

Discrimination of the Heart Ventricular and Atrial Abnormalities via a Wavelet-Aided Adaptive Network Fuzzy Inference System (ANFIS) Classifier

A. Ghaffari*, M. R. Homaeinezhad** and M. Akraminia**

Abstract: The aim of this study is to address a new feature extraction method in the area of the heart arrhythmia classification based on a metric with simple mathematical calculation called Curve-Length Method (CLM). In the presented method, curve length of the under study excerpted segment of signal is considered as an informative feature in which the effect of important geometric parameters of the original signal can be found. To show merits of the presented method, first the original electrocardiogram (ECG) in lead I is pre-processed by removing its baseline wander then by scaling it in the [-1,1] interval. In the next step, using à trous method, discrete wavelet scales 2^3 and 2^4 and smoothing function scale 2^2 are extracted. Afterwards, segments including samples of the QRS complex, P and T waves are estimated via an approximation criterion and CLM is implemented to extract corresponding features from aforementioned scales, smoothing function and also from each original segment. The resulted feature vector (including 12 components) is used to tune an Adaptive Network Fuzzy Inference System (ANFIS) classifier. The presented strategy is applied to classify four categories found in the MIT-BIH Arrhythmia Database namely as Atrial Premature Beat (APB), Left Bundle Branch Block (LBBB), Right Bundle Branch Block (RBBB) and Premature Ventricular Contraction (PVC) and average values of Se = 99.81%, P+ = 99.80%, Sp = 99.81% and Acc = 99.72% are obtained for sensitivity, positive predictivity, specificity and accuracy respectively showing marginal improvement of the heart arrhythmia classification performance.

Keywords: Feature Extraction, Curve Length Method, Adaptive Network Fuzzy Inference System, Arrhythmia Classification, Left Bundle Branch Block, Right Bundle Branch Block, Atrial Premature Beat, Premature Ventricular Beat.

1 Introduction

Heart is a special muscle which its constitutive cells (myocytes) possess two important characteristics namely as nervous excitability and mechanical tension with force feedback. The superposition of all myocytes electrical activity on the skin surface results a detectable potential difference which its detection and registration is called electrocardiography [1]. If according to any happening, nervous and/or mechanical function of a region of myocytes encounter with failure, the corresponding abnormal effects will appear in the ECG

signal [2-5] and in the heart hemodynamic performance [6]. Today, the importance of elaboration and design of a classification algorithm with the ability of categorization of arrhythmias with acceptable accuracy within reasonable speed and computational burden is undeniable and correspondingly many researches have been conducted in this context.

Based on a comprehensive literature survey among many documented works, it is seen that several features and extraction (selection) methods have been created and implemented by authors. For example, original electrocardiogram (ECG) signal [17], preprocessed ECG signal via appropriately defined and implemented transformations such as discrete wavelet transform (DWT), continuous wavelet transform (CWT) [21], Hilbert transform (HT) [64], fast Fourier transform (FFT) [48-49], short time Fourier transform (STFT) [10], power spectral density (PSD) [51-52], higher order spectral methods [46-47], statistical

Iranian Journal of Electrical & Electronic Engineering, 2010.

Paper first received 23 Aug. 2009 and in revised form 21 Dec. 2009.

* The Author is with the Department of Mechanical Engineering, K. Nassir Toosi University of Technology.

E-mail: ghaffari@kntu.ac.ir

** The Authors are Members of Cardio Vascular Research Group (CVRG), Department of Mechanical Engineering, K. Nassir Toosi University of Technology.

E-mails: mrezaahmadi@yahoo.com, mahdi.akraminia@gmail.com

moments [24], nonlinear transformations such as Liapunov exponents and fractals [43-45] have been used as appropriate sources for feature extraction. In order to extract feature(s) from a selected source, various methodologies and techniques have been introduced. To meet this end, the first step is segmentation and excerption of the specific parts of the preprocessed trend (for example, in the area of the heart arrhythmia classification, ventricular depolarization regions are the most used segments). Afterwards, appropriate and efficient features can be calculated from excerpted segments via a useful method. Up to now, various techniques have been proposed for the computation of feature(s). For example mean, standard deviation, maximum value to minimum value ratio, maximum-minimum slopes, summation of point to point difference, area, duration of events, correlation

coefficient with a pre-defined waveform template, statistical moments of the auto (cross) correlation functions with a reference waveform [32], bi-spectrum [46], differential entropy [37], mutual information [39], nonlinear integral transforms and some other more complicated structures [33-45] may be used as an instrument for calculation of features.

After generation of the feature source, segmentation, feature selection and extraction (calculation), the resulted feature vectors should be divided into two groups "train" and "test" to tune an appropriate classifier such as a neural network, support vector machine or ANFIS, [30-40]. The general block diagram of the heart arrhythmia recognition-classification algorithms is shown in Fig. 1.

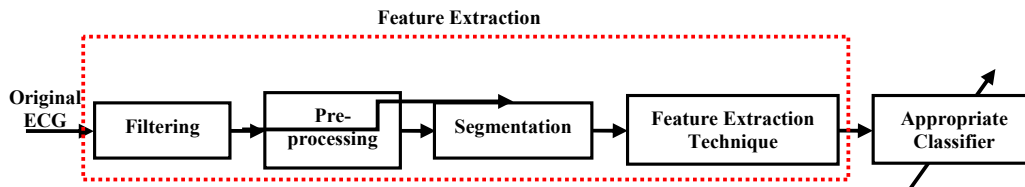


Fig. 1 General overview of the heart arrhythmia recognition-classification procedure.

As previous researches show, occurrence of arrhythmia(s) affects RR-tachogram and Heart Rate Variability (HRV) in such a way that these quantities can be used as good features to classify arrhythmias. Using RR-tachogram or HRV analysis in feature extraction and via simple if-then or other parametric or nonparametric classification rules [7-9], artificial neural networks, fuzzy or ANFIS networks [10-14], support vector machines [15] and probabilistic frameworks such as Bayesian hypotheses tests [16], arrhythmia classification would be fulfilled with acceptable accuracies. Up to now, the main concentration of the arrhythmia classification schemes has been on morphology assessment and/or geometric parameters of the ECG events. Traditionally, in the studies based on morphology and wave geometry, first, during a preprocessing stage, some corrections such as baseline wander removal, noise-artifact rejection and a suitable scaling are applied. Then, using an appropriate mapping for instance, filter banks, discrete or continuous wavelet transform in different spatial resolutions and etc., more information is derived from the original signal for further processing and analyses. In some researches, original and/or preprocessed signal are used as appropriate features and using artificial neural network or fuzzy classifiers [17-25], parametric and probabilistic classifiers [26-28], the discrimination goals are followed. Although in such classification approaches, acceptable results may be achieved, however, due to

implementation of the original samples as components of feature vector, computational cost and burden especially in high sampling frequencies will be very high and the algorithm may take long time to be trained for a given database. In some other researches, geometric parameters of QRS complexes such as maximum value to minimum value ratio, area under the segment, maximum slope, summation (absolute value) of point to point difference, ST-segment, PR and QT intervals, statistical parameters such as correlation coefficient of a assumed segment with a template waveform, first and second moments of original or preprocessed signal and etc. are used as effective features [29-35]. The main definition origin of these features is based on practical observations and a priori heuristic knowledge whilst conducted researches have shown that using these features convincing results may be reached. On the other hand, some of studies in the literature focus on the ways of choosing and calculating efficient features to create skillfully an efficient classification strategy [36-39]. In the area of nonlinear systems theory, some ECG arrhythmia classification methods on the basis of fractal theory [40, 41], state-space, trajectory space, phase space, Liapunov exponents [42-44] and nonlinear models [45] have been innovated by researches. Amongst other classification schemes, structures based on higher order statistics in which to analyze features, a two or more dimensional frequency space is constructed can be mentioned [46,

47]. According to the concept that upon appearance of changes in the morphology of ECG signal caused by arrhythmia, corresponding changes are seen in the frequency domain, some arrhythmia classifiers have been designed based on appropriate features obtained from signal fast Fourier transform (FFT), short-time Fourier transform (STFT), auto regressive (AR) models and power spectral density (PSD), [48-53]. Finally, using some polynomials such as Hermite function which has specific characteristics, effective features have been extracted to classify some arrhythmias, [54, 55].

In this study, based on simple mathematical calculations, a new feature extraction method for ECG arrhythmia classification namely as CLM is introduced. In study [63] it was shown that many geometric parameters of a complex such as mean, variance, maximum value to minimum value ratio, (absolute) area under waveform, summation of point to point difference and duration of events lie on. To show merits of the presented method, first the original ECG in lead I is pre-processed by removing its baseline wander and then by scaling it in the $[-1,1]$ interval. In the next step, using a trous method, discrete wavelet scales 2^3 and 2^4 and smoothing function in scale 2^2 are extracted. Afterwards, segments including samples of the QRS complex, P and T waves are estimated via an approximation criterion and CLM is implemented to extract corresponding features from aforementioned scales, smoothing function and also from each original segment. The resulted feature vector (including 12 components) is used to tune an ANFIS classifier. The presented strategy is applied to classify four categories found in the MIT-BIH Arrhythmia Database namely as Atrial Premature Beat (APB), Left Bundle Branch Block (LBBB), Right Bundle Branch Block (RBBB) and Premature Ventricular Contraction (PVC) and average values of $Se = 99.81\%$, $P+ = 99.80\%$, $Sp = 99.81\%$ and $Ac = 99.72\%$ are obtained for sensitivity, positive predictivity, specificity and accuracy, respectively. Acceptable performance, low training computational burden and quick responding in test are important merits of the presented algorithm. The organization of the manuscript consists of the following sections: in section 2, preparation of required data for training and testing including choice of target database, baseline correction (and projection), segmentation, in section 3, DWT implementation, in section 4 feature extraction via CLM, in section 5 feature space dimension reduction including PCA, LDA and GDA, in section 6 ANFIS classifier design, in section 7 results and discussions including multi-stage tests to assess performance characteristics of the classifier and comparison with the previous works and finally in part 8, obtained conclusions are presented.

Abbreviations are as follows:

ECG: Electrocardiogram
 CLM: Curve Length Method

ANFIS: Adaptive Network Fuzzy Inference System
 RBBB: Right Bundle Branch Block
 LBBB: Left Bundle Branch Block
 APB: Atrial Premature Beat
 PVC: Premature Ventricular Contraction
 DWT: Continuous Wavelet Transform
 SMF: Smoothing Function
 Se: Sensitivity
 P+: Positive Predictivity
 Sp: Specificity
 Acc: Accuracy
 PCA: Principal Component Analysis
 LDA: Linear Discriminant Analysis
 GDA: Generalized Discriminant Analysis
 FIR: Finite-Duration Impulse Response
 MITDB: MIT-BIH Arrhythmia Database

2 Procedure of the Presented Algorithm

The procedure of the arrhythmia classification method presented in this study primarily is composed of 11 steps as (a) choosing train and test sets, (b) preprocessing, (c) ECG events detection, (d) beat cumulation, (e) noisy beats rejection, (f) beat segmentation, (g) implementation of discrete-wavelet transform, (h) application of curve-length method for feature extraction, (i) application of Principal Component Analysis (PCA), (j) classification and (k) evaluation and validation of the algorithm. The general block-diagram of this research is depicted in Fig. 2.

2.1 Procedure of the Presented Algorithm

In this study, In order to train, test, and performance evaluation of the proposed classification algorithm, MIT-BIH Arrhythmia Database (MITDB) [69] is used. The MITDB includes 48 records with sampling frequency 360 Hz and 11-bits of resolution. For each record, annotation (*.atr), header (*.hea) and signal (*.dat) files are found. Four different categories of QRS complexes labeled with 65 (APB), 76 (LBBB), 86 (PVC) and 82 (RBBB) codes have been chosen from MITDB to assess the ideas of this study. In this part, first, the throughout of MITDB is explored to find the complexes with aforementioned codes. It should be noted that each QRS complex of MITDB was already annotated by experts and therefore it is not necessary to implement a QRS detection algorithm. However, if it is needed, one can find reliable QRS detectors based on Hilbert or Wavelet transforms.

2.2 Baseline Wander Removal and Scaling Between $[-1,1]$

In this step, baseline wander is removed from all detected complexes based on an algorithm with simple calculations. It is noted that in order to remove baseline wander, various methods have previously proposed such as Hidden Markov Models and cubic splines [62]. In this work, first, samples of a signal is divided to identical length segments (an appropriate length of

windows can be chosen between 600 to 1000 msec. [6] in rest state) and the mean value of the samples in each segment is subtracted from corresponding segment and uttermost number of samples then are moved to vicinity of zero and eventually baseline wander is eliminated. This method is an acceptable approximation to remove baseline wander and also to remove disturbing factors generating similarities. After removing baseline wander, the resulted signal is scaled to [-1,1] interval. Although scaling process destroys the quantitative information of the signal, however the original morphology of the signal is restored and the efficiency of the pattern recognition algorithm will increase.

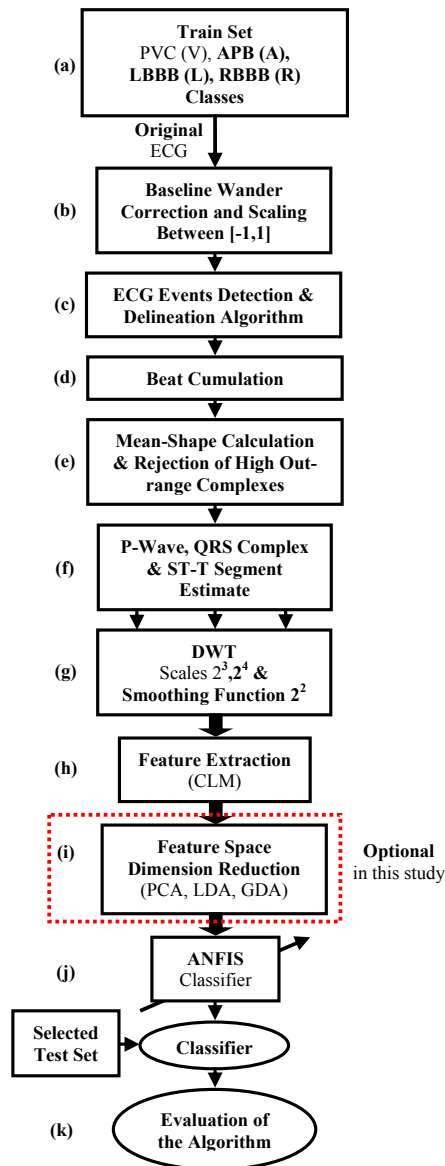


Fig. 2 The general over-view of the procedure of the presented arrhythmia classification algorithm.

Suppose that a hypothesized signal $x[n]$, $n = 0, \dots, N$ is divided into M identical segments with length of W_L samples. Therefore, each segment $Y(k)$, $k=1, \dots, M$ is obtained from $x[n]$ signal as follows

$$Y(k) = x[(k-1) \times W_L : k \times W_L] \quad (1)$$

Then the baseline wander removed segment is obtained as

$$Y_{BWR}(k) = Y(k) - \text{mean}[Y_{BWR}(k)] \quad (2)$$

where $Y_{BWR}(k)$ is the baseline wander removed segment. From the above equation, it can be proven that the mean value of the segment $Y_{BWR}(k)$ is zero. Therefore, by applying this method and augmenting all segments, the mean value of the resulted signal will be zero. It should be noted that although more accurate methods such as cubic spline or Hidden Markov Models (HMM) can be found to correct baseline, however, in the area of the heart arrhythmia recognition, it has been proven that the aforementioned approximation causes acceptable results from classification algorithm.

Although, duration of QRS complexes in different subjects follow some rather regularity, however this property does not hold about the amplitude of QRS complexes. Many different factors may affect the height of the QRS complexes in diverse number of subjects. It should be noted that the height of each measured and detected QRS complex depends strongly on the ECG machine specifications and level of noise and disturbances. Also, the height of QRS complexes depends on the pathological and physical characteristics of the subject but on the other hand, voltage of the QRS peaks (R-waves) relative to their baseline can be cast in a regular probabilistic framework similar to their duration for different subjects. Scaling is an operation that maps vertical shape of all segments into a common interval (for example [-1,+1] interval) and can the classification network to be trained more accurately. At last, it should be noted that scaling does not affect the morphology of the segmented waveform; the item that is determinant in the classification of the heart arrhythmias. This simple methodology has been implemented in some works such as Martinez et al. [62] and Arif et al. [32].

2.3 ECG Events Detection and Delineation

In this step, QRS complexes are detected and delineated. Today reliable QRS detectors based on Hilbert [64, 65] and Wavelet [62, 63] transforms can be found in literature. However, MITDB includes expert annotation files for each of 48 records, therefore, implementation of QRS detection and delineation algorithm can be ignored.

2.4 Beats Cumulation

Suppose that time duration between $(k-1)$ -th to k -th QRS peak (R-peak) is indicated by RR_k . If N number of registered beats are available, the mean value of RR_k

time-series indicated by RR_{mean} is obtained from following equation

$$RR_{mean} = \frac{1}{N-1} \sum_{k=2}^N RR_k \quad (3)$$

Then in order to excerpt a reasonable segment as a complete beat from whole preprocessed signal, each R-peak is considered as a local origin and the samples located in positions $(RR_{mean})/2$ before and after local origin, are assumed as an estimation of beat onset and offset, respectively. After segmentation of each beat, the corresponding samples are put into \mathbf{X}_{Arr} (Arr=APB, LBBB, PVC, RBBB) matrix called beat cumulation matrix, in a row-wise fashion.

2.5 Mean-Shape Calculation & Rejection of High Out-range Complexes

The aim of this step is to reject beats with very out of range and misleading morphologies. It is noted that although all beats of all records of the MITDB have been annotated, however, some beats morphologies are really unknown and can not be assumed to belong to a specific category. So, removing these types of beats which with high probability will be classified falsely (FP and/or FN increase), help us to have an accurate judgment about performance of the arrhythmia classification algorithm. To meet this end, first to find a mean-shape morphology for each arrhythmia class, the mean value of each column of \mathbf{X}_{Arr} is calculated as follows

$$\mathbf{Med}_{Arr} = \text{mean}(\mathbf{X}_{Arr}) \quad (4)$$

where, $\text{mean}(\cdot)$ is column-wise mean value operator and \mathbf{Med}_{Arr} is mean shape vector of arrhythmia with index Arr. In the next step, each row of the matrix \mathbf{X}_{Arr} indicated by $(\mathbf{X}_{Arr})_j$, $j = 1, 2, \dots, N_{Arr}$ (N_{Arr} is number of cumulated beats of arrhythmia with index Arr) is subtracted from \mathbf{Med}_{Arr} and second moment of the resulted difference vector is calculated as follows

$$\sigma_{Arr,j} = \text{std}[(\mathbf{X}_{Arr})_j - \mathbf{Med}_{Arr}] \quad (5)$$

If σ_{Arr} is a vector with $\sigma_{Arr,j}$ components, then this vector has first (μ_σ) and second (σ_σ) moments. A threshold with $\tau_\sigma = \mu_\sigma + \alpha \sigma_\sigma$ ($3.5 \leq \alpha \leq 5.5$) structure is then used to remove very out of range morphologies from \mathbf{X}_{Arr} cumulation matrix. The decision law $\delta_{Arr}(j)$ is "reject" if $\sigma_{Arr,j}$ is greater than τ_σ and is "hold" if $\sigma_{Arr,j}$ is less than τ_σ , i. e.,

$$\delta_{Arr}(j) = \begin{cases} \text{Reject} & \sigma_{Arr,j} \geq \tau_\sigma \\ \text{Hold} & \sigma_{Arr,j} < \tau_\sigma \end{cases} \quad (6)$$

In Fig. 3, \mathbf{Med}_{Arr} waveforms and cumulated beats given Arr=APB, LBBB, PVC, RBBB obtained from arbitrary subjects of MITDB are shown simultaneously.

The admissible range of tuning parameter α was obtained empirically after performing numerous simulations and manual checking ups and it was seen that when α is chosen less than 3.5, some determinable morphologies are rejected, however, on the other hand, when α is chosen greater than 5.5, uttermost morphologies even with undeterminable shape are held and rejection criterion will be superfluous. The appropriate value of α can be interpreted as confidence interval (CI) for normal (Gaussian) distribution of cardiac events. More details can be seen in [6]. In Fig. 4, curve of σ_{Arr} and calculated threshold for all four under study classes are depicted. Also in table 1, the total number of beats of each class and number of rejected beats using the aforementioned criterion are shown.

2.6 Estimation of P-Wave, QRS Complex & ST-T Segment

After rejection of undeterminable morphologies from cumulation matrix \mathbf{X}_{Arr} , for each beat, ranges of atrial depolarization (P-wave if exists), ventricular depolarization (QRS complex if be normal) and ventricular repolarization (T-wave if exists) are appropriately estimated. In this study, it is assumed that all ECG signals have been registered in rest state of each subject and therefore duration of each complex can be approximately assumed to be about 180 msec.. Namely, by considering each R-peak as a local origin, 90 msec. before and after of origin can be assumed as QRS onset and offset respectively. Also to incorporate P and T waves, about 100 msec. before from QRS onset and 100 msec. after of QRS offset can be assumed as P-wave and ST-T segment estimation, respectively [6]. It should be noted that if R-peak location of a QRS complex is assumed as a reference, in the rest state, the heart rate is usually between 70 to 90 beats/min therefore, duration of a complete heart beat will be between 620 msec. to 1000 msec. [6]. In this situation because of the fact that duration of each QRS complex is between 180~350 msec. (due to the refractory phase duration of myocytes [1]), the remained time that should be assigned to P and T-waves in the worst case is about 270 msec. therefore coming back and going forward about 100 msec. from each R-peak, will certainly cover some samples of P-wave and T-wave of the corresponding heart-beat, respectively. This simple logic is also applied in the following works of Martinez et al. [62], Ghaffari et al.[63] and Arif et al. [32].

In Fig. 5, for each class of arrhythmias, each cumulated segmented beats are shown. In this figure, using black, red and blue colors, estimation of P-wave, QRS complex and ST-T segment are specified respectively.

3. DWT Scales $2^3, 2^4$ & Smoothing Function 2^2

In this step, using à trous discrete wavelet transform the main basic information of feature extraction step is

generated. Generally, it can be stated that the wavelet transform is a quasi-convolution of the hypothetical signal $x(t)$ and the wavelet function $\psi(t)$ with the

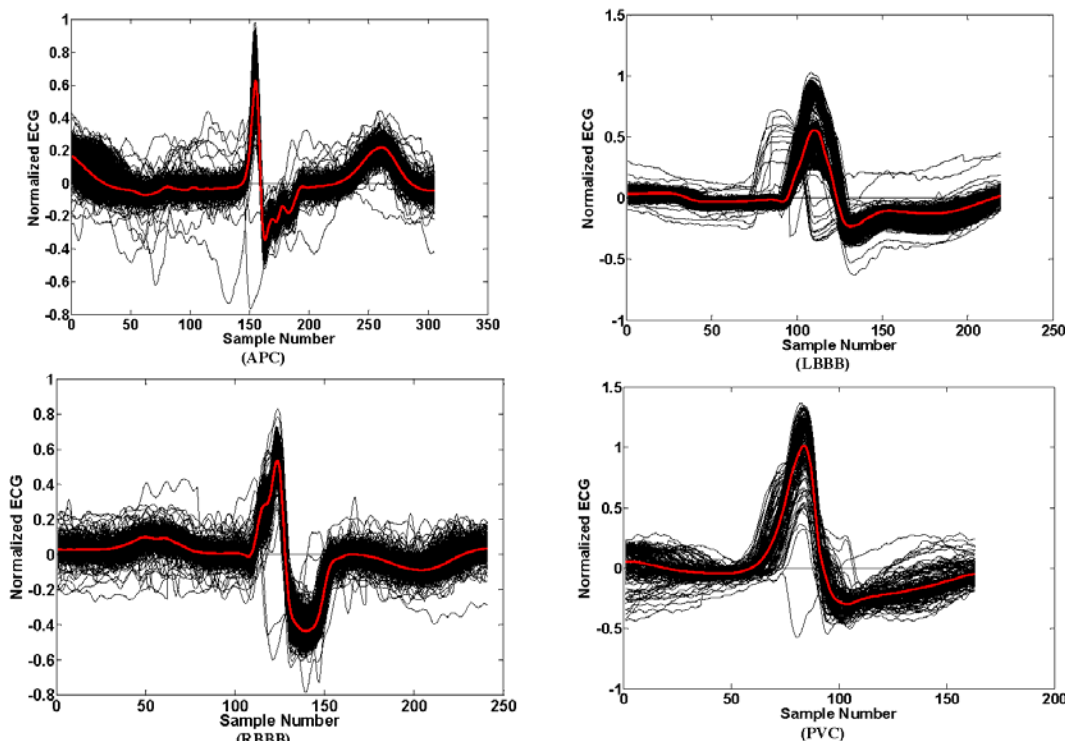


Fig. 3. Cumulated beats and calculated mean template waveform for arbitrary records of the MITDB including APB, LBBB, PVC and RBBB classes.

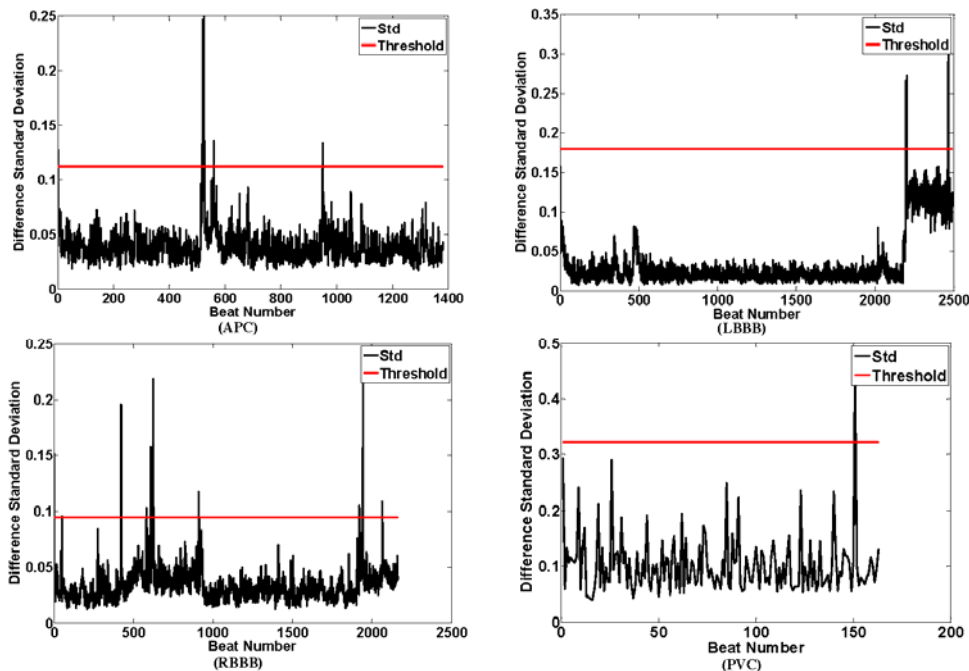


Fig. 4. Standard deviation of difference between each morphology and corresponding template waveform and calculated rejection threshold for arbitrary records of the MITDB including APB, LBBB, PVC and RBBB classes.

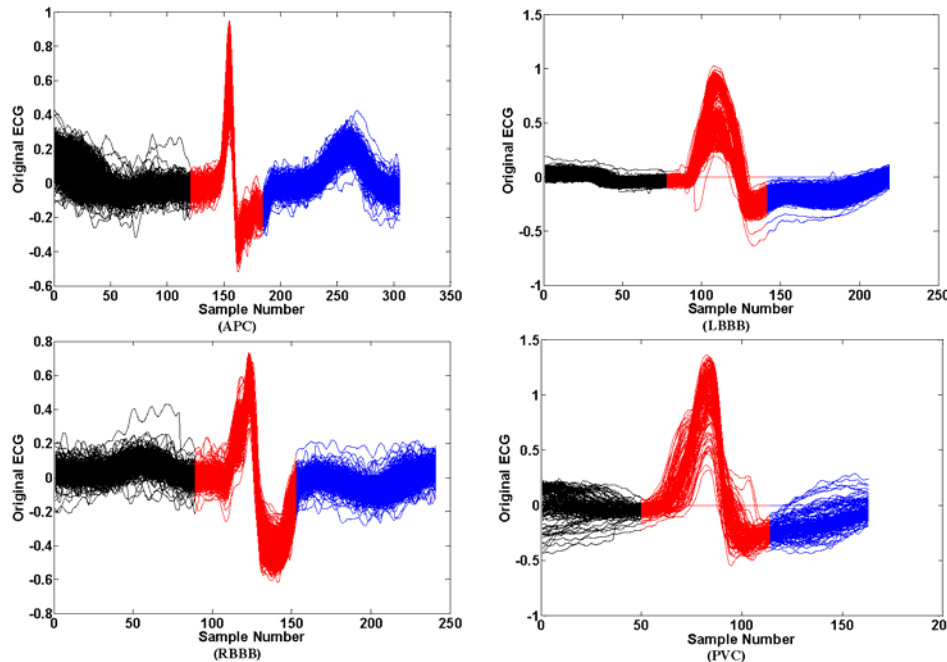


Fig. 5 Segmented beats into atrial depolarization (black), ventricular depolarization (red) and ventricular repolarization (blue) samples for arbitrary records of the MITDB including APB, LBBB, PVC and RBBB classes.

dilation parameter a and translation parameter b , as follows

$$W_{a^x}(b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} x(t) \psi((t-b)/a) dt, \quad a > 0 \quad (7)$$

The parameter a can be used to adjust the wideness of the basis function and therefore the transform can be adjusted in temporal resolutions. Suppose that the function $Y_{a^x}(b)$ is obtained based on a quasi-convolution of signal $x(t)$ and function $\theta(t)$, as follows

$$Y_{a^x}(b) = \int_{-\infty}^{+\infty} x(t) \theta((t-b)/a) dt \quad (8)$$

If the derivative of $Y_{a^x}(b)$ is calculated relative to b , then

$$\frac{\partial Y_{a^x}(b)}{\partial b} = -\frac{1}{a} \int_{-\infty}^{+\infty} x(t) \theta'((t-b)/a) dt \quad (9)$$

On the other hand, if $\psi(t)$ is the derivative of a smoothing function $\theta(t)$, i.e. $\psi(t) = \theta'(t)$, then

$$W_{a^x}(b) = -\frac{1}{\sqrt{a}} \frac{\partial Y_{a^x}(b)}{\partial b} \quad (10)$$

Accordingly, it can be concluded that wavelet transform at the scale a is proportional to the quasi-convolution derivative of the signal $x(t)$ and the smoothing function $\theta(t)$. Therefore, if wavelet transform of the signal crosses of zero, it will be an indicative of local extremum(s) existence in the

smoothed signal and the absolute maximum value of the wavelet transform in different scales represents a maximum slope in the filtered signal. Thus, useful information can be obtained using wavelet transform in different scales. If the scale factor a and the translation parameter b are considered as $a = 2^k$ and $b = 2^k l$, the dyadic wavelet with the following basis function will be resulted [62],

$$\psi_{k,l}(t) = 2^{-k/2} \psi(2^{-k}t - l); \quad k, l \in \mathbb{Z}^+ \quad (11)$$

To implement the à trous wavelet transform algorithm, filters $H(z)$ and $G(z)$ should be used according to the block diagram represented in Fig. 6, [63]. According to this block diagram, each smoothing function (SMF) is obtained by sequential low-pass filtering (convolving with $H(z^{2^k})$ filters), while after high-pass filtering of a SMF (convolving with $G(z^{2^k})$ filters), the corresponding DWT at appropriate scale is generated.

For a prototype wavelet $\psi(t)$ with the following quadratic spline Fourier transform,

$$\Psi(\Omega) = j \Omega \left(\frac{\sin(\Omega/4)}{\Omega/4} \right)^4 \quad (12)$$

the transfer functions $H(z)$ and $G(z)$ can be obtained from the following equation [63],

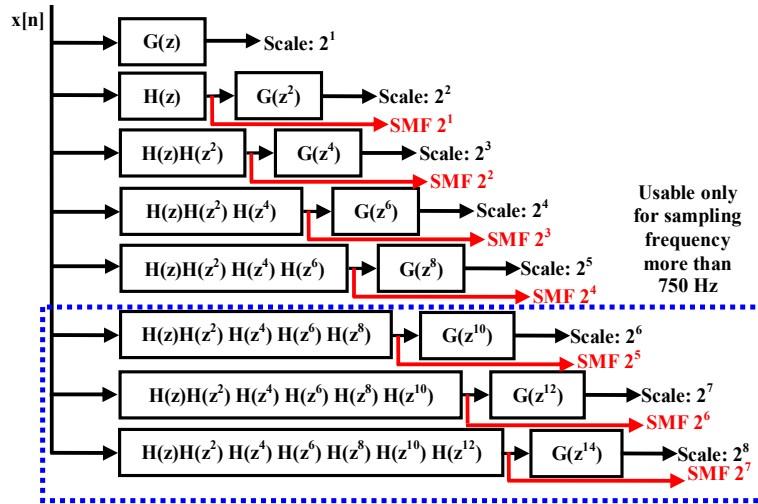


Fig. 6 FIR filter-bank implementation to generate discrete wavelet transform based on à trous algorithm.

$$H(e^{j\omega}) = e^{j\omega/2} (\cos(\omega/2))^3 \quad (13)$$

$$G(e^{j\omega}) = 4je^{j\omega/2} (\sin(\omega/2))$$

and therefore,

$$h[n] = (1/8)\{\delta[n+2] + 3\delta[n+1] + 3\delta[n] + \delta[n-1]\} \quad (14)$$

$$g[n] = 2\{\delta[n+1] - \delta[n]\}$$

It should be noted that for frequency contents of up to 50 Hz, the à trous algorithm can be used in different sampling frequencies. Therefore, one of the most prominent advantages of the à trous algorithm is the approximate independency of its results from sampling frequency. This is because of the main frequency contents of the ECG signal concentrate on the range less than 20 Hz [62, 63]. After examination of various databases with different sampling frequencies (range between 136 to 10 kHz), it has been concluded that in low sampling frequencies (less than 750 Hz), scales 2^λ ($\lambda=1,2,\dots,5$) are usable while for sampling frequencies more than 750 Hz, scales 2^λ ($\lambda=1,2,\dots,8$) contain profitable information that can be used for the purpose of wave detection, delineation and classification, [63].

In the presented study, scaled ECG, smoothing function scale 2^2 , DWT in scales 2^3 and 2^4 are used in feature extraction stage. In Fig. 6 DWT scales 2^3 and 2^4 and smoothing function scale 2^2 are shown obtained from arbitrary MITDB subjects. In this figure corresponding DWT and smoothing function of each segment of beats are illustrated in different colors.

4. Feature Extraction via Newly Proposed CLM

4.1. Implementation Preliminaries

Curve Length Method (CLM) is a new feature extraction (calculation) technique that is newly developed by authors of this work. From mathematical point of view, if $y(x)$ indicates a continuous and differentiable function in the interval $[a,b]$, then the

curve length of the function $y(x)$ in the interval $[a, b]$ can be obtained as

$$Curve(y(x), a, b) = \int_a^b \sqrt{1 + (y'(x))^2} dx$$

$$y'(x) = \frac{d y(x)}{d x} \quad (15)$$

For example, if $y(x) = \sqrt{r^2 - x^2}$ (which in this case, $y(x)$ represents a half circle as shown in Fig. 7.

Therefore through conducting required calculations it is obtained that $y'(x) = \frac{x}{\sqrt{r^2 - x^2}}$. By substituting the

$y'(x)$ into $Curve(y(x), a, b)$ equation, the following result is obtained

$$Curve(y(x), a, b) = \int_{-r}^{+r} \frac{1}{\sqrt{1 - (x/r)^2}} dx \quad (16)$$

$$= r \text{Arc sin}(x/r) \Big|_{-r}^{+r} = r(\pi/2 + \pi/2) = \pi r$$

The principle idea of this study focuses on the definition of a metric containing large amount of information about signal geometric properties. To this end, curve length of DWT in scales 2^3 and 2^4 , smoothing function scale 2^2 and original scaled ECG for P, QRS and T segments are used as components of feature vector. Suppose that $\mathbf{Y}_{N \times 1}$ is a vector with $y(t)$, ($t=1, 2, \dots, N$) components and $\dot{y}(t)$ indicate components of $d\mathbf{Y}/dt$ vector, so it can be shown that the curve length of vector $\mathbf{Y}_{N \times 1}$ is determined as follows

$$CL(\mathbf{Y}) = \int_{t_0}^{t_1} \sqrt{1 + \dot{y}^2(t)} dt \quad (17)$$

where $\dot{y}(t) = (y(t) - y(t-1))F_s$, F_s is sampling frequency and t_0 and t_1 segment start and finish time, respectively. Therefore, dimension of feature space is 12 i.e., 4 (curve length of original scaled ECG + curve length of DWTs in scales 2^3 and 2^4 + curve length of

smoothing function scale $2^2) \times 3$ (P, QRS and T segments) = 12. In study [63], it has been shown that the curve length is an informative metric in which some important geometric parameters such as maximum to minimum ratio, area under curve, mean, standard deviation, summation of point to point difference and some other statistical parameters simultaneously lie on and therefore this feature selection method may be a good solution in the area of arrhythmia recognition.

After conducting a complete search over throughout of the MITDB, according to table 1, 2526, 8031, 6987, and 7007 beats are obtained for APB, LBBB, PVC, and RBBB classes respectively, (some arbitrary feature sources of these classes are shown in Fig. 8). Then, using the CLM feature vector of all beats of each class are calculated and are shown in Fig. 9. According to this figure, for each class some branches can be seen and in this study to assign a code to each branch (Ind_{Arr} , Arr=APB, LBBB, PVC, RBBB) the integer part of the first component of each feature vector is used as its branch identification code. Accordingly, for APB $51 \leq Ind_{APB} \leq 119$, LBBB $76 \leq Ind_{LBBB} \leq 97$, for PVC $78 \leq Ind_{PVC} \leq 136$ and for RBBB $67 \leq Ind_{RBBB} \leq 136$.

4.2 Important Characteristics of Feature Obtained from CLM

Accurate studies performed on all obtained features from all classes show that features obtained via CLM possess two important characteristics:

- CLM is many to one mapping. Loosely speaking, features of each MITDB record are mapped only to one branch of corresponding class and there is no record that features belong to more than one branch of the corresponding class.
- In each class, branches do not intersect each other. Therefore the CLM divide each class into separate regions.

We called two above important properties as *a priori* knowledge resulted from application of CLM.

Table 1 Total number of beats of each APB, LBBB, PVC and RBBB classes in MITDB with number of out of range rejected beats.

Rejection Number	Total Number	Arrhythmia Class
20	2546	APB
44	8075	LBBB
48	7129	PVC
52	7259	RBBB
666	107,090	Total

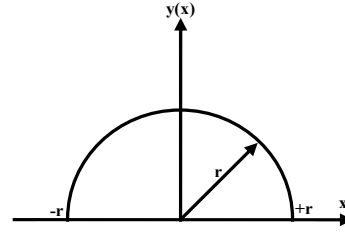


Fig. 7 The half circle used to calculate the curve length of it via Eq. (15).

5. Reduction of Feature Space Dimension to Optimize Computational Burden of the Classifier

5.1. Linear Discriminant Analysis (LDA)

LDA is one of the feature dimensionality reduction technique in which to achieve maximum extent of features separability, the ratio of between-class variance to the within-class variance in an understudy dataset is maximized. As a comparison, after application of PCA transformation, the formation (shape and location) of feature space changes whilst this change does not happen by application of LDA. In the following paragraphs, a brief mathematical description for illustration of LDA structure is presented.

Suppose that $\tau_{FG} = \{(\mathbf{f}_1, g_1), (\mathbf{f}_2, g_2), \dots, (\mathbf{f}_p, g_p)\}$ with feature vectors $\mathbf{f}_i \in \mathbf{R}^n$, and $g \in G = \{1, 2, \dots, q\}$ which is label set for training feature vectors. The within-class \mathbf{S}_w and between class \mathbf{S}_b scatter matrices can be obtained as

$$\mathbf{S}_w = \sum_{g \in G} \mathbf{S}_g, \quad \mathbf{S}_g = \sum_{i \in U_g} (\mathbf{f}_i - \boldsymbol{\mu}_i)(\mathbf{f}_i - \boldsymbol{\mu}_i)^T \quad (18)$$

$$\mathbf{S}_b = \sum_{g \in G} n_g (\boldsymbol{\mu}_g - \boldsymbol{\mu})(\boldsymbol{\mu}_g - \boldsymbol{\mu})^T$$

where U_g consists of training features attributed to the new feature space and n_g is the number of training features in the subsets of U_g . $\boldsymbol{\mu}$ and $\boldsymbol{\mu}_g$ vectors can be obtained from the following summations as

$$\boldsymbol{\mu} = \frac{1}{p} \sum_{i=1}^p \mathbf{f}_i \quad ; \quad \boldsymbol{\mu}_g = \frac{1}{n_g} \sum_{i \in U_g} \mathbf{f}_i \quad (19)$$

If a linear transformation is defined to project the features to a new feature space as

$$\mathbf{z} = \boldsymbol{\Omega}^T \mathbf{f} \quad (20)$$

Then, LDA is obtained by maximizing the below separability criterion which is a function of weight matrix $\boldsymbol{\Omega}$

$$\mathcal{E}(\boldsymbol{\Omega}) = \frac{\det(\tilde{\mathbf{S}}_b)}{\det(\tilde{\mathbf{S}}_w)} \quad (21)$$

where $\tilde{\mathbf{S}}_b$ and $\tilde{\mathbf{S}}_w$ are between-class and within-class matrices respectively obtained for the transformed data \mathbf{z} . In order to maximize the Eq. (21) to obtain the optimal weight matrix $\boldsymbol{\Omega}$, several methods based on generalized singular value decomposition (GSVD) have been proposed by authors and would be seen in [70].

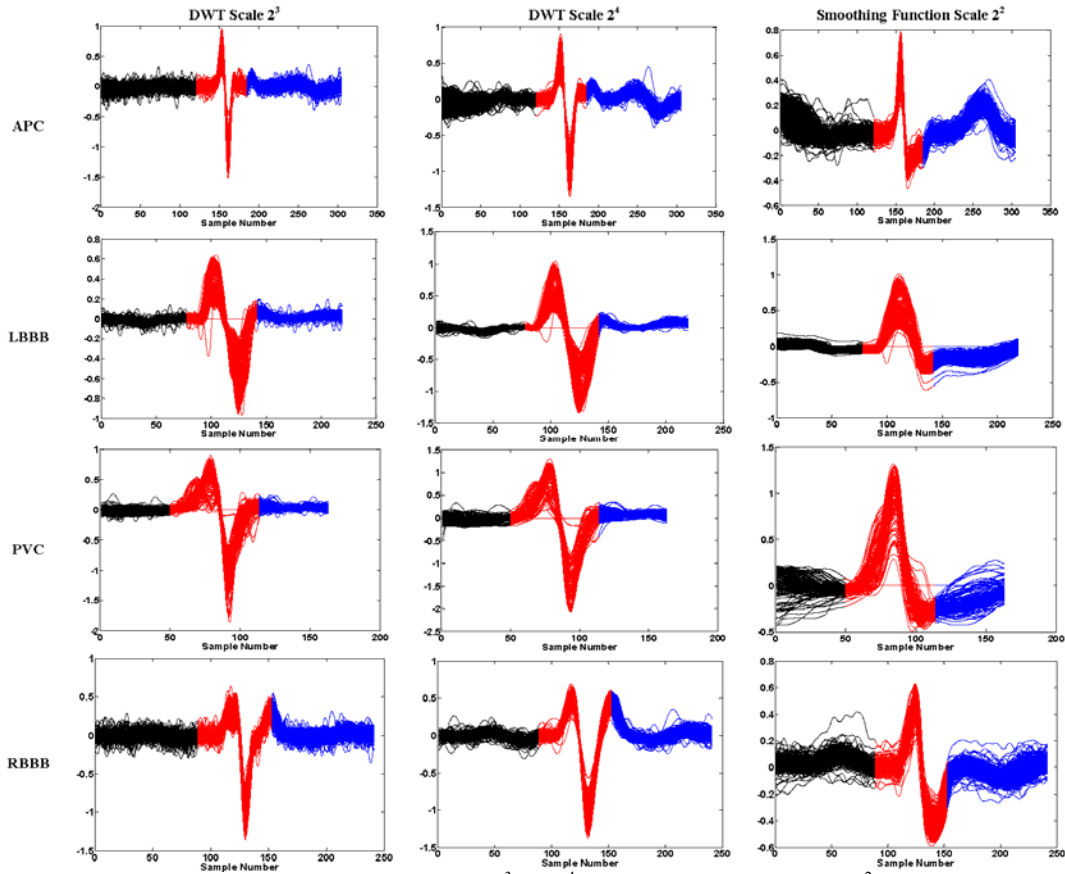


Fig. 8 Discrete wavelet transform of the scaled ECG in scales 2^3 and 2^4 including smoothing function 2^2 for arbitrary records of the MITDB including APB, LBBB, PVC and RBBB classes.

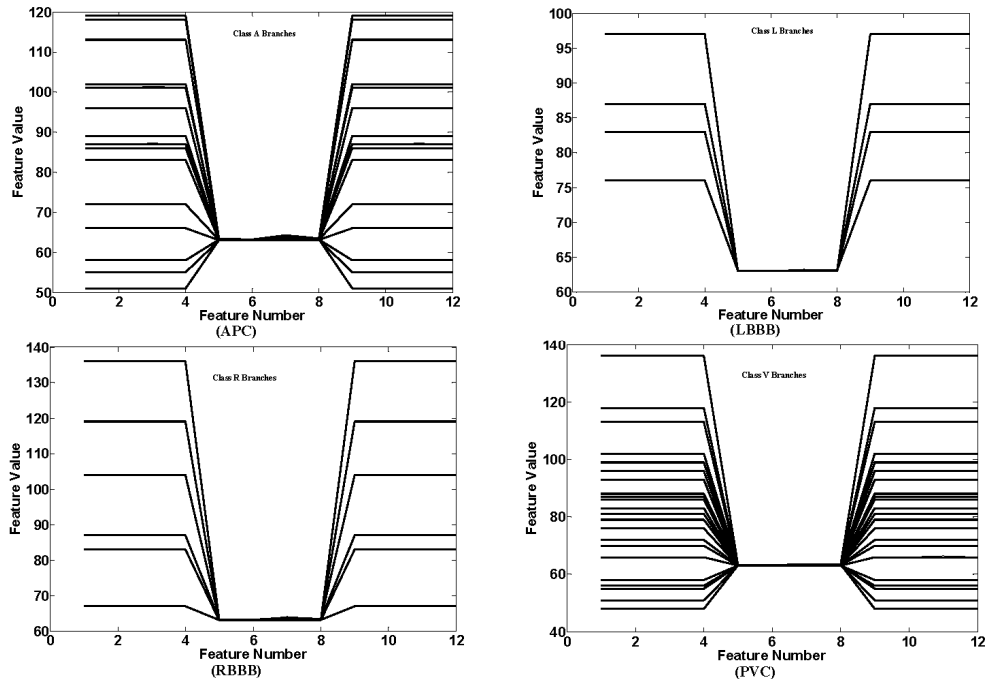


Fig.9 Features extracted via CLM for all APB, LBBB, PVC and RBBB classes of the MITDB.

5.2 Generalized Discriminant Analysis (GDA)

The GDA is the nonlinear version of the conventional LDA and in this case $\tau_{FG} = \{(\mathbf{f}_1, g_1), (\mathbf{f}_2, g_2), \dots, (\mathbf{f}_p, g_p)\}$ with feature vectors $\mathbf{f}_i \in \mathbf{R}^n$, $g \in G = \{1, 2, \dots, q\}$. In GDA projection, the input training set is mapped from feature space F to a high-dimensional feature space H. Then the conventional LDA is applied on the mapped data. During this transformation, the features of the space F are projected by a kernel function $k(\cdot)$ and bias term \mathbf{b} as follows

$$\mathbf{z} = \mathbf{\Omega}^T k(\mathbf{f}) + \mathbf{b} \quad (22)$$

where $\mathbf{\Omega}$ is the weight matrix and \mathbf{b} is bias vector which both of them should be determined via an optimization approach. It should be noticed that the optimum parameters $(\mathbf{\Omega}, \mathbf{b})$ can be calculated by maximizing of the difference between-class and within-class scatter of the $\tau_{ZG} = \{(\mathbf{z}_1, g_1), (\mathbf{z}_2, g_2), \dots, (\mathbf{z}_p, g_p)\}$ and $\mathbf{z}_i \in \mathbf{R}^m$. More details about the methods to solve this optimization problem can be seen in [70].

5.3 Principal Component Analysis (PCA)

Suppose that $\mathbf{F}_{Arr} = [\mathbf{f}_1, \dots, \mathbf{f}_N]$ is a matrix includes feature vector $(\mathbf{f}_j)_{12 \times 1}$ $j = 1, \dots, N$ and the aim is to reduce dimension of \mathbf{F}_{Arr} and reconstruct a matrix $\mathbf{G}_{Arr} = [\mathbf{g}_1, \dots, \mathbf{g}_N]$ in which $(\mathbf{g}_j)_{p \times 1}$ $j = 1, \dots, N$ and $p \leq 12$. Using a linear orthonormal projection as follows

$$\mathbf{g} = \mathbf{\Omega}^T \mathbf{f} + \mathbf{u} \quad (23)$$

where $\mathbf{\Omega}_{12 \times p}$ is weight matrix and $\mathbf{u}_{p \times 1}$ is offset vector. By application of linear back projection, the estimation of vector \mathbf{f} i.e., $\hat{\mathbf{f}}$ is obtained from following equation

$$\hat{\mathbf{f}} = \mathbf{\Omega}(\mathbf{g} - \mathbf{u}) \quad (24)$$

and $(\mathbf{\Omega}, \mathbf{u})$ is the optimal solution of the following constrained optimization equation

$$(\mathbf{\Omega}, \mathbf{u}) = \underset{(\mathbf{\Omega}_0, \mathbf{u}_0)}{\text{Arg min}} [\mathcal{E}(\mathbf{\Omega}_0, \mathbf{u}_0)] \quad (25)$$

$$\text{s.t. } \mathbf{\Omega}^T \mathbf{\Omega} = \mathbf{I}$$

where $\mathbf{\Omega}^T \mathbf{\Omega} = \mathbf{I}$ condition imposes orthonormality constraint to optimization problem. To solve the Eq. (26), several methods such as Cholesky decomposition and linear algebra methods can be found, [70]. The main role of PCA application is to reduce feature space dimension so train time and computational burden decreases. However, for low dimension feature vectors such as this study, implementation of PCA doesn't make significant improvement in train accuracy and computation load decrement.

6 Adaptive Network Fuzzy Inference System (ANFIS) Heart Arrhythmia Classifier Design

For Computer Implementation of classification algorithm, an adaptive network-based fuzzy inference system (ANFIS) is trained using the information obtained from the section 4. To consider the mutual influence of features in the evaluation of error probability, an ANFIS network is implemented. Using this network, it will be possible to incorporate the possible mutual influences between risk features to the classification of arrhythmias. Adaptive neuro-fuzzy networks are actually fuzzy inference systems which are translated to a neural network language and using the neural network features, the actual system is trained based on the input-output data and then turned into a fuzzy model. More details about ANFIS can be seen at [71].

According to Fig. 10, inputs are elements of feature vector included in the first layer and incorporated to fuzzy membership functions. The outputs of neurons are the membership values of inputs in fuzzy values. This part of the neuro-fuzzy network is equivalent to the IF part of the fuzzy inference system. In the Input Rules Layer, resulted membership values are passed through a T-norm and the firing strength of each rule is determined. In the output membership layer, using the final relationships, the output of each IF-THEN rule is determined. For each of the feature vector elements, two Gaussian memberships are considered with a Sugeno Difuzzification algorithm. It should be mentioned that, triangular memberships can be selected instead of Gaussian memberships, by this choice some performance quality will be dedicated as the expense of decreasing the computational burden. The structure of this network is illustrated in Fig. 10.

7 Results and Discussions

7.1 Evaluation and of the ANFIS-Based Arrhythmia Classification Algorithm

In this section, evaluation process of the proposed classification algorithm is demonstrated. According to section 4, by application of CLM to feature extraction, in each class of under study arrhythmias, some branches are seen which for APB $51 \leq \text{Ind}_{APB} \leq 119$ (total number of 2526), LBBB $76 \leq \text{Ind}_{LBBB} \leq 97$ (total number of 8031), for PVC $78 \leq \text{Ind}_{PVC} \leq 136$ (total number of 7081) and for RBBB $67 \leq \text{Ind}_{RBBB} \leq 136$ (total number of 7207) were previously specified. Validation of the presented classification algorithm is fulfilled in four steps.

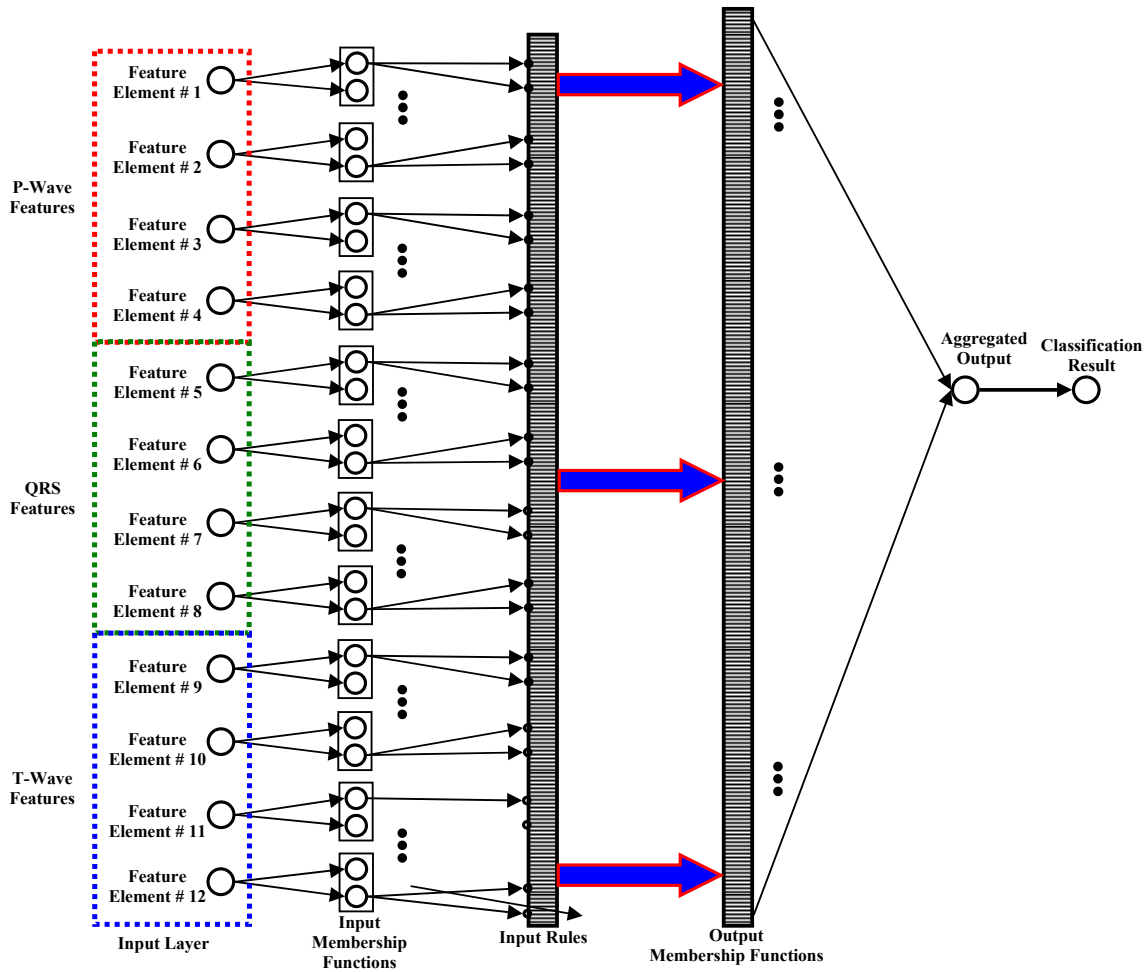


Fig. 10 The ANFIS structure used to classify four LBBB, RBBB, PVC and APB arrhythmia classes.

7.2 Justification of Features Selection

In this study, to determine the optimum number of DWT scales and the corresponding SMF that should be incorporated in the feature vector to reach the acceptable accuracy, first, an examination group possessing 100 beats excerpted from each of APB, LBBB, RBBB, and PVC groups is provided and 50% of this group is used to train the classifier while the rest 50% is used for the classification performance testing. In this study the examination of the effect of the increasing number of included DWTs and SMFs on the classification algorithm accuracy is performed via six steps:

- Feature vector including DWTs and SMFs of scales 2^1 and 2^2 (3×4 feature vector elements).
- Feature vector including DWTs and SMFs of scales 2^1 and 2^3 (3×4 feature vector elements).
- Feature vector including DWTs and SMFs of scales 2^1 and 2^4 (3×4 feature vector elements).

- Feature vector including DWTs and SMFs of scales 2^2 and 2^3 (3×4 feature vector elements).
- Feature vector including DWTs and SMFs of scales 2^2 and 2^4 (3×4 feature vector elements).
- Feature vector including DWTs and SMFs of scales 2^3 and 2^4 (3×4 feature vector elements- best accuracy).

In each of aforementioned steps, the accuracy of the classification algorithm is calculated and is compared with its previous step(s). The conducted studies show that by increasing the number of scales from one to four, the accuracy of the classification algorithm increases monotonically (Fig. 11). However, by adding the features with the scale orders more than three, almost no improvement can be achieved. Therefore, the fifth step of the above testing process, can be assumed as the optimum bound for the DWT and SMF feature order selection for which the best trade-off between algorithm accuracy and dimension of the feature vector (correspondingly the computational burden and run-

time) can be achieved. In Fig. 11, the performance accuracy of an ANFIS classifier versus incorporated scales and SMFs is shown. As, it can be seen, the accuracy of the classification algorithm saturates by passing the DWT scales from three.

7.3 Analysis of the Proposed Arrhythmia Classification Performance by Implementation Feature Space Dimension Reductions Algorithms

In order to performance assessment of the classification algorithm along with the application of feature space dimension reduction methodologies like PCA, LAD and GDA, using the aforementioned dimensionality reduction algorithms, the dimension of feature space is reduced one by one, and in each case the performance of the classifier on a test database defined in section 7.2 is evaluated. The general procedure of the performance evaluation of the classification algorithm in the presence of dimension reduction methods is shown in Fig. 12.

The conducted studies show that by reduction of feature space dimension one by one via PCA algorithm, accuracy of the trained ANFIS classifier decreases with a rather high rate where by reaching the dimension of the feature space to the number 6, the accuracy of the classification algorithm drops to 65% and lower showing a partial failure in the classification performance. On the other hand, by application of the LDA or GDA (with Radial Basis Function-RBF kernel), the accuracy of the algorithm drops rather smoothly relative to the PCA-based classification algorithm screening higher stability (robustness) that by reaching of the feature dimension to the number 6, accuracy of the classification algorithm attains to 81% which is a better record rather than the accuracy of the PCA-based ANFIS classifier. A graphical representation of this investigation is shown in Fig. 13. It should be noted that because of the fact that the feature dimension being equal to 12 obtained in this study is not a high dimensionality to impose a heavy computational burden and cost, therefore logically, the application of feature reduction algorithms such as PCA, LDA or GDA can be ignored. Exploitation of dimension reduction algorithms should be involved especially in the cases that feature space dimension exceeds than 50 or more which in such situations, the computational burden (and therefore train and test run-time) will be high and may cause very long or even impossible responding of the computerized algorithm.

7.4 Case That in Which Train and Test Features Are Chosen Only From One Branch of Each Arrhythmia Class

In this case, only a specific branch from each arrhythmia class is considered and all corresponding

features are extracted. Then using stochastic Roulette Wheel shuffling algorithm [68], obtained feature vectors are shuffled completely and 40% of feature vectors are used for training and the rest are used for testing. In this study, we have examined more than 2500 cases and to show performance of the classification algorithm, 15 cases are randomly chosen and obtained results are shown in table 2. As it can be seen, due to a priori knowledge obtained from CLM, the accuracy level of the classification algorithm is very close to 100%.

7.5 Case That in Which Train and Test Features Are Chosen From More Than One Branch of Each Arrhythmia Class

In this experiment simulation, from each arrhythmia class, more than one branch (3, 4 or 5 branches) are chosen and the corresponding features are extracted.

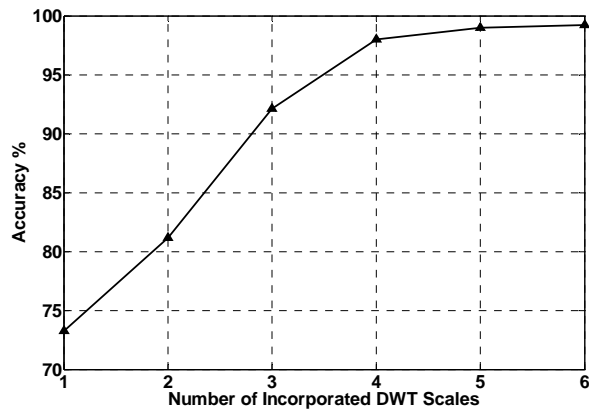


Fig. 11 The accuracy variation of a trial ANFIS versus increasing the number of incorporated higher order moments for the Group of test arrhythmias.

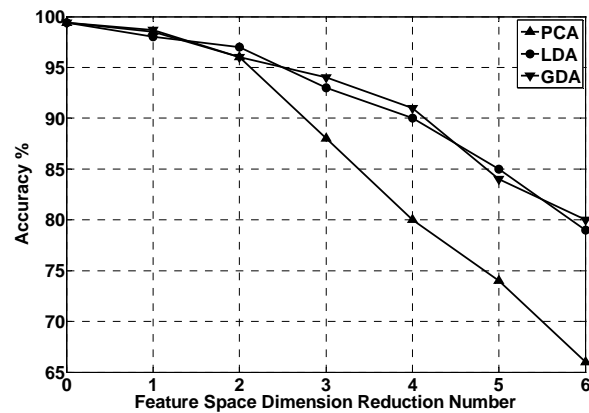


Fig. 13 Variations of arrhythmia classification performance in the presence of feature space dimensionality reduction by PCA, LDA and GDA.

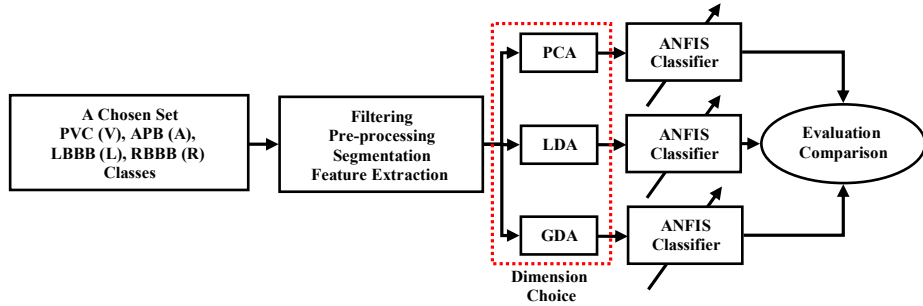


Fig. 12 The block diagram of the arrhythmia classification performance assessment versus reduction of feature space dimension via PCA, LDA and GDA.

Table 2 Statistical parameters of classification algorithm performance obtained during first test (section 7.4).

Accuracy (%)	# of Error	# of Tested Feature Vectors	# of Corresponding Feature Vectors	Branches Code [Ind _{Arr}] Arr=APB,LBBB,PVC,RBBB
100	0	7696	[26,2038,1219,285]	[66,97,104,70]
100	0	5782	[371,2038,1468,155]	[58,87,136,48]
100	0	2891	[43,1916,382,200]	[113,87,119,51]
100	0	6017	[29,1418,1219,427]	[102,83,104,56]
100	0	8403	[26,2257,2109,380]	[66,76,87,79]
100	0	4885	[88,1418,382,774]	[101,83,119,66]
100	0	5180	[105,2257,2109774]	[83,76,87,66]
100	0	5369	[105,2308,80,194]	[83,97,83,102]
100	0	7580	[266,1916,382,811]	[72,87,119,55]
100	0	7112	[29,1418,1219,427]	[102,83,104,56]
100	0	8746	[29,6772,1493,382,40]	[102,76,119,118]
100	0	7106	[88,2038,382,103]	[101,97,119,81]
99.99	1	7137	[48,2038,382,194]	[89,97,119,102]
100	0	4723	[371,2257,1219,103]	[58,76,104,81]
100	0	3818	[93,2038,1219,103]	[87,97,104,81]

Final Result of Stage (section 7.4)								
Acc (%)	Sp (%)	P+ (%)	Se (%)	RBBB	PVC	LBBB	APB	
100	100	100	100	0	0	0	1030	ABP
100	100	100	100	0	0	20476	0	LBBB
100	100	100	100	0	9158	0	0	PVC
99.97	99.99	100	99.97	3187	0	0	0	RBBB

Then the obtained feature vectors are shuffled randomly and 40% of the resulted dataset is used as train set and rest of dataset as test set. In this case, 450 different arrangements are examined and 5 of randomly chosen cases are shown in table 3. In this stage, the mean values of Se=99.80%, P+=99.79%, Sp=99.84% and Acc=99.81% were obtained for all 450 examinations.

7.6. Case That in Which Train and Test Features Are Chosen From All Branches of Each Arrhythmia Class

In this stage of evaluation process, all branches of each arrhythmia are used to evaluate generalization power of the algorithm. According to table 1, there are 2526, 8031, 6987, 7207 feature vectors for APB, LBBB, PVC, and RBBB categories, respectively. First, features

of each arrhythmia class are shuffled in a random fashion and for the purpose of training, 400, 600, 800, 1000 and 1200 feature vectors are chosen from each class. Then 800 feature vectors which don't belong to train set are chosen randomly as test features from main database and for each above training sets, test evaluation is performed. The results obtained in this stage are shown in table 4. Obtained results came up us with the conclusion that when number of subjects incorporated into train and test sets increases (i.e., variety of morphology increases), accuracy of the classification algorithm saturates and cannot go up a threshold. According to table 4, the final accuracy value Acc=99.72% has been obtained using CLM feature extraction and implementation of ANFIS classifier.

Table 3 Statistical parameters of classification algorithm performance obtained during second test (section 7.5).

Accuracy (%) (Mean Value)	# of Error	# of Tested Feature Vectors	# of Corresponding Feature Vectors	Range of Branches Codes [Ind _{Arr}]
99.68	19	25643	[1510,8031,2532,7207]	$76 \leq \text{Ind}_{LBBB} \leq 97, 67 \leq \text{Ind}_{RBBB} \leq 136, 80 \leq \text{Ind}_{APB} \leq 119, 40 \leq \text{Ind}_{PVC} \leq 60$
99.71	23	30445	[560,8031,2532,7207]	$76 \leq \text{Ind}_{LBBB} \leq 97, 67 \leq \text{Ind}_{RBBB} \leq 136, 60 \leq \text{Ind}_{APB} \leq 80, 40 \leq \text{Ind}_{PVC} \leq 60$
99.74	21	30523	[691,8031,2532,7207]	$76 \leq \text{Ind}_{LBBB} \leq 97, 67 \leq \text{Ind}_{RBBB} \leq 136, 51 \leq \text{Ind}_{APB} \leq 60, 40 \leq \text{Ind}_{PVC} \leq 60$
99.68	25	42211	[691,8031,2012,7207]	$76 \leq \text{Ind}_{LBBB} \leq 97, 67 \leq \text{Ind}_{RBBB} \leq 136, 51 \leq \text{Ind}_{APB} \leq 60, 99 \leq \text{Ind}_{PVC} \leq 136$
99.77	11	23586	[691,8031,2012,7207]	$76 \leq \text{Ind}_{LBBB} \leq 97, 67 \leq \text{Ind}_{RBBB} \leq 136, 51 \leq \text{Ind}_{APB} \leq 60, 99 \leq \text{Ind}_{PVC} \leq 136$

Final Result of Stage (section 7.5)								
Acc (%)	Sp (%)	P+ (%)	Se (%)	RBBB	PVC	LBBB	APB	
99.32	99.99	99.64	99.68	0	2	0	2478	ABP
99.84	99.98	99.90	99.94	4	0	24079	0	LBBB
99.35	99.97	99.54	99.81	5	6929	9	4	PVC
99.89	99.99	99.96	99.94	21607	5	0	0	RBBB

Table 4 Statistical parameters of classification algorithm performance obtained during third test (section 7.6).

Accuracy (%) (Mean Value)	# of Error	# of Tested Feature Vectors	# of Corresponding Feature Vectors	Range of Branches Codes [Ind _{Arr}] (All Branches)
99.35	10	800	[400,400,400,400]	$76 \leq \text{Ind}_{LBBB} \leq 97, 67 \leq \text{Ind}_{RBBB} \leq 136, 51 \leq \text{Ind}_{APB} \leq 119, 78 \leq \text{Ind}_{PVC} \leq 136$
99.63	8	800	[600,600,600,600]	$76 \leq \text{Ind}_{LBBB} \leq 97, 67 \leq \text{Ind}_{RBBB} \leq 136, 51 \leq \text{Ind}_{APB} \leq 119, 78 \leq \text{Ind}_{PVC} \leq 136$
99.68	7	800	[800,800,800,800]	$76 \leq \text{Ind}_{LBBB} \leq 97, 67 \leq \text{Ind}_{RBBB} \leq 136, 51 \leq \text{Ind}_{APB} \leq 119, 78 \leq \text{Ind}_{PVC} \leq 136$
99.77	4	800	[1000,1000,1000,1000]	$76 \leq \text{Ind}_{LBBB} \leq 97, 67 \leq \text{Ind}_{RBBB} \leq 136, 51 \leq \text{Ind}_{APB} \leq 119, 78 \leq \text{Ind}_{PVC} \leq 136$
99.78	4	800	[1200,1200,1200,1200]	$76 \leq \text{Ind}_{LBBB} \leq 97, 67 \leq \text{Ind}_{RBBB} \leq 136, 51 \leq \text{Ind}_{APB} \leq 119, 78 \leq \text{Ind}_{PVC} \leq 136$

Final Result of Stage (section 7.6)								
Acc (%)	Sp (%)	P+ (%)	Se (%)	RBBB	PVC	LBBB	APB	
99.73	99.98	99.89	99.83	0	2	0	3993	ABP
99.77	99.97	99.85	99.93	1	2	3997	1	LBBB
99.57	99.95	99.73	99.85	2	3994	2	3	PVC
99.80	99.98	99.93	99.88	3995	1	0	0	RBBB

7.7. Arrhythmia Classification Performance Comparison with Other Works

In the final step, in order to show the marginal performance improvement of the proposed arrhythmia classification algorithm, the method is assessed relative to other high-performance recent works. The result of comparison of the proposed method and other works is shown in table 5.

8 Conclusion

Introducing a new feature extraction method in the area of the heart arrhythmia classification based on a metric with simple mathematical calculation called Curve-Length Method (CLM) is the aim of this study. In the presented method, curve length of the under study excerpted segment of signal is considered as an informative feature in which the effect of important geometric parameters of the original signal such as

Table 5 Comparison of the CLM-based classification algorithm with previous recent works (section 7.7).

Accuracy (%)	Feature Space Dimension	Number of Incorporated Morphologies (Beat Type)	Study
99.70	10	4	This Study (with PCA)
99.76	10	4	This Study (with LDA)
99.76	10	4	This Study (with RBF-GDA)
99.72	12	4	This Study (without PCA)
96.06	18	7	Osowski et al. [18]
97.78	24	5	Guler et al. [60]
96.94	19	4	Guler et al. [61]
98.00	5	3	Minami et al. [57]
97.50	6	4	Al-Fahoum et al. [56]
97.00	15	5	Dokur et al. [67]
96.77	25	12	Prasad et al. [58]
99.51	27	6	Sung-Nien et al. [16]
97.53	30	8	Chen et al. [59]
99.70	30	6	Chen et al. [59]
99.65	11	6	Yu et al. [66]
99.4	11	6	Yu et al. [66]
99.49	11	6	Minhas et al. [32] (without PCA)
99.47	6	6	Minhas et al. [32] (with PCA)
99.40*	9	5	de Chazel et al. [9]
98.50**	15	15	de Chazel et al. [24]

* Maximum accuracy value of this study was adopted.

** Average accuracy value of the method was calculated (application to all subjects of MITDB)

mean, standard deviation, maximum value to minimum value ratio, (absolute value of) area under the signal, summation of point to point difference and events duration can be found. To show merits of strategy, it was applied to classify four categories found in the MIT-BIH Arrhythmia Database namely as Atrial Premature Beat (APB), Left Bundle Branch Block (LBBB), Right Bundle Branch Block (RBBB) and Premature Ventricular Contraction (PVC) via an ANFIS classifier and average values of $Se = 99.81\%$, $P+ = 99.80\%$, $Sp = 99.81\%$ and $Acc = 99.72\%$ are obtained for sensitivity, positive predictivity, specificity and accuracy respectively showing marginal improvement of the heart arrhythmia classification performance. Acceptable performance, low training computational burden and quick responding in test can be mentioned as important merits of the algorithm.

References

- [1] Ghaffari A. and Homaeinejad M. R., "Fading Parameters of Sodium, Potassium and Leakage Ionic Channels the Best Linear Unbiased Sequentially Estimation (BLUE) Via Voltage Clamp Technique Noisy Measurement," *16th Annual (International) Conference on Mechanical Engineering-ISME 2008*, Shahid Bahonar University of Kerman, Iran, 2008.
- [2] Ghaffari A., Atarod M., Homaeinejad M. R., Ahmady Y. and Rahmani R., "Detecting and Quantifying T-wave Alternans Using the Correlation Method and Comparison with the FFT-based Method," *34th Annual Conference of Computers in Cardiology (CinC)*, September 14-17, Bologna, Italy, 2008.
- [3] Ghaffari A., Homaeinezhad M. R., Akraminia M., Atarod M. and Davaecha M., "Detecting and Discriminating Premature Atrial and Ventricular Contractions: Application to Prediction of Paroxysmal Atrial Fibrillation," *35th Annual Conference of Computers in Cardiology (CinC)*, September 13-16, Lake City-Utah, USA, 2009.
- [4] Ghaffari A., Homaeinezhad M. R., Akraminia M., Atarod M. and Davaecha M., "Detecting and Quantifying T-Wave Alternans in Patients with Heart Failure and Non- Ischemic Cardiomyopathy via Modified Spectral Method," *35th Annual Conference of Computers in Cardiology (CinC)*, September 13-16, Lake City-Utah, USA, 2009.
- [5] Ghaffari A., Atarod M., Homaeinejad M. R. and Rahmani R., "On-Line Identification of the Heart Hemodynamic Parameters via the Discrete-Time Kalman-Bucy Filter Using Invasive Noisy Blood Pressure Waveform Observations," *34th Annual Conference of Computers in Cardiology (CinC)*, September 14-17, Bologna, Italy, 2008.
- [6] Ghaffari A., Homaeinezhad M. R., Ahmadi Y. and Rahnavard M., "An Open-Source Computer Model for Visualization of Artificial Abnormal Multi-Lead Electrocardiographic Phenomena," *World Journal of Modelling and Simulation*, In-Press, 2009.
- [7] Tsiouras M. G., Fotiadis D. I. and Sideris D., "An arrhythmia classification system based on the

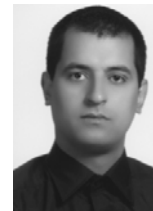
- RR-interval signal," *Artificial Intelligence in Medicine*, Vol. 33, pp. 237-250, 2005.
- [8] Nilsson M., Funk P., Olsson E. M. G., von Scheele B. and Xiong N., "Clinical decision-support for diagnosing stress-related disorders by applying psychophysiological medical knowledge to an instance-based learning system," *Artificial Intelligence in Medicine*, Vol. 36, pp. 159-176, 2006.
- [9] de Chazal P. and Reilly R. B., "A Patient-Adapting Heartbeat Classifier Using ECG Morphology and Heartbeat Interval Features," *IEEE Transactions on Biomed. Eng.*, Vol. 53, No. 12, pp. 2535-2543, Dec. 2006.
- [10] Tsipouras M. G. and Fotiadis D. I., "Automatic arrhythmia detection based on time and time-frequency analysis of heart rate variability," *Computer Methods and Programs in Biomedicine*, Vol. 74, pp. 95-108, 2004.
- [11] Yu Sung-Nien and Chou Kuan-To, "Integration of independent component analysis and neural networks for ECG beat classification," *Expert Systems with Applications*, Vol. 34, pp 2841-2846, 2008.
- [12] Yu S. N. and Chou K. T., "Selection of significant independent components for ECG beat classification," *Expert Systems with Applications*, Vol. 36, pp 2088-2096, 2009.
- [13] Acharya U. R., Sankaranarayanan M., Nayak J., Xiang C. and Tamura T., "Automatic identification of cardiac health using modeling techniques: A comparative study," *Information Sciences*, Vol. 178, pp. 4571-4582, 2008.
- [14] Kannathal N., Lim C. M., Rajendra Acharya U. and Sadasivan P. K., "Cardiac state diagnosis using adaptive neuro-fuzzy technique," *Medical Engineering & Physics*, Vol. 28, pp. 809-815, 2006.
- [15] Mohammadzadeh Asl B., Kamaleddin Setarehdan S. and Mohebbi M., "Support vector machine-based arrhythmia classification using reduced features of heart rate variability signal," *Artificial Intelligence in Medicine*, Vol. 44, pp. 51-64, 2008.
- [16] Yu S. N. and Chou K. T., "A switchable scheme for ECG beat classification based on independent component analysis," *Expert Systems with Applications*, Vol 33, pp.824-829, 2007.
- [17] Ozbay Y., Ceylan R. and Karlik B., "A fuzzy clustering neural network architecture for classification of ECG arrhythmias," *Computers in Biology and Medicine*, Vol. 36, pp. 376-388, 2006.
- [18] Osowski S., Markiewicz T. and Tran Hoai L., "Recognition and classification system of arrhythmia using ensemble of neural networks," *Measurement*, Vol. 41, pp. 610-617, 2008.
- [19] Ceylan R., Uzbay Y. and Karlik B., "A novel approach for classification of ECG arrhythmias: Type-2 fuzzy clustering neural network," *Expert Systems with Applications*, Vol. 36, pp. 6721-6726, 2009.
- [20] Polat K., Sahan S. and Gune S., "A new method to medical diagnosis: Artificial immune recognition system (AIRS) with fuzzy weighted pre-processing and application to ECG arrhythmia," *Expert Systems with Applications*, Vol. 31, pp. 264-269, 2006.
- [21] Lin C. H., Du Y. C. and Chen T., "Adaptive wavelet network for multiple cardiac arrhythmias recognition," *Expert Systems with Applications*, Vol. 34, pp. 2601-2611, 2008.
- [22] Wen C., Lin T. C., Chang K. C. and Huang C. H., "Classification of ECG complexes using self-organizing CMAC," *Measurement*, Vol. 42, pp. 399-407, 2009.
- [23] Ebrahimzadeh A. and Khazaei A., "Detection of Premature Ventricular Contractions Using MLP Neural Networks: A Comparative Study," *Measurement* (2009), DOI: 10.1016/j.measurement, 2009. 07. 002.
- [24] de Chazal P., O'Dwyer M. and Reilly R. B., "Automatic Classification of Heartbeats Using ECG Morphology and Heartbeat Interval Features," *IEEE Transactions on Biomed. Eng.*, Vol. 51, No. 7, pp. 1196-1206, Jul. 2004.
- [25] Inan O. T., Giovangrandi L. and A. Kovacs G. T., "Robust Neural-Network-Based Classification of Premature Ventricular Contractions Using Wavelet Transform and Timing Interval Features," *IEEE Transactions on Biomed. Eng.*, Vol. 53, No. 12, pp. 2507-25-15, Dec. 2006.
- [26] Wiggins M., Saad A., Litt B. and Vachtsevanos G., "Evolving a Bayesian classifier for ECG-based age classification in medical applications," *Applied Soft Computing*, Vol. 8, pp. 599-608, 2008.
- [27] Bartolo A., Clymer B. D., Burgess R. C., Turnbull J. P., Golish J. A. and Perry M. C., "An Arrhythmia Detector and Heart Rate Estimator for Overnight Polysomnography Studies," *IEEE Transactions on Biomed. Eng.*, Vol. 48, No. 5, pp. 513-521, May, 2001.
- [28] Polat K. and Gunes S., "Detection of ECG Arrhythmia using a differential expert system approach based on principal component analysis and least square support vector machine," *Applied Mathematics and Computation*, Vol. 186, pp. 898-906, 2007.
- [29] Yeh Y. C., Wang W. J. and C. W. Chiou, "Cardiac arrhythmia diagnosis method using linear discriminant analysis on ECG signals," *Measurement*, Vol. 42, pp. 778-789, 2009.
- [30] Tsipouras M. G., Voglis C. and Fotiadis D. I., "A Framework for Fuzzy Expert System Creation-

- Application to Cardiovascular Diseases," *IEEE Transactions on Biomed. Eng.*, Vol. 54, No. 11, pp. 2089-2105, Nov., 2007.
- [31] Christov I., Jekova I. and Bortolan G., "Premature ventricular contraction classification by the Kth nearest-neighbours rule," *Physiol. Meas.*, Vol. 26, pp. 123-130, 2005.
- [32] Minhas F. A. and Arif M., "Robust electrocardiogram (ECG) beat classification using discrete wavelet transform," *Physiological Measurement*, Vol. 29, pp. 555-570, 2008.
- [33] Chudacek V., Georgoulas G., Lhotska L., Stylios C., Petrik M. and Cepek M., "Examining cross-database global training to evaluate five different methods for ventricular beat classification," *Physiol. Meas.*, Vol. 30, pp. 661-677, 2009.
- [34] Exarchos T. P., Tsiouras M. G., Exarchos C. P., Papaloukas C., Fotiadis D. I. and Michalis L. K., "A methodology for the automated creation of fuzzy expert systems for ischaemic and arrhythmic beat classification based on a set of rules obtained by a decision tree," *Artificial Intelligence in Medicine*, Vol. 40, pp. 187-200, 2007.
- [35] Christov I. and Bortolan G., "Ranking of pattern recognition parameters for premature ventricular contractions classification by neural networks," *Physiol. Meas.*, Vol. 25, pp. 1281-1290, 2004.
- [36] Polat K., Kara S., Güven A. and Günes S., "Usage of class dependency based feature selection and fuzzy weighted pre-processing methods on classification of macular disease," *Expert Systems with Applications*, Vol. 36, pp. 2584-2591, 2009.
- [37] Liu H., Sun J., Liu L. and Zhang H., "Feature selection with dynamic mutual information," *Pattern Recognition*, Vol. 42, pp. 1330 - 1339, 2009.
- [38] Abe N. and Kudo M., "Non-parametric classifier-independent feature selection," *Pattern Recognition*, Vol. 39, pp. 737 - 746, 2006.
- [39] Peng H., Long F. and Ding C., "Feature Selection Based on Mutual Information: Criteria of Max-Dependency, Max-Relevance, and Min-Redundancy," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 27, No. 8, pp. 1226-1238, Aug. 2005.
- [40] Lin C.H., Du Y.C. and Chen T., "Nonlinear interpolation fractal classifier for multiple cardiac arrhythmias recognition," *Chaos, Solitons and Fractals*, Vol. 42, pp. 2570-2581, 2009.
- [41] Wang Y., Zhu Y., Thakor N. V. and Xu Y., "A Short-Time Multifractal Approach for Arrhythmia Detection Based on Fuzzy Neural Network," *IEEE Transactions on Biomed. Eng.*, Vol. 48, No. 9, Sep. 2001.
- [42] Rohani Sarvestani R., Boostani R. and Roopaei M., "VT and VF classification using trajectory analysis," *Nonlinear Analysis*, In-Press, 2009.
- [43] Nopone K., Kortelainen J. and Seppanen T., "Invariant trajectory classification of dynamical systems with a case study on ECG," *Pattern Recognition*, Vol. 42, pp. 1832 - 1844, 2009.
- [44] Povinelli R. J., Johnson M. T., Lindgren A. C., Roberts F. M. and Ye J., "Statistical Models of Reconstructed Phase Spaces for Signal Classification," *IEEE Transactions on Signal Processing*, Vol. 54, No. 6, pp. 2178-2186, Jun., 2006.
- [45] Owis M. I., Abou-Zied A. H., Youssef A. M. and Kadah Y. M., "Study of Features Based on Nonlinear Dynamical Modeling in ECG Arrhythmia Detection and Classification," *IEEE Transactions on Biomed. Eng.*, Vol. 49, No. 7, Jul. 2002.
- [46] Yu S. N. and Chen Y. H., "Noise-tolerant electrocardiogram beat classification based on higher order statistics of subband components," *Artificial Intelligence in Medicine*, Vol. 46, pp. 165-178, 2009.
- [47] Khadra L., Al-Fahoum A. S. and Binajjaj S., "A Quantitative Analysis Approach for Cardiac Arrhythmia Classification Using Higher Order Spectral Techniques," *IEEE Transactions on Biomed. Eng.*, Vol. 52, No. 11, pp. 1840-1845, Nov. 2005.
- [48] Christov I., Gomez-Herrero G., Krasteva V., Jekova I., Gotchev A. and Egiazarian K., "Comparative study of morphological and time-frequency ECG descriptors for heartbeat classification," *Medical Engineering & Physics*, Vol. 28, pp. 876-887, 2006.
- [49] Lin C.H., "Frequency-domain features for ECG beat discrimination using grey relational analysis-based classifier," *Computers and Mathematics with Applications*, Vol. 55, pp. 680-690, 2008.
- [50] Chen S. W., "A Two-Stage Discrimination of Cardiac Arrhythmias Using a Total Least Squares-Based Prony Modeling Algorithm," *IEEE Transactions on Biomed. Eng.*, Vol. 47, No. 10, pp. 1317-1327, Oct., 2000.
- [51] Kar S. and Okandan M., "Atrial fibrillation classification with artificial neural networks," *Pattern Recognition*, Vol. 40, pp. 2967-2973, 2007.
- [52] Stridh M., Sörnmo L., Meurling C. J. and Olsson S. B., "Sequential Characterization of Atrial Tachyarrhythmias Based on ECG Time-Frequency Analysis," *IEEE Transactions on Biomed. Eng.*, Vol. 51, No. 1, pp. 100-114, Jan. 2004.
- [53] Jekova I., Bortolan G. and Christov I., "Assessment and comparison of different methods for heartbeat classification," *Medical Engineering & Physics*, Vol. 30, pp. 248-257, 2008.
- [54] Lagerholm M., Peterson C., Braccini G., Edenbrandt L. and Sörnmo L., "Clustering ECG

- Complexes Using Hermite Functions and Self-Organizing Maps," *IEEE Transactions on Biomed. Eng.*, Vol. 47, No. 7, pp. 838-848, Jul. 2000.
- [55] Jiang W. and Kong S. G., "Block-Based Neural Networks for Personalized ECG Signal Classification," *IEEE Transactions on Neural Network*, Vol. 18, No. 6, pp. 1750-1761, Nov. 2007.
- [56] Al-Fahoum A. S., "IH combined wavelet transform and radial basis neural networks for the classifying life threatening cardiac arrhythmias," *Medical, Biological Engineering and Computing*, Vol. 37, pp. 566-573, 1999.
- [57] Minami K., Nakajima H. and Toyoshima T., "Real-time discrimination of ventricular tachyarrhythmia with Fourier transform neural network," *IEEE Transactions on Biomedical Engineering*, Vol. 46, pp. 179-185, 1999.
- [58] Prasad G. K. and Sahambi J. S., "Classification of ECG arrhythmias using multi resolution analysis and neural networks," *Conf. Convergent Technologies Bangalore*, India, 2003.
- [59] Chen Y. H. and Yu S. N., "Subband features based on higher order statistics for ECG beat classification," *29th Annual International Conference of IEEE Engineering in Medicine and Biology*, pp. 1859-1862, 2007.
- [60] Guler I. and Ubeyli E. D., "A modified mixture of experts network structure for ECG beats classification with diverse features," *Engineering Applications of Artificial Intelligence*, Vol. 18, pp. 845-856, 2005.
- [61] Guler I. and Ubeyli E. D., "ECG beat classifier designed by combined neural network model," *Pattern Recognition*, Vol. 38, pp. 199-208, 2005.
- [62] Martinez J. P., Almeida R., Olmos S., Rocha A. P. and Laguna P., "A Wavelet-Based ECG Delineator: Evaluation on Standard Databases," *IEEE Transactions on Biomedical Engineering*, Vol. 51, No. 4, pp.570-581, 2004.
- [63] Ghaffari A., Homaeinezhad M. R., Akraminia M., Atarod M. and Davaecha M., "A Robust Wavelet-based Multi-Lead Electrocardiogram Delineation Algorithm," *Medical Engineering & Physics*, In-Press, 2009.
- [64] Benitez D., Gaydecki P. A., Zaidi A. and Fitzpatrick A. P., "The use of the Hilbert transform in ECG signal analysis," *Computers in Biology and Medicine*, Vol. 31, pp. 399-406, 2001.
- [65] Ghaffari A., Homaeinezhad M. R., Atarod M. and Akraminia M., "Parallel Processing of ECG and Blood Pressure Waveforms for Detection of Acute Hypotensive Episodes: A Simulation Study Using a Risk Scoring Model," *Computer Methods in Biomechanics and Biomedical Engineering*, Taylor & Francis Publishing, In-Press, 2009.
- [66] Yu S. N. and Chen Y. H., "Electrocardiogram beat classification based on wavelet transformation and probabilistic neural network," *Pattern Recognition Letters*, Vol. 28, pp. 1142-1150, 2007.
- [67] Dokur Z., Olmez T. and Yazgan E., "Comparison of discrete wavelet and Fourier transforms for ECG beat classification," *Electronic Letters*, Vol. 35, pp. 1502-4, 1999.
- [68] Montgomery Douglas C. and Runger George C., *Applied Statistics and Probability for Engineers*, Third Edition, John Wiley & Sons, 2003.
- [69] Moody G. B. and Mark R. G., "The MIT-BIH Arrhythmia Database on CD-Rom and Software for it," *The Proceeding of Computers in Cardiology*, pp. 185-188, 1990.
- [70] Bishop Christopher M., *Pattern Recognition and Machine Learning*, Springer Publishing, 2006.
- [71] Jang J. S. R., "ANFIS: Adaptive-Network-Based Fuzzy Inference System," *IEEE Trans. Systems, Man, Cybernetics*, Vol. 23, No. 5, pp. 665-685, 1993.



Ali Ghaffari Professor of Mechanical Engineering at K. Nassir Toosi univ. of Technology, Tehran, Iran. Ph. D.: University of California at Berkeley, Dept. of Mechanical Eng., California, USA. M. Sc.: Georgia Institute of Technology, Dept. of Mechanical Eng., USA. B. Sc.: Sharif University of Technology Dept. of Mechanical Eng., Tehran, Iran. His research interests include nonlinear dynamics and control, analysis of stochastic phenomena, Fuzzy systems, Artificial Neural Networks, ANFIS, and biomedical signal processing and control.



Mohammad Reza Homaeinezhad was born in Shiraz, Iran, in 1980. He received the B.S. and the M.S. degrees (with the best honors) both in Mechanical engineering- Dynamic systems and control- from K. Nassir Toosi univ. of Technology, Tehran, Iran in 2003 and 2005 respectively. He is currently Ph. D. candidate of Mechanical engineering in K. Nassir Toosi univ. of Technology. His research interests include nonlinear dynamics and control, statistical signal analysis and parameter estimation, automatic decision making (detection & modulation) theory and biomedical waveforms (BP, ECG & PCG) processing.



Mahdi Akraminia was born in Tehran, Iran in 1985. He received the B.S. degree in Mechanical engineering- Dynamic systems and control- from K. Nassir Toosi univ. of Technology, Tehran, Iran in 2007. Now, he is pursuing toward his M. Sc. degree at K. Nassir Toosi univ. of Technology.