Classification of ECG signals using Hermite functions and MLP neural networks

A. Ebrahimzadeh[*], M. Ahmadi and M. Safarnejad

Faculty of Electrical & Computer Engineering, Babol University of Technology.

Abstract

Classification of heart arrhythmia is an important step in developing devices for monitoring the health of individuals. This paper proposes a three module system for classification of electrocardiogram (ECG) beats. These modules are: denoising module, feature extraction module and a classification module. In the first module the stationary wavelet transform (SWF) is used for noise reduction of the ECG signals. The feature extraction module extracts a balanced combination of the Hermit features and three timing interval feature. Then a number of multi-layer perceptron (MLP) neural networks with different number of layers and eight training algorithms are designed. Seven files from the MIT/BIH arrhythmia database are selected as test data and the performances of the networks, for speed of convergence and accuracy classifications, are evaluated. Generally all of the proposed algorisms have good training time, however, the resilient back propagation (RP) algorithm illustrated the best overall training time among the different training algorithms. The Conjugate gradient back propagation (CGP) algorithm shows the best recognition accuracy about 98.02% using a little amount of features.

1. Introduction

Development of accurate and quick methods for automatic ECG classification is vital for clinical diagnosis of heart disease [1]. An arrhythmia is any abnormal cardiac rhythm [2]. Heart arrhythmias result from any disturbance in the rate, regularity, and site of origin or conduction of the cardiac electric impulse [1]. Among the various abnormalities related with functioning of the human heart, Premature Ventricular Contraction (PVC) is one the most important arrhythmias. PVC is the contraction of the lower chambers of the heart (the ventricles) that occur earlier than usual, because of abnormal electrical activity of the ventricles [2]. This paper investigates the detection and classification of PVC arrhythmias.

In the literature, several methods have been proposed for the automatic classification of ECG signals. The recently published papers are presented in [3–20]. In [3], the authors used the discrete wavelet transform as the feature extractor and linear discriminants as the classifier for PVC beat classification and achieved the recognition accuracy (RA) about 95.6%. In [5] the authors used a feed forward neural network as classifier. They derived five features including the QRS width and offset, amplitude of R segment, the T segment slope and the R-R interval duration for PVC beat classification. In their work RA is about 96.85%. In [6], the authors used wavelet feature extraction in tandem with fuzzy neural network classification for PVC beat classification with RA about 98.20%. In [7], the authors used a multilayer perceptron (MLP) neural network classifier and achieved an accuracy of 88.3% in their testing set. In [8], the authors used morphological information as the features and a neural network classifier for differentiating the ECG beats including PVC beats. They achieved RA about 97%. The method presented in [15] is based on a hybrid fuzzy neural network that consists of a fuzzy self-organizing sub-network

connected in a cascade with a multilayer perceptron. The authors proposed to use high-order statistics as input features for feeding their classifier for recognition of PVC beat from the others. They achieved RA about 98.20%. In [18], an automatic online beat segmentation and classification system based on a Markovian approach is proposed.

According to these works, some important issues of designing an ECG classification system become clear. Addressing these factors suitably lead to development of more robust and efficient recognizers. One of these issues is related to the preprocessing module. Another issue is selection of the classification approach. In particular, the training algorithms of the neural network classifiers. Also, selection of the proper features plays an important role. Here, we present an efficient system for classification of the normal heartbeats, PVC and other abnormalities. In the preprocessing module, an un-decimated wavelet transform is used to provide an informative representation that is both robust to noise and tuned to the morphological characteristics of the waveform features. For the feature extraction module we have used the Hermite functions to represent ECG signals. Then we investigated several Multi-Layer Perceptron neural networks by varying number of hidden layers and neurons of each layer. Multiple algorithms were employed for training these neural networks.

The performance of these algorithms was compared in several experiments. Figure 1 shows a general scheme of the proposed method. The paper is organized as follows. Section 2 describes the preprocessing module. Section 3 explains the feature extraction module. Section 4 presents the classifier. Section 5 describes the database and performance metrics. Section 6 shows some simulation results. Section 7 discusses the results and concludes the paper.

2. Signal preprocessing

Noise reduction is an important problem for analysis of ECG signals. The most troublesome noise sources are the electrical activity of muscles (EMG) and instability of electrode-skin contact [21]. For removal of such noise, an advanced signal processing method, such as discrete wavelet transform (DWT) denoising technique [22], should be used. However, DWT is a time-variant transform. To overcome this problem, we used the stationary wavelet transform (SWT) which is also known as the undecimated wavelet transform (UWT) or translation-invariant wavelet transform. SWT uses the average of several

denoised signals that are obtained from the coefficients of ε-decimated DWT [23].
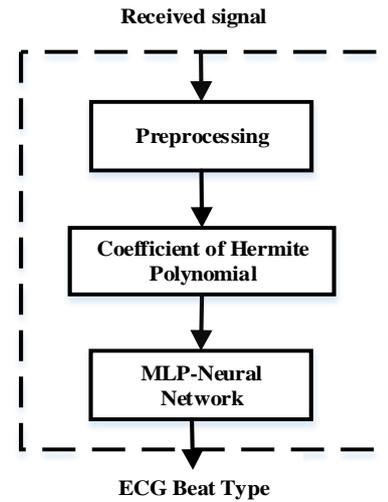
Figure 1. General scheme of the proposed method.

Suppose the signal $s \in L^2(\mathrm{R})$. The SWT is given by:

$$w_v(\tau) = \frac{1}{\sqrt{v}} \int_{-\infty}^{+\infty} s(t) \psi^* \left(\frac{t - \tau}{v} \right) dt \qquad (1)$$

where, $v = 2^k, k \in \square, \tau \in \square$, and ψ^* is the complex conjugate of the mother wavelet.

Figure 2 shows the block diagram of SWT. In this figure, the blocks of H(z) and Hr(z) are the decomposition and reconstruction high pass filters, respectively and the blocks of G(z) and Gr(z) are low pass filters. $d(\cdot, \cdot)$ denotes the decomposition coefficients and $a(\cdot, \cdot)$ denotes the approximation coefficients. This figure shows a three level decomposition. For denoising, we have used the Daubechies wavelet functions (db1) with decomposition level of five, based on our extensive experiments. Figure 3 shows ECG signals before and after denoising.

For smoothing the ECG signals, we have used the Savitsky-Golay (SG) filtering method [23].The filter coefficients are achieved by the un-weighted linear least-squares fitting method using a polynomial function. For this reason, the Savitzky-Golay filter is also called a digital smoothing polynomial filter or a least-squares smoothing filter. A higher degree of polynomial makes it possible to achieve a high level of smoothing without attenuation of the data features. The Savitzky-Golay filtering method is often used for frequency information or spectroscopic information. For the first type, it

preserves the high-frequency components of the signal and for the second type it preserves higher

moments of the Peak. In this paper we have used the first type of SCG.
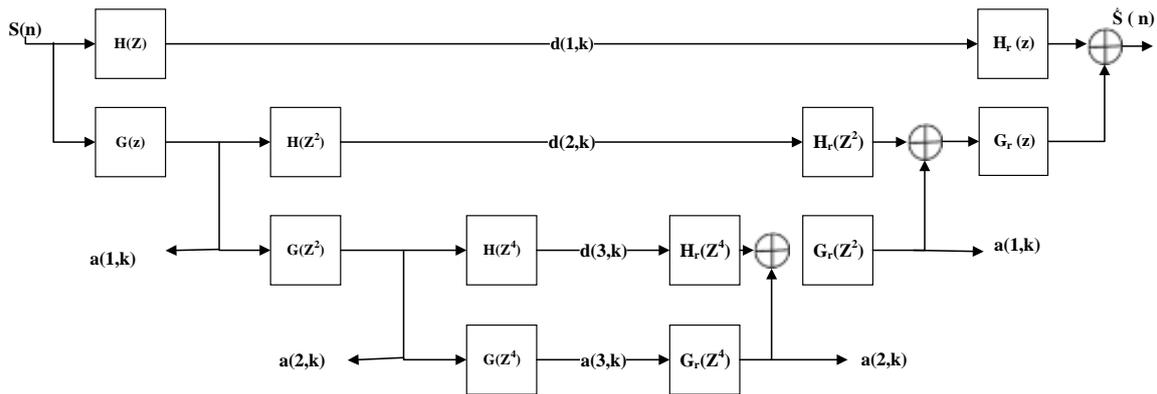
Figure 2. Block diagram of SWT. H(z) and Hr(z) are the decomposition and reconstruction high pass filters. G(z) and Gr(z) are low pass filters. Term *d(·, ·)*denotes the decomposition coefficients and *a(·, ·)*denotes the approximation coefficients.
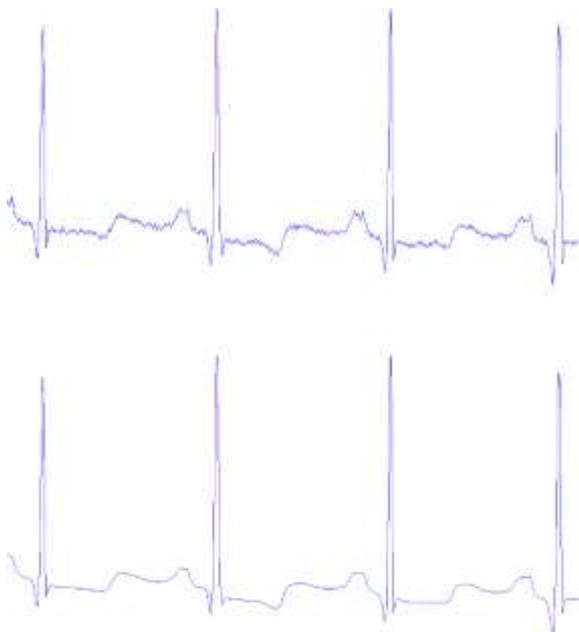
Figure3. Part of ECG signal from MIT-BIH database before and after denoising.

3. Feature extraction

Feature extraction plays an important role in any classification task. It is clear that there is a great variety of morphologies among the heartbeats which belongs to one class, even for the same patient. Moreover, beats belonging to different classes are morphologically similar to each other. They occupy a similar range of values and frequencies; thus, it is difficult to recognize one from the other on the basis of only time or frequency representations.

Different feature extraction techniques have been applied. Traditional representations include features describing the morphology of the QRS complex, such as RR intervals, width of the QRS

complex [3,14,15,16,2], wave interval and wave shape features [15]. Some authors have processed features resulting from Fourier [25] or wavelet transformations [3] of the ECG. Clustering of the ECG data, using methods such as self-organizing maps [16] or learning vector quantization [4], as well as internal features resulting from the neural preprocessing stages [4] have been also exploited. Other important feature extraction methods generate statistical descriptors [24] or orthogonal polynomial representations [16,25]. None of these methods is of course perfect and fully satisfactory. In this paper we will illustrate supervised classification applications that rely on the processing of features originating from the description of the QRS complex by using the higher-order statistics and Hermite basis functions expansion. The HOS description exploits the fact that the variance of cumulant functions is usually lower than the variance of the original signals.

On the other hand, the Hermite expansion takes advantage of the similarity of the individual Hermite functions and different fragments of QRS complex of the ECG waveform. Coefficients of the Hermit function can describe the ECG signals. In the Hermite basis function expansion method, the QRS complex is represented by a series of Hermite functions. This approach successfully exploits existing similarities between the shapes of Hermite basis functions and QRS complexes of the ECG waveforms under analysis. Moreover, this characterization includes a width parameter, which provides good representation of beats with large differences in QRS duration. Let us denote the QRS complex of the ECG curve by $x(t)$. Its expansion into Hermite series may be written in the following way:

$$x(t) = \sum_{n=0}^{N-1} c_n \varphi_n(t, \sigma) \tag{2}$$

where, c_n shows the expansion coefficients, σ is the width parameter, and $\varphi_n(t, \sigma)$ shows the Hermite basis functions of the nth order defined as follows [16]:

$$\varphi_n(t, \sigma) = \frac{1}{\sqrt{\sigma 2^n n! \sqrt{\pi}}} e^{\frac{-t^2}{2\sigma^2}} H_n\left(\frac{t}{\sigma}\right) \tag{3}$$

And $H_n(t/\sigma)$ is the Hermite polynomial of the nth order. The Hermite polynomials satisfy the following recurrence relation:

$$H_n(x) = 2x H_{n-1}(x) - 2(n-1) H_{n-2}(x) \tag{4}$$

For example:

$$H_0(x) = 1; \tag{4-1}$$
$$H_1(x) = x;$$
$$H_2(x) = x^2 - 1;$$
$$H_3(x) = x^3 - 3x;$$
$$H_4(x) = x^4 - 6x^2 + 3$$

The higher order of the Hermite polynomials shows better quick changes in the time. The coefficients of Hermite basis functions expansion, c_n, may be treated as the features used in the recognition process. They may be obtained by minimizing the sum square error, defined as:

$$E = \sum_i \left[x(t_i) - \sum_{n=0}^{N-1} c_n \varphi_n(t_i, \sigma) \right]^2 \tag{5}$$

This error function represents the set of linear equations with respect to the coefficients c_n. They have been solved by using singular value decomposition (SVD) and the pseudo-inverse technique [26]. In numerical calculations, we have represented the QRS segment of the ECG signal by 181 data points around the R peak (90 points before and 90 points after). A data sampling rate equal to 360 Hz generates a window of 250 ms, which is long enough to cover a typical QRS complex. The data have been additionally expanded by adding 45 zeros to the end of each QRS segment. This additional information is added to reinforce the idea that the beats do not exist outside the QRS complex. Subtracting the mean level of the first and the last points normalizes the ECG signals. The width σ was chosen proportional to the width of the QRS complex. These modified QRS complexes of the ECG have been decomposed onto a linear combination of Hermite basis functions. We can solve the (2) with this matrix:

$$\varphi A = x \tag{6}$$

$$\begin{bmatrix} \varphi_0(t_1, \sigma) & \cdots & \varphi_{N-1}(t_1, \sigma) \\ \vdots & \ddots & \vdots \\ \varphi_0(t_s, \sigma) & \cdots & \varphi_{N-1}(t_s, \sigma) \end{bmatrix}_{s*N} \tag{7}$$

$$= \begin{bmatrix} a_0 \\ \vdots \\ a_{N-1} \end{bmatrix}_{N*1}$$

S is number of data which are selected from one signal. t_i (i=1,2,..s) shows the ith datum of one signal. Because the above matrix is not squared, we have used the pseudo-inverse technique.

4. Neural network classifier

We have used MLP neural networks as classifiers. An MLP neural network has several layers. One input layer of source nodes and one or more hidden layers of computation nodes (neurons) and one output layer. The problem of learning algorithm is very crucial for MLP neural network. Back-propagation (BP) learning algorithm is still one of the most popular algorithms [27]. In BP the weight values are updated by a simple gradient descent algorithm:

$$w_{k+1} = w_k - a_k g_k \tag{8}$$

where, g_k is the current gradient, a_k is the learning rate and w_k is a vector of current weights and biases. With standard steepest descent, the learning rate is held constant throughout training, that causes the algorithm be very sensitive to the correct setting of the learning rate. If the learning rate is set too high, the algorithm may oscillate and we may have an unstable algorithm. On the other hand, if the learning rate is set too small, it may take a long time to converge. However, by allowing changes in learning rate change during the training process, the performance of the algorithm can be improved. An adaptive learning rate will attempt to keep the learning step size as high as possible while keeping the learning stable [28]. In Variable Learning Rate Back-propagation (GDX) algorithm the adaptive learning rate with momentum training is used. Gradient descent with momentum is one of the versions of gradient descent that not only allows a network to respond to the local gradient, but also allows it to response to the recent trends in the error surface. Multilayer networks usually use sigmoid transfer functions in the hidden layers. These functions can compress an infinite input range into a finite output range.

Using the steepest descent training method causes a big problem, because the gradient may have a very small magnitude and, therefore, cause small changes in the weights and biases. To solve these problems, the resilient back-propagation (RP) algorithm is proposed [27]. The training phase of the BP algorithm network is usually time-consuming. The conjugate gradient algorithms are proposed to overcome this issue. In the conjugate gradient algorithms, in order to determining the step size, search is made along the conjugate gradient path, which will minimize the performance function along that line. This will make generally faster convergence than steepest descent directions. The conjugate gradient algorithms start out by searching in the steepest descent direction (negative of the gradient) on the first iteration:

$$p_0 = -g_0 \tag{9}$$

Then, a line search is performed to determine the optimal distance to move along the current search direction:

$$w_{k+1} = w_k + a_k p_k \tag{10}$$

The next search direction is then determined so that it is conjugate of the previous search directions. The global method for determining the new search direction is to combine the new steepest descent direction with the previous search direction:

$$p_k = -g_k + \beta_k p_{k-1} \tag{11}$$

The way in which the β_k constant is computed and distinguishes the various versions of conjugate gradient. For the Fletcher-Reeves update the method is:

$$\beta_k = \frac{g_k^T g_k}{g_{k-1}^T g_{k-1}} \tag{12}$$

Equation (12), shows the ratio of the norm squared of the current gradient to the norm squared of the previous gradient. Another version of the conjugate gradient algorithm was introduced by Polak and Ribiére [29]. As with the Fletcher-Reeves algorithm, the search direction at iteration is determined by:

$$p_k = -g_k + \beta_k p_{k-1} \tag{13}$$

For the Polak-Ribiére update, the constant β_k is caculated by:

$$\beta_k = \frac{\Delta g_{k-1}^T g_k}{g_{k-1}^T g_{k-1}} \tag{14}$$

For all conjugate gradient algorithms, the search direction is periodically reset to the negative of the gradient. The standard reset point occurs when the number of iterations is equal to the number of network parameters (weights and biases), but there are other reset methods that can improve the efficiency of training. One such reset method was proposed in [30], the line search requires that the network response to all training inputs be computed several times for each search so that it is computationally expensive. The scaled conjugate gradient algorithm (SCG) was developed by Moller [29]. Newton's method is an alternative to the conjugate gradient methods for fast optimization. The basic step of Newton's method is:

$$w_{k+1} = w_k - A_k^{-1} g_k \tag{15}$$

where, A is the Hessian matrix (second derivatives) of the performance index at the current values of the weights and biases. Because of the expensive calculation of the Hessian matrix, usually some of algorithms which don't require the calculation of second derivatives are introduced. These are called quasi-Newton (or secant) methods. The quasi-Newton method, which has been most successful in published studies, is the Broyden, Fletcher, Goldfarb, and Shanno (BFGS) update [29].The one step secant (OSS) method is an attempt to bridge the gap between the conjugate gradient algorithms and the quasi-Newton (secant) algorithms. This algorithm does not store the complete Hessian matrix. The Levenberg-Marquardt (LM) algorithm [31] uses the approximation to the Hessian matrix in the following Newton-like update:

$$w_{k+1} = w_k - \left[J^T J + \mu I \right]^{-1} J^T e \tag{16}$$

where, J is the Jacobian matrix, e is a vector of network errors, and μ is a constant.

5. Database and performance metrics
5.1. Performance metrics
One of the significant issues in ECG beat classification is how to appropriately evaluate the performance of a proposed method. In this study, we have considered three statistical indices: Accuracy *(Acc)*, Sensitivity *(Se)* and positive Predictivity *(P_P)* which are computed in (17)-(19), respectively. Accuracy is the most important metric for determining overall system performance. We defined the overall accuracy of the classifier as follows:

$$Acc = \frac{N_T - N_E}{N_T} \times 100 \tag{17}$$

where, N_E is the total number of errors in classification and N_T is the total number of beats

in classification. Equation (18), represents the ratio of the number of correctly detected events, *TP* (true positives), to the total number of events:

$$Se = \frac{TP}{(TP + FN)} \times 100 \qquad (18)$$

Pp, is the ratio of the number of correctly detected events, *TP*, to the total number of events detected by the analyzer and is given by:

$$Pp = \frac{TP}{(TP + FP)} \times 100 \qquad (19)$$

Sensitivity measures how successfully a classifier recognises beats of a certain class without missing them, whereas positive predictivity measures how exclusively it classifies beats of a certain type [32].

5.2. MIT-BIH arrhythmia database

The MIT-BIH arrhythmia database [33] was used as the data source in this study. The database contains 48 recordings. Each record has a duration of 30 minutes and includes two leads; the modified limb lead II and one of the modified leads V1, V2, V4 or V5. The sampling frequency is 360 Hz, the data are bandpass filtered at 0.1–100 Hz and the resolution is 200 samples per mV. Twenty-three of the recordings are intended to serve as a representative sample of routine clinical recordings and 25 recordings contain complex ventricular, junctional, and supraventricular arrhythmias. There are over 109,000 labelled ventricular beats from 15 different heartbeat types. There is a large difference in the number of examples in each heart beat type. The largest class is "Normal beat" with about 75,000 examples and the smallest class is "Supraventricular premature beat" (SP) with just two examples. The database is indexed both in timing information and beat classification. For more details about MIT-BIH Arrhythmia database see [34].We used a total of seven records marked as: 100, 101, 102, 104, 105, 106, and 107 in the database. We extracted a total of 15,506 beats; 8,405 normal beats, 625 abnormal PVC arrhythmia beats, and 6,476 other arrhythmic beats. We used the database index files from database to locate beats in ECG signals.

6. Results

6.1. MLP neural network architectures and training algorithms

Various network architectures were evaluated to find an optimum solution for ECG signal diagnosis problem. In order to evaluate system performance, the number of hidden layers as well as the number of neurons in the hidden layers was varied in different experiments.

The output (target) vector is defined with a combination of 1s or 0 s to represent each of the classes being recognized. To assign the ECG waveforms into one of three different classes, the number of neurons in the output layer is set to be three. Table 1 shows the ECG classes and representation of the desired neural network outputs. We have investigated four architectures of neural networks. Two architectures, NET1 and NET2 have a single hidden layer with X neurons which was set to 35 and 45 respectively.

A tan-sigmoid transfer function was selected for hidden layers. The output layer transfer functions for fitting problems are generally selected to be linear, whereas for pattern recognition problems it may be a sigmoid function. However, we propose two other network architectures NET3 and NET4 with two hidden layer. NET3 and NET4 also have three neurons in output layer.

Hidden layers of NET3 have 30 and five neurons respectively and these values for NET4 are 20 and five neurons. Other parameters for these four architectures are the same. In order to facilitate the performance comparison of different training algorithms, we have used their acronyms. Table 2 shows these training algorithms and their acronyms.

Table 1. ECG classes and representation of desired neural network outputs.

Classes	ECG Description	Neural Outputs	Network	
1	Normal	1	0	0
2	Premature Ventricular Contraction Arrhythmia	0	1	0
3	Other Arrhythmias	0	0	1

Table 2. Eight different training algorithms used for training of MLP neural networks.

Algorithm	Acronym
Resilient	RP
Scaled Conjugate Gradient	SCG
Polak-Ribiére Conjugate Gradient	CGP
Conjugate Gradient with Powell/Beale Restarts	CGB
Fletcher-Powell Conjugate Gradient	CGF
Broyden, Fletcher, Goldfarb, and Shanno (BFGS) Quasi-Newton	BFGS
One Step Secant Quasi-Newton	OSS
Variable Learning Rate	GDX

We randomly selected 200 beats from each class, and used these beats for training of neural networks. Total number of beats in our database was 15,506. Clearly, the number of our training set is less than 3.8% of all beats. Thus, our study is well generalized.

6.2. Experiments

To select the last layer's transfer function (output layer), and to comprise hermit features with shape features we performed a set of simulations whose results are shown in table 3. We trained four neural network architectures using two training algorithms: SCG and CGB. Respective recognition accuracies were then achieved. The training dataset included 200 beats from each class and the testing data set included remain beats. Bold numbers shows the higher accuracies in table 3. It can be seen that neural networks with linear transfer functions in their outputs perform better than for the sigmoid transfer

functions. Hence, we chose a linear transfer function in output layers. In order to compare the convergence speed of different training algorithms, the classifiers were trained for a predetermined value of mean-squared error (MSE). The number of training iterations and the time required to achieve the mentioned value of error were recorded. MSE was set to 0.001.

The results are shown in table 4. In order to simplify the comparison among the different results in table 4, we have highlighted in bold type the best results for each of the network structures. In order to evaluate the recognition accuracies (Acc) of the classifiers with different training algorithms and structures, the classifiers were trained for a pre-defined number of iterations set at 1000. In order to improve the generalization ability, validation sets was also used. Two previously defined performance metrics were used for each neural network.

Table 3. Comparative results for networks with linear or sigmoid transfer function in the output layer.

		Accuracy with hermit features		Accuracy with hermit features	
		Output layer with Linear TF	Output layer with sigmoid TF	Output layer with Linear TF	Output layer with sigmoid TF
30*5		97.23	95.37	94.22	93.1
20*5	SCG	97.4	94.35	93.84	92.72
35		97.31	95.19	93.87	91.68
45		97.37	95.10	94.08	91.56
30*5		97.94	92.69	93.93	92.38
20*5	CGB	97.24	93.63	93.9	92.05
35		96.99	94.71	93.8	93.7
45		97.1	93.53	93.7	91.27

Table 5 shows the results. We have compared the performance of MLP NN with PNN neural network. Comparative results of PNN and MLP are shown in table 6.

7. Conclusion

We have proposed a number of efficient methods to accurately classify ECG beats for a relatively large set of data. These methods include three modules: an efficient preprocessing module, a feature extraction module and a classifier. We used a stationary wavelet transform for denoising and a Savitsky-Golay filter for smoothing and normalization in the preprocessing module. The feature extraction module extracts a balanced combination of the Hermit features and three timing interval feature. Then a number of multi-layer perceptron (MLP) neural networks with different number of layers and eight training algorithms are designed. Classification abilities of

these algorithms were compared for three different classes of ECG signals. Four architectures with one hidden or two hidden layers were chosen for this investigation. As our experimental results indicate the neural network with linear transfer function at its output layer performs better than a sigmoid one. At the same time, other layers have sigmoid transfer functions. Generally, it is a difficult problem to determine which training algorithm is the fastest for a given task. To compare the speed of the training methods, consider table 4. As shown, most of the algorithms have good accuracies of test and they have good time for training but the RP algorithm shows the best time for training stage. The SCG algorithm offers a slightly better. It can be seen from the table 4 that for the NET3, with CGP algorithm has recognition accuracy about 98.02%. The Savitsky-Golay filter has important role. We

have tested the system without it. For example for SCG algorithm the RA decrease to 93.50%. We have used another set of features for comparison with Hermitian features. These are shape features. Simulation results show that the NET3 with CGB algorithm has highest RA about 97.94%. We have compared the performance of MLP NN with PNN neural network. The performance of PNN is about 97.1%.

Table 4. Comparison among the speed of convergence of different training algorithms for detection of PVC arrhythmia.

Hidden Layers	Training Function	Accuracy in Train	Accuracy in Test	Training Time(s)	No of Epochs
30*5		96.63	96.26	**1.7**	80
20*5	RP	96.93	96.15	**1.8**	86
35		97.07	95.8	**1.62**	100
45		96.96	96.28	**1.69**	95
30*5		97.45	97.23	3.19	80
20*5	SCG	96.87	97.4	2.8	74
35		97.45	**97.31**	2.6	78
45		97.8	**97.37**	3.4	61
30*5		98.1	**98.02**	3.9	71
20*5	CGP	97.29	97.03	3.02	62
35		96.96	96.18	3.9	70
45		97.06	96.68	4.64	78
30*5		97.83	**97.94**	4.19	65
20*5	CGB	97.75	**97.24**	3.8	61
35		97.45	96.99	2.9	51
45		97.62	**97.1**	4.3	68
30*5		97.06	96.98	4.05	65
20*5	CGF	96.15	95.91	3.31	60
35		97.70	**97.34**	4.12	55
45		97.57	97.09	2.98	65
30*5		97.42	95.25	270	54
20*5	BFGS	97.12	94.7	245	52
35		96.21	94.73	32	60
45		97.47	95.66	49	62
30*5		97.26	97.01	4.3	89
20*5	OSS	96.61	96.48	3.9	111
35		96.48	95.05	3.7	114
45		96.89	95.79	4.7	103
30*5		93.54	90.6	3.5	89
20*5	GDX	93.39	90.45	3.375	93
35		92.82	88.78	4.5	114
45		90.72	86.77	4	150

Table 5. Comparative results of two performance metrics Se and Pp.

Hidden Layers	Training Algorithm	Normal		PVC	
		Se	Pp	Se	Pp
30*5		94.16	97.17	96.6	63.55
20*5	RP	94.25	96.45	95.48	62.04
35		93.33	95.63	96.67	59.43
45		92.73	95.9	99.65	55.78
30*5		93.59	97.74	97.16	67.62
20*5	SCG	94.78	97.53	97.27	68.66
35		94.26	98.34	99.85	71.52
45		94.5	98.5	99.86	72.05
30*5		94.58	97.94	97.04	66.34
20*5	CGP	94.52	97.79	97.66	67.34
35		94.21	98.34	99.88	70.79
45		94.41	98.43	99.84	73.46
30*5		94.54	97.83	97.7	67.13
20*5	CGB	94.41	97.79	97.65	67.34
35		94.32	98.72	99.91	72.63
45		94.31	98.6	99.89	72.91
30*5		94.07	97.42	96.24	65.48
20*5	CGF	94.49	98.2	96.17	65.8
35		94.19	98.48	99.9	71.27
45		94.4	98.29	99.88	72.9
30*5		94.52	98.07	97.86	67.61
20*5	OSS	94.46	97.52	96.41	65.44
35		94.26	98.34	99.87	72.61
45		94.42	98.23	99.84	72.86
30*5		94.17	98.29	96.58	57.59
20*5	BFGS	94.78	97.37	96.44	59.74
35		94.96	98.10	95.32	61.89
45		94.50	97.68	95.26	64.09
30*5		93.09	93.8	90.36	71.6
20*5	GDX	94.78	94	90.82	68.78
35		90.7	92.97	96.46	68.52
45		93.43	94.83	99.38	69.31

Table 6. Comparative results of PNN and MLP neural network.

Classifier	Parameters	Accuracy Test	Normal		PVC	
			Se	Pp	Se	Pp
PNN	Spread=1	96.71	95.27	99.01	96.71	79.6
PNN	Spread=0.7	97.1	95.94	99.1	98.35	80.2
MLP(CGP)	NET3	98.02	94.58	97.94	97.04	66.34
MLP (CGB)	NET3	97.94	94.54	97.83	97.7	67.13

References

[1] Thaler, M. S. (2010). The only EKG book you'll ever need. Philadelphia, PA: Williams & Wilkins.

[2] Clifford, G. D., Azuaje, F. & McSharry, P. (2006). Advanced methods and tools for ECG data analysis. Artech House, Inc.

[3] Senhadji, L., et al. (1995). Comparing wavelet transforms for recognizing cardiac patterns. Engineering in Medicine and Biology Magazine, IEEE, vol. 14, no. 2, pp. 167-173.

[4] Hu, Y. H., Palreddy, S. & Tompkins, W. J. (1997). A patient-adaptable ECG beat classifier using a mixture of experts approach. IEEE Transactions on Biomedical Engineering, vol. 44, no. 9, pp. 891-900.

[5] Yeap, T., Johnson, F. & Rachniowski, M. (1990). ECG beat classification by a neural network. 12th Annual International Conference of the IEEE on Engineering in Medicine and Biology Society, pp. 1457-1458, 1990.

[6] Shyu, L.-Y., Wu, Y.-H & Hu, W. (2004). Using wavelet transform and fuzzy neural network for VPC detection from the Holter ECG. IEEE Transactions on Biomedical Engineering, vol. 51, no. 7, pp. 1269-1273.

[7] Hosseini, H. G., Reynolds, K. & Powers, D. (2001). A multi-stage neural network classifier for ECG events. 23rd Annual International Conference of the IEEE Engineering on Medicine and Biology Society, pp. 1672-1675, 2001.

[8] Inan, O. T., Giovangrandi, L., & Kovacs, G. T. (2006). Robust neural-network-based classification of premature ventricular contractions using wavelet transform and timing interval features. IEEE Transactions on Biomedical Engineering, vol. 53 , no. 12, pp. 2507-2515.

[9] Yu, S.-N & Chou, K.-T. (2009). Selection of significant independent components for ECG beat classification. Expert Systems with Applications, vol. 36, no. 2, pp. 2088-2096.

[10] Acharya, U. R., et al. (2008). Automatic identification of cardiac health using modeling techniques: A comparative study. Information Sciences, vol. 178, no. 23, pp. 4571-4582.

[11] Lin, C.-H. (2008). Frequency-domain features for ECG beat discrimination using grey relational analysis-based classifier. Computers & Mathematics with Applications, vol. 55, no. 4, pp. 680-690.

[12] Übeyli, E. D. (2009). Features for analysis of electrocardiographic changes in partial epileptic patients. Expert Systems with Applications, vol. 36, no. 3, pp. 6780-6789.

[13] Sarvestani, R. R., Boostani, R. & Roopaei, M. (2009). VT and VF classification using trajectory analysis, Nonlinear Analysis: Theory. Methods & Applications, vol. 71, pp. e55-e61.

[14] Piao, M. , et al. (2012). Evolutional diagnostic rules mining for heart disease classification using ECG signal data. Advances in Control and Communication, ed: Springer, pp. 673-680.

[15] Chazal, P. De, Dwyer, M. O. & Reilly, R. B. (2004). Automatic classification of heartbeats using ECG morphology and heartbeat interval features. IEEE Transactions on Biomedical Engineering, vol. 51, no. 7, pp. 1196-1206.

[16] Özbay, Y. , Ceylan, R. & Karlik, B. (2011). Integration of type-2 fuzzy clustering and wavelet transform in a neural network based ECG classifier. Expert Systems with Applications, vol. 38, no. 1, pp. 1004-1010.

[17] Zadeh, A. E. & Khazaee, A. (2011). High efficient system for automatic classification of the electrocardiogram beats. Annals of Biomedical Engineering, vol. 39, no. 3, pp. 996-1011.

[18] Andreão, R. V., Dorizzi, B. & Boudy, J. (2006). ECG signal analysis through hidden Markov models. IEEE Transactions on Biomedical Engineering, vol. 53, no. 8, pp. 1541-1549.

[19] Mitra, S., Mitra, M. & Chaudhuri, B. B. (2006). A rough-set-based inference engine for ECG classification. IEEE Transactions on Instrumentation and Measurement, vol. 55, no. 6, pp. 2198-2206.

[20] de Chazal, P. and Reilly, R. B. (2006). A patient-adapting heartbeat classifier using ECG morphology and heartbeat interval features. IEEE Transactions on Biomedical Engineering, vol. 53, no. 12, pp. 2535-2543.

[21] Kania, M., Fereniec, M. & Maniewski, R. (2007). Wavelet denoising for multi-lead high resolution ECG signals, Measurement Science Review, vol. 7, no. 2, pp. 30-33.

[22] Donoho, D. L. & Johnstone, I. M. (1995). Adapting to unknown smoothness via wavelet shrinkage. Journal of the American Statistical Association, vol. 90, pp. 1200-1224.

[23] Donoho, D. L. (1995). De-noising by soft-thresholding, IEEE Transactions on Information Theory, vol. 41, no. 3, pp. 613-627.

[24] Osowski, S., Hoai, L. T. & Markiewicz, T. (2004). Support vector machine-based expert system for reliable heartbeat recognition, IEEE Transactions on Biomedical Engineering, vol. 51, no. 4, pp. 582-589.

[25] Minami, K. i., Nakajima, H. & Toyoshima, T. (1999). Real-time discrimination of ventricular tachyarrhythmia with Fourier-transform neural network, IEEE Transactions on Biomedical Engineering, vol. 46, no. 2, pp. 179-185.

[26] Hoai Linh, T., Osowski, S. & Stodolski, M. (2002). On-line heart beat recognition using hermite polynomials and neuro-fuzzy network, 19th IEEE Conference of Instrumentation and Measurement Technology, Proceedings of the IMTC, pp. 165-170, 2002.

[27] Hossienzadeh, M. & Ebrahimzadeh, A. (2010). An efficient system for automatic sorting of the ceramic tiles, Multime dia Technology and its Applications, vol. 6, no. 1, pp. 372-374.

[28] Demuth, H., Beale, M. & Hagan, M. (1998). Neural network toolbox user's guide. The MathWorks, Inc. Natrick, USA.

[29] Grippo, L. & Lucidi, S. (1997). A globally convergent version of the Polak-Ribiere conjugate gradient method. Mathematical Programming, vol. 78, no. 3, pp. 375-391.

[30] Powell, M. J. (1977). Restart procedures for the conjugate gradient method. Mathematical Programming, vol. 12, no. 1, pp. 241-254.

[31] Hagan, M. T. & Menhaj, M. B. (1994). Training feedforward networks with the Marquardt algorithm. IEEE Transactions on Neural Networks, vol. 5, no. 6, pp. 989-993.

[32] Khazaee, A. & Zadeh, A. E. (2014). ECG beat classification using particle swarm optimization and support vector machine. Frontiers of Computer Science, vol. 8, no. 2, pp. 217-231.

[33] Mit-bih arrhythmia database (1997), Available: http://ecg. mit. edu/dbinfo. html.

[34] Moody, G. B. & Mark, R. G. (2001). The impact of the MIT-BIH arrhythmia database. Engineering in Medicine and Biology Magazine, IEEE, vol. 20, no. 3, pp. 45-50.

طبقه‌بندی سیگنال‌های قلب به کمک توابع هرمیت و شبکه عصبی MLP

عطاالله ابراهیم زاده*، ملیحه احمدی و مونا صفرنژاد

دانشگاه صنعتی نوشیروانی بابل، بابل، ایران.

چکیده:

کلاسه‌بندی آریتمای قلبی گام مهمی در دستگاه‌هایی در حال توسعه برای نظارت بر سلامت افراد است. این مقاله سیستمی متشکل از سه ماژول برای کلاسه‌بندی ضربان‌های قلب پیشنهاد می‌کند. این سه ماژول شامل ماژول رفع نویز، استخراج ویژگی و کلاسه‌بندی می‌باشد. در بخش اول تبدیل موجک ایستا جهت کاهش نویز سیگنال‌های قلب استفاده شده است. در بخش استخراج ویژگی، ویژگی‌های هرمیت و سه ویژگی زمانی بکار گرفته شده‌اند سپس شبکه عصبی پرسپترون چندلایه با تعداد لایه‌های مختلف و الگوریتم‌های آموزش متفاوت جهت کلاسه‌بندی استفاده شدند. هفت فایل از پایگاه داده‌ی MIT/BIH جهت ارزیابی عملکرد شبکه بکار گرفته شدند. در حالت کلی همه‌ی الگوریتم‌های بکار گرفته شده زمان آموزش خوبی داشتند با این وجود الگوریتم RP زمان بهتر و الگوریتم CGP دقت بالاتری نسبت به سایر الگوریتم‌ها نشان دادند.

کلمات کلیدی: کلاسه‌بندی ضربان‌های قلب، شبکه عصبی MLP، الگوریتم‌های آموزش، تبدیل ویولت، ویژگی‌های هرمیت.