Research Article
Combination of Nonlinear Feature Extraction Techniques for ECG Arryhtmia Detection System

[1, 2]Fatin Ahmed Elhaj, [1, 3]Arief R. Harris, [3]Tan Tian Swee and [2]Naomie Salim
[1]Neural Engineering Lab, Centre for Biomedical Engineering,
[2]Soft Computing Research Group, Faculty of Computing, Universiti Teknologi Malaysia,
Skudai 81310, Malaysia
[3]Faculty of Biosciences and Medical Engineering, Universiti Teknologi Malaysia, Johor, Malaysia

Abstract: Electrocardiogram (ECG) is a noninvasive method used to detect arrhythmias or heart abnormalities. It describes the electrical activity of the heart. Physicians are faced with difficulties in detecting irregular heartbeats due to the presence of noise and subtle changes in the signal amplitude and duration. Depending on human visual detection alone may lead to misdiagnosis or insignificant detection of cardiovascular diseases. A computer-aided diagnosis of the ECG will assist physicians to significantly detect the cardiovascular diseases. Non-linear method is useful to extract and capture hidden information in the ECG signal. In this study we present a combination of two nonlinear methods; Higher Order Statistics (HOS) cumulants and Independent Component Analysis (ICA), performed on the dynamics ECG signals for arrhythmia detection. The abnormal heartbeats focused in this study are Right Bundle Branch Block (RBBB), Left Bundle Branch Block (LBBB), Atrial Premature Contraction (APC), Ventricular Premature Contraction (VPC) and Paced beat (P). The HOS and ICA features were fed to the Support Vector Machine (SVM) with kernel functions for automatic classification and neural network. In our study, we obtained the highest accuracy of 96.34% with Radial Basis Function (RBF) kernel and neural network.

Keywords: ECG, feature extraction, HOS, ICA, supervised learning algorithm

INTRODUCTION

The occurrence of Cardiovascular Diseases (CVD) has increased in recent years (Thom *et al*., 2006). The increase in death rates due to CVD is associated with the epidemiological transition of unhealthy lifestyles such as obesity, diabetes mellitus and smoking habit (Thom *et al*., 2006). The American Heart Association (AHA) stated that in 2006 over 70 million people were diagnosed with CVD. The common causes of CVD are high blood pressure, insufficient physical exercise, poorly balanced diet, smoking and abnormal glucose levels (Sidney *et al*., 2013). Two of the main factors of CVD are atrial and ventricular arrhythmias which occur during cardiac rhythm. Arrhythmia is a cardiac condition caused by the abnormal electrical activity of the heart. Cardiac arrhythmias generally occur in people with diseases, such as cardiomyopathy, hypertension and coronary artery disease. The electrical activity of the heart can be recorded using surface electrodes. The recording describes the main direction of electrical impulses throughout the heart at any portion of time and is known as Electrocardiogram (ECG). The ECG is a noninvasive method used to detect arrhythmias or heart abnormalities. Due to the presence of noise and

irregular heartbeat, physicians are faced with difficulties in the diagnosing arrhythmias. Furthermore, visual inspection alone may lead to misdiagnosis or insignificant detection of CVD (Goldberger, 2012). Therefore, a computer-aided diagnosis of ECG will assist physicians to significantly detect the cardiovascular diseases (Shiyovich *et al*., 2010; Hadhoud *et al*., 2006). There are three main procedures in ECG arrhythmia classification, namely feature extraction, feature selection and classifier construction. Feature extraction plays an important role in detecting most of the heart diseases, extract hidden information present in the ECG signals and improve the classification performance. Higher Order Statistics (HOS) cumulants and Independent Component Analysis (ICA) are popular methods for feature extraction (Osowski and Linh, 2001; Yu and Chou, 2008).

The nature of ECG signal is non-stationary (i.e., the morphology of ECG signal changes with respect to time) and these variations are present not only between different patients but also within the same patients (Osowski and Linh, 2001). In addition, the noise such as power line interference, baseline drift, muscle tremor and movement electrodes always exist in the ECG

signals (Martis *et al*., 2014). Based on (Osowski and Linh, 2001; Martis *et al*., 2014) the nonlinear methods characterize the signal in a more significant sense and achieve well under noisy situations, for example using HOS cumulants provide a very good recognition and less effect to the morphological differences of the ECG signal.

Furthermore, a nonlinear method of dimensionality reduction ICA is a computational statistical technique that is very effective to disclose hidden factors underlying in mixed samples of random variable measurements (Jutten and Herault, 1991; Martis *et al*., 2014). This technique was originally developed to deal with situations that are closely related to Blind Source Separation (BSS) (e.g., the 'cocktail party' problem (Jutten and Herault, 1991; Mantini *et al*., 2008). In fact, it is known to be an efficient method in performing the extraction of the fundamental features from multiple aggregated observations (Jutten and Herault, 1991).

In this study, we propose a combination of extracted hidden features from ICA and HOS cumulants computation approach for pattern classification of normal and abnormal ECG signals. We tested the effect of the combination features on SVM and Neural Network classifiers to classify six types of

arrhythmia classes such as (Normal (N), Right Bundle Branch Block (RBBB), Left Bundle Branch Block (LBBB), Atrial Premature Contraction (APC), Ventricular Premature Contraction (VPC) and Paced beat (P).

MATERIALS AND METHODS

MIT-BIH arrhythmia database were used in this study where the signals were sampled at 360 Hz. The database consists of 48 signals, each with half an hour duration of ambulatory ECG recording. Twenty-three recordings were selected randomly from a set of 4000 24-h ECG recordings. The remaining 25 recordings were selected from the same set to include other arrhythmias that may not be well represented in a small and random sample group because they are less common. In this study the ECG beat types obtained from the MIT-BIH include Normal (N), Right Bundle Branch Block (RBBB), Left Bundle Branch Block (LBBB), Atrial Premature Contraction (APC) and Ventricular Premature Contraction (VPC) and Paced beat (P) arrhythmias. Table 1 summarizes data distribution of the six types of arrhythmia and Fig. 1 shows the six types of ECG beat samples. The R-peak

Table 1: Distribution of the six types of arrhythmia

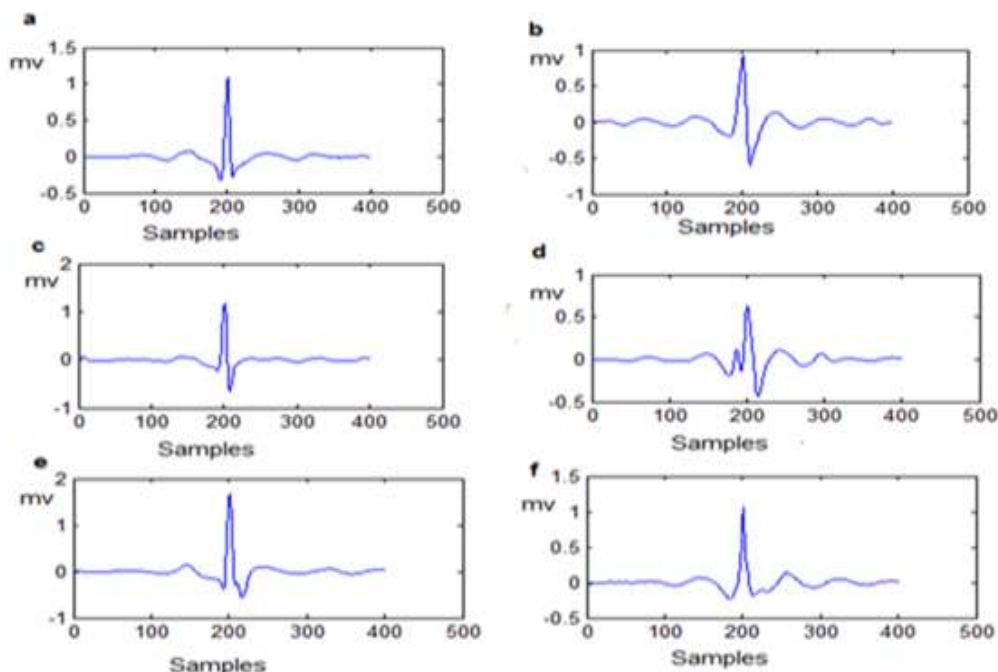| Type	Subject	Total no of beats
Normal	100, 101, 103, 105, 108, 112, 113, 114, 115, 117, 121, 122, 123, 202, 205, 219, 230, 234	1800
PVC	106, 119, 200, 203, 20, 221, 228, 233, 201, 210, 215, 116	1230
APC	209, 222, 232, 220, 223	382
LBBB	109, 111, 207, 214	400
RBBB	118, 124, 212, 231	400
Paced	102, 104, 107, 217	400

Fig. 1: Six types of ECG beat; (a): Normal; (b): PVC; (c): APC; (d): LBBB; (e): RBBB; (f): Paced
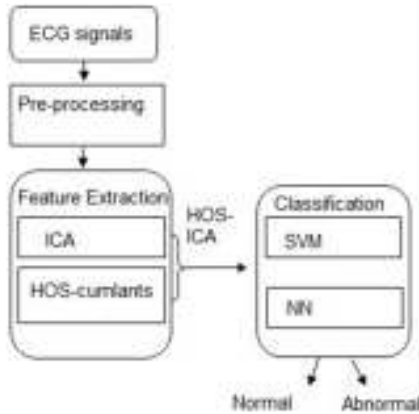
Fig. 2: Diagram of the proposed model

detection involves the pre-processing step. The R peak is the most important peak in the ECG detection due to its high amplitude (Wang *et al*., 2013; Shet *et al*., 2012).

Using Fast Fourier Transform (FFT) in order to eliminate the low frequency components. The inverse of FFT was used to rebuild signal from frequency domain to time domain.

Furthermore, using double filtering technique in R peak detection. The first one filter is apply a windowed filter that involves finding only the maximum peaks in the window. The second filter uses the output of the first filter and calculate a threshold to obtain a optimum set of the R peaks. Once the R peak is detected, the ECG is then segmented into 200 samples before and 200 samples after R peak. These segmented signals are passed to feature extraction algorithm to locate more hidden data. The features are splitting into training and testing using 10 fold cross validation.

Figure 2 shows the proposed system used for the dynamic examination of six types of beats in ECG of arryhtmia. The ECG arrhythmia classification system is involves data acquisition, feature extraction, feature reduction and classification. Each part is explained in the following sections.

In this study, the ECG arrhythmia recordings are acquired from the MIT-BIH arrhythmia database.

Independent Component Analysis (ICA): ICA is a computational method that involves a multi-variant analysis to a multi-source signal into additive subcomponents. It can be represented mathematically as follows:

$$x = A \cdot S \tag{1}$$

where, x is the matrix of n observed signals:

$$x = [x_1, x_2, \ldots, x_n]^T$$

S is the matrix of m underlying signals:

$$S = [S_1, S_2, \ldots, S_m]^T$$

And A is the mixing matrix of (n×m). The number of Independent Components (ICs) can be calculated by finding an un-mixing matrix named as W where:

$$S = w \cdot x \tag{2}$$

In order to calculate w, the data must be centered and whitened (10). Centering is obtained by subtracting the jth average value from each mixed signal x_j, so that the underlying signals S_i become zero-mean.

Whitening is a linear transformation of matrix x into another matrix x^-. The most common method for whitening is the eigenvalue decomposition of the covariance matrix:

$$Exx^T = EDE^T \tag{3}$$

where, E is the orthogonal matrix of eigenvectors of Exx^T, D is the diagonal matrix of its eige values $D = diag (d_1, d_2, \ldots, d_n)$. The whitened matrix x^- can be calculated as:

$$x^- = ED^{1/2}E^T x \tag{4}$$

Whitening transforms the un-mixing matrix w for which:

$$S^\wedge = w \cdot x = wED^{-1/2}E^T x^- = w^- x^- \tag{5}$$

Since w^- is orthogonal, this minimizes the number of parameters to be estimated. Therefore, the un-mixing matrix w:

$$w = w^- ED^{-1/2}E^T \tag{6}$$

Subsequently, the matrix A is obtained from the equation:

$$A = (w^T w)^{-1} w^T \tag{7}$$

High Order Statistics (HOS) cumulants: HOS is a technique used for the analysis of nonlinear, non-stationary and non-Gaussian signals (Martis *et al*., 2013a; Osowski and Linh, 2001). The nonlinearity of HOS has been generated from the Gaussian deviation and phase mutual relations between various frequency components. Moreover, HOS removes Gaussian noise. As Gaussian process is categorized by the first two order statistics, the first and second order statistics are not significant to represent the nonlinear model. Therefore, the third and fourth order statistics are used in this analysis. Higher order cumulants are derived

from higher order moments to represent deviation from Gaussianity and can construct the dynamic non-stationary natures of ECG signals. Suppose that $x(n)$ is a stationary signal of a discrete time and its moments are available up to the order n. Then the nth order moment function is defined by:

$$m_n^x(\tau_1, \tau_2, ..., \tau_{n-1}) = E[x(n)x(n+\tau_1)...x(n+\tau_{n-1})] \qquad (8)$$

The moment function depends only on the time lags $\tau_1, \tau_2, ..., \tau_{n-1}$, $\tau_i = 0, \pm1, \pm2,...$ for all i. The 2nd order moment $m_2^m(\tau_1)$ is the autocorrelation sequence of x (n) and is calculated as follows:

$$m_2^m(\tau_1) = E[x(n)x(n+\tau_1)] \qquad (9)$$

whereas the 3rd and 4th orders moment functions are given by $m_3^x(\tau_1, \tau_2)$ and $m_4^x(\tau_1, \tau_2, \tau_3)$ respectively. E [.] is the statistical expectation operator. Using the nth order moment, the nth order cumulant can be computed as:

$$C_n^x(\tau_1, \tau_2, ..., \tau_{n-1}) = m_n^x(\tau_1, \tau_2, ..., \tau_{n-1}) - m_n^G(\tau_1, \tau_2, ..., \tau_{n-1})$$

where, $m_n^x(\tau_1, \tau_2,..., \tau_{n-1})$ is the nth order moment function and $m_n^G(\tau_1, \tau_2,..., \tau_{n-1})$ is the nth order moment of Gaussian process. The first four order cumulants for a zero mean process are given by:

$$
\begin{aligned}
C_1^x &= m_1^x \\
C_2^x &= m_2^x(\tau_1) \\
C_3^x &= m_3^x(\tau_1, \tau_2) \\
C_4^x &= m_4^x(\tau_1, \tau_2, \tau_3) - m_2^x(\tau_1)m_2^x(\tau_2 - \tau_3) \\
&\quad - m_2^x(\tau_2)m_2^x(\tau_3 - \tau_1) - m_2^x(\tau_3)m_2^x(\tau_1 - \tau_2)
\end{aligned}
\qquad (10)
$$

Neural Network (NN): Neural Network (NN) plays an important role in a wide variety of applications, such as pattern recognition and classification tasks. In NN model, each neuron computes the weighted sum of its inputs and applies the sum to a nonlinear function called activation functions. Multi-Layer Perception MLP is capable of recognizing and classifying ECG signals more accurately than other methods of NN. In general, the performance of MLP depends on the number of hidden layers, the number of hidden neurons, the learning algorithm and the activation function for each neuron (Bishop, 1995). The input layer consists of 16 nodes, corresponding to the 16 features used. A hidden layer of 10 neurons and the output layer of 6 neurons corresponding to six classes were used.

The 10 neurons in the hidden layers were chosen by trial and error function. Initially, we obtained highest accuracy using 12 neurons in the hidden layer but this resulted in high mean square error. To improve learning procedure the error back propagation method is used. Based on the computation of mean square error between the desired response and actual response of NN, the weights in the network were updated and the process is continued until the MSE is below specific threshold. The testing data is then forwarded to the trained NN to acquire the output and the testing patterns were classified.

Support Vector Machine (SVM): SVM classifier for single layer can supervise classification problems due to its capability of generalization (Vapnik, 1995). It converts the input vector patterns to higher dimension feature space through some non-linear mapping and obtains optimal separating hyper-plane which is built to separate two classes of samples. In particular, the SVM classifier shows a favorable generalization capability when using a Maximal Margin Principle (MMP). It maximizes the distance between the patterns and the class separating hyper-plane. An objective function is formulated based on the distances of the class separating hyper-plane and the optimization process is carried out (13). Different kernel transformations are used to map the data into high dimensional functions such as quadratic, polynomial and Radial Basis Function (RBF). The performance of SVM can be effected by hyper-parameter (C parameter and the kernel parameter), where these parameters determine the number of support vectors and the maximization margin of the SVM.

RESULTS

We have performed the ICA method to reduce the input features from 400 to 13 independent components that represent an ECG beat sample. The 2nd, 3rd and 4th cumulants coefficients were computed and used as features for our further study.

The transformed data points were formed into feature vectors inputted to the NN and SVM using various kernel functions including linear, polynomial and RBF. The SVM classification using RBF kernel

Table 2: Accuracy of the SVM and NN classifiers with ICA and HOS cumulants features

Features	Classifier	Accuracy (%)
ICA	SVM-polynomial	90.90
	SVM-RBF	95.45
	NN	95.45
HOS	SVM-polynomial	94.87
	SVM-RBF	92.30
	NN	96.00
ICA-HOS	SVM-polynomial	95.12
	SVM-RBF	96.34
	NN	96.34

Table 3: Classification performance comparisons of the proposed scheme with some existing schemes

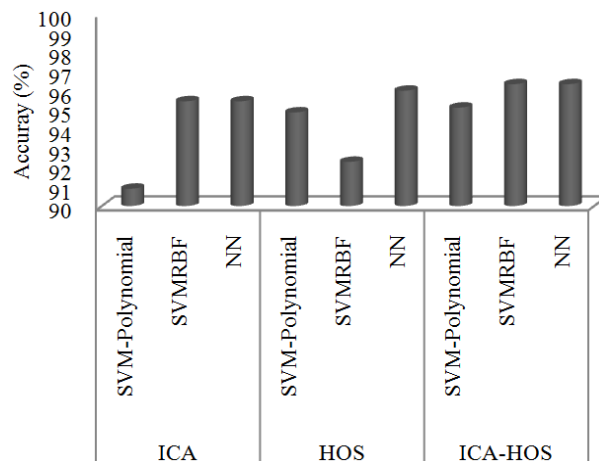| Literature	Features	Classifier	Class	Accuracy (%)
Oswski and Linh (2001)	HOS	Hybrid fuzzy NN	7	96.00
Martis *et al.* (2013b)	Bispectrum+PCA	SVM-RBF	5	93.48
Martis *et al.* (2013a)	Cumulant+PCA	NN	5	94.52
Proposed model	HOS-ICA	SVM-RBF	6	96.34

Fig. 3: Accuracy of the SVM and NN classifiers based on ICA and HOS cumulants features

and NN performed better than the other kernel functions. It was found that with 13 features of ICA, the accuracy of SVM-RBF is (95.45%), SVM-Polynomial is (90.9%) when the C parameter is equal to 78 and gamma equal to 0.7. In NN, the accuracy of (95.45%) is obtained using the 10 number of hidden layer. When three features were chosen from 2nd, 3rd and 4th cumulants, the accuracy of SVM-RBF is (92.30%), SVM-Polynomial is (94.87%) when the C parameter is equal to 78 and gamma equal to 0.7. In NN the accuracy of (96.0%) is obtained when the number of hidden layer is equal to 10. In total there were 16 features Including 13 (ICs) and 3 high order cumulants. The ten-fold cross validation was used to choose training and testing sets. It was found that the combination of ICA features and high order cumulates features provide the highest accuracy (96.34%) with SVM-RBF and NN. The accuracy of classification over ten folds is tabulated in Table 2 and represented in Fig. 3 (Table 3).

DISCUSSION

In this study, the features of the ECG beat were extracted using HOS cumulants method and ICA component. The combination of three features of HOS cumulants with 13 ICA features gave maximum classification accuracy compared to the other methods. It can be seen from our results that the proposed method is able to classify the normal and abnormal classes with accuracy of more than 95% for all kernel functions of the SVM classifier and NN classifier. In fact, RBF kernel and NN performed better than the SVM-Polynomial. HOS was applied to recognize the normal and different types of arrhythmias using fuzzy NN.

The cumulants of the 2nd, 3rd and 4th orders have been used for the feature selection. The recognition results of seven beat types show 96.06% accuracy (6). HOS bispectrum which is the third order spectra has been used to capture information beyond mean and standard deviation.

The Principal Component Analysis (PCA) used the extracted bi-spectrum features for dimensionality reduction. Using feed forward NN and Least Square-Support Vector Machine (LS-SVM), five types of beats were classified, namely: N, RBBB, LBBB, APC and VPC. Accuracy of 93.48% is obtained (Martis *et al.*, 2013b). HOS features which are the higher order cumulants have been used to recognize five types of ECG beats, namely: N, RBBB, LBBB, APC and VPC cumulant features and cumulants of discrete wavelet transform coefficients were used for classification. The features were subjected to data reduction using PCA and further classified using feed-forward NN and LS-SVM classifiers. The average accuracy of 94.52% was obtained, using cumulants features with NN classifier (11).

CONCLUSION

The ECG signal describes the electrical activity of the heart and can provide information to detect heart diseases. Physicians are faced with difficulties to diagnose heart diseases due to presence of noise and subtle changes in the morphology of ECG signal due to the variations of signal amplitude and duration with respect to time. In this study, a nonlinear dimensionality reduction method (ICA) on ECG signal combined with a nonlinear HOS features is proposed. It was shown that

the combination of ICA features and HOS along with SVM-RBF and NN classifiers gave the highest accuracy of 96.34%. In the future, the performance of the system can be improved with more diverse data and better features.

ACKNOWLEDGMENT

REFERENCES

Bishop, C.M., 1995. Neural Networks for Pattern Recognition. Oxford University Press, Inc., New York, NY, USA, ISBN: 0198538642.

Goldberger, A.L., 2012. Clinical Electrocardiography: A Simplified Approach. Elsevier Health Sciences, Philadelphia.

Hadhoud, M.M., M.I. Eladawy and A. Farag, 2006. Computer aided diagnosis of cardiac arrhythmias. Proceeding of the IEEE International Conference on Computer Engineering and Systems, pp: 262-265.

Jutten, C. and J. Herault, 1991. Blind separation of sources, Part I: An adaptive algorithm based on neuromimetic architecture. Signal Process., 24(1): 1-10.

Mantini, D., F. Petrucci, P. Del Boccio, D. Pieragostino, M. Di Nicola, A. Lugaresi, G. Federici, P. Sacchetta, C. Di Ilio and A. Urbani, 2008. Independent component analysis for the extraction of reliable protein signal profiles from maldi-tof mass spectra. Bioinformatics, 24(1): 63-70.

Martis, R.J., U.R. Acharya, C.M. Lim, K.M. Mandana, A.K. Ray and C. Chakraborty, 2013a. Application of higher order cumulant features for cardiac health diagnosis using ECG signals. Int. J. Neural Syst., 23(04).

Martis, R.J., U.R. Acharya, K. Mandana, A. Ray and C. Chakraborty, 2013b. Cardiac decision making using higher order spectra. Biomed. Signal Proces., 8(2): 193-203.

Martis, R.J., U.R. Acharya and H. Adeli, 2014. Current methods in electrocardiogram characterization. Comput. Biol. Med., 48: 133-149.

Osowski, S. and T.H. Linh, 2001. ECG beat recognition using fuzzy hybrid neural network. IEEE T. Bio-Med. Eng., 48(11): 1265-1271.

Shet, M.S., M. Patel, A. Rao, C. Kantharaj and K.V. Suma, 2012. ECG arrhythmia classification using R-peak based segmentation, binary particle swarm optimization and absolute euclidean classifier. In: Aswatha Kumar, M. *et al*. (Eds.), Proceeding of ICAdC. AISC 174, Springer, India, pp: 303-317.

Shiyovich, A., A. Wolak, L. Yacobovich, A. Grosbard and A. Katz, 2010. Accuracy of diagnosing atrial flutter and atrial fibrillation from a surface electrocardiogram by hospital physicians: Analysis of data from internal medicine departments. Am. J. Med. Sci., 340(4): 271-275.

Sidney, S., W.D. Rosamond, V.J. Howard and R.V. Luepker, 2013. The "heart disease and stroke statistics-2013 update" and the need for a national cardiovascular surveillance system. Circulation, 127(1): 21-23.

Thom, T., N. Haase, W. Rosamond, V.J. Howard, J. Rumsfeld, T. Manolio, Z.J. Zheng, K. Flegal, C. O'donnell and S. Kittner, 2006. Heart disease and stroke statistics-2006 update a report from the American heart association statistics committee and stroke statistics subcommittee. Circulation, 113(6): e85-e151.

Vapnik, V.N., 2000. The Nature of Statistical Learning Theory. Springer-Verlag, New York.

Wang, J.S., W.C. Chiang, Y.L. Hsu and Y.T.C. Yang, 2013. ECG arrhythmia classification using a probabilistic neural network with a feature reduction method. Neurocomputing, 116: 38-45.

Yu, S.N. and K.T. Chou, 2008. Integration of independent component analysis and neural networks for ECG beat classification. Expert Syst. Appl., 34(4): 2841-2846.