Improving Ischemic Beat Classification Using Fuzzy-Genetic Based PCA And ICA

S. Murugan
Assistant Professor, Department of ICE, Arulmigu Kalasalingam College of Engineering,Krishnankoil-626190, Srivilliputhur, TamilNadu, India.

Dr.S. Radhakrishnan
Senior Professor and Head, Department of CSE, Arulmigu Kalasalingam College of Engineering,Krishnankoil-626190 Srivilliputhur, TamilNadu, India.

***Abstract*—In this paper, an improved version of Principal Component Analysis (PCA) and Independent Component Analysis (ICA) is proposed for feature extraction to classify the ischemic beats from electrocardiogram (ECG) signal. The Fuzzy C-Means (FCM) and Genetic Algorithm (GA) is combined with PCA and ICA to extract more relevant features; the proposed methods are named as Fuzzy-Genetic based PCA (FGPCA) and Fuzzy-Genetic based ICA (FGICA). Least Square Support Vector Machine (LSSVM) is used to classify the beats into ischemic or non-ischemic, with the features from the FGPCA and FGICA. The ECG beats used in this paper are collected from European ST-T database. There is totally 2040 beats extracted from 17 different patients. The performance of our proposed method is compared with the linear PCA and ICA, shown that the proposed methods improve the sensitivity of ischemic classification.**

Keywords-Principal Component Analysis, Independent Component Analysis, Fuzzy Logic, Genetic Algorithm.

I. INTRODUCTION

Electrocardiography is a significant tool in analyzing the condition of the heart. The electrocardiogram (ECG) is the record of discrepancy of bioelectric potential with respect to time as the human heart beats. It provides most valuable information about the functional characteristics of the heart and cardiovascular system. Early detection of heart diseases can protract life through appropriate treatment. Therefore, numerous research and work analyzing the ECG signals have been reported [1–5]. For effective diagnostics, the study of ECG pattern and heart rate variability signal may have to be carried out over several hours. Thus, the volume of the data being enormous, the study is tedious and time consuming. Naturally, the possibility of the analyst missing (or misreading) vital information is high. Therefore, computer-based analysis and classification of diseases can be very helpful in diagnostics [1–5].

Various methodologies of automated diagnosis have been adopted; however the entire process can generally be subdivided into a number of disjoint processing modules: beat detection, feature extraction/selection, and classification. In this paper, we have proposed a fuzzy-genetic based component analysis FGPCA (Fuzzy-Genetic based Principal Component Analysis) and FGICA Fuzzy-Genetic based Independent Component Analysis) for feature extraction from ECG signals. The major limitation in component analysis is, we can't expect that the chosen components have equal number of vectors from each class; in that case the overall performance may be decreased. Hence, it is better to choose equal number of components from each class. Here we have applied Fuzzy C-Means to cluster the components first, and then the principal components have been selected using Genetic Algorithm (GA). Least Square Support Vector Machine (LSSVM) is used for classification, since it is a robust and reliable classifier system and has the ability to perform fast classification. The computation time of LSSVM classifier is lower than the other classifier algorithms such as artificial neural network, decision tree and an artificial immune system.

Until now, ECG recordings that are used for the diagnosis of ischemic episodes are affected by noise, which deteriorates significantly the diagnostic accuracy. In this paper, the noises are removed by the method as discussed in [6, 7]. Better handling of the noisy ECGs can improve the accuracy of the diagnostic methods and increase their applications in every day practice. The performance of the system was analyzed with regard to the classification accuracy and we generated Receiver Operating Characteristic (ROC) curves to present our results. Our proposed system obtained 94% classification accuracy. This performance exceeds that of other studies applied to the ECG dataset classification problem so far.

The paper is organized as follows: the following section explains the standard PCA and ICA. Section 3 explains the fuzzy-genetic based feature selection. Section 4 presents the Least Square Support Vector Machine (LSSVM) architecture for beats classification. The experiments and results are discussed in Section 5. And the paper is concluded at Section 6.

II. FEATURE EXTRACTION

This section describes the feature extraction process from the beat signals extracted from the electrocardiograms using PCA and ICA.

A. Principal Component Analysis

Principal Components Analysis (PCA) is an exploratory multivariate statistical technique for simplifying complex data sets [8]. Given n observations on m variables, the goal of PCA

is to reduce the dimensionality of the data matrix by finding r new variables, where r is less than m. Principal components project high dimensional data into the subspace spanned by the eigenvectors with the r largest eigenvalues while remaining mutually uncorrelated and orthogonal. Each principal component is a linear combination of the original variables. The algorithm to obtain the Principal Components of a vector set X represented by a $X_{N\times M}$ matrix, where N represents the number of segments, and M represents the dimension of the vectors that constitute the vector set.

The algorithm of PCA is explained as below:

a. Obtain the Mean vector (Ψ): $\mu_x = \frac{1}{N}\sum_{i=0}^{N-1} x_i$

b. Obtain the Covariance Matrix:

$$C_x = \frac{1}{N}\sum_{i=0}^{N-1}(x_i - \mu_i)(x_i - \mu_i)^T \quad (1)$$

c. Obtain the eigenvectors and eigenvalues: $C_x e = \lambda e$ where e is eigenvector and λ is eigenvalue.

d. After creating the eigenspace we can proceed to recognition. Given a new beat of an individual Γ, the signals are concatenated the same way as the training, the mean vector Ψ is subtracted and the result is projected into the face space:

$$\omega_k = e_k^T(\Gamma - \Psi) \quad (2)$$

for k=1,..,M'. These calculated values of ω together form a vector $\Omega^T = [\omega_1, \omega_2, \ldots, \omega_{M'}]$. Ω is then used to establish which of the pre-defined classes best describes the new signal. The simplest way to determine class k that minimizes the Euclidian distance:

$$\varepsilon_k = \sqrt{\|\Omega - \Omega_k\|^2} \quad (3)$$

Where Ω_k is a vector describing the k^{th} signal class. A signal is classified as belonging to a certain class when the minimum ε_k (i.e. the maximum matching score) is below some certain threshold.

Choosing components and forming a feature vector: We get 2040 components corresponding to the dimensionality of the input sequence. Components that are significant from the point of view of contribution to the total energy of the signal are selected. The selected components together must constitute about 99% of the total energy of the signal. This procedure decreases the data dimensionality without significant loss of information. There are at least three proposed ways to eliminate eigenvectors.

- First is the mentioned elimination of eigenvalues with smallest eigenvalues. This can be accomplished by discarding the last 60% of total number of eigenvectors.
- The second way is to use the minimum number of eigenvectors to guarantee that energy E is greater than a threshold. A typical threshold is 0.9 (90% of total energy). If we define E_i as the energy of the i^{th} eigenvector, it is the ratio of the sum of all eigenvalues up to and including i over the sum of all the eigenvalues: where k is the total number of eigenvectors.

$$E_i = \left(\sum_{j=1}^{i}\lambda_j\right) \Big/ \left(\sum_{j=1}^{k}\lambda_j\right) \quad (4)$$

- The third variation depends upon the stretching dimension. The stretch for the i^{th} eigenvector is the ratio of that eigenvalue over the largest eigenvalue (λ_I):

$$\Sigma_\iota = \lambda_\iota / \lambda_I \quad (5)$$

- In our proposed method, Fuzzy-Genetic based algorithm is used to select the best eigenvectors.

B. Independent Component Analysis

Independent component analysis (ICA) is a statistics method that searches from multivariate statistical data for underlying factors or components that are statistically independent [9,10,11]. In recent years, ICA technique has found various applications, especially in feature extraction and blind source separation (BSS). However, the ICA procedure usually creates a great number of independent components (ICs) that are generated in an arbitrary order [9]. This property usually leads to the requirement for dimension reduction in the feature space. Moreover, with the random order of the ICs, it is tricky to determine the relative significance of each IC to be recruited in the task. Therefore, while applying ICA to solve problem, two basic questions must first be dealt with: (i) how many ICs are sufficient and (ii) what is the appropriate combination of ICs for the encountered problem. The solution of these questions belongs to the technique of subset feature selection that extract significant features of lower dimension to achieve approximately as well as, or sometimes better than, the performance by using the entire features.

To rigorously define ICA [12], we can use a statistical "latent variables" model. Assume that we observe n linear mixtures $x_1,\ldots,x_n$ of n independent components

$$x_j = a_{j1}s_1 + a_{j2}s_2 + \ldots + a_{jn}s_n, \text{ for all j.} \quad (6)$$

Where a_j represents the time variance and s represents the signals. We have now dropped the time index t; in the ICA model, we assume that each mixture x_j as well as each independent component s_k is a random variable, instead of a proper time signal. Without loss of generality, we can assume that both the mixture variables and the independent components have zero mean: If this is not true, then the

observable variables x_i can always be centered by subtracting the sample mean, which makes the model zero-mean.

It is convenient to use vector-matrix notation instead of the sums like in the previous equation. Let us denote by x the random vector whose elements are the mixtures $x_1,...,x_n$, and likewise by **s** the random vector with elements $s_1,...,s_n$. Let us denote by **A** the matrix with elements a_{ij}. Generally, bold lower case letters indicate vectors and bold upper-case letters denote matrices. All vectors are understood as column vectors; thus $\mathbf{x}^T$, or the transpose of **x**, is a row vector. Using this vector-matrix notation, the above mixing model is written as

$$\mathbf{x} = \mathbf{As} \tag{7}$$

Sometimes we need the columns of matrix **A**; denoting them by $\mathbf{a}_j$ the model can also be written as

$$x = \sum_{i=1}^{n} a_i s_i \tag{8}$$

The statistical model in the above equation is called independent component analysis, or ICA model. The ICA model is a generative model, which means that it describes how the observed data are generated by a process of mixing the components si. The independent components are latent variables, meaning that they cannot be directly observed.

Also the mixing matrix is assumed to be unknown. All we observe is the random vector **x**, and we must estimate both **A** and **s** using it. This must be done under as general assumptions as possible. The starting point for ICA is the very simple assumption that the components s_i are statistically independent. Then, after estimating the matrix **A**, we can compute its inverse, say **W**, and obtain the independent component simply by:

$$\mathbf{s} = \mathbf{Wx} \tag{9}$$

In this paper, these two algorithms has been improved using fuzzy-genetic based approach as discussed in the following section.

III. FUZZY-GENETIC BASED PCA & ICA

Initially the feature vectors have been clustered using Fuzzy C-Means (FCM) clustering. The FCM algorithm, also known as Fuzzy ISODATA, is one of the most frequently used methods in pattern recognition. It is based on minimization of the given objective function to achieve good classifications.

$$J(U,V) = \sum_{i=1}^{n}\sum_{j=1}^{c} (\mu_{ij})^m \parallel x_i - v_j \parallel^2 \tag{10}$$

J(U,V) is a squared error clustering criterion, and solutions of minimization of (1) are least-squared error stationary points of J(U,V). The expression, $X = \{x_1, x_2,\ldots, x_n\}$ is a collection of data, where n is the number of data points. $V = \{ v_1, v_2,\ldots, v_c\}$ is a set of corresponding cluster centers in the data set **X**, where c is the number of clusters. μ_{ij} is the membership degree of data x_i to the cluster centre v_j. Meanwhile, μ_{ij} has to satisfy the following conditions:

$$\mu_{ij} \in [0,1], \forall i = 1,...,n, \forall j = 1,...,c$$
$$\sum_{j=1}^{c} \mu_{ij} = 1, \forall i = 1,...,n \tag{11}$$

Where $U = (\mu_{ij})_{n*c}$ is a fuzzy partition matrix, $\parallel x_i - v_j \parallel$ represents the Euclidean distance between x_i and v_j, parameter m is the "fuzziness index" and is used to control the fuzziness of membership of each datum in the range $m \in [1,\infty]$. In this experimentation the value of m=2.0 was chosen. Although there is no theoretical basis for the optimal selection of m, this has been chosen because the value has been commonly applied within the literature. The FCM algorithm is described in, for example, and can be performed by the following steps:

a. Initialize the cluster centres $V = \{ v_1, v_2,\ldots, v_c\}$, or initialize the membership matrix μ_{ij} with random value and make sure it satisfies the above conditions and then calculate the centres.
b. Calculate the fuzzy membership μ_{ij} using

$$\mu_{ij} = \frac{1}{\sum_{k=1}^{c}\left(\frac{d_{ij}}{d_{ik}}\right)^{\frac{2}{m-1}}}$$
$$where, d_{ij} = \parallel x_i - v_j \parallel, \forall i = 1,...,n, \forall j = 1,...c \tag{12}$$

c. Compute the fuzzy centres v_j using

$$v_j = \frac{\sum_{i=1}^{n} (\mu_{ij})^m x_i}{\sum_{i=1}^{n} (\mu_{ij})^m}, \forall j = 1,...,c \tag{13}$$

d. Repeat steps (2) and (3) until the minimum J value is achieved.
e. Finally, defuzzification is necessary to assign each data point to a specific cluster (i.e. by setting a data point to a cluster for which the degree of the membership is maximal).

In the next step, Genetic Algorithm is used to choose the more relevant features. The input data is transformed to higher dimension using a non-linear transfer function (polynomial

function) and GA is used to select the optimal subset of the component clusters with the fitness function taken as the recognition performance. The limitations in PCA & ICA are, (i) We cannot determine the variances (energies) of the components, (ii) We cannot determine the order of the components, and (iii) we can't expect the selected components has equal number of vectors from each class. Here, the GA is used to select the optimum subset of components from each cluster; hence the above said limitations can be overcome.

In our proposed method, we are going to choose only F number of components, for each cluster, the reduced feature set will contain S=NC×F number of features, where NC is the number of clusters. In our case we have two clusters: ischemic and non-ischemic. Initially an n number of components are selected from each cluster at random. The index of each component is used to construct one chromosome. Similarly N number of chromosomes is generated. (N=10). For example, consider the chromosome:

980	726	657	807	240	825…..
24	252	220	92	534	155……

Each integer represents one component. The first 980 stands for the 980th components from the first cluster, the 24 in the second row represents the 24th component from the second cluster. The total length of the chromosome is equal to the total number of independent components required. Here, we kept the size as 600. For each chromosome, the Euclidian distance within the class (W) and between the classes (B) has been calculated. The fitness value is calculated as:

$$f(x) = B / W \tag{14}$$

The chromosome which has the minimum fitness value (Gmin) is stored as the best independent component set. Then the genetic operators are applied to search for the optimum set.

Reproduction (selection) – The selection process selects chromosomes from the mating pool directed by the survival of the fittest concept of natural genetic systems. In the proportional selection strategy adopted in this article, a chromosome is assigned a number of copies, which is proportional to its fitness in the population that goes into the mating pool for further genetic operations. Roulette wheel selection is one common technique that implements the proportional selection strategy.

Crossover – is a probabilistic process that exchanges information between two parent chromosomes for generating two child chromosomes. In this paper, single point crossover with a fixed crossover probability of p_c=0.6 is used. For chromosomes of length *l*, a random integer, called the crossover point, is generated in the range [1, *l*-1]. The portions of the chromosomes lying to the right of the crossover point are exchanged to produce two offspring.

Mutation – Each chromosome undergoes mutation with a fixed probability p_m=0.003. For binary representation of chromosomes, a bit position (or gene) is mutated by simply flipping its value. Since we are considering real numbers in this paper, a random position is chosen in the chromosome and replace by a random number between 0-9.

The new population is generated after the genetic operators are applied. The current best IC set is (Lmin) selected from the new population and compared with the global one. If the global set contains minimum fitness value then the local, the next iteration is continued with the old population. Otherwise, the current population is considered for the next iteration. This process is repeated for k number of iterations. The algorithm is given as:

1. Construct the initial population (p1) with random ICs.
2. Calculate the fitness value $f(x) = B / W$.
3. Find out the Global minimum (Gmin)
4. For i = 1 to k do
    a. Perform reproduction
    b. Apply the crossover operator between each parent.
    c. Perform mutation and get the new population. (p2)
    d. Calculate the local minimum (Lmin).
    e. If Gmin < Lmin then
        i. Gmin = Lmin;
        ii. p1 = p2;
5. Repeat

The similar approach is applied for both PCA & ICA to select optimum feature set, these feature vectors are fed to LSSVM to classify the ECG signals.

IV. Least Square Support Vector Machine (LSSVM) Classifier

SVM is a reliable classification technique, which is based on the statistical learning theory. This technique was firstly proposed for classification and regression tasks by [13]. A linear SVM was developed to classify the data set which contains two separable classes such as {+1,−1}. Let the training data consist of n datum $(x_1, y_1), \ldots, (x_n, y_n)$, $x \in R_n$ and $y \in \{+1,-1\}$. To separate these classes, SVMs have to find the optimal (with maximum margin) separating hyperplane so that SVM has good generalization ability. All of the separating hyperplanes are formed with

$$D(x) = (w * x) + w_0 \tag{15}$$

and provide following inequality for both y =+1 and y =−1.

$$y_i [(w*x_i) + w_0] \geq 1, i = 1 , \ldots , n. \tag{16}$$

The data points which provide above formula in case of equality are called the support vectors. The classification task in SVMs is implemented by using these support vectors. Margins of hyperplanes obey following inequality:

$$[(y_k \times D(x_k)) / \| w \|] \geq \Gamma, k = 1, \ldots, n \quad (17)$$

To maximize this margin (Γ), norm of w is minimized. To reduce the number of solutions for norm of w, following equation is determined:

$$\Gamma \times \|w\| = 1 \quad (18)$$

Then the following formula is minimized to $1/2 \|w\|^2$

When we study on the non-separable data, slack variables ξ_i, are added and the formulas are revised as

$$y_i [(w.x_i) + w_0] \geq 1 - \xi_i \quad (19)$$

$$C \sum_{i=1}^{n} \xi_i + \frac{1}{2} \|w\|^2 \quad (20)$$

Since originally SVMs classify the data in linear case, in the nonlinear case SVMs do not achieve the classification tasks. To overcome this limitation on SVMs, kernel approaches are developed. Nonlinear input data set is converted into high dimensional linear feature space via kernels. In SVMs, following kernels are most commonly used.

- Dot product kernels: $K(x, x') = x.x'$;
- Polynomial kernels: $K(x, x') = (x.x' + 1)^d$; where d is the degree of kernel and positive integer number;
- RBF kernels: $K(x, x') = \exp (- \| x - x' \|^2 / \sigma^2)$; where σ is a positive real number.

LSSVMs are firstly proposed by [14]. The most important difference between SVMs and LSSVMs is that LSSVMs use a set of linear equations for training while SVMs use a quadratic optimization problem [15]. While equation (20) is minimized subject to formula (19) in Vapnik's standard SVMs, in LSSVMs formula (22) is minimized subject to formula (21).

$$y_i [(w.x_i) + w_0] = 1 - \xi_i, i = 1,\ldots,n, \quad (21)$$

$$\frac{1}{2}\|w\|^2 + \frac{C}{2} \sum_{i=1}^{n} \xi_i^2 \quad (22)$$

According to these formulas, their dual problems are built as following:

$$Q(w, b, \alpha, \xi) = \frac{1}{2} \|w\|^2 + \frac{C}{2} \sum_{i=1}^{n} \xi_i^2 - \sum_{i=1}^{n} \alpha_i \{ y_i [(w.x_i) + w_0] - 1 + \xi_i \} \quad (23)$$

Another difference between SVMs and LSSVMs is that α_i (Lagrange multipliers) are positive or negative in LSSVMs but they must be positive in SVMs [14,15].

V. EXPERIMENTS AND RESULTS

The European ST-T Database is used for evaluation of our proposed algorithm. This database consists of 90 annotated excerpts of ambulatory ECG recordings from 79 subjects. The subjects were 70 men aged 30 to 84, and 8 women aged 55 to 71. The database includes 367 episodes of ST segment change, and 401 episodes of T-wave change. Each record is two hours in duration and contains two signals, each sampled at 250 samples per second with 12-bit resolution over a nominal 20 millivolt input range. The sample values were rescaled after digitization with reference to calibration signals in the original analog recordings, in order to obtain a uniform scale of 200 ADC units per millivolt for all signals. (The calibration signals are not included in the signal files.) The header files include information about the leads used, the patient's age, sex, and medications, the clinical findings, and the recording equipment. Each of the signal files is 5,400,000 bytes long.

Two cardiologists worked independently to annotate each record beat-by-beat and for changes in ST segment and T-wave morphology, rhythm, and signal quality. ST segment and T-wave changes were identified in both leads (using predefined criteria which were applied uniformly in all cases), and their onsets, extrema, and ends were annotated. Annotations made by the two cardiologists were compared, disagreements were resolved by the coordinating group in Pisa, and the reference annotation files were prepared; altogether, these files contain 802,866 annotations. Over half (48 of 90 complete records, and reference annotation files for all records) of this database is freely available from PhysioNet. In this paper, we have taken the full length ECG signals from 17 patients. And the signals are sliced with 60 seconds time interval; each signal will be translated into 120 samples and totally 2040 beats. This dimensionality is reduced by GICA as discussed in the earlier section.

The performance of our proposed method is compared with linear PCA, ICA with LSSVM classifier. Also the performance is studied with Genetic based PCA and ICA, here the clustering step is excluded and the components have been chosen directly with GA. The performance is studied with the measurements like classification accuracy, sensitivity and specificity analysis and ROC curves. The classification accuracies for the datasets are measured using the equation:

$$Accuracy\ (T) = \frac{\sum_{t=1}^{|T|} assess(t_t)}{|T|}, t_t \epsilon\ T, \quad (24)$$

$$Assess(t) = \begin{cases} 1, & if\ classify(t) = t.c, \\ 0, & otherwise \end{cases} \quad (25)$$

where T is the set of data items to be classified (the test set), t ∈ T , t.c is the class of item t , and classify(t) returns the classification of t by LSSVM.

The Sensitivity, specificity, TP rate, FP rate, and accuracy are calculated as:

Sensitivity = (TP / (TP+FN)) (%)
Specificity = (TN / (FP+TN)) (%)
FP_{rate} = (FP / (FP+TN))
TP_{rate} = (TP / (TP+FN))
Accuracy = (TP+TN) / (TP+FN+FP+TN)

where TP, TN, FP, and FN denotes true positives, true negatives, false positives, and false negatives, respectively.

A Receiver Operating Characteristic (ROC) graph is a procedure for visualizing and analyzing the classifier performance [16]. ROC graphs are two-dimensional graphs in which TP (true positive) rate is plotted on the Y axis and FP (false positive) rate is plotted on the X axis. The area under the ROC curve is an important criterion for evaluating diagnostic performance. Usually it is referred as the A_Z index. The A_Z value of ROC curve is just the area under the ROC curve. The estimation of the A_Z value can be obtained with the trapezoidal rule that can underestimate areas under the curve. More operating points are generated; less underestimation error will be obtained. The A_Z value can also be computed by fitting a continuous binomial curve to the operating points. Here we have conducted ten-fold validation method for analyzing the performance of our proposed method with the existing methods. Figure 1 shows the ROC curves for comparison of classification performances for the proposed method.

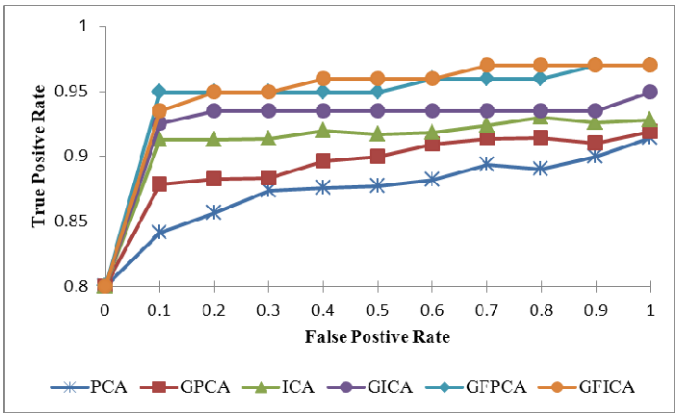

Figure 1. Performance Analysis of Feature Extraction Methods

Table 1 shows the A_z value of existing and the proposed methods for automated ischemic beat classification. Table 2 shows the sensitivity at each fold.

TABLE I. PERFORMANCE ANALYSIS OF ISCHEMIC BEAT CLASSIFICATION

Methods	PCA	GPCA	ICA	GICA	FGPCA	FGICA
Sensitivity (%)	88	90	92	94	94	96
Specificity (%)	85	89.7	90	92.5	94.8	92.6
FP rate (%)	15	10.3	10	7.5	5	7
TP rate (%)	88	90	92	94	94	96
Accuracy (%)	86.7	90	91	93.3	94.4	94.7
A_z Value	0.87	0.90	0.92	0.94	0.95	0.95

TABLE II. THE VALUE OF SENSITIVITY AT EACH FOLD

	PCA	GPCA	ICA	GICA	FGPCA	FGICA
1	84	88	91	94	96	97
2	86	88	91	94	96	95
3	87	88	91	93	95	97
4	88	90	93	94	95	96
5	88	91	92	94	96	95
6	88	91	92	94	95	97
7	89	91	92	94	96	94
8	90	91	92	94	95	96
9	89	90	93	94	97	97
10	91	92	93	95	97	96

VI. CONCLUSION

In this paper, we have proposed an improved version of Principal Component Analysis (PCA) and Independent Component Analysis (ICA) in ischemia beat classification. The Fuzzy C-Means (FCM) clustering and Genetic Algorithm (GA) is combined with PCA and ICA to select optimum components from the feature set vector of ECG signals. Initially the features are extracted from the ECG signals, and FCM is applied to cluster the signals into ischemic and non-ischemic. GA is applied to select optimum number of features from each cluster. The extracted features from GFPCA, GFICA are fed into Least Square Support Vector Machine (LSSVM) to classify the beats into ischemic or non-ischemic. The experimental results shown that our proposed GFPCA and GFICA methods can extract more relevant features than the other methods proposed in the literature with highest classification accuracy.

REFERENCES

[1] S.C. Saxena, V. Kumar, S.T. Hamde, Feature extraction from ECG signals using wavelet transforms for disease diagnostics, Int. J. Syst. Sci. 33 (13) (2002) 1073–1085.

[2] S.Y. Foo, G. Stuart, B. Harvey, A. Meyer-Baese, Neural network-based EKG pattern recognition, Eng. Appl. Artif. Intell. 15 (2002) 253–260.

[3] N. Maglaveras, T. Stamkopoulos, K. Diamantaras, C. Pappas, M. Strintzis, ECG pattern recognition and classification using non-linear transformations and neural networks: A review, Int. J. Med. Inform. 52 (1998) 191–208.

[4] K. Sternickel, Automatic pattern recognition in ECG time series, Comput. Methods Programs Biomed. 68 (2002) 109–115.

[5] I. Güler, E.D. Übeyli, ECG beat classifier designed by combined neural network model, Pattern Recogn. 38 (2) (2005) 199–208.

[6] Papaloukas C, Fotiadis D.I, Likas A, Liavas A.P, and Michalis L.K, A robust knowledge based technique for ischemia detection in Noisy ECGs, 4th International conference on knowledge-based Intelligent Engineering Systems & Allied Technologies, IEEE, 2000.

[7] S. Radhakrishnan, and S. Murugan, Automated Ischemic Beat Classification Using Genetic Algorithm Based Principal Component Analysis, International Journal on Healthcare Technology and Management, 2010. vol 3, No. 11,PP.151-162

[8] Castells F, Laguna P, Ornmo PS, Bollmann A, and Roig J, Principal Component Analysis in ECG Signal Processing, EURASIP Journal on Advances in Signal Processing, 2007;1-29.

[9] Hyvarinen, A., Karhunen, J., & Oja, E. (2001). Independent component analysis. John Wiley & Sons.

[10] Hyva¨rinen, A., & Oja, E. (1997). A fast fixed-point algorithm for independent component analysis. Neural Computation, 9, 1483–1492.

[11] Hyva¨rinen, A. (1999). Fast and robust fixed-point algorithms for independent component analysis. IEEE Transactions on Neural Networks, 10(3), 626–634.

[12] Comm, P. (1994). Independent component analysis – A new concept? Signal Processing, 36, 287–314.

[13] V. Vapnik, The Nature of Statistical Learning Theory, Springer, New York, 1995.

[14] J.A.K. Suykens, J. Vandewalle, Least squares support vector machine classifiers, Neural Process. Lett. 9 (3) (1999) 293–300.

[15] D. Tsujinishi, S. Abe, Fuzzy least squares support vector machines for multi-class problems, Neural Networks Field 16 (2003) 785–792.

[16] F. Fawcett, An introduction to ROC analysis, Pattern Recogn. Lett. 27 (8) (2006) 861–874.

Biographical notes: S. Murugan received his MTech in Biomedical Engineering from Indian Institute of Technology, Madras in 2002. He is currently working as an Assistant Professor in the Department of Instrumentation and Control Engineering, Arulmigu Kalasalingam College of Engineering, Krishnankoil, Srivilliputhur (via), Virudhunagar District, Tamil Nadu, India. His current research interests include biomedical instrumentation, bio-signal processing and analysis.

S. Radhakrishnan received his PhD in Biomedical Engineering from Banaras Hindu University in 1993. At present he is working as a Senior Professor and Head, Department of Computer Science Engineering, Arulmigu Kalasalingam College of Engineering, Krishnankoil, Srivilliputtur (via), Virudhunagar District, Tamil Nadu, India. He has rich teaching experience and published more than 100 papers in various journals and conferences in both national and international level. His current research focuses on sensor networks, bio signal processing and image processing.