VOL. 6, NO. 10, OCTOBER 2011 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

FUZZY C-MEANS CLUSTERING, NEURAL NETWORK, WT, AND HRV FOR CLASSIFICATION OF CARDIAC ARRHYTHMIA

A. Dallali, A. Kachouri and M. Samet
Laboratory of Electronics and Technology of Information (LETI), National School of Engineers of Sfax, Sfax, Tunisia
E-Mail: dallali_a@voila.fr

ABSTRACT

The classification of the electrocardiogram registration into different pathologies diseases devises is a complex pattern recognition task. The traditional methods of diagnosis and classification present some inconveniences; seen that the precision of credit note one diagnosis exact depends on the cardiologist experience and the rate concentration. Due to the high mortality rate of heart diseases, early detection and precise discrimination of ECG arrhythmia is essential for the treatment of patients. In this paper, a new cardiology system has been proposed for diagnosis, consultation, and treatment. The aim of this method is to help to practitioner doctor. During the recording of ECG signal, different forms of noise can be superimposed in the useful signal. This model consists of three subsystems. The first subsystem divides into suppression of base line and filtering the ECG recorded from different forms of noise that can be superimposed in the useful signal. The second subsystem realizes the extraction of RR interval using wavelet transform, and pre-classification based on FCMC technique. The third subsystem classifies the output clusters centers of the second using artificial neural network (ANN). In addition, FCMC-HRV is a new method proposed for classification of ECG. In this study, a combined classification system has been designed using fuzzy c-means clustering (FCMC) algorithm and neural networks. FCMC was used to improve performance of neural networks which was obtained very high performance accuracy to classify RR intervals of ECG signals. The ECG signals taken from MIT-BIH ECG database are used in training and testing data to classify four different arrhythmias (Atrial Fibrillation Termination). The test results suggest that HRV-FCMCNN structure can generalize better and is faster than other structures. Correct classification rate was found as 99.99% using proposed combination of Fuzzy C-Means Clustering Neural Networks (FCMCNN) method.

Keywords: cardiac arrhythmia, fuzzy C-means clustering, WT, HRV, MCN, neural network, classification.

1. INTRODUCTION

Electrocardiography deals with the electrical activity of the central of the blood circulatory system, i.e., the heart. Monitored by placing sensors at the limb extremities of the subject, electrocardiogram (ECG) is a record of the origin and the propagation of the electrical potential through cardiac muscles [1]. Thus, ECG offers cardiologists with useful information about the rhythm and functioning of the heart. Therefore, its analysis represents an efficient way to detect and treat different kinds of cardiac diseases. [2].

The state of cardiac heart is generally reflected in the shape of ECG waveform and heart rate. A typical structure of the ECG signal is shown in Figure-1. It may contain important pointers to the nature of diseases afflicting the heart. Early and quick detection and classification of ECG arrhythmia are important, especially for the treatment of patients in the intensive care unit. In the last four decades, computer-aided diagnostic (CAD) systems have been applied to the classification of the ECG resulting in several techniques [1-3]. This is mainly due to the fact that ECG signal provides cardiologists with useful and important information concerning the dysfunctions and physical condition of human heart. In designing of CAD system, the most important is the integration of suitable features extractor and pattern classifier such that they can operate in coordination to make an effective and efficient system [2].

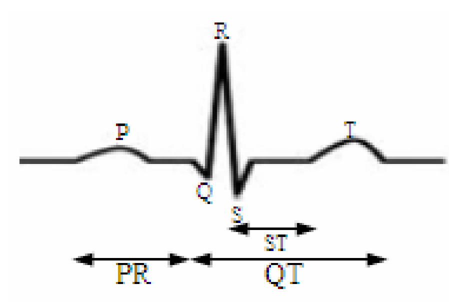

Figure-1. Structure of the ECG signal.

Up to now; numerous algorithms have been developed in the literature for the recognition and classification of ECG signal. Some of them use time, some use frequency, and other use frequency-level domain for depiction. The ECG waveforms may be different for the same patient to such extent that they are unlike each other and at the same time alike for different types of beats [1]. Artificial neural network [4-5], and fuzzy-logic based technique [6], were also employed to exploit their natural ability in pattern recognition task for successful classification of ECG beats. The hybrid system of neural network and fuzzy logic has been widely accepted for pattern recognition tasks. Yu *et al*., have implemented the integration of independent component analysis and neural network classifier (ICA - NN) along with R-R intervals to discriminate eight types of ECG beats [7].

In [8], Ozbay *et al*., had combined principal component analysis with neural network (PCA -NN) and compared it with wavelet transform technique for ECG signal classification. In [9], T.M. Nazmy had combined ICA and hybrid system (ICA - ANFIS) for ECG signal classification.

In this paper, the approach to ECG beat classification presented thorough experimental exploration of the FCMCNN capabilities for ECG classification. Further the performances of the FCMCNN approach in terms of classification accuracy are evaluated. In our previous works, we had showed clearly that neural network and fuzzy system with feature extraction methods had better performances than the traditional clustering and statistical methods.

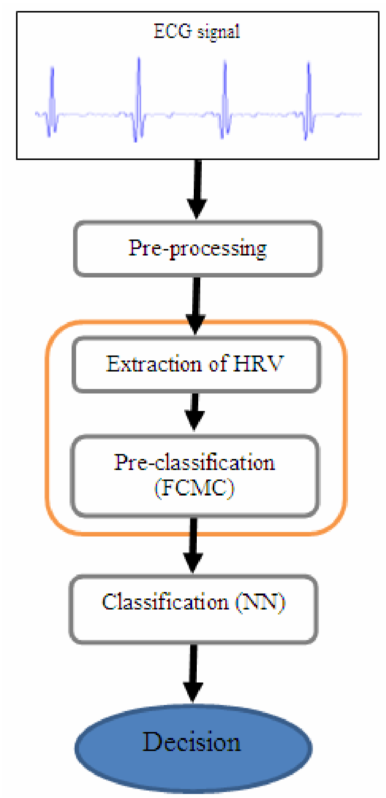

Figure-2. Block diagram of proposed arrhythmia classifier.

All the samples must be normalized in order to have the features at the same level. ECG signals can be contaminated with several kinds of noise, such as power line interference (A/C), baseline wandering (BW), and electromyography noise (EMG), which can affect the extraction of parameters used for classification, so we want to filter the signal. The unwanted noise of the signal must be removed. ECG was filtered using Low pass filter and high pass filter. The pre-treatment of ECG signals imposes the suppression of each perturbation signals, the noise high frequency electromyography and the low frequency drift. After that, the signal baseline may be shifted from zero line. The baseline of the ECG signal was adjusted at zero line by subtracting the median of the ECG signal.

2. LITERATURE STUDY

In the literature survey, several methods have been proposed for the automatic classification of ECG signals.

Among the most recently published works are those presented as follows. CUIWEI Li *et al*., (1995) showed that it is easy with wavelets transform decomposition to characterize the ECG waves. The difference between QRS, P and T waves, noise, baseline drift and interference were recognized. Using a raised cosine wavelet transform, Khadra *et al*., (1997) have undertaken a preliminary investigation of three arrhythmias-ventricular fibrillation (VF), ventricular tachycardia (VT) and atrial fibrillation (AF). They developed an algorithm based on the scale-dependent energy content of the wavelet decomposition to classify the arrhythmias, distinguishing them from each other and normal sinus rhythm. Again this study involved low numbers of data: 13 VF, 12 VT, 13 AF and 8 normal sinus rhythms. A. Kachouri, M. Ben Messaoud and A.Dallali (2003) compared wavelet transform for recognizing cardiac patterns. The choices of the wavelet family as well as the selection of the analyzing function into these families have been discussed; this study comprises four wavelet families: Daubechies, coiflet, haar, and symmetric. The choice of level decomposition has the same importance as the selection of wavelet function.

Using a support vector machines methodology, MG Tsipouras and al (2004), had classify cardiac arrhythmia. The results indicate high classification ability of the proposed method. The largest misclassification rates are between the PVC and NSR classes. This is due to the fact that PVC and NSR classes present high similarity. The method utilizes only the intervals of the ECG; therefore it is faster and more unaffected by the presence of noise than other proposed methods but also classifies a relatively small number of cardiac beats. (2007), Ceylan, R. and Ozbay, Y. proposed a classification of ECG arrhythmias using neural network architecture based on techniques of FCM, PCA and WT. Accurate and computationally efficient means of classifying electrocardiography (ECG) arrhythmias has been the subject of considerable research effort in recent years. This study presents a comparative study of the classification accuracy of ECG signals using a well-known neural network architecture named multi-layered perceptron (MLP) with backpropagation training algorithm, and a new fuzzy clustering NN architecture (FCNN) for early diagnosis. The results suggest that proposed FCNN architecture can generalize better than ordinary MLP architecture and also learn better and faster. The advantage of the proposed structure is a result of decreasing the number of segments by grouping similar segments in training data with fuzzy c-means clustering.

VOL. 6, NO. 10, OCTOBER 2011 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

(2010), S. Karpagachelvi *et al*., have developed a combined support vector machine and relevance vector machine analysis method for the automatic classification of electrocardiogram signals. The obtained experimental results show that the use of the RVM approach for classifying ECG signals on account of their superior generalization capability as compared to traditional classification techniques. The results confirm that the RVM classification system substantially boosts the generalization capability achievable with the SVM classifier, and its robustness against the problem of limited training beat availability, which may characterize pathologies of rare occurrence.

3. MATERIALS AND METHODS

This paper presents using fuzzy c-means clustering algorithm to develop the performance of neural network classifier. In many pattern recognition applications, the task of partitioning a pattern set can be considered to be the result of clustering algorithms in which the cluster prototypes are estimated from the information of the pattern set. In many cases, it may be impossible to obtain exact knowledge from a given pattern set. For recognition of the ECG arrhythmias, different methods were presented in the literature, such as the MLP approach, LVQ. In this paper, we present the combination of different forms of fuzzy c-means clustering, neural network, and wavelet transform; named as WT - FCMC - NN.

3.1 Wavelet transform

Physiological signal used for diagnosis are frequently characterized by a non-stationary time behavior. For such patterns, time and frequency representations are desirable. The ECG signals are considered as representative signals of cardiac physiology, which are useful in diagnosing cardiac disorders. The wavelet transform (WT) provides very general and powerful techniques, which can be applied to many tasks in signal processing. WT can represent signals in different resolutions. The most important application is the ability to compute and manipulate data in compressed parameters. Thus, the ECG records can be compressed into a few useful parameters. These parameters can be used for recognition and diagnosis. The most favorable choice of types of wavelet functions and level for pre-processing is problem dependent. The smoothing feature of the Daubechies wavelet of order 3 made it more suitable to detect variation on the ECG signals [8]. Therefore, the wavelet coefficients were computed using the Daubechies wavelet of level 3 in the present work.

$$x(n) = D_{j,k}[x(n)] + A_{j,k}[x(n)], n \in Z \quad (1)$$

where $D_{j,k}$ represents the detailed signal at j level. Note that j controls the dilation or contraction of the scale function $\varphi(t)$ and k denotes the position of the wavelet function $\Psi(t)$, and n represents the sample number of the x (n). The frequency spectrum of the signal is classified into high frequency and low frequency for wavelet decomposition as the level increases (j = 1… 6).

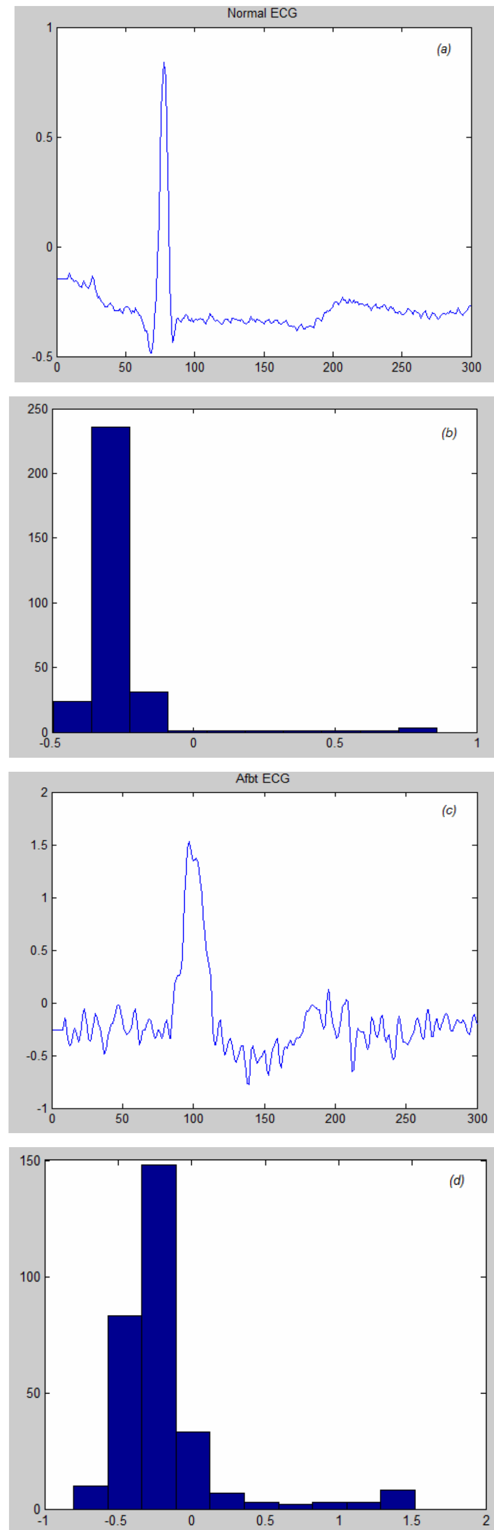

Figure-3. Top: Normal signal (a), and histogram of approximation after wavelet transform (b). Bottom: Same (c), but for signal with atrial fibrillation (d).

3.2 The fuzzy c-means clustering

The FCM algorithm has successfully been applied to a wide variety of clustering problems. The FCM algorithm attempts to partition a finite collection of elements $X = \{x_1, x_2 \ldots x_N\} \subset Rh$ where N represents the number of data vectors and h the dimension of each data vector, into a collection of C fuzzy clusters. C - Partition of X constitutes sets of (c. N) $\{u_{ij}\}$ member ship values can be conveniently arranged as a (c. N) matrix $u = [u_{ij}]$. The objective of fuzzy clustering is to find the optimum member ship matrix U. the most widely used objective function for fuzzy clustering is the weight within - groups sum of squared errors J_m, which is used to define the following constrained optimization problem [13].

$$J_m = \sum_{i=1}^{N} \sum_{j=1}^{c} u_{ij}^m \|x_i - c_i\|$$

Where $1 \leq m \leq \infty$.., i.e., m is any real number greater than 1, u_{ij} is the degree of member ship of x_i in the cluster j, x_i is the i[th] component of d-dimensional measured data, c_j is the d - dimension center of the cluster, and $\| \quad \|$ is any norm expressing the similarity between any measured data and the center. Fuzzy partition is carried out through an iterative optimization of the objective function shown above, with the update of member ship u_{ij}.

$$u_{ij} = \frac{1}{\sum_{k=1}^{c} \left(\left\| \frac{x_i - c_{ij}}{x_i - c_{jk}} \right\| \right)^{\frac{2}{m-1}}}$$

and the cluster cj by:

$$c_j = \frac{\sum_{i=1}^{N} u_{ij}^m . x_i}{\sum_{i=1}^{N} u_{ij}^m}$$

This iteration will stop when error$\{\|u_{ij}^{k+1} - u_{ij}^k\|\} \leq \varepsilon$, where ε is a termination criterion between 0 and 1; whereas k are the iteration steps. This procedure converges to a local minimum or a saddle point of Jm. The algorithm is composed of the following steps:

The steps are as follows.

I. Initialize the number of clusters (c), weighting exponent (m), iteration limit, termination criterion ($\varepsilon > 0$) and $U = [u_{ij}]$ matrix, U (0).

II. Guess initial position of cluster centers. $c^k = [c_j]$

III. At k step calculate the center vectors $c_j = \frac{\sum_{i=2}^{N} u_{ij}^m . x}{\sum_{i=2}^{N} u_{ij}^m}$

IV. Update U(k) to U(k+1) (2)

$$u_{ij} = \frac{1}{\sum_{k=1}^{c} \left(\left\| \frac{x_i - c_{ij}}{x_i - c_{jk}} \right\| \right)^{\frac{2}{m-1}}}$$

If $\{\|u_{ij}^{k+1} - u_{ij}^k\|\} \leq \varepsilon$, then stop; otherwise to step (i)

3.3 Artificial neural network (ANN)

Artificial Neural Network is biologically inspired network that are suitable for classification of biomedical signal. A combination of wavelets transform, FCMC and NNs is proposed to classify cardiac arrhythmias. The precision of classification results of the anomalies depends on the number of parameters selected; the number of neurons of input layer is equals to the numbers of FCMC clusters centers.

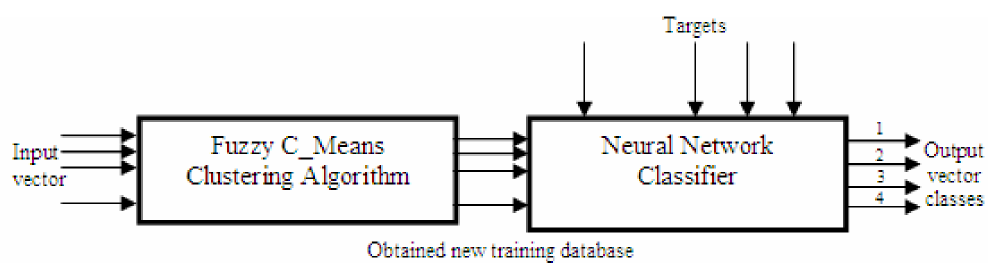

Figure-4. Structure of WT-FCMCNN classifier.

In this study, a three-layered feed-forward (MLP) neural network architecture was utilized and trained with the error back propagation algorithm with Log-sigmoid activation function that using the generalized back propagation for training which minimize the squared error between the desired outputs and the actual outputs of the ANNs. The desired output is being a real number in the interval [0-1]. The Back propagation algorithm is described step by step as [18]:

a) **Initialization:** Set all the weights and biases to small real random values.

b) **Presentation of input and desired outputs:** Present the input vector x(1), x(2),…,x(N) and corresponding desired response d(1),d(2),…,d(N), one pair at a time, where N is the number of training patterns.

c) **Calculation of actual outputs:** Use Equation (8) to calculate the output signals $y_1, y_2, \ldots, y_N$

$$y_i = \varphi\left(\sum_{j=1}^{N_{M-1}} w_{i,j}^{(M-1)} . x_j^{(M-1)} + b_i^{(M-1)}\right), \quad i = 1, \ldots, N_{M-1} \qquad (8)$$

d) **Adaptation of weights (wij) and biases (bi):**

$$\Delta w_{ij}^{(l-1)}(n) = \mu . x_j(n) . \delta_i^{(l-1)}(n) \qquad (9)$$

VOL. 6, NO. 10, OCTOBER 2011 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

$$\Delta b_i^{(l-1)}(n) = \mu.\delta_i^{(l-1)}(n) \tag{10}$$

And

$$\delta_i^{(l-1)}(n) = \begin{cases} \varphi'(net^{(l-1)})[d_i - y_i(n)], & l = M \\ \varphi'(net^{(l-1)})\sum_k w_{ki}.\delta_i^{(l)}(n), & 1 \le l \le M \end{cases} \tag{11}$$

where $x_j(n)$= output of node j at iteration n, l is layer, k is the number of output nodes of neural network, M is output layer, φ is activation function. The learning rate is demonstrated by μ. It may be noted here that a large value of the learning rate may lead to faster convergence but may also result in oscillation. With the purpose of achieving faster convergence with minimum oscillation, a momentum term may be added to the basic weight updating equation. The training of the artificial neural network ends if the sum of the square errors for all segments is less than 0.01. The number of data set used for training and testing of the ANNs classifier and the results obtained are tabulated in Table-2. The parameters extracted are used as inputs to system and the classification is done. The output of the classifier is a graphical representation.

4. TEST RESULTS OF NUMERICAL EXPERIMENTS

4.1 Dataset description

The experiment conducted on the basis of ECG data from the MIT-BIH arrhythmia database. ECG signal is filtered with low pass and high pass filters. After that signal is preprocessed to detect RR intervals; which is calculated from the location of the R points, of ECG signals which were sampled at 360 Hz. The detailed numbers of training and test beats are reported for each class in Table-1.

Table-1.
Numbers of training and test beats used in the experiments.

Type	MIT - BIH data base	Training file	Testing file
1	n01, n02, n03, n04	20	18
2	S01, s02, s03, s04	20	18
3	A01, a02, a03, a04	20	18
4	B01, b02, b03, b04	20	18
Total		**80**	**72**

4.2 Test results

In this paper, a fuzzy c-means clustering neural network was proposed to diagnose an abnormal heart beat on electrocardiogram. The improved FCMCNN was obtained by combination of fuzzy clustering layer and neural network classifier (Figure-2) we proposed an approach that the number of HRV segment in original training set was reduced by fuzzy c-means clustering algorithm and process of reducing was performed on each arrhythmia type individually.

Training max error was obtained as 1.3878e-016. Mis - Classification Number (MCN) represents number of misclassification ECG in testing. Rate of misclassification (RMC) is calculated using:

$$RMC(\%) = \left(\frac{Number\ of\ misclassification\ beat}{Number\ of\ total\ beat}\right) \tag{4}$$

The obtained results can be seen in Table-2. As it can be seen in this table, we found training and test errors calculated from tables according to the equations given in [8, 9].

Table-2. Pre-classification results for each arrhythmia in test.

Arrhythmia types of	Number	FCMC	
		MCN	RMC (%)
1	20	0	0
2	20	0	0
3	20	0	0
4	20	1	5
Total	**80**	**1**	**1.25**
Average test error (%)		0.0118	

The performance of WT - FCMC - NN technique is depicted as shown in Figure-6. It is observed that the % of recognition rates average is 99.99%.

ISSN 1819-6608

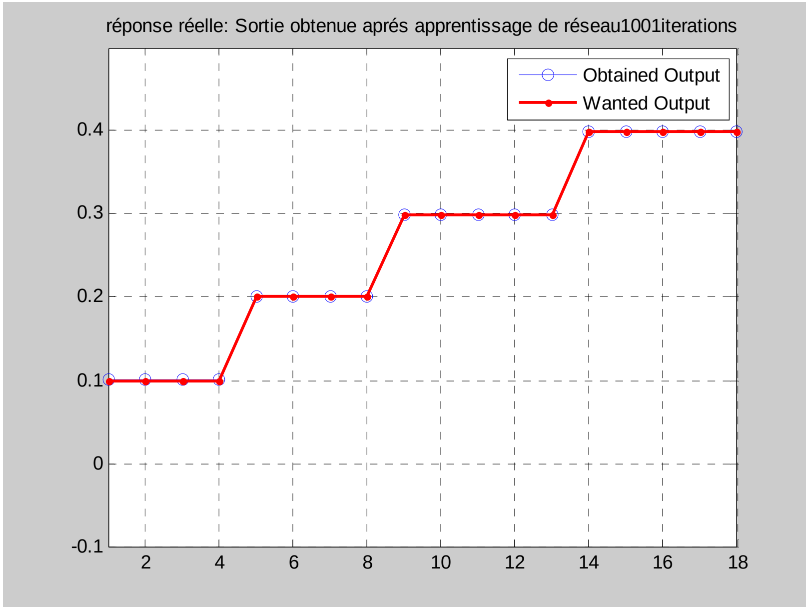

Figure-5. Classification results for different arrhythmias.

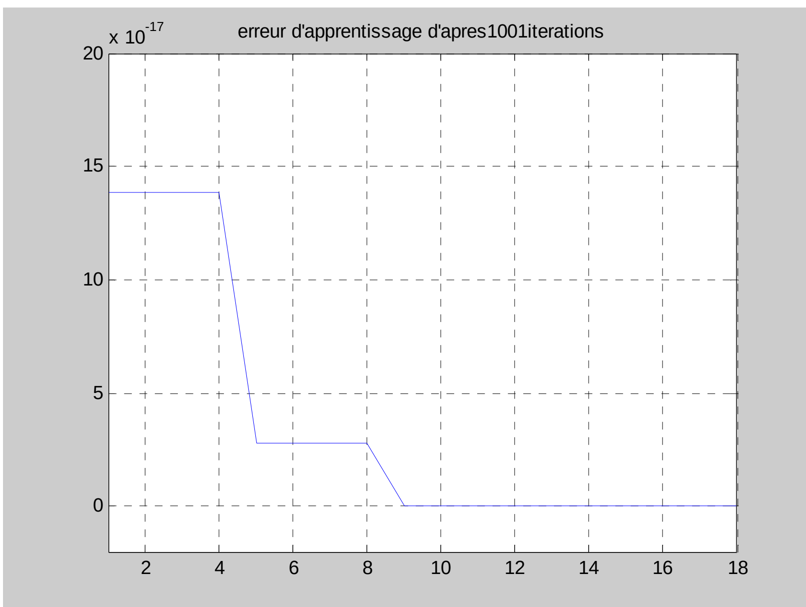

Figure-6. Error of classification for different arrhythmias.

5. CONCLUSIONS

In this piece of work, a novel ECG beat classification system using FCMC and NN are proposed and applied to MIT/BIH data base. The WT-FCMCNN was proposed and developed to classify electrocardiography signals. The wavelet transforms and HRV model parameters have been used for the features selection. For the conventional FCMC all patterns in the pattern space are assigned membership values, which are based on the Euclidean distance between the patterns to each cluster. The cluster centers obtained by FCMC are classified by neural network. The aim in developing WT - FCMCNN was to achieve more optimum cluster centers locations and to reduce the time of training of the neural network classifier.

This technique is obtained by incorporating the preprocessing method of ECG signal, extraction of RR intervals, fuzzy c-means clustering method, and neural network technique and combining their advantages for the classification of ECG arrhythmias. So, it can be said that the structure, which is a widely beneficial structure than conventional WT - NN to recognize and classify ECG signals, is obtained.

The results obtained confirm that the NN classification system substantially boosts the generalization capability achievable with the FCMC classifier. It can also be seen that NN accomplishes better and more balanced classification for individual categories as well in very less training time comparative to FCMC.

REFERENCES

[1] Owis M. I., Youssef A.-B. M. and Kadah Y. M. 2002. Characterization of ECG signals based on blind source separation. Medical and Biological Engineering and Computing. 40: 557-564.

[2] Rahime Ceylan, Yuksel Ozbay and Bekir Karlik. 2009. A novel approach for classification of ECG arrhythmias: Type-2 fuzzy clustering neural network. Expert Systems with Applications. pp. 6721-6726.

[3] Meau Y.P., Ibrahim F., Naroinasamy S.A.L. and Omar R. 2006. Intelligent Classification of Electrocardiogram (ECG) Signal Using Extended Kalman Filter (EKF) Based Neuro Fuzzy System. Elsevier Science Computer Methods and Programs in Biomedicine. 82: 157-168.

[4] A. Kachouri, M. Ben Messaoud and A. Dallali. 2003. Wavelet based on Electrocardiogram Signal Analyses for Classification and Diagnosis by Neural Networks. 2nd International Conference on Signals Systems Decision and Information Technology SSD'03. March 26-28, Sousse, Tunisia.

[5] S.Y. Foo, G. Stuart, B. Harvey and A. Meyer-Baese. 2002. Neural network-based ECG pattern recognition. Eng. Appl. Artif. Intell. 15: 253-260.

[6] Hu Y.H., Palreddy S. and Tompkins W. 1997. A patient adaptable ECG beat classifier using a mixture of experts approach. IEEE Trans. Biomed. Eng. 44(9): 891-900.

[7] Yu S. N. and Chou K.T. 2006. Combining Independent Component Analysis and Back propagation Neural Network for ECG Beat Classification. Proceedings of the 28th IEEE EMBS Annual International Conference, New York City, USA, Aug.30-Sept.3. pp. 3090-3093.

[8] Ceylan R. and Ozbay Y. 2007. Comparison of FCM, PCA and WT techniques for classification ECG arrhythmias using artificial neural network. Expert Systems with Applications. 33: 286-295.

[9] T. M. Nazmy, H. El-Messiry and B. Al-Bokhity. 2010. Adaptive Neuro-Fuzzy Inference System for Classification of ECG Signals. Journal of Theoretical and Applied Information Technology, JATIT. 12(2): 71-76.

[10] M. I. Owis, A. B. M. Youssef and Y. M. Kadah. 2002. Characterisation of electrocardiogram signals based on blind source separation. Medical and Biological Engineering and Computing, Cairo, Egypt.

[11] Yu S.N. and Chou K.T. 2008. Integration of independent component analysis and neural networks for ECG beat classification. Expert Systems with Applications. 34: 2841-2846.

[12] Guler I. and Ubeyli E. D. 2005. ECG beat classifier designed by combined neural network model. Pattern Recognition. 38: 199-208.

[13] Dipti Patra and Smita Pradhan. 2009. Integration of FCM, PCA and Neural networks for classification of ECG Arrhythmias. IAENG International Journal of Computer Science, IJCS_36_3_05, 01 August. 36:3.

[14] E. M. Tamil, N. H. Kamarudin, R. Salleh, M. Yamani Idna Idris, M. N. Noorzaily and A. M. Tamil. 2008. Heartbeat electrocardiogram (ECG) signal feature extraction using discrete wavelet transforms (DWT). In: Proceedings of CSPA. pp. 1112-1117.

[15] B.U. Kohler, C. Hennig and R. Orglmeister. 2002. The principles of software QRS detection. IEEE Engineering in Medicine and Biology Magazine. 21(1): 42-57. January-February.

[16] Yu S. N. and Chen Y. H. 2007. Electrocardiogram beat classification based on wavelet transformation and probabilistic neural network. Pattern Recognition Letters. 28: 1142-1150.

[17] B. Anuradha and V.C. Veera Reddy June 2008. ANN for Classification of Cardiac Arrhythmias. ARPN Journal of Engineering and Applied Sciences. 3(3): 1-6. June.

[18] Srinivasca K. G., Amrinder Singh, A. O. Thomas, Venugopal K. R. and L. M Patnaik. Generic Feature Extraction for Classification using Fuzzy C - Means Clustering. 0- 7803 - 9588 - 3/05/$20.00 ©2005 IEEE.

[19] S. Z. Mahmoodabadi, A. Ahmadian and M. D. Abolhasani. 2005. ECG Feature Extraction using Daubechies Wavelets. Proceedings of the 5th IASTED International conference on Visualization, Imaging and Image Processing. pp. 343-348.