Technical note

ECG beat classification using PCA, LDA, ICA and Discrete Wavelet Transform

Roshan Joy Martis [a,*], U. Rajendra Acharya [a,b], Lim Choo Min [a]

[a] Ngee Ann Polytechnic, Singapore
[b] Department of Biomedical Engineering, Faculty of Engineering, University of Malaya, Malaysia

ARTICLE INFO

ABSTRACT

Electrocardiogram (ECG) is the P-QRS-T wave, representing the cardiac function. The information concealed in the ECG signal is useful in detecting the disease afflicting the heart. It is very difficult to identify the subtle changes in the ECG in time and frequency domains. The Discrete Wavelet Transform (DWT) can provide good time and frequency resolutions and is able to decipher the hidden complexities in the ECG. In this study, five types of beat classes of arrhythmia as recommended by Association for Advancement of Medical Instrumentation (AAMI) were analyzed namely: non-ectopic beats, supra-ventricular ectopic beats, ventricular ectopic beats, fusion betas and unclassifiable and paced beats. Three dimensionality reduction algorithms; Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA) and Independent Component Analysis (ICA) were independently applied on DWT sub bands for dimensionality reduction. These dimensionality reduced features were fed to the Support Vector Machine (SVM), neural network (NN) and probabilistic neural network (PNN) classifiers for automated diagnosis. ICA features in combination with PNN with spread value (σ) of 0.03 performed better than the PCA and LDA. It has yielded an average sensitivity, specificity, positive predictive value (PPV) and accuracy of 99.97%, 99.83%, 99.21% and 99.28% respectively using ten-fold cross validation scheme.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

The incidence and prevalence of cardiovascular diseases (CVD) have increased in recent years [1]. As per the recent WHO report, the overall death rates due to CVD have declined, however the burden of the disease still remains high [2]. In 2007, the overall death rate due to CVD was 251.7 per 100,000 population [2]. The death rate was different for male and female genders and also for black and white races. The rate was 294.0 per 100,000 among white males, 405.9 per 100,000 among black males, 205.7 per 100,000 white females and 286.1 per 100,000 black females [2]. This increase in death rates due to CVD in the modern world is due to the epidemiological transition due to obesity, diabetes mellitus, smoking habit and other lifestyle changes. One of the complication of CVD among many others is atrial and ventricular arrhythmias which occur due to cardiac rhythm disturbances. Arrhythmia is a collective term for a heterogeneous group of conditions in which there would be abnormal electrical activity. There are many causes for arrhythmias, majority of them are related to CVD. Arrhythmias like ventricular fibrillation and flutter are life threatening medical emergencies which result in cardiac arrest, hemodynamic collapse and sudden cardiac death [3]. The electrical activity of the heart could be non-invasively monitored using ECG. It will be very difficult to decipher the hidden information present in the ECG data due to its small amplitude and duration. Therefore a computer assisted tool could help the physicians in their diagnosis. This computer assisted diagnosis can be used as an adjunct tool for the physicians in their practice and interpretation and can play a major role in the management of cardiovascular diseases [4].

Using local fractal dimension, six types of ECG beats (normal, left bundle branch block (LBBB), right bundle branch block (RBBB), atrial premature contraction (APC), ventricular premature contraction (VPC) and paced beats) were classified with more than 97% of sensitivity for normal, LBBB, RBBB, paced and VPC beats; and more than 86% sensitivity for APC beats [5]. Five types of ECG beats (normal, LBBB, RBBB, APC and VPC) were classified with 93.97% of accuracy [6]. Wavelet transform and particle swarm optimization technique were used to classify six types of ECG beats (normal, LBBB, RBBB, APC, VPC and paced beats) and obtained 88.84% of accuracy [7].

A minicomputer system was designed for analyzing 24 h ambulatory ECG and automatically classified normal, supra-ventricular ectopic beats and ventricular ectopic beats using time intervals between different characteristic peaks of ECG [8]. Using linear prediction method, the VPC was detected with a sensitivity of 92% [9]. Using Hidden Markov Model (HMM), various segments of ECG were modeled and subsequently classified the ECG beats into normal, supra-ventricular ectopic beats (SVEB) and ventricular

---

* Corresponding author. Tel.: +65 6460 8393.
E-mail address: roshaniitsmst@gmail.com (R.J. Martis).

Table 1
MIT-BIH arrhythmia database beats classified as per ANSI/AAMI EC57:1998 standard [17].

Non ectopic beat (N)	Supra-ventricular ectopic beats (S)	Ventricular ectopic beats (V)	Fusion beat (F)	Unknown beat (Q)
Normal beat	Atrial Premature (AP) beat	Premature ventricular contraction (PVC) beat	Fusion of ventricular and normal (fVN) beat	Paced beat
Left bundle branch block (LBBB) beat	aberrated Atrial Premature (aAP)	Ventricular escape (VE) beat		Fusion of paced and normal (fPN) beat
Right bundle branch block (RBBB) beat	Nodal (junctional) Premature (NP) beat			Unclassifiable (U) beat
Atrial escape (AE) beat	Supra-ventricular premature beat (SP) beat			
Nodal (junctional) escape (NE) beat				

ectopic beats (VEB) using the time domain features [10]. The ECG morphology and RR interval features were used for the classification of normal and five types of arrhythmia beats using particle swarm optimization and obtained average accuracy of 93.27% [11]. Using auto-regressive model and generalized linear model classifier, normal, APC, supra ventricular tachycardia (SVT), ventricular tachycardia (VT), VPC and ventricular flutter (VF) ECG beats were classified with 93.2% of accuracy [12]. Using Hermite coefficients of ECG beats and neuro-fuzzy technique, the ECG beats were classified with 96% of accuracy [13]. A hardware system was developed with four signals (ballistocardiogram, ECG, lower body impedance plethysmogram and lower body electromyogram) for monitoring the cardiovascular health at home [14]. Two kinds of abnormalities in the ECG were classified using Gaussian Mixture Model (GMM) and reported more than 94% accuracy [15]. Using wavelet transform and PCA, normal and abnormal beats in ECG were classified using different classifiers and reported 95.6% of accuracy with SVM classifier [16]. However all these methods have following drawbacks.

(i) All these methods use time domain features where subtle changes in the ECG signal and the hidden complexities cannot be clearly deciphered.
(ii) Most of these methods were tested only on limited data sets and the generalization performance of these methods on large databases was not tested.
(iii) All these methods were tested only on few classes of ECG beats and there is a need to test the methods and algorithms on a standard classification scheme of arrhythmia beats such as ANSI/AAMI EC57:1998 [17].

The subtle changes in the amplitude and duration of ECG represent the function of the heart [21,23]. In our previous work, we have computed the principal components of time domain ECG signal and subjected them for statistical Student's t-test. Also in parallel the principal components of DWT coefficients of ECG beats were computed and subjected to Student's t-test. It was seen from our experiments that the later method provided higher value for the t-statistic, which implies that the features were more discriminating in DWT domain than time domain for normal and arrhythmia classes. So the subtle changes in the amplitude and duration in the ECG as it was in the time domain did not provide good discrimination, hence the DWT transform domain features were used. There are a large number of coefficients, and hence a dimensionality reduction algorithm needs to be applied. In our proposed method the ECG beats were transformed using DWT and subsequently three dimensionality reduction methods were applied independently to extract the features. Three dimensionality reduction methods were used: Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA) and Independent Components Analysis (ICA). These features with reduced dimensionality were fed to neural network (NN), SVM with different kernel functions and probabilistic neural network (PNN) for automated classification.

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

2. Materials used

In this study, the MIT-BIH arrhythmia database [18] is used, where the signals were sampled at 360 Hz. The database consists of 48 signals, each of half an hour duration of Holter recording. In this analysis, we have used the *entire data of MIT-BIH arrhythmia database* as recommended by ANSI/AAMI EC57:1998 standard [17]. The data used consisted of 90, 580 non-ectopic beats, 2973 supra-ventricular ectopic beats, 7707 ventricular ectopic beats, 1784 fusion beats and 7050 unknown and paced beats. In each of the 48 signal files, if insufficient samples were present to the left of the first detected QRS complex, the corresponding beat is neglected. Also if there were insufficient samples after the last detected QRS complex in any of the signal files, then also the corresponding beat is neglected. Table 1 shows the details of the different beats of MIT-BIH database grouped into five main classes.

3. Methodology

Fig. 1 shows the proposed automated system for classification of 5 types of beats in ECG of arrhythmia as recommended by ANSI/AAMI EC57:1998 standard [17]. The working of each block is explained in detail in the following sections.

3.1. Discrete Wavelet Transform (DWT) and denoising of ECG

Unlike Fourier transform, the wavelet transform offers resolution in both time and frequency domains. The DWT is obtained from the continuous wavelet transform by sampling it on a dyadic grid. The DWT decomposes a signal successively into low frequency and high frequency components. The low frequency component is called approximation, and the high frequency component is called the detail. Fig. 2 depicts the DWT decomposition using filter banks. The ECG signal $x(n)$ is passed through a low pass filter $h(n)$, and then down sampled by a factor of two to obtain the approximation coefficients at level one. The high pass filter is derived from the low pass filter as,

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

where L is the length of the filter in number of points. The detail coefficients were obtained by passing the signal through $g(n)$ and then down sampling by a factor of two. The two filters $h(n)$ and $g(n)$ were called quadrature mirror filters. The DWT filtering along with sub sampling were given by,

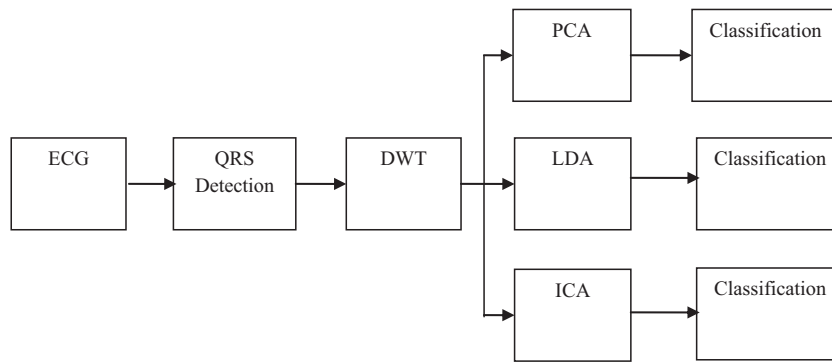$$y_{\text{low}}(k) = \sum_n x(n)h(-n + 2k) \tag{2}$$

Fig. 1. The proposed system.

and

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

As seen from Fig. 2, the original signal is passed through low pass, $h(n)$ and high pass, $g(n)$ half band filters and then both signals were down sampled by a factor of two. The low pass signal is again successively filtered by $h(n)$ and $g(n)$ and sub sampled by a factor of two to obtain the next level approximation and detail coefficients. A detailed review is provided by Addison [45].

The ECG signal downloaded from MIT-BIH database was subjected to wavelet based denoising using Daubechies D_6 ('db6') wavelet basis function [19]. The ECG signals sampled at 360 Hz were decomposed up to 9 levels using db6 wavelet. The 9th level approximation sub band contains the frequency range of 0–0.351 Hz which is mainly the baseline wander, was not used for reconstructing the denoised signal. Also the ECG would not contain much information after 45 Hz. Therefore the first and second level detail coefficients consisting frequency band of 90–180 Hz and 45–90 Hz respectively were not used for reconstructing the denoised ECG. The required sub bands, the 3rd, 4th, 5th, 6th, 7th, 8th and 9th level detail signals were only used (all other sub-band coefficients were replaced with zeros) for computing the inverse wavelet transform and the denoised ECG signal is obtained [19]. The required sub band coefficients were kept and in the sub bands which are not required, the coefficients were all replaced with zeros and the inverse DWT is computed to obtain the denoised ECG.

The QRS complex in the ECG is detected using Pan Tompkins algorithm [20] on the denoised ECG signals. The Pan Tompkins algorithm consists of taking derivatives, absolute/rectification operation, squaring, moving average integration and threshold operations [20]. After detection of QRS complex, 99 samples before the QRS peak and 100 samples after the peak and the QRS peak itself are considered as 200 samples segment as a single beat for the subsequent analysis.

Each beat consisting of 200 samples was decomposed into four levels using FIR approximation of Mayer's wavelet ('dmey'). The
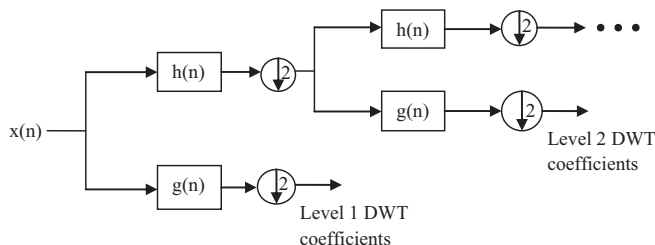

Fig. 2. DWT decomposition.

fourth level approximation sub band consist of frequency range of 0–11.25 Hz, whereas the fourth level detail consist of frequency range of 11.25–22.5 Hz. In a previous work [22], it was shown that the power spectral density of different beats from each of the classes has discriminatory information in these two sub bands (4th level approximation and detail). These two sub bands were subjected to dimensionality reduction using PCA, LDA and ICA methods independently.

3.2. Principal Component Analysis (PCA)

Principal Component Analysis (PCA) is a linear dimensionality reduction technique that seeks projection of the data into the directions of highest variability [24]. It computes the principal components which are the basis vectors of directions in decreasing order of variability. The first principal component provides basis vector for the direction of highest variability. The second principal component provides the basis vector for the next direction orthogonal to the first principal component and so on. A percentage of total variability of the data is set as the threshold in order to select the number of principal components. Computation of principal components involves computation of covariance matrix of the data, its eigenvalue decomposition, sorting of eigenvectors in the decreasing order of eigenvalues and finally projection of the data into the new basis defined by principal components by taking the inner product of the original signals and the sorted eigenvectors. The detailed procedure of PCA is provided below:

Step 1: Compute the covariance matrix from the data as,

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

where X is the data matrix of DWT coefficients in a sub band of $N \times 200$ dimension, N is the total number of patterns, $\bar{x}$ represents mean vector of X.

Step 2: Compute the matrix of eigenvectors V and diagonal matrix of eigenvalues D as

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

Step 3: The eigenvectors in V are sorted in the descending order of eigenvalues in D and the data is projected on these eigenvector directions by taking the inner product between the data matrix and sorted eigenvector matrix as,

$$\text{Projected data} = [V^T(X - \bar{x})^T]^T \tag{6}$$

where V is of 200×200 dimension, each row of it is a eigenvector. The first six columns of the projected data were considered as the six features for subsequent classification. The PCA was applied on both sub band coefficients of 4th level approximation and detail independently. From each of the sub band, six principal

components were selected based on containment of 99% of the total variability. In total 12 features (six each from the two sub bands) were used for subsequent pattern identification using classifiers.

3.3. Linear Discriminant Analysis (LDA)

The PCA projects the data into the directions of highest variability which helps to represent the data in less number of components. However these components need not be the best directions for providing highest possible discrimination so as to classify into different classes. Fisher's Linear Discriminant Analysis [24] provides highest possible discrimination between different classes of data which helps us to classify the data accurately. In order to reduce the dimensionality of data using LDA, the within class covariance matrix was defined as,

$$S_W = \sum_{k=1}^{K} S_k \tag{7}$$

where

$$S_k = \sum_{n \in C_k} (x_n - m_k)(x_n - m_k)^T \tag{8}$$

and

$$m_k = \frac{1}{N_k} \sum_{n \in C_k} x_n \tag{9}$$

and N_k is the number of patterns in class C_k, x_n is the DWT coefficient vector of nth pattern and K is the total number of classes present in the data.

Also the between class covariance matrix is defined as,

$$S_B = \sum_{k=1}^{K} N_k (m_k - m)(m_k - m)^T \tag{10}$$

where

$$m = \frac{1}{N} \sum_{n=1}^{N} x_n = \frac{1}{N} \sum_{k=1}^{K} N_k m_k \tag{11}$$

is the global mean of the data. The total covariance matrix is defined as,

$$S_T = S_W + S_B \tag{12}$$

Finally the projection matrix is computed as,

$$W = arg_W max\{(WS_W W^T)^{-1}(WS_B W^T)\} \tag{13}$$

From the projection matrix the LDA coefficients were obtained as,

$$y = W^T x \tag{14}$$

where x is the DWT coefficient vector in a sub band for a given pattern and y is the vector of LDA coefficients. The first six columns of y were considered as six features for subsequent pattern classification. Similar to PCA, the LDA algorithm is applied independently on both sub bands, 4th level approximation and detail. From each of these two sub bands six LDA coefficients were selected. In total these twelve features were used for the subsequent pattern classification.

3.4. Independent Component Analysis (ICA)

Independent Component Analysis (ICA) is a non-linear dimensionality reduction method. If x is a vector with mixtures $\{x_1, x_2, \ldots, x_n\}$ and let s be the source vector with $\{s_1, s_2, \ldots, s_n\}$. Let A denote the weight matrix with elements a_{ij}. The ICA model assumes that the signal x (the DWT coefficients in a sub band) which we were observing is linearly mixed with the source signals. The ICA model is given by,

$$x = As = \sum_{i=1}^{n} a_i s_i \tag{15}$$

We need to solve for the elements of A to solve the ICA problem. Also the source signal is expressed in terms of mixed signals as,

$$s = Wx \tag{16}$$

where A and W are the inverses of each other. We have used FastICA algorithm [25] for computation of independent components. Following is the step by step method of ICA.

Step P: Pre-processing

(a) *Centering*: Here the mean of the data was subtracted so that the average value of the signal would be zero as,

$$x = x - E\{x\} = x - \frac{1}{N} \sum_{i=1}^{N} x_i \tag{17}$$

where $E\{\cdot\}$ is the statistical expectation operator and N is the total number of patterns present in the data.

(b) *Whitening*: If the data is not Gaussian distributed, it is made Gaussian by the whitening transformation,

$$\tilde{x} = VD^{-1/2}V^T x \tag{18}$$

where

$$E\{xx^T\} = VDV^T \tag{19}$$

The left hand side of Eq. (19), is the covariance matrix of the data, x, V is the matrix of Eigenvectors and $D = diag\{d_1, d_2, \ldots, d_n\}$ is the diagonal matrix of eigenvalues. The right hand side of Eq. (19) is the eigenvalue decomposition of covariance matrix.

Step 1: Choose an initial weight vector w.
Step 2: Let

$$W^+ = E\{xg(W^T x)\} - E\{g'(W^T x)\} \cdot W \tag{20}$$

where

$$g(u) = \frac{1}{a_1} \log \cosh(a_1 u) \tag{21}$$

and $g'(u)$ is the derivative of g.
Step 3: Normalize W^+ as,

$$W = \frac{W^+}{\|W^+\|} \tag{22}$$

Step 4: If W is not converged (W is said to be converged if its value does not change over next iteration), go to step 2.

After W is found from the above method, its inverse is computed to get matrix A. The weights in matrix A were used as features for subsequent pattern recognition. ICA method was applied independently on two DWT sub bands, 4th level approximation and detail. From each of the sub bands six ICA components were used. So in total twelve features from the two sub bands were used for subsequent pattern recognition.

3.5. Neural network (NN) classifier

In this study a fully connected feed-forward neural network [26] [27] was used for pattern classification. The input layer consisted of 12 nodes corresponding to the 12 features used; a hidden layer

of 10 neurons and an output layer of 5 neurons corresponding to the five classes was used. The choice of 10 neurons in the hidden layer is by trial and error method. We have also tested the three layer neural network with different number of hidden neurons. We obtained highest accuracy with 10 neurons in the hidden layer. The NN weights were updated using error back propagation method of learning. Based on the class labels of each pattern in the training set, the MSE (mean square error) between the desired response and the actual response of NN was computed. Based on this error, the weights in the network were updated, and the process is continued until the MSE would be below a threshold (0.0001). After that the testing data is fed to the trained NN and the output is noted and the testing patterns were classified.

3.6. Support Vector Machine (SVM) classifier

SVM is a highly nonlinear and single layer network which has higher generalization ability in the sense that it can classify unseen data accurately [28,29]. It maximizes the distance between the patterns and the class separating hyper-plane. It minimizes the structural risk than the empirical risk. An objective function is formulated based on the distances to the class separating hyper-plane and the optimization process is carried out. Generally the patterns are not linearly separable, therefore first they were mapped into a high dimensional space using a proper kernel, and then the optimization was carried out. For mapping the data into high dimensional space, various kernel transformations namely: quadratic, polynomial and radial basis function (RBF) were used. In the current study Least Square SVM (LS-SVM) was used which was proposed by Suykens and Vandewalle [30].

3.7. Probabilistic neural network

The PNN [24] classifier consists of as many input nodes as the number of features used for classification. Each input node is connected to each node in the pattern layer. Each pattern node is connected to only one node in the category layer. There would be as many nodes in the category layer as the number of classes in the data. The modifiable weights are trained on the training data. Once the PNN network is trained on the training data, the test samples are fed to the PNN and it was classified. Based on the correct or wrong classification the performance measures were computed.

4. Results

The entire experiments were simulated using MATLAB. The ECG signal from MIT BIH arrhythmia database was denoised using wavelet based denoising technique. The denoised signal was subjected to QRS complex detection using Pan-Tompkins method. After QRS complex detection, 100 samples from the right of QRS peak, 99 samples to the left of QRS complex and the QRS peak itself were chosen as a segment of ECG beat. The choice of 200 samples around the R peak as a signal window length is such that it consists approximately one cycle of cardiac activity. Based on the variability, the dimension reduction method (PCA, LDA or ICA) will discard unnecessary dimensions. This duration was used in our previous studies as well [15,16,22,23,31,38,42,44]. In total 110,094 beats were considered for analysis from the MIT BIH database. The DWT was computed on each of the ECG beat using finite impulse response (FIR) approximation of Mayer's wavelet ('dmey') [31]. In a previous work [31], we tried 54 wavelet basis functions independently for decomposing the ECG for the features extraction. It was found experimentally that the FIR approximation of Mayers wavelet ('dmey') provided highest discrimination among various classes and hence highest classification accuracy. Therefore in the current work also we have chosen dmey wavelet for feature extraction. Based on the signal characteristics of ECG, 4th level approximation and 4th level detail were considered for subsequent dimensionality reduction and pattern classification. On each of the sub-band, three dimensionality reduction methods: PCA, LDA and ICA were applied. After dimensionality reduction, six components from each of the sub-band were considered for subsequent pattern identification. The *mean* and standard deviation of different principal components due to PCA are summarized in Tables 2 and 3. Table 2 provides summary statistics of principal components on 4th level approximation sub-band, whereas Table 3 provides summary statistics of principal components of 4th level detail sub-band. It can be seen from the tables that, all *twelve* features (six from each sub-band) were statistically significant with low *p* value (<0.0001).

Twelve features after PCA were fed to different classifiers (NN, SVM with different kernels and PNN) for automated classification. In this study we used 10-fold cross validation technique for training and testing of the classifiers. In this technique, the entire dataset (110,094 beats) was sub-sampled into ten sets each having almost same distribution of samples from each of the classes. Nine sets (99,085 beats) were used for training the classifiers and

Table 2
Summary of six principal components extracted from 4th level approximation sub-band (*p* value <0.0001).

Feature	N	S	V	F	Q
PC1	0.12 ± 7.24	1.96 ± 10.52	-2.55 ± 5.84	0.01 ± 5.13	0.21 ± 8.37
PC2	-0.16 ± 3.34	0.73 ± 6.08	1.62 ± 7.63	0.16 ± 2.97	0.10 ± 4.29
PC3	-0.04 ± 2.41	1.03 ± 4.87	0.11 ± 3.47	0.03 ± 1.80	-0.03 ± 2.86
PC4	$-1 \times 10^{-3} \pm 1.64$	0.03 ± 2.55	-0.11 ± 2.14	-0.32 ± 1.12	0.22 ± 1.78
PC5	$6.3 \times 10^{-3} \pm 1.30$	$3.1 \times 10^{-3} \pm 2.48$	-0.10 ± 1.78	0.03 ± 0.95	0.01 ± 1.44
PC6	$-1.49 \times 10^{-2} \pm 1.04$	0.08 ± 1.71	0.15 ± 1.18	-0.25 ± 0.72	0.05 ± 1.17

Table 3
Summary of principal components extracted from 4th level detail sub-band (*p* value <0.0001).

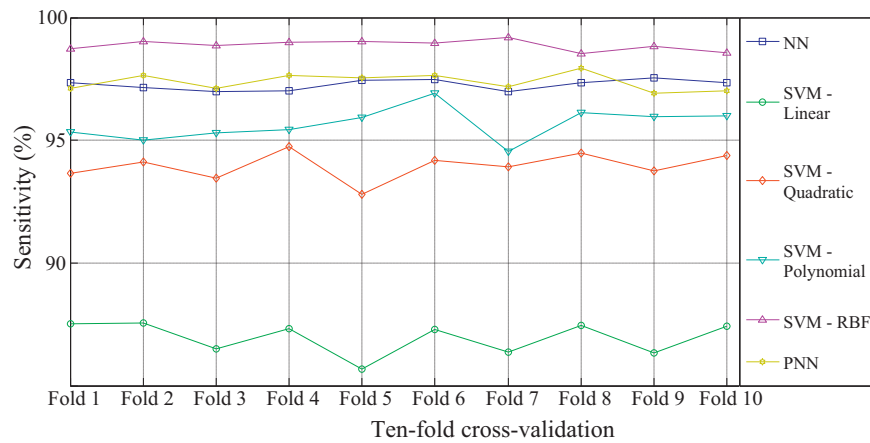| Feature	N	S	V	F	Q
PC1	-0.01 ± 0.88	0.27 ± 1.44	0.03 ± 1.06	-0.06 ± 0.62	-0.04 ± 1.02
PC2	-0.02 ± 0.70	0.23 ± 1.23	0.17 ± 0.96	0.11 ± 0.56	-0.01 ± 0.78
PC3	-0.01 ± 0.55	0.08 ± 1.03	0.12 ± 0.83	-0.03 ± 0.40	0.01 ± 0.61
PC4	0.01 ± 0.44	-0.11 ± 0.78	-0.08 ± 0.65	0.04 ± 0.33	-0.02 ± 0.53
PC5	-0.04 ± 1.09	0.31 ± 1.29	0.32 ± 0.94	0.06 ± 1.10	0.06 ± 1.12
PC6	0.02 ± 0.47	-0.14 ± 0.79	-0.21 ± 0.97	0.04 ± 0.50	-0.01 ± 0.58

Fig. 3. Results of sensitivity for different folds and classifiers using principal components.
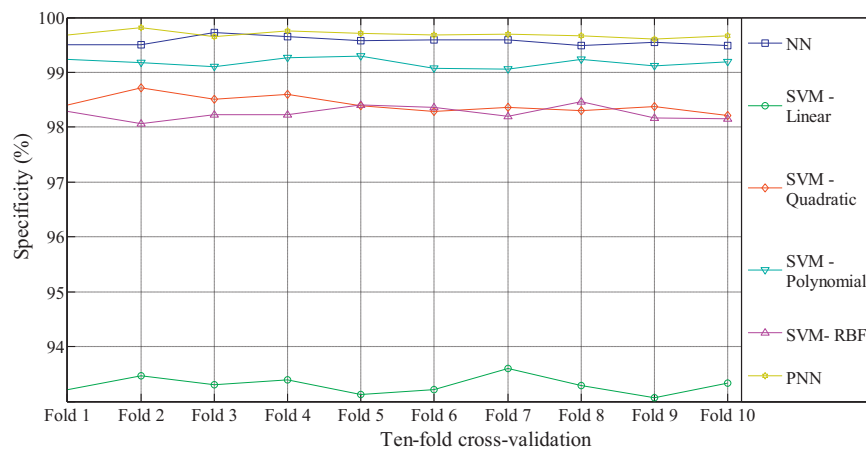

Fig. 4. Results of specificity for different folds and classifiers using principal components.

the remaining one set (11,009 beats) was used for testing and performance evaluation of the classifier. The process was repeated ten times so that each subset is used into testing. The overall performance of the classifier is evaluated by taking the average of ten folds. The correct classification or misclassification was assessed as True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN). Based on these measures the sensitivity, specificity, PPV and accuracy were determined. Fig. 3 shows the sensitivity of different classifiers during each fold of the classification using

PCA components. It can be observed from Fig. 3, that the SVM with linear kernel function provided consistently less accuracy, and the SVM with RBF kernel provides highest sensitivity. The NN and PNN yielded nearly same sensitivity. Fig. 4 shows specificity in each of the ten folds for different classifiers. It can be observed from this figure, that SVM with linear kernel provides less specificity, PNN provided highest specificity. Also NN provided the specificity close to PNN classifier. Fig. 5 shows the accuracy for different folds of different classifiers using principal components. It can be noted from
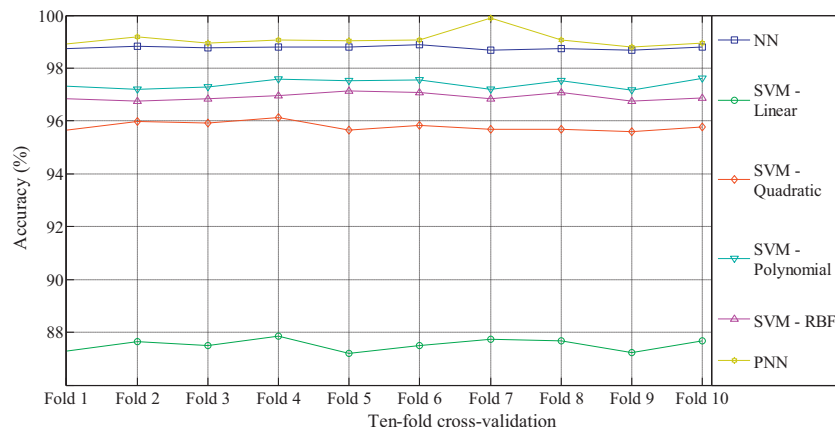

Fig. 5. Results of accuracy for different folds of classifiers using principal components.
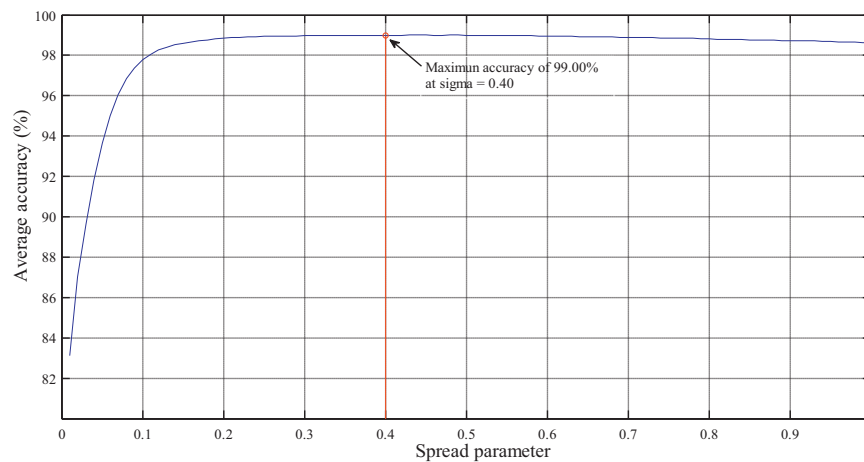
Fig. 6. Variation of average accuracy with respect to different spread values during PNN classification using PCA coefficients.

this figure, that SVM with linear kernel provided less accuracy, whereas PNN provided highest accuracy. The accuracy provided by the NN was lesser than PNN classifier. Also the PNN was tested for all possible values of spread parameter, and the average accuracy of PNN was plotted with respect to different spread values as shown in Fig. 6. This figure shows, that PNN provided highest average accuracy at spread = 0.40 and the corresponding accuracy is 99.00%.

The average performance from each of the classifiers is tabulated in Table 4. It can be seen from Table 4 that PNN with spread of 0.40 provided highest average performance with sensitivity of 97.38%, specificity of 99.69%, PPV of 98.56% and accuracy of 99.00%.

In the same way, LDA was used for dimensionality reduction. Tables 5 and 6 show the summary of results of LDA features extracted from DWT. Table 5 shows the LDA features on 4th level approximation sub-band, and Table 6 shows the LDA features extracted from 4th level detail sub-band. All these features were clinically significant ($p < 0.0001$).

The performance of sensitivity due to LDA features extracted from DWT coefficients for different folds of classifiers is shown in Fig. 7. It is seen from Fig. 7 that SVM with RBF kernel provided highest sensitivity. Fig. 8 shows the performance of specificity of classification of LDA features. It could be seen from Fig. 8 that NN provided highest specificity. Fig. 9 shows accuracy of each folds of classification due to LDA features. It is seen from figure, that NN provided consistently higher accuracy, and PNN provided accuracy close to NN classifier. Fig. 9 depicts the accuracy for different folds of the classifiers using LDA features. The figure shows that the PNN provided highest accuracy of 98.36% for spread = 0.32. The best spread ($\sigma = 0.32$) value of PNN was identified using brute-force method for average accuracy was shown in Fig. 10. The average performance of different classifiers was shown in Table 7 for LDA features. It can be seen from table that NN provided highest performance of 96.00% of sensitivity, 99.58% of specificity, 98.03% of PPV and 98.59% of accuracy respectively.

Tables 8 and 9 show the summary of ICA features extracted from 4th level approximation sub-band and 4th level detail sub-band.

It was found that all the ICA features were statistically significant ($p < 0.0001$). The sensitivity of classification of ICA components for different folds of the classifiers is shown plotted in Fig. 11. It can be seen from this figure that, SVM with RBF kernel provided highest sensitivity consistently for all folds. Fig. 12 shows specificity of ICA features with different classifiers for different folds. This figure shows that PNN provided highest specificity for ICA features. Fig. 13 shows the accuracy of ICA features for different folds. This figure shows that PNN provided consistently high accuracy for all folds. Also the best spread of PNN was identified by brute-force method. Fig. 14 shows variation of average accuracy of PNN for different spread values. It was seen that PNN provided highest accuracy of 99.28% at spread of 0.03. Table 10 summarizes the average performance of different classifiers for extracted ICA features. It is seen from Table 10 that ICA features provided highest performance of 97.97% of sensitivity, 99.83% of specificity, 99.21% of PPV and 99.28% of accuracy.

To carry out the computations, MATLAB 2012A was used. All the methods including DWT denoising, PCA, LDA were developed in MATLAB with the custom software. To implement ICA, the fastICA toolbox in MATLAB was installed and linked to the developed programs. The MATLAB routines were used for classifiers.

5. Discussion

The aim of this paper is to investigate various dimensionality reduction techniques on compactly supported Discrete Wavelet Transform basis space. The hidden complexities in the ECG signal were visible more clearly in the transform domain than the conventional time domain. Moreover the DWT representation is sparse, so more information is contained in few coefficients of DWT. When the dimensionality reduction method was applied on a sparse representation, the features would contain a compact representation. When these features were classified by automated classifiers, it would result in higher accuracy. Three commonly used dimensionality reduction methods, PCA, LDA and ICA were used to reduce the dimensionality of DWT coefficients.

Table 4
Classification results using principal components.

Classifier	Average sensitivity	Average specificity	Average PPV	Average accuracy
NN	97.26	99.57	97.99	98.78
SVM – linear	86.95	93.30	77.01	87.53
SVM – quadratic	93.95	98.41	92.86	95.79
SVM – polynomial	95.65	99.18	96.23	97.40
SVM – RBF	98.87	98.26	92.59	96.92
PNN with $\sigma = 0.40$	97.38	99.69	98.56	99.00

Table 5
Summary of six linear discriminant features extracted from 4th level approximation coefficients (*p* value <0.0001).

Feature	N	S	V	F	U
LD1	0.23 ± 0.21	−0.18 ± 0.34	−0.72 ± 0.48	0.94 ± 0.27	1.76 ± 0.95
LD2	−0.67 ± 0.36	−0.56 ± 0.61	−3.65 ± 0.81	−0.39 ± 0.44	−4.36 ± 1.29
LD3	−2.55 ± 0.78	−0.95 ± 0.60	−8.39 ± 1.41	−4.21 ± 0.75	−18.89 ± 3.40
LD4	0.04 ± 0.63	0.88 ± 0.35	−5.26 ± 1.36	−0.37 ± 1.07	−7.65 ± 2.86
LD5	2.67 ± 1.31	1.03 ± 0.50	6.43 ± 3.09	2.46 ± 1.10	19.05 ± 4.81
LD6	1.02 ± 0.74	−2.69 ± 1.71	7.68 ± 1.56	4.77 ± 0.97	17.34 ± 2.58

Table 6
Summary of six linear discriminant features extracted from 4th level detail coefficients (*p* value <0.0001).

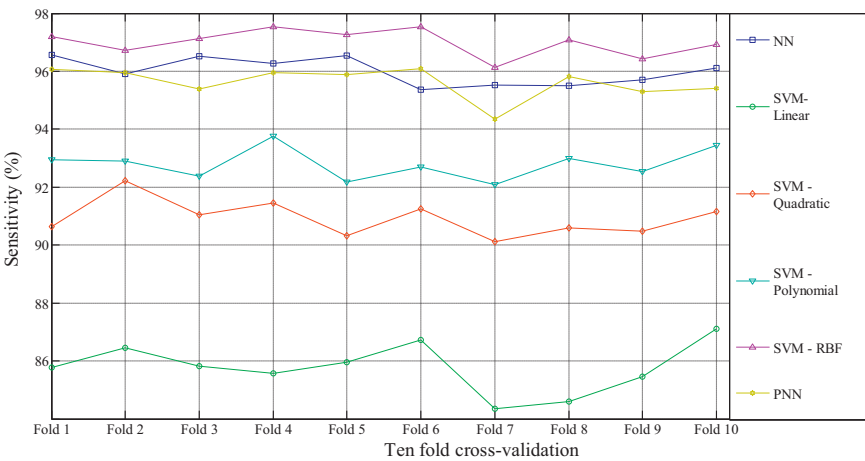| Feature	N	S	V	F	U
LD1	−0.45 ± 0.41	−0.48 ± 0.30	0.58 ± 0.20	−0.83 ± 0.41	−0.50 ± 0.13
LD2	−0.71 ± 0.34	−0.12 ± 0.18	0.79 ± 0.37	−0.48 ± 0.24	−0.65 ± 0.15
LD3	−0.83 ± 1.08	0.32 ± 0.80	−0.88 ± 0.46	0.59 ± 1.06	0.48 ± 0.34
LD4	−0.56 ± 0.80	0.03 ± 0.54	−1.44 ± 0.55	0.20 ± 0.72	0.43 ± 0.34
LD5	−1.82 ± 1.13	−0.01 ± 0.77	0.25 ± 0.62	−0.54 ± 1.12	−1.11 ± 0.41
LD6	−2.26 ± 1.24	−0.82 ± 0.65	0.05 ± 0.43	−1.09 ± 1.19	−0.29 ± 0.40

Fig. 7. Results of sensitivity for different folds and classifiers using linear discriminant coefficients.
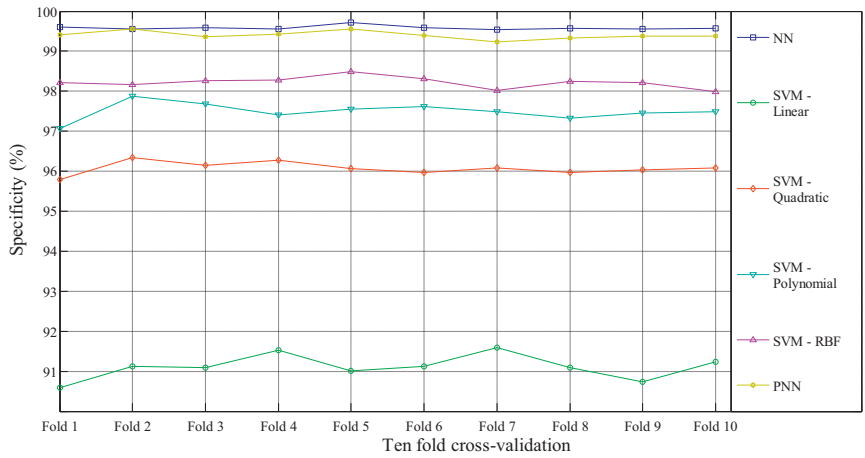

Fig. 8. Results of specificity for different folds and classifiers using linear discriminant coefficients.

Table 7
Classification results using linear discriminant components.

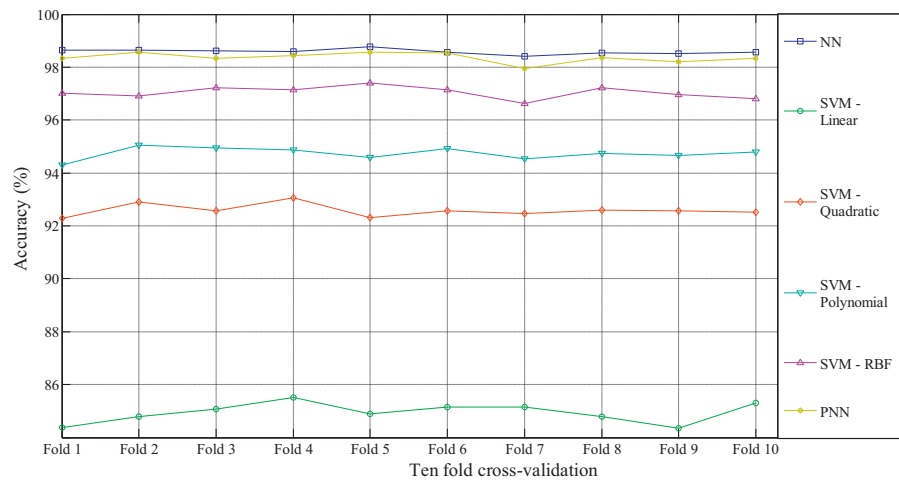| Classifier	Average sensitivity	Average specificity	Average PPV	Average accuracy
NN	96.00	99.58	98.03	98.59
SVM – linear	85.79	91.11	73.07	84.94
SVM – quadratic	90.93	96.08	85.07	92.58
SVM – polynomial	92.79	97.49	89.47	94.74
SVM – RBF	97.00	98.22	92.31	97.04
PNN with σ = 0.32	95.62	99.40	97.18	98.36

Fig. 9. Results of accuracy for different folds and classifiers using linear discriminant coefficients.
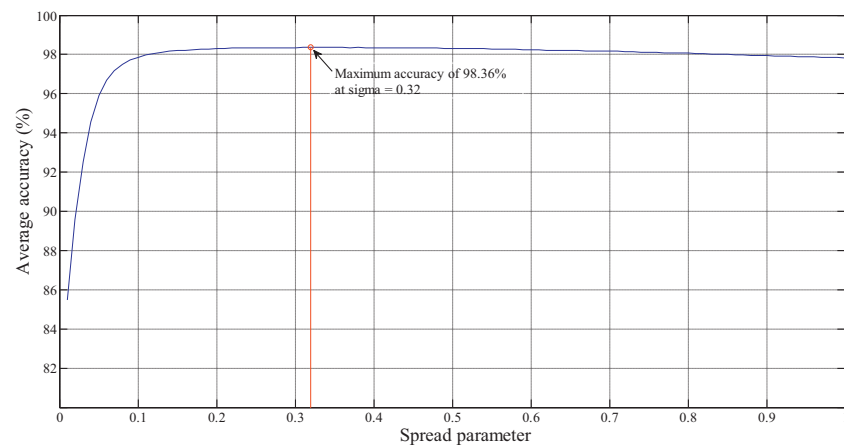

Fig. 10. Variation of average accuracy with respect to different spread values during PNN classification using linear discriminant coefficients.

Table 8
Summary of six independent component features extracted from 4th level approximation coefficients (p value <0.0001).

Feature	N	S	V	F	U
IC1	0.0084 ± 0.1386	-0.0176 ± 0.0925	0.0028 ± 0.2689	0.0022 ± 0.1535	0.0122 ± 0.2845
IC2	0.0532 ± 0.1958	0.0359 ± 0.1367	-0.0171 ± 0.2966	0.0384 ± 0.2620	0.0965 ± 0.2553
IC3	0.0000 ± 0.1735	-0.0082 ± 0.1120	-0.0032 ± 0.2748	-0.0232 ± 0.2609	0.0549 ± 0.3472
IC4	-0.0383 ± 0.1602	0.0011 ± 0.1028	-0.1097 ± 0.2923	-0.0089 ± 0.2471	0.0277 ± 0.3350
IC5	-0.0104 ± 0.1682	0.0139 ± 0.2025	-0.0623 ± 0.3776	0.0069 ± 0.2549	-0.1256 ± 0.2856
IC6	-0.0491 ± 0.1775	-0.0093 ± 0.1697	0.0122 ± 0.3124	0.0146 ± 0.3743	-0.0239 ± 0.2769

Table 9
Summary of six independent component features extracted from 4th level detailed coefficients (p value <0.0001).

Feature	N	S	V	F	U
IC1	-0.0176 ± 0.1111	-0.0140 ± 0.0795	-0.0070 ± 0.0631	-0.0138 ± 0.1125	-0.0101 ± 0.0688
IC2	-0.0384 ± 0.0829	-0.0093 ± 0.0719	-0.0248 ± 0.0619	-0.0168 ± 0.0794	0.0057 ± 0.0555
IC3	0.0122 ± 0.0823	0.0040 ± 0.0528	-0.0062 ± 0.0735	0.0106 ± 0.0875	-0.0196 ± 0.0578
IC4	0.0076 ± 0.1097	0.0197 ± 0.0806	0.0097 ± 0.0753	0.0279 ± 0.1049	0.0036 ± 0.0790
IC5	0.0136 ± 0.0640	0.0112 ± 0.1314	0.0056 ± 0.0521	0.0169 ± 0.0446	0.0127 ± 0.0603
IC6	0.0067 ± 0.0284	0.0079 ± 0.0703	-0.0004 ± 0.0453	0.0031 ± 0.0278	0.0014 ± 0.0360

Table 10
Classification results using independent component coefficients.

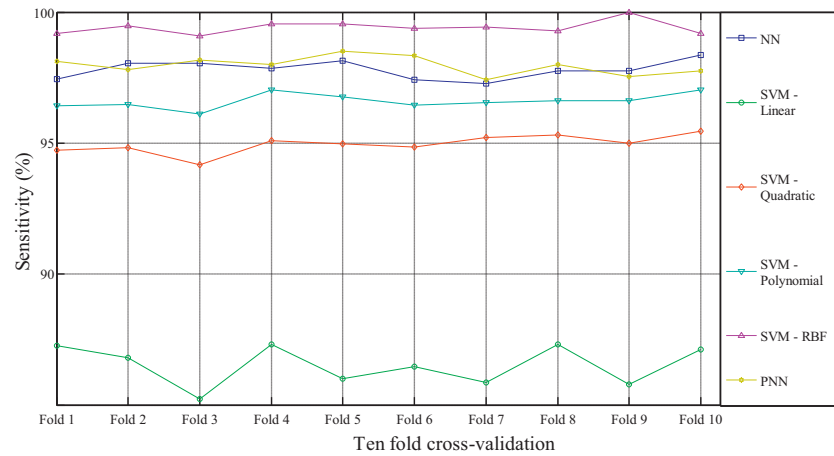| Classifier	Average sensitivity	Average specificity	Average PPV	Average accuracy
NN	97.81	99.70	98.59	99.00
SVM – linear	86.51	95.07	81.41	88.89
SVM – quadratic	94.96	99.03	95.54	96.74
SVM – polynomial	96.61	99.39	97.24	97.84
SVM – RBF	99.31	99.01	95.69	98.36
PNN with $\sigma = 0.32$	97.97	99.83	99.21	99.28

Fig. 11. Results of sensitivity for different folds and classifiers using Independent component coefficients.
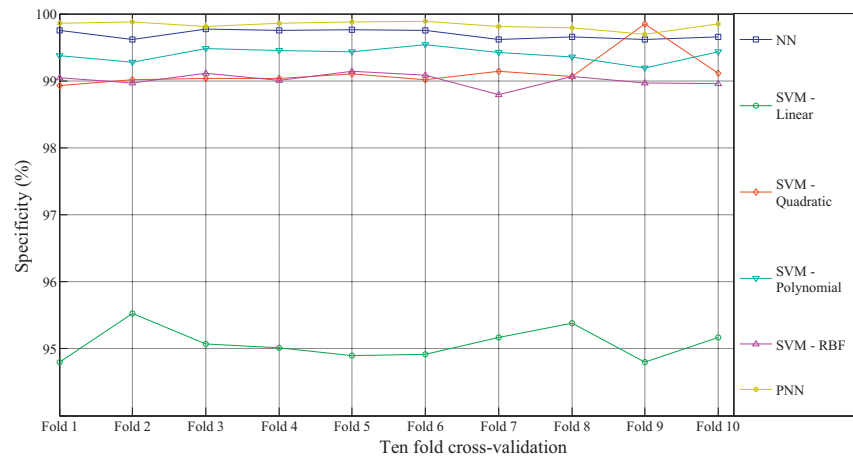

Fig. 12. Results of specificity for different folds and classifiers using Independent component coefficients.

Table 11, provides a comprehensive summary of automated classification of ECG beats using the MIT-BIH arrhythmia database. Normal, SVEB, VEB and fusion beats were classified using mixture of experts approach and reported an accuracy of 94% [32]. They used morphology and RR interval features for classification. The multi-scale wavelet features along with timing information were used for classification of three types of ECG beats [33]. They have achieved a classification accuracy of 95.16%. The ECG beats were decomposed into finite characteristic waveforms using a sum of Gaussian kernels and reported 99.1% of accuracy for three types of ECG beats (normal, VPC and other possible beats) using Bayesian filtering [34]. SVEB and VEB were classified with 95.9% and 99.2% accuracy respectively using a patient adapting scheme [35]. Five types of heartbeats were classified using Hermite transform coefficients and RR interval features with block based neural network and obtained 96.6% of accuracy [36]. The five types of beats recommended by AAMI were
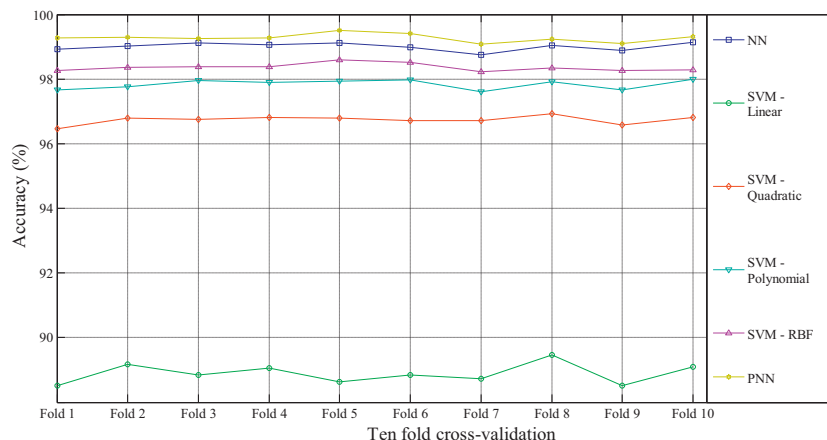

Fig. 13. Results of accuracy for different folds and classifiers using Independent component coefficients.
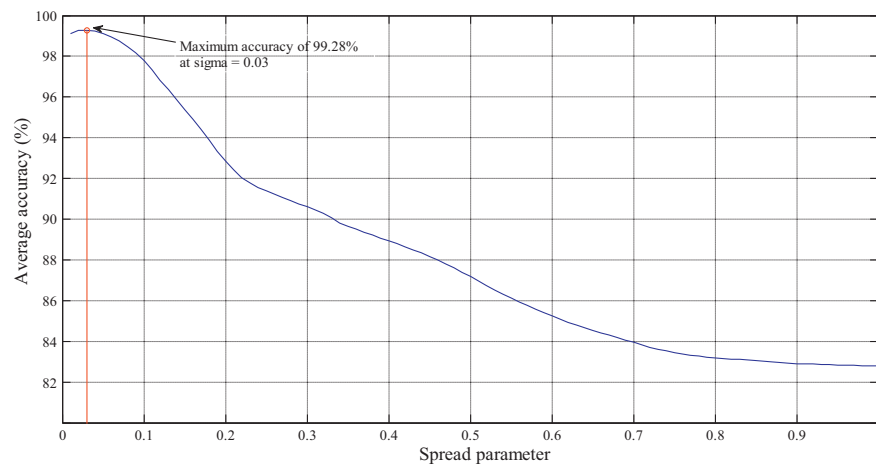
Fig. 14. Variation of average accuracy with respect to different spread values during PNN classification using independent component coefficients.

Table 11
Comparison of ECG beat classification methods on MIT BIH arrhythmia database.

Literature	Features	Classifier	Classes	Accuracy
Two class classification				
Hu et al. [32]	Time domain features	Mixture of experts	2	94.00
Inan et al. [37]	WT and timing interval	Neural network	2	95.20
Sayadi et al. [40]	Innovation sequence of EKF	Bayesian filtering	2	99.10
Three class classification				
Llamedo and Martinez [41]	RR interval and its derived features	Linear discriminant and expectation maximization	3	98.00
Five class classification				
de Chazal and Reilly [35]	Morphology and heartbeat interval	Linear discriminant	5	85.9
Jiang and Seong [36]	Hermite function parameters and RR interval	Block based NN	5	96.6
Ince et al. [39]	WT +PCA	Multidimensional particle swarm optimization	5	95.58*
Martis et al. [22]	PCA	SVM with RBF kernel	5	98.11
Martis et al. [38]	Bispectrum + PCA	SVM with RBF kernel	5	93.48
More than five class classification				
Osowski and Linh [34]	HOSA	Hybrid fuzzy NN	7	96.06
Lagerholm et al. [33]	Hermite functions	Self organizing map	25	98.51
Proposed methodology				
Martis et al. (2013)	DWT+ICA	PNN	5	99.28

classified using morphological wavelet transform and time interval features and obtained 98.3% and 97.4% of accuracy for the detection of VEB and SVEB respectively [37].

The five types of ECG beats (normal, right bundle branch block, left bundle branch block, atrial premature contraction and ventricular premature contraction) were classified with 98.11% of accuracy using the Principal Component Analysis of time domain ECG samples [22]. The same five types of beats were classified with an accuracy of 93.48% using principal components of bispectrum with SVM and RBF kernel [38].

Using fuzzy hybrid neural network classifier and Higher Order Spectra Analysis (HOSA) features, the seven ECG beats were classified [39]. Their system yielded a classification accuracy of 96.06%. Twenty-five types of ECG beats were classified using Hermite function decomposition coupled with self organizing map (SOM) and reported an accuracy of 98.51% [40]. A patient adaptable algorithm was developed and tested on more than 9,000,000 ECG beats and reported maximum accuracy of 98% [41].

In our present study, the performance of three dimensionality reduction methods was tested on DWT coefficients for the five classes as recommended by AAMI on the MIT-BIH arrhythmia database. The DWT provides energy compaction and which dimensionality reduction method would provide higher discrimination in the dimensionality reduced space depends on the data. There is no hypothesis saying which method would provide higher discrimination in order to get higher classification performance. In this study we have experimentally demonstrated that application of PCA on

DWT coefficients has provided 99.00% accuracy, using PNN classifier of spread parameter as 0.40. Also the LDA on DWT coefficients has provided 98.59% of accuracy with NN classifier. However the ICA on DWT coefficients has yielded highest performance with 99.28% accuracy using PNN classifier with spread parameter of 0.03.

The developed methodology has many practical applications including Holter monitoring with automated alarm generation during the onset of disease, telemedicine applications, cardiac pacemakers, remote patient monitoring etc. With the demonstration of the proposed methodology, the ICA on DWT coefficients would be more robust and yielded good classification accuracy. The classification accuracy is tested on large database (110,094 beats), the errors are very less as evidenced by the 99.28% of accuracy using ten-fold cross validation, one can use it for these practical applications. The performance of the developed method may be increased further using better feature sets and more robust classifiers.

6. Conclusion

The ECG signal depicts the electrical activity of the heart providing vital information about the cardiac state. In this study three dimensionality reduction methods on DWT coefficients were compared and the performance was analyzed. It was shown that ICA with combination of PNN classifier ($\sigma = 0.03$) performed the highest average accuracy, sensitivity and specificity of 99.28%, 97.97%, and 99.83% respectively. The best spread for providing

highest accuracy was searched by the brute-force method for each of the dimensionality reduction method. The developed methodology in this study can be used in arrhythmia monitoring systems, telemedicine applications, cardiac pacemakers and mass screening for cardiac health.

References

[1] S.S. Anand, S. Yusuf, Stemming the global tsunami of cardiovascular disease, Lancet 377 (February (9765)) (2011) 529–532.

[2] D. Lloyd-Jones, R. Adams, M. Carnethon, G. De Simone, T.B. Ferguson, K. Flegal, E. Ford, K. Furie, A. Go, K. Greenlund, N. Haase, S. Hailpern, M. Ho, V. Howard, B. Kissela, S. Kittner, D. Lackland, L. Lisabeth, A. Marelli, M. McDermott, J. Meigs, D. Mozaffarian, G. Nichol, C. O'Donnell, V. Roger, W. Rosamond, R. Sacco, P. Sorlie, R. Stafford, J. Steinberger, T. Thom, S. Wasserthiel-Smoller, N. Wong, J. Wylie-Rosett, Y. Hong, American Heart Association Statistics Committee and Stroke Statistics Subcommittee, heart disease and stroke statistics – 2009 update: a report from the American Heart Association Statistics Committee and Stroke Statistics Subcommittee, Circulation 119 (January (3)) (2009) 480–486.

[3] H.V. Huikuri, A. Castellanos, R.J. Myerburg, Sudden death due to cardiac arrhythmias, New England Journal of Medicine 345 (20) (2001) 1473–1482.

[4] M. Hadhoud, M. Eladawy, A. Farag, Computer aided diagnosis of cardiac arrhythmias, in: Proc. IEEE Int. Conf. Computer Engineering and Systems, 2006, pp. 262–265.

[5] A.K. Mishra, S. Raghav, Local fractal dimension based ECG arrhythmia classification, Biomedical Signal Processing and Control 5 (April (2)) (2010) 114–123.

[6] A. Khazaee, A. Ebrahimzadeh, Classification of electrocardiogram signals with support vector machines and genetic algorithms using power spectral features, Biomedical Signal Processing and Control 5 (October (4)) (2010) 252–263.

[7] A. Daamouche, L. Hamami, N. Alajlan, F. Melgani, A wavelet optimization approach for ECG signal classification, Biomedical Signal Processing and Control 7 (4) (2011) 342–349.

[8] T. Fancott, D.H. Wong, A minicomputer system for direct high speed analysis of cardiac arrhythmia in 24 h ambulatory ECG tape recordings, IEEE Transactions on Biomedical Engineering BME-27 (December (12)) (1980) 685–693.

[9] K.-P. Lin, W.H. Chang, QRS feature extraction using linear prediction, IEEE Transactions on Biomedical Engineering 36 (October (10)) (1989) 1050–1055.

[10] D.A. Coast, R.M. Stern, G.G. Cano, S.A. Briller, An approach to cardiac arrhythmia analysis using hidden Markov models, IEEE Transactions on Biomedical Engineering 37 (September (9)) (1990) 826–836.

[11] F. Melgani, Y. Bazi, Classification of electrocardiogram signals with support vector machines and particle swarm optimization, IEEE Transactions on Information Technology in Biomedicine 12 (September (5)) (2008) 667–677.

[12] D. Ge, N. Srinivasan, S.M. Krishnan, Cardiac arrhythmia classification using autoregressive modelling, Biomedical Engineering Online 1 (5) (2002) 1–12.

[13] T.H. Linh, S. Osowski, M. Stodolski, On-line heart beat recognition using Hermite polynomials and neuro-fuzzy network, IEEE Transactions on Instrumentation and Measurement 52 (August (4)) (2003) 1224–1231.

[14] O.T. Inan, D. Park, L. Giovangrandi, G.T.A. Kovacs, Noninvasive measurement of physiological signals on a modified home bathroom scale, IEEE Transactions on Biomedical Engineering 59 (August (8)) (2012) 2137–2143.

[15] R.J. Martis, C. Chakraborty, A.K. Ray, A two-stage mechanism for registration and classification of ECG using Gaussian mixture model, Pattern Recognition 42 (November (11)) (2009) 2979–2988.

[16] R.J. Martis, M.M.R. Krishnan, C. Chakraborty, S. Pal, D. Sarkar, A.K. Ray, Automated screening of arrhythmia using wavelet based machine learning techniques, Journal of Medical Systems 36 (2) (2012) 677–688.

[17] ANSI/AAMI EC57: Testing and Reporting Performance Results of Cardiac Rhythm and ST Segment Measurement Algorithms (AAMI Recommended Practice/American National Standard). Available: http://www.aami.org, Order Code: EC57-293, 1998.

[18] R. Mark, G. Moody, MIT-BIH Arrhythmia Database, 1997, May, Available at: http://ecg.mit.edu/dbinfo.html

[19] B.N. Singh, A.K. Tiwari, Optimal selection of wavelet basis function applied to ECG signal denoising, Digital Signal Processing 16 (May (3)) (2006) 275–287.

[20] J. Pan, W.J. Tompkins, A real-time QRS detection algorithm, IEEE Transactions on Biomedical Engineering BME-32 (March (3)) (1985) 230–236.

[21] U.R. Acharya, P.K. Joseph, N. Kannathal, C.M. Lim, J.S. Suri, Heart rate variability: a review, IFMBE Journal of Medical & Biological Engineering & Computing Journal 44 (12) (2006) 1031–1051.

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

[23] R.J. Martis, C. Chakraborty, A.K. Ray, An integrated ECG feature extraction scheme using PCA and wavelet transform, in: IEEE INDICON-2009, 2009, ISBN: 978-1-4244-4859-3/09.

[24] R.O. Duda, P.E. Hart, D.G. Stork, Pattern Classification, 2nd ed., Wiley, New York, 2001.

[25] A. Hyvärinen, E. Oja, Independent component analysis: algorithms and applications, Neural Networks 13 (June (4–5)) (2000) 411–430.

[26] C.M. Bishop, Neural Networks for Pattern Recognition, Clarendon Press, Oxford, 1995.

[27] S.S. Haykin, Neural Networks: A Comprehensive Foundation, 2nd ed., Prentice Hall, New Jersey, 1999.

[28] S. Gunn, Support Vector Machines for Classification and Regression, Technical Report, University of Southampton, 1998, May.

[29] N. Christianini, J.S. Taylor, An Introduction to Support Vector Machines and Other Kernel Based Learning Methods, Cambridge University Press, Cambridge, MA, 2000.

[30] J.A.K. Suykens, J. Vandewalle, Least square support vector machine classifiers, Neural Processing Letters 9 (1999) 293–300.

[31] R.J. Martis, U.R. Acharya, A.K. Ray, C. Chakraborty, Application of higher order cumulants to ECG signals for the cardiac health diagnosis, in: Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE, August 30 2011–September 3 2011, 2011, pp. 1697–1700.

[32] Y.H. Hu, S. Palreddy, W.J. Tompkins, A patient-adaptable ECG beat classifier using a mixture of experts approach, IEEE Transactions on Biomedical Engineering 44 (September (9)) (1997) 891–900.

[33] M. Lagerholm, C. Peterson, G. Braccini, L. Edenbrandt, L. Sornmo, Clustering ECG complexes using Hermite functions and self-organizing maps, IEEE Transactions on Biomedical Engineering 47 (July (7)) (2000) 838–848.

[34] S. Osowski, T.H. Linh, ECG beat recognition using fuzzy hybrid neural network, IEEE Transactions on Biomedical Engineering 48 (November (11)) (2000) 1265–1271.

[35] P. de Chazal, R.B. Reilly, A patient-adapting heartbeat classifier using ECG morphology and heartbeat interval features, IEEE Transactions on Biomedical Engineering 53 (December (12)) (2006) 2535–2543.

[36] W. Jiang, S.G. Kong, Block-based neural networks for personalized ECG signal classification, IEEE Transactions on Neural Networks 18 (6) (2007) 1750–1761.

[37] O.T. Inan, L. Giovangrandi, G.T.A. Kovacs, Robust neural-network-based classification of premature ventricular contractions using wavelet transform and timing interval features, IEEE Transactions on Biomedical Engineering 53 (December (12)) (2006) 2507–2515.

[38] R.J. Martis, U.R. Acharya, K.M. Mandana, A.K. Ray, C. Chakraborty, Cardiac decision making using higher order spectra, Biomedical Signal Processing and Control (2012), http://dx.doi.org/10.1016/j.bspc.2012.08.004, Available online 4 September 2012, ISSN 1746-8094.

[39] T. Ince, S. Kiranyaz, M. Gabbouj, A generic and robust system for automated patient-specific classification of ECG signals, IEEE Transactions on Biomedical Engineering 56 (May (5)) (2009) 1415–1426.

[40] O. Sayadi, M.B. Shamsollahi, G.D. Clifford, Robust detection of premature ventricular contractions using a wave-based bayesian framework, IEEE Transactions on Biomedical Engineering 57 (February (2)) (2010) 353–362.

[41] M. Llamedo, J.P. Martinez, An automatic patient adapted ECG heartbeat classifier allowing expert assistance, IEEE Transactions on Biomedical Engineering 59 (August (8)) (2012) 2312–2320.

[42] R.J. Martis, C. Chakraborty, Arrhythmia disease diagnosis using neural network, SVM, and genetic algorithm-optimized k-means clustering, Journal of Mechanics in Medicine and Biology 11 (2011) 897–915.

[44] S. Banik, R.J. Martis, D. Nayak, Code excited linear prediction codec for electrocardiogram, in: Engineering in Medicine and Biology Society, 2004. IEMBS'04. 26th Annual International Conference of the IEEE, 2004, pp. 160–163.

[45] P.S. Addison, Wavelet transforms and the ECG: a review, Physiological Measurement 26 (5) (2005) R155–R199.