computer methods and programs in biomedicine 127 (2016) 52–63

journal homepage: www.intl.elsevierhealth.com/journals/cmpb

Arrhythmia recognition and classification using combined linear and nonlinear features of ECG signals

Fatin A. Elhaj[a,*], *Naomie Salim*[a,*], *Arief R. Harris*[b,c,**], *Tan Tian Swee*[b,c,**], *Taqwa Ahmed*[a]

[a] *Soft Computing Research Group, Faculty of Computing, Universiti Teknologi Malaysia, Skudai 81310, Malaysia*
[b] *Neural Engineering Lab, Centre for Biomedical Engineering, Universiti Teknologi Malaysia, Malaysia*
[c] *Faculty of Biosciences and Medical Engineering, Universiti Teknologi Malaysia, Johor Malaysia, Malaysia*

ARTICLE INFO

Article history:
Received 16 August 2015
Received in revised form
13 December 2015
Accepted 14 December 2015

Keywords:
ECG
Linear features
Nonlinear features
Classification

ABSTRACT

Arrhythmia is a cardiac condition caused by abnormal electrical activity of the heart, and an electrocardiogram (ECG) is the non-invasive method used to detect arrhythmias or heart abnormalities. Due to the presence of noise, the non-stationary nature of the ECG signal (i.e. the changing morphology of the ECG signal with respect to time) and the irregularity of the heartbeat, physicians face difficulties in the diagnosis of arrhythmias. The computer-aided analysis of ECG results assists physicians to detect cardiovascular diseases. The development of many existing arrhythmia systems has depended on the findings from linear experiments on ECG data which achieve high performance on noise-free data. However, nonlinear experiments characterize the ECG signal more effectively sense, extract hidden information in the ECG signal, and achieve good performance under noisy conditions. This paper investigates the representation ability of linear and nonlinear features and proposes a combination of such features in order to improve the classification of ECG data. In this study, five types of beat classes of arrhythmia as recommended by the Association for Advancement of Medical Instrumentation are analyzed: non-ectopic beats (N), supra-ventricular ectopic beats (S), ventricular ectopic beats (V), fusion beats (F) and unclassifiable and paced beats (U). The characterization ability of nonlinear features such as high order statistics and cumulants and nonlinear feature reduction methods such as independent component analysis are combined with linear features, namely, the principal component analysis of discrete wavelet transform coefficients. The features are tested for their ability to differentiate different classes of data using different classifiers, namely, the support vector machine and neural network methods with tenfold cross-validation. Our proposed method is able to classify the N, S, V, F and U arrhythmia classes with high accuracy (98.91%) using a combined support vector machine and radial basis function method.

* *Corresponding authors.*
** *Corresponding authors at*: Neural Engineering Lab, Centre for Biomedical Engineering, Universiti Teknologi Malaysia, Malaysia.
E-mail addresses: fatin_elhaj@hotmail.com (F.A. Elhaj), naomie@utm.my (N. Salim), arief@biomedical.utm.my (A.R. Harris), tantswee@utm.my (T.T. Swee).

http://dx.doi.org/10.1016/j.cmpb.2015.12.024

1. Introduction

Cardiovascular disease (CVD) has become one of the main causes of death in the world [1]. The increase in death rates due to CVD is associated with the epidemiological transition of unhealthy lifestyles such as obesity, diabetes mellitus and smoking habits [2]. The American Heart Association stated that, in 2006, over 70 million people around the world were diagnosed with CVD. The common causes of CVD are high blood pressure, insufficient physical exercise, poorly balanced diet, smoking and abnormal glucose levels [3]. Two of the main factors of CVD are atrial and ventricular arrhythmias which disturb the cardiac rhythm. Arrhythmia is a cardiac condition caused by the abnormal electrical activity of the heart. Cardiac arrhythmias generally occur in people with diseases, such as cardiomyopathy, hypertension and coronary artery disease. The electrical activity of the heart can be recorded using surface electrodes. The recording describes the main direction of electrical impulses throughout the heart for any length of time and is known as the electrocardiogram (ECG). The ECG is a non-invasive method used to detect arrhythmias or heart abnormalities. Due to the presence of noise and the irregularity of the heartbeat, physicians face difficulties in the diagnosis of arrhythmias. Furthermore, visual inspection alone may lead to misdiagnosis or insignificant detection of arrhythmias [4]. Therefore, the computer-aided analysis of ECG data assists physicians to efficiently detect arrhythmia [5,6].

There are three main procedures in an ECG arrhythmia detection system, namely, feature extraction, feature selection, and classifier construction [7]. As a one key procedure, feature extraction transforms the input data into a set of features and plays an important role in detecting most heart diseases.

2. Related work

It is vital to extract the hidden information present in the ECG signals and improve the classification performance. Several feature extraction techniques are used for the analysis of ECG signals, and these are categorized into time, frequency and time–frequency techniques.

In the time domain, the extracted features are the heartbeat interval, duration parameters (QRS, QT, and PR) and; amplitude parameters (QRS, ST) [8]. Due to the subtle changes in the amplitude and duration in the ECG, time methods do not provide good discrimination [9]. Therefore, frequency methods such as the Fourier transform and the power spectral density (PSD) are used. However, the frequency methods do not provide temporal information from the ECG signals. A proper time–frequency technique can tackle this problem. The wavelet transform is the most widely used time–frequency method [10–13]. It provides a good resolution in both the time and frequency domains.

In wavelet analysis, the selection of an optimal dimensionality reduction method is important before applying the feature vector to a classifier. Linear and non-linear transformation methods have been critical to the success of the time–frequency-based feature sets developed during the last decade. Martis et al. [14] applied a linear method of discrete wavelet transform (DWT) coefficients with the feature reduction technique of principal component analysis (PCA) to discriminant features between normal and arrhythmia classes. A good classification accuracy of 98.78% with a neural network (NN) was obtained to classify the five main beat classes recommenced by the ANSI/AAMI EC57:1998. Most of the linear experiments on ECG data are noise-free, providing good classification accuracy. However, these linear methods may not obtain the same maximum accuracy in the presence of noise [15].

The ECG signal is non-stationary, that is, the morphology of the ECG signal changes with respect to time, and these variations are present not only between different patients but also within the same patients [16]. The ECG signal can be represented using nonlinear experiments which characterize the signal more effectively, extract the hidden information in the ECG signal, and perform well under noisy conditions [17]. For example, the use of higher order spectra (HOS) cumulants provides very good recognition and is less affected by the morphological changes of the ECG signal [15,16]. In one study [16], the HOS was applied to recognize the normal and different types of arrhythmias using the fuzzy NN, and the cumulants of the second, third and fourth orders were used as the selected features. The recognition result for seven beat types showed 96.06% accuracy. In another study [18], the HOS bispectrum (which is the third order spectra) was used to capture information beyond the mean and standard deviation. The PCA used the extracted bispectrum features for dimensionality reduction. Using the feed-forward NN and least square-support vector machine (SVM), five types of beats were classified. Accuracy of 93.48% was obtained in that study. However, nonlinear methods are computationally complex, do not include the symmetry and reflection properties, and do not follow the principle of superposition; hence, additivity and homogeneity principles are not followed by these methods [15].

In the literature, no special feature or features have been determined to differentiate one ECG beat from another. The available classifiers are mostly based on artificial neural networks (ANNs), a mixture of the expert approach and fuzzy logic, support vector machine, and the probability neural network (PNN). Details of the classifiers most commonly used in the literature are shown in Table 1.

Kutlu and Kuntalp [19] proposed an arrhythmia recognition system based on a combination of diverse features including higher order statistics, morphological features, Fourier transform coefficients, and higher order statistics of the wavelet package coefficients. Using a wrapper type feature selection algorithm to determine the optimal features, the optimal features are selected similarly and then only a vector of data is used to classify the beats into the five main groups defined by the AAMI. The classification accuracy values of the k-nearest neighbor were 85.59%, 95.46%, and 99.56%, for average sensitivity, average selectivity and average specificity, respectively. Das and Ari [20] combined two features, namely, the S-transform (ST) and wavelet transform, to select the features more effectively than using them independently. Using the NN classifier of the appending feature vector to classify the five classes of ECG beats recommended by the AAMI, Übeyli [21] used Lyapunov exponents, wavelet coefficients and

Table 1 – Most commonly used classifiers in the literature.

Classifier	Ref.
Fuzzy hybrid NN	Osowski and Linh [16]
Linear discriminant classifier	De Chazal and Reilly [22]
NN	Inan et al. [38]
Block-based NNs (BbNNs)	Jiang and Kong [39]
Adaptive wavelet NN (WNN)	Lin et al. [40]
Adaptive neuro-fuzzy Inference system (ANFIS)	Übeyli [41]
Extreme learning machine (ELM)	Kim et al. [29]
Recurrent NN (RNN)	Derya Übeyli [42]
Bayesian filter	Sayadi et al. [43]
Fuzzy-SVM (FSVM)	Ozcan and Gurgen [44]
Probability NN (PNN)	Wang et al. [7]
SVM + ELM	Banupriya and Karpagavalli [45]

power spectral density (PSD) methods as a set of features and achieved average accuracy of 93.89%. The time domain features such as morphological and temporal features have also been combined to classify the five classes of ECG signal using a linear discriminant classifier, obtaining 85.9% accuracy [22].

However, all of the above techniques have the following disadvantages:

(i) It is probable that the ECG signal is masked by noise and artifacts and most of the maximum accuracy is achieved in noise-free conditions [15].
(ii) Some of the feature extraction methods are computationally complex, do not include the symmetry and reflection properties, and do not follow the principle of superposition; hence, additivity and homogeneity principles are not followed by these methods [15].

Hence, a combination of linear features with nonlinear features would serve as an excellent solution for these drawbacks.

This paper proposes a new technique for arrhythmia recognition and classification of ECG signals based on a combination of linear and nonlinear feature extraction techniques for noisy and noise-free conditions, compared with the linear techniques that have less computational complexity. This combined technique is able to identify the most discriminant features between normal and abnormal classes of the input signals. A linear PCA of the DWT method is applied, transforming the ECG beats using DWT and then applying the dimensionality reduction method of PCA to extract the features. Additionally, the nonlinear HOS cumulants method and a nonlinear dimensionality reduction technique ([ICA]) are used for the analysis of nonlinear ECG signals. These features with reduced dimensionality are fed to the NN and SVM with the radial basis kernel function for automated discrimination between normal and abnormal classes of ECG signals.

The paper is organized as follows: Section 2 presents the materials used, Section 3 deals with the methods and classifiers used, Section 4 provides the experimental results, and Section 5 discusses the results. Finally, Section 6 concludes the paper.

Table 2 – MIT-BIH arrhythmia beats classification per ANSI/AAMI EC57:1998 standard database.

AAMI classes	MIT-BIH heartbeat classes
Non-ectopic beat (N)	Normal beat (N)
	Left bundle branch block (LBBB)
	Right bundle branch block (RBBB)
	Nodal (junctional) escape (j)
	Atrial escape beat (e)
Supra-ventricular ectopic beat (S)	Aberrated atrial premature (A)
	Atrial premature (a)
	Supraventricular premature (S)
	Nodal (junctional) premature (J)
Ventricular ectopic beat (V)	Ventricular escape (V)
	Premature ventricular contraction (E)
Fusion beat (F)	Fusion of ventricular and normal (F)
Unknown beat (U)	Unclassifiable (U)
	Paced (p)
	Fusion of paced and normal (f)

3. Materials

The source of the ECG records used in this study is the MIT-BIH arrhythmia database [23]. The database contains 48 signals sampled at 360 Hz, and each of the 48 signals is a little more than 30 min long. Twenty-three recordings were selected randomly from a set of 4000 24-h ambulatory ECG recordings collected from patients at Boston's Beth Israel Hospital; the remaining 25 recordings were selected from the same set to include other clinically significant arrhythmias that may not be well-represented in a small random sample group because they are less common. An ECG beat is considered to be one cycle of the electrical activity of the heart. The irregularities of the beats are called ectopic beats. The ANSI/AAMI EC57:1998 standard recommends the entire MIT-BIH arrhythmia data should be grouped into five main beat classes [14]. The five classes are non-ectopic beats (N), fusion beats (F), supraventricular ectopic beats (S), ventricular ectopic beats (V) and unknown beats (U). The data used in this study involved (90, 580) N, (2973) S, (7707) V, (1784) F and (7050) U. Table 2 shows the details of the different beats of the MIT-BIH database and summarizes the five types of ECG beat samples used in this paper.

4. Methods

The proposed automated recognition system adopts different methods following different pre-processing, feature extraction and classifier methods. Fig. 1 shows a general diagram of the constructed system. The work concerning each part is explained in detail in the following sections.

4.1. *Pre-processing*

The ECG signal contains different natures of noise. Contact noise, baseline drift, muscle artifacts, powerline interference, electrode motion artifact, electromyography artifact, data-collecting device noise and quantization noise are examples of noise. The quality of the ECG classification depends on the accuracy of the detection of each cardiac cycle [24]. In this work, the pre-processing module is decomposed into three

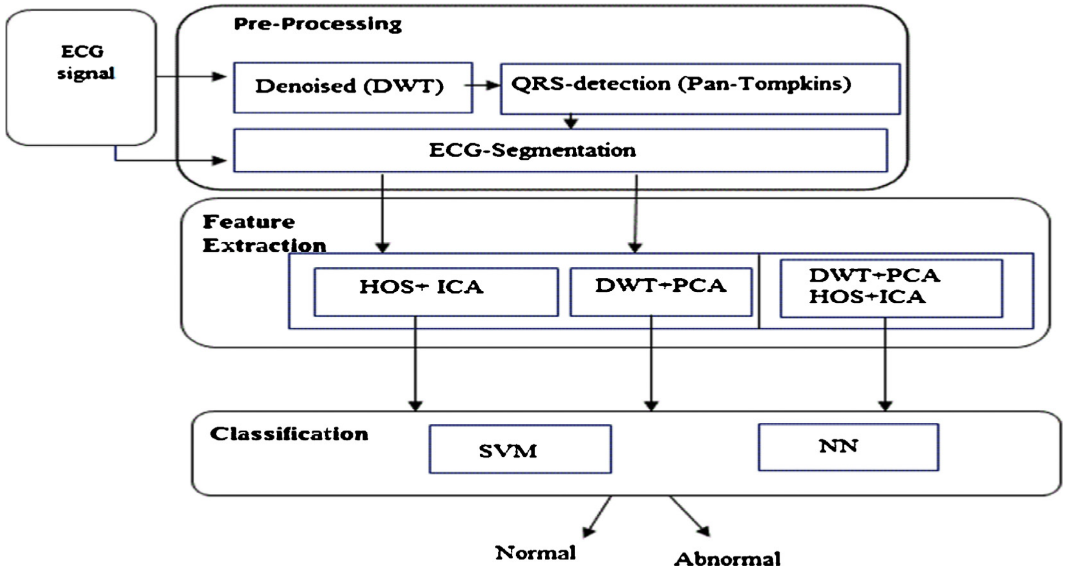

Fig. 1 – Proposed automated recognition system.

components as shown above in Fig. 1. The first component involves the denoising of the ECG signal based on the DWT, the second component is the detection of the QRS complex (i.e. the combination of three of the graphical deflections seen on a typical ECG, namely, the Q, R and S waves), and the last component is the segmentation of the signal. Each component is described in brief as follows.

4.1.1. *Denoising*

Since DWT is efficient in analyzing non-stationary signals, it is used in this paper. The Daubechies D6 (db6) wavelet basis function is used to denoise the data, with the ECG signals decomposed to nine levels [25]. A frequency range from 0 to 0.351 Hz in the ninth level approximation sub-band is mostly the baseline wander, and is not used for reconstructing the denoised signal. Furthermore, the frequency ranges of 90–180 and 45–90 were not considered because the information after 45 Hz is not important for arrhythmia detection. The inverse wavelet transform is computed from the third to the ninth level detail sub-bands to obtain the denoised and smoothed ECG signal [25].

come capisco le frequenze?

4.1.2. *QRS complex-detection*

The Pan–Tompkins algorithm was used on the denoised ECG signal algorithm for detection of the QRS complex [26]. It analyzes the slope, amplitude and width. The algorithm consists of three steps. In the first step, the low pass and high pass filters form a band pass filter, which reduces noise in the ECG signal like muscle noise. In the second step, to distinguish QRS complexes from low-frequency ECG components such as the P and T waves, the signal is passed through a differentiator to highlight the high slopes. The third step is the squaring operation, which places stress on the higher values that are mainly present because of QRS complexes. Then, the squared signal passes through a Moving-Window Integrator of the length of the window. The result is a smooth peak ECG cycle. The output of the Moving-Window Integrator may be used to find QRS complexes.

4.1.3. *Segmentation*

After detection of the QRS complex, 99 samples were chosen from the left side of the QRS mid-point and 100 samples after QRS mid-point and the QRS mid-point itself as a segment or beat of 200 samples [14]. Fig. 2 represents the five types of beat classes in the ANSI/AAMI EC57:1998 standard database in two pre-processing categories: (a) smoothing signal beats using DWT + QRS detection of 200 samples; and (b) QRS detection of 200 samples. The circle illustrates the presence of non-smoothing signal beats.

4.2. *Feature extraction techniques*

Automatic ECG beat recognition and classification depends on various features [27]. The feature sets from the ECG data set were created by extracting the linear and nonlinear features. In the following sub-sections, the linear and nonlinear methods are discussed.

4.2.1. *Linear method based wavelets and principle component analysis*

DWT is used to extract hidden information from the ECG signal [14]. In this study, it was applied after ECG signal denoising, QRS complex detection and ECG segmentation. The PCA was applied on both the sub-band coefficients of the third level detail and fourth level detail. The third level detail and fourth level detail of the DWT are shown in Fig. 3. From each of the sub-bands, the first six columns of the p variable were regarded as the six features for subsequent classification. In total, 12 features (six each from the two sub-bands) were used for subsequent pattern identification using classifiers. The wavelet transform displays good temporal localization in both time and frequency domains. It provides discrimination between two normal and abnormal classes. The DWT

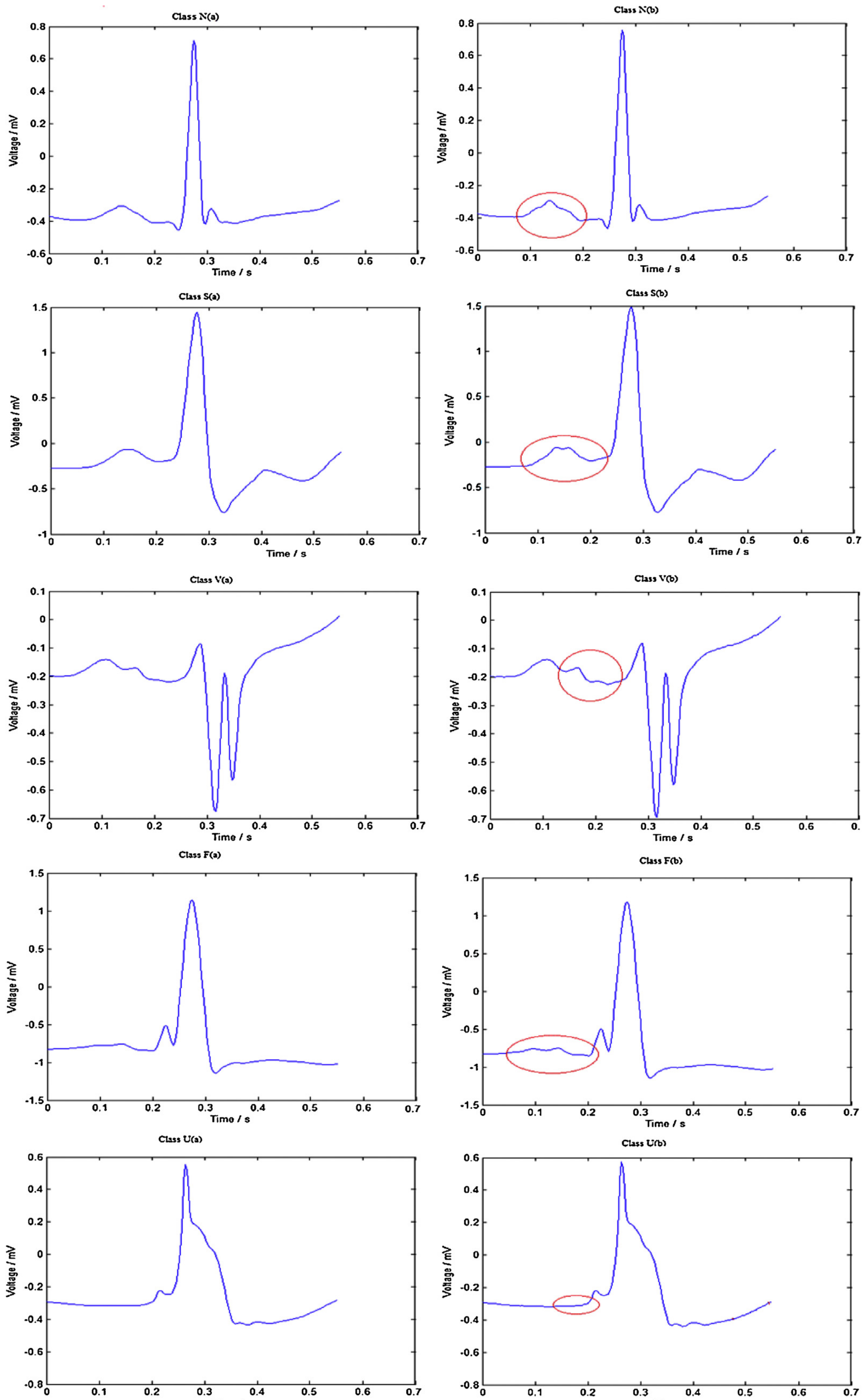

Fig. 2 – Five types of beat classes in the ANSI/AAMI EC57:1998 standard database in two pre-processing categories – (a) smoothing signal beats using DWT, and (b) signal beats without using DWT.

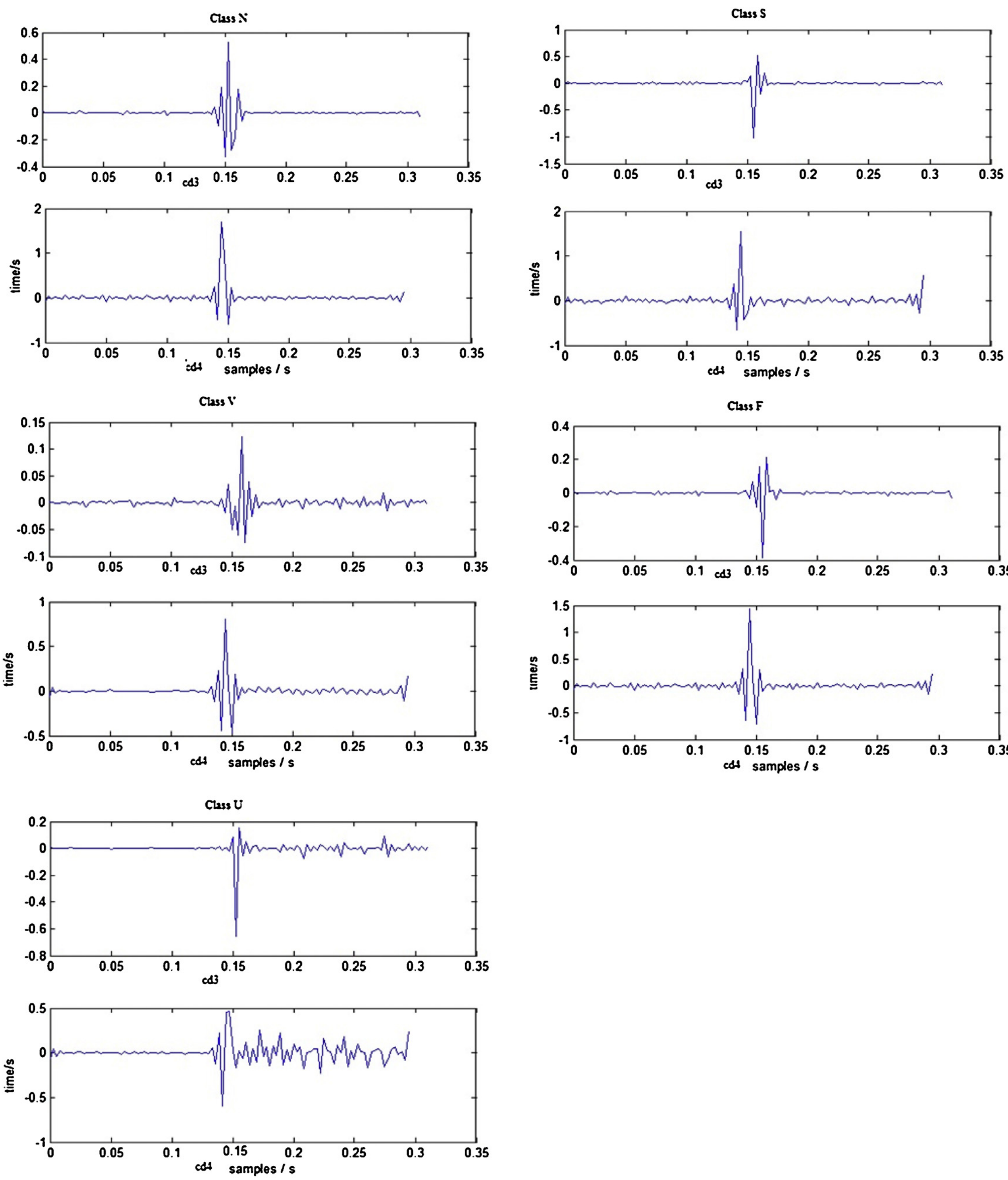

Fig. 3 – Third level detail (cd3) and fourth level detail (cd4) of DWT for the five types of beat classes in the ANSI/AAMI EC57:1998 standard database.

technique decomposes a signal into low frequency and high frequency components. The low frequency component is called the approximation, and the high frequency component is called the detail.

Using a filter bank, a low pass filter $h(n)$ down-samples the ECG signal $x(n)$ by a factor of two to obtain the approximation coefficients at level one. The high pass filter is given by:

$$g(L-1-n) = (-1)^n h(n) \tag{1}$$

where L is the length of the filter. The detail coefficient was obtained by passing the signal through $g(n)$ and then down-sampling by a factor of two. The DWT filtering and the sub-sampling are given by:

$$y_{\text{low}}(k) = \sum_n x(n)h(-n+2k) \tag{2}$$

and

$$y_{\text{high}}(k) = \sum_n x(n)g(-n+2k) \tag{3}$$

A detailed review is provided by Addison [28].

After implementing the pre-processing, 200 samples as one ECG beat were decomposed into four level approximations of the frequency range from 0 to 11.25 Hz using Mayer's wavelet. The fourth level detail consists of the 11.25–22.5 Hz frequency range.

PCA is one of the main linear dimensionality reduction techniques for extracting effective features from high dimensions. Usually PCA is done by projecting the data into the feature space and finding the correlation among those features. It computes the principal components as a percentage of the total variability of the data used to select a number of them. The first principal component provides the vector of the highest variability (uncorrelated). The second principal component provides the vector for the next direction orthogonal to the first principal component and so on. PCA is widely used to reduce the dimensions of ECG arrhythmia detection (e.g. Kim et al. [29] and Polat and Güneş [30]). Computation of the principal components includes a computation of the covariance matrix of the data, its eigenvalue decomposition, sorting of eigenvectors in the decreasing order of eigenvalues and projection of the data into the new basis defined by the principal components by taking the inner product of the original signals and the sorted eigenvectors. The PCA involves the following three phases:

Phase 1: Calculate covariance matrix from the data as,

$$C = (x - x^{\sim})(x - x^{\sim})^{T} \tag{4}$$

where x is the data matrix, and $x^{\sim}$ represents mean vector of x.

Phase 2: Calculate the array of eigenvectors V and diagonal matrix of eigenvalues D as

$$V^{-1}CV = D \tag{5}$$

Phase 3: Sort the eigenvectors in V in descending order of eigenvalues in D and project the data on these eigenvector directions by taking the inner product between the data matrix and the sorted eigenvector matrix as,

$$p = [V^{T}(x - x^{\sim})^{T}]^{T} \tag{6}$$

PCA was applied on both sub-band coefficients of the third level detail and fourth level detail. From each of the sub-bands, the first six columns of the p variable were regarded as the six features for subsequent classification. In total, 12 features (six each from the two sub-bands) were used for subsequent pattern identification using classifiers.

4.2.2. *Nonlinear method-based high order statistic and multivariate analysis*

In order to illustrate the nonlinear properties in ECG preprocessing, we utilized only two components, namely, the QRS complex detection and ECG segmentation of the ECG signals. HOS cumulant features and nonlinear feature reduction techniques such as ICA were applied.

The HOS technique is used for the analysis of nonlinear, non-stationary and non-Gaussian signals [16,31]. The nonlinearity of HOS has been generated from the Gaussian deviation and phase mutual relations between various frequency components. Moreover, HOS removes Gaussian noise. As the Gaussian process is categorized by the first two order statistics, the first and second order statistics are not significant to represent the nonlinear model. Therefore, the third and fourth order statistics are used in this analysis. Suppose that $x(n)$ is a stationary signal of a discrete time and its moments are available up to the order n. Then, the nth order moment function is defined by:

$$m_n^x(\tau_1, \tau_2, \ldots, \tau_{n-1}) = E[x(n)x(n+\tau_1)\ldots x(n+\tau_{n-1})] \tag{7}$$

The moment function depends only on the time lags τ_1, τ_2, ..., τ_{n-1}, $\tau_i = 0, \pm 1, \pm 2$ for all i. The second order moment $m_2^m(\tau_1)$ is the autocorrelation sequence of $x(n)$ and is calculated as follows:

$$m_2^m(\tau_1) = E[x(n)x(n+\tau_1)] \tag{8}$$

Whereas the third and fourth order moment functions are given by $m_3^x(\tau_1, \tau_2)$ and $m_4^x(\tau_1, \tau_2, \tau_3)$, respectively. $E[\cdot]$ is the statistical expectation operator. Using the nth order moment, the nth order cumulant can be computed as:

$$C_n^x(\tau_1, \tau_2, \ldots, \tau_{n-1}) = m_n^x(\tau_1, \tau_2, \ldots, \tau_{n-1}) - m_n^G(\tau_1, \tau_2, \ldots, \tau_{n-1})$$

where $m_n^x(\tau_1, \tau_2, \ldots, \tau_{n-1})$ is the nth order moment function and $m_n^G(\tau_1, \tau_2, \ldots, \tau_{n-1})$is the nth order moment of the Gaussian process. The first four order cumulants for a zero mean process are given by:

$$C_1^x = m_1^x$$

$$C_2^x = m_2^x(\tau_1)$$

$$C_3^x = m_3^x(\tau_1, \tau_2)$$

$$C_4^x = m_4^x(\tau_1, \tau_2, \tau_3) - m_2^x(\tau_1)m_2^x(\tau_2 - \tau_3) - m_2^x(\tau_2)m_2^x(\tau_3 - \tau_1) - m_2^x(\tau_3)m_2^x(\tau_1 - \tau_2)$$

A nonlinear method of dimensionality reduction, ICA is a computational statistical technique that is very effective in disclosing the hidden factors underlying mixed samples of random variable measurements [15,32]. This technique was originally developed to deal with situations that are closely related to blind source separation (e.g., the 'cocktail party' problem [32,33]). ICA involves a multi-variant analysis to reduce a multi-source signal into additive subcomponents. It can be represented mathematically as follows:

$$X = A \cdot S \tag{9}$$

where X is the matrix of n observed signals, $X = [x_1, x_2, \ldots, x_n]^T$, S is the matrix of m underlying signals, $S = [s_1, s_2, \ldots, s_n]^T$ and A is the mixing matrix of $[n^X m]$. The number of independent components (ICs) can be calculated by finding an un-mixing matrix named w where:

$$S = w \cdot X \tag{10}$$

In order to calculate w, the data must be centered and whitened. The most common method for whitening is the eigenvalue decomposition of the covariance matrix:

$$\mathrm{EXX}^{\mathrm{T}} = EDE^{\mathrm{T}} \tag{11}$$

where E is the orthogonal matrix of eigenvectors of $\mathrm{EXX}^{\mathrm{T}}$, and D is the diagonal matrix of its eigenvalues $D = \mathrm{diag}(d_1, d_2, \ldots, d_n)$. The whitened matrix X^{-} can be calculated as:

$$X^{-} = ED^{1/2}E^{\mathrm{T}}X \tag{12}$$

Whitening transforms the unmixing matrix for which:

$$\hat{S} = w \cdot X = wED^{-1/2}E^{\mathrm{T}}X^{-} = w^{-}X \tag{13}$$

Since w^{-} is orthogonal, this minimizes the number of parameters to be estimated. Therefore, the unmixing matrix w:

$$w = w^{-}ED^{-1/2}E^{\mathrm{T}} \tag{14}$$

Subsequently, the matrix A is obtained from the following equation:

$$A = (w^{\mathrm{T}}w)^{-1}w^{\mathrm{T}} \tag{15}$$

4.3. Combination of linear and nonlinear methods

The feature sets from the ECG data set were created by combining the linear and nonlinear features. The combined feature set is formed by appending the twelve PCA of DWT features, sixteen ICA features and HOS cumulant features. It may be reasonable to consider a feature vector to be composed of a linear feature and a nonlinear feature. That is:

$$Z_t = L_t + N_t \tag{16}$$

where L_t is the linear feature and N_t is the nonlinear feature of the combination method.

4.4. Classification

For each class (N, S, V, F, U), a total of 28 features of the combined feature vector were extracted for testing. The feature vector was assessed using two different classifiers.

4.4.1. Neural network classifier

The NN plays an important role in a wide variety of applications, such as pattern recognition and classification tasks. In the NN model, each neuron computes the weighted sum of its inputs and applies the sum to a nonlinear function called the activation function. The feed-forward NN is capable of recognizing and classifying ECG signals more accurately. In general, the performance of the feed-forward NN depends on the number of hidden layers, the number of hidden neurons, the learning algorithm and the activation function for each neuron [34].

In this study, the input layer consisted of 28 nodes, corresponding to the 28 features used. A hidden layer of 40 neurons and an output layer of five neurons corresponding to five classes were used. The 40 neurons in the hidden layers were chosen by trial and error. To improve the learning procedure, the back propagation method was used. Based on the computation of the mean square error (between the desired response and actual response of NN), the weights in the network were updated and the process was continued until the mean square error was below the specific threshold. The testing data was then forwarded to the trained NN to acquire the output, and the testing patterns were classified.

4.4.2. Support vector machine

The SVM classifier for a single layer can supervise classification problems due to its capability for generalization [35]. It converts the input vector patterns to higher dimension feature space through some nonlinear mapping and obtains an optimal separating hyper-plane which is built to separate two classes of samples. In particular, the SVM classifier shows a favorable generalization capability when using the maximal margin principle. It maximizes the distance between the patterns and the class separating hyper-plane. An objective function is formulated based on the distances of the class separating hyper-plane and the optimization process is carried out [35]. Different kernel transformations are used to map the data into high dimensional functions such as the quadratic, polynomial and radial basis function (RBF). The performance of the SVM can be affected by the hyper-parameter (C parameter and the kernel parameter), as these parameters determine the number of support vectors and the maximization margin of the SVM.

4.4.3. Evaluation method

To evaluate the performance of the classification, we used the standard statistical indices of sensitivity (Se), specificity, and accuracy (Acc) derived from four parameters: correctly detected beats (true positives = TP), undetected beats (false negatives = FN), correctly undetected beats (true negatives = TN) and falsely detected beats (false positives = FP).

These statistical indices are defined as follows:

$$\mathrm{Se} = \frac{\mathrm{TP}}{(\mathrm{TP}+\mathrm{FN})} \times 100$$

$$\mathrm{SP} = \frac{\mathrm{TN}}{(\mathrm{TN}+\mathrm{FP})} \times 100$$

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

5. Results

The experiments on the automated arrhythmia detection system were conducted using the MIT-BIH arrhythmia database. We used a combination of linear and nonlinear methods, and the data point consisted of 28 features. The transformed data points were formed into feature vectors inputted to the SVM-RBF kernel and NN. The nonlinear SVM was based on the popular Gaussian kernel (referred to as the SVM-RBF). Parameter C for this kernel was chosen to cover high and small regularization of the classification model, C was equal to 70

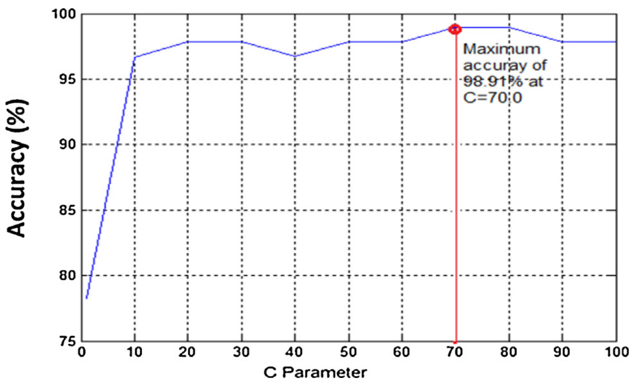

Fig. 4 – Variation of average accuracy with respect to different C parameter values during SVM-RBF classification using combination linear and nonlinear features.

and the gamma parameter was equal to 0.7 as shown in Fig. 4. In addition, for comparison purposes, in the first experiment, the feed-forward NN classifier was applied. The number of hidden layers and the learning rate value of the feed-forward NN that were chosen by trial and error were equal to 40 and 0.3, respectively.

The tenfold cross-validation technique was used for training and testing the classifiers. The overall performance of the classifiers was evaluated by taking the average of the tenfold cross-validations.

The corresponding overall accuracy and average accuracy for all classes of the proposed model were 98.91% and 99.57%, respectively. Comparing these results with those achieved with the NN classifier, there was a slight increase of 0.1% in both the overall accuracy and average accuracy for all classes, as presented in Table 3.

Fig. 5 illustrates the accuracy of classifying the ECG signals by using the SVM-RBF classifier for the linear features, nonlinear features and the proposed features. This shows that the proposed features (PCA-DWT) + (ICA-HOS) give much better accuracy when they are given as the input to the SVM-RBF classifier compared with linear features and nonlinear features, achieving the maximum average accuracy of 97.83% and 99.13% for all classes. The worst class accuracy of the proposed features was obtained for the N and V beats (98.91%), while 100% was obtained for the S, F and U beats.

The results from classifying the ECG signals by using the NN of the linear features, nonlinear features and proposed features are shown in Fig. 6. This shows that NN is much more accurate when the proposed features (PCA-DWT) + (ICA-HOS) are given as the input, achieving the maximum accuracy of 98.90% and 99.56% average accuracy for all classes. Its worst class accuracy was obtained for S and F beats (98.90%), while 100% was obtained for the remaining beats.

Table 4 – Classification results of the combined linear and nonlinear methods with the two different classifiers.

Classifier	Sensitivity (%)	Specificity (%)	Accuracy (%)
SVM-RBF	98.91%	97.85%	98.91%
NN	98.90%	98.90%	98.90%

Based on the performance evaluation, the sensitivity, specificity and accuracy of the combined linear and nonlinear methods with the SVM-RBF and NN classifiers are summarized in Table 4. It is noted that the NN and SVM-RBF provided equal sensitivity, specificity and accuracy (98.9%). To carry out the computations, MATLAB 2012A was used. All the methods including DWT denoising, PCA, ICA, HOS and classification algorithms were developed in MATLAB with the custom software. To implement ICA, the fast ICA toolbox was installed and linked to the developed programs.

6. Discussion

In this paper, we demonstrated that our proposed combination of linear and nonlinear methods as a feature extraction technique is able to extract the hidden information from the non-stationary ECG signal and discriminant features very well using the experimental MIT-BIH data. The technique provides a good extraction of the most discriminant features and clearly discriminates the N, S, V, F and U arrhythmia classes. The hidden information in the ECG signal was noticeable more obviously in the transform domain than the time domain. The features of the ECG beat were extracted using the HOS cumulants method and DWT. Moreover, the ECG signal and DWT representations contained more information, and the dimensionality reduction method was applied on DWT and the ECG signals. PCA was applied to reduce the dimensionality of DWT coefficients and ICA was used to reduce the representation of the ECG signals. The combination of 16 HOS cumulants and ICA and 12 PCA of DWT features gave maximum classification accuracy compared to the other methods.

It can be seen from our results that the proposed method is able to classify the S, F and U arrhythmia classes with 100% accuracy using the SVM-RBF classifier. In this section, we compare the classification performance of the proposed

Table 3 – Overall accuracy, average accuracy of all classes and class percentage accuracies achieved with the SVM-RBF and NN classifiers.

Classifiers	Features	Overall Acc. (%)	Avg. acc. of all classes	N	S	V	F	U
SVM-RBF	PCA-DWT	88.04%	94.97%	91.01%	94.19%	92.05%	97.59%	100%
NN		93.48%	97.34%	98.85%	94.51%	100%	98.85%	94.5%
SVM-RBF	ICA-HOS	97.83%	99.13%	98.90%	100%	97.83%	98.90%	100%
NN		94.57%	97.76%	97.75%	98.86%	97.75%	97.75%	96.6%
SVM-RBF	(PCA-DWT) + (ICA-HOS)	98.91%	99.57%	98.91%	100%	98.91%	100%	100%
NN		98.90%	99.56%	100%	98.90%	100%	98.90%	100%

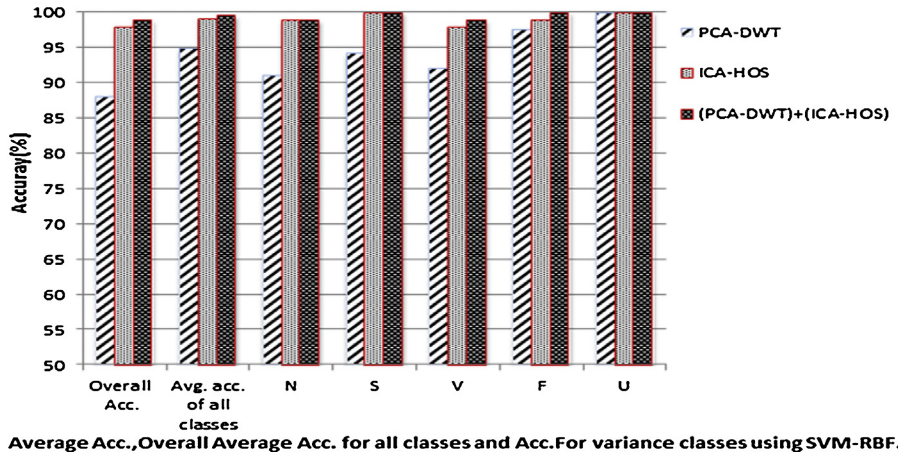

Fig. 5 – Comparison of linear, nonlinear and proposed features using the SVM-RBF classifier.

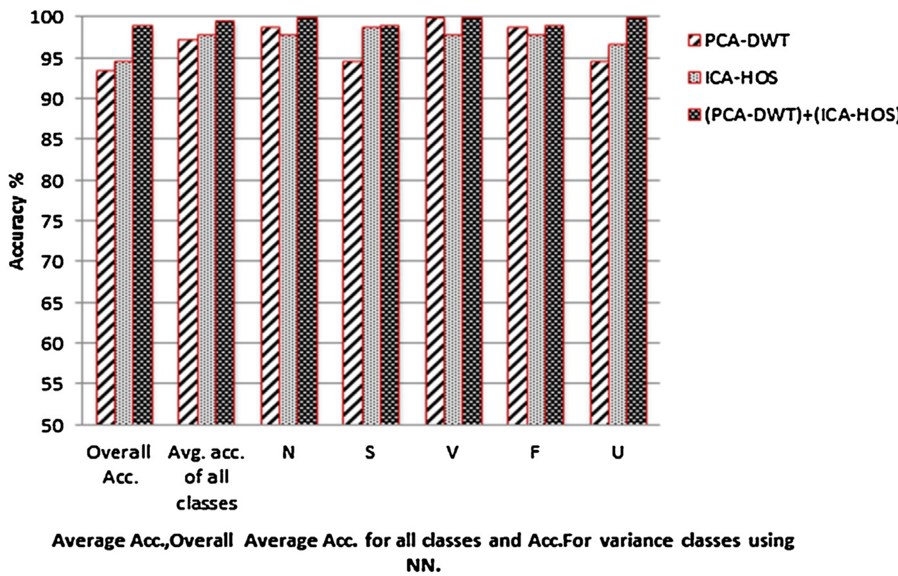

Fig. 6 – Comparison of linear, nonlinear and proposed features using the NN classifier.

PCA of DWT, HOS cumulants and ICA with SVM, NN for ECG arrhythmia classification with those of other systems in the literature utilizing different feature reduction methods and classifiers. We selected eight existing ECG beat classification systems using the MIT-BIH arrhythmia database for comparison with our proposed system. The number of classified arrhythmia types and the accuracies of our proposed system and eight existing systems are summarized in Table 5.

Table 5 – Comparison of the classification performance of the proposed system and some existing systems.

Literature	Features	Classifier	Classes	Accuracy (%)
Linear methods				
Hu et al. [36]	Time domain features	Mixture of experts	2	94.00%
Ince et al. [37]	DWT + PCA	MDPSO	5	95.58%
Martis et al. [17]	PCA	SVM-RBF	5	98.11%
Martis et al. [14]	DWT + PCA	SVM-RBF	5	96.92%
	DWT + PCA	NN	5	98.78%
Nonlinear methods				
Osowski and Linh [16]	HOS	Hybrid fuzzy NN	7	96.06%
Martis et al. [18]	Bispectrum + PCA	SVM-RBF	5	93.48%
Martis et al. [31]	Cumulant + PCA	NN	5	94.52%
Proposed Methodology	PCA + DWT + HOS + ICA	SVM-RBF	5	98.91%
		NN	5	98.90%

The results showed that the proposed system comprising the PCA of DWT, HOS cumulants and ICA with different classifiers (SVM-RBF and NN) can provide better classification accuracy compared to the other systems, which validates that the proposed system can serve as an effective tool for cardiologists to diagnose heart diseases based on ECG signals.

In one study, a mixture of expert approaches with morphology and the inter-beat (RR) interval features was used for the classification of normal, supra-ventricular ectopic, ventricular ectopic and fusion beats [36], and 94.0% accuracy was obtained. Five types of ECG beats were classified using PCA of the time domain of ECG signals, and 98.11% accuracy was obtained [17]. Ince et al. [37] proposed a arrhythmia recognition system to classify two types of classes, namely, (V) and (S), on the MIT-BIH database using DWT coefficients, PCA as feature extraction and feature reduction, and NN for pattern recognition, which were optimized by the multidimensional particle swarm optimization (MDPSO) technique. The average classification accuracy obtained was 95.58%.

The five classes recommended by the AAMI on the MIT-BIH arrhythmia database were classified by researchers using DWT and dimensionality reduction methods. They obtained 96.92% and 98.78 accuracy using SVM-RBF and NN, respectively [14]. The same five types of beats were classified with an accuracy of 93.48% using the principal components of bispectrum with the SVM and RBF kernel [14].

Using the fuzzy hybrid NN classifier and HOS analysis features, the seven ECG beats were classified [16] and a classification accuracy of 96.06% was obtained. Martis et al. [31] applied HOS cumulants with PCA; the average classification accuracy obtained was 94.52% using NN.

7. Conclusion

The ECG signal shows the electrical activity of the heart, providing important information about the cardiac state. In this paper, an effective ECG arrhythmia classification system that consists of the combination of PCA of DWT, ICA and HOS feature extraction methods and two different classifiers (SVM-RBF and NN) is proposed. The experiments on the proposed system showed that the combination of PCA-DWT, ICA and HOS feature extraction methods with SVM-RBF and NN provided equal average accuracy, sensitivity and specificity of 98.9%. The expanded methodology in this study can be used in the research and development of arrhythmia detection systems, telemedicine applications and cardiac pacemakers.

Acknowledgement

The work is supported by Universiti Teknologi Malaysia (UTM), Malaysia, under FRGS/2/2013/SKK10/UTM/03/1.

REFERENCES

[1] C. Mathers, A. Lopez, C. Stein, D. Fat, C. Rao, Deaths and Disease Burden by Cause: Global Burden of Disease Estimates for 2001 by World Bank Country Groups, 2005 (revised 2005).

[2] T. Thom, N. Haase, W. Rosamond, V.J. Howard, J. Rumsfeld, T. Manolio, Z.-J. Zheng, K. Flegal, C. O'donnell, S. Kittner, Heart Disease and Stroke Statistics – 2006 update a report from the American Heart Association Statistics Committee and Stroke Statistics Subcommittee, Circulation 113 (6) (2006) e85–e151.

[3] S. Sidney, W.D. Rosamond, V.J. Howard, R.V. Luepker, The “Heart Disease and Stroke Statistics – 2013 Update” and the need for a national cardiovascular surveillance system, Circulation 127 (1) (2013) 21–23.

[4] A.L. Goldberger, Clinical Electrocardiography: A Simplified Approach, Elsevier Health Sciences, 2012.

[5] M.M. Hadhoud, M.I. Eladawy, A. Farag, Computer aided diagnosis of cardiac arrhythmias, in: The 2006 International Conference on Computer Engineering and Systems, IEEE, 2006, pp. 262–265.

[6] A. Shiyovich, A. Wolak, L. Yacobovich, A. Grosbard, A. Katz, Accuracy of diagnosing atrial flutter and atrial fibrillation from a surface electrocardiogram by hospital physicians: analysis of data from internal medicine departments, Am. J. Med. Sci. 340 (4) (2010) 271–275.

[7] J.-S. Wang, W.-C. Chiang, Y.-L. Hsu, Y.-T.C. Yang, ECG arrhythmia classification using a probabilistic neural network with a feature reduction method, Neurocomputing 116 (2013) 38–45.

[8] I. Jekova, G. Bortolan, I. Christov, Assessment and comparison of different methods for heartbeat classification, Med. Eng. Phys. 30 (2) (2008) 248–257.

[9] C.-H. Lin, Frequency-domain features for ECG beat discrimination using grey relational analysis-based classifier, Comput. Math. Appl. 55 (4) (2008) 680–690.

[10] M. Arif, Robust electrocardiogram (ECG) beat classification using discrete wavelet transform, Physiol. Meas. 29 (5) (2008) 555.

[11] B. Gramatikov, I. Georgiev, Wavelets as alternative to short-time Fourier transform in signal-averaged electrocardiography, Med. Biol. Eng. Comput. 33 (3) (1995) 482–487.

[12] C. Li, C. Zheng, C. Tai, Detection of ECG characteristic points using wavelet transforms, Biomed. Eng. IEEE Trans. 42 (1) (1995) 21–28.

[13] Q.B. Zhao, L.Q. Zhang, ECG feature extraction and classification using wavelet transform and support vector machines, in: Proceedings of the 2005 International Conference on Neural Networks and Brain, vol. 1–3, 2005, pp. 1089–1092.

[14] R.J. Martis, U.R. Acharya, L.C. Min, ECG beat classification using PCA, LDA, ICA and discrete wavelet transform, Biomed. Signal Process. Control 8 (5) (2013) 437–448.

[15] R.J. Martis, U.R. Acharya, H. Adeli, Current methods in electrocardiogram characterization, Comput. Biol. Med. 48 (2014) 133–149.

[16] S. Osowski, T.H. Linh, ECG beat recognition using fuzzy hybrid neural network, Biomed. Eng. IEEE Trans. 48 (11) (2001) 1265–1271.

[17] R.J. Martis, U.R. Acharya, K. Mandana, A.K. Ray, C. Chakraborty, Application of principal component analysis to ECG signals for automated diagnosis of cardiac health, Expert Syst. Appl. 39 (14) (2012) 11792–11800.

[18] R.J. Martis, U.R. Acharya, K. Mandana, A. Ray, C. Chakraborty, Cardiac decision making using higher order spectra, Biomed. Signal Process. Control 8 (2) (2013) 193–203.

[19] Y. Kutlu, D. Kuntalp, A multi-stage automatic arrhythmia recognition and classification system, Comput. Biol. Med. 41 (1) (2011) 37–45.

[20] M.K. Das, S. Ari, ECG beats classification using mixture of features, Int. Sch. Res. Not. (2014) 2014.

[21] E.D. Übeyli, Statistics over features of ECG signals, Expert Syst. Appl. 36 (5) (2009) 8758–8767.
[22] P. De Chazal, R.B. Reilly, A patient-adapting heartbeat classifier using ECG morphology and heartbeat interval features, Biomed. Eng. IEEE Trans. 53 (12) (2006) 2535–2543.
[23] G.B. Moody, R.G. Mark, The impact of the Mit-Bih arrhythmia database, Eng. Med. Biol. Mag. IEEE 20 (3) (2001) 45–50.
[24] E.D. Übeyli, Combining recurrent neural networks with eigenvector methods for classification of ECG beats, Digit. Signal Process. 19 (2) (2009) 320–329.
[25] B.N. Singh, A.K. Tiwari, Optimal selection of wavelet basis function applied to ECG signal denoising, Digit. Signal Process. 16 (3) (2006) 275–287.
[26] J. Pan, W.J. Tompkins, A real-time QRS detection algorithm, Biomed. Eng. IEEE Trans. (3) (1985) 230–236.
[27] H.-S. Park, S.-M. Woo, Y.-S. Kim, B.-J. Kang, S.-W. Ban, ECG pattern classification based on generic feature extraction, in: Proceedings of the 3rd WSEAS International Conference on Circuits, Systems, Signal and Telecommunications (CISST'2009), 2009, pp. 21–25.
[28] P.S. Addison, Wavelet transforms and the ECG: a review, Physiol. Meas. 26 (5) (2005) R155.
[29] J. Kim, H.S. Shin, K. Shin, M. Lee, Robust algorithm for arrhythmia classification in ECG using extreme learning machine, Biomed. Eng. Online 8 (2009) 31.
[30] K. Polat, S. Güneş, Detection of ECG arrhythmia using a differential expert system approach based on principal component analysis and least square support vector machine, Appl. Math. Comput. 186 (1) (2007) 898–906.
[31] R.J. Martis, U.R. Acharya, C.M. Lim, K. Mandana, A.K. Ray, C. Chakraborty, Application of higher order cumulant features for cardiac health diagnosis using ECG signals, Int. J. Neural Syst. 23 (04) (2013).
[32] C. Jutten, J. Herault, Blind separation of sources. Part I: an adaptive algorithm based on neuromimetic architecture, Signal Process. 24 (1) (1991) 1–10.
[33] D. Mantini, F. Petrucci, P. Del Boccio, D. Pieragostino, M. Di Nicola, A. Lugaresi, G. Federici, P. Sacchetta, C. Di Ilio, A. Urbani, Independent component analysis for the extraction of reliable protein signal profiles from Maldi-Tof mass spectra, Bioinformatics 24 (1) (2008) 63–70.
[34] C.M. Bishop, Neural Networks for Pattern Recognition, 1995.
[35] V. Vapnik, The Nature of Statistical Learning Theory, Springer, 2000.
[36] Y.H. Hu, S. Palreddy, W.J. Tompkins, A patient-adaptable ECG beat classifier using a mixture of experts approach, Biomed. Eng. IEEE Trans. 44 (9) (1997) 891–900.
[37] T. Ince, S. Kiranyaz, M. Gabbouj, A generic and robust system for automated patient-specific classification of ECG signals, Biomed. Eng. IEEE Trans. 56 (5) (2009) 1415–1426.
[38] O.T. Inan, L. Giovangrandi, G.T. Kovacs, Robust neural-network-based classification of premature ventricular contractions using wavelet transform and timing interval features, Biomed. Eng. IEEE Trans. 53 (12) (2006) 2507–2515.
[39] W. Jiang, S.G. Kong, Block-based neural networks for personalized ECG signal classification, Neural Netw. IEEE Trans. 18 (6) (2007) 1750–1761.
[40] C.-H. Lin, Y.-C. Du, T. Chen, Adaptive wavelet network for multiple cardiac arrhythmias recognition, Expert Syst. Appl. 34 (4) (2008) 2601–2611.
[41] E.D. Übeyli, Adaptive neuro-fuzzy inference system for classification of ECG signals using Lyapunov exponents, Comput. Methods Programs Biomed. 93 (3) (2009) 313–321.
[42] E. Derya Übeyli, Recurrent neural networks employing Lyapunov exponents for analysis of ECG signals, Expert Syst. Appl. 37 (2) (2010) 1192–1199.
[43] O. Sayadi, M.B. Shamsollahi, G.D. Clifford, Robust detection of premature ventricular contractions using a wave-based Bayesian framework, Biomed. Eng. IEEE Trans. 57 (2) (2010) 353–362.
[44] N.O. Ozcan, F. Gurgen, Fuzzy support vector machines for ECG arrhythmia detection, in: 2010 20th International Conference on Pattern Recognition (ICPR), IEEE, 2010, pp. 2973–2976.
[45] C. Banupriya, S. Karpagavalli, Electrocardiogram beat classification using support vector machine and extreme learning machine, in: ICT and Critical Infrastructure: Proceedings of the 48th Annual Convention of Computer Society of India – vol. I, Springer, 2014, pp. 187–193.