Research Article

An Innovative Machine Learning Approach for Classifying ECG Signals in Healthcare Devices

Kishore B,1 A. Nanda Gopal Reddy,2 Anila Kumar Chillara,3 Wesam Atef Hatamleh,4 Kamel Dine Haouam,5 Rohit Verma,6 B. Lakshmi Dhevi,6 and Henry Kwame Atiglah7

1Department of Computer Science and Engineering, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, Karnataka 576104, India
2Mahaveer Institute of Science and Technology, Hyderabad, India
3Birla Institute of Technology & Science, Hyderabad, India
4Department of Computer Science, College of Computer and Information Sciences, King Saud University, P. O. Box 51178, Riyadh 11543, Saudi Arabia
5School of Electronics, Dublin City University, Dublin, Ireland
6Institute of Artificial Intelligence and Machine Learning, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai, India
7Department of Electrical & Electronics Engineering, Tamale Technical University, Tamale, Ghana

Correspondence should be addressed to Henry Kwame Atiglah; hkatiglah@tatu.edu.gh

Received 22 January 2022; Revised 20 February 2022; Accepted 23 February 2022; Published 13 April 2022

Academic Editor: Bhagyaveni M. A

Copyright © 2022 Kishore B et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

An ECG is a diagnostic technique that examines and records the heart’s electrical impulses. It is easy to categorise and prevent computational abstractions in the ECG signal using the conventional method for obtaining ECG features. It is a significant issue, but it is also a difficult and time-consuming chore for cardiologists and medical professionals. The proposed classifier eliminates all of the following limitations. Machine learning in healthcare equipment reduces moral transgressions. This study’s primary purpose is to calculate the R-R interval and analyze the blockage utilising simple algorithms and approaches that give high accuracy. The MIT-BIH dataset may be used to rebuild the data. The acquired data may include both normal and abnormal ECGs. A Gabor filter is employed to generate a noiseless signal, and DCT-DOST is used to calculate the signal’s amplitude. The amplitude is computed to detect any cardiac anomalies. A genetic algorithm derives the main highlights from the R peak and cycle segment length underlying the ECG signal. So, combining data with specific qualities maximises identification. The genetic algorithm aids in hereditary computations, which aids in multitarget improvement. Finally, Radial Basis Function Neural Network (RBFNN) is presented as an example. An efficient feedforward neural network lowers the number of local minima in the signal. It shows progress in identifying both normal and abnormal ECG signals.

1. Introduction

Automatic electrocardiogram analysis is the best practice utilized by clinicians for scrutinizing and recording the functions of the heart by positioning the electrodes at the external area of the skin membrane can be observed by electrocardiogram device and a greater number of researches are focused by scientists in recent years [1, 2]. Advancement in the technology provides enhancement in visualizing heart abnormalities at regular interval. It is most helpful in diagnosing cardiac disorders such as myocardial infarction. The extended ischemia will continue till the cells start to die, which is called myocardial infarction. In India, compiling accurate data on sudden cardiac death, 5.5% of all-out mortality happened, and around 1-fifth of all the cardiovascular passing and 6 lakh heart passing in the nation were suddenly occurred [3, 4]. A determination of myocardial dead tissue is produced by incorporating the historical
backdrop of the displaying ailment and physical investigation with electrocardiogram discoveries. Free wall rupture is a complicated one. In [5], it occurs in 1% of patients of acute myocardial infarction and it accounts for up to 7% of all infarct-related deaths. Automatic ECG analysis works well in identification of cardiac-related problems in an enhanced manner and for the better treatment. We mean to decide the measure of heart tissue harm by multigoals examination of electrocardiogram signals [6]. The most critical component of the ECG signal is the QRS complex, the pinnacle of which is indicated as R-peaks [7]. The R-R intermission means the time space between the two successive R tops. It is utilized to find the abnormalities in the heart normal operation called arrhythmia. The diagnosis involves an estimation of the size of infarct and to identify the acute complications [8]. In ECG, Q and T waves play a major role in electrocardiogram signal. In P wave, if any problem occurs, it does not lead to any complication. So, QRS detection is necessary to achieve our target. T-wave change occurs in larger area, and it denotes ischemia and ST segment change occurs in lesser number of leads and it indicates the myocardial injury and Q-wave overlie and it denotes the main area of myocardial necrosis [9, 10]. Many researchers have worked in the area of medical field in performing analysis of cancer detection, electrocardiogram analysis, and so on.

Sudden Cardiac Arrest (SCA) or Sudden Cardiac Death (SCD) is one of the most common causes of cardiac mortality in the world, accounting for about one-third of all cardiac deaths. If the danger of SCD can be detected at an early stage, it may be feasible to preserve the lives of patients by administering suitable treatment at the right time. The risk of SCD may be detected by analyzing the conventional 12-lead Electrocardiogram (ECG) data, which are available in most hospitals. Various studies have shown that differences in the shape of the ECG, particularly in the ST-T-wave and QT segments, are directly associated with the risk of SCD and Ventricular Arrhythmias (VA). Some of these changes are so minute that they are not detectable just by looking at the ECG for a short period of time. As a result, advanced-level computerized ECG algorithms are needed for this new field of investigation. The single and multilead approaches used in this thesis are proved to be effective in the analysis of ST and QT segments. By using the multilead idea, the goal is to enhance the quantitative and qualitative performances of the currently existing methodologies. T-Wave Alternans (TWA) and the QT interval, two noninvasive SCD indicators, are investigated in depth in this study. For the categorization of MI and healthy people, a novel Stationary Wavelet (SWT) method is proposed. Multilead QT interval analysis is also carried out using three frank leads, designated as X, Y, and Z. Multilead and single-lead approaches are used to analyze patients with a variety of cardiac problems as well as healthy subjects. In addition, as a consequence of participation in the worldwide-level challenge, the methods for measuring the foetal QRS and QT intervals were explained in the dissertation. PhysioNet/CinC 2013 is a collaborative effort. Brief description of the work given in the thesis is provided in the paragraphs following the thesis. The creation of a multilead TWA detection idea is the first step in the process. Because of the limits of single-lead ECG analysis, which is lead dependent, a new multilead approach is suggested to address these issues. For the purpose of translating the alternans-related information from the ST segment into a new signal known as the derived lead, the Principal Component Analysis (PCA) approach is utilized in this procedure. With the use of calibrated alternans records, the algorithm has been confirmed. Figure 1 represents the basic signal.

In [11], the authors focus on determining the blockage and R-R interval to achieve with good accuracy. Systematic finding of QRS complex is essential to extract the R-R interval from the electrocardiogram recordings. To accurately analyze the cardiac rate variation, RR series plays a significant role and it is helpful to provide a quantitative evaluation of heart autonomic capacity in wellbeing and in sickness states [12]. In the past decades, wide collections of algorithm and techniques were used in understanding automatic regulation of heart beat. But the ECG recording may contain fictitious occurrences because of multiple disruptions like commotion interference in the signal, unexpected change in amplitude of QRS, and so on [13, 14]. Since there are so many methods for detecting a QRS signal as well as preventing its propagation, it is important to pick a method that works in real time and can handle big datasets while requiring little computational effort [15, 16]. In this study, the noise interference that is present in the electrocardiogram will be removed by handling preprocessing and then it is split up into samples by using the algorithm DCT-based DOST [17] and amplitude is computed in each interval. If there is any complication found in computing amplitude, it detects a block in such area. Initially it is of 100 hertz. It is split up into 5 intervals PQRST and amplitude is of 1 millivolt. Frequency is computed by $f = \frac{1}{T}$ and the next step involved in our work is feature extraction [18]. It helps to compute the mean and average of each interval and finally, Radial Basis Function Neural Network (RBFNN) is used to analogize the trained and test data. The data is collected from the MIT-BIH dataset. The collected information has normal dataset and abnormal dataset [19]. The trained and test dataset is analogous with the ratio 1:6 and the expected accuracy is met. The last objective of this work is to decide the perfect calculation for analogizing various classes of ECG oddities by quantitatively looking at the different QRS identification method to detect the blockage and R-R interval and delineating their failure instance [20, 21]. By this study, it achieves 98.5% accuracy.

2. Proposed Works

There are many databases accessible for public use, including the MIT-BIH arrhythmia database, which contains standard investigative material for the identification of cardiac arrhythmias. It has been in use since 1980 for the purposes of basic research and medical device development in the field of cardiac rhythm and associated illnesses.

There are many databases accessible for public use, including the MIT-BIH arrhythmia database, which contains standard investigative material for the identification of
cardiac arrhythmias. It has been in use since 1980 for the purposes of scientific research and medical device development in the field of cardiac rhythm and associated illnesses. The goal of creating the database is to develop automated arrhythmia detectors that read the variety of the signal and, on the basis of that, can perform automated cardiac diagnostics. The many intricacies of the ECG, such as the variation of the waveform of the pulse and the accompanying cardiac beat, as well as the baffling strength of artefacts and noise, combine to make signal analysis difficult. As a result, automation of the recording of the Electrocardiogram (ECG) signal is clear, and many publicly accessible databases exist that store the recorded ECG signal for future medical use. The MIT-BIH arrhythmia database is primarily utilized for medical and scientific purposes including the identification and analysis of various cardiac arrhythmias. It is the goal of this database to create a completely automated environment in which precise information may be obtained for the diagnosis of ventricular arrhythmias.

Electrocardiograms (ECGs) are very popular because they are a low-cost and noninvasive method of examining the physiologic function of the heart. Initially developed in 1961, Holter introduced techniques for continuous recording of the ECG in ambulatory subjects for extended periods of time. The long-term ECG (Holter recording), which typically lasts 24 hours, has since become the standard technique for observing transient aspects of cardiac electrical activity.

Since the mid-1970s, our research group has investigated irregularities in heart rhythm (arrhythmias) as seen in long-term electrocardiograms (ECGs), as well as automated approaches for detecting arrhythmias in real time. Other research groups in academia and business have pursued topics that are comparable to this one. Until 1980, anyone seeking to pursue such a career were required to gather their own information. Despite the fact that the recordings themselves are copious, access to this data is not ubiquitous, and comprehensive characterization of the recorded waveforms is a time-consuming and costly procedure. Aside from that, there is a great deal of variation in ECG rhythms and features of waveform morphology, both across subjects and within persons over time, and therefore a meaningful representative collection of long-term ECGs for study must comprise a large number of recordings.

Development of automated arrhythmia analysis algorithms was slowed throughout the 1960s and 1970s due to a scarcity of data that could be accessed by all researchers. When doing such work, each group gathered its own collection of recordings and often utilized the same data that had been used to construct the algorithms in order to self-evaluate their algorithms. From the beginning, it was evident that the performance of these algorithms was inevitably data-dependent, and that the use of different data for the assessment of each algorithm made it impossible to make objective comparisons between algorithms belonging to various groups of algorithms.

Figure 1: Structure of ECG signal.
Data Collection and Selection. As soon as we realised we would require a suitable set of well-characterized long-term ECGs for our own research, we began collecting, digitising, and annotating long-term ECG recordings obtained by the Arrhythmia Laboratory of Boston’s Beth Israel Hospital (BIH; now known as the Beth Israel Deaconess Medical Center), which was established in 1975. However, we intended to make these recordings accessible to the wider research community from the beginning, in order to spur more study in this area and to promote rigorously repeatable and objectively comparable assessments of various methods [3]. We anticipated that the availability of a shared database would be a positive development.

Our proposed system focuses on the blockage area to detect the R-R interval from the ECG signal. This DCT-DOST segmentation with adaptive threshold is used in this paper to determine the QRS complex and R peak from the recorded signals of the MIH-BIH database. The distortion in the ECG is filtered by a Gabor filter and therefore the QRS complex information was preserved. After denoising, the signal gets segmented into 256 constituent parts and the magnitude is found to compare with trained data. It is performed to identify the cardiac abnormality. The difference in the amplitude and time period of the ECG signal helps to analyse the abnormality. Nearly 50,000 samples of ECG signals were considered for analysis. The sampling frequency is split into 5 intervals to detect RR interval. The mean, variance, and entropy are evaluated to extract the features. Genetic algorithm is used to select significant features and is labelled with specific class. The R peak, variance, and entropy are evaluated to identify the R-R interval from the ECG signal. This DCT-DOST segmentation was preserved. After denoising, the signal gets segmented into 256 constituent parts and the magnitude is found to compare with trained data. It is performed to identify the structural abnormality. The difference in the amplitude and time period of the ECG signal helps to analyse the abnormality. Nearly 50,000 samples of ECG signals were considered for analysis. The sampling frequency is split into 5 intervals to detect RR interval. The mean, variance, and entropy are evaluated to extract the features. Genetic algorithm is used to select significant features and is labelled with specific class. The R peak, variance, and entropy are evaluated to identify the R-R interval from the ECG signal.

Figure 2 represents the block diagram of proposed work.

2.1. Preprocessing. Gabor filter is a type of linear filters whose response for impulse signal is characterized as a Gaussian function paired with a coherence function [22, 23]. The requirement of minimal space bandwidth product makes this filter highly suitable for our proposed work.

Figure 3 represents the frequency domain. To define the result of signal propagation in frequency domain, the unpredictable theory should surpass or equals the constant value.

\[
\Delta t \Delta f = c,
\]

where \( c \) is a constant, \( \Delta t, \Delta f \Delta \) is the time and frequency space measurement.

In 2D type, the time variable \( t \) is supplanted by spatial coordinates \((x, y)\), and the frequency \( f \) is superseded by space variables \((u, v)\). In most cases, the 2D Gabor function is evaluated as follows:

\[
g(u, v) = \exp \left\{ \frac{1}{2} \left[ \frac{(u - w)^2}{\sigma_u^2} + \frac{v^2}{\sigma_v^2} \right] \right\},
\]

where \( \sigma_u = 1/2\pi\sigma_x \) and \( \sigma_v = 1/2\pi\sigma_y \), while the standard deviation of the elliptical Gaussian is represented as \( \sigma_x \) and \( \sigma_y \) in the \( x \)- and \( y \)-axes. For exact amplitude estimation, the DC values of a 2D Gabor filter were used to minimize the higher order harmonics. The formula used to calculate the filter parameter is

\[
a = \left( \frac{u_0}{u_1} \right)^{-1/s-1},
\]

and \( \sigma_u \) is computed by using the equation

\[
\sigma_u = \left( N_u / (a + 1)^{1/2} \right).
\]

Figure 4 represents the Gabor filter in the proposed work.

2.2. DCT-DOST-Based Segmentation. This method uses the DCT-DOST scheme to examine the time domain representation of the ECG signal and to naturally distinguish the R peak. The result of signal propagation in frequency domain, the unpredictable theory should surpass or equals the constant value.

\[
s(\tau, f) = \int_{-\alpha}^{\alpha} [h(t) e^{-(t-t_0)^2/2\sigma_t^2} e^{-j2\pi f t}] dt.
\]

\[
\delta(\tau, f) = \int_{-\alpha}^{\alpha} [h(t) e^{-(t-t_0)^2/2\sigma_t^2} e^{-j2\pi f t}] dt.
\]

\[
\sigma(f) = T = \frac{1}{|f|},
\]

\[
\delta(\tau, f_0) = \int_{-\alpha}^{\alpha} [h(t) e^{-(t-t_0)^2/2\sigma_t^2} e^{-j2\pi f t}] dt.
\]

where

\[
G(m, n) = e^{-2\pi i m n / n},
\]

where \( n \) extends from 1, 2, \ldots, N - 1.
The proposed work’s main goal is to automatically find the peak value of R. To detect the R peak in this proposed work, every heartbeat segment is assumed to consist of 105 patterns prior to the R top identification and 151 patterns were generated after the retrieval of R peak. A sum of 256 patterns were taken to find the extension of cardiac pulse. The advantage of determining the length of every cardiac pulse is to accurately detect the R top corresponding to P and T signs due to its minimum magnitude and are noise sensitiveness. Figure 5 represents the R peak detection using DCT.

After the retrieval of noiseless image, DCT-DOST approach is applied for performing operation and the peak identification is performed. Initially the sample frequency is of 100 hertz. It is split up into five intervals to accurately locate the R-R interval. DCT is chosen because of the following reason. It is a true value transformation, and it is strategically placed in space to reduce the amount of time required. It does not include any negative frequency. Only positive frequencies are used, and there is no symmetry coefficient as a result, higher frequencies are needed to convert to frequency space during segmentation. Since the DCT-DOST contains no negative frequencies, the frequency width for any signal of length 2N is as follows:

\[ N_1 = 1, \]
\[ N_i = 2^{i-2}, \quad \text{for} \, 2 \leq i \leq N - 1. \]  

(11)

The DCT-DOST method is as follows.

Initially, the info ECG signal propagates via N point DCT. This level produces the coefficients \( A_1, A_2, \ldots, A_N \). The acquired coefficients are split into subbands \([2^0, 2^1, 2^2, \ldots, 2^{n-1}]\). For each subband, \( \beta \) point inverse DCT operation is performed to ensure the \( \beta \) bandwidth that is generated in the frequency and decomposition is perpendicular.

Figure 6 shows input ECG signal example.

2.3. Feature Extraction. In ECG signal, feature extraction helps to figure out the amplitude and interval values of P-QRS-T segment present in the ECG signal. The primary goal of this suggested work is to identify the R-R interval and extract the transitory and morphological highlights from the data. By utilising highlight extraction, 19 transient highlights including PQ, RR, and PT interim and 3 morphological highlights were extricated from the ECG signal. Figure 7 represents the feature extraction of the proposed work.
The maximal and minimal points of each beat of the ECG signal were captured using morphological highlights. The equation is

\[ f(t) = f(t) - \frac{\min(t)}{\max(t)} - \min(t). \]  \hspace{1cm} (12)

The least value and most value point were figured out in the first R peak and the next R peak and then it normalized by taking the estimates between 0 and 1.

Feature describing the position of P, Q, R, S, T peak and QRS duration has been computed by using the initial position of the Q-wave to the end of the S-wave. The QRS
complex is computed, which has a significant role in the detection of abnormality.

2.4. Algorithm Used to Compute Duration of QRS Complex

(i) Step 1: Read the signal.

(ii) Step 2: Identify the duration of QRS complex waveform.

(iii) Step 3: Execute the wavelet analysis.

(iv) Step 4: Calculate the coefficients by using wavelet decomposition.

(v) Step 5: Identify R peak location in the signal by taking 60% of its value as threshold.

(vi) Step 6: Identify Q point by finding the smallest value ranging from Rloc − 50 to Rloc − 10.

(vii) Step 7: Identify S point by finding the smallest value ranging from Rloc + 5 to Rloc + 50.

(viii) Step 8: Identify T point by finding the highest value ranging from Rloc + 25 to Rloc + 100.

(ix) Step 9: Compute the duration of QRS complex by using the equation

\[ QRS(i, j) = \text{ceil}((\text{SOFF}(i, j) − \text{QON}(i, j))). \]  

(x) Step 10: Find \( X = \text{QRS} \).

False negative detection of QRS complex is carried out by using the following.

(a) Premature ventricular complexes

(b) Low amplitude

False positive detection is carried out by using the following.

(a) Negative QRS complexes

(b) Low SNR
This QRS algorithm is helpful to extract the R-R interval. It was performed by using heart rate variability (HRV). It is defined as the interval among two successive R peaks. The R-R interval was computed by using the equation
\[
rr(i) = rr(i + 1) - r(i); \quad 1, 2, \ldots, m - 1.
\] (14)

where \( r(i) \) is the peak time of \( i \)th wave.

Figure 8 represents the structure of genetic algorithm. The next step was to reduce the number of features. It is done with the aid of a genetic algorithm. Recently, there has been a surge in the use of genetic algorithms to reduce enhancement issues. This algorithm is used in high-complexity executions and large sets of arrangements. It was utilized to improve the features for identifying ECG signals. It assists in extracting the most desirable characteristics and is incorporated into the following generation. The next generation would choose the best conditions, while the others would be ignored. It begins to repeat and build a population by producing a new population at each stage through selection, crossover, and mutation and then continues in this manner.

And finally, it applies a fitness function, which is computed by
\[
f.f = \frac{1}{n} \sum_{i=1}^{n} (t - \text{out}).
\] (15)

\( n \) stands for the number of outputs, \( t \) stands for the goal output, and \( \text{out} \) stands for the actual output. Positive and negative values may be present in the fitness function. As a result, we cannot use fitness benefit directly. The selection operator is used to identify the best features associated with the highest fitness value and passes them over to the next generation. The crossover operator swaps the selected individuals chromosomes to produce offspring chromosomes.

Chromosome \( i \) reproduce
\[
\text{Chromosome } i \text{ reproduce } = \frac{f(x_i)}{\sum_{k=1}^{N} f(x_k)}
\] (16)

The final operator is then used to notify the bits in the chromosome. The probability that the chromosome in the \( n \)th position will be estimated is calculated using
\[
P_n = \frac{N - N + 1}{\sum_{i=1}^{N} i}.
\] (17)

The GA algorithm aids in the optimization of neural network results, and it works well to achieve high precision, sensitivity, and specificity, as well as providing output with better classification. The classification is performed by RBFNN.

2.5. Radial Basis Function Neural Network. RBFNN is a function that is used in time series prediction, classification, and approximation of function. It can be used for any type of model, including linear and nonlinear, as well as any network. The three layers are input layer, hidden layer, and output layer. The input to the hidden layer is converted nonlinearly by the hidden layer. The hidden layer’s activation is combined in a linear way by the output layer. The input layer is represented as an \( x \in \mathbb{R}^n \) vector of real numbers. The network’s result is \( R^n \rightarrow R \), which is given by
\[
\varphi(x) = \sum_{i=1}^{N} a_i x(\|x - c_i\|),
\] (18)

where the neurons present in the hidden layer are represented as \( N, C_i \) is the centre vector, and \( a_i \) is the neuron’s weight. The parameters \( a_i, c_i, \) and \( \beta_i \) aid in optimizing the fitness between \( \varphi \) and the signal. Figure 9 represents the RBNN network.

A typical RBF of the scalar input vector that is a first layer is
\[
h(x) = \exp\left(-\frac{(x - c)^2}{\rho^2}\right)
\] (19)

Normalized and denormalized forms of the generated input are also possible. But it is discovered to be in a nonnormalized state. The equation is
\[
\varphi(x) = \frac{\sum_{i=1}^{N} a_i \rho(\|x - c_i\|)}{\sum_{i=1}^{N} \rho(\|x - c_i\|)}
\] (20)

where
\[
\rho(\|x - c_i\|) = \frac{\rho(\|x - c_i\|)}{\sum_{i=1}^{N} \rho(\|x - c_i\|)}
\] (21)

This input layer expression can also be expressed as
\[
\varphi(x) = \sum_{i=1}^{2N} \sum_{j=1}^{n} e_{ij} v_{ij}(x - c_i),
\]

where
\[
e_{ij} = \begin{cases} 
a_i, & \text{if } i \in [1, N], 
\end{cases}
\]
\[
b_{ij}, & \text{if } i \in [N + 1, 2N], 
\]
\[
v_{ij}(x - c_i) = \begin{cases} 
\delta_{ij} \rho(\|x - c_i\|), & \text{if } i \in [1, N], 
\end{cases}
\]
\[
(x_{ij} - c_{ij}) \rho(\|x - c_i\|), & \text{if } i \in [N + 1, 2N]. 
\]

In the denormalized form,
\[
v_{ij}(x - c_i) = \begin{cases} 
\delta_{ij} u(\|x - c_i\|), & \text{if } i \in [1, N], 
\end{cases}
\]
\[
(x_{ij} - c_{ij}) u(\|x - c_i\|), & \text{if } i \in [N + 1, 2N]. 
\]

In the normalized form,
\[
\delta_{ij} = \begin{cases} 
1, & \text{if } i = j, 
\end{cases}
\]
\[
0, & \text{if } i \neq j. 
\]

The output of the training dataset is \(Y_s\) and time prediction is done by predicting the successive value and features of a sequence:
\[
\ldots, y_{t-3}, y_{t-2}, y_t - 1, \ldots
\]

3. Results and Discussion

The proposed ECG classification method discussed in this paper is implemented in MATLAB to analyze ECG signals. The proposed methodology is implemented in MATLAB and the MIT-BIH dataset is used to validate [24]. The RBFNN classifier is trained with the data from the previous section, and its performance is evaluated using the sample ECG signal as an example. The expected performance for the ECG signals at each subsequent stage of the proposed method is exhibited for detailed analysis. The ECG specimen image taken for analyzing has been elaborated for 50,000 samples. One of the sample ECG signals is shown in Figure 10.

The process of the proposed methodology starts with filtering of noises using Gabor filter. The two types of noises in the ECG signal are high-frequency noises such as electromyogram noise and Gaussian noise and low-frequency noises like baseline wandering, and power line interference causes misinterpretation [25]. To eliminate all these noises, orientation-specific encoding schemes like Gabor filter is used for analyzing the texture features of ECG signal. Analogous to input signal, the output of Gabor is more precise and accurate [26].
Figure 11 represents the Gabor filter output. For further processing with minimum data redundancy and to constrain the dataset integration, the filtered output is normalized.

The distance between the R-peak values is estimated by finding the absolute values, as shown in Figure 12.

When the heart’s electrical function is assumed as a vector, it is easy to analyze the trajectory of the vectors peak. The signal ECG is considered as projection of the heart’s electrical vector on the corresponding lead vector as a time function (amplified by the absolute magnitude of the lead vectors). It is depicted in Figure 13 below.

Generally, the coefficients are dispersed based on the bandwidth. The energies in the ECG signal are gathered together using DCT-DOST so as to represent the most important coefficient at the low frequency. The features that are extracted using the DCT-DOST approach indicate the time-recurrence attributes of the ECG signal and are unsymmetrical in nature. Also the peak values in QRS polarity and the unexpected variations in QRS amplitude are detected. Figure 14 represents the energy results.

The traditional filtering minimizes signal noise by delaying the QRS components. As QRS complex represents the ventricular activity of heart, it is necessary to preserve them. The zero-phase filtering minimizes phase distortion and provides a compromise among filtering and data retention. The output of the zero-phase filter is depicted in Figure 15.

It is composed of 112 patterns before the R top occurs and 144 patterns after the R top occurs; an aggregate of 256 patterns are chosen to find the length of every occasion relating to window size. The ECG portion is composed of 112 patterns before the R top occurs and 144 patterns after the R top occurs. The duration of each event is determined in order to condense the great majority of the data collected in relation to each cardiac event as much as possible. The benefit of establishing the duration of each heart event is that it allows you to discover the R top with more precision when compared to the P and T waves, which have a low magnitude and are vulnerable to turbulence. Those uneven time-recurrence coefficients must be processed for the ECG signal in order to describe their morphological characteristics, which are then employed for further investigation. As illustrated in Figure 16, the DCT-DOST segmentation method produced the following results.

The moving average filter is dedicated to removing high-frequency noises from the ECG signal by computing the running mean on the predetermined window length. This is a moderately straightforward estimation that will smoothen both the signal and its anomalies. The R top in the ECG sign is smoothed to around 33% of its unique height. The low-frequency contents of the ECG signal are represented in Figure 17.

The QRS wave of the ECG is detected using zero crossing point detection approach. The dominant and low-frequency contents in the ECG are roughly estimated. Ideally the number of zero crossing points should be low for QRS, while it can be high at other times. The number of zero crossing points is used to determine the QRS with low computational cost. Figure 18 represents the zero crossing output.

The R top in the QRS interim is the most significant component for examining the ECG signal. R top discovery in ECG is a strategy that is generally used to analyze heart anomalies and gauge pulse fluctuation. It is natural that the magnitudes of genuine R tops are more than those for bogus pinnacles. The primary request separation of the sign is
utilized to store the incline data of the genuine pinnacles yet diminishes the slant data of the bogus pinnacles. The proposed strategy can proficiently recognize $R$ tops under different conditions like pattern float, uproarious sign, tall $T$ waves, or a quite delayed waves. Figure 19 represents the peak detection.

To detect ischemia, the slope index is preferred, which outperforms the higher recurrence index model of the bandpass filtered QRS signal as the average relative factor of variation is much higher. The superior performances can be achieved with the slope index when compared with the high recurrence index. This is depicted in Figure 20.

The QRS detection ensures the efficient extraction of beat interval and the abnormalities in the heart function. The improvement in the QRS sections is executed by the proposed technique to eliminate the pattern meandering. In this
paper, the QRS fiducial focuses are detected to perceive the R point using QRS complex so that heart function classification can be accomplished simultaneously. Figure 21 represents the QRS detection.

The RR-interim is resolved to obtain the dynamic qualities of the ECG signal. The 4 RR attributes that are discussed in this paper are pre-RR, post-RR, neighborhood RR, and mean RR interim. The interim between a past R top and the present R top is processed to find the pre-RR attribute, while the interim between a specific R top and the successive R top is estimated to find the post-RR highlight. The combined features of the pre- and post-RR interim represent the momentary cadence characteristics. The mean RR interim features are determined by averaging the RR interims of the previous 3-minimum RR interval of a specific occasion. Figure 22 represents the RR interval.

Similarly, the neighborhood RR features are inferred by averaging all the RR-interims of the previous episodes of a specific occasion [27]. The neighborhood and mean highlights indicate the mean qualities. These 4 highlights are connected to the morphological list of the ECG signal.

The proposed method’s performance is compared with the traditional methods such as CNN and SVM. With a maximum accuracy of 98.5% for different numbers of test samples, our system outperforms the competition [28]. The proposed method’s reliability is guaranteed since its efficiency is consistently high and without compromise. Figure 23 represents the accuracy comparison.

The sensitivity shows the true positive value of the classification. It is calculated as the percentage of positives that are correctly categorised [29]. With a maximum sensitivity of 98.3%, it outperforms the current system, while CNN and SVM have maximum sensitivity of 92 percent and 86 percent, respectively. Figure 24 illustrates the sensitivity relation. Figure 24 represents the sensitivity comparison.

The proposed method’s specificity values change in a zigzag pattern as the number of samples increased [30], with a maximum specificity of 99% for the proposed method and 93 percent and 95.6 percent for CNN and SVM classifiers, respectively [31].

Figure 25 represents the comparison of specificity. The measure of various contents in the ECG signal [32] such as class, sinus rhythm, artifact, ventricular tachycardia, atrial fibrillation, bigeminy, and PVC is computed in terms of R, P, S, and F1. From the comparison table, it is clear that the estimation [33] by the proposed RBFNN is more than the conventional methods. Table 1 represents the accuracy comparison [34].

The training, validation, and testing efficiencies of the proposed method are compared with conventional methods. The training efficiency of our method is much higher than
Figure 23: Accuracy comparison of different methods.

Figure 24: Sensitivity comparison.

Figure 25: Comparison of specificity.
the other methods [35]. Table 2 represents the classification metrics.

From Table 3, the overall $F_1$ score of the proposed method is 90.2%, which is more than the existing methods in which the least performance is shown by the residual method [36].

By concatenating the classification methods, the performance can be improved, which is shown in Table 4.

Table 1: Comparison of aggregate accuracy.

| Model                        | Training (%) | Validation (%) | Test (%) |
|------------------------------|--------------|----------------|----------|
| Baseline—LSTM               | 66.8         | 66.3           | 65.6     |
| Baseline—CNN                | 68.6         | 72.2           | 68.8     |
| Stacked unidirectional—LSTM | 80.5         | 78.1           | 79.2     |
| Stacked bidirectional—LSTM  | 82.2         | 79.5           | 80.2     |
| Stacked unidirectional—LSTM | 80.4         | 79.4           | 79.3     |
| Deep residual—CNN           | 84.7         | 75.3           | 74.7     |
| Combined unidirectional LSTM—CNN | 83.4     | 77.7           | 79.6     |
| Combined bidirectional LSTM—CNN | 93.2     | 74.8           | 76.8     |
| Proposed RBFNN              | 99           | 84.4           | 98.5     |

Table 2: Comparison of classification metrics.

| Class                        | BDLSTM | Residual | LSTM-CNN | Proposed RBFNN |
|------------------------------|--------|----------|----------|----------------|
| Sinus rhythm                 | 0.82   | 0.83     | 0.84     | 0.88           |
| Artifact/noise               | 0.88   | 0.94     | 0.94     | 0.81           |
| Ventricular tachycardia      | 0.16   | 0.51     | 0.26     | 0.92           |
| Atrial fibrillation          | 0.81   | 0.83     | 0.82     | 0.93           |
| Bigeminy                     | 0.72   | 0.65     | 0.82     | 0.98           |
| PVC                          | 0.78   | 0.76     | 0.78     | 0.93           |

Table 3: $F_1$ score class comparison.

| Rhythm class                 | BDLSTM | Residual | LSTM-CNN | Proposed RBFNN |
|------------------------------|--------|----------|----------|----------------|
| Sinus rhythm                 | 0.812  | 0.734    | 0.793    | 0.883          |
| Artifact/noise               | 0.834  | 0.818    | 0.843    | 0.923          |
| Ventricular tachycardia      | 0.265  | 0.169    | 0.417    | 0.721          |
| Atrial fibrillation          | 0.837  | 0.763    | 0.764    | 0.852          |
| Bigeminy                     | 0.663  | 0.136    | 0.553    | 0.754          |
| PVC                          | 0.769  | 0.821    | 0.724    | 0.912          |
| Overall                      | 0.813  | 0.728    | 0.742    | 0.902          |

Table 4: $F_1$ score class comparison.

| Rhythm class                 | Multi | Single | Multi | Single | Multi | Single | Multi | Single |
|------------------------------|-------|--------|-------|--------|-------|--------|-------|--------|
| Sinus rhythm                 | 0.812 | 0.612  | 0.734 | 0.692  | 0.793 | 0.702  | 0.883 | 0.813  |
| Artifact/noise               | 0.834 | 0.734  | 0.818 | 0.746  | 0.843 | 0.774  | 0.923 | 0.874  |
| Ventricular tachycardia      | 0.265 | 0.065  | 0.169 | 0.085  | 0.417 | 0.145  | 0.721 | 0.835  |
| Atrial fibrillation          | 0.837 | 0.337  | 0.763 | 0.797  | 0.764 | 0.717  | 0.852 | 0.857  |
| Bigeminy                     | 0.663 | 0.263  | 0.136 | 0.073  | 0.553 | 0.523  | 0.754 | 0.873  |
| PVC                          | 0.769 | 0.669  | 0.821 | 0.709  | 0.724 | 0.709  | 0.912 | 0.879  |

4. Conclusion

Our proposed work enhances the diagnosis accuracy by eliminating the redundant and noise highlights. The algorithm presented here provides sensitivity and accuracy above 98.5%. These are computationally facile algorithms that can be applied for practical application and aids in processing of a massive set of databases. By this work, the
objective gets achieved and the artifacts can be detected by analogizing with the results from the algorithm for additional analysis. It gives better acknowledgement exactness when compared with various other existing frameworks. In defining the path of development of arrhythmia detectors, the experience of the last 20 years, since the release of the MIT-BIH Arrhythmia Database and the AHA Database soon afterwards, may be considered a big experiment in shaping the direction of development. Performance statistics were of little or no value until databases became available, because it was widely accepted that each manufacturer designed its products using its own data and designed its statistics to present the products in a favourable light, so performance statistics were of little or no value until databases became available. The temptation to match their rivals' products feature for feature, rather of investing time and money on enhancements that could not be measured and hence did not add perceived value to the product, challenged conscientious developers who attempted to increase the accuracy of their algorithms.

Was the experiment a success or a failure, and what were the outcomes? The introduction of databases in the early 1980s signaled a sea shift in the development community’s efforts. End users and regulatory bodies started to inquire of manufacturers about the effectiveness of their gadgets when subjected to conventional testing. Manufacturing companies could not avoid doing the tests and reporting the results, and those whose algorithms did not perform as well as their rivals spent their development expenditures on focused attempts to enhance performance.

As a result of the availability of databases, the overall level of performance of commercial arrhythmia detectors has increased significantly in recent years.

Although it would be incorrect to suggest that manufacturers were capable of producing significantly better products in the late 1970s and instead chose to add bells and whistles in response to their customers’ apparent lack of interest in performance, it is possible to infer this from the fact that manufacturers were capable of producing significantly better products in the late 1970s.

Data Availability

The data that support the findings of this study are available upon request from the corresponding author.

Conflicts of Interest

The authors declare that they do not have any conflicts of interest regarding the publication of this paper.

Acknowledgments

The authors extend their appreciation to the researchers supporting project number (RSP-2021/384), King Saud University, Riyadh, Saudi Arabia.

References

[1] Y. Miao, Y. Tian, L. Peng, M. S. Hossain, and G. Muhammad, “Research and implementation of ECG-based biological recognition parallelization,” IEEE Access, vol. 18, no. 3, pp. 161–164, 2018.
[2] L. Sun, Y. Lu, K. Yang, and S. Li, “ECG analysis using multiple instance learning for myocardial infarction detection,” IEEE Transactions on Biomedical Engineering, vol. 59, no. 12, pp. 3348–3356, 2012.
[3] G. Da Poian, R. Bernardini, and R. Rinaldo, “Separation and analysis of fetal-ECG signals from compressed sensed abdominal ECG recordings,” IEEE Transactions on Biomedical Engineering, vol. 63, no. 6, pp. 1269–1279, 2016.
[4] N. Bayasi, T. Tekeste, H. Saleh, B. Mohammad, A. Khandoker, and M. Ismail, “low-power ECG-based processor for predicting ventricular arrhythmia,” IEEE Transactions on Very Large Scale Integration Systems, vol. 24, no. 5, pp. 1962–1974, 2016.
[5] X. Zhai and C. Tin, “Automated ECG classification using dual heartbeat coupling based on convolutional neural network,” IEEE Access, vol. 6, no. 5, pp. 27466–27472, 2018.
[6] M. J. Mollakazemi, F. Asadi, H. E. Orimi, S. A. Atyabi, I. Uzelac, and A. Ghaffari, “Estimation of extent damage tissue by multi resolution analysis of the electrocardiogram and arterial blood pressure,” in Proceedings of the 2015 Computing in Cardiology Conference (Cinc), vol. 42, no. 9, pp. 1113–1116, Nice, France, September 2015.
[7] R. He, Y. Liu, K. Wang et al., “Automatic cardiac arrhythmia classification using combination of deep residual network and bidirectional LSTM,” IEEE Access, vol. 7, no. 7, pp. 102119–102135, 2019.
[8] S. K. Yadav, R. Sinha, and P. K. Bora, “Electrocardiogram signal denoising using non-local wavelet transform domain filtering,” IET Signal Processing, vol. 9, no. 2, pp. 88–96, 2015.
[9] H. D. Hesar and M. Mohhebbi, “A multi rate marginalized particle extended kalman filter for P and T wave segmentation in ECG signals,” IEEE Journal of Biomedical and Health Informatics, vol. 23, no. 1, pp. 112–122, 2019.
[10] S. Xu, J. Li, K. Liu, and L. Wu, “A parallel GRU recurrent network model and its application to multi-channel time-varying signal classification,” IEEE Access, vol. 7, no. 8, pp. 118739–118748, 2019.
[11] S. K. Bashar, E. Ding, A. J. Walkey, D. D. McManus, and K. H. Chon, “Noise detection in electrocardiogram signals for intensive care unit patients,” IEEE Access, vol. 7, no. 7, pp. 88357–88368, 2019.
[12] H. Liu, D. Chen, and G. Sun, “Detection of fetal ECG R wave from single-lead abdominal ECG using a combination of RR time-series smoothing and template-matching approach,” IEEE Access, vol. 7, no. 5, pp. 66633–66643, 2019.
[13] S. Jain, V. Bajaj, and A. Kumar, “Riemann liouville fractional integral based empirical mode decomposition for ECG denoising,” IEEE Journal of Biomedical and Health Informatics, vol. 22, no. 4, pp. 1133–1139, 2018.
[14] J. Wang, M. She, S. Nahavandi, and A. Kouzani, “Human identification from ECG signals via sparse representation of local segments,” IEEE Signal Processing Letters, vol. 20, no. 10, pp. 937–940, 2013.
[15] F. Khosrow-Khavar, K. Tavakolian, A. Blaber, and C. Menon, “Automatic and robust delineation of the fiducial points of the seismocardiogram signal for noninvasive estimation of cardiac time intervals,” IEEE Transactions on Biomedical Engineering, vol. 64, no. 8, pp. 1701–1710, 2017.
F. Alonso-Atienza, E. Morgado, L. Fernandez-Martinez, and A. Garcia-Alberola, J. L. Rojo-Alvarez, Detection of life-threatening arrhythmias using feature selection and support vector machines,” *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 3, pp. 832–840, 2014.

S. Raj and K. C. Ray, "ECG signal analysis using DCT-based DOST and PSO optimized SVM,” *IEEE Transactions on Instrumentation and Measurement*, vol. 66, no. 3, pp. 470–478, 2017.

S. R. Moosavi, E. Nigussie, M. Levorato, S. Virtanen, and J. Isoaho, “Low-latency approach for secure ECG feature based cryptographic key generation,” *IEEE Access*, vol. 6, no. 2, pp. 428–442, 2018.

A. O. Bicen, N. Z. Gurel, A. Dorier, and O. T. Inan, “Improved pre-ejection period estimation from ballistocardiogram and electrocardiogram signals by fusing multiple timing interval features,” *IEEE Sensors Journal*, vol. 17, no. 7, pp. 4172–4180, 2017.

U. Satija, B. Ramkumar, and M. S. Manikandan, "A new automated signal quality-aware ECG beat classification method for unsupervised ECG diagnosis environments,” *IEEE Sensors Journal*, vol. 19, no. 1, pp. 277–286, 2019.

Q. A. Rahman, L. G. Tereshchenko, M. Kongkatong, T. Abraham, M. R. Abraham, and H. Shatky, “Utilizing ECG-based heartbeat classification for hypertrophic cardiomyopathy identification,” *IEEE Transactions on Nanobioscience*, vol. 14, no. 7, pp. 505–513, 2018.

R. G. Sutar and A. G. Kothari, “Intelligent electrocardiogram pattern classification and recognition using low-cost cardiology system,” *IET Science, Measurement & Technology*, vol. 9, no. 5, pp. 134–143, 2015.

H. E. Saadawy, M. Tantawi, H. A. Shedeed, and M. F. Tolba, “Hybrid hierarchetical method for electrocardiogram heartbeat classification,” *IET Signal Processing*, vol. 12, no. 5, pp. 506–513, 2018.

P. K. Shukla, M. Zakariah, W. A. Hatamleh, H. Tarazi, and B. Tiwari, “AI-DRIVEN novel approach for liver cancer screening and prediction using cascaded fully convolutional neural network,” *Journal of Healthcare Engineering*, vol. 2022, Article ID 4277436, 14 pages, 2022.

P. K. Shukla, V. Roy, P. K. Shukla et al., "An Advanced EEG Motion Artifacts Eradication Algorithm,” The Computer Journal, p. bxab170, 2021.

N. K. Rathore, N. K. Jain, P. K. Shukla, U. Rawat, and R. Dubey, “Image forgery detection using singular value decomposition with some attacks,” *National Academy Science Letters*, vol. 44, no. 4, pp. 331–338, 2021.

A. Kumar, M. Saini, N. Gupta et al., “Efficient stochastic model for operational availability optimization of cooling tower using metaheuristic algorithms,” *IEEE Access*, vol. 10, p. 1, 2022, https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9682734&isnumber=6514899.

D. Singh, V. Kumar, M. Kaur, M. Y. Jarabulla, and H.-N. Lee, "Screening of COVID-19 suspected subjects using multi-crossover genetic algorithm based dense convolutional neural network,” *IEEE Access*, vol. 9, pp. 142566–142580, 2021.

H. Kaushik, D. Singh, M. Kaur, H. Alshazly, A. Zaguia, and H. Hamam, "Diabetic retinopathy diagnosis from fundus images using stacked generalization of deep models,” *IEEE Access*, vol. 9, pp. 108276–108292, 2021.

D. B. Saadi, G. Tanve, M. Flintring et al., “Automatic real-time embedded QRS complex detection for a novel patch-type electrocardiogram recorder,” *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 59, no. 12, pp. 3348–3356, 2015.

C. Venkatesan, P. Karthigaikumar, A. Paul, S. Satheekumaran, and R. Kumar, “ECG signal preprocessing and SVM classifier-based abnormality detection in remote healthcare applications,” *IEEE Access*, vol. 6, no. 1, pp. 9767–9773, 2018.

Y. Park, D. YunII, and S.-H. Kang, "Preprocessing method for performance enhancement in CNN-based STEMI detection from 12-lead ECG,” *IEEE Access*, vol. 7, no. 7, pp. 99964–99977, 2019.

W. Zhu, X. Chen, Y. Wang, and L. Wang, "Arrhythmia recognition and classification using ECG morphology and segment feature analysis,” *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 16, no. 2, pp. 131–138, 2019.

S. Kiranyaz, T. Ince, and M. Gabbouj, “Real-time patient-specific ECG classification by 1-d convolutional neural networks,” *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 3, pp. 664–675, 2016.

L. Han, X. Pu, and X. Chen, "Method of fetal electrocardiogram extraction based on v-support vector regression,” *IET Signal Processing*, vol. 9, no. 7, pp. 430–439, 2015.

A. Smruthy and M. Suchetha, “Real-time classification of healthy and apnea subjects using ECG signals with variational mode decomposition,” *IEEE Sensors Journal*, vol. 17, no. 5, pp. 3092–3099, 2017.