Proceedings of the 28th IEEE
EMBS Annual International Conference
New York City, USA, Aug 30-Sept 3, 2006

Intelligent Arrhythmia Detection and Classification Using ICA

Asad Azemi, *Senior Member IEEE*, Vahid R. Sabzevari, Morteza Khademi, Hossein Gholizade, Arman Kiani, and Zeinab S. Dastgheib

Abstract— **In this paper a novel approach for cardiac arrhythmias detection is proposed. The proposed method is based on using Independent Component Analysis (ICA) and wavelet transform to extract important features. Using the extracted features different machine learning classification schemas, MLP and RBF neural networks and K- nearest neighbor, are used to classify 274 instance signals from the MIT-BIH database. Simulations show that multilayer neural networks with Levenberg-Marquardt (LM) back propagation algorithm provide the optimal learning system. We were able to obtain 98.5% accuracy, which is an improvement in comparison with the similar works.**

I. Introduction

The classification of an electrocardiogram (ECG) into different pathophysiological disease categories is a complex pattern recognition task. Computer based classifications of the ECG can achieve high accuracy and offer the potential of an affordable mass screening for cardiac abnormalities. Successful classification is achieved by finding the characteristic shapes of the ECG that discriminate effectively between the required diagnostic categories. Conventionally, a typical heart beat is identified from the ECG and the component waves of the QRS, T and possibly P waves are characterized using measurements such as magnitude, duration and area. Datasets that are used for heart diseases involve different features. Some of them are based on laboratory experiments, while others include clinical symptoms. However, one of the most popular and useful databases is the MIT-BIH [1]. Researchers have used this database to test their various algorithms for arrhythmia detection and classification. Two of the most popular methods are artificial neural networks (ANNs) and wavelet transforms, and their variations. For example, Lee [2] classified three types of cardiac arrhythmias with accuracies of 99.55%, 97.75%, 57.1%, respectively using ANNs. Chi et al. [3] using triple neural networks classified ventricular arrhythmia with the average accuracy of 95.1%. Karlik et al. [4] were able to classify 10 types of arrhythmia with average accuracy of 95%. Yu et al. [5], using self organizer neural networks and the Linear Vector Quantization, were able to classify heart beats with average accuracy of 91.3% and 90.3%, respectively. Hosseini et al. [6] implemented a multi stage neural network with two MLPs and were able to classify five kinds of arrhythmia. At the first stage, they achieved an average accuracy of 81.8% and at the second stage, 88.3%. The algorithms developed in these works are based on analyzing the ECG signals, which requires a large simulation time.

Selected examples of using wavelet transforms and their variations for arrhythmia classification include the following. Yang et al. [7], using dyadic wavelets to extract features and Kohonen self organizing neural networks for classification, were able to obtain an average precision of 97.77% for heart disease diagnosis. Dokur et al. [8] used discrete wavelet transforms to classify ten types of arrhythmias with a precision of 97%. Chazal et al. [9], using a set of 500 records with 345 abnormal cases, classified arrhythmias by using 15 feature sets of three Daubiches wavelets decomposition level and reached the maximum precision of 74.2%. Finally Dokur et al. [10], using wavelet transforms and ANNs trained by genetic algorithms back propagation, was able to classify 10 types of arrhythmia with a precision of 96%.

The goal of this paper is to optimize the feature extraction process by using ICA and wavelet transform, apply the obtained set to several different machine learning schemes, and compare their performances. The paper is structured as follows. Section II describes our proposed method for cardiac arrhythmias detection. Section III covers an overview of different classifier types that were used in this work. Sections IV and V summarizes our simulation scheme and results. Finally, section VI presents the concluding remarks.

II. Proposed Method

Figure 1, presents the block diagram of the proposed detection and classification process. First, the appropriate components of the ECG signal are obtained by using the ICA algorithm. Next, these components are used to calculate the coefficients of the Daubechies wavelets. Based on this step, proper features are selected and fed into the classifier. For comparison purposes three different machine learning methods have been implemented, namely, RBF, MLP and KNN.

A. Azemi is with Engineering Department, Penn State University, Delaware County Campus, Media, PA 19063 (phone: 610-892-1421; fax: 610-892-1490; e-mail: azemi@psu.edu)
V.R. Sabzevari is a M.S. student in Bioengineering at the Azad University of Mashhad, Iran (email: vahidreza.sabzevari@gmail.com)
M. Khademi is with Electrical Engineering Department, Ferdowsi University of Mashhad, Iran (email: khademi@ferdowsi.um.ac.ir)
H. Gholizade is a PhD. student in Electrical Engineering at the Ferdowsi University of Mashhad (email: h_golizade@yahoo.com)
A. Kiani is a B.S. student in Electrical Engineering at the Ferdowsi University of Mashhad (email: kiani@kiaeee.org)
Z.S. Dastgheib is a M.S. student in Electrical Engineering at the Ferdowsi University of Mashhad (email: z.dastgheib@gmail.com)

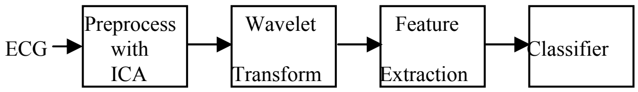


Fig1: Block diagram of the proposed detection-classification system

Next, we will present an overview of the ICA and wavelet transform.

A. Independent Component Analysis with a Time Structure Method

To rigorously define ICA, we can use a statistical "latent variables" model. We observe n random variables $x_1,\ldots,x_n$, which are modeled as linear combination of n random variables $s_1,\ldots,s_n$:

$$x_i = a_{i1}s_1 + a_{i2}s_2 + \ldots + a_{in}s_n \qquad i = 1,\ldots,n \qquad (1)$$

Where a_{ij}, $i,j=1, \ldots, n$ are some real coefficients. By definition, the s_i are statistically mutually independent.

This is the basic 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 s_i. The independent components s_i (often abbreviated as ICs) are latent variables, meaning that they cannot be directly observed. Also the mixing coefficients a_{ij} are assumed to be unknown. All we observe are the random variables x_i, and we must estimate both the mixing coefficients a_{ij} and the ICs s_i using the x_i.

Here we have dropped the time index t, because in this basic ICA model, we assume that each mixture x_i, as well as each independent component s_i, is a random variable, instead of a proper time signal or time series.

In this research, we consider the estimation of the ICA model when the ICs are time signals, $s_i(t)$, $t=1,\ldots,T$, where t is the time index. Here t has a more precise meaning, since it defines an order between the ICs. The model is then expressed by

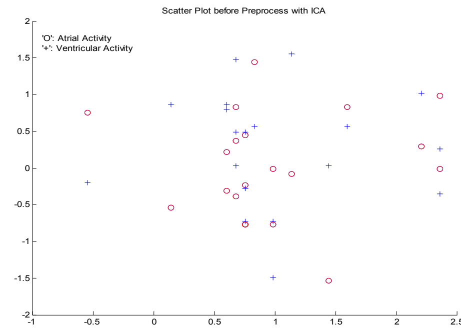
$$x(t) = As(t) \qquad (2)$$

where A is assumed to be square as usual and the ICs are of course independent. We shall make some assumptions on the time structure of the ICs that allow for the estimation of the model. First, we shall assume that the ICs have different autocovariances (in particular, they are all different from zero). Second, we shall consider the case where the variances of the ICs are nonstationary [11].

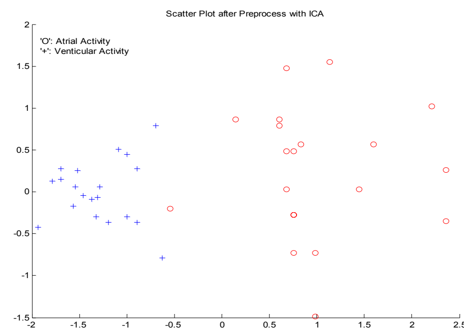
In this paper, we have implemented "Fast ICA" algorithm to calculate ICs [12]. According to the previous research [13, 14], at least 3 factors make ICA suitable for ECG signal analysis. These factors are as follow:

1) Atrial and ventricular activity (AA, VA) are generated by sources of independent bioelectric activity;
2) AA and VA present non-Gaussian distributions and
3) the generation of the surface ECG potentials from the cardioelectric sources can be regarded as a narrow-band linear propagation process.

Figure 2 shows results of the ICA processing and the original signal, which was generated by two independent sources of atrial and ventricular electrical activity (in such a ventricular arrhythmia). As shown in this figure ICA with time structure successfully described these two sources that are characterized by 'o' and '+' symbols in this figure.



(a)



(b)

Fig. 2: Scatter Plots (a) before denotation of independent sources by ICA and (b) after denotation by Matlab ICA pack.

B. Wavelet Analysis

Wavelet Analysis of a signal consists of breaking up a signal into a translated (shifted) and dilated (scaled) version of a reference (mother) wavelet. A wavelet is a signal of effectively limited duration that has an average value of zero. In determining the wavelet (decomposition) coefficients of a signal, the correlation of the mother wavelet at different translations and dilations with the signal is computed. Hence, the wavelet coefficients represent measures of similarity of the local shape of the signal to the mother wavelet under different translations and dilations. We utilize the feature extraction ability of this analysis for ECG. In this work we have used Daubechies wavelets; this choice is explained in section IV.

III. CLASSIFIERS

The results obtained from the previous section were fed into three different machine learning algorithms, namely MLP, RBF and K- nearest neighbor. An overview of each of these methods follows:

A. Multi-layer Perceptron

The most widely used neural classifier today is a Multilayer Perceptron (MLP) network which has also been extensively analyzed and for which many learning algorithms have been developed. The MLP belongs to the class of supervised neural networks. The network has a simple interpretation as a form of an input-output model, with the weights and thresholds (biases) the free parameters of the model. Such networks can model functions of almost arbitrary complexity, with the number of layers, and the number of units in each layer, determining the function complexity. Multi-layer networks use a variety of learning techniques, the most popular being back-propagation. Here the output values are compared with the correct answer to compute the value of some predefined error-function.

B. *Radial Basis Function*

Radial-Basis Function Networks used for pattern classification are based on Cover's theorem on the separability of patterns. It contains an input layer, a hidden layer with nonlinear activation functions and an output layer with linear activation functions. The RBF network is a popular alternative to the MLP which, although it is not as well suited to larger applications, can offer advantages over the MLP in some applications. The RBF network has a similar form to the MLP in that it is a multi-layer, feed-forward network. However, unlike the MLP, the hidden units in the RBF are different from the units in the input and output layers: they contain the "Radial Basis Function," a statistical transformation based on a Gaussian distribution from which the neural network's name is derived.

C. *K-nearest neighbor classifier*

The nearest neighbor classifier algorithm is a method for classifying phenomena based upon observable features. In this algorithm, each feature is assigned a dimension to form a multidimensional feature space. A training set of objects, with a priori known class are processed by feature extraction and plotted within the multi-dimensional feature space. The offsets in each dimension are referred to as the feature vector. This is the training or learning stage. The geometric distance is computed between the new feature vector and each a priori feature vector from the training set. The shortest distance thus computed is to the nearest neighbor. The a priori class of the nearest neighbor is now assigned to the phenomena to be classified.

The simulation results of using the aforementioned profilers showed that the MLP provides the best classification. Details of the MLP classifier are presented in the next section. Table 1 summarizes the comparative simulation results:

TABLE 1
A comparison of the Different Profiler Performance

Method	Overall Accuracy
RBF ($\sigma = 0.3$)	71.25%
KNN (k=5)	85.00%
MLP (3 layers)	98.56%

IV. SIMULATION

For the simulation part we have used MATLAB's Neural Networks and Wavelet toolboxes, along with its Fast ICA package for pattern recognition and classification.

Five data sets, one normal and four abnormal, with modified lead II (MLII) and lead V1 from MIT-BIH were used in training and testing stages. The first set was from the NSR (Normal Sinus Rhythm), the second set from PVC (Premature Ventricular Contraction), the third set from LBBB (Left Bundle Branch Block), the fourth set from RBBB (Right Bundle Branch Block) and finally, the fifth set from P (Paced Beat) data bases. Dilation coefficients of Daubechies wavelets levels 6, 7, 8 of MLII and dilation coefficient levels 6, 7 of V1 samples after denotation by ICA were implemented as features. Next the wavelet coefficients were normalized and then fed into MLP and the classification results were obtained. Table 2 illustrates the data distribution and the number of records used for this part of the simulation. Wavelet levels were chosen based on the following considerations: power spectrum for normal beat, LBBB, RBBB, and PVC indicates that the frequency information are located in mid frequencies and high frequencies, respectively, and low order Daubechies wavelets have good time resolution but poor frequency resolution whereas high order wavelets have good frequency resolution and poor time resolution. Based on these reasons, levels 3 to 8 were considered, where the best results were obtained by using the aforementioned levels.

The MLP, which is implemented in this paper, is a three layered neural networks. The optimum topology was obtained with the following structure: five input neurons, eight hidden neurons, five output neurons, and the Levenberg-Marquardt as the training algorithm. The desired error was set at 0.01. The system was tested for 1-10 hidden neurons and four different discriminator functions (activation functions).

TABLE 2
Data Set Descriptions and Numbers Used in the Simulation

Class No.	Record no. used from MIT-BIH	No. of beats used in training	No. of beats used in testing	Description
1	101, 105	100	50	Normal Beat
2	200, 213	23	12	PVC
3	207, 214	20	10	LBBB
4	118, 212	21	8	RBBB
5	107, 217	20	10	P
Total =274		184	90	

V. SIMULATION RESULTS

Table 3 illustrates the comparison of different activation functions that have been used in the MLP network that was described in the previous section. An entry such as "tansig-logsig" means that "tansig" is the activation function of the hidden neuron and "logsig" is the activation function of the output layer. As the table indicates optimum results can be obtained with the "tansign-tansig" activation functions with an overall accuracy of 98.56%, during testing. This is an improvement in comparison to the ones obtained in [15].

TABLE 3
The Overall Accuracy in Testing of MLP for Different Activation Function (100% accuracy in training)

Activation functions	MLP Accuracy (%)
tansig-tansig	98.56
logsig-tansig	91.667
tansig-logsig	90.278
logsig-logsig	93.056

VI. Conclusion

In this paper we have shown a new method for cardiac arrhythmias detection and classification based on using Independent Component Analysis and wavelet transforms to extract important features. The extracted features were then fed to an MLP with the Levenberg-Marquardt back propagation algorithm. For the simulation part, 274 instances of MIT-BIH database from five disease categories were used and a 98.5% accuracy was obtained, which is an improvement with respect to similar works

According to the simulations, the type of activation function and the number of hidden nodes are the important factors in the topology of neural networks. We note that for extending these results we need to use a larger number of samples from the database.

References

[1] "MIT-BIH arrhythmia database directory", [Online]. Available: http://physionet.fri.uni-lj.si/physiobank /database/mitdb/.

[2] S.C. Lee, "Using A Translation-Invariant Neural Network to Diagnose Heart Arrhythmia," *Proc. of Annual International Conference of the IEEE Eng. of Med.& Biology Society*, 12, pp.2025, 1990.

[3] Z. Chi, and M. A. Jaberi, "Identification of Supraventricular and Ventricular Arrhythmias Using a Combination of 3NNs," *Dept. of Electrical Eng. Univ. of Sydney*, 1992.

[4] Bekir Karlik, and Yuksel Ozbey, "A New Approach for Arrhythmia Classification", *Proc. of Annual International Conference of the IEEE Eng. of Med. & Biology Society*, 1996.

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

[6] H. Gholam Hosseinie, K.J. Rainolds, and D. Powers, "A Multi-Stage NN. Classifiers for ECG Events," *Proc. of Annual International Conference of the IEEE Eng. of Med. & Biology Society*, 2001.

[7] Ming-Yao Yang, Wei-Chih Hu and Liang-Yu Shyu, "ECG Events Detection and Classification Using Wavelet and Neural Networks," *Proceedings of 19th International Conference IEEE/EMBS*, Oct. 30-Nov. 2, 1997, Chicago, IL. USA.

[8] Z. Dokur, T. Olmez and E. Yazgan, "Comparison of discrete wavelet and Fourier transforms for ECG beat classification," *IEEE Electronic Letters, September 1999*, vol. 35, no. 18.

[9] P. de Chazal, B. G. Celler and R. B. Reilly, "Using wavelet coefficients for the classification of the electrocardiogram," *Proceedings of the 22nd Annual EMBS International Conference*, July 23-28, 2000, Chicago, IL. USA.

[10] Zumray Dokur, and Tamer Olmez**,** "ECG beat classification by a novel hybrid neural network," *ELSEVIER, Computer Methods and Programs in Biomedicine 66*, pp. 167–181, 2001.

[11] Aapo Hyvarinen, Juha Karhunen, and Erikki Oja, *"Independent Component Analysis,"* Wiley-Interscience, 2001.

[12] J. Hurri, J. Gavert, J. Sarela and A. Hyvarinen. The Fastica Package for Matlab, (2005). [Online]. http://www.cis.hut.fi/projects/ica/fastica

[13] José Joaquín Rieta, Francisco Castells, César Sánchez, Vicente Zarzoso and José MilletAtrial, "Activity Extraction for Atrial Fibrillation Analysis Using Blind Source Separation," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 7, July 2004.

[14] F. Castells, J. J. Rieta, J. Millet and V. Zarzoso "Spatiotemporal Blind Source Separation Approach to Atrial Activity Estimation in Atrial Tachyarrhythmias," *IEEE Trans. on Biomedical Engineering*, vol. 52, no. 2, February 2005.

[15] Liang-Yu Shyu, Ying-Hsuan Wu, and Weichih Hu, "Using Wavelet Transform and Fuzzy Neural Network for VPC Detection From the Holter ECG," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 7, July 2004.