An Automatic Diagnostic Machine for ECG Arrhythmias classification Based on Wavelet Transformation and Neural Networks

A. R. Sahab and Y. Mehrzad Gilmalek

***Abstract*—** The objective of this paper is to design a heart arrhythmias diagnosis instrument that has very low complicated computations. Therefore, a ECG classifier system based on discreet wavelet (DW) transformation and multi layer Perceptron neural network is presented. There is a new Idea in this paper in which signal is pre-processed in order to omit its noises firstly, then, using DW, 6db of signal is divided into eight levels and the minimum, maximum, variance and standard deviation of the signal are obtained. In addition, time features of the signal are obtained. Then combining time features with discrete wavelet output features an array of them are made to be used as final features in order to teach and test a 3-layer MLP neural network. Finally, using 255 heart signal samples existed in MIT-BIH data base, designed Classifier is taught and tested and in its best performance accuracy of 98% percentage have been obtained for three different heart arrhythmias of ECG signals include; RBBB,LBBB and normal heart rhythm.

***Keywords*— Diagnostic Machine, ECG, Classifier, Wavelet Transformation, Heart Arrhythmia, Neural Networks, Features.**

I. Introduction

Bioelectrical signals represent human different organs electrical activities and Electrocardiogram or ECG is one of the important signals among bioelectrical ones that represent heart electrical activity. Deviation and distortion in any parts of ECG that is called Arrhythmia can illustrate a specific heart disease.

These signals carry crucial information about heart operations and conditions that should be extracted and analyzed. The process of extracting and analyzing of ECG can be done by human. In this method, Amplitude and time distances among waves are considered by an operator so that, is a limited and time wasting method, and potentially can have errors. In order to remove mentioned disadvantages, automatic analysis way of ECG has been proposed. Today, variety of methods is introduced for classification and diagnosis of heart arrhythmias. The main differences among them are the way of characteristics extraction and the type of their classifier.

Manuscript received October 29, 2010. This work was supported in part by the Islamic Azad University Lahijan Branch.

A. R. Sahab is a staff member of Electrical Group, Engineering Faculty of Islamic Azad University, Lahijan Branch, Iran. (Phone: 0098-911-1832163; Fax: 0098-141-2222150; e-mail: sahab@iau.lahijan.ac.ir).

Y. Mehrzad Gilmalek is now with the Islamic Azad University Roudsar and Amlash Branch (e-mail: jaber.mehrzad@gmail.com).

In reference [1], Chi et al using 3 neural networks, have classified the arrhythmias with accuracy of 95.1 percent. In reference [2], Karlik et al using neural network, have classified 10 types of arrhythmias with the accuracy of 91.3 and 90.3 percent. In reference [3], Yu et al using 2 RVQ neural networks, have classified the arrhythmias with the accuracy of 91.3 and 90.3 percent.

The common problem in all these proposed methods is that they have used ECG signal itself for heart arrhythmia (HA) classification. In this work, to solve this problem, in addition to extract time and morphology features of ECG signal, wavelet transformation (WT) is used to extract ECG signal features and then the signals are classified by a MLP Neural Network.

II. Selected Normal and Arrhythmia ECG Signals Characteristics

Fig.1 shows a single period of normal ECG signal. Each normal ECG has 4 main sections include; P wave, QRS complex, T wave and U wave. It is necessary to mention that U wave is existed in 50 to 75 percentages of signals. Distortions, changes or deformations of any main section of ECG signal represent an arrhythmia [4], [5] and [15].

A. Normal ECG Signal Characteristics

A normal ECG signal is illustrated in Fig.2. The P wave that is the first part of normal ECG signal has the height of 2 until 3 mm, PR length of 0.12 s. Complex QRS has the height of 5 until 30 mm, time span PR length between 0.06 until 0.12 s and T wave is positive with height of approximately between 0.5 to 10 mm [4], [5] and [15].

B. RBBB Arrhythmia ECG Signal Characteristics

Fig.3 represents a sample LBBB arrhythmia signal. In this signal P wave is identical to normal one and PR length is almost normal but sometimes is more than normal. Complex QRS changes in shape and is wider, more than 0.12 s, comparing with normal signal. Here, T wave is inversed or negative [4], [5] and [15].

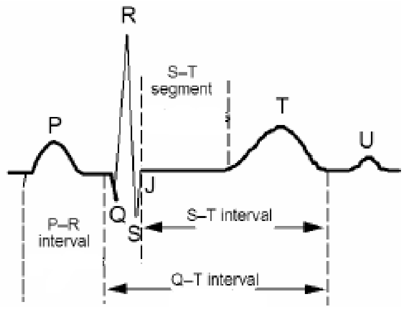

Fig. 1 one period of normal ECG signal

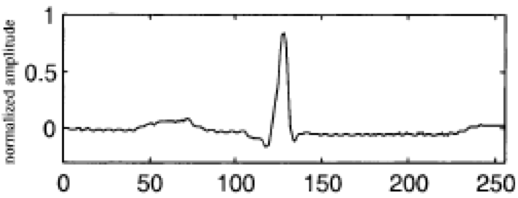

Fig. 2 wave form of normal ECG

C. *LBBB Arrhythmia ECG Signal Characteristics*

Fig.4 represents a sample LBBB signal. These signal properties are like RBBB but its form in complex QRS is different, height of the R wave increases and also its peak is wider [4], [5] and [15].

As it can be inferred from Fig.1, Fig.2 and Fig.3 and their descriptions, these signals have different maximums and minimums so that utilizing these differences and some other characteristics vector can be extracted.

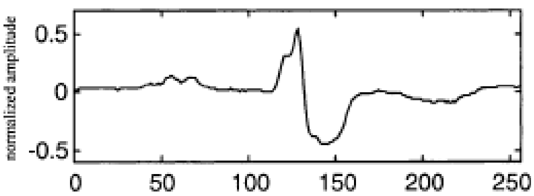

Fig. 3 wave form of RBBB ECG

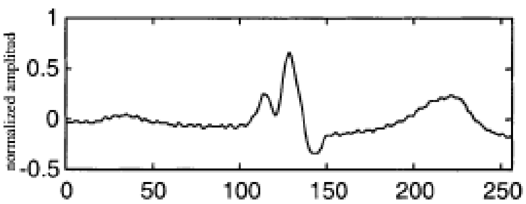

Fig. 4 wave form of LBBB ECG

III. Wavelet Transform

As a result of the infinite extent of the Fourier integral, analysis is time averaged. Thus it contains only globally averaged information and so has the potential to obscure transientor location specific features within the signal.

This limitation can be partly overcome by introducing a sliding time window of fixed length to localize the analysis in time. This local or short time Fourier transform (STFT) provides a degree of temporal resolution by high lighting changes in spectral response with respect to time. A number of alternative time–frequency methods are now available for signal analysis. Of these, the wavelet transform has emerged over recent years as the most favoured tool by researchers for analysing problematic signals across a wide variety of areas in science, engineering and medicine [6].

It is especially valuable because of its ability to elucidate simultaneously local spectral and temporal information from a signal in a more flexible way than the STFT by employing a window of variable width. Thus wavelet transforms produce a time–frequency decomposition of the signal which separates individual signal components more effectively than the traditional short time Fourier transform (STFT). This flexible temporal–spectral aspect of the transform allows a local scale-dependent spectral analysis of individual signal features. In this way both short duration, high frequency and longer duration, lower frequency information can be captured simultaneously.

Hence the method is particularly useful for the analysis of transients, aperiodicity and other non-stationary signal features where, through the interrogation of the transform, subtle changes in signal morphology may be highlighted over the scales of interest. Another key advantage of wavelet techniques is the variety of wavelet functions available, thus allowing the most appropriate to be chosen for the signal under investigation [6].

This is in contrast to Fourier analysis which is restricted to one feature morphology: the sinusoid. In its discrete form using orthogonal wavelet bases, the wavelet transform is particularly usefulin signal coding, allowing information within the signal to be localized within a number of pertinent coefficients for compression purposes. Wavelet transform analysis has now been applied to a wide variety of biomedical signals including: the EMG, EEG, ECG. [6], [12] and [13].

Time–frequency signal analysis methods offer simultaneous interpretation of the signal in both time and frequency which allows local, transient or intermittent components to be elucidated. Such components are often obscured due to the averaging inherent within spectral only methods, i.e. the FFT. A number of time–frequency methods are currently available for the high resolution decomposition in the time–frequency plane useful for *signal analysis*, including the short time Fourier transform (STFT), Wigner–Ville transform (WVT), Choi–Williams distribution (CWD) continuous wavelet transform (CWT) and the discrete wavelet transform (DWT) [6], [12] and [13].

Of course these the wavelet transform has emerged as the most favored tool by researchers as it does not contain the cross terms inherent in the WVT and CWD methods while possessing frequency-dependent windowing which allows for arbitrarily high resolution of the high frequency signal components (unlike the STFT) [6]. Wavelet transforms as they are in use today come in essentially two distinct varieties or classes: the continuous wavelet transform (CWT) and the discrete wavelet transform (DWT) [6].

IV. Discrete Wavelet Transform

The wavelet transform is a linear distribution and continues wavelet transform (CWT) is defined as the sum over all time of the signal multiplied by scaled, shifted versions of the wavelet function Ψ [6] and [18].

$$C(Scale, Position) = \int f(t)\,\Psi(Scale, Position, t)\,dt \tag{1}$$

The results of the CWT are wavelet coefficients C, which are a function of scale and position. Multiplying each coefficient by the appropriately scaled and shifted wavelet yields the constituent wavelets of the original signal $f(t)$ [6] and [18].

The wavelet transform is a technique for analyzing signals. It was developed as an alternative to the Short Time Fourier Transform (STFT) to overcome problems related to its frequency and time resolution properties. More specifically, unlike the STFT that provides uniform time resolution for all frequencies the DWT provides high time resolution and low frequency resolution for high frequencies and high frequency resolution and low time resolution for low frequencies. In that respect it is similar to the human ear which exhibits similar time frequency resolution characteristics. The Discrete Wavelet Transform (DWT) is a special case of the WT that provides a compact representation of a signal in time and frequency that can be computed efficiently [6], [12] and [18].

The DWT is defined by the following equation [6] and [18].

$$W(j,k) = \sum_j \sum_k x(k)\, e^{-\frac{j}{2}}\, \Psi(2^{-j} n - k) \tag{2}$$

Where $\Psi(t)$ is a time function with finite energy and fast decay called the mother wavelet. Fig.5 shows a number of examples of mother wavelets.

In this research Discrete Wavelet transformation (DWT) have been utilized. In the case of Signal processing using wavelet, the type of wavelet is important, as an instance, for a ECG signal, one of the best choices is Daubechies wavelet .in this paper among all Daubechies types of wavelet, 6db is selected based on its similarities to ECG signal.

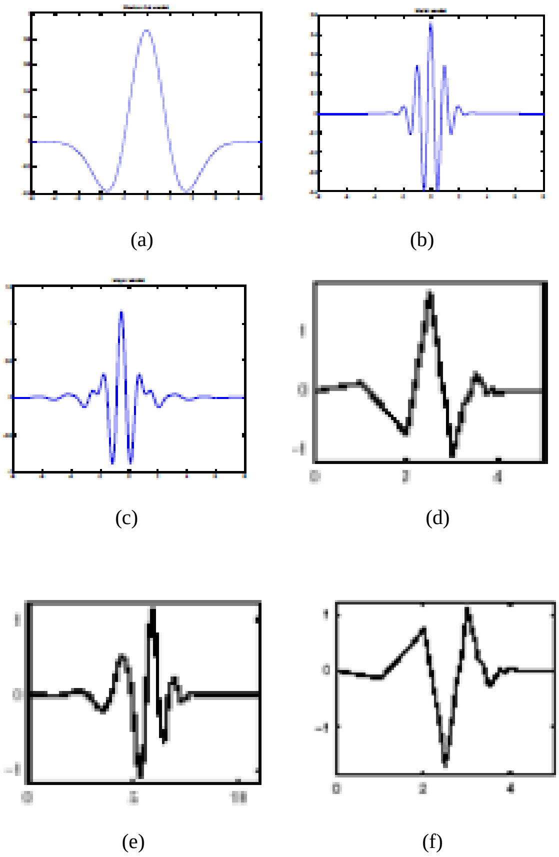

Fig.5 examples of discrete mother wavelets:
(a) mexican hat wavelet (b) Morlet wavelet. (c) Meyer wavelet
(D) Daubechies db4 wavelet (e) Daubechies db6 wavelet
(f) Symlets (sym3) wavelet

V. Neural Network

Neural networks are composed of simple elements operating in parallel. These elements are inspired by biological nervous systems. As in nature, the network function is determined largely by the connections between elements. We can train a neural network to perform a particular function by adjusting the values of the connections (weights) between elements [7].

Commonly neural networks are adjusted, or trained, so that a particular input leads to a specific target output. Such a situation is show in Fig.6. There, the network is adjusted, based on a comparison of the output and the target, until the network output matches the target [7].

Typically many such input/target pairs are needed to train a network. Neural networks have been trained to perform complex functions in various fields, including pattern recognition, identification, classification, speech, vision, and control systems. A classification scheme is developed in which a neural network is used as a classification [7], [8] and [11].

An elementary neuron with *R* inputs is shown Fig.7. Each input is weighted with an appropriate *w*. The sum of the weighted inputs and the bias forms the input to the transfer function *f*. Neurons can use any differentiable transfer function *f* to generate their output [7].

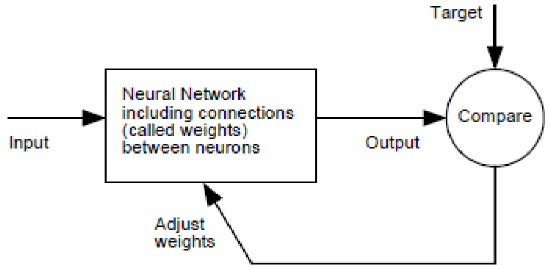

Fig.6 operation neural network

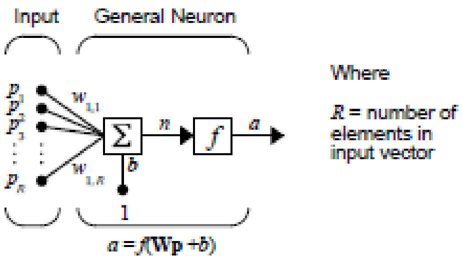

Fig.7 neuron with R inputs

Multilayer networks often use the log-sigmoid transfer function log-sigmoid that is shown in Fig.8 [7].

The function log-sigmoid generates outputs between 0 and 1 as the neuron's net input goes from negative to positive infinity. Alternatively, multilayer networks can use the tan-sigmoid transfer function Tansig that is shown Fig.9 [7].

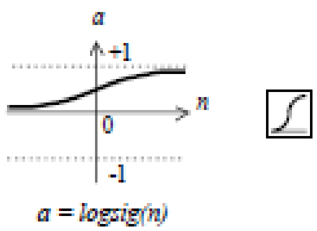

Fig.8 log-sigmoid transfer function

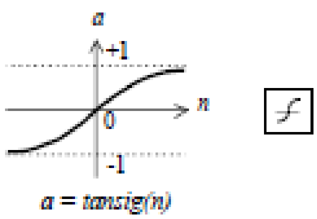

Fig.9 tan-sigmoid transfer function

Occasionally, the linear transfer function purelin that is shown in Fig.10 is used in back propagation Networks [7].

If the last layer of a multilayer network has sigmoid neurons, then the outputs of the network are limited to a small range. If linear output neurons are used the network outputs can take on any value [7].

In this research we want use purelin function and Tansig function as transfer function of network neurons.

A. Back-propagation Algorithm

The back-propagation algorithm (BP) allows experiential acquisition of input/output mapping knowledge within multi-layer networks. BP performs the gradient descent search to reduce the mean square error (MSE) between the actual output of the network and the desired output through the adjustment of the weights.

It is highly accurate for most classification problems because of the property of the generalized data rule [7] and [8].

In the traditional BP training, the weights are adapted using a recursive algorithm starting at the output nodes and working back to the first hidden layer. The above algorithm could be performed using the following equation [8]:

$$W_{ij}(t+1) = W_{ij}(t) + \eta\delta_j x_i + \alpha\big(W_{ij}(t) - W_{ij}(t-1)\big) \qquad (3)$$

Where W_{ij} is the weight value of node i connected to node j from previous layer, x_i is the output at node i, η and α are the learning rate and the momentum term respectively. δ_j is an error term for node j. If node j is an output node, then [8]:

$$\delta_j = f(x_j)(d_j - x_j) \qquad (4)$$

Where $f(x)$ is the nonlinear sigmoid logistic function and d_j is the desired output of node j. If node j is an internal hidden node, then [8]:

$$\delta_j = f(x_j)\sum \delta_k w_{jk} \qquad (5)$$

Where k is over all nodes in the layer above node j.

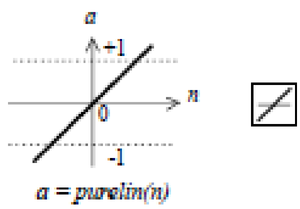

Fig.10 linear transfer function

(4) and (5) show that the error term depends basically on the activation function which is given by (6). The sigmoid function is chosen because it is a continuous function whose derivatives exist [8].

$$f(x)=\frac{1}{1+\exp(-x)} \tag{6}$$

B. *Batch Gradient Descent (Traing d)*

The batch steepest descent training function is traingd. The weights and biases are updated in the direction of the negative gradient of the performance function.

There are seven training parameters associated with traingd[7]:

- epochs
- show
- goal
- time
- min_grad
- max_fail
- lr

The learning rate lr is multiplied times the negative of the gradient to determine the changes to the weights and biases. The larger the learning rate, the bigger the step. If the learning rate is made too large, the algorithm becomes unstable. If the learning rate is set too small, the algorithm takes a long time to converge [7] and [8].

The training status is displayed for every show iterations of the algorithm. (If show is set to NaN, then the training status is never displayed.) The other parameters determine when the training stops. The training stops if the number of iterations exceeds epochs, if the performance function drops below goal, if the magnitude of the gradient is less than mingrad, or if the training time is longer than time seconds [7] and [8].

VI. Methods and Techniques

The block diagram of the utilized method is shown in Fig.11. As it can be seen, it has 3 stages. Firstly, the ECG signal should be pre-processor in order to eliminate existed noises of ECG and preparing a processed signal for the next stage. Secondly, there is main processor that extracts feature and produces a feature vector for the next usage. Finally, there is a classifier that determines the type of arrhythmia.

Fig. 11 block diagram of the arrhythmia diagnosis system

A. *Achieving Needed Data*

In order to teach and test the neural network some ECG signals have been downloaded from MIT-BIH data base [9] and 17 files of them, with time length of 30 minutes and sampled frequencies of 300 HZ, have been utilized in this work.

255 sample signals from different times of these files have been chosen incidentally to provide needed data. Table.1 shows utilized files and their classifications.

B. *Pre-Process of ECG Signal*

In this stage, it is necessary to eliminate noises from input signals using WT [16]. This pre-process of ECG signal before extracting its feature can resulted in better extracted features which in turn can resulted in a increase of system efficiency in HA diagnosis.

More explanations are given in section.7 of paper. Fig.12 and Fig.13 illustrate a sample signal of ECG before and after noise omission.

C. *Main Processor (Characteristics Extraction)*

After noise elimination, it is necessary to extract the features from the signal in order to use it in the next stage. Pattern identification is a basic step in features extraction and classification parameters. There are varieties of feature extraction ways.

In this work two categories of features are extracted from ECG signals:

1. Features resulted from WT applying.
2. Time and morphology features of ECG signal itself.

Utilized WT in this work is DWT [10], [11], [12] and [16] that will be described in section 7.

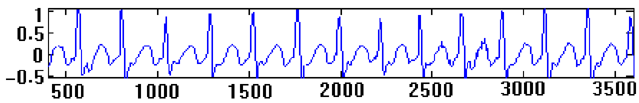

Fig. 12 a main ECG signal after noise omission

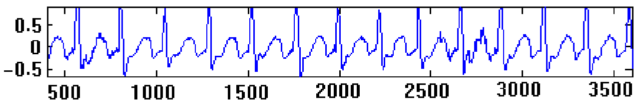

Fig. 13 noised signal

TABLE I
UTILIZED FILES AND THEIR CLASSIFICATIONS THAT ARE OBTAINED FROM MIT-BIH DATA BASE [9]

Class	Records
Normal	100-101-103-112-115-117-121-123-202-220-222-234
RBBB	212-118-124
LBBB	109-111

D. Classifier

A part of researches in this work is devoted to consideration of different neural networks in order to determine their accuracy in identification and separation of categories or classes.

Among all neural networks MLP [17], which is illustrated in Fig.14 and Fig.15, has been chosen based on the below mentioned reasons:

1. It has 3 layers including; input layer, hidden layer and output layer.
2. As result of this fact that the numbers of existed neurons in hidden layer is an effective parameter for improvement of learning results, variety of neuron numbers was chosen in order to achieve the optimum number based on output results.
3. Tansig and Purelin function and also their combination function have been compared as transfer function of network neurons and finally, the effective one has been chosen.
4. For training utilized of BP algorithm and traingd function.
5. Lr parametr 0.05 has been chosen.
6. For teaching of mentioned neural network, mean squared error (MSE) or goal parameter criterion was utilized in which error of 0.0001 was the stopping point of teaching and maximum repetitions was 500 times.

VII. SIMULATION ENVIRONMENT

The simulations have been done by MATLAB software [13] because of its various capabilities in recognition of the pattern. As it mentioned before, firstly, the ECG signal should be pre-processed to eliminate its noise that is done by Wavelet. In this research Discrete Wavelet transformation (DWT) have been utilized. In the case of Signal processing using wavelet, the type of wavelet is important, as an instance, for a ECG signal, one of the best choices is Daubechies wavelet [13]. Among all Daubechies types of wavelet, 6db is selected based on its similarities to ECG signal. DWT can be utilized as a bank filter [16], [18] and [19].

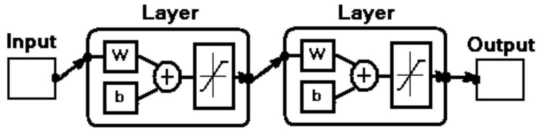

Fig.14 illustrates the structure of the utilizes neural network

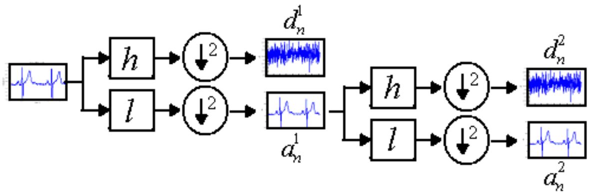

Fig.15 bank filter proportional to two stages WT

Fig. 14 illustrates the structure of the utilized neural network. Fig. 15 shows a sample of a bank filter proportional to two stages Wavelet Transformation.

In above mentioned bank filter h is a high pass filter or is the wavelet, l is a low pass filter or scaling function. It should be mentioned that in this work, in order to achieve better results, a 8 stage wavelet have been used to eliminate signal noise [14], [16] and [17]. After noise elimination signal features should be extracted. Here, two groups of features are extracted and their combination had used as feature vector for neural network.

A. Characteristic Vector Resulted from Applying Wavelet

In this stage, by the use of DWT signal has been divided in to 8 levels, and in each level low frequency coefficients and high frequency, is obtained. Direct applying of wavelet coefficient as neural network inputs has resulted in an increase in neuron numbers in hidden layer which in turn has a negative impact on network operation.

Thus, in this stage, variance, maximum, minimum and standard deviation of the signal in each level and for each high frequency and low frequency coefficient has been calculated. Finally for each of signals 64 feature have been obtained.

B. Time and Morphology Characteristics of ECG Signal

Time and morphology extracted features of ECG signals include; variance, maximum, minimum, standard deviation and R-R distance in 5 beats of the ECG signal.

By extracting these features, five feature of a signal will be achieved. These 5 time and morphology feature are combined with 64 obtained feature of applied DWT and finally 69 feature are assumed as total feature vector of neural network. This total vector is normalized to obtain optimum results and also is used to recognize three categories of heart arrhythmias. First category is people with normal heart operation. Second category is people who suffer from Right Bundle Branch Block (RBBB) and third category is people who suffer from Left Bundle Branch Block (LBBB).

As it is mentioned before to classify this three categories of arrhythmias three MLP neural network is utilized this has 10 neurons with Tansig transfer function in hidden layer and Purelin transfer function in its output layer. Total feature vectors of some signals are used to teach the neural network and some others are accidentally chosen to test it.

VIII. Simulation Results

To simulate and teach the network, 45 data, 15 data for each class, are used and two groups of feature have been extracted from them. Combining these feature based on mentioned descriptions in last section, a 69×45 matrix has been obtained as teaching input data and also 210 data, 70 data for each classified categories, is used for testing. Then, simulation has been done by different numbers of neurons in hidden layer. Using try and error method has given us 10 numbers of neurons in hidden layer as optimum state.

Table.2 shows the results of applying two groups of features to neural network in accuracy state of 100% for teaching stage. In this state of 100% teaching accuracy, when testing data are applied, test results will give us the best accuracy result of 98%. Defining the percentage of efficiency and correct recognition of arrhythmia by neural network is done by below parameter:

$$A = \frac{N_c}{N_t} \times 100 \tag{7}$$

In which A is percentage of efficiency, N_c is the number of correct classified samples and N_t is number of total samples. In Table.2 percentage of efficiency is shown for each classified categories.

IX. Comparison Between Proposed Method And Other ECG Signal Classifier Systems

To provide comparison, some ECG classifier systems are chosen such as: [20], [21], [22] and [23]. Summary of these systems accuracies (percentages of efficiency) are given in Table.3. A simple comparison shows that proposed algorithm has higher accuracy. It should be mentioned that this comparison is not very fair because of variety of diseases and different classifications are done in related papers.

X. Results and Discussions

In this paper, utilizing different signal processing techniques and re-determining pattern, a powerful method has been achieved for diagnosis of HA. The crucial point that should be considered in this method is that it needs more data for expanding its decisions. Based on experiments results, it has been understood that determining feature vector, transfer function type and hidden layer neurons are the main influential factors in teaching and testing of the network in order to classify HA. In this work it is shown that combining the time and morphology features of ECG signal with features obtained from WT, results in network efficiency.

TABLE II
Results of Applying Two Groups' Characteristics to Neural Network and Their Camparison

Type of feature	Classified categories	Number of teaching data	Number of testing data	Number of correct classified samples	Network Efficiency %
Combination of time and morphology feature with ECG from WT	Normal	15	70	69	98.57
	RBBB	15	70	68	97.14
	LBBB	15	70	69	98.57
	Total	45	210	206	98.00
Feature resulted from applied wavelet	Normal	15	70	67	95.71
	RBBB	15	70	66	94.28
	LBBB	15	70	66	94.28
	Total	45	210	199	94.76

TABLE III
Comparison Among Different Methods of ECG Classifications Accuracies

Methods	Accuracy (%)
Proposed Method	98.00
MME	97.78
SOM-SVD	92.20
MLP-LVQ	96.80
FHYB-HOSA	96.06

Acknowledgment

This work is supported by Electrical group, department of engineering, Islamic Azad university Lahijan branch, Lahijan, Iran.

References

[1] Z. Chi and M. A. Jaberi “Identification of supraventricular and ventricular arrhythmias using a combination of 3NNs” *Department of Electrical Engineering University of Sydney, 1992.*

[2] B. Karlik and Y. Ozbey “A New Approach for Arrhythmias Classification” *Proc. Of Annual International Conference of IEEE of Medicine and Biology Society, 1996.*

[3] Yu Hen Hu “Customization of ECG Beat Classifiers Developed Using SOM&LVQ” *Proc. of IEEE, 1997.*

[4] Porterfield, Linda M. “ECG Interpretation Made Incredibly Easy” Ones Edition, Spring House Co, Pennsylvania, 1997.

[5] Goldberger, Ary L. “Clinical Electrocardiography A Simplified Approach” , 6th edition, Mosby co , St.Louis , 1999.

[6] P. S. Addison “Wavelet transforms and the ECG: a review” *Institute of Physics Publishing, Physiol. Meas., Vol. 26, 2005, pp .155–199.*

[7] Neural Network Toolbox. Available: http://www.mathworks.com

[8] L. Khadra, A. Fraiwan, W. Shahab, **“**Neural-wavelet analysis of cardiac arrhythmias**”**, *Proceedings of the WSEAS International Conference on Neural Network and Applications (NNA '02), Interlaken, Switzerland, February 11-15, 2002, pp.3241-3244.*

[9] MIT-BIH Arrhythmia Database Directory. Available:

http://www.physionet.fri,uni-1j.si/phsiobank/database/mitdb

[10] S. G. Mallat “A theory for multi resolution signal ecomposition: the wavelet representation” *IEEE Trans. Pattern Recognition and Machine Intelligence, Vol. 11, No.7, 1989, pp. 674-693.*

[11] G. Qiang, Y. Jun, P. H. Chuan and et al "Wavelet Neural Network for ECG signal classification" *International Conference BME' 96, On Biomedical Engineering, Hong Kong, 1996*.

[12] Y. Jun, P. H. Chuan and G. Qiang. "Learning algorithm for wavelet neural network based on discrete wavelet transforms and its application in ECG classification" *CNNC97, Nanjing, China, 1997.*

[13] Available: http://www.mathworks.com

[14] K. Anant, F. Dowla and G. Rodrigue "Vector quantization of ECG wavelet coefficients" *IEEE Trans. On Signal Processing, Vol. 2, No. 7, 1995, pp. 129-130.*

[15] P. Welling, Z. Cheng, M. Semling and G. S. Moschytz "Electromyogram data compression using single-tree and modified zero-tree wavelet encoding" *Proc. Of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Vol. 20, No. 3, 1998, pp. 1303-1306.*

[16] M. Rizzi, M. D. Aaloia and B. Castagnolo "Fast parallelized algorithm for ECG analysis", *WSEAS Trans. On Biology and Biomedicine, issue.8, Vol.5, August 2008, pp. 210-219.*

[17] Abdel-Rahman Al-Qawasmi, Khaled Daqrouq "ECG Signal Enhancement Using Wavelet Transform", *WSEAS Trans. On Biology and Biomedicine, issue.2, Vol.7, August 2010, pp. 210-219.*

[18] Salama Meghriche, Mohammed Boulemden and Amer Draa "Agreement Between Multi-Layer Perceptron and a Compound Neural Network on ECG Diagnosis of Aatrioventricular Blocks", *WSEAS Trans. On Biology and Biomedicine, issue.1, Vol.5, January 2008, pp. 12-22.*

[19] Y. Zandi Mehran "A New Application of Wavelet Transform in Classification the Arterial Pulse Signals", *Proc. Of the 5th WSEAS International Conference on System Science and Simulation in Engineering, Tenerife, Canary Island, Spain, December 2006, pp. 448-453.*

[20] G. Selvakumar, K. Bhoopathy Bagan, B. Chidambararajan, "Wavelet Decomposition foe Detection and Classification of Critical ECG Arrhytmias", *Proc. Of the 8th WSEAS International Conference on Mathematics and Computers in Biologo and Chemistry, Vancouver, Canada, June 2007, pp. 80-84.*

[21] I. Guler, E.D. Ubeyli "A Modified Mixture of Experts Network Structure for ECG Beats Classification with Diverse Features", *Engineering Applied on Artificial Intell., Vol.18, No.7, pp. 845-856*.

[22] Y. H. Hu, S. Palreddy and W. J. Tompkins "A Patient-Adaptivable ECG Beat Classifier Using a Mixture of Expert Epproach", *IEEE Trans. On Biomedical Engineering Vol.44, No.9, 1997, pp. 891-900.*

[23] G. E. Oien, N. A. Bertelsen, T. Eftestol and J. H. Husoy "ECG Rhythm Classification Using Artificial Neural Networks", *IEEE Digital Signal Processing Workshop, 1996, pp. 514-517.*

[24] S. Osowski and T. H. Linh "ECG Beat Recognition Using Fuzzy Hybrid Neural Network", *IEEE Trans. On Biomedical Engineering, Vol.48, No.11, 2001, pp.1265-1271.*

Ali Reza Sahab Received the B.S. in control engineering from KNT the university of technology, Tehran, Iran in 2001 and the M.S. and Ph.D. degrees in control engineering from Shahrood university of technology, Shahrood, Iran in 2003 and 2009 respectively.

He is a staff member of Electrical Group, Engineering Faculty of Islamic Azad University, Lahijan Branch. His research interests include nonlinear control and intelligent systems.

Yadollah Mehrzad Gilmalek Received the B.S. in electronic engineering from Islamic Azad University, Lahijan Branch in 2008 and the M.S. in electronic engineering from Islamic Azad University, Arak Branch in 2010 respectively.