Journal of Engineering Science and Technology
Vol. 15, No. 2 (2020) 1277 - 1295


# MULTISTAGE CLASSIFICATION OF ARRHYTHMIA AND ATRIAL FIBRILLATION ON LONG-TERM HEART RATE VARIABILITY

KIRTI*, HARSH SOHAL, SHRUTI JAIN

Department of Electronics and Communication,
Jaypee University of Information Technology,
Waknaghat, Solan, 173234, Himachal Pradesh, India
*Corresponding Author: kirtitripathifzk@gmail.com

**Abstract**

This article proposes a Multi-Stage Heart Rate Variability Classification (MSHVC) system to diagnose Normal, Arrhythmia (AR) and Atrial Fibrillation (AF) for Long-Term ECG analysis. The MSHVC methodology comprises of ECG pre-processing, QRS detection, HRV feature extraction, statistical analysis and classification. The frequency-domain, time-domain, and geometrical-domain HRV features were extracted and accuracy was improved using Analysis of Variance (ANOVA) test. Artificial Neural Network (ANN), Support Vector Machine (SVM), and k-Nearest Neighbor (kNN) classifiers are utilized at various levels to classify in two-stage and three-stage classification. The MSHVC classification system demonstrates a higher accuracy compared to that of other state of the art methods when applied to MIT/ BIH Normal Sinus Rhythm (NSR), MIT/ BIH Arrhythmia (AR) and MIT/ BIH Atrial Fibrillation (AF) databases. To classify normal ECG from abnormal, proposed system attained maximum overall accuracy of 98.36% by ANN at 2-stage classification. Multi-stage classification of abnormal ECG further divided into AR and AF attains 99% of overall accuracy by ANN after statistical analysis.

Keywords: Analysis of variance, Artificial neural network, Heart rate variability, K-nearest neighbor, Support vector machine.

## 1. Introduction

Irregular heartbeats are the early warning sign to diagnose various types of heart diseases including Congestive Heart Failure, Arrhythmia (AR), Atrial Fibrillation (AF), etc. [1]. Early detection of irregular heartbeats leads to towards significant reduction in the expensive cardiac operation [2]. The indications of Arrhythmia and Atrial Fibrillation can be misled with normal sinus rhythm, which can lead to false diagnosis by clinicians [3, 4]. A patient may not be alert about its symptoms of palpitations, dizziness, or chest ache; these signs can also be indicated by a healthy heart. Hence, only clinicians are not enough to diagnose cardiac diseases by its symptoms itself. Various techniques are available to detect these irregular heartbeats including Electrocardiogram (ECG), Blood test, Holter Monitoring, Electro-Physiology Studies (EPS) and Echocardiogram [5]. Among them, ECG is the best technique to detect irregular heartbeats as it detects the electrical activity of cardiac muscles by placing the electrodes nearby the heart.

To analyse the difference between healthy and unhealthy patient Heart Rate Variability (HRV) is considered as the practical approach. The HRV analysis standards are published by the European Society of Cardiology and The North American Society of Pacing Electrophysiology [6]. Its monitoring also provides a measure of the sympathetic and parasympathetic nervous system, which is controlled by the primary part of our brain, i.e., Autonomic Nervous System (ANS). HRV analyses notify the working of Heart Rate (HR), blood pressure, breathing and digestion. Nowadays, HRV is used to provide information about our mental and physical health by observing one’s lifestyle. HRV parameter analysis assists clinicians to monitor their patients for various disorders such as sleep, stress, physical activity, emotional health, etc. HRV features can be computed in the time domain, geometrical domain and frequency domain. These domains are also analysed by applying a different length of the ECG data termed as an ultra-short term (less than 5 minutes), short term (5 minutes) and long term (24 hours) [7].

## 2. Related Work

To classify heart diseases on the basis of HRV analysis numerous studies have been already completed [8-10]. Classification based on HRV analysis provides optimistic results at various levels of classification. Automatic computer-based machine learning algorithms reduce the manual error arising during medical analysis and diagnosis [11, 12]. Mental health, schizophrenia, anxiety, arrhythmia, congestive heart failure and atrial fibrillation are the main categories of diseases, which can be accurately analysed by HRV [13-16]. Nine features from various domains are extracted on short-term HRV using SVM [17]. The extraction of non-linear parameters of HRV is done using SVM to categorize normal patient and congestive heart failure patient [7]. He achieved 94.4% accuracy on 2-hour time-length and 96.7% accuracy by applying non-standard features of HRV. The proposal of using ensemble classifier is less reliable as compared to other classifiers due to the varying nature of training dataset [18].

Incremental ANOVA and Functional Networks-Feature Selection (IAFN-FS) technique to deal with the complex nature of already existing methods such as decision tree and naive bayes with better accuracy is introduced [19]. Classification using HRV analysis in the long and short term has been done to detect the mental stress on students during examination [14, 15]. The automatic predictive model for short term HRV is proposed to classify the mental pressure present during viva-voce [15]. Study to evaluate the risk of falling in hypertensive patients as per short-term HRV is carried out by Castaldo et al. [13]. A prediction based method using HRV analysis is presented and obtained 68.9% and 79.3% accuracy for 10 min and 15-minute respectively [10]. Yang and Yin [20] obtained a low accuracy of 57% during short-term HRV signal established on footprint analysis. To classify Paroxysmal Atrial Fibrillation (PAF) genetic algorithm is utilized to attain better accuracy [3]. The kNN classifier using short term HRV is employed by Isler and Kuntalp [21] to classify the HRV feature of 54 normal and 29 abnormal subjects of congestive heart failure. The same classifier is also presented using non-standard HRV parameters on the MIT/BIH database [22]. To overcome the problem of selection of centroids in kNN classifier researcher associates the weights along with them [23]. Also, by using four HRV features for 1000 RR-intervals, a significant accuracy has been obtained [24].

Classification employing Artificial Neural Network (ANN) is a broad topic in the literature. Previously, enormous work has already been done on ANN to classify various types of heart diseases using HRV analysis. Several authors have evaluated the results from ANNs with existing classifiers and concluded that ANNs gives better accuracy than the others. Isler et al. [5] mentioned that a multistage classification for short term HRV analysis using 3-stage classifier is proposed and achieved 98.8% accuracy for congestive heart failure database.

In this paper, designing of a Multistage Heart rate Variability Classification (MSHVC) technique is done that consists of two-stage and three-stage classifications. For two-stage classification normal (NSR) and abnormal ECG (AR and AF) are separated, which is followed by the distribution of the abnormal ECG into AR and AF for three-stage classification. ECG pre-processing techniques are applied on the ECG databases and R-peak is detected using the Pan-Tompkins algorithm [25]. Various HRV features are extracted in time-domain, frequency-domain and geometrical domain [4]. Later, classification is performed using SVM, kNN and ANN classifiers.

The novelty of this article lies in extensive feature selection applied on the extracted feature set to select the most significant features that give additional useful information than others. The contribution of this work lies in selecting the most accurate classifier by comparing various classifiers at various level of abstraction of classification. This work is organized as: Section 3 explains the experimental dataset description and different approaches used for the proposed system. In section 4 the results and discussions of the presented system are explained with percentage improvement at various level of experiments and comparison of the proposed approach with recent approaches. Lastly, conclusions are drawn, and discussion is carried out for a further extension of the current work in future.

## 3. Materials and Methods

Generally, the flow of ECG signal processing comprises of dataset collection, pre-processing, QRS detection, feature extraction, statistical analysis and classification. The proposed methodology for MSHVC is broadly divided in six phases namely: ECG Database Description, ECG pre-processing, QRS detection, HRV Feature Extraction, HRV Statistical analysis and lastly the proposed classification approach is demonstrated in Fig. 1.

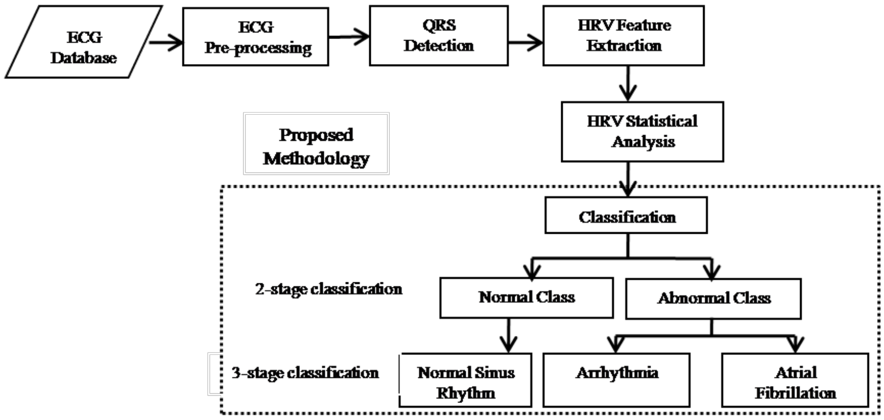


**Fig. 1. Proposed methodology for classification of heart diseases.**

### 3.1. ECG dataset description

The online platform is used to obtain the standard MIT/ BIH database for the proposed methodology. In this work, three types of databases namely NSR, AR and AF are considered. The database comprises of the 24-hour length of ECG of patients having different age and gender. The experimental database contains a total number of 64 ECG waveforms having different values of Heart Rate (HR) [26]. To improve the accuracy of classification technique outliers have been removed using the statistical method. The proposed classification technique considers 61 ECG waveforms from three datasets in the experiment. Table 1 illustrates the description of ECG database used in the experiment.

**Table 1. ECG database description [26].**

| Database description | No. of ECG database with outliers | No. of ECG database without outliers | Class description |
|---|---|---|---|
| **MIT/BIH normal sinus rhythm** | 18 | 18 | 60 bpm < HR < 00 bpm |
| **MIT/BIH Arrhythmia** | 21 | 20 | 60 bpm > HR > 100bpm |
| **MIT/BIH atrial fibrillation** | 25 | 23 | HR > 100 bpm |

Heart rate (HR), beats per minute (bpm)

### 3.2. ECG pre-processing

ECG is a biomedical signal corrupted with various high and low-frequency noises. Signal pre-processing is an important step to eliminate unwanted signal from the useful ECG signal. The valuable information of ECG lies in between 0.5 Hz to 100 Hz. This useful range of frequency corrupted from high-frequency noises and low-frequency noises labelled as Electromyography (EMG), Base Line Wander (BLW) and Power Line Interference (PLI). Pre-processing of the ECG signal is a crucial phase before further signal processing. Because these unwanted frequencies produce incorrect information and lead to false diagnosis by clinicians. Digital filters are generally used for pre-processing of the ECG signal [4].

In this work, we are using FIR Kaiser Window filter due to its linear phase and ease of use. Kaiser window technique for pre-processing filters is applied during ECG signal pre-processing, due to its better frequency response as compared to other classical windowing techniques. EMG is a high-frequency noise produced due to the interference of the electrical activity of muscles, existed above 100 Hz of frequency. A low pass filter (LPF) having cut-off-frequency (‘$f_c$’) of 100 Hz is used to eliminate EMG noise. BLW is a low-frequency noise generated due to respiration, loose electrodes and movement during acquisition, present below 0.5 Hz of frequency. A high pass filter (HPF) having $f_c$ of 0.5 Hz is used to remove BLW noise [25]. PLI is also a high-frequency noise induced by cause of the electromagnetic interference produced by other biomedical instruments present nearby ECG equipment, occurs at 50/60 Hz of frequency. A band-stop filter (BSF) having $f_c$ of 49.5 Hz to 50.5 Hz is used to get rid of PLI noise.

### 3.3. QRS detection

QRS wave detection is an important step before HRV feature extraction. It is a prominent positive wave distinguished easily in ECG waveform. Detection of QRS wave includes squaring and moving window integration. Squaring helps to eradicate all the negative peaks and enhance the detection accuracy of the QRS wave. Moving window integration helps to detect the waveform nature of QRS wave by sliding a window around each QRS wave present in the ECG [25]. Further, adaptive thresholding is used to eliminate abnormal peaks and extracting the R-peak from the QRS complex. Several techniques have been developed for beat detection such as fixed thresholding, adaptive thresholding, k-mean clustering, segment specific thresholding and classification. Among all these methods adaptive thresholding technique is commonly used for decision making. An adaptive thresholding technique modification of threshold is done according to the morphology of the QRS wave. In any case, if there is missing QRS wave then search back technique is used to detect that missed wave. Nowadays, some soft computing methods such as neural network-based techniques are used for classification.

### 3.4. HRV feature extraction

Feature extraction is a decisive part of the classification process. Better features will lead to enhanced classification performance, however, if the feature is not chosen properly it will yield to incorrect result for the classification system. In the HRV feature extraction, the R-peak detection assists to calculate the Heart Rate (HR) of the patient. HR helps to calculate the various HRV features in three domains: time domain, frequency domain and geometric domain. In this work, consideration of long-term (24 hour ECG recordings) HRV feature extraction in all the domains is taken [4].

Time-domain HRV features The fluctuations in the time intervals between adjacent heartbeat termed as HRV. In this work, eight HRV time-domain features are calculated. All these time-domain features are calculated by using the interval between two R-R peaks or $RR_{interval}$. These eight HRV features are labelled as mean of RR interval (mean RR), Standard deviation of RR interval (SDNN), Standard deviation of Successive Difference (SD) of RR interval (SDSD), Root mean square of SD (RMSSD), a number of consecutive RR intervals differ by more than 50 minutes (NN50), a sum of NN50 divided by the total number of the RR interval (pNN50), the difference between maximum to minimum HR (Max-Min HR) and mean of HR (mean HR). The interval between two specific *R* waves termed as RR Interval. Normal RR interval lies within the range of 0.6-1.2 seconds [4, 27].

Frequency-domain HRV features: These features are measured through the estimated Power Spectral Density (PSD) of RR intervals termed a periodogram. It estimates the frequency of occurrence of RR intervals in time-series data, similar to FFT but optimized for unevenly time-sampled data. The Welch PSD estimation '$P_{WE}$' is used to compute frequency-domain HRV where $P_j$ ($f$) is the PSD of $RR_{interval}$ samples having '$L$' segments is illustrated by Eq. (1) [28].

$$P_{WE} = {}^{1}/_{L} \sum_{j=0}^{L-1} P_j \ (f) \tag{1}$$

There are mainly eleven HRV frequency domain features calculated from the periodogram [28]. Long term frequency domain spectrum consists of four frequency bands: ultra-low-frequency (ULF), very-low-frequency (VLF), low-frequency (LF) and high-frequency (HF) bands. HRV frequency components encompass of absolute values of power computed from the power spectrum of these bands specified as the absolute power of ULF (aULF), the absolute power of VLF (aVLF), the absolute power of LF (aLF), the absolute power of HF (aHLF), the absolute power of total power (aTotal), the relative power of VLF (pVLF), the relative power of LF (pLF), the relative power of HF (pHF), Normalized change of Total power-on LF (Lfnorm), Normalized change of Total power on HF (Hfnorm), Ratio of LF power to HF power (LF/HF).

Geometrical-domain HRV features: These features are computed by generating the histogram of the successive difference of RR interval. HRV Triangular Index (HRVTI) is produced in this domain using the ratio of the maximum value in the histogram to the length of the histogram.

### 3.5. Statistical analysis

Statistical analysis helps to select the most significant features among the different calculated time-domain, frequency-domain and geometrical domain features. In ECG signal processing we extract a large number of features in various domains but only some of the features show a notable effect in the final result. In this work, a total number of 20 HRV features are obtained from frequency, geometrical and time domain. Statistical analysis is done to extract the significant features using Analysis of Variance (ANOVA) test. It illustrates the variation among the mean of NSR, AR and AF diseases.

In the ANOVA test, hypothesis testing is done on HRV features for different databases as mentioned earlier. In this paper, we have considered two hypotheses, null hypothesis ($H_o$) and the alternative hypothesis ($H_1$). $H_o$ means no variation and $H_1$ signifies at least one of the means is unequal.

$H_o$: $\mu_1 = \mu_2 = \mu_3$
$H_1$= at least one of the means is unequal

where, $\mu_1$ = Mean of NSR HRV feature

$\mu_2$ = Mean of AR HRV feature
$\mu_3$ = Mean of AF HRV feature

Comparative results of the ANOVA test, which represents the significant variation among parameters of different databases represented in Mean ± Standard Deviation (SD) form. The p-value illustrates a significance level that can be demonstrated by Eqs. (2) and (3):

$p > 0.05$ = Weak Significance (WS) (2)

$p \leq 0.05$ = Strongly Significance (SS) (3)

Feature selection process using ANOVA test that provides the significance values between two groups. By employing this technique nine significance features have been selected, which is illustrated in Table 2.

The HRV feature values depend on the sampling frequency utilized while processing. In Table 2, authors have generated the results at 1000 Hz sampling frequency for a long-term database.

**Table 2. Comparison of HRV features using ANOVA.**

| Domain | HRV features | MIT/BIH AR database | MIT/BIH NSR database | MIT/BIH AF database | p-value |
|---|---|---|---|---|---|
| **Time domain** | **SDSD** | 1.84e2 ± 20.76 | 1.98e2 ± 5.53 | 1.97e2 ± 13.47 | p < 0.05 |
| | **RMSSD** | 1.84e2 ± 20.75 | 1.98e2 ± 5.64 | 1.97e2 ± 13.47 | p < 0.05 |
| | **NN50** | 5.37e2 ± 28.54 | 7.88e2 ± 19.23 | 8.09e2 ± 50.45 | p < 0.05 |
| | **pNN50** | 55.51 ± 2.65 | 52.28 ± 0.89 | 54.80 ± 3.12 | p < 0.05 |
| | **mean HR** | 93.78 ± 4.05 | 97.05 ± 1.83 | 94.89 ± 5.32 | p < 0.05 |
| **Frequency domain** | **aULF** | 0.073 ± 0.007 | 0.07 ± 0.008 | 0.76 ± 0.008 | p < 0.05 |
| | **aVLF** | 1.33 ± 0.058 | 1.34 ± 0.03 | 1.37 ± 0.041 | p < 0.05 |
| | **aHF** | 9.13 ± 0.25 | 9.34 ± 0.082 | 8.89 ± 1.90 | p < 0.05 |
| | **aTotal** | 14.47 ± 0.37 | 14.76 ±0.14 | 14.81 ± 0.29 | p < 0.05 |

## 3.6. Classification system

Several classification methods are used to investigate the performance of HRV features to distinguish between normal and abnormal ECG signal [29]. They are trained using five-fold cross-validation technique. In this proposed methodology, classification phase has been divided into two levels, the first level consists of 2-stage classification that separates the normal ECG signal from an abnormal ECG signal. Whereas, the second level comprises of 3- stage classification, which further divides the abnormal ECG signal into two types of heart diseases that depend on HR namely Arrhythmia and Atrial Fibrillation. In order to classify the ECG signals at both levels, k-Nearest Neighbour (kNN), Artificial Neural Network (ANN) and Support Vector Machine (SVM) techniques are employed [13-16, 26].

### 3.6.1. Support vector machine

Hsu et al. [30] introduced a supervised learning machine algorithm termed as SVM, which is based on statistical learning theory. SVM classification is a two-level concept: training and testing. The best feature among trained data is chosen, then it applies to unseen data during the testing period. A hyperplane is a three-dimensional plane that helps to separate two classes in SVM. The training data, which is nearest to the hyperplane termed as Support Vectors (SV). The largest distance to SV intends to maximize the margin between the separating hyperplane and reduction of generalized error also called as functional margin [31]. The given set of training data ($X_i$) lies in between ($X_1$, $X_2$, $X_3$...... $X_n$) of n point is represented by Eq. (4):

$$(\vec{X}_1, O_1), \dots \dots \dots, (\vec{X}_n, O_n) \tag{4}$$

The output ($O_1$, $O_2$, $O_3$...... $O_n$) can be -1 or 1 represent different classes for input data ($X_1$, $X_2$, $X_3$...... $X_n$) during binary classification. The generalized equation for hyperplane, which separates the classes is given by Eq. (5).

$$\vec{W} \cdot \vec{X} - B = \tag{5}$$

Annotations: $\vec{W}$ is weight vector, $\vec{X}$ is the input vector and $B$ is the bias

SVM classification depends on the kernel selection as there are various SVM kernels like linear, polynomial, Gaussian radial basis function (RBF) kernel, etc. All these kernels are tested on HRV features but Gaussian Radial Basis Function (GRBF) gives better accuracy due to higher dimensional space to separate the training data in their respective classes. The non-linear equation of GRBF is represented by Eqs. (6) and (7) as:

$$f(X) = \sum_{i=1}^{n} W_i K(X - X_i) + B \tag{6}$$

$$K(X, X_i) = \exp\left(-\frac{\sum\|\overline{X - X_i}\|^2}{2\sigma^2}\right) \tag{7}$$

where $X$ is the input vector, $X_i$ is the support vector, $B$ is bias and weight vector, $W$ is weight, $\sigma$ is the Gaussian distribution.

### 3.6.2. k-nearest neighbor

The k-Nearest Neighbor (kNN) is the simplest and widely used supervised classification algorithm. It works on the principle of the Nearest Neighbor rule because it allocates a point '$x$' to the class maximum present amongst the '$k$' points in the training set nearest to '$x$'. The decision is made based on which, points are nearest as per Euclidean distance. The entire points within the neighbourhood add equally to the final decision for '$x$'. Although the kNN has various advantages such as ease of implementation and effectiveness it also has less accuracy in the classification process. To improve the accuracy of classification, we have used weighted kNN. According to Bicego and Loog [32], the general equation for weighted kNN is illustrated in Eq. (8) as:

$$arg\max_{c} \sum_{i=1}^{K} I_c(n_i)\, w_{n_i} \rightarrow x \tag{8}$$

In the weighted kNN, every neighbor $n_i \in n_e K(x)$ is equipped with a weight $w_{n_i}$.Where $I_c(n_i)$ is the function to determine the class '$c$'. Conceptualization of kNN was firstly considered by Royall [33] but in the classification method, Dudani [34] commenced a distance weighted kNN.

### 3.6.3. Artificial neural network

To develop an intelligent system for the classification numerous research has been done in the literature [35, 36]. Most of these systems are inspired by the human brain including how they process the information. High performance of the ANN classification process required a high number of interconnected neurons. A node is the basic computational unit in artificial neurons; it receives the input from the environment or the external source. Each node is associated with some weight that calculates the function of the weighted sum of its input. The essential structure of the neural network is presented by selecting the learning rule, activation function architecture.

The feed-forward networks and feedback networks are two types of NN architectures. MSHVC classification system used feedback network termed as backpropagation method because it takes the system output into consideration, which helps the system to adjust the performance as per the desired response. It is a supervised learning neural network algorithm. The neurons in the input layer assign the input data '$x_{ij}$' (input vector) to neurons in the hidden layers. Each neuron '$j$' in the hidden layer adds its input signals '$x_i$' providing them with respective weight vector '$w_{ij}$'. The net input of the activation block, *netj*, is computed by the summation of the inner product of the input, weight vector, and bias having '$n$' number of neurons in Eq. (9) as:

$$net_j = bias_j + \sum_{i=1}^{n} w_{ij} x_i \tag{9}$$

$$O_j = \varphi(net_j) \tag{10}$$

The output of the activation function $O_j$ is given by Eq. (10) [36], where $\varphi$ represents is the activation function of the neuron.

### 3.7. Parameter evaluation

The final classification stage is evaluated by employing evaluation parameters described in terms of true positives (TP), true negatives (TN), false positives (FP) and false negatives (FN). The Positive Predictive Value (PPV) signifies the proportion of positive results that are TP. The Area Under Curve (AUC) or receiver operating characteristics compares sensitivity versus specificity to predict the outcome of the classification [2, 10].

## 4. Results and Discussion

The results of the proposed system are divided into six phases: ECG dataset description, ECG pre-processing, QRS detection, HRV feature extraction, HRV statistical analysis and classification. Figure 2 depicts the qualitative results obtained from the four phases of HRV signal processing.

In this paper, authors have worked on sixty-one (61) ECG database of NSR, AR and AF. For pre-processing, the Kaiser Window FIR filter is applied to eliminate EMG, BLW and PLI. The possible limitation in the pre-processing phase is the consideration of three types of the digital filter; it can also be done by using one bandpass filter having '$f_c$' 0.5 to 100 Hz and BSF having '$f_c$' 50/60 Hz. The QRS complex detection leads to computing the HR of the patient. Figure 2(a) represents the ECG waveforms of normal sinus rhythm, Figs. 2(b) to (d) illustrates the waveform after removing EMG, BLW and PLI respectively. Figure 2(e) demonstrates the QRS detection obtained from the Pan-Tompkins algorithm. During HRV feature extraction twenty features of time-domain, frequency-domain and geometrical domain are extracted. To select the most significant features among extracted feature ANOVA is applied. After monitoring the results of ANOVA, the inference is drawn that some HRV features like SDSD, RMSSD, NN50, pNN50, HRVTI, aULF, aVLF, aHF, and aTotal gives a strong indication to reject $H_o$. It signifies that these features show a significant difference in their mean value. Whereas, all other parameters show weak significant against to decline $H_o$, except mean HR, which indicates a moderately significant change. After analysing the p-value of ANOVA that gives the prominent difference among the databases got selected. Based on studies by Kirti et al. [4], only 9 out of 20 features are selected using statistical analysis resulted for MSHVC system.

Other than statistical analysis optimization techniques are also popular nowadays to improve classification accuracy. The proposed system is evaluated based on evaluation parameters. The authors have performed a list of experiments tabulated in Table 3. In these experiments, we have compared the evaluation parameters at both levels; 2-stage classification and at 3-stage classification. After HRV feature extraction, statistical analysis is applied to select the prominent features that give a noteworthy effect on the output of the system. Nowadays, HRV features are popularly used to classify patients with cardiac diseases using numerous classification techniques. The comparison between various classification techniques like SVM, kNN and ANN is illustrated. The comparison of the evaluation parameters before HRV statistical analysis and after HRV statistical analysis is done.

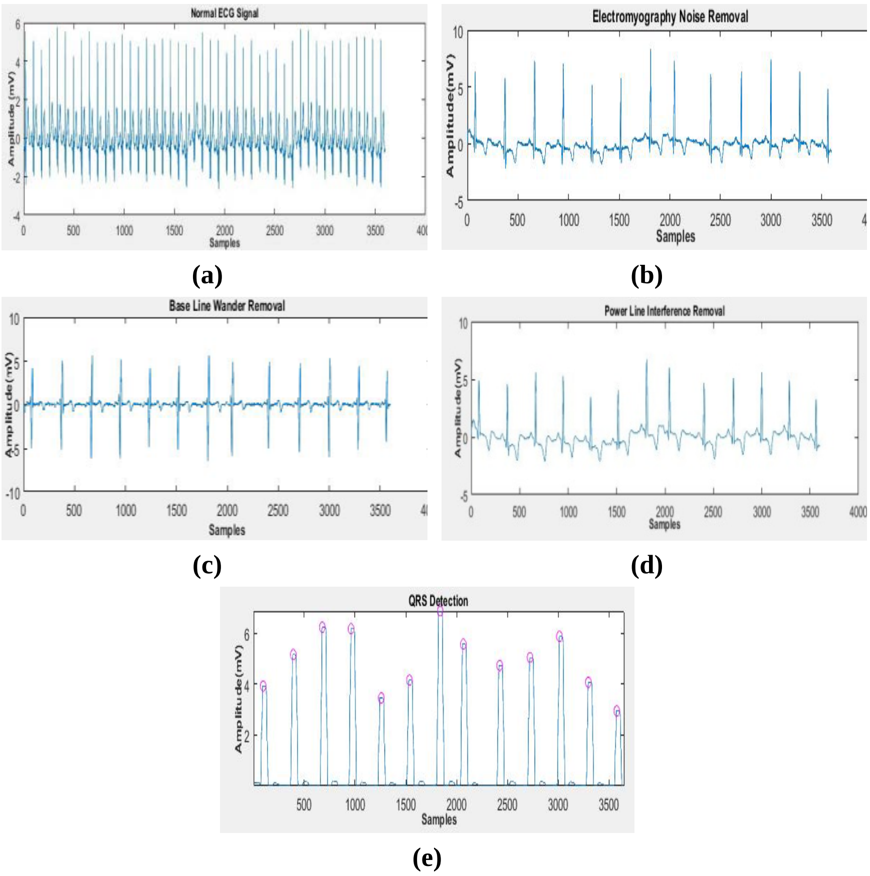


**Fig. 2. ECG waveforms: (a) Normal ECG signal, (b) EMG removal, (c) BLW removal, (d) PLI removal, (e) QRS detection.**

## 4.1. HRV classification levels

The classification performance of the HRV classification illustrates the performance of the proposed MSHVC system. The 2-stage classification and 3-stage classification is applied in this article. Table 3 signifies the list of experiments performed during methodology.

**Table 3. List of experiments performed for MSHRV system.**

| Case | Level | Experiments | Description |
|---|---|---|---|
| **1** | **2-stage classification** | Experiment 1 | Performance measure of MSHVC before statistical analysis |
| | | Experiment 2 | Performance measure of MSHVC after statistical analysis |
| **2** | **3-stage classification** | Experiment 3 | Performance measure of MSHVC before statistical analysis |
| | | Experiment 4 | Performance measure of MSHVC after statistical analysis |

### 4.1.1. Case 1. 2-stage classification

Experiment 1: Performance measure of MSHVC before statistical analysis: Two-stage classification classifies the dataset into normal and abnormal ECG signal. The first level ECG waveforms are signified as Normal ECG and Abnormal ECG. The outcomes of 2-stage classification before statistical analysis using SVM, kNN and ANN is tabulated in Table 4.

**Table 4. Performance measure of classification system for two stage classification before HRV statistical analysis.**

| Features | Classification technique for 2-stage without feature selection | Confusion matrix | | | Sensitivity | Specificity | Accuracy | PPV | AUC |
|---|---|---|---|---|---|---|---|---|---|
| | | | N | AN | | | | | |
| **HRV features (TD, FD, GD)** | **SVM** | N | 10 | 8 | 100% | 55.5% | 86.9% | 84.3% | 0.94 |
| | | AN | 0 | 43 | | | | | |
| | **kNN** | N | 18 | 1 | 95.23% | 94.73% | 95% | 94% | 0.95 |
| | | AN | 2 | 40 | | | | | |
| | **ANN** | N | 18 | 0 | 95.3% | 100% | 96.7% | 100% | 0.96 |
| | | AN | 2 | 41 | | | | | |

Normal (N), Abnormal (AN), Positive Prediction Value (PPV), Area under the curve (AUC)

For two-class classification before statistical analysis, the features determining the normal and abnormal ECG gives 86.9% using SVM classifier, 95 % using kNN and 96.7 % accuracy using the ANN technique. The best results were obtained using ANN having 100% of specificity and 95.3% of sensitivity.

Experiment 2: Performance measure of MSHVC after statistical analysis: Two-stage classification classifies the dataset into normal and abnormal ECG signal. The first level ECG waveforms are signified as Normal ECG and Abnormal ECG. The outcomes of 2- stage classification after statistical analysis using SVM, kNN and ANN is tabulated in Table 5.

For two-class classification after statistical analysis, the features determining the normal and abnormal ECG gives 95% employing SVM classifier, 96.7 % using kNN and 98.36% accuracy by ANN technique. The best results were obtained using ANN having 100% of specificity and PPV along with 97.67% of sensitivity. It can be seen that classification accuracy has been improved after feature selection.

**Table 5. Performance measure of classification system for two stage classification after HRV statistical analysis.**

| Features | Classification technique for 2-stage without feature selection | Confusion matrix | | | Sensitivity | Specificity | Accuracy | PPV | AUC |
|---|---|---|---|---|---|---|---|---|---|
| | | | ***N*** | ***AN*** | | | | | |
| **HRV features (TD, FD, GD)** | **SVM** | N | 18 | 1 | 95.23% | 94.73% | 95% | 94% | 0.95 |
| | | AN | 2 | 40 | | | | | |
| | **kNN** | N | 18 | 1 | 97.6% | 94.7% | 96.7% | 97.6% | 0.96 |
| | | AN | 1 | 41 | | | | | |
| | **ANN** | N | 18 | 0 | 97.67% | 100% | 98.36% | 100% | 0.98 |
| | | AN | 1 | 42 | | | | | |

Normal (N), Abnormal (AN), Positive Prediction Value (PPV), Area under the curve (AUC)

### 4.1.2. Case 2: 3-stage classification

Experiment 3: Performance measure of MSHVC before statistical analysis: The result of 3-stage classification before statistical analysis using SVM, kNN and ANN are tabulated in Table 6.

**Table 6. Performance measure of classification system for 3-stage classification before HRV statistical analysis.**

| Classification technique without feature selection / Performance metrices | | SVM | | | kNN | | | ANN | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | ***N*** | ***AR*** | ***AF*** | ***N*** | ***AR*** | ***AF*** | ***N*** | ***AR*** | ***AF*** |
| **Confusion matrix** | ***N*** | 9 | 2 | 7 | 15 | 1 | 2 | **18** | **0** | **0** |
| | ***AR*** | 0 | 3 | 17 | 0 | 6 | 14 | **1** | **17** | **0** |
| | ***AF*** | 0 | 0 | 23 | 0 | 0 | 23 | **2** | **1** | **22** |
| **PPV** | **$PPV_0$** | 100% | | | 100% | | | 85.71% | | |
| | **$PPV_1$** | 60% | | | 35.29% | | | 94.44% | | |
| | **$PPV_2$** | 48.93% | | | 58.97% | | | 100% | | |
| **Sensitivity** | **$Sen_0$** | 50% | | | 83.33% | | | 100% | | |
| | **$Sen_1$** | 15% | | | 30% | | | 94.44% | | |
| | **$Sen_2$** | 100% | | | 100% | | | 88% | | |
| **Specificity** | **$Sp_0$** | 100% | | | 100% | | | 93.02% | | |
| | **$Sp_1$** | 92.85% | | | 97.56% | | | 97.60% | | |
| | **$Sp_2$** | 36.84% | | | 57.89% | | | 100% | | |
| **Accuracy** | **$Acc_0$** | 85.24% | | | 95.08% | | | 95.08% | | |
| | **$Acc_1$** | 68.85% | | | 75.40% | | | 96.72% | | |
| | **$Acc_2$** | 60.65% | | | 35.29% | | | 95.08% | | |
| **Overall accuracy** | | **59%** | | | **72.13%** | | | **93.4%** | | |
| **AUC** | | **0.91** | | | **0.93** | | | **0.96** | | |

The 3-stage classification divides the abnormal ECG into two types of heart disease: Arrhythmia and Atrial Fibrillation. The evaluation parameters of classification are tabulated in Tables 7 and 8.

For 3-stage HRV classification, the accuracy using SVM employs only 59% of overall accuracy, weighted kNN depicts 72.13 % and ANN provides 93.4% overall accuracy for classification. The better results are obtained from ANN having more AUC of 0.96 than other classifiers.

Experiment 4: Performance measure of MSHVC after statistical analysis: The result of 3-stage classification after statistical analysis using SVM, kNN and ANN are tabulated in Table 7.

For 3- stage HRV classification, SVM employs 73.77% of overall accuracy, weighted kNN gives 93.4% and ANN provides 99% overall accuracy for classification. Table 7 states that the best results are acquired using ANN and a significant improvement are observed in overall accuracy after feature selection.

**Table 7. Performance measure of classification system for three stage classification after HRV statistical analysis.**

| Classification technique without feature selection / Performance metrices | | SVM | | | kNN | | | ANN | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | ***N*** | ***AR*** | ***AF*** | ***N*** | ***AR*** | ***AF*** | ***N*** | ***AR*** | ***AF*** |
| **Confusion matrix** | ***N*** | 15 | 2 | 1 | 18 | 0 | 0 | 18 | 0 | 0 |
| | ***AR*** | 0 | 7 | 13 | 1 | 17 | 0 | 0 | 20 | 0 |
| | ***AF*** | 0 | 0 | 23 | 2 | 1 | 22 | 1 | 0 | 22 |
| **PPV** | **$PPV_0$** | 100% | | | 85.71% | | | 94.73% | | |
| | **$PPV_1$** | 77.77% | | | 94.44% | | | 100% | | |
| | **$PPV_2$** | 62.16% | | | 100% | | | 100% | | |
| **Sensitivity** | **$Sen_0$** | 83.33% | | | 100% | | | 100% | | |
| | **$Sen_1$** | 23.33% | | | 94.44% | | | 100% | | |
| | **$Sen_2$** | 100% | | | 88% | | | 95.65% | | |
| **Specificity** | **$Sp_0$** | 100% | | | 93.02% | | | 100% | | |
| | **$Sp_1$** | 95.12% | | | 97.60% | | | 100% | | |
| | **$Sp_2$** | 63.15% | | | 100% | | | 100% | | |
| **Accuracy** | **$Acc_0$** | 95.08% | | | 95.08% | | | 99% | | |
| | **$Acc_1$** | 75.40% | | | 96.72% | | | 100% | | |
| | **$Acc_2$** | 77.04% | | | 95.08% | | | 99% | | |
| **Overall accuracy** | | **73.77%** | | | **93.4%** | | | **99%** | | |
| **AUC** | | **0.92** | | | **0.94** | | | **1.00** | | |

## 4.2. Percentage improvement in MSHVC system at various level of abstraction

The percentage improvement is computed to verify the performance at different levels of the proposed classification system. The authors have calculated the percentage improvement for three categories: percentage improvement of 2-stage classification, percentage improvement of 3-stage classification and percentage improvement of MSHVC system.

Percentage improvement of 2-stage classification: For the 2-stage classification phase, significant percentage improvement in the performance parameters for all the classification approaches before HRV statistical analysis and after HRV statistical analysis using three classifiers is observed. SVM provides the maximum percentage improvement of 11.7% in the overall accuracy and ANN depicts a maximum of 3 % improvement for Area under the Receiver Operating Characteristics (ROC) curve. Figure 3 visually represents the percentage improvement in accuracy and AUC.

For 2-stage classification system SVM, kNN and ANN give 11.7%, 1.6% and 3.3% of improvement respectively. These classifiers provide 2%, 1% and 3% percentage improvement in area under the curve.

Percentage improvement during 3-stage classification: For the 3-stage classification phase noteworthy percentage improvement in the accuracy and AUC parameters are seen before HRV statistical analysis and after HRV statistical analysis. The kNN achieved the maximum percentage improvement of 26.63% in the overall accuracy and SVM gives 7.14% improvement for Area under the Receiver Operating Characteristics (ROC) curve. Figure 4 depicts the percentage improvement in accuracy and AUC for 3-stage classification.

For 3-stage classification system SVM, kNN and ANN give 20.02%, 26.63% and 6.86 % of improvement respectively. These classifiers also provide 7.14%, 1.01% and 1% percent improvement in area under the curve.

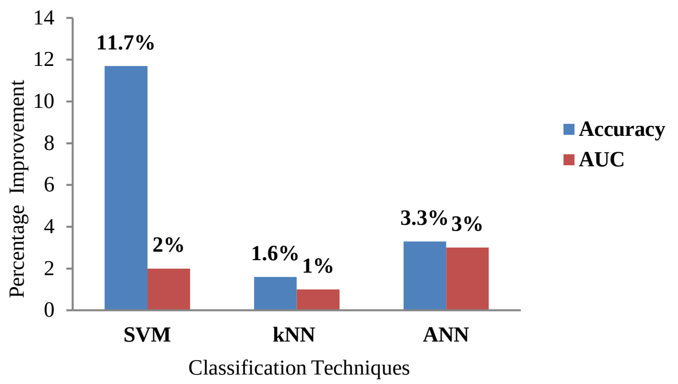


**Fig. 3. Percentage improvement for accuracy and AUC during 2-stage classification.**

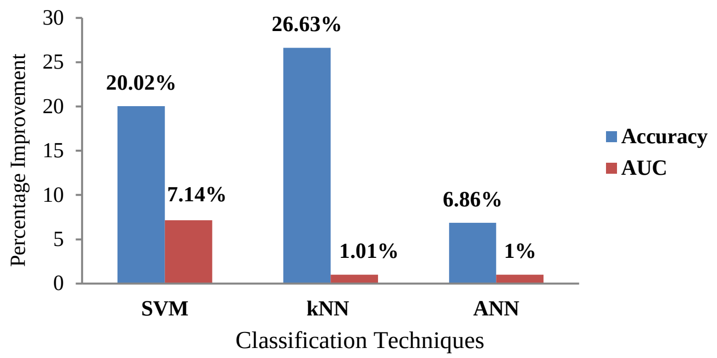


**Fig. 4. Percentage improvement for accuracy and AUC during 3-stage classification.**

Percentage improvement during MSHVC system: A considerable percentage improvement in all the classifiers have also been observed from 2-stage to 3-stage classification for the proposed methodology illustrated in Fig. 5. The SVM attained 8.52% improvement in overall accuracy and kNN achieved 22.77% of improvement in area under the curve.

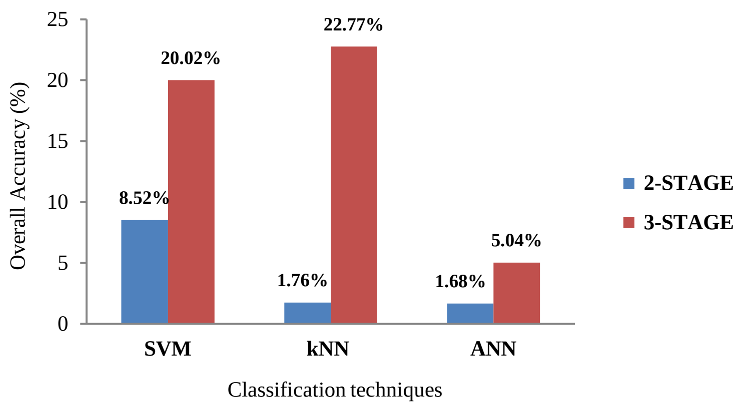


**Fig. 5. Percentage improvement during MSHVC classification.**

## 4.3. Comparison with existing state of art techniques

The suggested MSHVC system is compared with recent classification techniques. The presented methodology implemented multi-stage classification to diagnose AR ECG, AF ECG and NSR ECG. Table 8 presents the comparison based on the database used, number of features, time-length, classifier used and performance parameters. The presented methodology contributes multistage classification of AR, AF and NSR long-term ECG by extracting twenty HRV features from frequency-domain, time-domain and geometrical-domain. Popular classification techniques namely SVM, kNN and ANN are applied on nine significant features computed from ANOVA. It is observed that ANN gives the best performance among them for MSHVC classification.

Isler et al. proposed a 3-stage classifier to detect Congestive Heart Failure (CHF) from NSR. He extracted thirty-four features for 5-min of time length and obtained 98.8%, 98.1% and 100% of accuracy, specificity and sensitivity respectively [5]. Hu et al. [7] classifies similar database using SVM for 2-hour and 5-minutes of time length and achieve 94.44% and 96.67% accuracy respectively. Wang et al. [17] practice SVM with CHF and NSR on nine feature and attained 90.95% of accuracy, 90.03 % of specificity and 91.31% of sensitivity. K.H. Boon et al. used Atrial Fibrillation Prediction (AFPDB) in 2016 and 2018 to classify paroxyml atrial fibrillation using SVM and achieved 79.3% accuracy for 15-minute time length and improved accuracy of 87.7% for a 5-minute time length of ECG [3, 10]. Ensemble classifier is applied by Mahajan et al. [18] to diagnose CHF, NSR and long-term ST-T database based on three HRV features, he obtained 98.1%, 100% and 98.57 % of accuracy, specificity and sensitivity respectively. Lastly, Cornforth and Jelinek [24] applied kNN to detect CHF and NSR using four HRV feature for 1000 RR-intervals and attained 79.3 % of accuracy, 81.1% of specificity and 79.3 % of sensitivity [24]. Our proposed system outperforms among all the recent techniques and attained 99% overall accuracy.

We have presented a MSHVC system to detect AR, AF and NSR ECG. Detection of these diseases is done based on long-term HRV analysis. We have extracted twenty features in time-domain, frequency-domain and geometrical domain. Only 9 features are selected namely SDSD, RMSSD, NN50, pNN50, HRVTI, aULF, aVLF, aHF, and a total for classification. Four experiments have been performed during 2-stage classification and 3-stage classification. All the performance parameters are evaluated before statistical analysis and after statistical analysis.

The SVM, kNN and ANN classifiers are compared with MSHVC system. The conclusion has been drawn that for proposed system ANN has attained 99% overall accuracy at 2-stage and 3-stage classification. The innovation of this article manifested in the statistical study, which is used to extract the most significant features via ANOVA technique. It is observed that classification performance parameters of our proposed MSHVC system are significantly improved after statistical analysis. The future work will focus on the hardware implementation of MSHVC system on FPGA board using VIVADO tool.

**Table 8. Comparison with existing state of art techniques.**

| Previous work (year) | Database used | Number of features | Time length | Classification technique | Performance parameters |
|---|---|---|---|---|---|
| **Proposed MSHVC technique** | **NSR, AR and AF** | **9** | **24-hour** | **ANN** | Accuracy = 99%<br>Specificity = 100%<br>Sensitivity = 98.55% |
| **[5]** | NSR, CHF | 34 | 5-minutes | 3-stage classifier | Accuracy = 98.8 %<br>Specificity = 98.1%<br>Sensitivity =100% |
| **[7]** | NSR, CHF | 9 | 2- hour | SVM | Accuracy = 94.44%<br>Specificity = 98.33%<br>Sensitivity =86.67% |
| **[7]** | NSR, CHF | 3 | 5 minutes | SVM | Accuracy = 96.67%<br>Specificity = 98.33%<br>Sensitivity =93.33% |
| **[17]** | NSR, CHF | 9 | 5- minutes | SVM | Accuracy = 90.95%<br>Specificity = 90.03%<br>Sensitivity = 91.31% |
| **[3]** | AFPDB | 25 | 5- minutes | SVM | Accuracy = 87.7%<br>Specificity = 88.7%<br>Sensitivity = 86.8% |
| **[18]** | NSR, CHF, Long term ST-T | 3 | 24-hour | Ensemble of bagged decision tree | Accuracy = 98.1%<br>Specificity = 100%<br>Sensitivity = 98.57% |
| **[24]** | NSR, CHF | 4 | 1000 RR-intervals | kNN | Accuracy = 87.9%<br>Specificity = 94.4%<br>Sensitivity = 80% |
| **[10]** | AFPDB | 26 | 15- minutes | SVM | Accuracy = 79.3%<br>Specificity = 81.1%<br>Sensitivity = 79.3% |

## References

1. Kirti; Sohal, H.; and Jain, S. (2018). Interpretation of cardio vascular diseases using electrocardiogram: A study. *Proceedings of the Fifth International Conference on Parallel, Distributed and Grid Computing (PDGC).* Solan Himachal Pradesh, India, 159-164.
2. Borchini, R.; Veronesi, G.; Bonzini, M.; Gianfagna, F.; Dashi, O.; and Ferrario, M.M. (2018). Heart rate variability frequency domain alterations among healthy nurses exposed to prolonged work stress. *International Journal of Environmental Research and Public Health,* 15(1), 11 pages.
3. Boon, K.H.; Khalil-Hani, M. and Malarvili, M.B. (2018). Paroxysmal atrial fibrillation prediction based on HRV analysis and non-dominated sorting genetic algorithm III. *Computer Methods and Programs in Biomedicine*, 153, 171-184.
4. Kirti; Sohal, H. and Jain, S. (2018). Comparative analysis of heart rate variability parameters for arrhythmia and atrial fibrillation using ANOVA. *Biomedical and Pharmacology Journal*, 11(4), 1841-1849.
5. Isler, Y.; Narin, A.; Ozer, M.; and Perc, M. (2019). Multi-stage classification of congestive heart failure based on short-term heart rate variability. *Chaos, Solitons & Fractals,* 118, 145-151.
6. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. (1996). Heart rate variability: Standards of measurement, physiological interpretation, and clinical use. *European Heart Journal*, 17, 354-381.
7. Hu, B.; Wei, S.; Wei, D.; Zhao, L.; Zhu, G.; and Liu, C. (2019). Multiple time scales analysis for identifying congestive heart failure based on heart rate variability. *IEEE Access,* 7, 17862-17871.
8. Kirti; Sohal, H. and Jain, S. (2019). FPGA implementation of power-efficient ECG pre-processing block. *International Journal of Recent Technology and Engineering (IJRTE)*, 8(1), 2988-2904.
9. Prashar, N.; Sood, M.; and Jain. S. (2018). Semiautomatic detection of cardiac diseases employing dual tree complex wavelet transform. *Periodicals of Engineering and Natural Sciences*, 6(2), 129-140.
10. Boon, K.H.; Khalil-Hani, M.; Malarvili, M.B.; and Sia, C.W. (2016). Paroxysmal atrial fibrillation prediction method with shorter HRV sequences. *Computer Methods and Programs in Biomedicine*, 134, 187-196.
11. Acharya, U.R.; Fujita, H.; Adam, M.; Lih, O.S.; Sudarshan, V.K.; Hong, T.J.; Koh, J.E.W.; Hagiwara, Y.; Chua, C.K.; Poo, C.K.; and San, T.R. (2017). Automated characterization and classification of coronary artery disease and myocardial infarction by decomposition of ECG signals: A comparative study. *Information Sciences,* 377, 17-29.
12. Kim, Y.J.; Heo, J.; Park, K.S. and Kim, S. (2016). Proposition of novel classification approach and features for improved real-time arrhythmia monitoring. *Computers in Biology and Medicine*,75,190-202.
13. Castaldo, R.; Melillo, P.; Izzo, R.; De Luca, N.; and Pecchia, L. (2017). Fall prediction in hypertensive patients via short-term HRV analysis. *IEEE Journal of Biomedical and Health Informatics,* 21(2), 399-406.

14. Castaldo, R.; Montesinos, L.; Melillo, P.; James, C. and Pecchia, L. (2019). Ultra-short term HRV features as surrogates of short term HRV: A case study on mental stress detection in real life. *BMC Medical Informatics and Decision Making*, 19(1), 13 pages.
15. Castaldo, R.; Xu, W.; Melillo, P.; Pecchia, L.; Santamaria, L. and James, C. (2016). Detection of mental stress due to oral academic examination via ultra-short-term HRV analysis. *Proceedings of the 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. Orlando, Florida, 3805-3808.
16. Fonseca, P.;. Long, X.; Radha, M.; Haakma, R.; Aarts, R.M.; and Rolink, J. (2015). Sleep stage classification with ECG and respiratory effort. *Physiological Measurement*, 36(10), 2027-2040.
17. Wang, Y.; Wei, S.; Zhang, S.; Zhang, Y.; Zhao, L.; Liu, C.; and Murray, A. (2018). Comparison of time-domain, frequency-domain and non-linear analysis for distinguishing congestive heart failure patients from normal sinus rhythm subjects. *Biomedical Signal Processing and Control*, 42, 30-36.
18. Mahajan, R.; Viangteeravat, T.; and Akbilgic, O. (2017). Improved detection of congestive heart failure via probabilistic symbolic pattern recognition and heart rate variability metrics. *International Journal of Medical Informatics*, 108, 55-63.
19. Mufassirin, M.M.M.; and Ragel, R.G. (2018). A novel filter-wrapper based statistical analysis approach for cancer data classification. *Proceedings of the IEEE International Conference on Information and Automation for Sustainability (ICIAfS).* Columbo, Sri Lanka, 6 pages.
20. Yang, A.C.C.; and Yin, H.W. (2001). Prediction of paroxysmal atrial fibrillation by footprint analysis. *Computers in Cardiology*, 28, 401-404.
21. Isler, Y.; and Kuntalp, M. (2007). Combining classical HRV indices with wavelet entropy measures improves to performance in diagnosing congestive heart failure. *Computers in Biology and Medicine,* 37(10), 1502-1510.
22. Liu, G.; Wang, L.; Wang, Q.; Zhou, G.M.; Wang, Y.; and Jiang, Q. (2014). A new approach to detect congestive heart failure using short-term heart rate variability measures. *PloS One*, 9(4), 8 Pages.
23. Pandey, A.K.; Pandey, P.; and Jaiswal, K.L. (2014). Classification model for the heart disease diagnosis. *Global Journal of Medical Research: F Diseases,* 14(1), 9-14.
24. Cornforth, D.J.; and Jelinek, H.F. (2016). Detection of congestive heart failure using Renyi entropy. *Proceedings of the Computing in Cardiology Conference (CinC).* Vancouver, Canada, 669-672.
25. Pan, J.; and Tompkins, W.J. (1985). A real-time QRS detection algorithm. *IEEE Transactions on Biomedical Engineering*, 32(3), 230-236.
26. Goldberger, A.L.; Amaral, L.A.; Glass, L.; Hausdorff, J.M.; Ivanofff, P.C.; Mark, R.G.; Mietus, J.E.; Moody, G.B.; Peng, C.K.; and Stanley, H.E. (2000). PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation*, 101(23), 6 pages.
27. Poddar, M.G.; Birajdar, A.C.; Virmani, J.; and Kriti. (2019). Automated classification of hypertension and coronary artery disease patients by PNN,

KNN, and SVM classifiers using HRV analysis. *Machine Learning in Bio-Signal Analysis and Diagnostic Imaging*, 99-125.

28. Rahi, P.K.; and Mehra, R. (2014). Analysis of power spectrum estimation using welch method for various window techniques. *International Journal of Emerging Technologies and Engineering (IJETE)*, 2(6), 106-109.
29. Melillo, P.; Izzo, R.; Orrico, A.; Scala, P.; Attanasio, M.; Mirra, M.; De Luca, N.; and Pecchia, L. (2015). Automatic prediction of cardiovascular and cerebrovascular events using heart rate variability analysis. *PloS One*, 10(3), 14 pages.
30. Hsu, C.-W.; Chang, C.-C. and Lin, C.-J. (2003). A practical guide to support vector classification. *Technical Report*, Department of Computer Science and Information Engineering, University of National Taiwan, Taipei, 12 pages.
31. Park, D.; Lee, M.; Park, S.E.; Seong, J.-K.; and Youn, I. (2018). Determination of optimal heart rate variability features based on SVM-recursive feature elimination for cumulative stress monitoring using ECG sensor. *Sensors*, 18(7), 15 Pages.
32. Bicego, M.; and Loog, M. (2016). Weighted K-nearest neighbor revisited. *Proceedings of the 23rd International Conference on Pattern Recognition (ICPR).* Cancun, Mexico, 1643-1648.
33. Royall, R.M. (1966). A class of non-parametric estimates of a smooth regression function. *Technical Report No. 14,* Stanford University, Department of Statistics, 68 pages.
34. Dudani, S.A. (1976). The distance-weighted k-nearest-neighbor rule. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-6(4), 325-327.
35. Parsons, S.P.; and Huizinga, J.D. (2019). Robust and fast heart rate variability analysis of long and noisy electrocardiograms using neural networks and images. *arXiv:arXiv:1902.06151,* 21 pages.
36. Sieciński, S.; Kostka, P.S.; and Tkacz, E.J. (2019). Comparison of atrial fibrillation detection performance using decision trees, SVM and Artificial Neural Network. *International Conference on Information Technology & Systems (ICITS).* Quito, Ecuador, 693-701.