International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 5
DOI: https://doi.org/10.17762/ijritcc.v10i5.5542
***Article Received**: 10 March 2022 **Revised**: 22 March 2022 **Accepted**: 25 April 2022 **Publication**: 31 May 2022*

Identifying Arrhythmias Based on ECG Classification Using Enhanced-PCA and Enhanced-SVM Methods

Akhil Mathew Philip[1], Dr. S Hemalatha[2]
[1]Research Scholar, Karpagam Academy of Higher Education (Deemed to be University),
Coimbatore, India
[2]Research Guide, Karpagam Academy of Higher Education (Deemed to be University),
Coimbatore, India

Abstract: The "Cardio Vascular Diseases (CVDs)" had already attained worrisome proportions in both advanced and emerging nations in recent times. Physically inactive behaviors, altered eating, and occupational routines, and reduced daily fitness were all recognized as crucial contextual elements, in addition to genetics. Considering CVDs have such a significant morbidity and mortality, accurate and early diagnosis of cardiac disease by "ElectroCardioGram (ECG)" allows clinicians to decide suitable therapy for a multitude of cardiovascular disorders. The interpretation of ECG signal is an important bio-signal processing area that involves the application of computer science and engineering to detect and visualize the functional status of the heart. Therefore, in the present work, a detailed study on ECG signals denoising and abnormalities detection using different techniques were performed. Annoying distortions and noisy particles are common in ECG signals. The "Biased Finite Impulse Response (BFIR)" preprocessing filtering is employed in this research to eliminate the noises in the raw ECG signals. The "Nonlinear-Hamilton" segmentation method is employed to segment the 'R' peak signals. To decrease the extraneous features included in the segmented ECG data, the innovative "Enhanced Principal Component Analysis (EPCA)" was applied for feature extraction. A unique "Enhanced version of the Support Vector Machine (ESVM)" framework with a "Weighting Kernel" based technique is proposed for classifying the ECG data. The 'Q', 'R', and 'S' waves in the given ECG data will be identified by this framework, allowing it to characterize the cardiac rhythm. The evaluation metrics of the EPCA-ESVM proposed method is comparatively analyzed with our previous approach EPSO. To estimate the results for the dataset from MIT-BIH it was experimented with by the EPSO and the EPCA-ESVM methods focused upon different parameters such as Accuracy, F1-score, etc. The final findings of the EPCA-ESVM method were good than the EPSO method in which the accuracy is higher even though unbalanced data were present.

Keywords: Cardio Vascular Diseases, ECG, Enhanced Principal Component Analysis, Enhanced Support Vector Machine.

1. INTRODUCTION

Globally, CVDs and stroke are the leading causes of mortality, as per the "World Health Organization (WHO)". CVDs resulted in the mortality of approximately 20.3 million individuals in 2019, accounting for 33% of all fatalities globally. Heart health studies have become more important for medicinal scientists, particularly those who are interested in technical, preventative, or therapeutic breakthroughs in the field of CVDs. As a result, experts recently focused their attention on conventional methods of cardiovascular assessment deployed in the household, health centers, and emergency rooms [1].

Many heart disorders may be detected using a basic, affordable, and risk-free ECG assessment, which was a much more frequent medical cardiac assessment. There have been a lot of studies done on ECG interpretation in a previous couple of decades. The majority of ECG signal reflects a crucial indicator for cardiac functioning evaluation since it shows the electrical events that occur with a trigeminal and a trigeminal-atrioventricular rhythm. In this way, the ECG offers adequate data about a patient's cardiac condition [2].

The CVDs include a multitude of illnesses distinguished by abnormalities in the heart's electrical activity, such as rapidity or slowness, or indeed waveform deformity. ECG waveform analysis may identify any abnormality, whether it is a change in heartbeat or morphology structure, that may point to a pathological condition. Medical professionals who use ECG signals for medical processing and analysis take a lengthy time frame and are often impractical or impossible in distant places, such as those requiring long-term surveillance. As a result, albeit difficult for real-time ECG processing, an "Automated Arrhythmia Beat Classification" is critically needed [3].

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 5
DOI: https://doi.org/10.17762/ijritcc.v10i5.5542
Article Received: 10 March 2022 Revised: 22 March 2022 Accepted: 25 April 2022 Publication: 31 May 2022

Segmentation, extraction of features, and Classification process are the three primary phases in abnormality beat categorization. Although successful and consistent classifications depend on good feature approximations, extraction of features from ECG signals seems to be an essential and preparatory process [4]. Measurements taken from a heartbeat cycle may be employed to identify the sort of ECG features it contains. There should be an expectation that features reflect sequences so that important details are not discarded. To find alterations in the spectral of ECG waves, investigations are often done whether in "Time-Domain" or the "Frequency-Domain" [5]. Alternatively, they may be employed in the "Time-Frequency" domains also to represent morphology and spectrum features concurrently. Research in this area has led to the development of a wide range of feature extraction methods [6].

To minimize the abundance of details in a huge collection of linked variables to a manageable number of "Principal Components (PCs)", the statistical approach of "Principal Component Analysis (PCA)" is used. As a Linear-Function of a set of the items in the dataset, PCs were created with weights selected such that the PCs are not connected at all. The primary several components make up a large portion of the variability in the data collection, and each one of them adds additional information about it. A collection of time samples, instead of a set of variables, is used for PCA in "Time & Frequency" domains. While evaluating a signal that occurs repeatedly, such as ECG, the samples are often taken from the original segment position at several points in the signal's life cycle.

A broad range of cardiac disorders may be better diagnosed with the use of signal analysis, which can be encountered in almost any ECG analysis system these days [7]. Like Quantization, Beat-Recognition and Beat-Categorization, Distortion-Removal, Signal-Segmentation, and extraction of features are a few of the many applications for signal analysis in ECG assessment. That most of these concerns have been effectively addressed using the PCA, which was originally developed for the aim of searching and accessing ECGs more efficiently. Whereas the motivating factor behind these studies has shifted from small storage devices to inefficient communication boundaries [8], this problem has continued critical throughout the decades.

A low-noise subset of eigenvectors is often used in the restoration of the source signal, suggesting that noise removal and data reduction go hand in hand. Such a decrease, on the other hand, is most beneficial for noisy signals that originate in muscle contraction. To classify the shape of waveforms with arrhythmia tracking, PCA was first used to discriminate across normal heart-beats and aberrant-beats waveforms, such as "Premature-Ventricular-Beats (PVBs)" [9].

The reliable extraction of features from multiple waveform components for the goal of detecting temporal changes caused by Myocardial-Ischemia is a new trend of PCA in ECG signal processing. Local measures from the "ST-T" segment have traditionally been used for this tracking but these observations are inaccurate when the examined signal seems distorted. By using correlation as the primary signal analysis procedure, it has also become obvious that the "ST-T" segment may be better classified using PCs. With "Atrial-Fibrillation", PCA has recently been used to separate atrial function from the ventricular activity so that the arrhythmia's features may be examined without distraction from ventricular function. Due to their distinct bioelectrical origins, these functions could be separated using temporal and spatial redundancy while evaluating multi-lead records [10].

According to the research's problem statement, it is discovered that ECG data augmentation has been thoroughly investigated. The challenge of extracting features, on the other hand, has received only minimal attention. Classifiers receive support from having a high collection of features to build a classifying framework with a thorough understanding of the training data. A huge range of features, on the other hand, raises the computing burden. An effective feature array would have fewer redundant items that represent a signal's most important features. Even though, there is a further concern that is sometimes discussed in the academic literature. As a rule, several ECG feature extraction algorithms employ classifications to assess the feature effectiveness, either by selecting a random sample of heartbeats from the dataset for training and testing or by selecting a certain percentage for every class to serve as training and testing sets. It is important to highlight, therefore, that all this classification technique is not a practical measurable statistic in practical uses of heartbeat detection.

The prime contribution of this research article is to develop a unique enhancement process for established PCA and SVM approaches that uses a composite of actual and synthesized heartbeats to enhance the categorization of ECG heartbeats from the MIT-BIH arrhythmias database. Preprocessing, segmentation, extraction of features, and classifying are typical components of a computerized ECG classifying framework. We go into great detail on extracting ECG optimal features and ECG classification in this article. We develop an EPCA to retrieve the necessary features from the segmented ECG data for extracting features. In an attempt to classify EPCA characteristics as normal or abnormal, we develop an ESVM classification approach.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 5
DOI: https://doi.org/10.17762/ijritcc.v10i5.5542
***Article Received**: 10 March 2022 **Revised**: 22 March 2022 **Accepted**: 25 April 2022 **Publication**: 31 May 2022*

The following is a summary of the remaining article: A few recent publications on ECG classifiers research are reviewed in Section 2, followed by Section 3 covers the description of the proposed methodology module-by-module and the existing approach, Section 4 deals with the comparison of the results, and Section 5 finish up with the conclusion of this research article with suggestions for future improvement.

2. RELATED WORKS

The "Google-Inception V3" architecture relying on the "ImageNet" database was utilized by the researchers in [11] to update the CNN classifier, which recognized ischemia abnormalities in ECG. They were able to categorize ECG participants with 97 percent accuracy.

To characterize the ECG waveform, the researchers in [12] used "AlexNet" weighted indexing that had been performed on the "ImageNet" and provided with training "2D-CNN". Transferred-Learning algorithms are used to distinct "Neural-Networks" in an attempt to bring out eventual categorization in their framework.

In [13] the researchers recommended a "Neural-Network" framework for exact categorization of heartbeats based on "AAMI inter-patient" criteria. There are 2 phases in this process. Features are retrieved from waveform in the initial phase of preprocessing. Throughout stage 2, a two-layer classification model is used, also every layer consists of two fully-connected Neural-Networks. The developed framework can accurately recognize arrhythmia situations, as shown in their studies.

The deep training framework presented by the researchers in [14] includes a CNN and an LSTM. Classifying six categories of ECG signals from the MIT-BIH arrhythmias database, this framework processes 10 secs of ECG segments. An arrhythmia detection methodology developed in this paper has been tested and found to be useful by cardiologists.

A CNN-based framework for the correct diagnosis of Ischemic-Heart disease through ECG was described by the researchers in [15]. The presented framework was validated and trained using publicly accessible ECG samples. The developed model's ability to accurately predict heart problems is shown by its effectiveness.

3. METHODOLOGIES

3.1 ECG SIGNALS FROM MIT-BH

Together, in summary, the ECG readings are a series of voltage levels sampled on a specific portion of the human anatomy at a frequency of hundreds of times per second in a standardization form. Various leads are used to understand the attributes collected from numerous parts of the body. There are often multiple cardiac phases in a normal ECG recording, these may be subdivided into the individual beats of the body. A sample ECG recording is shown in Figure 1.

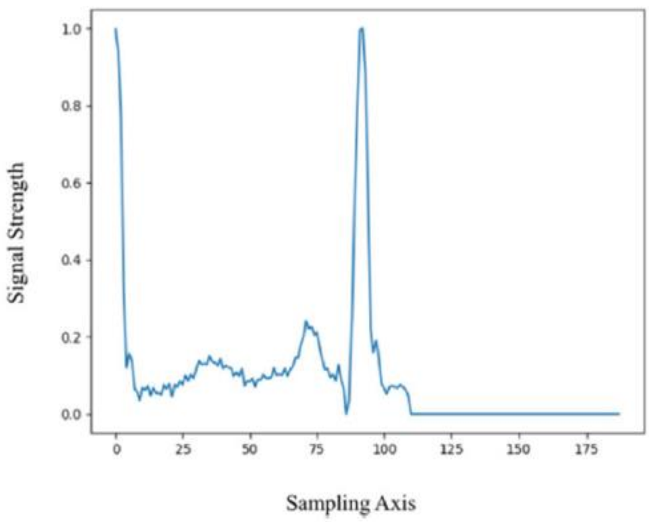

FIGURE 1: SAMPLE ECG SIGNAL

According to the characteristics available in an ECG, a classifying model's function is to determine the proper kind of heartbeat. Let's assume that 's' is an ECG recording. We have to identify a function called 'H', which takes 's' as an argument, in the perspective of classification. Because it's capable of returning a result of the form "h = H(s)," it allows us to collect as several accurate findings as appropriate from actual ECG samples.

3.2 ECG SIGNAL PREPROCESSING

The noise in the unprocessed ECG readings is common. An inconsistent extracted feature is induced by the

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 5
DOI: https://doi.org/10.17762/ijritcc.v10i5.5542
***Article Received**: 10 March 2022 **Revised**: 22 March 2022 **Accepted**: 25 April 2022 **Publication**: 31 May 2022*

presence of higher frequencies. To normalize the signal, a preprocessing methodology is necessary. Under this research, we employ the "Biased Finite Impulse Response (BFIR)" preprocessing approach. Compared to the usual polynomial extrapolation technique, it gives superior noise removal and superior extraction of features in terms of signal-to-noise ratio. As a result, detecting the durations and amplitudes of the 5 waves per cardiac-cycle, the "P-wave", the "QRS-complex", and the "T-wave", becomes simpler.

Regarding signal preprocessing, we employ the BFIR, specifically the impulse response, which seems to be the outcome in reactions to the original signal's Kronecker delta. As a result of this, it does have an ordering of 'N' and seems to last "N + 1" data points until it reaches zero. That each data point within outcome sequences is a weighted combination among the most latest entries in a BFIR filter of order 'N' as per Equation 1:

$$y[n] = \sum_{i=0}^{N} b_i x[n-i]$$

Eq→1

As shown in Figure 2, the signals have been preprocessed. Prior and following the BFIR filter preprocessing for per cardiac-cycle. A lot of the distortion in a signal could well be reduced by preprocessing. Voltage samples are represented by microvolts on the Vertical-Axis of the signal, while time is represented by milliseconds on the Horizontal-Axis.

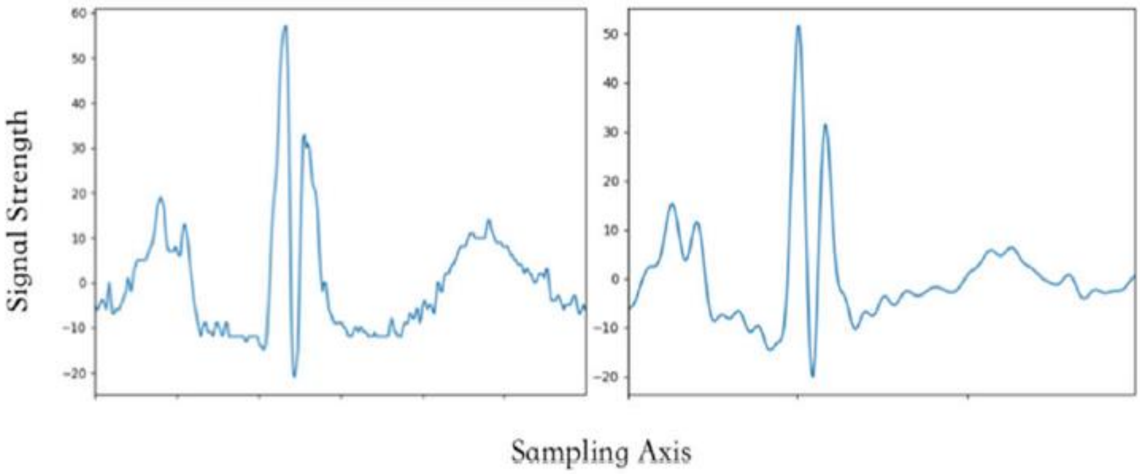

FIGURE 2: LEFT (RAW SIGNAL) RIGHT (PREPROCESSED SIGNAL)

3.3 ECG SIGNAL SEGMENTATION

Although heartbeats might vary from individual to individual, maybe from each cardiac-cycle to another, the overall pattern of heartbeats is consistent. To segment heartbeats, specialists may use this patterning in the cycle morphology. The "P-wave", the "QRS-complex", the "T-wave", the 2-segments such as "PR-segment" and the "ST-segment", and the 2-intervals such as "PR-interval" and the "QT-interval", make up a typical cardiac-cycle morphologically as shown in Figure 3. The "P, Q, R, S, and T" are the five periods of a cardiac-cycle. Voltage estimated in microvolts is shown on the Vertical-Axis, while time is shown on the Horizontal-Axis.

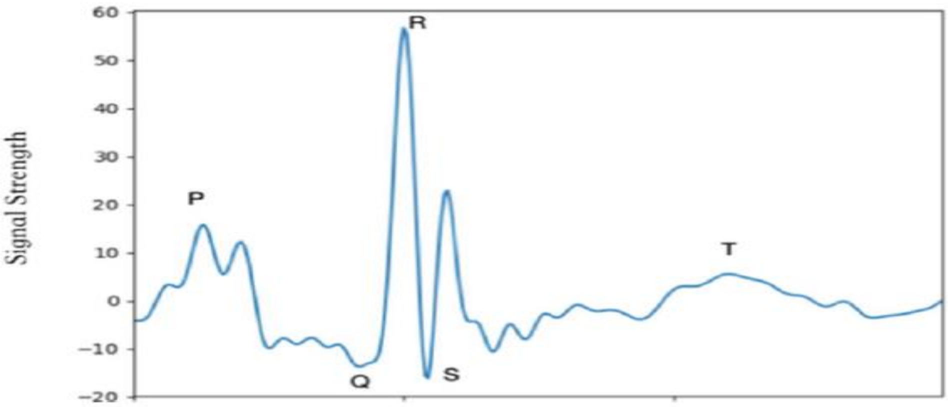

Figure 3: Anatomy of a cardiac-cycle

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 5
DOI: https://doi.org/10.17762/ijritcc.v10i5.5542
***Article Received**: 10 March 2022 **Revised**: 22 March 2022 **Accepted**: 25 April 2022 **Publication**: 31 May 2022*

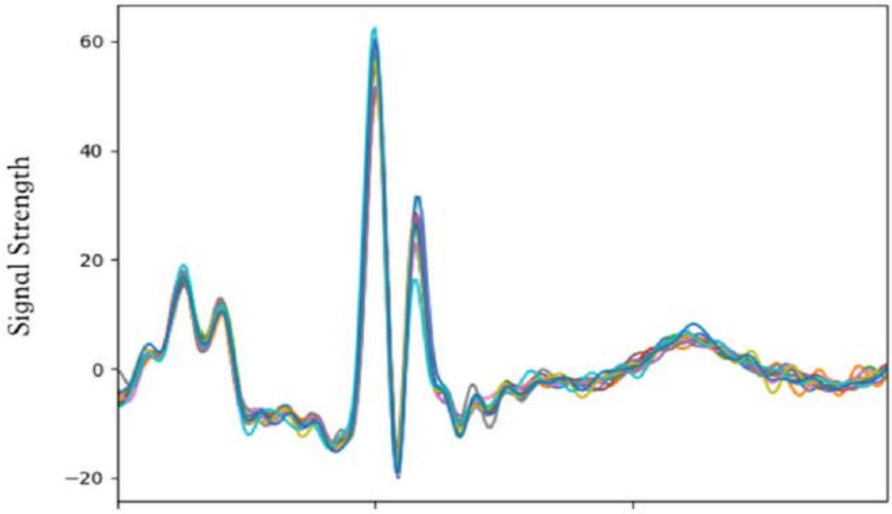

Figure 4: Output of cardiac segmentation

For segmenting the "R-Peaks" from the ECG signal obtained in this research, we used a "Nonlinear-Hamilton" segmentation technique. For the cardiac-cycle segmentation, a mixed "Adaptive-Threshold" is used to identify the "QRS-Complex". The sample that was shown in Figure 4 has been segmented output. The ECG samples are segmented into heartbeat cycles, which are shown in Figure 4 by the various colors of the coordinated cycles of one sample.

3.4 ECG FEATURE EXTRACTION

The PCA is a technique for reducing the dimension of a collection of data while maintaining its heterogeneity. Almost every collection of ECG data includes information expressed as vectors of single parameters with integer, binary, or real values in most cases. A geometric point in three-dimensional space, e.g., may be described by a vector of three parameters each of which is aligned with one of the three coordinates axis 'x', 'y', and 'z'. An ECG sample could be described in particular by a vector made up of a set of parameters. The length of the vectors found in the collection, and therefore the set's dimension, is determined by the number of parameters. Furthermore, a continuum of variability may be specified for each parameter, that defines the range of values that only the individual parameter could accept. For example, if the data set includes three-dimensional points delimited by a cube with side 1 and centered in (0, 0, 0), the three parameters describing Cartesian coordinates were bound to [-1/2, 1/2]. The spectrum of heterogeneity of the three parameters is described by this interval. The objective of PCA is always to uncover hidden patterns in data and turn them in just such a manner that its differences and similarities are exposed. When the patterns have been discovered, the data may be interpreted as components that are sorted by importance, allowing low-level components to be discarded without losing valuable details.

ENHANCED PCA:

When the actual parameters were chosen to describe the data that are associated, the traditional PCA will minimize the dimension of a collection of data. Now let reconsider the illustration of the three ECG parameters describing the three positions of ECG signal points inside a cube by EPCA. The three ECG parameters are associated if, for example, any of the points in the collection fall on an appropriate plane.

As EPCA has been used to solve this issue, one of the three parameters is transformed into a void parameter. The points in the fresh transformed space could consequently be defined by just two parameters, resulting in a space with a smaller dimension than the first. Since the points are in a two-dimensional region, the details about the third dimension, which is the rejected dimension, are meaningless. This is an oversimplification of the case. The explanations that follow go into the EPCA process in greater depth.

Imply that perhaps the ECG data collection under consideration includes points in a two-dimensional region with coordinates of (-2, -1), (-1, 0), (0, 1), (1, 2), (2, 3). The values of 'x' differ in the range [-2, 2], whereas the values of y vary in the range [-1, 3]. The variation of the parameters 'x' and 'y' is described by these two intervals. These two factors are associated, as can be shown. As the 'y' coordinates rise, the 'x' coordinates rise as well, and a straight line runs between them all. As a result, if one of the two coordinates is identified, the other may be obtained.

Whereas traditional PCA aims to convert certain parameters so that they are no longer associated. By

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 5
DOI: https://doi.org/10.17762/ijritcc.v10i5.5542
***Article Received**: 10 March 2022 **Revised**: 22 March 2022 **Accepted**: 25 April 2022 **Publication**: 31 May 2022*

achieving this, the dimension of the collection of data could be minimized by only considering the parameters with the highest uncertainty and discarding the others. Principal-Components (PC) are still the factors with the most variability. Here in EPCA, the first PC of the ECG data could have been considered to reflect the results since they are normally ordered by their heterogeneity. The fact that perhaps a lower order PC exhibits lower variance within the ensemble however does not mean this is irrelevant in regression models. Through EPCA it is possible to find out the extra ECG features that are omitted by the traditional PCA.

Extraction of ECG Statistical Features through EPCA

EPCA is most commonly used for locating a lower-dimensional representation of ECG data. That has 2 distinguishing characteristics. During computing, it initially keeps shrinking the measurements of the provided ECG data to a rational and accurate scale. Foremost, it separates the number of distinguishing ECG features from the segmented ECG input in a quiet manner that the overall dimension is reduced. The important feature characteristics were always present and could be used to identify the actual ECG input details.

The covariance-matrix could be found again from the matrix through leveraging the collection of optimal ECG features. The Eigen-values are then calculated using this covariance-matrix. Eigen-vectors were effective in representing complete ECG databases in their nature. Merely some small Eigen-values were considered toward being substantially preferable and greater in importance, whereas the others are substantially quite minimal, and its exposure to data variations is indeed quite limited. As a result, after computing the inner product of the ECG data well with respective Eigen-vectors for its respective Eigen-values, the preferred and greater variance paths were simply maintained in this proposed EPCS.

The following are the general measures of EPCA methodology:

Step-1: Compute the covariance-matrix of the specified ECG input from segmented ECG signal using the following Equation 2:

$$\sum V = \frac{1}{Num}\left\{\left(diag(m) - \overline{diag}(m)\right)\left(diag(n) - \overline{diag}(n)\right)^T\right\}$$

Eq→2

Where, $1 \leq m, n \leq Num$

Step-2: As a result of the Eigen-vector matrix (V) and diagonal-matrix (D) of computed ECG Eigen-values are:

$$V^{-1} \sum V = D \qquad \textbf{Eq→3}$$

Step-3: To achieve the PC parameter, organize the Eigen-vectors in decreasing order with the accompanying magnitude of ECG Eigen-values.

Step-4: At last, ECG data is transformed in the form of PCs by measuring the inner product of data with relevant Eigen-vectors.

Through specific, the EPCA of a given vector ‘v’ associated with the group ‘V’ is achieved by mapping vector v onto the subspaces with the length or gaps of corresponding e' Eigen-vectors that relate to the top e' Eigen-values of the auto-correlation matrix ‘R’ in downwards sequence, where e' is lower than e. The above transformation generates a vector of e' coefficients c1,..., ce'. Even so, a linear structure of the Eigen-vectors with its corresponding weights c1,..., ce' is used to describe the vector v.

Table 1: Extracted Features from the Segmented signals

Leads/ Feature Type	Peak Value and Duration
I	Segment P, Q, R, S, and T
II	Segment P, Q, R, S, and T
V1	Segment P, Q, R, S, and T
V2	Segment P, Q, R, S, and T
V3	Segment P, Q, R, S, and T
V4	Segment P, Q, R, S, and T
V5	Segment P, Q, R, S, and T
V6	Segment P, Q, R, S, and T

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 5
DOI: https://doi.org/10.17762/ijritcc.v10i5.5542
Article Received: 10 March 2022 Revised: 22 March 2022 Accepted: 25 April 2022 Publication: 31 May 2022

Thus the statistical features from the segmented ECG signal are derived efficiently by our proposed EPCA method. Using such techniques, we may determine the maximum values of each wave and the length of each phase to which they belong. Because the ECG sample we used has eight separate leads, the feature we were able to extract from a single sample had a total of 80 dimensions. Feature extractions output is included in Table 1.

3.5 ECG CLASSIFICATION

3.5.1 EXISTING METHOD (EPSO)

The classifying work is dependent on a reconfigured EPSO model training. The features utilized for classification are the inputs, and the six-beat classifications analyzed are the outputs.'R' and 'T' waves, amplitudes ("R-ampli", "T-ampli"), QRS-complex and T-wave duration ("QRS-dura", "T-dura"), "ST" and "QT" segments, and "RR" intervals such as "RR_p" and "RR_f", which have been measured by calculating the distance among both the present and the prior 'R' peak and also the distance among both the present and also the next R-peak, mostly between, are among the morphological extracted features in this research. The ratios of the "RR" intervals ($ratio_1 = RR_p/RR_f$) and the 'R' peak amplitude to the "QRS-complex" duration ($ratio_2$ = R-ampl/QRS-dur). The EPSO method ($w = 0.2$, $c_1 = c_2 = 2$) is used to explore the number of elements in the solution space. The fitness value is calculated as the difference among direct and immediate outcomes well before the training period in this stage. Initially, a population of 50 particles was evaluated. The frequencies of elements in the search space are represented by the particles ("n_H" varies from 1 to 50). For each vector input, the process is performed 100 times. The optimized range of elements in the search area seems to be equivalent to 12 for both the vector consisting of 6 morphological features [R-ampl, QRS-dur, RR_p, RR_f, $ratio_1$, $ratio_2$] and the least value of error has been recognized for the vector consisting of 6 morphological features [R-ampl, QRS-dur, RR_p, RR_f, $ratio_1$, $ratio_2$]. This was simple to determine if the input ECG signal was normal or abnormal based on these numbers.

3.5.2 PROPOSED METHOD (ESVM)

The SVM has been the more commonly employed pattern classification-based machine learning method. Vapnik first launched the SVM during the year 1995. This approach is mostly known as mathematical learning techniques, and it creates a medium for retrieving facts, making forecasts, and making a decision for a situation. This supports the selection of the appropriate hyperplane for the ultimate purpose. The core principle underlying the whole strategy is to use various forms of kernel functions to portray the nonlinearly separable sample onto a higher-dimensional region. This is a discriminating classifier is represented either by separating hyper plane in the feature space that is controlled by a kernel function throughout the dimensional space.

Mostly in the context of high-dimensional content of ECG features, trying to isolate the exact ECG feature vectors towards the appropriate information is impossible for traditional SVM. As a consequence, we propose an Enhanced based SVM (ESVM) which employs a weighted kernel method to apply a nonlinear separation strategy, that resulted in the translation of the actual extracted features into high dimensional, and the projection of all these dimensions leads to effective separation.

The major difference between traditional SVM and ESVM is that the computational distance would be affected by traditional SVM training and identification whereas the ESVM utilizes the weighted kernel feature to train the ECG features and statistically learned parameters for the classification process. Such iterations may be performed in a sequence that also increases the method's convenience of usage while dealing with issues including huge ECG features in the testing and a broad range of optimal ECG features mostly in vector.

ESVM Process

It is a binary-based classification that could, furthermore, have been used to solve the various classification tasks that are often encountered in this analysis for particular arrhythmia from the ECG signals. The ESVM classification model receives the chosen feature subset from training samples and obtained testing features as classification data. During the training phase, two groups were identified: the cardiac diseased class (c1) and the healthy class (c2).

ESVM Algorithm:

(i) The training data and also its labeling was provided as feedback during the training process.

(ii) The training samples would then be normalized by subtracting the learned data's average significance.

(iii) At last, the trained significance is determined by finding the determination value for every label.

(iv) The following process is used to measure the decision function.

$$d(v) = w_i L_i v_i^T v + n \qquad \textbf{Eq}\rightarrow\textbf{4}$$

(v) The test data has been normalized and also the margin value is calculated.

(vi) By using the equation below, such decision values were compared to the qualified data to determine the appropriate label.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 5
DOI: https://doi.org/10.17762/ijritcc.v10i5.5542
***Article Received**: 10 March 2022 **Revised**: 22 March 2022 **Accepted**: 25 April 2022 **Publication**: 31 May 2022*

$$v \in \begin{cases} C1, & d\,v > 0.5 \\ C2, & \text{otherwise} \end{cases}$$ **Eq→5**

(vii) The performance is then provided as the related class labels as "c1" or "c2" respectively.

The flowchart of the proposed ESVM classification for ECG signals is given in the following Figure 5.

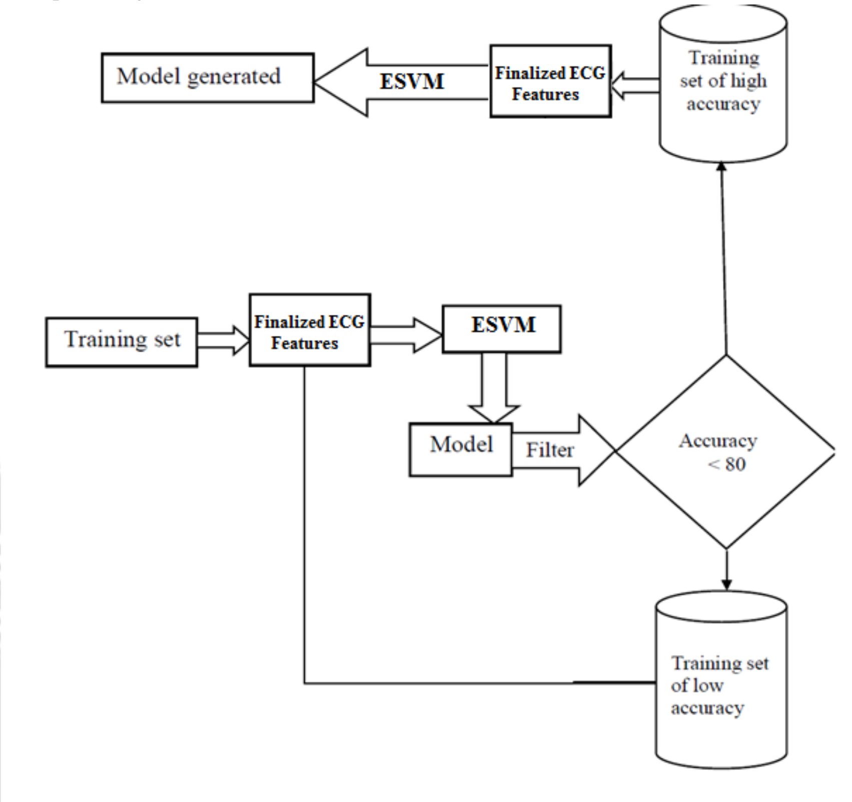

Figure 5: Flow Chart of Proposed ESVM ECG Classification

4. RESULTS AND DISCUSSIONS

Arrhythmia Datasets from MIT-BIH: These resources, which were retrieved from the Physio-net website, comprise 48 1/2-hour samples of 2 channel peripheral ECG readings from 47 participants. The participants comprised 25 males and 22 females, ranging in age from 32 to 89 years of age. The males ranged 32–89 years of age, while the females have been 23–89 years of age. Every recording has a frequency of 360 Hz and an 11-bit resolution across a 10mV range. There are 14 different kinds of rhythms and 17 different kinds of heartbeats in the datasets. It chooses the more common components (4 majority-classes) and the least common items (4 minority-classes) from the datasets to exhibit unbalancing activity.MATLAB 2013b has been used to simulate the system.

False-Positives (FP) and False-Negatives (FN) must always be avoided while diagnosing a patient since an inaccurate diagnostic might result in the patient suffering an impairment if the improper treatment is prescribed. In this research, FP and FN are significant. The ultimate metrics have been calculated by taking the average of all class metric scores.

It uses accuracy as a metric for the classifier model's entire performance

$$Acc = \frac{TP + TN}{(TP + TN + FP + FN)} \cdot 100$$

Eq→6

Table 2: Numerical Comparison of Accuracy

ECG-SIGNALS	EPSO	ESVM
ECG-1	95.31	97.31
ECG-2	96.99	98.99
ECG-3	96.11	98.11
ECG-4	94.21	96.21
ECG-5	93.71	95.71

Table 2 and Figure 6 show the comparison of the accuracy level for both EPSO and the ESVM classification. Hence it proves the accuracy is better for the ESVM as for ECG-2 is 98.99 when compared with EPSO as for ECG-2 is 96.99. This proves the role of EPCA based feature extraction plays a significant effect on classification performance.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 5
DOI: https://doi.org/10.17762/ijritcc.v10i5.5542
Article Received: 10 March 2022 Revised: 22 March 2022 Accepted: 25 April 2022 Publication: 31 May 2022

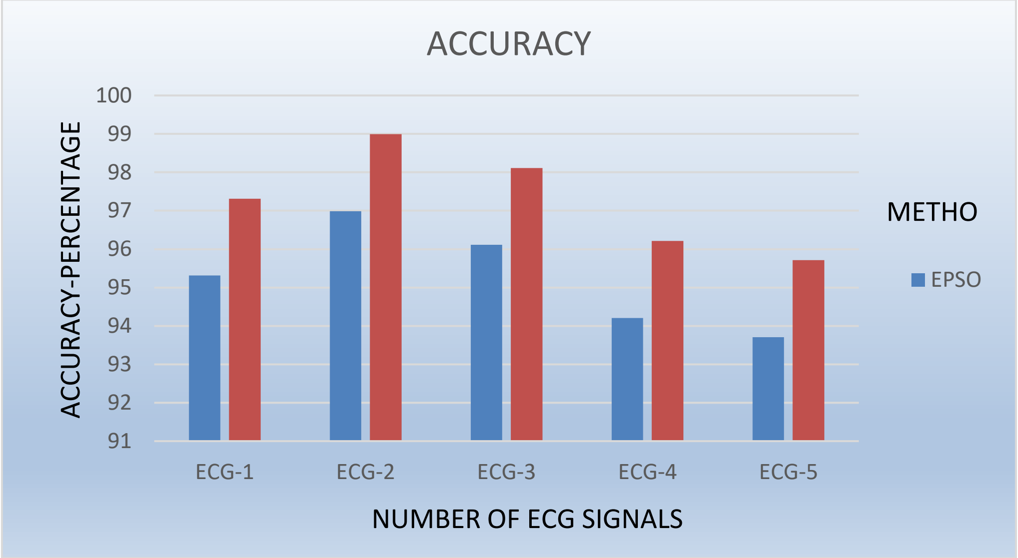

Figure 6: Graphical Comparison of Accuracy

Sensitivity is often a measure that is connected to the FN rate.

$$Sen = \frac{TP}{(TP + FN)} \cdot 100 \qquad \textbf{Eq} \rightarrow \textbf{7}$$

Table 3 and Figure 7 show the comparison of the Sensitivity level for both EPSO and the ESVM classification. Hence it proves the sensitivity is better for the ESVM as for ECG-2 is 98.87 when compared with EPSO as for ECG-2 is 98.87.

Table 3: Numerical Comparison of Sensitivity

ECG-SIGNALS	EPSO	ESVM
ECG-1	95.22	97.22
ECG-2	96.87	98.87
ECG-3	96.02	98.02
ECG-4	94.11	96.11
ECG-5	93.62	95.62

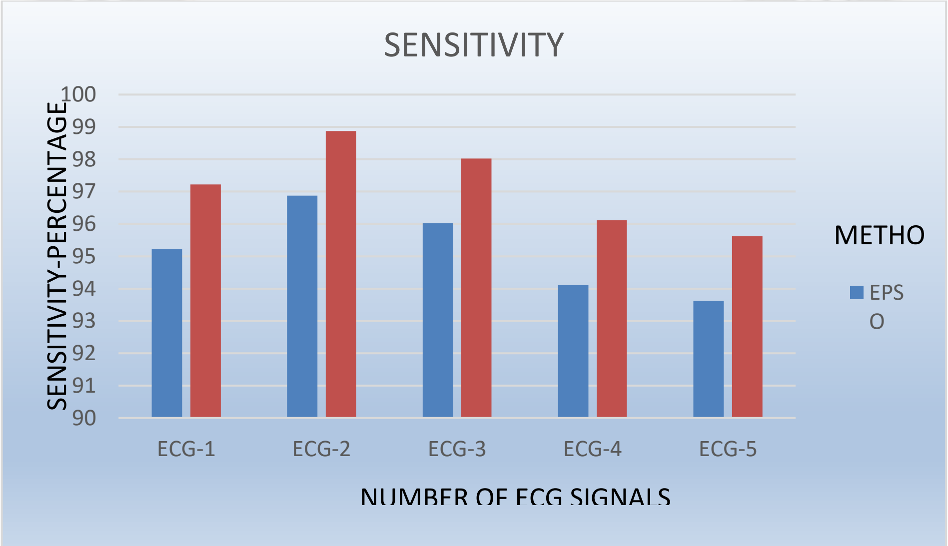

Figure 7: Graphical Comparison of Sensitivity

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 5
DOI: https://doi.org/10.17762/ijritcc.v10i5.5542
***Article Received**: 10 March 2022 **Revised**: 22 March 2022 **Accepted**: 25 April 2022 **Publication**: 31 May 2022*

Precision is often a metric that is connected to the FP rate.

$$Pre = \frac{TP}{(TP + FP)} \cdot 100$$

Eq→8

Table 5: Numerical Comparison of Precision

ECG-SIGNALS	EPSO	ESVM
ECG-1	95.12	97.12
ECG-2	96.77	98.77
ECG-3	95.92	97.92
ECG-4	94.01	96.01
ECG-5	93.52	95.52

Table 5 and Figure 8 show the comparison of the Precision level for both EPSO and the ESVM classification. Hence it proves the precision is better for the ESVM as for ECG-2 is 98.77 when compared with EPSO as for ECG-2 is 96.77.

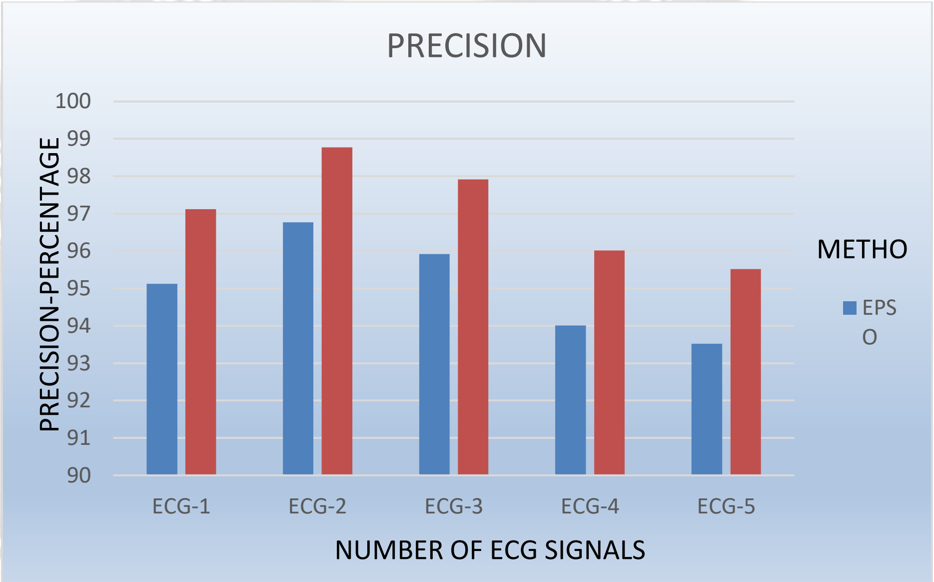

Figure 8: Graphical Comparison of Precision

While there exist unbalancing class issues, the F1-Score is computed as the harmonic mean of sensitivity and precision, considering these measures into the evaluation to measure the efficiency of the classification algorithm.

$$F1Score = 2 \cdot \frac{precision \cdot sensitivity}{precision + sensitivity} \cdot 100$$

Eq→9

Table 5: Numerical Comparison of F1-Score

ECG-SIGNALS	EPSO	ESVM
ECG-1	95.18	97.18
ECG-2	96.82	98.82
ECG-3	95.97	97.97
ECG-4	94.06	96.06
ECG-5	93.57	95.57

Table 5 and Figure 9 show the comparison of the F1-Score level for both EPSO and the ESVM classification. Hence it proves the F1-Score is better for the ESVM as for ECG-2 is 98.82 when compared with EPSO as for ECG-2 is 96.82.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 5
DOI: https://doi.org/10.17762/ijritcc.v10i5.5542
***Article Received**: 10 March 2022 **Revised**: 22 March 2022 **Accepted**: 25 April 2022 **Publication**: 31 May 2022*

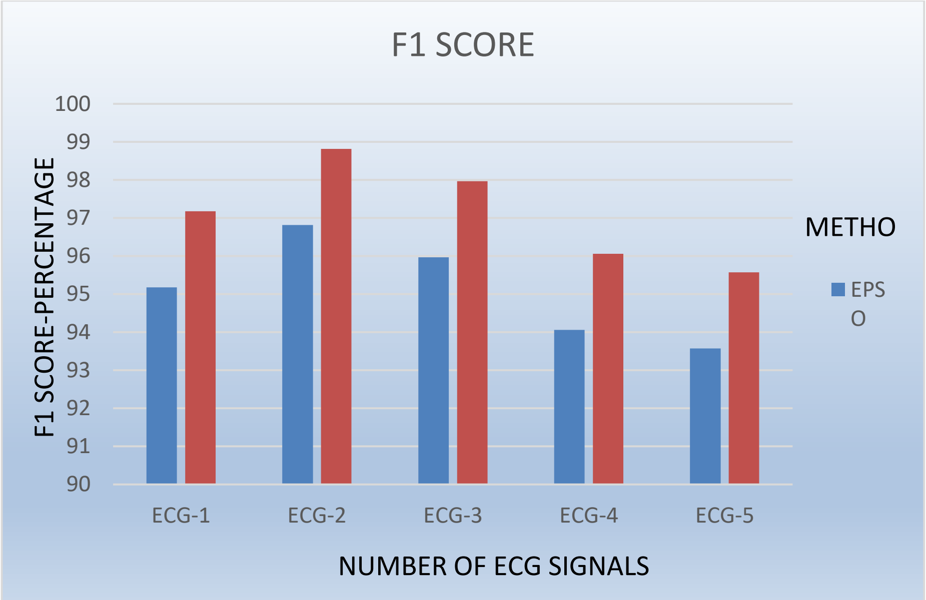

Figure 9: Graphical Comparison of F1-Score

5. CONCLUSION

ECG monitoring has been the simplest significant and effective method of evading cardiac problems. Our research integrates a unique "Enhanced version of SVM (ESVM)" classification approach with a unique "Enhanced version of PCA (EPCA)" feature extraction approach to diagnose various kinds of cardiac irregularities. The MIT-BIH Arrhythmia Datasets were used to undertake classification studies on various cardiac arrhythmias and healthy beats. Following feature extraction, the classifications were trained, cross-validated, and evaluated on the actual heartbeat. We use an ESVM classification to categorize the signals into healthy and cardiac disease categories. The ESVM specifications are chosen using a 5-fold cross-validation technique throughout the training phase. Our studies demonstrate that the presented ESVM, when integrated with EPCA extracting features, considerably increases classification performance over EPSO. Considering Accuracy, Sensitivity, Precision, and F1-Scores of 98.99 percent, 98.87 percent, 98.77 percent, and 98.82 percent, it demonstrated good effectiveness in distinguishing anomalous beats from regular beats. In future work, we plan to classify the ECG signals with advanced bio-inspiring classification models to improve the accuracy rate.

REFERENCES

[1]. World Health Organization. (2018). World Health Statistics 2018: Monitoring health for the SDGs, Sustainable Development Goals. Geneva, Switzerland. Accessed: Jun. 5, 2020. [Online]. Available: https://www.who.int/health-topics/cardiovascular-diseases/

[2]. S. H. Jambukia, V. K. Dabhi, and H. B. Prajapati, ‘‘Classification of ECG signals using machine learning techniques: A survey,’’ in Proc. Int. Conf. Adv. Comput. Eng. Appl., Mar. 2015, pp. 714–721.

[3]. U. R. Acharya, Y. Hagiwara, J. E. W. Koh, S. L. Oh, J. H. Tan, M. Adam, and R. S. Tan, ‘‘Entropies for automated detection of coronary artery disease using ECG signals: A review,’’ Biocybern. Biomed. Eng., vol. 38, no. 2, pp. 373–384, 2018.

[4]. M. Kachuee, S. Fazeli, and M. Sarrafzadeh, ‘‘ECG heartbeat classification: A deep transferable representation,’’ in Proc. IEEE Int. Conf. Healthcare Informat. (ICHI), Jun. 2018, pp. 443–444

[5]. X. Gao, ‘‘Diagnosing abnormal electrocardiogram (ECG) via deep learning,’’ in Practical Applications of Electrocardiogram, U. Lakshmanadoss, Ed. Rijeka, Croatia: InTech, 2019

[6]. P. Plawiak and U. R. Acharya, ‘‘Novel deep genetic ensemble of classifiers for arrhythmia detection using ECG signals,’’ Neural Comput. Appl., pp. 1–25, Jan. 2019.

[7]. A. M. Delaney, E. Brophy, and T. E. Ward, ‘‘Synthesis of realistic ECG using generative adversarial networks,’’ 2019, arXiv:1909.09150. [Online]. Available: https://arxiv.org/abs/1909.09150

[8]. V. V. Kuznetsov, V. A. Moskalenko, and N. Y. Zolotykh, ‘‘Electrocardiogram generation and feature extraction using a variational autoencoder,’’ 2020, arXiv:2002.00254. [Online]. Available: http://arxiv.org/abs/2002.00254

[9]. W.-H. Jung and S.-G. Lee, ‘‘ECG identification based on non-fiducial feature extraction using window removal method,’’ Appl. Sci., vol. 7, no. 11, p. 1205, Nov. 2017, DOI: 10.3390/app7111205.

[10]. L. B. Marinho, N. D. M. M. Nascimento, J. W. M. Souza, M. V. Gurgel, P. P. R. Filho, and V. H. C. de Albuquerque, ‘‘A novel electrocardiogram feature extraction approach for

cardiac arrhythmia classification,'' Future Gener. Comput. Syst., vol. 97, pp. 564–577, Aug. 2019.

[11] R. Xiao, Y. Xu, M. M. Pelter, D. W. Mortara, and X. Hu, ''A deep learning approach to examine ischemic ST changes in ambulatory ECG recordings,'' AMIA Summits Transl. Sci. Proc., vol. 2017, pp. 256–262, May 2018.

[12] Y. Wu, F. Yang, Y. Liu, X. Zha, and S. Yuan, ''A comparison of 1-D and 2-D deep convolutional neural networks in ECG classification,'' in Proc. Comput. Vis. Pattern Recognit., 2018.

[13] H. Wang, H. Shi, K. Lin, C. Qin, L. Zhao, Y. Huang, and C. Liu, ''A high-precision arrhythmia classification method based on dual fully connected neural network,'' Biomed. Signal Process. Control, vol. 58, Apr. 2020, Art. no. 101874.

[14] C. Chen, Z. Hua, R. Zhang, G. Liu, and W. Wen, ''Automated arrhythmia classification based on a combination network of CNN and LSTM,'' Biomed. Signal Process. Control, vol. 57, Mar. 2020, Art. no. 101819.

[15] M. Porumb, E. Iadanza, S. Massaro, and L. Pecchia, ''A convolutional neural network approach to detect congestive heart failure,'' Biomed. Signal Process. Control, vol. 55, Jan. 2020, Art. no. 101597.