Classification of Arrhythmias with LDA and ANN using Orthogonal Rotations for Feature Reduction

Manpreet Kaur[1] and A.S.Arora[2]

[1] **Department of EIE, SLIET**
Longowal-148106, Distt. Sangrur (PB) India

[2] **Department of EIE, SLIET**
Longowal-148106, Distt. Sangrur (PB) India

Abstract

This paper presents a new approach for feature reduction by using orthogonal rotations. Wavelet coefficients for beat segments are taken as features which are reduced by factor analysis method using orthogonal rotations. LDA (Linear Discriminant Analysis) and ANN (Artificial Neural Network) classifiers are used for classification. The signals are taken from MIT-BIH arrhythmia database to classify into Normal, PVC, Paced, LBBB and RBBB. The performance of classification output has been compared by the performance parameters. Both the classifiers have given best overall accuracy for 'equimax' rotation. 96% accuracy is achieved with LDA classifier,99.2% accuracy is achieved using ANN.

Keywords: *ECG, Linear Discriminant Analysis, Artificial Neural Network, Holdout Method, Orthogonal Rotations*

1. Introduction

ECG is an important measure for diagnosing heart diseases. Automatic analysis and interpretation of ECG signal is difficult task as large variations in the morphology of ECG waveform exists not only of the different patients or patient groups but also within the same patient. With the advances in computer aided diagnosis many methods have been proposed by researchers to provide such a diagnosis system that can assist expert in making decision.

Ozbay Y., Ceylan R., Karlik B.[1] have used fuzzy clustering neural network architecture for classification of ECG arrhythmias. They made a comparative study of classification accuracy of ECG signals using MLP NN and a new fuzzy clustering architecture. Average recognition accuracy is 99.9%. Gholam Hosseini H., Luo D. and Reynolds K.J.[2] have compared different feed forward neural networks architectures for ECG signal diagnosis. Korurek M. and Nizam A.[3] presents an efficient Arrhythmia Clustering and Detection algorithm based on Ant Colony optimisation technique for QRS complex. ACO based clustering technique has also been improved using nearest neighbourhood. The method is tested with MIT-BIH database to classify six different arrhythmias i.e. normal sinus rhythm, premature ventricular contraction, atrial premature contraction right bundle branch block, ventricular fusion and fusion. Kim J., Shin H.S., Shin K. and Lee M.[4] have classified six different beats using extreme learning machine taking PCA and other statistical features for classification. Yeh Y-C., Wang W-J. and Chiou C.W.[5] have used linear disciminant analysis for classification of ECG signals. Zadeh A.E., Khazaee A. and Ranaee V.[6] have considered only one morphological feature that is timing to classify three different beats. They have used neural networks, MLP, RBF and SVM classifiers. Ozbay Y., Ceylan R. and Karlik B.[7] proposed a new automated diagnostic system using type-2 fuzzy c-means clustering (T2FCM), wavelet transform and neural network to classify arrhythmias. The T2FCWNN architecture is realised in three stages. First stage is formed by the new training set obtained by selection of the best segments for each arrhythmia class using T2FCM.Second stage is feature extraction by WT on the new training set. Third set is classification of the extracted features using neural network. Sufi F., Khalil I. and Mahmood A.N.[8] have proposed attribute selection method that selects only few features from compressed ECG. They used expected maximization (EM) clustering technique to create normal and abnormal ECG clusters. Twenty different segments (13 normal and 7 abnormal) of compressed ECG from MIT-BIH subject are tested with 100% success. They also presented an algorithm to identify initiation of abnormality. Karpagachelvi S., Arthanari M. and Siva Kumar M.[9] have compared the performance of relevance vector machine (RVM) with extreme learning machine (ELM) to classify ECG beats and concluded that sensitivity of RVM classifier is more than ELM. They concluded that RVM accomplishes better and more balanced classification for individual classifications as well in very less training time.

The proposed work use orthogonal rotations: quartimax, varimax and equimax for feature reduction which are then classified with Linear Discriminant Analysis and Artificial Neural Network.

2. Wavelet for Feature Extraction

The wavelet transform is a time-scale representation method that decompose signal f(t) into basis function of time and scale which are dilated and translated version of a basis function $\psi(t)$ which is called mother wavelet. Translation is accomplished by considering all possible integer translation of $\psi(t)$ and dilation is obtained by multiplying t by a scaling factor which is usually 2.The following equation shows how wavelets are generated from mother wavelet:

$$\Psi_{j,k}(t) = 2^{j/2}\Psi(2^{j/2}\,t - k)$$

where j indicates the resolution level and k is the translation in time. This is called dyadic scaling. Wavelet decomposition is a linear expansion and it is expressed as:

$$F(t) = \sum_{k=-\infty}^{+\infty} c_k \varphi(t-k) + \sum_{k=-\infty}^{+\infty}\sum_{j=0}^{+\infty} d_{j,k}\,\psi(2^j t - k)$$

Where $\varphi(t)$ is called the scaling function or father wavelet. c_k and $d_{j,k}$ are the coarse and detail level expansion coefficients, respectively[10].

For current analysis, the detailed coefficients cd4, cd5, cd6, cd7 are taken as features for classifications .These sub-band signals have representative components and different distributions to each other for various types of ECG beats[11].

3. Factor Analysis

The idea underlying factor analysis[12] is that *p,* the observed random variables, can be written as linear functions of d(less than p), the unobserved latent variables or common factors f_j as follows:

$$x_1 = \lambda_{11} f_1 + \cdots + \lambda_{1d} f_d + \varepsilon_1$$
$$\ldots$$
$$x_p = \lambda_{p1} f_1 + \cdots + \lambda_{pd} f_d + \varepsilon_p$$

The λ_{ij} (i=1..p, j=1..d) in the above model are called factor loadings, and the error terms ε_i are called the specific factors. The error terms ε_i are specific to each of the original variables, while f_j are common to all of the variables. The sum of the squared factor loadings for i-th variable

$$\lambda_{i1}^2 + \cdots + \lambda_{id}^2$$

is called communality of x_i.

Matrix form of the factor analysis model is $x = \Lambda f + e$, where Λ is a factor loading matrix. Some assumptions are made regarding this model, which are E[e]=0, E[f]=0, E[x]=0. Where E [•] denotes the expected value. If the last of these assumptions is violated, the model can be adjusted to accommodate this, yielding x= Λ f + e+μ, where E[x] =μ. It is also assumed that the error terms ε_i are uncorrelated with each other and that the common factors are uncorrelated with the specific factors f_j. Given these assumptions, the sample covariance (or correlation) matrix is of the form:

$$S = \Lambda^T\Lambda + \Psi$$

where Ψ is a diagonal matrix representing E[ee^T]. The variance of ε_i is called the specificity of x_i so the matrix Ψ is also called the specificity matrix. Estimation of the parameters in the factor analysis model is usually accomplished via the matrices Λ and Ψ. The estimation proceeds in stages, where an initial estimate is found by placing conditions on Λ. Once this initial estimate is obtained, other solutions can be found by rotating Λ. These factor rotation methods can either be orthogonal or oblique. The orthogonal rotation methods include quartimax, varimax and equimax.

The goal of the quartimax rotation is to simplify the rows of the factor matrix by getting a variable with a high loading on one factor and small loadings on all other factors. The varimax rotation focuses on simplifying the columns of the factor matrix. For the varimax approach, perfect simplification is obtained if there are only ones and zeros in a single column. The output from this method tends to have high loadings close to ±1 and some near zero in each column. The equimax rotation is a compromise between these two methods, where both rows and columns of the factor matrix are simplified as much as possible [12].

In this paper ‘Principal Component Method’ is used to estimate the component loadings.

4. Linear Discriminant Analysis

Discriminant functions(z) are the linear combinations of the variables which best separate the pre-determined groups[13]; they are uncorrelated with each other. Discriminant functions are obtained from the eigenvalues (λ) and eigenvectors (a) of a matrix relationship (E-λH) a=0 for the "between sums of squares" covariance matrix (H) and "within sums of squares" covariance matrix (E). The discriminant function coefficients are vectors (z_1, z_2, … , z_s) calculated from the eigenvectors (a_1, a_2, … , a_s).

The discriminant function loadings (or structure coefficients) are the correlations of the original variables with the new discriminant functions. The discriminant function scores are obtained as the discriminant function coefficients multiplied by the original *Y* data matrix.

IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 4, No 2, July 2012
ISSN (Online): 1694-0814
www.IJCSI.org

An improved estimate of error rate can be obtained by Holdout method, where all data but one case is used to calculate the linear classification functions, which are then used to classify the omitted case either correctly or incorrectly. This holdout procedure is repeated for every individual case (requiring considerable computation effort) and the error rate is determined from the cumulated classifications/ misclassifications of the holdout cases[13].

5. Artificial Neural Networks

Presently, artificial neural networks have been widely used in the field of classification. Developing a classifier involves choosing an appropriate classifier model, and then training algorithm to train and then test the input signal to classify them into different categories[2]. Back propagation algorithm is used to train feed-forward NN to classify arrhythmias. In this paper for classification, Levenberg-Marquardt algorithm has been used as it is the fastest among all back-propagation algorithms although it requires more memory than other algorithms. A tan-sigmoid function is used for hidden layer and a linear transfer function is used for the output layers

Table 1: ECG classes and desired NN output

Classes	Type of Beats	Neural Networks output				
1	Normal	1	0	0	0	0
2	PVC	0	1	0	0	0
3	Paced	0	0	1	0	0
4	LBBB	0	0	0	1	0
5	RBBB	0	0	0	0	1

6. Performance Parameters

There are many approaches in the literature which judge the performance of a classifier. For this paper, three statistical indices Accuracy(Acc), Sensitivity(Se) and Positive predictivity (Pp) are taken as performance parameters[14].

Accuracy: Overall accuracy of the classifier has been defined as:

$$Acc = \frac{N_T - N_E}{N_T}\ 100$$

Where N_E represents total number of classification errors and N_T represents total number of beats.

Sensitivity: It is the ratio of number of correctly classified beats (TP) to the total number of beats (TP+FN)

$$Se = \frac{TP}{TP + FN}\ 100$$

Where TP represents the true positive beats and FN represents the false negative beats.

***Positive Predictivity*:** It is the ratio of number of correctly classified beats (TP) to sum of (TP+FP)

$$Pp = \frac{TP}{TP + FP}\ 100$$

FP represents false positive beats which is number of falsely detected events.

7. Methodology

7.1 Preparation of datatsets

Considering R point at middle, a total of 144 samples have been taken as one beat segment[11]. The samples are extracted from the raw dataset. Since the MIT-BIH database comes with annotations, each R wave location and the extracted beat are categorized from annotations. Total of 1060 beats are extracted with 212 of each type.

7.2 Extraction of features

db4 wavelet transform has been used to extract features. The detailed coefficients cd4, cd5, cd6, cd7 are taken as features for classifications. Total of 43 features have been selected for classification.

7.3 Features Reduction

These features have been reduced with the factor analysis method. The factor analysis with 'principal components method' has been used with orthogonal rotation, varimax rotation, quartimax rotation and equimax rotation and result is compared with factors obtained with no rotation. The criteria for determining number of factors is 'variance'[13]. Then total of 23 features have been used for classification.

7.4 Classification Using LDA

For LDA classification, whole data is used for classification. After classification, improved estimate of error rate is obtained with holdout method.

7.5 Classification using ANN

For ANN, data set of 162 signal of each type is used for training and 50 samples of each type are used for testing the network. The classification has been done for all the four categories and results have been shown in Tables 10-14. So a total of 810×23 matrix is used for learning of the neural network and 250×23 matrix is used for testing of network. For the learning of the network, 70% data is used for training, 15% is used for validation and 15% is used for testing the network.

8. Classification Results

8.1 Outputs using LDA

Table 2: No Rotation Classification

S No	Type of Beats	Total No. of beats	No of beats correctly classified	No. of FN beats	No. of FP beats	Se (%)	Pp (%)
1	Normal	212	206	6	25	97.2	89.17
2	PVC	212	207	5	2	97.6	99.04
3	Paced	212	207	5	5	97.6	97.64
4	LBBB	212	205	7	10	96.7	95.35
5	RBBB	212	189	23	4	89.2	97.92
	%Acc					95.7	

Table 3:Varimax Rotation Classification

S No	Type of Beats	Total No. of beats	No of beats correctly classified	No. of FN beats	No of FP beats	Se (%)	Pp (%)
1	Normal	212	192	20	23	90.6	89.30
2	PVC	212	207	5	2	97.6	99.04
3	Paced	212	207	5	5	97.6	97.64
4	LBBB	212	207	5	17	97.6	92.41
5	RBBB	212	181	31	19	85.4	90.50
	%Acc					93.8	

Table 4: Quartimax Rotation Classification

S No	Type of Beats	Total No. of beats	No of beats correctly classified	No. of FN beats	No of FP beats	Se (%)	Pp (%)
1	Normal	212	194	18	22	91.5	89.81
2	PVC	212	207	5	2	97.6	99.04
3	Paced	212	208	4	5	98.1	97.65
4	LBBB	212	208	4	16	98.1	92.85
5	RBBB	212	181	31	17	85.4	91.41
	%Acc					94.2	

Table 5: Equimax Rotation Classification

S No	Type of Beats	Total No. of beats	No of beats correctly classified	No. of FN beats	No of FP beats	Se (%)	Pp (%)
1	Normal	212	204	8	20	96.2	91.07
2	PVC	212	207	5	2	97.6	99.04
3	Paced	212	206	6	5	97.2	97.63
4	LBBB	212	209	3	8	98.6	96.31
5	RBBB	212	192	20	7	90.6	96.48
	%Acc					96	

8.2 Output using Holdout Method

Table 6: No Rotation HoldOut

S No	Type of Beats	Total No. of beats	No of beats correctly classified	No. of FN beats	No of FP beats	Se (%)	Pp (%)
1	Normal	212	206	6	28	97.2	88.03
2	PVC	212	207	5	4	97.6	98.10
3	Paced	212	204	8	5	97.2	97.60
4	LBBB	212	205	7	11	96.7	94.90
5	RBBB	212	186	26	4	87.7	97.89
	%Acc					95.1	

Table 7: Varimax Rotation HoldOut

S No	Type of Beats	Total No. of beats	No of beats correctly classified	No. of FN beats	No of FP beats	Se (%)	Pp (%)
1	Normal	212	190	22	24	89.6	89.78
2	PVC	212	207	5	2	97.6	99.04
3	Paced	212	206	6	5	97.2	97.63
4	LBBB	212	207	5	18	97.6	92
5	RBBB	212	180	32	21	84.9	89.55
	% Acc					93.4	

Table 8:Quartimax Rotation HoldOut

S No	Type of Beats	Total No. of beats	No of beats correctly classified	No. of FN beats	No of FP beats	Se (%)	Pp (%)
1	Normal	212	192	20	24	90.6	88.89
2	PVC	212	207	5	3	97.6	98.57
3	Paced	212	206	6	5	97.2	97.63
4	LBBB	212	207	5	21	97.6	90.78
5	RBBB	212	177	35	18	83.5	90.77
	%Acc					93.3	

Table 9:Equimax Rotation HoldOut

S No	Type of Beats	Total No. of beats	No of beats correctly classified	No. of FN beats	No of FP beats	Se (%)	Pp (%)
1	Normal	212	199	13	23	93.9	89.64
2	PVC	212	207	5	2	97.6	99.04
3	Paced	212	206	6	5	97.2	97.63
4	LBBB	212	209	3	8	98.6	96.31
5	RBBB	212	189	23	12	89.2	94.03
	%Acc					95.3	

8.3 Output of Training data using ANN

Table 10: No Rotation Training

No Rotation	Hidden Layers	Accuracy (%)	Time (sec)	Normal		PVC		Paced		LBBB		RBBB	
				Se	Pp	Se	Pp	Se	Pp	Se	Pp	Se	Pp
NN1	10	99.6	5	100	99.4	99.4	99.4	100	99.4	100	100	98.8	100
NN2	15	99.6	6	100	99.4	99.4	100	99.4	99.4	100	99.4	99.4	100
NN3	20	99.4	9	98.8	99.4	99.4	99.4	99.4	99.4	100	99.4	99.4	99.4

Table 11: Varimax Rotation Training

No Rotation	Hidden Layers	Accuracy (%)	Time (sec)	Normal		PVC		Paced		LBBB		RBBB	
				Se	Pp	Se	Pp	Se	Pp	Se	Pp	Se	Pp
NN1	10	99.3	4	99.4	100	99.4	99.4	98.1	98.8	99.4	98.8	100	99.4
NN2	15	99.1	6	100	97.6	99.4	99.4	96.9	100	100	98.8	99.4	100
NN3	20	99.4	15	98.8	98.8	100	100	100	100	100	99.4	98.1	98.8

Table 12: Quartimax Rotation Training

No Rotation	Hidden Layers	Accuracy (%)	Time (sec)	Normal		PVC		Paced		LBBB		RBBB	
				Se	Pp	Se	Pp	Se	Pp	Se	Pp	Se	Pp
NN1	10	99	4	99.4	98.8	100	100	97.5	100	98.8	98.8	99.4	97.6
NN2	15	99.4	4	100	99.4	99.4	99.4	99.4	100	99.4	99.4	98.8	98.8
NN3	20	99.6	13	100	100	100	99.4	99.4	99.4	99.4	100	99.4	99.4

Table 13: Equimax Rotation Training

No Rotation	Hidden Layers	Accuracy (%)	Time (sec)	Normal		PVC		Paced		LBBB		RBBB	
				Se	Pp	Se	Pp	Se	Pp	Se	Pp	Se	Pp
NN1	10	99.4	6	100	98.2	99.4	100	99.4	99.4	99.4	100	98.8	99.4
NN2	15	99.8	12	100	100	99.4	100	100	99.4	100	100	99.4	99.4
NN3	20	99.8	12	100	98.8	100	100	100	100	99.4	100	99.4	100

8.4 Output of Test data using ANN

Table 14: Testing data(for 20 hidden layers)

	Accuracy (%)	Normal		PVC		Paced		LBBB		RBBB	
		Se	Pp	Se	Pp	Se	Pp	Se	Pp	Se	Pp
No Rotation	98	100	100	100	96.2	94	100	96	98	100	96.2
Varimax	98.8	100	100	100	100	100	98	96	100	98	96.1
Quartimax	98	100	92.6	100	100	96	98	94	100	100	100
Equimax	99.2	100	98	100	98	100	100	98	100	98	100

9. Conclusions

Factor analysis method gives reduction in dimensionality without and with orthogonal rotations. In terms of overall accuracy, equimax rotation provides the maximum %age value i.e. 96% with LDA classifier and 95.3% with Holdout procedures. In case of ANN, %age accuracy for training data is 99.8% and for testing data it is 99.2% (the test data result is calculated for 20 hidden layers as it provides the more accuracy than the 10 and 15 hidden layer model). The results obtained are comparable with the results obtained by other researchers and hence the method is proposed for feature reduction.

References

[1] Ozbay Y., Ceylan R., Karlik B., "A fuzzy clustering neural network architecture for classification of ECG arrhythmias", Computers in Biology and Medicine, Vol.36, Issue 4, 2006, pp.376-388.

[2] Gholam Hosseini H., Luo D. and Reynolds K.J., "The comparison of different feed forward neural networks architectures for ECG signal diagnosis", Medical Engineering and Physics, Vol. 28, 2006, pp. 372-378.

[3] Korurek M. and Nizam A., "A new arrhythmia clustering technique based on Ant Colony optimization", Journal of Biomedical Informatics, Vol.41, 2008, pp. 874-881.

[4] Kim J., Shin H.S., Shin K. and Lee M., "Robust algorithm for arrhythmia classification in ECG using extreme learning machine", Biomedical Engineering Online, Vol.8, Issue 1, 2009.

[5] Yeh Y-C., Wang W-J. and Chiou C.W., "Cardiac arrhythmia diagnosis method using linear disciminant analysis on ECG signals", Measurement, Vol.42, 2009, pp.778-789.

[6] Zadeh A.E., Khazaee A. and Ranaee V., "Classification of the electrocardiogram signals using supervised classifiers and efficient features", Computer Methods and Programs in Biomedicine, Vol.99, 2010, pp.179-194.

[7] Ozbay Y., Ceylan R. and Karlik B., "Integration of type-2 fuzzy clustering and wavelet transform in a neural network based ECG classifier", Expert System with Applications Vol. 38, 2011, pp.1004-1010.

[8] Sufi F., Khalil I. and Mahmood A.N., "A clustering based system for instant detection of cardiac abnormalities from compressed ECG", Expert System with Applications, Vol.38, pp.4705-4713.

[9] Karpagachelvi S., Arthanari M. and Siva Kumar M., "Classification of Electrocardiogram signals with extreme learning machine and relevance vector machine", International Journal of Computer Science Issues, Vol.8, Issue 1, pp.338-345.

[10] Tinati M.A. and Mozaffary B., "A wavelet packets approach to electrocardiograph baseline drift cancellation", International Journal of Biomedical Imaging, 2006, pp. 1-9.

[11] Song M.H., Lee J., Cho S.P., Lee K.J. and Yoo S.K., "Support Vector Machine based arrhythmia classification using reduced features", International Journal of Control Automation and Systems, Vol. 3, No.4, 2005, pp.571-579.

[12] Wendy L. Martinez, Angel R. Martinez and Jeffrey L. Solka, "Exploratory Data Analysis with Matlab", Second Edition, CRC Press.

[13] http://www.statistixl.com/

[14] Zadeh A.E., Khazaee A., Ranaee V., "Classification of the electrocardiogram signals using supervised classifiers and efficient features", Computer Methods and Programs in Biomedicine, Vol.99, Issue 2, 2010, pp.179-194.

Manpreet Kaur She is working as Associate Professor in Department of Electrical and Instrumentation Engineering, Sant Longowal Institute of Engineering and Technology, Longowal. She has completed her B.Tech. from NIT Jalandhar and M.Tech. from Punjab University Chandigarh. Her area of interest is biomedical engineering.

A.S.Arora is Professor in Department of Electrical and Instrumentation Engineering, Sant Longowal Institute of Engineering and Technology, Longowal. He has completed his B.Tech., M.Tech and Ph.D. from IIT Roorkee.