International Journal of Mechanical, Industrial and Aerospace Sciences
ISSN: 2517-9950
Vol:5, No:3, 2011

Combination of Different Classifiers for Cardiac Arrhythmia Recognition

M. R. Homaeinezhad, E. Tavakkoli, M. Habibi, S. A. Atyabi, A. Ghaffari

Abstract—This paper describes a new supervised fusion (hybrid) electrocardiogram (ECG) classification solution consisting of a new QRS complex geometrical feature extraction as well as a new version of the learning vector quantization (LVQ) classification algorithm aimed for overcoming the stability-plasticity dilemma. Toward this objective, after detection and delineation of the major events of ECG signal via an appropriate algorithm, each QRS region and also its corresponding discrete wavelet transform (DWT) are supposed as virtual images and each of them is divided into eight polar sectors. Then, the curve length of each excerpted segment is calculated and is used as the element of the feature space. To increase the robustness of the proposed classification algorithm versus noise, artifacts and arrhythmic outliers, a fusion structure consisting of five different classifiers namely as Support Vector Machine (SVM), Modified Learning Vector Quantization (MLVQ) and three Multi Layer Perceptron-Back Propagation (MLP–BP) neural networks with different topologies were designed and implemented. The new proposed algorithm was applied to all 48 MIT–BIH Arrhythmia Database records (within–record analysis) and the discrimination power of the classifier in isolation of different beat types of each record was assessed and as the result, the average accuracy value Acc=98.51% was obtained. Also, the proposed method was applied to 6 number of arrhythmias (Normal, LBBB, RBBB, PVC, APB, PB) belonging to 20 different records of the aforementioned database (between–record analysis) and the average value of Acc=95.6% was achieved. To evaluate performance quality of the new proposed hybrid learning machine, the obtained results were compared with similar peer–reviewed studies in this area.

Keywords—Feature Extraction, Curve Length Method, Support Vector Machine, Learning Vector Quantization, Multi Layer Perceptron, Fusion (Hybrid) Classification, Arrhythmia Classification, Supervised Learning Machine.

I. Introduction

SIGNAL processing and data mining tools have been developed to enhance the computational capabilities so as to help clinicians in diagnosis and treatment. The ECG is a representative signal containing information about the condition of the heart. The shape and size of the P–QRS–T cycles and the time intervals between various peaks possess useful information about the nature of disease afflicting the heart. However, the human observer cannot directly monitor these subtle details. Besides, since biosignals are highly subjective, the symptoms may appear at random in the timescale. The presence of cardiac abnormalities is generally reflected in the shape of ECG waveform and heart rate. Therefore, study of ECG pattern and heart rate variability has to be carried out over extended periods of time, [1]. If according to any happening, the electro–mechanical function of a region of myocytes fails, the corresponding abnormal effects appear in the ECG which is an important part of the preliminary evaluation of a patient suspected to have a heart–related problem. Generally, each heart–beat type classification algorithm consists of four sequential computational levels: a) pre–processing stage, b) detection–delineation stage, c) appropriate signal segmentation and feature selection–computation stage and d) selection and design of an efficient recognition algorithm as well as feature space dimension enhancement. 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, short time Fourier transform (STFT) [7], fast Fourier transform (FFT) [8–9], original ECG signal [10], preprocessed ECG signal via appropriately defined and implemented wavelet transform [11], statistical moments [12], nonlinear transformations such as Liapunov exponents and fractals [13–15], higher order spectral methods [16–17], power spectral density (PSD) [18–19], Hilbert transform (HT) [20], 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 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 [21], mutual information [22–23], bi–spectrum [24] 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" for tuning of an appropriate classifier such as a neural network [25–28],

M. R. Homaeinezhad is an Assistant Professor of Mechanical Engineering at K. N. Toosi University of Technology, Tehran, Iran (email:mrezahomaei@yahoo.com).
E. Tavakkoli is MS.C. Student of Mechanical Engineering, K. N. Toosi University of Technology, Tehran, Iran (email:ehsan.tavakkoli@yahoo.com).
M. Habibi is MS.C. Student of Mechatronic Engineering, K. N. Toosi University of Technology, Tehran, Iran (email:majid.hbi@gmail.com).
S. A. Atyabi is MS.C. Student of Mechatronic Engineering, Islamic Azad University of Tehran, south branch, Tehran, Iran (email:abbas.atyabi@yahoo.com).
A. Ghaffari is Professor of Mechanical Engineering at K. N. Toosi University of Technology, Tehran, Iran (email:ghaffari@kntu.ac.ir).
Manuscript received **; revised **.

support vector machine [29], pnn [30], knn [31], fuzzy network [32–35] or ANFIS, [36–37].

Organization of This Study. The general block diagram of the proposed heart arrhythmia recognition–classification algorithm including two train and test stages is shown in Fig. 1. According to this figure, first, the events of the ECG signal are detected and delineated using a robust wavelet-based algorithm [38–39]. Then, each QRS region and also its corresponding DWT are supposed as virtual images and each of them is divided into eight polar sectors. Next, the curve length of each excerpted segment is calculated and is used as the element of the feature space for the aim of increasing the robustness of the proposed classification algorithm versus noise, artifacts and arrhythmic outliers, a fusion structure consisting of five different classifiers namely as SVM, MLVQ and three MLP–BP neural networks with different topologies were designed and implemented. The new proposed algorithm was applied to all 48 records of the MITDB and the average value of Acc=98.18% was obtained. Also, the proposed hybrid classifier was applied to 6 number of arrhythmias (Normal, left bundle branch block–LBBB, right bundle branch block–RBBB, premature ventricular contraction–PVC, atrial premature beat–APB, paced beat–PB) belonging to 20 number of the MITDB and the average value of Acc=95.6% was achieved. To compare the outcomes of this study with previous peer–reviewed studies and to show the generalization power of the proposed classification algorithm, 500 samples have been selected for training and 40,438 samples for testing groups. Finally, some comparisons between existing peer–reviewed studies and the presented work aimed for validating the proposed Neuro–SVM–MLVQ learning machine and showing its merit were conducted.

II. Materials and Methods

A. The Discrete Wavelet Transform (DWT)

Generally, it can be stated that the wavelet transform is a quasi–convolution of the hypothetical signal $x(t)$ and the wavelet function (mother wavelet) $\psi(t)$ with the dilation parameter a and translation parameter b, as the following integration

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

The parameter a can be used to adjust the wideness of the basis function and therefore the transform can be adjusted in several temporal resolutions. In Eq. 1, for dilation parameter "a" and the translation parameter "b", the values $a = q^k$ and $b = q^k lT$ can be used in which q is the discretization parameter, l is a positive constant, k is the discrete scale power and T is the sampling period. By substituting the new values of the parameters "a" and "b" into the wavelet function , the following result is obtained

$$\psi_{k,l}(t) = q^{-k/2}\psi(q^{-k}t - lT); \qquad k,l \in Z^+ \qquad (2)$$

The scale index k determines the width of wavelet function, while the parameter l provides translation of the wavelet function.

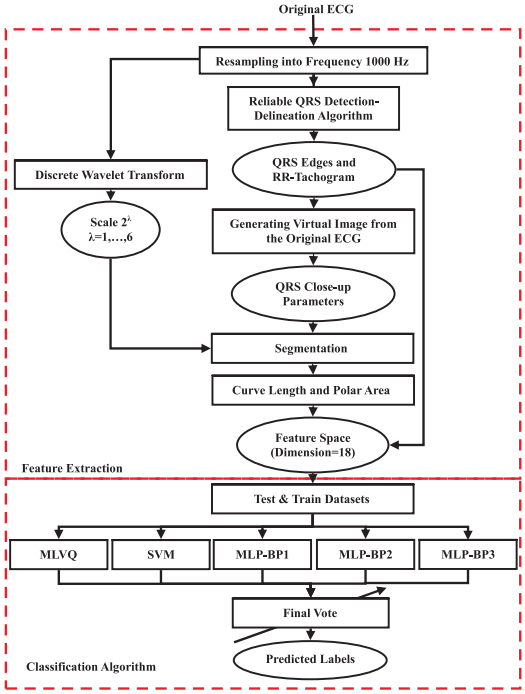

Fig. 1. The general block diagram of an ECG beat type recognition algorithm supplied with the virtual image–based geometrical features

If the scale factor a and the translation parameter b are chosen as $q = 2$ i.e., $a = 2^k$ and $b = 2^k l$, the dyadic wavelet with the following basis function will be resulted [40],

$$\psi_{k,l}(t) = 2^{-k/2}\psi(2^{-k}t - l); \qquad k,l \in Z^+ \qquad (3)$$

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. 2–a, [40]. According to this block diagram, each smoothing function (SMF) is obtained by sequential low–pass filtering (convolving with $G(z)$ filters), while after high–pass filtering of a SMF (convolving with $H(z)$ filters), the corresponding DWT at appropriate scale is generated. In order to decompose the input signal $x(t)$ into different frequency passbands, according to the block diagram of Fig. 2–b, sequential high–pass low– pass filtering including down–sampling should be implemented. The filter outputs $x_H(t)$ and $x_L(t)$ can be obtained by convolving the input signal $x(t)$ with corresponding high–pass and low-pass finite–duration impulse responses (FIRs) and contributing the down–sampling as

$$\begin{cases} x_L(t) = \sum_{k=-\infty}^{k=+\infty} g(k)\ x(2t-k) \\ x_H(t) = \sum_{k=-\infty}^{k=+\infty} h(k)\ x(2t-k) \end{cases} \qquad (4)$$

$$t = 0, 1, \ldots, N-1$$

On the other hand, to reconstruct the transformed signal, the obtained signals $x_H(t)$ and $x_L(t)$ should first be up–sampled by following simple operation

$$\begin{cases} x_L^*(2t) = x_L(t),\ x_L^*(2t+1) = 0 \\ x_H^*(t) = x_H(t),\ x_H^*(2t+1) = 0 \end{cases} \quad (5)$$
$$t = 0, 1, \ldots, N-1$$

If the FIR lengths of the $H(z)$ and $G(z)$ filters are represented by L_H and L_G , respectively, then the reconstructing high–pass and low-pass filters are obtained as

$$\begin{cases} g^*(t) = g(L_G - 1 - t) \\ h^*(t) = H(L_H - 1 - t) \end{cases} \quad (6)$$

Then, the reconstructed signal $x_R(t)$ is obtained by superposition of the up– sampled signals convolved with their appropriately flipped FIR filters as follow

$$x_R(t) = \sum_{k=-\infty}^{k=+\infty} h^*(k) x_H^*(t-k) + \sum_{k=-\infty}^{k=+\infty} g^*(k) x_G^*(t-k) \quad (7)$$

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

$$\Psi(\Omega) = j\Omega \left(\frac{\sin(\Omega/4)}{\Omega/4} \right)^4 \quad (8)$$

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

$$\begin{cases} H(e^{j\omega}) = e^{j\omega/2} \big(cos(\omega/2)\big)^3 \\ G(e^{j\omega}) = 4j e^{j\omega/2} \big(sin(\omega/2)\big) \end{cases} \quad (9)$$

and therefore,

$$\begin{cases} h[n] = (1/8)\delta[n+2] + 3\delta[n+1] + 3\delta[n] + \delta[n-1] \\ g[n] = 2(\delta[n+1] + \delta[n]) \end{cases} \quad (10)$$

After numerous empirical investigations, it was concluded that for frequency contents of up to 50 Hz, à trous algorithm can be used in different sampling frequencies. Therefore, one of the most prominent advantages of 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 40 Hz [38–39]. 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 1000 Hz), scales $2^\lambda (\lambda = 1, 2, \ldots, 5)$ are usable while for sampling frequencies more than 1000 Hz, scales $2^\lambda (\lambda = 1, 2, \ldots, 8)$ contain profitable information that can be used for the purpose of wave detection, delineation and classification.

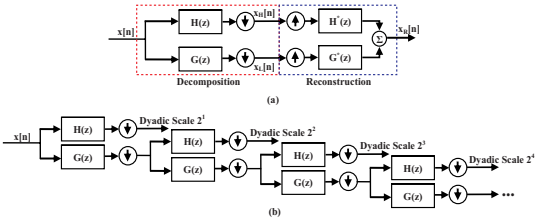

Fig. 2. FIR filter–bank implementation to generate discrete wavelet dyadic scales and smoothing functions transform based on à trous algorithm. (a) one–step generation of detail coefficient scales and reconstruction of the input signal, (b) four–step implementation of DWT for extraction of dyadic scales

III. Modified Learning Vector Quantization Algorithm

1) Conventional LVQ: The conventional LVQ algorithm is a learning machine which requires no hidden layer and possesses a m-neuron and a n-neuron input and output layers, respectively. The number of input layer neurons is equal to feature space dimension while the number of output layer neurons is equal to the number of classes forming the feature space. Each neuron of the input layer is attached to all neurons of the output layer via a connection and a scalar weight is associated with each connection (Fig. 3). The weight between node i of the input layer and node j of the output layer is indicated by w_{ij}. According to the LVQ algorithm, to fulfill the train stage, if the k-th input feature vector x_k is applied to the network, then an appropriately defined distance of the feature vector with the weights terminating to the j-th output layer neuron is calculated as follows

$$\begin{cases} D(j,k) = f(x_k, w_j) \\ w_j = \{w_{ij} \mid i = 1, 2, \ldots, m\} \end{cases} \quad (11)$$

Where $f(x_i, x_j)$ is a scalar distance function. For instance, $f(x_i, x_j)$ can be defined as

$$\begin{cases} (a)\ f(x_i, x_j) = (x_i - x_j)^T \Sigma (x_i - x_j) \\ (b)\ f(x_i, x_j) = \big(\Sigma_{k=1}^p \big(x_i(k) - x_j(k)\big)^r\big)^{1/r} \\ (c)\ f(x_i, x_j) = \frac{1}{p}\ \Sigma_{k=1}^p abs\big(x_i(k) - x_j(k)\big) \end{cases} \quad (12)$$

Where the first term of the Eq. 12 called generalized distance and for the weight matrix $\Sigma = I$ the famous Euclidean norm will be achieved. While the second term of the Eq. 12 is called Minkovski distance of degree r and for $r = 2$, again the Euclidean distance appears. The third term of Eq. 12, is called the City Block distance and is used in many pattern recognition cases. If $D_T(k)$ indicates an array including distances of the feature vector x_k from all output layer neurons, then, the label of this feature vector is predicted by the following criterion

$$\begin{aligned} \hat{\delta}(k) &= \delta(\min\{D_T(k)\}) \\ D_T(k) &= \{D(j,k) \mid j = 1, 2, \ldots, n\} \end{aligned} \quad (13)$$

If the predicted label $\hat{\delta}(k)$ is true, the minimum distance of the array $D_T(k)$ is decreased by learning rate proportionally. On the other hand, if $\hat{\delta}(k)$ is false, then the minimum distance of the array $D_T(k)$ is increased by the same learning rate as

$$\begin{cases} \min\left(D_T(k)\right) = \min\left(D_T(k)\right) - \eta \min\left(D_T(k)\right) \\ \text{if } \hat{\delta}\text{(k) is True} \\ \min\left(D_T(k)\right) = \min\left(D_T(k)\right) + \eta \min\left(D_T(k)\right) \\ \text{if } \hat{\delta}\text{(k) is False} \end{cases} \quad (14)$$

2) Modified Learning Vector Quantization (MLVQ): Suppose that w_{ij} indicates an array including l_j scalar weights and the indices i , j are pointers to the i–th neuron in the input and j–th neuron in the output layers, respectively. If each array w_{ij} is put into a matrix, the weight matrix $(W_j)_{p\times l_j}$ will be resulted. In order to formulate the MLVQ algorithm, first it is noted that $j = 1, 2, \ldots, N_c$ shows the class index, N_c is the number of classes and p is the dimension of the feature space. If each column of the weight matrix W_j is indicated by $C_k^{(j)}$ $(k = 1, 2, \ldots, l_j)$ and f_n is a train feature vector, then the distance between f_n and all $C_k^{(j)}$ arrays can be obtained as

$$d_{kn}^{(j)} = (C_k^{(j)} - f_n)^T \, \Sigma (C_k^{(j)} - f_n) \quad (15)$$

Where, Σ is a weighting matrix and for $\Sigma = I$, the quadratic form is obtained. In this case, the array $D_n^{(j)}$ including all distances between vector f_n and $C_k^{(j)}$ s is created. So, the predicted label $\hat{\delta}$ of feature vector f_n can be determined as follow

$$D_n^{(j)} = \left\{ d_{1n}^{(j)}, \ldots, d_{l_j n}^{(j)} \right\} \quad (16)$$

$$\hat{\delta} = \delta \left\{ \min \left\{ (D_n^{(1)}), \min(D_n^{(2)}), \ldots, \min(D_n^{(N_C)}) \right\} \right\} \quad (17)$$

Where $\hat{\delta}(0)$ is the associated true label operator of the input argument. If the predicted label is true, then the column $C_q^{(\hat{\delta})}$ is decreased by learning rate η while if the predicted label is not true, that column is increased by the same learning rate and can be written as

$$d_{qn}^{(\hat{\delta})} = Arg \min(d_{1n}^{(\hat{\delta})}, \ldots, d_{\hat{\delta} n}^{(\hat{\delta})}) \quad (18)$$

$$\begin{cases} C_q^{(\hat{\delta})} = C_q^{(\hat{\delta})} - \eta \; d_{qn}^{(\hat{\delta})} \text{ if } \hat{\delta}\text{(k) is True} \\ C_q^{(\hat{\delta})} = C_q^{(\hat{\delta})} + \eta \; d_{qn}^{(\hat{\delta})} \text{ if } \hat{\delta}\text{(k) is False} \end{cases} \quad (19)$$

As an interpretation for Eqs. 15 to 19, by inserting feature vector f_n, all pre–defined distances between this vector and all weight vectors between input and output layers are calculated and as the result, the f_n is considered to belong to the class including minimum distance between all weights and all output neurons. If this classification is true, the minimum distance is decreased by learning rate η while if the outcome of the classification is false, the minimum distance will be increased by η. By this learning strategy, desirable results for the selection of the best weight vector and error increasing rate versus epoch number will be attained. The accuracy of the MLVQ network depends upon the following parameters:

- **The Number of Train Epochs.** Generally, more epoch number results better accuracy and the epochs can be considered to have inverse proportionality with number of weights l_j, i.e., the network with larger l_j will requires fewer epochs for reaching an acceptable accuracy. Although, a trade-off between the number of l_j and epochs number can be found for covering the stability-plasticity dilemma.
- **The Number of Weights Assigned to Each Connection.** In the conventional LVQ method if the number of train data is a large value, the weights lying in connections should adapt themselves with several data types and probable outliers and therefore, a weak performance might be expected. In other words, if by entering a new feature vector to the network, the training strategy pushes the incorporated weights toward forgetting the previously learned information, the stability-plasticity dilemma will appear. To solve this problem, more weights can be assigned to each connection. To this end, one way is to increase the number of the output layer neurons and considering more than one node for each class. Although, by this modification the overall accuracy of the network may increase, however, a malformed topology with high computational burden will be achieved. As the second way, instead of assigning a scalar weight to each connection, a vector including some weights is considered between each input-output neurons connection.
- As final comment for the MLVQ method, if the a priori probabilities associated with the feature space classes are not equal, in regulation of the weight vector ending to the class with maximum probability, the corresponding neuron of this class will win predominatingly and correspondingly the winning Euclidean norm is permanently decreased. Thus, after passing a large number of train data from the network, in the test stage, inputted features will falsely being guided to this node and consequently the cumulative accuracy is corrupted. To solve this problem, a modified learning rate is proposed as follow

$$\eta_m = \begin{cases} \eta \; \frac{M_L}{M}(1 - \frac{M_m}{M}) \text{ if } \hat{\delta}\text{(k) is True} \\ \eta \; \frac{M_m}{M}(1 - \frac{M_L}{M}) \text{ if } \hat{\delta}\text{(k) is False} \end{cases} \quad (20)$$

Where, M, M_m and M_L are the data numbers of the train, the largest and the smallest classes, respectively.

A. Radial Basis Function based Support Vector Machine (RBF–SVM) Classifier

In this work, RBF–SVM is implemented as arrhythmias classification method. According to Vapnik formulation [41], if couple $(x_i, \delta(x_i))$ (in which $\delta(x_i)$) is class function, $i = 1, 2, \ldots, N$) describing data elements and their corresponding categories which are linearly separable in the feature space, then

$$f(x) = w^T \phi(x) + b \quad (21)$$

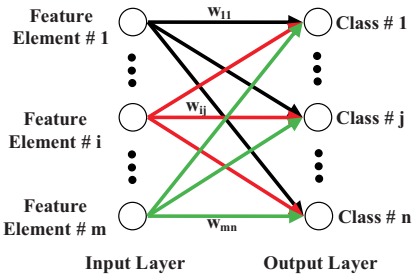

Fig. 3. An mn MLVQ network topology for a m–dimensionality feature space and n–type categories

where w is weight vector, b is bias term and the condition $f(x_i)\delta(x_i) > 0$ holds. On the other hand, if train data is not linearly separable in the feature space to find a suitable separating hyper plane, the following constrained optimization problem should be solved

$$CoF(w,\zeta) = \frac{1}{2}\ ||w||^2 + C\sum_{j=1}^{N}\zeta_i$$
$$s.t.\ \ \delta(x_i)\ (w^T\phi(x_i)+b) \geq 1-\zeta_i\ ,\ \ i=1,2,\ldots,N \tag{22}$$

Where CoF is a cost function. Upon solving the above constrained optimization problem, separating hyper plane will be obtained. In the above equation, C is called regularization parameter which its value generates a trade- off between hyper plane margin and classification error. ζ_i is stack parameter corresponds to x_i. Introducing Lagrange multipliers as

$$CoF(\alpha) = \sum_{j=1}^{N}\alpha_i - \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\left(\alpha_i\alpha_j\delta(x_i)\delta(x_j)\right.$$
$$\left.K(x_i,x_j)\right)\ s.t.\ \sum_{j=1}^{N}\alpha_j\delta(x_j)=0,\ 0\leq\alpha_j\leq C \tag{23}$$

Where $K(x_i,x_j)$ is kernel function obtained from following equation

$$K(x_i,x_j) = \phi^T(x_i)\phi(x_i) \tag{24}$$

for example, $K(x_i,x_j)=(x_i^Tx_j+1)^\lambda$ is polynomial kernel of degree λ and $K(x_i,x_j) = exp(-\gamma||x_i-x_j||^2)$ is RBF kernel. In the Eq. 23, if $\alpha_i > 0$, x_i s are called support vectors. In specific cases, if $\alpha_i = 0$, x_i s are bounded support vectors and if $0 \leq \alpha_i \leq C$, x_i s will be called unbounded support vectors. To solve the constrained Eq. 23, several approaches can be found in the literature, [41]. After solving Eq. 23, the decision function $f(x)$ is obtained as follows

$$\begin{cases} f(x) = \sum_i\ \alpha_i\delta(x_i)K(x_i,x)+b \\ w = \sum_j\ \delta(x_j)\alpha_j\phi(x_j) \end{cases} \tag{25}$$

and margin Λ is obtained as

$$\Lambda = \frac{1}{||w||}\frac{1}{\sqrt{\Sigma_i\Sigma_j\delta(x_i)\delta(x_j)\alpha_i\alpha_jK(x_i,x_j)}} \tag{26}$$

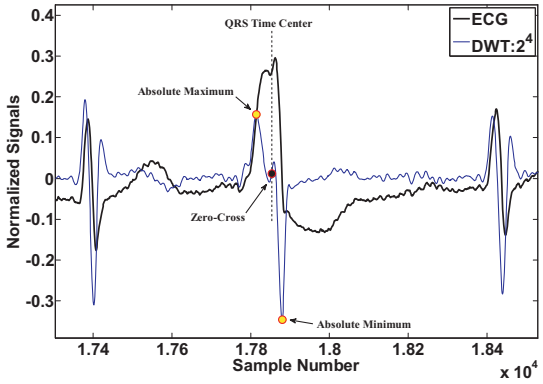

Fig. 4. Determination of the time center of a detected QRS complex using excerpted DWT scale 2^4

More details about fundamental concepts of SVM can be found in [41].

IV. QRS Geometrical Features Extraction

A. ECG Events Detection and Delineation

In this step, QRS complexes are detected and delineated. Today reliable QRS detectors based on Hilbert [42–43] and Wavelet [38–39] transforms can be found in literature. In this study, an ECG detection–delineation method with the sensitivity and positive predictivity Se = 99.95% and P+ = 99.94% and the average maximum delineation error of 6.1 msec, 4.1 msec and 6.5 msec for P–wave, QRS complex and T–wave, respectively is implemented [38]. By application of this method, detecting the major characteristic locations of each QRS complex i.e., fiducial, R and J locations, becomes possible.

B. Detected QRS Complex Geometrical Features Extraction [44]

In order to compute features from the detected QRS complexes either normal or arrhythmic via the proposed method, first a reliable time center should be obtained for each QRS complex. To find this point, the absolute maximum and the absolute minimum indices of the excerpted DWT dyadic scale 2^4 using the onset–offset locations of the corresponding QRS complex, are determined. It should be noted that according to comprehensive studies fulfilled in this research, the best time center of each detected QRS complex is the mean of zero–crossing locations of the excerpted DWT (see Fig. 4).

To make a virtual close–up from each detected QRS complex, a rectangle is built on the complex with following specifications:

- The left–side mid–span (longitudinal direction) of the rectangle is the fiducial location of the QRS complex.
- The maximum absolute vertical distance of the complex from the fiducial point is the half of the rectangle height.
- The center of rectangle is the time-center of the QRS complex.
- The right-hand abscissa of the rectangle is the distance between QRS time center and its J–location.

Afterwards, Each QRS region and also its corresponding DWT are supposed as virtual images and each of them is divided into eight polar sectors. Next, the curve length of each excerpted segment is calculated and is used as element of the feature space, (therefore, for each detected QRS complex, 16 features are computed). The quantity curve–length of a hypothetical time series x(t) in a window with length W_L samples can be estimated as

$$M_{CL} \approx \frac{1}{F_s} \sum_{t=k}^{k+W_L-1} \sqrt{1+\left[\left(x(t+1)-x(t)\right)F_s\right]^2} \quad (27)$$

Where, F_s is sampling frequency of the time series x(t). The curve length is suitable to measure the duration of the signal x(t) events, either being strong or weak. Generally, the M_{CL} measure indicates the extent of flatness (smoothness or impulsive peaks) of samples in the analysis window. This measure allows the detection of sharp ascending/descending regimes occurred in the excerpted segment [39]. A generic example of a holter ECG and its corresponding 2^4 DWT dyadic scale with the virtual images of the complexes provided for feature extraction as well as two quantities obtained from the RR–tachogram are shown in Fig. 5.

V. Design of the Hybrid (Fusion) Neuro-SVM-MLVQ Classification Algorithm

Several differences exists in the structure and operating mechanisms of diverse classification algorithms such as Artificial Neural Network (ANN), MLVQ and SVM. Reasonably, achieving exactly similar result from them for given common train and test feature spaces, cant be expected. Assessments confirm that in the arrhythmia classification of the MITDB beat–types, if two classifiers belonging to different recognition families are appropriately trained, a uniform difference between their operating characteristics versus records cannot be found. In other words, the performance characteristics of two different classifiers show recognition diversity opposed to changing the record number. For instance, suppose that a SVM and a MLP–BP are trained with record 105 of the MITDB. In this case the SVM and the MLP–BP accuracies are calculated as 97.02% and 94.32%, respectively. However, by using of record 221 of the MITDB, the obtained accuracies are 97.97% and 98.17%, respectively indicating the existence of diversity between the SVM and the MLP classifiers. In order to increase the total accuracy of the proposed classification algorithm, one way is to synthesize the output of several classification algorithms with different inherent structures to achieve the best possible accuracy leading to higher robustness against uncertainties and probable arrhythmia or outliers. In this study, to build a fusion (hybrid) classification scheme, five types of different classification methods namely as SVM, MLVQ and three MLP-BP networks with different topologies were properly regulated using the train dataset. The specifications of each classification algorithm are described below.

- **SVM Classifier.** According to section B.3, each SVM includes two parameters C and γ that should be tuned properly to achieve satisfactory accuracies. In this study the best

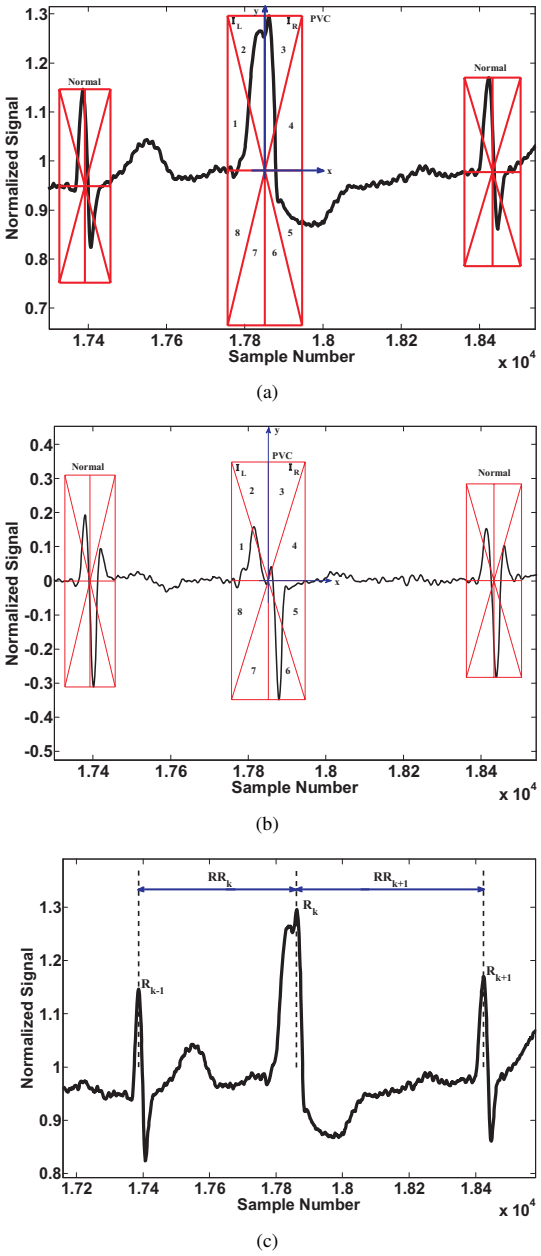

Fig. 5. Extraction of the geometrical features from a delineated QRS complex via segmentation of each complex into 8 polar sectors by generating of a virtual image from the complex. (a) original ECG, (b) DWT of the original ECG and (c) RR–interval

choices for the parameters C and γ were concluded to be 10 and 0.000001, respectively. The predicted labels of the input feature vector were considered as the output of this classifier in the fusion structure.

- **MLVQ Classifier.** As mentioned previously in section B.2, this classifier doesnt require any hidden layer. For this topology, learning rate (LR), number of weights assigned to each connection and maximum epoch number (MEN) were chosen 0.01, 4 and 3000, respectively.
- **MLP*-BP1.** The first MLP–BP classifier includes one

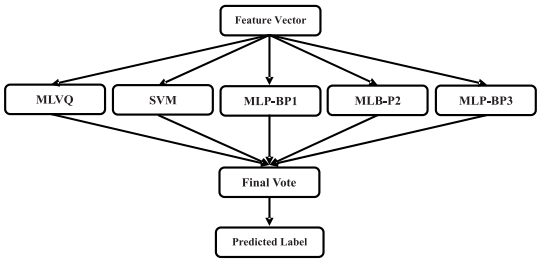

Fig. 6. Design of the fusion classification algorithm consisting of five particle classifiers

hidden layer with number of hidden layer neurons (NHLN) equal to 17, tangent sigmoid and the logarithmic sigmoid as the activation functions of the hidden layer and output layer, respectively. Also, for this ANN, MEN is chosen to be 200.

- **MLP–BP2.** This classifier possesses one hidden layer with NHLN=15. The tangent sigmoid was chosen as the activation function for both hidden and output layers, respectively. For this ANN, MEN = 300 was assigned.
- **MLP–BP3.** The third ANN-type classifier has one hidden layer with NHLN=18 and logarithmic sigmoid for both the hidden and the output layers. In this case MEN=1000 was chosen.

It should be noticed that several parameters such as types of activation functions and several values for NHLN, MEN were examined and were altered based on trying–and–error method and suitable ranges and types were chosen for these parameters. From each classifier embedded into the fusion structure, following outputs are processed

- Predicted labels for train and test feature space.
- Accuracy of the classifier The predicted labels of each particle classification algorithm are used for creation of a hybrid classifier consisting of a MLVQ, a SVM and three MLP–BP types ANN classifiers. To build a fusion classification, in this study, the predicted label of each classifier for the k-th test input is put in the vote array $G(k)$ as follows

$$G(k) = \{p(i,k) \mid i = 1, 2, \ldots, 5\} \tag{28}$$

Where $p(i,k)$ is the label predicted by the i- th classifier of the fusion algorithm for the k-th test input. To estimate the label of test input, the value with the most iteration in $G(k)$ is selected as a label of test input (see Fig 6).

To evaluate performance of the proposed feature extraction method and the fusion classification algorithm, the following steps are pursued

- Evaluation of the discrimination power of the selected features.
- Design of the particle classifiers and their implementation to all MITDB records.
- Design of the fusion classifier for each MITDB record and comparing the obtained results with each particle classifier.
- Selection of some rhythms from the MITDB records and designing of the particle and fusion classifiers.
- Comparison of the final obtained results with the previous similar peer-reviewed studies.

TABLE I
THE DIFFERENT RHYTHM TYPES AND THE CORRESPONDING EQUIVALENT ASCII CODE INTEGER NUMBERS

Numeric Code	Rhythm	Numeric Code	Rhythm
33	Ventricular Flutter wave	83	Supraventricular Premature Ectopic Beat
34	Comment Annotation	86	Premature Ventricular Contraction
43	Rhythm Change	91	Start of Ventricular Flutter/Fibrillation
47	Paced Beat	93	End of Ventricular Flutter/Fibrillation
65	Atrial Premature Beat	97	Aberrated Atrial Premature Beat
69	Ventricular Escape Beat	101	Atrial Escape Beat
70	Fusion of Ventricular and Normal Beat	102	Fusion of Paced and Normal Beat
74	Nodal (junctional) premature Beat	106	Nodal (junctional) Escape Beat
76	Left Bundle Branch Block Beat	120	Non-Conducted P-wave (Blocked APC)
78	Normal Beat	124	Isolated QRS-Like Artifact
81	Unclassifiable Beat	126	Change in Signal Quality
82	Right Bundle Branch Block Beat		

VI. VALIDATION OF HYBRID NEURO-SVM-MLVQ CLASSIFICATION ALGORITHM

In table 1, the numeric codes of the 23 MITDB rhythms (beat–types) and their corresponding annotations are illustrated. After implementation of the MLVQ, SVM and three MLP–BP neural network classifiers and the corresponding fusion classifier to all 48 MITDB records and calculation of the accuracy, the obtained results are shown in table 2. According to this table, the fusion classifier yielded the average accuracy of Acc=98.51% given all data and all rhythms of the MITDB records (within–record analysis). As it can be seen in this table, the overall performance quality associated with the fusion classification algorithm is superior rather than the structural classifiers embedded in the body of the hybrid algorithm. It should be noted that although in some records of the MITDB, one or more particle classifiers might have better performance rather that the fusion classifier, but this behavior doesnt continue uniformly for all records and hence the superiority of the fusion scheme is justified.

In order to be able for comparing the obtained results of this study with Melgani–Bazi [29] utilizing exactly the same train and test databases is mandatory. To this end, records 100, 102, 104, 105, 106, 107, 118, 119, 200, 201, 202, 203, 205, 208, 209, 212, 213, 214, 215 and 217 are selected from the MITDB records and the rhythms Normal, LBBB, RBBB, PVC, APB and PB are extracted according to the MITDB annotation files. In table 3, the name of the MITDB records as well as the selected beat–types and their corresponding beat numbers are presented.

A. Error-Rate Analysis

It should be noted that if some diversely designed classification algorithms show error rate diversity relative to each

TABLE II
PERFORMANCE ILLUSTRATION OF THE PARTICLE CLASSIFIERS SVM, MLVQ, AND THREE MLP-BP NETWORKS AS WELL AS THE CORRESPONDING PERFORMANCE OF THE FUSION CLASSIFICATION ALGORITHM

MIT Records	# of Beats	# of classes	SVM	MLVQ	MLP–BP1	MLP–BP2	MLP–BP3	Final Vote
100	2274	2	100	99.89	99.22	98.67	99.89	100
101	1874	3	99.86	100	99.73	99.73	99.86	99.86
102	2192	5	98.51	99.54	94.62	96.22	98.85	99.54
103	2091	2	99.76	100	99.88	100	99.76	100
104	2311	6	91.63	91.41	95.32	95	92.17	95.65
105	2691	5	97.01	97.86	94.31	94.03	97.85	97.67
106	2098	4	96.29	95.94	96.14	98.44	96.65	98.74
107	2140	2	100	100	100	99.88	100	100
108	1824	6	98.20	98.89	95.30	97.37	98.34	98.93
109	2535	2	100	100	99.30	98.51	100	100
111	2133	2	99.64	100	99.76	100	99.88	100
112	2550	2	99.60	100	99.60	99.60	99.90	99.90
113	1796	2	100	100	99.86	100	100	100
114	1890	5	99.06	99.73	99.20	99.33	99.33	99.45
115	1962	2	99.61	100	99.74	99.48	99.87	99.74
116	2421	3	99.48	99.86	99.48	99.48	99.38	99.48
117	1539	1	100	100	100	100	100	100
118	2301	5	98.25	98.36	96.72	92.14	98.36	98.87
119	2094	4	97.96	99.28	99.88	100	98.08	99.88
121	1876	2	100	100	99.46	99.46	100	99.46
122	2479	1	100	100	100	100	100	100
123	1519	1	100	100	100	100	100	100
124	1634	6	95.21	95.67	96.91	96.45	95.91	97.11
200	2792	5	93.44	93.44	95.32	95.59	93.80	95.95
201	2039	8	92.45	91.71	90.11	96.04	92.71	97.52
202	2146	5	97.65	98.36	95.66	95.68	97.89	98.43
203	3107	6	95.23	96.36	90.46	92.02	95.72	97.55
205	2672	4	98.96	99.25	99.15	98.96	98.96	99.25
207	2385	10	95.66	94.58	92.07	96.51	95.88	96.88
208	3040	6	96.94	93.54	93.64	97.44	96.04	97.02
209	3052	5	97.53	97.7	87.83	93.5	98.11	98.73
210	2685	6	97.19	97.29	95.13	96.08	97.85	98.66
212	2763	3	98.36	98.91	94.37	96.07	98.73	98.91
213	3294	5	92.46	94.97	96.11	94.79	92.38	95.58
214	2297	5	98.57	99.45	95.73	99.01	99.23	99.23
215	3400	4	99.18	99.48	93.57	92.25	99.34	99.77
217	2280	6	85.57	86.66	95.81	96.76	85.35	97.13
219	2312	6	97.39	97.71	98.69	98.91	97.28	98.11
220	2069	4	99.75	98.54	98.30	96.36	100	99.27
221	2462	4	97.96	98.78	98.16	83.51	98.07	99.08
222	2634	5	86.47	85.33	87.04	86.85	85.9	87.04
223	2643	7	95.05	92.49	98.09	97.71	95.63	97.85
228	2141	5	96.47	95.53	76.49	97.64	97.18	96.59
230	2466	2	100	99.8	99.89	100	100	100
231	2011	4	97.75	99. 34	99.34	99.12	98.12	99.25
232	1816	3	96.27	94.89	97.23	96.82	96.96	97.37
233	3152	5	97.29	98.57	85.12	98.96	98.73	99.36
234	2764	3	99.72	99.82	99.27	99.63	99.55	99.91
Total # of Records	48	Total # of Beats			**112646**	Average Accuracy (%)		98.51

other for a given common database, then the utilization of them in a vote-based fusion classification structure is justified. In Fig. 7, the error rate diversities of structural classifiers including SVM, MLVQ, and three MLP–BP networks and also the proposed hybrid classifier for each class are demonstrated. In table 4 and table 5, result of Melgani–Bazi [29] and result of this study are presented, respectively. In table 6, the performance of the fusion classification algorithm has been described by the obtained confusion matrix. For instance, the third row of this table shows that 37, 1, 6, 1 and 2 beat numbers were falsely classified into the N, RBBB, PVC, APB and PB categories, respectively. In this way the number of fusion classifier false negative (FN) detections for the Normal class equals to 47. On the other hand, for instance, the third column of this table illustrates that 21, 0, 5, 0 and 2 beat numbers from the Normal, RBBB, PVC, APB and PB categories were falsely classified as LBBB class, i.e., the number of fusion classifier false positive (FP) detections for the LBBB class equals to 28.

VII. ARRHYTHMIA CLASSIFICATION PERFORMANCE COMPARISON WITH OTHER WORKS

In the final step, in addition to comparison the result of this study with Melgani–Bazi [29] (previous section), the method is assessed relative to other high–performance recent works in order to show the marginal performance improvement of the proposed arrhythmia hybrid classification algorithm. The

TABLE III
THE NAME OF SELECTED MITDB RECORDS WITH THEIR RHYTHM TYPES CONTENTS FOR THE AIM OF PERFORMANCE EVALUATION AND COMPARISON WITH OTHER STUDIES

	Normal		LBBB		RBBB		PVC		APB		PB	
record	train	test	train	test	train	test	train	test	train	test	train	test
100	11	1833	0	0	0	0	0	0	5	12	0	0
102	0	0	0	0	0	0	0	0	0	0	14	1954
104	0	0	0	0	0	0	0	0	0	0	10	1329
105	13	2069	0	0	0	0	1	35	0	0	0	0
106	6	1233	0	0	0	0	12	467	0	0	0	0
107	0	0	0	0	0	0	1	51	0	0	15	2003
118	0	0	0	0	27	2029	0	0	15	36	0	0
119	8	1263	0	0	0	0	10	398	0	0	0	0
200	9	1427	0	0	0	0	19	747	5	11	0	0
201	8	1330	0	0	0	0	4	175	5	11	0	0
202	11	1688	0	0	0	0	0	0	5	13	0	0
203	14	2071	0	0	0	0	10	398	0	0	0	0
205	14	2106	0	0	0	0	2	62	0	0	0	0
208	8	1298	0	0	0	0	22	896	0	0	0	0
209	14	2147	0	0	0	0	0	0	61	146	0	0
212	5	755	0	0	23	1710	0	0	0	0	0	0
213	13	2163	0	0	0	0	5	195	4	9	0	0
214	0	0	50	1751	0	0	6	227	0	0	0	0
215	16	2617	0	0	0	0	4	145	0	0	0	0
217	0	0	0	0	0	0	4	143	0	0	11	1485
Total	150	24000	50	1751	50	3739	100	3939	100	238	50	6771

TABLE IV
RESULT OF THE STUDY CONDUCTED BY MELGANI–BAZI [29]

	Accuracy for each class (%)						
Classifier	Normal	LBBB	RBBB	PVC	APB	PB	Total Accuracy (%)
SVM-linear	81.42	72.58	82.53	74.84	80.25	81.79	80.55
SVM-poly	85.74	89.94	92.03	84.48	83.19	79.11	85.25
SVM-rbf	88.69	87.49	95.98	81.48	87.39	83.47	87.76
RBF	85.17	81.95	75.66	88.16	86.55	74.95	82.74
KNN	81.18	94.8	95.29	76.38	62.18	74.36	81.36
PCA-SVM	86.78	92.63	93.31	83.62	89.07	88.11	87.57
PCA-RBF	84.79	79.15	78.95	92.2	85.29	77.68	83.54
PCA-KNN	86.71	94.63	92.35	81.13	73.1	69.08	84
PSO-SVM	**89.12**	**94.8**	**95.31**	**89.97**	**92.85**	**91.98**	**90.52**

TABLE V
RESULTS OBTAINED FROM SEVERAL CLASSIFICATION ALGORITHMS IMPLEMENTED IN THIS STUDY INCLUDING SVM–RBF, MLP–BP, MLVQ AND THE NEURO–SVM–MLVQ CLASSIFIERS.

	Accuracy for each class (%)						
Classifier	Normal	LBBB	RBBB	PVC	APB	PB	Total Accuracy (%)
SVM-RBF	95.63	93.6	94.57	85.94	56.3	87.74	92.95
MLP-BP1	90.52	96.63	92.67	83.78	74.79	89.84	90.12
MLP-BP2	92.2	96.8	87.3	82.94	74.37	89.87	90.55
MLP-BP3	90.73	96.74	89.89	81.24	80.67	93.8	90.44
MLVQ	93.06	92.23	95.03	69.38	67.65	87.7	89.85
Fusion (Neuro-SVM-MLVQ)	**96.97**	**97.32**	**95.45**	**88.6**	**83.19**	**94.83**	**95.6**

TABLE VI
PERFORMANCE EVALUATION OF THE HYBRID NEURO-SVM-MLVQ CLASSIFICATION ALGORITHM FOR THE SELECTED MITDB RECORDS THE CONFUSION MATRIX

	Normal	LBBB	RBBB	PVC	APB	PB
Normal	23273	21	62	299	205	140
LBBB	37	1704	1	6	1	2
RBBB	59	0	3569	28	24	59
PVC	101	5	39	3490	288	16
APB	19	0	1	20	198	0
PB	53	2	38	242	15	6421

result of comparison of the proposed method and other works is shown in table 7.

VIII. CONCLUSION

In this study, a new supervised heart arrhythmia hybrid (fusion) classification algorithm based on a new QRS complex geometrical features extraction technique as well as an appropriate choice from each beat RR–tachogram was described. In the proposed method, first, the events of the ECG signal were detected and delineated using a robust wavelet–based algorithm. Then, each QRS region and also its corresponding DWT were supposed as virtual images and each of them was divided into eight polar sectors. Next, the curve length of each excerpted segment was calculated and is used as the element of the feature space. To increase the robustness of

TABLE VII
APPROXIMATE PERFORMANCE EVALUATION OF THE PRESENTED NEURO–SVM–MLVQ FUSION CLASSIFICATION ALGORITHM WITH A PREVIOUS STUDIES AIMED FOR VALIDATING THE OBTAINED RESULTS AND SHOWING MERIT OF THE METHOD.

Authors	Method	Arrhythmia types	Dataset	Accuracy
Gholam Hosseini, Luo and Reynolds [27]	NET–BST	6 Classes	1700 dataset from MIT–BIH; 850 training-850 testing; [Normal: 800, PVC: 260, APB:260, RBBB: 260, F:260]	93.0
Mrs. B.Anuradha and V.C.Veera Reddy [45]	Feature extraction; Spectral entropy, Poincare plot geometry, Largest Lyapunov exponent and and Detrended fluctuation analysis Classification; fuzzy	8 Classes	233 beats from MIT–BIH; used for testing	93.13
N.Kannathal and C.M. Lim [36]	Feature extraction: Largest Lyapunov exponent, Spectral entropy, Poincare geometry Classification: anfis	10 Classes	600 dataset from MIT–BIH; 320 training–280testing;	94.0
Tsipouras et al [32]	Knowledge- based system	10 Classes	109880 beats from MIT–BIH; [N, P, f, P, L, R, Q: 102793, V: 6183, [,!,]: 484, BII: 420]	94.26
Omer T.Inan et al [46]	Feature extraction: Wavelet transform and timing interval features Classification: neural networks	3 Classes	NSR,PVC,others	95.16
This study	Feature extraction: Geometrical properties obtained from segmentation of each detected–delineated QRS complex virtual image as well as RR–tachogram Classification: Fusion (Neuro–SVM–MLVQ)	6 Classes	40938 beats from MIT–BIH; 500 training – 40438 testing [Normal: 24150, LBBB: 1801, RBBB: 3789, PVC:4039, APB:338, PB:6821]	95.6

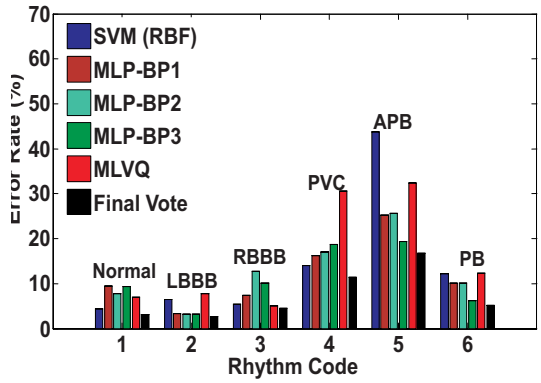

Fig. 7. Error–rate diversity analysis for justification of the fusion of SVM, MLVQ and three MLP–BP classifiers

the proposed classification algorithm versus noise, artifacts and arrhythmic outliers, a fusion structure consisting of five different classifiers namely as a SVM, a MLVQ and three MLP–BP neural networks with different topologies were designed. To show the merit of the new proposed algorithm, it was applied to all 48 MITDB records and the discrimination power of the classifier in isolation of different beat types of each record was assessed (within–record analysis). As the result, the average accuracy value Acc=98.51% was obtained. Also, the method was applied to 6 number of arrhythmias namely as Normal, LBBB, RBBB, PVC, APB, PB belonging to 20 number of the MITDB (between–record analysis) and the average value of Acc=95. 6% was achieved showing marginal improvement in the area of the heart arrhythmia classification. To evaluate performance quality of the new proposed hybrid learning machine, the obtained results were compared with several similar peer–reviewed studies.

APPENDIX A
ABBREVIATIONS

LVQ: learning vector quantization
MLVQ: modified LVQ
KNN: K nearest neighbors
PNN: probabilistic neural networks
SVM: support vector machine
ECG: electrocardiogram
DWT: discrete wavelet transforms
SNR: signal to noise ratio
ANN: artificial neural network
MEN: maximum epochs number
NHLN: number of hidden layer neurons
RBF: radial basis function
MLP-BP: multi-layer perceptron back propagation
LR: learning rate
FP: false positive
FN: false negative
TP: true positive
P+: positive predictivity
Se: sensitivity
MITDB: MIT-BIH arrhythmia database
SMF: smoothing function
FIR: finite-duration impulse response
LBBB: left bundle branch block
RBBB: right bundle branch block
PVC: premature ventricular contraction
APB: atrial premature beat
PB: paced beat
N: normal (rhythm)

REFERENCES

[1] U.R Achraya, *Advances in Cardiac Signal Processing*. New York: Springer Publishing, 2007.

[2] A. Ghaffari , M. Atarod , M. R. Homaeinejad, Y. Ahmady and R. Rahmani, "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)*, Bologna, Italy, 2008, pp.14–17.

[3] A. Ghaffari, M. R. Homaeinezhad, M. Akraminia, M. Atarod, and M. Davaeeha, "Detecting and Discriminating Premature Atrial and Ventricular Contractions: Application to Prediction of Paroxysmal Atrial Fibrillation," *35th Annual Conference of Computers in Cardiology (CinC)*, Lake City-Utah, USA, 2009, pp.13–16.

[4] A. Ghaffari, M. R. Homaeinezhad, M. Akraminia, M. Atarod, and M. Davaeeha, "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)*, Lake City-Utah, USA, 2009, pp.13–16.

[5] A. Ghaffari, M. Atarod, M. R. Homaeinejad, and R. Rahmani, "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)*, Bologna, Italy, 2008, pp.14–17.

[6] A. Ghaffari, M. R. Homaeinejad, Y. Ahmadi and R. Rahmani, "An Open-Source Computer Model for Visualization of Artificial Abnormal Multi-Lead Electrocardiographic Phenomena," *World Journal of Modelling and Simulation*, In–Press, 2009.

[7] M. G. Tsipouras and D. I. Fotiadis, "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.

[8] I. Christov, G. Gomez-Herrero, V. Krasteva, I. Jekova, A. Gotchev and K. Egiazarian, "Comparative study of morphological and time-frequency ECG descriptors for heartbeat classification," *Medical Engineering and Physics*, vol.28, pp. 876–887, 2006.

[9] Chia-Hung. Lin, "Frequency-domain features for ECG beat discrimination using grey relational analysis-based classifier," *Computers and Mathematics with Applications*, vol.55, pp. 680–690, 2008.

[10] Y. Ozbay, R. Ceylan and B. Karlik, "A fuzzy clustering neural network architecture for classification of ECG arrhythmias," *Computers in Biology and Medicine*, vol.36, pp. 376–388, 2006.

[11] Chia-Hung. Lin, Yi-Chun. Du and Tainsong. Chen, "Adaptive wavelet network for multiple cardiac arrhythmias recognition," *Expert Systems with Applications*, vol.34, pp. 2601–2611, 2008.

[12] P. de Chazal, M.O. Dwyer and R. B. Reilly , "Automatic Classification of Heartbeats Using ECG Morphology and Heartbeat Interval Features," *IEEE Transactions on Biomed. Eng*, vol.51, 2004.

[13] K. Nopone, J. Kortelainen and T. Seppanen, "Invariant trajectory classification of dynamical systems with a case study on ECG," *Pattern Recognition*, vol.42, pp.1832–1844, 2009.

[14] R.J. Povinelli, M.T. Johnson, A.C. Lindgren, F.M. Roberts and J. Ye, "Statistical Models of Reconstructed Phase Spaces for Signal Classification," *IEEE Transactions on Signal Processing*, vol.54, 2006.

[15] M.I. Owis, A.H. Abou-Zied, A.M. Youssef and Y. M. Kadah, "Study of Features Based on Nonlinear Dynamical Modeling in ECG Arrhythmia Detection and Classification," *IEEE Transactions on Biomed. Eng*, vol.49, 2002.

[16] S.N. Yu and Y.H. Chen, "Noise-tolerant electrocardiogram beat classification based on higher order statistics of subband components," *Artificial Intelligence in Medicine* , vol.46, pp.165–178, 2009.

[17] L. Khadra, A.S. Al-Fahoum and S. Binajjaj, "A Quantitative Analysis Approach for Cardiac Arrhythmia Classification Using Higher Order Spectral Techniques," *IEEE Transactions on Biomed. Eng*, vol.52, 2005.

[18] S. Kara, M. Okandan, "Atrial fibrillation classification with artificial neural networks," *Pattern Recognition*, vol.40, pp.2967–2973, 2007.

[19] M. Stridh, L. Srnmo, C.J. Meurling and S.B. Olsson, "Sequential Characterization of Atrial Tachyarrhythmias Based on ECG Time-Frequency Analysis," *IEEE Transactions on Biomed. Eng*, vol.51, 2004.

[20] D. Benitez, P.A. Gaydecki, A. Zaidi and A.P. Fitzpatrick, "The use of the Hilbert transform in ECG signal analysis," *Computers in Biology and Medicine*, vol.31, pp.399–406, 2001.

[21] F.A. Minhas and M. Arif, "Robust electrocardiogram (ECG) beat classification using discrete wavelet transform," *Physiological Measurement*, vol.29, pp.555–570, 2008.

[22] H. Liu, J. Sun, L. Liu and H. Zhang, "Feature selection with dynamic mutual information," *Pattern Recognition* , vol.42, pp.1330–1339, 2009.

[23] H. Peng, F. Long and C. Ding, "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, 2005.

[24] S.N. Yu and Y.H. Chen, "Noise-tolerant electrocardiogram beat classification based on higher order statistics of subband components," *Artificial Intelligence in Medicine*, vol.46, pp.165–178, 2005.

[25] S.N. Yu and Kuan-To. Chou, "Integration of independent component analysis and neural networks for ECG beat classification," *Expert Systems with Applications*, vol.34, pp.2841–2846, 2008.

[26] S. Osowski, T. Markiewicz, L. Tran Hoai, "Recognition and classification system of arrhythmia using ensemble of neural networks," *Measurement*, vol.41, pp.610–617, 2008.

[27] H. Gholam Hosseini, D. Luo and K.J. Reynolds, "The comparison of different feed forward neural network architectures for ECG signal diagnosis," *Medical Engineering and Physics*, vol.28, pp.372–378, 2006.

[28] S. Kara, M. Okandan, "Atrial fibrillation classification with artificial neural networks," *Pattern Recognition*, vol.40, pp.2967–2973, 2007.

[29] F. Melgani and Y. Bazi, "Classification of Electrocardiogram Signals With Support Vector Machines and Particle Swarm Optimization," *IEEE Transactions on Information Technology in Biomedicine*, vol.12, pp.667–677, 2008.

[30] Chia-Hung. Lin, Yi-Chun. Du and Tainsong. Chen, "Adaptive wavelet network for multiple cardiac arrhythmias recognition," *Expert Systems with Applications*, vol.34, pp.2601–2611, 2008.

[31] I. Christov, G. Gomez-Herrero, V. Krasteva, I. Jekova, A. Gotchev and K. Egiazarian, "Comparative study of morphological and time-frequency ECG descriptors for heartbeat classification," *Medical Engineering and Physics* , vol.28, pp.876–887, 2006.

[32] M.G. Tsipouras, D.I. Fotiadis and D. Sideris, "An arrhythmia classification system based on the RR-interval signal," *Artificial Intelligence in Medicine*, vol.33, pp.237–250, 2005.

[33] R. Ceylan, Y. Uzbay and B. Karlik, "A novel approach for classification of ECG arrhythmias: Type-2 fuzzy clustering neural network," *Expert Systems with Applications*, vol.36, pp.6721–6726, 2009.

[34] K. Polat, S. Sahan and S. Gune, "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.

[35] T.P. Exarchos, M.G. Tsipouras, C.P. Exarchos, C. Papaloukas, D.I. Fotiadis and L.K. Michalis, "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.

[36] N. Kannathal, C.M. Lim, U. Rajendra Acharya and P.K. Sadasivan, "Cardiac state diagnosis using adaptive neuro-fuzzy technique," *Medical Engineering and Physics*, vol.28, pp.809–815, 2006.

[37] L. Tran Hoai, S. Osowski and M. Stodolski, "On-Line Heart beat recognition using Hermite polynomials and neuro-fuzzy network," *IEEE Transactions on Biomedical Engineering*, vol.52, pp.1224–1231, 2003.

[38] A. Ghaffari, M.R. Homaeinezhad, M. Khazraee and M. Daevaeiha, "Segmentation of Holter ECG Waves via Analysis of a Discrete Wavelet-Derived Multiple Skewness-Kurtosis Based Metric," *Annals of Biomedical Engineering, Springer Publishing*, vol.38, pp.1497–1510, 2010.

[39] A. Ghaffari, M.R. Homaeinezhad, M. Akraminia, M. Atarod, and M. Daevaeiha, "A Robust Wavelet-based Multi-Lead Electrocardiogram Delineation Algorithm," *Medical Engineering and Physics*, vol.31, pp.1219–1227, 2009.

[40] S. Mallat, *A Wavelet Tour of Signal Processing*. Academic Press, 1999.

[41] C.M. Bishop, *Pattern Recognition and Machine Learning*. New York: Springer Publishing, 2006.

[42] D. Benitez, P.A. Gaydecki, A. Zaidi and A.P. Fitzpatrick, "The use of the Hilbert transform in ECG signal analysis," *Computers in Biology and Medicine*, vol.31, pp.399–406, 2001.

[43] A. Ghaffari, M.R. Homaeinezhad, M. Atarod, and M. Akraminia, "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 Engineeing, Taylor and Francis Publishing*, In-Press, 2009.

[44] M.R. Homaeinezhad, H. Najjaran Toosi, A. Ghaffari, M. Tahmasebi and M.M. Daevaeiha, "Long-Duration Ambulatory Holter ECG QRS Complex Geometrical Templates Extraction by Non-Parametric Clustering of the QRS Virtual Close-up Extracted Feature Space," *36th Annual Conference of Computers in Cardiology (CinC)*, Belfast, UK, 2010, pp.24–27.

[45] Mrs.B. Anuradha and V.C. Veera Reddy, "cardiac arrhythmia classification using fuzzy classifiers," *Journal of Theoretical and Applied Information Technology*, vol.4, pp.353–359, 2008.

[46] O.T. Inan, L. Giovangrandi and 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*, vol.53, pp.2507–2515, 2006.

Mohammad Reza Homaeinezhad was born in Shiraz, Iran, in 1980. He received his BSc, MSc and Ph.D. degrees (with the best honors) all in Mechanical Engineering, Dynamic systems and control, in 2003, 2005, 2010, respectively from K. N. Toosi University of Technology, Tehran, Iran. Since September 2010, he has been an assistant professor of Mechanical Engineering (bio-mechatronics) at K. N. Toosi University of Technology and his research interests include nonlinear dynamics and control, statistical signal analysis and parameter estimation, automatic decision making (detection and modulation) theory and biomedical waveforms (ABP, ECG and PCG) processing.

Ehsan Tavakkoli was born in Damavand in 1985. He received the BSc and the MSc degree in Mechanical engineering from Mazandaran University and K.N. Toosi University of technology, respectively in 2008 and 2011. Since 2009 he has been a member of the CardioVascular Research Group (CVRG). His research interests include artificial intelligence, signal processing, control and pattern recognition.

Majid Habibi was Born in Tehran, Iran in 1979. He received his B.S. degree in Software engineering from Payame Noor University in 2004. He also received his M.S. degree in Mechatronics engineering in 2011 from K.N. Toosi University of technology.His research interests include artificial intelligence and control.

Abbas Atyabi was born in Golpayegan, Iran in 1986. He received the BSc and the MSc degrees in Mechanical engineering, and Machatronics engineering from Golpayegan College of Engineering (Joint program with Sharif University of Technology) and Islamic Azad University – South Tehran Branch respectively in 2008 and 2011. Since 2009 he has been a member of the CardioVascular Research Group (CVRG) – K. N. Toosi University of Technology. His research interests include Intelligent Patient Monitoring, Biomedical Image and Signal Processing and machine learning, artificial intelligence and pattern recognition.

Ali Ghaffari was born in Neyshabour in 1947. He received the BSc, MSc and Ph.D. all in Mechanical Engineering from Sharif University of Technology, Georgia Institute of Technology and University of California at Berkeley in 1971, 1974 and 1978, respectively. Since, 1979 he has been with the department of Mechanical Engineering of K. N. Toosi University of Technology. Professor Ghaffaris research is mainly focused on dynamic systems and control including analysis of stochastic phenomena, dynamics and control of nonlinear systems, application of fuzzy set theory and artificial neural networks to mechanical systems, and biomedical signal processing, specifically ECG.