Original Research Paper

Cardiovascular Diseases Identification Using Wavelet Optimization and Modify Cuckoo Search Algorithm

[1]Banit Negi, [3]Abhilekh Bartwal, [2]Agya Ram Verma, [3]Surjeet Singh Patel, [4]Yatendra Kumar, [1]Abhishek Gupta, [1]Vivek Kumar Tamta and [5]Priti Kumari

[1]Department of Computer Science and Engineering, G.B. Pant Insitute of Engineering and Technology, India
[2]Department of Electronic and Communication Engineering, G.B. Pant Insitute of Engineering and Technology, India
[3]Department of Electrical Engineering, G.B. Pant Insitute of Engineering and Technology, India
[4]Department of Electronics and Instrumentation Engineering, FET, MJP Rohilkhand University, Bareilly, India
[5]Department of Computer Science and Engineering, NIT, Patna, India

Abstract: The problem the author aims to solve is the extraction of discriminatory features in ECG (Electrocardiogram) signals for classification purposes. The scope of the work is to propose a new method for building wavelets that best reflect the discriminatory capacity of ECG signals. The approach involves optimizing the wavelets specifically for the classification function under consideration. To address the problem, the author proposes a novel method for creating wavelets that optimize discriminatory feature extraction in ECG signals. The approach utilizes the poly-phase demonstration of the filter bank and incorporates the Modified Cuckoo Search (MCS) algorithm to project the problem context. The experiments are conducted using the MIT/BIH arrhythmia database to evaluate the performance of the proposed method against existing state-of-the-art techniques. The Support Vector Machine (SVM) classifier is used to demonstrate the effectiveness, precision, and robustness of the projected strategy on standard wavelets like Daubechies and Symlet. The extent of the author's work involves developing a new method for wavelet construction to enhance discriminatory feature extraction in ECG signals. Important variables controlled in the study include the choice of wavelet parameters, the application of the MCS algorithm, and the evaluation of results against standard wavelet-based classification methods. The experiment results demonstrate the superiority of the proposed wavelet construction method over traditional wavelets like Daubechies and Symlet for ECG signal classification. The new wavelet shows improved discriminatory capacity, leading to higher classification precision and accuracy. The findings imply that optimizing wavelets for the classification of ECG signals using the Modified Cuckoo Search algorithm can significantly enhance discriminatory feature extraction and improve classification precision. The results are potentially generalizable and can be applied to various ECG signal classification tasks, contributing to advancements in medical diagnostics and monitoring. Background: The extraction of discriminatory features from ECG signals is a critical task in the field of medical signal processing. Accurate classification of these features plays a crucial role in diagnosing various cardiac arrhythmias and abnormalities. Wavelet-based techniques have shown promise in this domain, but further optimizations are necessary to achieve higher classification precision and robustness. Current Status in the Field: The field of ECG signal classification continues to evolve, with researchers exploring various approaches to improve discriminatory feature extraction. Wavelet-based methods have gained popularity due to their ability to capture both frequency and time-domain information, but fine-tuning the

Science Publications

wavelets for specific classification tasks remains an on-going area of research. Study and Analysis: The experiments were conducted using the MIT/BIH arrhythmia database, which contains a diverse set of ECG signals. The proposed wavelet construction method was evaluated against existing standard wavelets using the Support Vector Machine (SVM) classifier. The Modified Cuckoo Search algorithm was used to optimize the wavelets for better discriminatory capacity. The results were then statistically analysed to demonstrate the effectiveness of the proposed approach.

Keywords: DWT, MCS, SVM, and ECG

Introduction

Bio-medical signal processing is always a pertinent field of research. One of the most important biomedical signals is Electrocardiogram (ECG) signal which depicts the electric cardiovascular cycle obtained by placing surface anodes in the body (Sörnmo and Laguna, 2006). The ECG signals investigation will give clinicians helpful insights regarding the condition of the cardiac patient (Sörnmo and Laguna, 2006; Chandra *et al.*, 2018). One of the major causes of death related to heart/cardiovascular diseases is arrhythmias, which can be detected by analyzing ECG signals. Detection of beat plays an important role in the analysis of ECG waveform. ECG waveform can be divided into normal and abnormal rhythm also known as arrhythmia. Almost 17.9 million people have died due to arrhythmia (Pandey *et al.*, 2020). Therefore, several studies have been reported to detect the beat for diagnosis of cardiovascular diseases using several techniques (Gupta *et al.*, 2022). Wavelet is one of the most popular and widely used techniques to detect beet from an ECG signal (Gupta *et al.*, 2022). Over the two decades, wavelets have accumulated an expanding enthusiasm for some signal handling and detection applications, having been the basis for various programmed and rapid arrhythmia-finding devices (Daqrouq *et al.*, 2022). The wavelets have their time recurrence portrayal of the signal as the most fascinating component. This allows in-depth knowledge of a signal at various scales and frequencies and has been powerful in both ECG and grouping (Gupta *et al.*, 2022; Daqrouq *et al.*, 2022; Toulni *et al.*, 2021).

In view of this, a few significant works can be found in the feeling of the ECG signal grouping that mirrors the subject of this study. Ince *et al.* proposed an extraction strategy, utilizing an interpretation invariant dyadic wavelet to separate morphological data efficiently from ECG information (Ince *et al.*, 2009). Further, in a study, dyadic wavelet is used to portray the ECG signal with the utilization of computerized signal handling add-on cards to diminish their high computational expenses (Sahambi *et al.*, 1997). Li *et al.*, have utilized wavelet coefficients to separate ordinary and PVC beats (Li *et al.*, 1995). Wavelet

transform can also use for Fetel ECG (FECG) analysis, which is very important to analyze the health of the baby and mother (Hao *et al.*, 2022). Khamene and Negahdaripour have introduced a wavelet transform-based method to extract the fetal ECG from the composite abdominal signal. Here, the detection of the singularities was obtained from the composite abdominal signal, using the modulus maxima in the wavelet domain (Khamene and Negahdaripour, 2000). To find the sharp variety of purposes of ECG, the neighborhood furthest reaches of the particles are utilized at various scales. The calculation proposed recognizes the QRS, the T-wave lastly the P wave complex. Khamene and Negahdaripour projected an answer dependent on the places of the ECG signs' single focuses (higher pinnacles) (Khamene and Negahdaripour, 2000). Their strategy endeavors, both in the composite stomach signal, to recognize the single purposes of the mother's and the fetal ECGs. All exploration happens in the changed ECG signal space of the wavelet.

The Principal Component Analysis (PCA) is a predominant procedure used for feature extraction for heartbeat classification in Gupta and Mittal (2021), the R-R interval is extracted using the Difference Operation Method (DOM) and PCA is employed for feature selection. The classification is done using Fuzzy Logic (FL) and Fisher's Linear Discriminant Analysis (LDA). The classification of ischemic and non-ischemic beats is carried out by the Least-squares Support Vector Machines (LSSVM) based Fuzzy Genetic Algorithm (FGA) (Gupta and Mittal, 2021). The principal and Independent Component Analysis (ICA) is combined with Genetic Algorithm (GA) for better results. The evolution of metaheuristic techniques happened with the usage of GA-based optimization in classifiers. This method also corresponded to labeling known and missing data. The validation was efficient enough as it labeled the training data and also helped in separating the invalidated samples (Ramkumar *et al.*, 2021). The change of classifier namely the Multilayer Perceptron (MP) along with GA was more efficient in detecting cardiac arrhythmias. The extraction of R-R intervals was done by using Symlet which also helped in the reduction of uncertain features (Kumari and Kumar, 2015). Inan *et al.* (2006) projected an approach for

grouping the beats of an enormous data set utilizing a wavelet and time function. The fourth phase of a dyadian wavelet was seen as profoundly fruitful in recognizing customary and PVC from different beats, alongside the pre/post-RR-stretch proportion (Inan *et al.*, 2006). In Khandoker *et al.* (2008), the contributions of a Support Vector Machine (SVM) arrangement to personality obstructive rest apnea disorder were utilized for attributes extricated from progressive wavelet coefficient rates after the wavelet decay of RR span and ECG inferred breath (EDR) from R rushes of QRS amplitudes. A significant inquiry was presented in which wavelet is used. The reaction was, no hypothetical reaction is conceivable at present, so a correlation between the impacts of various waves ought to be made experimentally (Senhadji *et al.*, 1995). Several other work has been done for beat detection of ECG signal (De Chazal and Reilly, 2006; Bazi and Melgani, 2007; Melgani and Bazi, 2008, Stone, 1974). All the works referenced above-utilized wavelet for general sign handling and examination. Nonetheless, it is believed that wavelets enhanced for this specific issue ought to be designed to build effectiveness in the ECG data analysis. ECG signal is a non-sinusoidal signal, therefore the use of non-sinusoidal signal processing approaches is always preferred. In Vaidyanathan (2006), different aspects to work with multirate signal processing are given.

The aim of this study is to propose a wavelet plan strategy guided by the exactness of arrangement process results. In light of the perplexing connection between the wavelet and the classifier accuracy, this study utilizes a stochastic structure technique dependent on Modified Cuckoo Search (MCS) advancement, which has demonstrated equipped for reacting adequately to the issues raised by a few applications (Jayaraman and Sultana, 2019; Goyal *et al.*, 2016; Sharma *et al.*, 2021; Dao, 2020; Kaya, 2018). The polyphase portrayal of the Discrete Wavelet Change (DWT), is utilized in the proposed procedure. This portrayal empowers a wavelet channel bank to be produced by methods for a lot of precise boundaries and consequently, the ECG signal classification wavelet plan issue to be defined as the issue to gauge such boundaries so that the classifier is amplified. The grouping technique utilized in this study is the cutting-edge SVM characterization perceived for its high limit of age. This study uses MIT/BIH database (Moody and Mark, 2001) to evaluate the performance of the proposed technique. The general audits of wavelets are given in segment 2.

Wavelets

The wavelet change is a direct activity that breaks down a sign into segments that show up at various scales (Chandra *et al.*, 2018). Wavelet capacities $\psi(t)$ are characterized in a space of quantifiable capacities that are total and square-integrable, i.e.:

$$\int_{-\infty}^{+\infty} |\psi(t)| \, dt < \infty \tag{1}$$

$$\int_{-\infty}^{+\infty} |\psi(t)|^2 \, dt < \infty \tag{2}$$

In such a space, they ought to fulfill states of zero mean and square standard one (Toulni *et al.*, 2021):

$$\int_{-\infty}^{+\infty} |\psi(t)| \, dt = 0 \tag{3}$$

$$\int_{-\infty}^{+\infty} |\psi(t)|^2 \, dt = 1 \tag{4}$$

The wavelet change of a capacity $f(t) \, \varepsilon L^2(R)$ at scale an and position τ is given by Toulni *et al.* (2021):

$$Wf(a, \tau) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f(t) \left| \psi * \left(\frac{t - \tau}{a} \right) \right| dt \tag{5}$$

Here asterisk * denotes the complex conjugation.

The signal investigation is done in Eq. (5) where $f(t)$ is convolved with extended/widened duplicates of the mother wavelet $\psi(t)$. The change creates minute subtleties for $f(t)$ when the wavelet is contracted for a <1. The coarse view for the signal is accomplished, for a >1 where the wavelet expands. On the off chance that the scaled boundary $a = 2^j$ with $j \epsilon Z$, Z is a whole number set, at that point the wavelet is known as a dyadic wavelet (Sahambi *et al.*, 1997). The wavelet change is applied on work in ceaseless time whereas in discrete time on vectors. In constant time, we decide the wavelet coefficients by means of assessing the essential in (5). While, in discrete time, the coefficients are found by means of passing a vector ($x(n)$, n number) through the bank of two channels, where one is a low-pass and the other is a high-pass. A total and intriguing portrayal of the DWT channel coefficients with reduced help was given by Daubechies in Li *et al.* (1995). Nevertheless, by and large, since looking for an ideal wavelet is an issue subordinate theme, DWT design can take a few structures. For this situation, a creative path by Sherlock and Monro (De Chazal and Reilly, 2006) was created to compute the coefficients of a channel bank. Vaidyanathan (2006); De Chazal and Reilly (2006) proposed a polyphase strategy whose premise is factorization. Their calculation permits determining any orthonormal impeccable reproduction limited drive reaction (FIR) channel of subjective length. The accompanying strategy quickly depicted it. The low-pass channel coefficients in the Z space are given by De Chazal and Reilly (2006):

$$H_0(Z) = \sum_{i=0}^{2N-1} h_i \, z^{-1} \tag{6}$$

and thus:

$$H_0(Z) = \sum_{i=0}^{N-1} h_{2i} \, z^{-2i} + z^{-1} \sum_{i=0}^{N-1} h_{2i+1} z^{-2i} \tag{7}$$

In (7), H_0 deteriorates into even and odd forces of z. Vaidyanathan proposed the accompanying factorization of the polyphase lattice (Melgani and Bazi, 2008):

$$H_p(z) = \begin{pmatrix} H_{00}(z) H_{01}(z) \\ H_{10}(z) H_{12}(z) \end{pmatrix} = \begin{pmatrix} c_0 s_0 \\ -s_0 c_0 \end{pmatrix} \prod_{i=1}^{N-1} \begin{pmatrix} 1 & 0 \\ 0 & z^{-1} \end{pmatrix} \begin{pmatrix} c_i \, s_i \\ s_i \, c_i \end{pmatrix} \tag{8}$$

where:

$$H_{00}(z) = \sum_{i=0}^{N-1} h_{2i} \, z^{-2i} \tag{9}$$

and:

$$H_{01}(z) = \sum_{i=0}^{N-1} h_{2i+1} \, z^{-2i} \tag{10}$$

$H_{00}(z)$ and $H_{01}(z)$ speak to the polyphase segments of the low-pass channel, while $H_{10}(z)$ and $H_{11}(z)$ are those of the high-pass channel. The coefficients c_i and s_i are figured as follows: $c_i = \cos(\theta_i)$ and $s_i = \sin(\theta_i)$. Sherlock and Monro built up another plan by reworking the factorization in a recursive structure (De Chazal and Reilly, 2006):

$$H_p^{(k+1)}(z) = H_p^k(z) \begin{pmatrix} 1 & 0 \\ 0 & z^{-1} \end{pmatrix} \begin{pmatrix} c_k s_k \\ -s_k c_k \end{pmatrix} \tag{11}$$

with, $k = 1, 2,\ldots, N$ and $H_p^{(a)} = \begin{pmatrix} c_0 s_0 \\ -s_0 c_0 \end{pmatrix}$. The superscript (k) alludes to channels of length $2k$. This structure prompts the accompanying recursive formulae for the even-numbered channel constants:

$$\begin{cases} h_0^{(k+1)} = c_k h_0^{(k)} \\ h_{2i}^{(k+1)} = c_k h_{2i}^{(k)} - s_k h_{2i-1}^{(k)} \; for \, i = 1, 2, 3 \ldots\ldots k-1 \\ h_{2k}^{(k+1)} = -s_k h_{2i-1}^{(k)} \end{cases} \tag{12}$$

With $h_0^{(1)} = c_0$ and $h_1^{(1)}$ s_0. The formulations for the odd coefficients are given by:

$$\begin{cases} h_1^{(k+1)} = s_k h_1^{(k)} \\ h_{2i+1}^{(k+1)} = s_k h_{2i}^{(k)} + c_k h_{2i-1}^{(k)} \; for \, i = 1, 2, 3 \ldots\ldots k-1 \\ h_{2k+1}^{(k+1)} = c_k h_{2k-1}^{(k)} \end{cases} \tag{13}$$

Equations (12-13) precise the low-pass constants $\{h_0, h_1, .., h_{2N-1}\}$ as far as N free picked rakish boundaries $\{\theta_0, \theta_1, \theta_2, \ldots \ldots \theta_{N-1}\}$ whose qualities are in the span $[0, 2\pi)$. The high-pass channel coefficients are found by rotating flip development, that Senhadji *et al.* (1995):

$$g_i = (-1)^{i+1} h_{2N-1-I} \tag{14}$$

The $2N$ low-pass channel constants alongside the comparing high-pass channel constants can be caused by N-free boundaries. In this manner, we can locate the design of an ideal DWT as an improvement issue in the RN space of the rakish boundaries $\theta_i's$.

Modify the Cuckoo Search Algorithm

The foundation of the CS algorithm is on the obligatory brood parasite behavior, together with the flight behavior L'evy of fruit flies and some birds. In L'evy, it is animals and birds who search for food in a random or almost random way and mostly go on a random path, as the next step is focused on the present location and the probability of movement to the next point. The behavior was used for CS optimization (Jayaraman and Sultana, 2019). Every egg is a solution in CS algorithms, whereas the Cuckoo egg is a new way of operating (Jayaraman and Sultana, 2019; Goyal *et al.*, 2016). The ultimate goal is to use a new and potentially improved approach to boost poor nest solutions instead of one egg each. While the existence is defined by several algorithms that optimize them, the CS algorithm is chosen in this analysis as a L'evy flight is best suited for a wide range of optimization procedures in finding a long-step search space. The three superlative rules on which CS relies are (i) At a time the single egg is laid by each cuckoo which is dumped in the selected random nest, (ii) It will be carried on to the next generations by the best nest of good egg quality (solutions), (iii) The possible number of host nests is set and the host will detect an alien egg with probability $p_a \varepsilon (0, 1)$. In a particular scenario, the host bird chooses to abandon the nest and shifts to an entirely new location to create a new nest or the egg is lob.

A L'evy flight is performed by the equation by Goyal *et al.* (2016):

$$X_i(t+1) = X_i(t) + \lambda * L'evy(\alpha) \tag{15}$$

where, λ is the step size and is related to the size of the optimization problem, multiplication ($*$), and $L'evy(\alpha)$ is L'evy flight distribution (u) which is defined as:

$$L'evy(\alpha) \approx u = t - \lambda \;\; 1 \leq \alpha \leq 3 \tag{16}$$

The CS algorithm is modified to get the MCS algorithm for improved optimization. The value of λ declines in the first adjustment, as the number of Generation (G) increases to find the quest. Therefore, in the MCS technique, the modified L'evy flight step size (λ_m) is denoted by:

$$\lambda m = \frac{\lambda}{\sqrt{G}} \tag{17}$$

The second alteration consists in linking information sharing between the egg's intention of reducing convergence. By choosing a portion of the eggs that are most appropriate, a top group is produced a second egg in this category is randomly taken for each of the top eggs and a new egg is then produced on the line of these top two eggs.

Separation along this line (at which the new egg is found) is determined utilizing the converse of the brilliant proportion $\phi = 1+\sqrt{5}/2$ to such an extent that it is nearer to the egg with the best wellness. On the off chance that, on the off chance that the two eggs have similar wellness, at that point, another egg is produced at the midpoint. Quite possibly in this progression, a similar egg is gotten once more. In such a case, a neighborhood L'evy flight search is performed from the haphazardly picked home with step size $\lambda_m = \lambda/G^{\wedge 2}$. There are two boundaries for example the part of homes to be deserted (p_{a1}) and the division of homes to make up the top homes ($p_{a2} = 1-p_{a1}$) (Bazi and Melgani, 2007). These parameters are adjusted in the MCS algorithm to achieve better λ_m. A brief overview of the steps that must be taken in designing the program algorithm for this MCS to construct a low pass filter prototype based on the above theory (Jayaraman and Sultana, 2019; Goyal *et al*., 2016; Sharma *et al*., 2021; Dao, 2020; Kaya, 2018):

Stride 1: Specify prototype filter design specifications like pass-band edge frequency (ω_p), stop-band edge frequency (ω_s), and order of filter (N)

Stride 2: Set the nest number. Nest's just various solutions. It is called $N/2$, in this problem. Set a discovery rate (probability) for the likelihood. Set the stop criteria, either fixed iteration numbers or tolerance values. The size of the issue, the dimension number is 3, and parameter limits are also established

Stride 3: Initialize the solution randomly by generating n different nesting solutions

Stride 4: Determine the health of each solution obtained. Find the best nest that matches the lowest fitness value

Stride 5: Launch iteration, create a new nest by flight with Levy, but keep up to date. Eq.15 is carrying out a Levy flight. Start cycle, produce a new home by Levy flight however keep the current best. A Levy flight is performed by the Eq. 15

Stride 6: Evaluate and achieve new fitness in this set of solutions. If the new exercise is different than the old one, equate the old exercise to this new fitness and substitute the old fitness. The best fitness-related nest update

Stride 7: Rehash the above procedure until certain halting measures give you the best wellness and the best home

In this study, the number of host nests (NS = N/2), step size ($\lambda = 1$) and probabilities ($p_{a1} = 0.75$) and ($p_{a2} = 0.25$), maximum Number of Iterations (NoI = 100), Maximum numbers of Generation (MG = 100), upper bound (UB = -0.5) and Lower Bound (LB = 0.5), are taken for MCS algorithm. It is clear from the overview of the CS and MCS technique that the controlling parameters for the CS are Nests (NS), the maximum number of iterations.

(NOI), λ and p_{a1} while in the case of the MCS algorithm, the controlling parameters are NS, Maximum Generation number (MG), λ, p_{a1}, and p_{a2}.

Support Vector Machines

In this study, we are using the cutting-edge SVM approach as an approach to classification. This approach has been especially useful in many fields of application including ECG signals (Ramkumar *et al*., 2021; Khandoker *et al*., 2008; Bazi and Melgani, 2007; Melgani and Bazi, 2008). The generally acknowledged dominance over conventional classifiers brings the emphasis on the SVM classifier. The approach described in this study is generic and any other grouping should also be taken into consideration. We will show this resource briefly below. The reader is directed to Ramkumar *et al*. (2021) for further information. Let us first consider a binary classification problem supervised for simplicity. Let us presume, in this training set, that the d-dimensional feature space X is composed of S vectors $x_i \in R^d$ ($i = 1, 2,\ldots, S$). The d-dimensional is, for instance, morphological and time attributes got from the ECG signal. To every vector xi, we partner an objective $y_i \in \{-1, +1\}$ (e.g., typical and unusual beats, separately). The direct way to deal with SVM arrangement comprises looking for a perfect hyperplane that boosts the isolating distance (Ramkumar *et al*., 2021) for a distinction between the two groups in X. The most common case since data is not always linearly separated is the first of the two classes transformed into a higher dimension, meaning that in a nonlinear case, the value of the features of $\Phi(x) \in R_d$ ($d'>d$) is dependent on the function sign, which determines the hyperplane-related distinguishing function in a transformed field. The judgment law of membership is the function sign ($f(x)$) (Khandoker *et al*., 2008):

$$f(x) = w * \phi(x) + b * \qquad (18)$$

It is a hyperplane that minimizes loss function, integrating two needs: Maximal margin and reduced error (Sharma *et al.*, 2021), defined by the weight vector *w* to R_d and Bias *B* to *R*. This is solved by a skewed location that is easily depicted as a data function in the original (lesser) scale area *X*:

$$f(x) = \sum_{i \in s} \alpha_i^* y_i K(x_i, x) + b^* \qquad (19)$$

where, $\alpha = [\alpha_1, \alpha_2 .... \alpha_N]$ is the vector of Lagrange multipliers and $K(\cdot, \cdot)$ is a kernel function. Set *S* is an index subset $\{1, 2, .., S\}$ is analogous to a non-zero α_i 's, multiplier Lagrange that determines the supporting vectors. The Kernel $K(\bullet, \bullet)$, in the transformed (higher) dimensional feature space $\Phi(X)$, must fulfill the condition defined in the Mercer Theorem in order to conform to a certain kind of internal product in Goyal *et al.* (2016). The famous Gaussian function is a typical example of such kernels:

$$\left(-K(x_i, x) = exp\gamma \left\| x_i - x \right\|^2 \right) \qquad (20)$$

where, γ represents a parameter inverse to the Gaussian kernel's width. SVMs are basically binary classification machines as above. But the diagnosis of ECG reciprocal segregation of various information types (arrhythmia) is also required. To order to solve this problem (Dao, 2020), a number of multi-class classification methods can be implemented. In this post, we have discussed the popular OAO. Let $\Omega = \{\omega_1, \omega_2,...,\omega_T\}$ can be the set of *T* groups. This approach is based on the creation, by using the partial SVMs, of a collection of SMSs for each group of the two. $T(T-1)/2$ SVM, which are the pathways for each of the above and lower classes are used in the OAO technique (Melgani and Bazi, 2008).

Methods

Several studies have employed this method to analyze ECG signals. However, this study assumes that by prioritizing the accuracy of ECG characterization, the efficiency of wavelets can be improved. Therefore, this study presents a novel wavelet technique that considers polyphase representation and addresses the optimization problem in a Multiscale (MCS) setting.

Figure 1 illustrates that the directions of the particles in the population encode precise parameters $\{\theta\_1, \theta\_2, \theta\_3, ..., \theta\_(N-1)\}$. To evaluate the fitness function, we utilize the sequence obtained through Cross-Validation (CV) on the training dataset of the state-of-the-art SVM classifier (Bazi and Melgani, 2007). This fitness function serves as an indicative measure of how accurately the MCS guides the wavelet candidate in the wavelet space. During the training phase, the SVM parameters (regularization and portion parameters) are determined using the m-fold CV technique, where m is typically chosen between 2 and 10. This involves randomly dividing the available training beats into m mutually exclusive subsets (folds) of equal size. The SVM classifier is then trained *m* times with predefined parameter values, using m-1 folds for training while keeping one fold aside for validation (as a set of validation tests). The average accuracy obtained from these computations reflects the predictive performance of the considered SVM classifier during *m* training iterations.

In order to maximize this prediction and obtain the final accuracy estimate, the optimal SVM classification parameter is selected. Another analytical technique that can be employed for this estimation is the Leave-One-Out (LOO) method. Generally, this method is more reliable but computationally intensive, as it is equivalent to S-fold CV.

Figure 1 outlines the essential steps of the proposed wavelet system methodology as follows:
Stage 1: Initiating

- Attached channel appeal N = order_min
- L = decLev_min level of attached disintegration

Step 2: Population original

- Initialize the MCS by producing various homes for different arranged nations with an irregular population
- Register the comparative well-being work for each house as an incentive through the associated progress
- Infuse the Vector of θ_i's in (12) and (13) recursively, so that the channel of 2N-low-pass constants can be generated;
- Determine the high-pass 2*N* channel coefficients by flipping flip creation (14)
- Apply DWT to every ECG preparation beat for the subsequent low-pass and high-pass channels
- Take charge of the SVM classifier with the wavelet highlights created (and, whenever you want, other types of components). Process its cross-approval exactness ($CVA_i(t)$) to set the wellness work estimation of the home, $X_i(0) = CVA_i(0) \, for \, I = 0,1,...,N-1$
- Hold each house's condition and mark it as the best location nearby. Save the situation of the molecule as the best position in the world by providing a

bigger wellness incentive.
$$X_{b_i}(0) = X_i(0) \, and \, Xg(0) = \max(X_{b_i}(0))$$

Step 3: Process of search

- Update each molecule's vector speed (15)

- Update organizes the molecule according to (16)
- Process the wellness work estimate again for every home in the new wavelet region, $X_i(t) = CVA_i(t)$
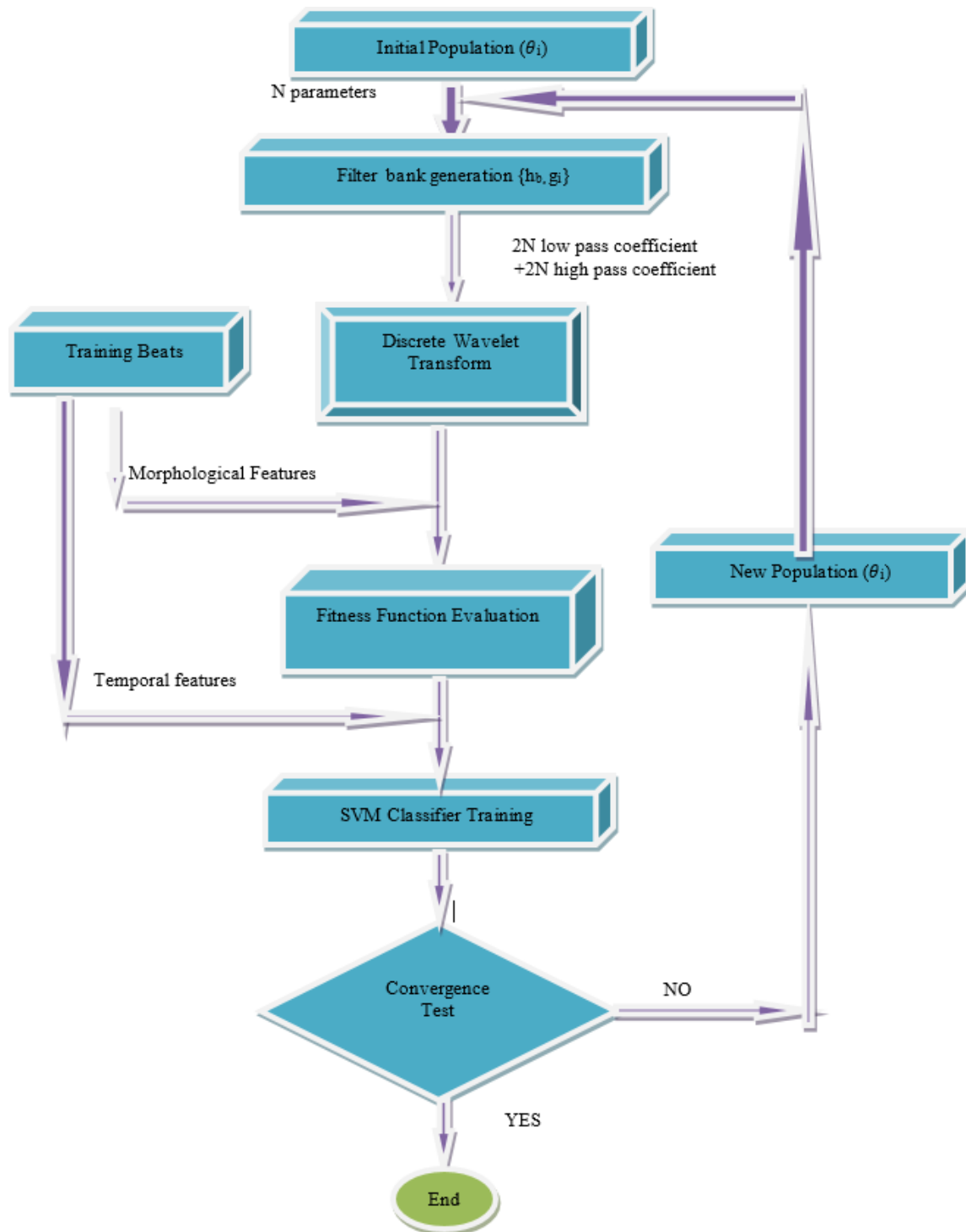- Change $Xbi(t)$ for each atom and $Xg(t) = \max(Xb_i(t))$ for the world's best location

Fig. 1: Flow chart of the proposed technique

Step 4: Convergence test In the case that the age quantity is not equal with the customer with a maximum number of ages or that the best-performing work is desired (the best position worldwide) $|Xg(t) - Xg(t -) > \varepsilon|$ (where ε is the customer's defined limit)

Step 5: Appeal filter and degree of decay

- Experiencing phases 2-4, on a single basis, for each channel operation and for each decay stage in the preset ranks (order min, request max) and (decLev min, decLev max), for a pre-defined moving rate (for instance = 1)
- Select some estimates (channel demand and level of deterioration) that provide an incentive for intermingling with the highest levels of wellness work

Results and Discussion

Dataset Explanation and Experimental Operation

We conducted our investigations on ECG data using the MIT-BIH arrhythmia database (Moody and Mark, 2001). The database contains beats from various classes, including Normal sinus rhythm (labeled as 'N'), premature Atrial beats ('A'), Ventricular premature beats ('V'), Right Bundle branch block ('RB'), left bundle branch block ('LB') and paced beats ('/'). These beats were extracted from the chronicle literature of 20 patients enrolled in classes: 100, 102, 104, 105, 106, 107, 118, 119, 200, 201, 202, 203, 205, 208, 209, 212, 213, 214, 215 and 217 (Moody and Mark, 2001).

Our contribution to this research consists of two parts: (i) The morphological features of the ECG and (ii) Three transient features, namely the QRS complex duration, the RR interval (time between two consecutive R peaks indicating the difference between the current and previous QRS peaks) and the typical 10-cycle average. To extract these features, we employed the widely used ecgpuwave software available at http:/www.physionet.org/physiotools/ecgpuwave/src/.

After removing these time-based features, the duration of the extracted ECG cycles was normalized following the protocol outlined in Lakhera *et al.* (2021). Consequently, a fixed duration of 300 samples was chosen for the normal beats. The complete set of morphological and transient analysis features resulted in 303 features per beat. From the MIT-BIH benchmark arrhythmic sample, we randomly selected 125 cases for the training trials. This limited number of pre-tests (less than 1% of the total tests) serves two purposes: (1) It reduces the processing time required for convergence and (2) It allows testing the strategy in scenarios where the availability of labeled beat samples is limited. Table 1 provides a detailed breakdown of the number of beats and test beats planned for each category. The performance of the classification system was evaluated based on four measures: (1) Overall accuracy, which represents the accuracy of correctly classified beats across all categories; (2) Class-specific accuracy, which measures the accuracy of correctly classified beats within each class; (3) Average Accuracy (AA), which considers the average accuracy of class-wise characterizations; and (4) McNemar's test, a statistical test to evaluate statistically significant differences between different classification approaches. This test follows the standard testing procedure (Gupta *et al.*, 2022):

$$Z_{ij} = \frac{f_{ij} - f_{ji}}{f_{ij} + f_{ji}} \tag{21}$$

The statistical significance of the difference in accuracy between the ith and j^{th} classifications on a pairwise basis is measured by Zij. The number of beats classified correctly and incorrectly by the i_{th} and j^{th} classifiers is presented separately, along with the number of beats that disagree with them. At the conventional 5% significance level, the difference in accuracy between the ith and j^{th} classifiers is considered statistically significant if $|Zij| > 1.96$.

Additionally, we investigated the effectiveness of this method in distinguishing Ventricular (V) beats from other types of beats. Rapid identification and treatment of V rhythm are crucial as they can be associated with localized myocardial necrosis, which can lead to increased mortality. Three standard measures were employed (Gupta *et al.*, 2022):

(1) Sensitivity (Se = TP/(TP + FN)): The ratio of correctly identified 'V' beats to the total number of 'V' beats. (2) Specificity (Sp = TN/(TN + FP)): the ratio of accurately discarded 'Non-V' beats. (3) Positive Predictivity (PP = TP/(TP + FP)): The proportion of accurately differentiated 'V' beats among the total number of recognized 'V' beats.

The above definitions refer to True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). In our MCS method, the plurality scale is set to 10 and the maximum importance is 20. The inertia weight (w) is 0.75 and $c_1 = c_2 = 1$. Figure 1 illustrates the convergence of the health indicator value at the highest position throughout the investigation process.

In all the experiments presented in this study, we employed an SVM classifier with a Gaussian kernel and performed five-fold cross-validation. It should be noted that, after decomposing the wavelet with the encoded channel, we considered three transient features alongside the 300 morphological features of the proposed wavelet structure technique. The optimal wavelet decomposition for ECG signals was determined based on previous works (Toulni *et al.*, 2021; Ince *et al.*, 2009; Sahambi *et al.*, 1997; Li *et al.*, 1995; Hao *et al.*, 2022), which involved decomposing signals up to the fourth level. Accordingly, the performance evaluation was conducted considering the first four levels of decomposition.

Table 1: No of activity check between exercise beats and examination beats

Class	N	A	V	RB	/Paced beats	LB	Total
Exercise beats	37	24	25	13	13	13	125
Examination beats	26000	236	3942	3786	6854	1867	44486

Table 2: Decomposition level representation for different order filter

Filter order	MCS-based Decomposition level				Daubechies decomposition level				Symlet decomposition level			
	1	2	3	4	1	2	3	4	1	2	3	4
4	92.36	89.64	89.42	89.48	79.26	79.38	75.56	76.35	79.65	84.07	84.73	82.45
6	92.69	89.56	87.76	87.62	76.54	78.48	76.72	77.75	79.79	83.45	83.76	81.21
8	91.63	88.86	87.83	88.80	74.66	76.75	76.43	77.46	78.45	84.32	83.54	81.33
10	90.88	88.77	88.84	87.44	73.44	73.45	78.68	78.09	78.33	84.22	84.24	81.11

Table 3: Fidelity parameters calculated using different methods

Class	OA	AA	N	A	V	RB	/Paced beats	LB
MCS-based	94.68	93.86	87.63	85.98	92.89	88.56	97.75	89.96
Daubechies	78.82	85.52	85.52	83.42	90.65	89.03	93.42	86.56
Symlet	76.89	83.79	84.79	82.42	87.36	90.42	93.85	83.89

Table 4: Contrast statistical centrality in the group precision of the three wavelets expressed in McNemar test methods. The variations apply to Table 3 tests

	Daubechies	Symlet
MCS-based	22.98	18.23
Daubechies	-	-2.56

Table 5: Statical parameters calculated using different techniques

	Se	Sp	Pp
Daubechies	89.72	94.84	70.54
Symlet	85.40	95.69	71.64
MCS-based	94.73	98.25	82.65
Gupta and Mittal (2021)	84.67	88.98	66.76
Goyal *et al.* (2016)	82.56	87.09	59.07

Results

The accuracy of test beats is demonstrated in Table 2 by feeding the SVM classification with the characteristics of the wavelet structure (MCS-based) and the Daubechies wavelet (Db) and the Symlet wavelet (Sym) to be evaluated. For an enhanced and fair-minded relationship, we noted that the three wavelets explored were evaluated on a comparative basis: A comparable number of features, comparable planning performances, a comparative classification, and a comparative test performance. At the end of the day, the MCS-based wavelet is less sensitive than normal wavelets to the option of channel requeston and decomposition. Table 3 indicates the best possible correctness for each form. The MCS-based technique impressively improves the sensitivity of regular waves as it behaves differently. Indeed, the OA (and AA) increases are independent of Daubechies and Symlet, at 3,32 percent (2.65 percent) and 3,42 percent (3,68 percent). Atrium trouble beats (A), with the correctness of 82,77%, 84,03%, and 81,09% only for MCS, Db, and Sym are the worst class

precision. Paced beats (/) with corrections of 96.19%, 92.33%, and MCS, Db, and Sym, are individually the highest class accuracy. A close analysis of Table 3 further reveals that the technique we recommend does not only confirm a dominant, i.e. typical beats (N). In any case, however, there are three distinct classes of five minority reminder classes. The McNemar test stresses that the overall relevance of the proposed strategy over the other two wavelets is considerably great (Table 4). Table 5 shows that the MCS-based wavelet derogates less than optimal beats from other normal wavelets in the experience of ventricular ventricles. In total, it received reports of 6.35, 2.03, 0.45 1.30, 2.62, and 3.72% of optimistic predictability, with a clear analysis of the waves of Symlet and Daubechies. It is important to remember that these modifications were only necessary to enhance the wavelet in multi-class request settings and not in two-way settings (i.e., "V" against each extraordinary class).

Conclusion

This study projects a new method of advancing the wavelet depending on the combination of wavelet and MCS polyphase representation. It can be seen that, the wavelet that best reflects the beats as far as the separation capacity is measured with a target accuracy estimate. The ECG signal portrayal for the being talked about grouping task is hence advanced. The procedure has been approved on the MIT/BIH arrhythmia benchmark by taking the state-of-the-raft SVM grouping as the classifier. The test outcomes show a critical increment in the order precision and strength of two normal wavelets (Daubechies and Symlet). While this study has utilized the SVM strategy, the method is conventional and should likewise be possible for some other arrangement method. Its key drawback is the significant preparing time to accomplish assembly, particularly while considering wide preparation sets.

Acknowledgment

Funding Information

Author's Contributions

Banit Negi and Priti Kumari: Contributed to literature review and research.

Abhilekh Bartwal, Surjeet Singh Patel and Vivek Kumar Tamta: Conducted the implementation and practical aspects of the study.

Agya Ram Verma, Yatendra Kumar and Abhishek Gupta: Authored the paper and contributed significantly to its writing.

Ethics

This article is original and contains unpublished material. The corresponding author confirms that all of the other authors have read and approved the manuscript and no ethical issues involved.

References

Bazi, Y., & Melgani, F. (2007). Semisupervised PSO-SVM regression for biophysical parameter estimation. *IEEE Transactions on Geoscience and Remote Sensing*, *45*(6), 1887-1895.
https://doi.org/10.1109/TGRS.2007.895845

Chandra, S., Sharma, A., & Singh, G. K. (2018). Feature extraction of ECG signal. *Journal of Medical Engineering & Technology*, *42*(4), 306-316.
https://doi.org/10.1080/03091902.2018.1492039

Dao, T. P. (2020). Cuckoo Search Algorithm: Statistical-Based Optimization Approach. *Applications of Cuckoo Search Algorithm and its Variants*, 79. ISBN: 9811551634.

Daqrouq, K., Alkhateeb, A., Ahmad, W., Khalaf, E., Awad, M., North, E., ... & Rushdi, A. M. (2022). A Universal ECG Signal Classification System Using the Wavelet Transform. *Neural Network World*, (1).
https://doi.org/10.14311/NNW.2022.32.003

De Chazal, P., & Reilly, R. B. (2006). A patient-adapting heartbeat classifier using ECG morphology and heartbeat interval features. *IEEE Transactions on Biomedical Engineering*, *53*(12), 2535-2543.
https://doi.org/10.1109/TBME.2006.883802

Goyal, S., Goswamy, S., Negi, A., Tomar, A., Verma, A. R., & Singh, Y. (2016). Design of ANC filter using modified cuckoo search technique for ECG signal enhancement. *Perspectives in Science*, *8*, 43-45.
https://doi.org/10.1016/j.pisc.2016.03.002

Gupta, V., & Mittal, M. (2021). R-peak detection in ECG signal using Yule–Walker and principal component analysis. *IETE Journal of Research*, *67*(6), 921-934.
https://doi.org/10.1080/03772063.2019.1575292

Gupta, V., Saxena, N. K., Kanungo, A., Gupta, A., Kumar, P., & Salim. (2022). A review of different ECG classification/detection techniques for improved medical applications. *International Journal of System Assurance Engineering and Management*, *13*(3), 1037-1051.
https://doi.org/10.1007/s13198-021-01548-3

Hao, J., Yang, Y., Zhou, Z., & Wu, S. (2022). Fetal electrocardiogram signal extraction based on fast independent component analysis and singular value decomposition. *Sensors*, *22*(10), 3705.
https://doi.org/10.3390/s22103705

Inan, O. T., Giovangrandi, L., & Kovacs, G. T. (2006). Robust neural-network-based classification of premature ventricular contractions using wavelet transform and timing interval features. *IEEE Transactions on Biomedical Engineering*, *53*(12), 2507-2515.
https://doi.org/10.1109/TBME.2006.880879

Ince, T., Kiranyaz, S., & Gabbouj, M. (2009). A generic and robust system for automated patient-specific classification of ECG signals. *IEEE Transactions on Biomedical Engineering*, *56*(5), 1415-1426.
https://doi.org/10.1109/TBME.2009.2013934

Jayaraman, V., & Sultana, H. P. (2019). Artificial gravitational cuckoo search algorithm along with particle bee optimized associative memory neural network for feature selection in heart disease classification. *Journal of Ambient Intelligence and Humanized Computing*, 1-10.
https://doi.org/10.1007/s12652-019-01193-6

Kaya, Y. (2018, May). Feature selection using a binary cuckoo search algorithm. In *2018 26th Signal Processing and Communications Applications Conference (SIU)* (pp. 1-4). IEEE.
https://doi.org/10.1109/SIU.2018.8404843

Khamene, A., & Negahdaripour, S. (2000). A new method for the extraction of fetal ECG from the composite abdominal signal. *IEEE Transactions on Biomedical Engineering*, *47*(4), 507-516.
https://doi.org/10.1109/10.828150

Khandoker, A. H., Palaniswami, M., & Karmakar, C. K. (2008). Support vector machines for automated recognition of obstructive sleep apnea syndrome from ECG recordings. *IEEE Transactions On Information Technology In Biomedicine*, *13*(1), 37-48. https://doi.org/10.1109/TITB.2008.2004495

Kumari, V. S. R., & Kumar, P. R. (2015). Optimization of multi-layer perceptron neural network using genetic algorithm for arrhythmia classification. *Communications*, *3*(5), 150-157. https://doi.org/10.11648/j.com.20150305.21

Lakhera, N., Verma, A. R., Patel, S. S., & Gupta, B. (2021). A novel approach of ECG signal enhancement using adaptive filter based on whale optimization algorithm. *Biomedical and Pharmacology Journal*, *14*(4), 1895-1903. https://doi.org/org/10.13005/bpj/2288

Li, C., Zheng, C., & Tai, C. (1995). Detection of ECG characteristic points using wavelet transforms. *IEEE Transactions on Biomedical Engineering*, *42*(1), 21-28. https://doi.org/10.1109/10.362922

Melgani, F., & Bazi, Y. (2008). Classification of electrocardiogram signals with support vector machines and particle swarm optimization. *IEEE Transactions on Information Technology in Biomedicine*, *12*(5), 667-677. https://doi.org/10.1109/TITB.2008.923147

Moody, G. B. & Mark, R. G. (2001). The Impact of the MIT-BIH Arrhythmia Database. *IEEE Engineering in Medicine and Biology Magazine*, *20*, 45-50. https://doi.org/10.1109/51.932724

Pandey, S. K., Janghel, R. R., & Vani, V. (2020). Patient-specific machine learning models for ECG signal classification. *Procedia Computer Science*, *167*, 2181-2190. https://doi.org/10.1016/j.procs.2020.03.269

Ramkumar, M., Babu, C. G., Priyanka, G. S., & Kumar, R. S. (2021, March). Ecg arrhythmia signals classification using particle swarm optimization-support vector machines optimized with independent component analysis. In *IOP Conference Series: Materials Science and Engineering* (Vol. *1084*, No. 1, p. 012009). IOP Publishing. https://doi.org/10.1088/1757-899X/1084/1/012009

Sahambi, J. S., Tandon, S. N., & Bhatt, R. K. P. (1997). Using wavelet transforms for ECG characterization. An online digital signal processing system. *IEEE Engineering in Medicine and Biology Magazine*, *16*(1), 77-83. https://doi.org/10.1109/51.566158

Senhadji, L., Carrault, G., Bellanger, J. J., & Passariello, G. (1995). Comparing wavelet transforms for recognizing cardiac patterns. *IEEE Engineering in Medicine and Biology Magazine*, *14*(2), 167-173. https://doi.org/10.1109/51.376755

Sharma, P., Dinkar, S. K., & Gupta, D. V. (2021). A novel hybrid deep learning method with cuckoo search algorithm for classification of arrhythmia disease using ECG signals. *Neural Computing and Applications*, *33*, 13123-13143. https://doi.org/10.1007/s00521-021-06005-7

Sörnmo, L., & Laguna, P. (2006). Electrocardiogram (ECG) signal processing. *Wiley Encyclopedia of Biomedical Engineering*. https://doi.org/10.1002/9780471740360.ebs1482

Stone, M. (1974). Cross-validatory choice and assessment of statistical predictions. *Journal of the Royal Statistical Society: Series B (Methodological)*, *36*(2), 111-133. https://doi.org/10.1111/j.2517-6161.1974.tb00994.x

Toulni, Y., Benayad, N., & Taoufiq, B. D. (2021). Electrocardiogram signals classification using discrete wavelet transform and support vector machine classifier. *Int J Artif Intell.* https://doi.org/10.11591/ijai.v10.i4.pp960-970

Vaidyanathan, P. P. (2006). *Multirate systems and filter banks*. Pearson Education India. ISBN: 8177589423.