IAENG International Journal of Computer Science, 36:3, IJCS_36_3_05

Integration of FCM, PCA and Neural Networks for Classification of ECG Arrhythmias

Dipti Patra, *Member IEEE*, Manab Kumar Das, Smita Pradhan

Abstract—The classification of the electrocardiogram (ECG) into different patho-physiological disease categories is a complex pattern recognition task. In this paper, we propose a scheme to integrate fuzzy c-means (FCM) clustering, principal component analysis (PCA) and neural networks (NN) for ECG beat classification. The PCA is used to decompose ECG signals into weighted sum of basic components that are statistically mutual independent. In addition, FCM clustering is among considerable techniques for data reduction. A back propagation neural network (BPNN) is employed as classifier. ECG samples attributing to six different beat types are sampled from the MIT-BIH arrhythmias database for experiments. In this paper comparative study of performance of four structures such as FCM-NN, PCA-NN, FCM-ICA-NN, and FCM-PCA-NN are investigated. The fuzzy self organizing layer performs the pre-classification task and it is among considerable techniques for data reduction. The aim of using fuzzy c-means (FCM) is to decrease the number of segments by grouping similar segments training data. The features of obtained clustered training patterns are extracted using principal component analysis (PCA).This layer performs elimination of inconsiderable features with PCA. The test results suggest that FCM-PCA-NN structure can perform better and faster than other techniques.

***Index Terms*—Principal Component Analysis, Fuzzy C-means Clustering, ECG, Arrhythmias, Artificial Neural Network.**

I. Introduction

Pattern recognition applied to biomedical signal processing is conventionally based on measurements directly taken from the time domain sampled data. This can lead to a redundant representation of the signal information. ECG classification is one of them [1, 2]. Due to the mortality rate of heart disease,
early detection and precise discrimination of ECG arrhythmias is essential for the treatment of patients. This leads to high-precision computer aided diagnosis (CAD) systems for ECG. An effective CAD system requires a powerful pattern classifier as well as feature extractor that is capable of extracting important information from the raw data. The most difficult problem faced by today's automatic ECG analysis is the large variation in the morphologies of ECG waveforms of different patients as well as in the same patients. The ECG waveforms may differ for the same patient at different time and may alike for different patient having different types of beats. This causes the beat classifiers performing well on the training data to perform poorly with ECG waveforms of different patients.

In last decade, a number of researchers have proposed various methods for ECG beat classification using neural network classifier [1,3,4]. Though conventional back-propagation neural networks (BPNN) is very popular among them which is able to recognize and classify ECG signals more accurately, but it suffers from slow convergence to local and global minima. In order to overcome this problem, many hybrid neural networks are proposed by researchers [5,7,8].

Owis *et al.* had applied independent component analysis (ICA) to extract the independent components of ECG signals in Fourier domain. These components are used for feature extraction to differentiate ECG beat types [6]. In [7], Yu *et al.* have implemented the integration of independent component analysis and neural network classifiers (ICA-NN) along with R-R intervals to discriminate eight types of ECG beats. Only one bank of independent components is calculated in the time domain and serves as bases to constitute the subspace for ECG signal representation. The capabilities of the neural networks in coordinate with the ICA features are justified. In [8], Ozbay *et al.* had combined PCA with neural network and compared with wavelet transform technique for ECG signal classification. In [9], Patra *et al.* had combined FCM, ICA, NN and compared with the structure ICA-NN for the same. The FCM-ICA-NN integrated structure could perform faster and more accurate than ICA-NN. Acin *et al.* had developed a support vector machine classifier for ECG beat classification designed by perturbation method [10]. The dimension of each feature set is reduced by using perturbation method and input dimension size could be reduced by their design. In [11], Engin *et al.* had used three statistical features which are third order cumulate, wavelet entropy and auto-regressive coefficients. They had applied these new features to different statistical classifiers for ECG beat recognition.

In this paper, we evaluate the integration of FCM-PCA-NN to discriminate six types of ECG beats. The proposed structure is composed of three sub networks: fuzzy classifier, layer of feature extraction with Principal component analysis, and classification by neural networks. The fuzzy self-organizing layer performs the

This work was supported by the Electrical Engineering Department of National Institute of Technology, Rourkela, India.

Dr. Dipti Patra is Associate Prof. with the Electrical Engineering Department, National Institute of Technology, Rourkela, India-769008 (phone: 91-661-2462410); e-mail: dpatra@ nitrkl.ac.in.

Manab Kumar Das is Senior Lecturer with Electronics and Communication Engineering Department in Mallabhum Institute of Technology, Braja Radhanagar, India-722122; email: manabster@gmail.com.

Smita Pradhan is research scholar of image processing and computer vision lab, Electrical Engg. Department of National Institute of Technology, Rourkela, (phone: 91-661-2464407); email: ssmita.pradhan@gmail.com.

(Revised online publication: 3 February 2010)

pre-classification task; layer of feature extraction performs the extraction of features with Principal component analysis, and the multilayer perceptron works as a final classifier. Back-propagation neural network (BPNN) is employed as the beat classifier in this study. The fuzzy stage is responsible for the analysis of the distribution of data and grouping them into clusters with different membership values. In the next stage, features of obtained clustered training patterns are extracted using principal component analysis. On the basis of these components, the MLP neural network classifies the applied input vector, representing the heart beat to the appropriate class. On the other hand, a number of segments in training patterns are reduced using fuzzy c-means clustering in fuzzy self-organizing layer and number of samples of clustered training patterns is decreased. Using PCA, principal components are obtained which are considered as the inputs to Neural Network (NN) classifier. Due to data reduction in fuzzy self organizing layer, training period of the neural network is decreased. In this paper the proposed method was used for electrocardiograph beat recognition and classification. The performance of the proposed structure FCM-PCA-NN is found to be more generalized, accurate and faster than the existing techniques presented in [7] and [9].

II. ECG BEAT CLASSIFICATION METHODS

For recognition of the ECG waveform type, different solutions were presented in the literature, such as the MLP approach, self organizing map and the LVQ. We present the combination of the fuzzy self-organizing layer, principal component analysis and the MLP connected in cascade, named as the FCM-PCA-NN and then compare this technique with the existing techniques. The self-organizing layer is responsible for the clustering of the input data. However, it is the fuzzy clustering, in which the input vector x is pre-classified to all sets with different membership values. The penetration of the data space is better and the localization of the input vector x in the data space is more precise. The outputs of all self organizing neurons (the cluster centers) form the input vector to independent component analysis. The output data vectors of independent component analysis form the input vector to the third subnet work (MLP). MLP sub-network is responsible for the final classification of the ECG beat.

A. *The fuzzy c-means clustering*

Structure identification of fuzzy systems is possible by constructing enough rules with appropriate input and output membership functions. The identified model can then be used to describe the behavior of the target system as well as for prediction purpose. In this paper, we have trained the fuzzy layer by using the fuzzy c-means clustering algorithm. The idea of fuzzy clustering is to divide the data into fuzzy partitions that overlap with one another. Therefore, the inclusion of data in a cluster is defined by a membership grade in [0, 1]. Formally, clustering an unlabeled data $X = \{X_1, X_2, \ldots, X_N\} \subset R^h$ where N represents the number of data vectors and h the dimension of each data vector, is the assignment of c-partition labels to the vectors in X. c-partition of X constitutes sets of (c.N) $\{u_{ij}\}$ membership values that can be conveniently arranged as a (cxN) matrix U = $[u_{ij}]$. The fuzzy clustering finds the optimum membership matrix U. The most widely used objective function for fuzzy clustering is the weighted within-groups sum of squared errors J_m, which is used to define the following constrained optimization problem [9,10].

$$J_m = \sum_{i=1}^{N} \sum_{j=1}^{C} u_{ij}^m \left\| x_i - c_j \right\| \qquad (1)$$

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

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

and the cluster centers c_j 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 J_m. The algorithm is composed of the following steps:

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

ii. Guess initial position of cluster centers.

iii. At k step calculate the center vectors $C^k = [C_j]$

$$c_j = \frac{\sum_{i=2}^{N} u_{ij}^m x}{\sum_{i=2}^{N} u_{ij}^m}$$

iv. Update $U^{(k)} to U^{(k+1)}$

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

If $\left\| U^{(k+1)} - U^{(k)} \right\| \leq \varepsilon$ then STOP ; Otherwise return to step (i).

B. *Principal Component Analysis*

PCA is a well-established technique for feature extraction and dimensionality reduction. It is based on the assumption that most information about classes is contained in the

directions along which the variations are the largest. The most common derivation of PCA is in terms of a standardized linear projection which maximizes the variance in the projected space. PCA is useful for data compression, by reducing the number of dimensions, without much loss of information.
The steps are as follows.

Step1: Get the input data.

Step2: Calculate the mean from the data set.

Step3: Subtract the mean.

Step4: Calculate the Covariance from the equation:

$$\operatorname{cov}(x,y)=\frac{\sum_{i=1}^{n}(X_i-\bar{X})(X_i-\bar{X})}{(n-1)}$$

where $\bar{X}$ is the mean of data sheet. Calculate the Covariance matrix.

Step5: Calculate the Eigen Vectors and Eigen Value of the Covariance matrix.

Step6: Choosing components and forming a feature vector.

Step7: Deriving the new data set by the following formula
Final data = Row Feature Vector * Row Data Adjust

where Row Feature Vector is the matrix with eigenvectors in the columns transpose and Row data Adjust is the mean adjusted transpose. An assumption made for feature extraction and dimensionality reduction by PCA is that most information of the observation vectors is contained in the subspace spanned by the first m principal axes, where m < p for a p-dimensional data space. Therefore, each original data vector can be represented by its principal component vector with dimensionality m.

C. Multilayer Perceptron

In this work, a three-layered feed-forward NN was used as the classifier and trained with the error back propagation. The input signals of NN were formed by cluster centers that generalized by fuzzy c-means clustering. The back propagation training with generalized delta learning rule is an iterative gradient algorithm designed to minimize the root mean square error between the actual output of a multilayered feed-forward NN and a desired output. Each layer is fully connected to the previous layer, and has no other connection.
Back-propagation algorithm:

1. Initialization: Set all the weights and biases to small real random values.
2. 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 training of training patterns.
3. Calculation of actual outputs: Use Eq. (3) to calculate the output signals $y_1, y_2, \ldots, y_{NM}$

$$y_i=\varphi\left(\sum_{j=1}^{N_{M-1}} w_{ij}^{(M-1)}x_j^{(M-1)}+b_i^{(M-1)}\right),\ i=1,\ldots,N_{M-1} \qquad (3)$$

4. Adaptation of weights (w_{ij}) and biases (b_i)

$$\Delta w_{ij}^{(l-1)}(n)=\mu\cdot x_j(n)\cdot\delta_i^{(l-1)}(n)$$
$$\Delta b_i^{(l-1)}(n)=\mu\cdot\delta_i^{(l-1)}(n)$$

where

$$\delta_i^{(l-1)}(n)=\begin{cases}\varphi'\left(net_i^{(l-1)}\right)\left[d_i-y_i(n)\right], & l=M\\ \varphi'\left(net_i^{(l-1)}\right)\sum_k w_{ki}\cdot\delta_k^{(l)}(n), & 1\le l\le M\end{cases}$$

In which $x_j(n)$=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 the activation function. The learning rate is represented 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. In order to achieve faster convergence with minimum oscillation, a momentum term may be added to the basic weight updating equation. After completing the training procedure of the neural network, the weights of MLP are frozen and ready for use in the testing mode.

III. Proposed Method

The method is divided into four steps: (i) ECG sampling and preprocessing, (ii) data reduction, (iii) calculation of feature vector, and (iv) classification by neural networks, which are described as follows.

A. ECG sampling and preprocessing

Table 1: ECG samples selected from MIT-BIH database

Type	MIT-BIH data file	Training (file)	Testing (file)
NORM	100,101,108,105	8	8
LBBB	109,214,111,207	8	8
RBBB	118,124,212,231	8	8
PVC	106,119,200,213	8	8
APB	209,222,220,223	9	9
PB	102,104,107,217	9	9
Total		50	50

For our experimental study, ECG signal samples are obtained from MIT-BIH arrhythmia database. The portions of ECG signal before and after R peak are considered for processing as this portion contains the most important information about the function of heart. ECG signals are sampled at 360 Hz taking 200 samples in the intervals of R-R for all arrhythmias which are called as a segment. We have selected 50 sample segments attributing to six ECG beat types from MIT-BIH arrhythmia database. The six beat types selected for our experiment are normal beat (NORM), left bundle branch block beat (LBBB), right bundle branch block beat (RBBB), artial premature beat (APB), premature ventricular contraction (PVC), paced beat (PB). The ECG beat types selected are presented in Table 1. In each category, half of the ECG beats are selected for training and other half for testing purpose. The locations of the R points provided in the annotation files of the source data is the key point for evaluation of R-R interval. After data sampling, the preprocessing of the samples is done by normalizing with zero mean and unity standard deviation in order to reduce the effect of bias due to signal amplitude.

B. ECG Data reduction and calculation of feature vector

In the first step of our proposed structure, training patterns are clustered using fuzzy c-means (FCM) algorithm. By this process, number of segments in each type of arrhythmia are reduced. Reduction of segments are made separately for each type of arrhythmias. In next step, the number of samples of training patterns, which are obtained to decrease the number of segments of training patters with FCM clustering are reduced to make feature extraction with principal component analysis (PCA). The number of input nodes of neural network is equal to the optimum number of principal components extracted in this step.

C. Classification by neural networks

The back-propagation neural network (BPNN) used in this study is a three-layer feed-forward structure. The first layer is the input layer that has the PCA features inputs. The second layer, also called the hidden layer, has 40 neurons and the output layer has six neurons, which is equal to the number of ECG beat types to be classified. There are three commonly used activation functions for the back-propagation multiplayer neural network, namely the logistic function, hyperbolic tangent function, and identity function. In this study, the hyperbolic tangent functions are used. The weight and bias values in the BPNN are updated with a learning rate of 0.02.

IV. SIMULATIONS AND RESULTS

We randomly select two samples from each of the records in the database. The selected samples constitute a 50x200 data matrix. In our proposed FCM-PCA-NN structure, the number of samples of training pattern, which were obtained to reduce the number of segments with fuzzy c-means clustering, was reduced to 38 segments.

The performance of FCM-PCA-NN technique is depicted as shown in Fig. 1. It is observed that the % of error in case of FCM-PCA-NN structure is least in comparison to FCM-ICA-NN, FCM-NN and PCA-NN structures keeping the number of iterations same. It also shows that % of accuracy is higher in our proposed technique FCM-PCA-NN which is comparable with the technique FCM-ICA-NN. The training and test results are also depicted for all structures in Table 2. It indicates that the training time in case of FCM- PCA-NN is 656.427 sec. On the other hand FCM-ICA-NN, PCA-NN and FCM-NN structures have the training time 2765.69 sec, 1525.57 sec and 1624.75 sec respectively which is much more than that of FCM-PCA-NN structure, keeping the % of average training error constant at $3.99x10^{-4}$. Further it is observed that the % of average test error is also less. Above all the number of iterations in case of FCM-PCA-NN is only 3028 as compared to all other existing structures. It is observed that FCM-PCA-NN is performing superior to other structures with high recognition rate at shorter time.

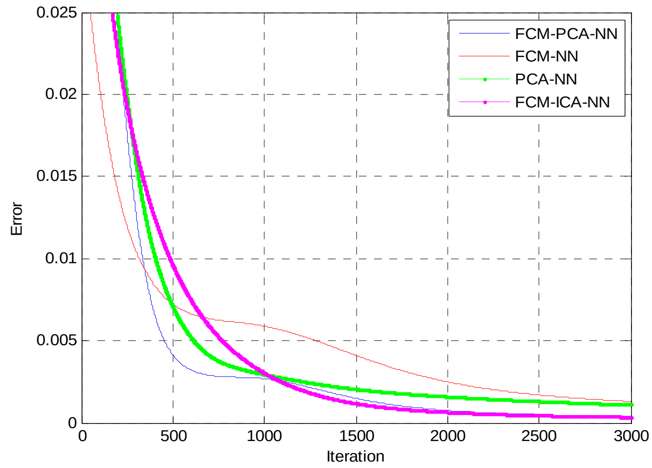

Fig. 1. Convergence performance of different techniques

Table 2: ECG signal classification data sheet for different techniques

	Archi-tecture	Size of training set	Average Training Error(%)	Averag -e Test Error (%)	No of Iterati ons	Training Time in Sec
1	FCM-ICA-NN	200x38	$3.99x10^{-4}$	0.45	12728	2765.69
2	FCM-PCA-NN	37x38	$3.99x10^{-4}$	0.13	3028	656.42
3	FCM-NN	200x38	$3.99x10^{-4}$	0.25	7087	1624.75
4	PCA-NN	37x50	$3.99x10^{-4}$	0.21	6883	1525.57

V. CONCLUSIONS

In this piece of work, a novel technique is developed for ECG beat classification. In this context, we have considered a neural network classifier for beat classification. In order to overcome the difficulty of intensive computational time taken using NN classifier, attempt has been made to reduce the number of input data points using FCM and PCA which is beneficial for automatic ECG beat classification in real time mode. The aim of using fuzzy c-means (FCM) is to decrease the number of segments by grouping similar segments training data. The features of obtained clustered training patterns are extracted using principal component analysis (PCA) which further reduces the total number of input nodes of the NN classifier. The proposed FCM-PCA-NN has been presented to classify electrocardiography signals. A comparative assessment of the performance of FCM-PCA-NN with FCM-NN, PCA-NN and FCM-ICA-NN shows that proposed hybrid classifier structure FCM-PCA-NN performs faster with high recognition rate of ECG beats. The advantages of FCM and PCA are incorporated with NN classifier which could able to outperform some of the existing classifiers. The proposed structure enhances the performance to recognize and classify ECG beats in terms of faster rate and better accuracy.

REFERENCES

[1] Acharya, R., Bhat, P. S., Iyengar, S. S., Roo, A. & Dua, S., "Classification of heart rate data using artificial neural network and fuzzy equivalence relation," *The Journal of the Pattern Recognition Society,* 2002.

[2] De Chazal, P., & Reilly, R. B., " Automatic classification of ECG beats using waveform shape and heart beat interval features," In *IEEE international conference on acoustic, speech and signal processing (ICASSP'03)*, vol. 2, pp. 269-272, Hong Kong, China, 2003.

[3] Osowski, S. & Linh, T. H., "ECG beat recognition using fuzzy hybrid neural network," *IEEE Transaction on Biomedical Engineering*, vol. 48, no. 11, pp. 1265–1271, 2001.

[4] Guler, I. & Ubeyli, E. D., "ECG beat classifier designed by combined neural network model," *Pattern Recognition*, vol. 38, pp. 199-208, 2005.

[5] Engin, M., "ECG beat classification using neuro-fuzzy network," *Pattern Recognition Letters*, vol. 25, pp. 1715-1722, 2004.

[6] Owis, M. I., Youssef, A.-B. M., & Kadah, Y. M., "Characterization of ECG signals based on blind source separation," *Medical and Biological Engineering and Computing*, vol. 40, pp. 557–564, 2002.

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

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

[9] Dipti Patra and Smita Pradhan, "Integration of FCM, ICA and Neural networks for ECG signal classification," In *Proceedings of International Conference on Soft Computing (ICSC 2008)*, Alwar, India, Nov. 2008.

[10] Acır, N., "A support vector machine classifier algorithm based on a perturbation method and its application to ECG beat recognition systems," *Expert Systems with Applications*, vol. 31, no. 1, pp.150-158, 2006.

[11] Engin, M., Fedakar, M., Engin, E.Z., Korürek, M., "Feature measurements of ECG beats based on statistical classifiers," *Measurement*, vol. 40, no. 9, pp. 904-912, 2007.

(Revised online publication: 3 February 2010)