall	95.64	97.75	95.97	100.00	96.21	95.64
		Specificity	92.76	96.21	98.99	100.00	97.75	94.12
		F-measure	94.95	96.13	95.97	100.00	97.32	95.46
3	**Whale Optimization Algorithm (WAO)**	Accuracy	99.19	99.09	100.00	100.00	99.09	100.00
		Precision	100.00	100.00	100.00	100.00	98.50	100.00
		Recall	98.54	97.75	100.00	100.00	100.00	100.00
		Specificity	100.00	100.00	100.00	100.00	97.75	100.00
		F-measure	99.26	98.86	100.00	100.00	99.24	100.00
4	**Grey Wolf Optimizer (GWO)**	Accuracy	95.56	97.29	98.55	100.00	97.29	100.00
		Precision	95.67	95.60	96.75	100.00	98.46	100.00
		Recall	96.36	97.75	95.97	100.00	96.97	100.00
		Specificity	94.57	96.97	99.19	100.00	97.75	100.00
		F-measure	96.01	96.67	96.36	100.00	97.71	100.00

Table 9 Comparison of Arrhythmia Detection Approaches and Performance Metrics

Paper	Approach	Models Used	Evaluation Metrics	Results
Celin and Vasanth[5]	Various ML techniques	SVM, AdaBoost, ANN, Naive Bayes	Classification Accuracy	SVM: 87.5%, ANN:94%, AdaBoost:93%, Naive Bayes:99.97
Singh and Kaul[6]	Multiclass SVM	SVM	Classification Accuracy	Two-stage SVM:98.3
Subramanian and Prakash[7]	SVM on pre-processed data	SVM	Overall Accuracy	SVM:91%
Sireesha et al.[8]	ML methods with K-fold cross-validation	DT, Gaussian Naive Bayes, SVC	Accuracy	DT:98.2%, Gaussian Naive Bayes: 97.3, SVC:80.3
Rao and Martis[9]	Atrial Fibrillation Detection	DT	Accuracy	DT:85.1%
Jothiramalingam et al.[10]	Left Ventricular Hypertrophy Detection	SVM, KNN, Ensemble	Classification Accuracy	SVM:86.6%, KNN:84.4%, Ensemble:93.3%
Kumari et al.[11]	SVM Classifier with DWT	SVM	Performance Accuracy	SVM:95.92%
Anupuram Pradeepkumar and Amit Kaul[12]	Ensemble of Features	ANN	Accuracy, Precision, Recall, Specificity, F1-Score	Overall Accuracy: 99.55
Pham et al.[13]	Random Forest Classifier	Random Forest	Accuracy, Sensitivity, Specificity	Accuracy:98.2%, Sensitivity:98.1%, Specificity:99.4%
Aziz et al.[14]	FrFT and TERMA for Feature Extraction	SVM	Sensitivity, Positive Predictivity	Sensitivity:99.83, Positive Predictivity:99.90%
Nurmaini et al.[15]	Deep Neural Network (DNN)	DNN with PCA	Accuracy, Sensitivity, Specificity	Accuracy:99.76, Sensitivity:91.80%, Specificity:99.78
Jiang et al.[16]	Multi-Module Neural Network System with Data Balancing Measures	DAE, CNN	Overall Accuracy	Overall Accuracy:96.6%
Sraritih et al.[17]	Supervised Learning with Inter-Patient Paradigm	SVM, RF, KNN, Ensemble	Overall Accuracy	Overall Accuracy:83%
Harrane and Belkhiri[18]	CNN and LSTM Network	CNN, LSTM	Training and Test Accuracy	Training Accuracy:99.9%, Test Accuracy:98.60
Marsa Gholamian et al.[19]	Modified Local Binary Pattern (MLBP)	MLBP	Classification Accuracy	Classification Accuracy:99.76%
Saad Irfan[20]	Unique Architecture for Overall Accuracy: 99.55Deep Learning	Combination of Many Networks	Model Performance	Overall Accuracy:93.33%
Mohammed Hammad[21]	Deep Neural Network (DNN)	DNN with Feature Selection(GA Procedure)	Accuracy, F1-Score	Average Accuracy:94%, F1-Score:95.3%
Proposed Approach	*Optimized Multi-Stage Sifting Approach*	*Combination of many ML models with optimizers.*	*Accuracy, Precision, Recall, Specificity, F1-Score*	*Accuracy:99.56%, Precision:99.75%, Recall:99.38%, Specificity:99.62%, F1-Score:99.56%*

F-measure, particularly for the misclassified CHF and VE classes. These improvements can be attributed to a significant decrease in FN, indicating that the classifier successfully identified more positive instances that were not detected using conventional classification approach. The reduction in FN played a crucial role in achieving higher precision, recall, and overall classification performance. While the proposed approach demonstrates its effectiveness, it is essential to acknowledge certain limitations and potential drawbacks. The multi-stage sifting process increases the complexity of the classification model, which may require additional computational resources and longer processing times. Future research should explore the use of alternative feature sets, classifiers, or ensemble techniques to further improve the performance and robustness of the classification model. Addressing the identified limitations and exploring future research avenues will further refine and advance the field of arrhythmia classification, ultimately benefiting patients and healthcare professionals alike.

Data availability The datasets utilized in the research are publicly available to everyone over the internet.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Cardiovascular diseases (CVDs). https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds). Accessed 30 Sep 2022
2. AlGhatrif M, Lindsay J (2012) A brief review: history to understand fundamentals of electrocardiography. J Community Hosp Intern Med Perspect 2(1):14383
3. Sharma LD, Sunkaria RK (2021) Detection and delineation of the enigmatic u-wave in an electrocardiogram. Int J Inf Technol 13:2525–2532
4. Kulkarni VT, Ross JS, Wang Y, Nallamothu BK, Spertus JA, Normand S-LT, Masoudi FA, Krumholz HM (2013) Regional density of cardiologists and rates of mortality for acute myocardial infarction and heart failure. Circ Cardiovasc Qual Outcomes 6(3):352–359
5. Celin S, Vasanth K (2018) Ecg signal classification using various machine learning techniques. J Med Syst 42(12):1–11
6. Singh A, Kaul A (2022) Analysis of ecg signal using machine learning approaches for detecting different heart abnormalities. In: 2022 2nd International Conference on emerging frontiers in electrical and electronic technologies (ICEFEET), pp 1–5. IEEE
7. Subramanian K, Prakash NK (2020) Machine learning based cardiac arrhythmia detection from ecg signal. In: 2020 Third International Conference on smart systems and inventive technology (ICSSIT), pp 1137–1141. IEEE
8. Moturi S, Vemuru S, Rao ST (2021) Ecg based decision support system for clinical management using machine learning techniques. IOP Conf Ser Mater Sci Eng 1085:012016
9. K SRS, Martis RJ (2020) Machine learning based decision support system for atrial fibrillation detection using electrocardiogram. In: 2020 IEEE International Conference on distributed computing, VLSI, Electrical Circuits and Robotics (DISCOVER), pp 263–266
10. Jothiramalingam R, Jude A, Patan R, Ramachandran M, Duraisamy JH, Gandomi AH (2021) Machine learning-based left ventricular hypertrophy detection using multi-lead ecg signal. Neural Comput Appl 33(9):4445–4455
11. Usha Kumari C, Sampath Dakshina Murthy A, Lakshmi Prasanna B, Pala Prasad Reddy M, Kumar Panigrahy A (2021) An automated detection of heart arrhythmias using machine learning technique: Svm. Materials Today Proceedings 45, 1393–1398. In: International Conference on Advances in Materials Research—2019
12. Pradeepkumar A, Kaul A (2022) Ecg arrhythmia classification using ensemble of features. In: 2022 2nd Asian Conference on innovation in technology (ASIANCON), pp 1–6. IEEE
13. Sree V, Mapes J, Dua S, Lih OS, Koh JE, Ciaccio EJ, Acharya UR et al (2021) A novel machine learning framework for automated detection of arrhythmias in ecg segments. J Ambient Intell Humaniz Comput 12(11):10145–10162
14. Aziz S, Ahmed S, Alouini M-S (2021) Ecg-based machine-learning algorithms for heartbeat classification. Sci Rep 11(1):1–14
15. Nurmaini S, Tutuko B, Rachmatullah MN, Darmawahyuni A, Masdung F (2021) Machine learning techniques with low-dimensional feature extraction for improving the generalizability of cardiac arrhythmia. IAENG Int J Comput Sci 48(2):369–378
16. Jiang J, Zhang H, Pi D, Dai C (2019) A novel multi-module neural network system for imbalanced heartbeats classification. Expert Syst Appl X 1:100003
17. Sraitih M, Jabrane Y, Hajjam El Hassani A (2021) An automated system for ecg arrhythmia detection using machine learning techniques. J Clin Med 10(22):5450
18. Harrane S, Belkhiri M (2021) Classification of ecg heartbeats using deep neural networks. Res Biomed Eng 37(4):737–747
19. Gholamian M, Yazdi M, Joursaraei A, Zeraatkar E (2021) An ecg classification based on modified local binary patterns: a novel approach. Res Biomed Eng 37(4):617–630
20. Irfan S, Anjum N, Althobaiti T, Alotaibi AA, Siddiqui AB, Ramzan N (2022) Heartbeat classification and arrhythmia detection using a multi-model deep-learning technique. Sensors 22(15):5606
21. Hammad M, Iliyasu AM, Subasi A, Ho ES, Abd El-Latif AA (2020) A multitier deep learning model for arrhythmia detection. IEEE Trans Instrum Meas 70:1–9
22. Goswami AD, Bhavekar GS, Chafle PV (2023) Electrocardiogram signal classification using vggnet: a neural network based classification model. Int J Inf Technol 15(1):119–128
23. Sengupta S, Mayya V, Kamath SS (2022) Detection of bradycardia from electrocardiogram signals using feature extraction and snapshot ensembling. Int J Inf Technol 14(6):3235–3244
24. Rao BM, Kumar A, Bachwani N, Marwaha P (2023) Detection of atrial fibrillation based on stockwell transformation using convolutional neural networks. Int J Inf Technol 15(4):1937–1947
25. Gupta PK, Avasthi V (2023) Person identification using electrocardiogram and deep long short term memory. Int J Inf Technol 15(3):1709–1717
26. Sharma LD, Sunkaria RK (2021) Detection and delineation of the enigmatic u-wave in an electrocardiogram. Int J Inf Technol 13:2525–2532
27. Bassiouni MM, El-Dahshan E-SA, Khalefa W, Salem AM (2018) Intelligent hybrid approaches for human ecg signals identification. SIViP 12(5):941–949
28. Kennedy J, Eberhart R (1995) Particle swarm optimization. In: Proceedings of ICNN'95-International Conference on neural networks, vol. 4, pp 1942–19484
29. Holland JH (1984) Genetic algorithms and adaptation. In: Selfridge OG, Rissland EL, Arbib MA (eds) Adaptive control of ill-defined systems. NATO Conference Series, vol 16. Springer, Boston, pp 317–333
30. Mirjalili S, Lewis A (2016) The whale optimization algorithm. Adv Eng Softw 95:51–67
31. Mirjalili S, Mirjalili SM, Lewis A (2014) Grey wolf optimizer. Adv Eng Softw 69:46–61