Journal of Applied Sciences

ISSN 1812-5654

Journal of Applied Sciences 10 (15): 1554-1562, 2010
ISSN 1812-5654

Automatic Heartbeats Classification based on Discrete Wavelet Transform and on a Fusion of Probabilistic Neural Networks

[1]M. Hendel, [1]A. Benyettou, [2]F. Hendel and [1]H. Khelil
[1]SIMPA Laboratory, Department of Computer Science, Faculty of Science,
University of Science Technology of Oran, Mohammed Boudiaf, USTO, Algeria
[2]LARESI Laboratory, Department of Electronic, Faculty of Electric,
University of Science Technology of Oran, Mohammed Boudiaf, USTO, Algeria

Abstract: In this study, we apply the wavelet transform and the fusion of two Bayesian neural networks to design an electrocardiogram (ECG) beat classification system; our main objective is to minimize at the maximum the miss diagnostic. Six ECG beat types namely: Normal beat (N), Left Bundle Branch Bloc (LBBB), Right Bundle Branch Block (RBBB), Premature Ventricular Contraction (PVC), Atrial Premature contraction (APB) and the Paced Beat (PB), obtained from the MIT-BIH ECG database were considered. First, five Discrete Wavelet Transform (DWT) levels were applied to decompose each ECG signal beat (360 samples centered on the peak R) into time-frequency representations. Statistical features relative to the three last decomposed signals and the last approximation, in addition to the original signal, were then calculated to characterize the ECG beat. Secondly, the fusion of RBFNN (Radial Basis Functions Neural Network) and BPNN (Back Propagation Neural Network) was employed as the classification system using as inputs, the calculated statistical features, in addition to the instantaneous RR interval. The proposed method achieves an equally well recognition rate of over 98% throughout all ECG beats type. These observations prove that the proposed beat classifier is very reliable and that it may be a useful practical tool for the automatic detection of heart diseases based on ECG signals.

Key words: Electrocardiogram, MIT-BIH database, Bayesian neural networks, Dempster-Shafer fusion, discrete wavelet transform

INTRODUCTION

The Electrocardiogram (ECG) is a bio-electric signal that records the electrical activities of the heart. It provides helpful information about the functional aspects of the heart and cardiovascular system. The state of cardiac health is generally reflected in the shape of ECG waveform and heart rate. It may contain important pointers to the nature of diseases afflicting the heart. Since, the ECG records are no stationary signals, this indication may occur at random in the time scale. In this situation, the disease symptoms may not come all the time, but would manifest at certain irregular intervals during the day. Therefore, 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 monotonous and time consuming. Naturally, the possibility of the analyst missing (or misreading) critical information is high. Therefore, computer based analysis and classification and automatic interpretation of the ECG signals can be very helpful to assure a continuous surveillance of the patients and to prepare the work of the cardiologist in the analysis of long recordings. It is a privileged domain in biomedical computer science applications. Numerous methods have been proposed during these last years for detection and identification assisted by computer of cardiac arrhythmias from the ECG signals (Ceylan and Ozbay, 2007; Yu and Chen, 2006).

Several approaches of extraction of ECG features have been proposed in the study. They include time domain (Challis and Kitney, 1990; Hu *et al*., 1997; De Chazel and Reilly, 2003; Jekova *et al*., 2008; Chen, 2007), frequency domain (Minami *et al*., 1999; Khadra *et al*., 2005; Moraes *et al*., 2002), or represented as statistical measure (Osowski and Linh, 2001). Although, these methods showed remarkable results in some classification tasks, they usually failed to demonstrate equally well discrimination ability throughout all ECG beats types (Yu and Chen, 2006). The WT can be

Corresponding Author: Hendel Mounia, SIMPA Laboratory, Department of Computer Science, Faculty of Science, University of Science Technology of Oran, Mohammed Boudiaf, USTO, Algeria

considered as an extension of the classic Fourier transform, except that, instead of working on a single scale (time or frequency), it works on a multi-scale basis. This multi-scale feature of the WT allows the signal decomposition into a number of scales, each scale representing a particular feature of the signal under study. In the consequence, DWT is frequently used (Ubeyli, 2007; Yu and Chen, 2006, 2009; Prasad and Sahamb, 2003).

As for classifiers, artificial neural networks have been extensively employed in computer aided diagnosis, because of their remarkable qualities and best results: capacity of adapting to various problems, training from examples and generalization, resoluteness to the noise. They are explored in real time processes and they are very efficient for training processes with high difficulty of modeling. The Multi-Layer Perceptron (MLP) is one of the most popular artificial neural networks used in ECG classification (Vargas *et al*., 2002; Minami *et al*., 1999), Self-Organizing Maps (SOMs) have also been applied for classifying ECG beats (Lagerholm *et al*., 2000). And recently, we have seen a growing number of combined neural network approaches to ECG classification. For example, the neuro-fuzzy network has been used to design the ECG classifier (Osowski and Linh, 2001; Engin, 2004). The neurofuzzy network is composed of two subnetworks connected in cascade: the fuzzy self-organizing layer performing the pre-classification task and the following MLP working as the final classifier.

In this study, we also combine and fuse two classifier architectures in order to discriminate the ECG beats. Indeed, the fact that having more than one primary classifier in the system, should improve its performances. In general, different classifiers work on different principles and tend to make independent errors. By combining the results of several such systems, we might assume that the errors will annul and that the overall performances of the system will improve.

First, five Discrete Wavelet Transform (DWT) levels were applied to decompose each ECG signal beat (360 samples centered on the peak R) into time-frequency representations. Statistical features relative to the three last decomposed signals and the last approximation, in addition to the original signal, were then calculated to characterize each ECG beat. For some heart disease, the RR interval provides useful information for clinical diagnosis. Thus, the instantaneous RR interval is added as another important feature for charaTcterizing an ECG beat.

Secondly, these parameters are used for training two Bayesian neural network classifiers: MLPNN and RBFNN for recognizing six ECG beat types namely: Normal beat (N), Left Bundle Branch Bloc (LBBB), Right Bundle Branch Block (RBBB), premature ventricular contraction (PVC), Atrial Premature Contraction (APC) and the Paced Beat (PB). The outputs of the two classifiers are then fused using the Dempster Shafer rule (Shafer, 1976). The main goal of this work is to realize a robust classifier able to identify all the above cited heartbeats types using neural networks based on a Bayesian framework and DWT.

FEATURE EXTRACTION

The discrete wavelet transformation (Ceylan and Ozbay, 2007; De Chazel and Reilly, 2003; Yu and Chen, 2009), the proposed wavelet-based feature extraction and feature normalization are described here.

Discrete wavelet transformation: The processing of the information by the heart is reflected in dynamical changes of the electrical activity in time, frequency and space. Therefore, the related studies require methods able of describing the qualitative variation of the signal both in time and frequency. Wavelet Transform (WT) is a powerful technology which works on a multi-scale basis. This multi-scale feature of the WT allows the decomposition of a signal into a number of scales, each scale representing a particular characteristic of the signal under study. The WT of signal x(t) is defined as:

$$\psi_a x(b) = \frac{1}{\sqrt{a}}\int_{-\infty}^{+\infty} x(t)\,\psi\left(\frac{t-b}{a}\right)dt \tag{1}$$

where, $\Psi(t)$, a, b$\in$R and a $\neq$ 0, are the mother wavelet, the dilatation and the translation factors, respectively. The dilatation and translation factors are used in wavelet transform to achieve different frequency and time localization.

Each stage of multi-resolution decomposition of a signal x(n) consists of two digital filters. The first filter g (n) is high-pass and the second h(x) is its mirror version, low-pass. The outputs of first high-pass and low-pass filters provide the detail, D1 and the approximation, A1, respectively. The first approximation, A1 is further decomposed and this process is continued (Fig. 1).

Feature vectors: Selection of the classifier inputs is the most important component of designing the neural network based on pattern classification; since even the best classifier will perform poorly if the inputs are not selected well. In this study, we proceeded as follows to build the vector parameter characterizing each ECG beat.

- First of all, each heart was isolated using a rectangular window, which was formed by 360 discrete data centered into the peak R, so that we are sure that all beat waves will be included and the morphology of the beat is preserved (Fig. 2)

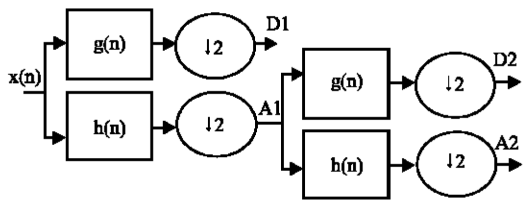

Fig. 1: Scheme for DWT implementation, in which g(x) is a high-pass filter, h (n) is a low-pass filter and (↓2) is a à trous algorithms

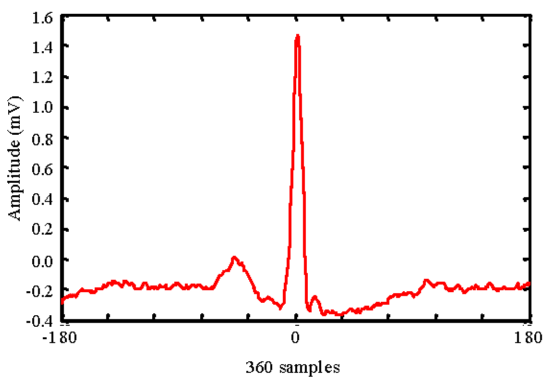

Fig. 2: ECG beat taken from patient 234 with 360 samples centred on R

- The spectral analysis of the ECG beat was then, performed using the DWT. The selection of appropriate wavelet and the number of decomposition levels is very important in signals analysis using the WT. The number of decomposition levels is chosen based on the dominant frequency components of the signal. In the present study, the number of decomposition levels was chosen to be 5. In the other hand, the smoothing feature of the Daubechies wavelet of order 2 (db2) made it more suitable to detect changes of the ECG signals. Therefore, the wavelet coefficients were computed using the db2 in the present study.

The bandwidths of the sub-band signals after five levels of DWT application are listed in Table 1.

Figure 3 shows the detail coefficients in five sub-bands (D1, D2, D3, D4, D5) and the five scale approximation A5.

- Since, ECG signal lies between 0.5 and 40 Hz. The energy of wavelet coefficients is concentrated mostly in the lower sub-bands. So we use only (D3, D4, D5

Table 1: Frequency ranges of the five-level discrete wavelet transform for digital signals sampled at 360 Hz

Band	Frequency ranges (Hz)
Signal	0-360
D1	180-360
D2	90-180
D3	45-90
D4	22.25-45
D5	11.25-22.50
A5	0-11.25

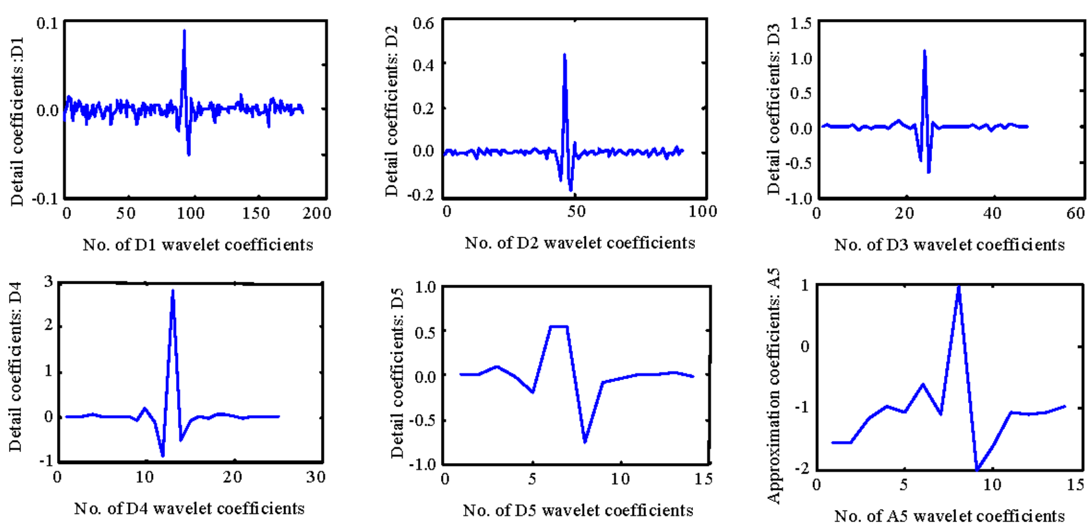

Fig. 3: Detail coefficients (D1, D2, D3, D4, D5) and the five scale approximation A5

and A5) to represent the beat. We extracted from these sub-bands and the original signal, three statistical features to characterizing each hearth beat (Yu and Chen, 2006)

Signal variance: The signal variance in a subband represents the average AC power in that band. With a discrete-time signal x of N samples, the signal variance is defined as:

$$\sigma_x{}^2 = \frac{1}{N}\sum_{n=1}^{N}\left(x(n)-\bar{x}\right)^2 \tag{2}$$

where, $\bar{x}$ is the sample mean of the signal.

Variance of the autocorrelation function of a signal: The autocorrelation function is considered to be a measure of similarity between a signal x(n) and its shifted version. Mathematically:

$$R_{xx}(l) = \sum_{n=1}^{N-|k-l|}\left[x(n)-\bar{x}\right]^2 \tag{3}$$

where, l is the time shift index, i = 1, k = 0 for l≥0 and i = 0, k = 1 for l≤0. The AC power of the autocorrelation function is calculated by determining the variance of the autocorrelation function.

Relative amplitude: The relative amplitude of the decomposed signal x(n) for each subband is defined as:

$$r_x = \frac{\mathrm{Min}(x(n))}{\mathrm{Max}(x(n))} \tag{4}$$

R intervals: Because the RR interval is crucial in characterizing arrhythmic ECG beat types, we recruited two RR intervals (RR1 and RR2) relative to the original signal and exploit them as two other features component. The RR1 interval is defined as the time duration between the current R peaks and the previous one. The RR2 is defined as the time duration between the current R peaks and the after one.

Normalization of feature vectors: As the quantities of the features can be quite, a normalization process is necessary to standardize all the features to same level. The relation used for normalization is defined as follows:

$$\tilde{x}_{ij} = \tan\mathrm{sig}\frac{(x_{ij}-\bar{x}_j)}{\sigma_{x_j}} \tag{5}$$

where, x_{ij} is the jth component of the ith feature vector, with $\bar{x}_j$ and σ_{x_j} being the mean and variance of the j.th component of the i.th feature vector computed over the training set and used throughout the computation. The applications of the tangent sigmoid function normalize the range of features to [-1, +1] (Yu and Chen, 2006).

CLASSIFICATION

Bayesian classification: The Bayesian statistic decision (TDSB) (Cid-Sueiro *et al*., 1999; De Chazel and Reilly, 2003) allows optimal decision-making by minimizing a test error. Assume we have the training set D, consisting of N input output pairs:

$$D = \left[(X_n, Y_n), n = 1,2,.....N\right] \tag{6}$$

where, X is an input vector consisting of L elements, every input vector corresponds to a heartbeat and (k = k1, k2, k3, k4, k5, k6) is the corresponding class label consisting of K classes.

The goal is to use an ANN to model the input-output relation which represents the membership of class. In our case we have six classes: where the first five classes correspond to the five types of arrhythmia. Any one of them takes a value (0) for the normal and (1) for the arrhythmic heart beat. The 6 class corresponds to the normal case and takes a value (1) for the normal beat and (0) in the case of arrhythmic. To realize a logistic regression model based on a Bayesian method, we estimated the class probability for the given input by:

$$P(Y = K/\mathrm{if}X) \tag{7}$$

The probabilistic classifier's configurations are shown schematically in Fig. 4 and 5.

RBFNN Bayesian classification: It is proven that RBFNN is an universal approximator (Cid-Sueiro *et al*., 1999) , it can approximate any non linear function if it is well dimensioned. In this study we use the RBFNN to estimate the posterior probability of the five classes of considered arrhythmias.

The goal of supervised RBF networks is to approximate the desired behaviour by a collection of functions, called kernel. The RBF network is constituted of three layers only:

- The input layer which relays the inputs without distortion and the neurons number is equal to the vector size (17 parameters in our case)
- The hidden layer is composed of Gaussian functions with center C_i and standard deviation σ_i. The output of the hidden neuron is given by:

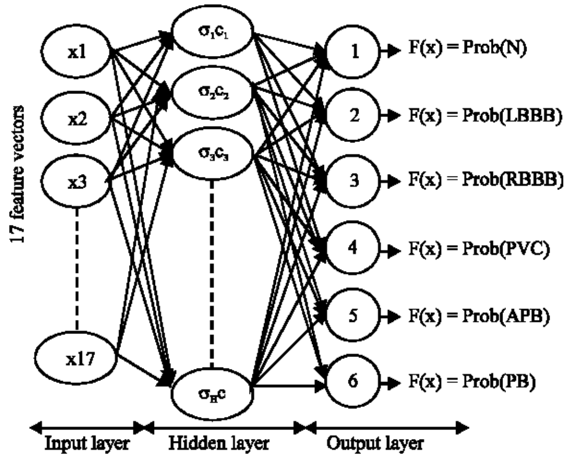

Fig. 4: Architecture of Bayesian RBFNN

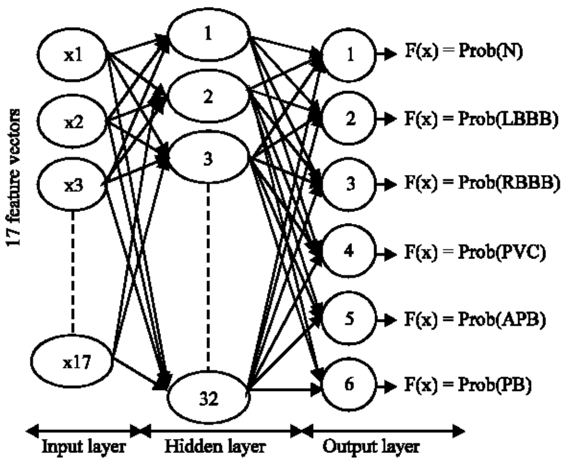

Fig. 5: BPNN network architecture

$$\Phi_i(X) = \exp\left|-\frac{1}{2}\sum_{j=1}^{n}\frac{(X_j - C_j)^2}{\sigma_{ij}^2}\right| \tag{8}$$

The response of Gaussian function is locale; its maximum value was in center Ci and decreases in a manner to fairly monotone to the distance.

- The output layer, the number of neurons is equal to the number of classes (in our case there is 6 classes). The output value is calculated by:

$$F_k(X) = w_{k0} + \sum_{i=1}^{i=H} w_{ki} \cdot \Phi(X) \tag{9}$$

$F_k(X)$ represents the posterior probability $\Phi(Y = k/X)$ of the class k knowing that X is the input vector; w_{Ki} is the synaptic weights between the hidden and output layer, with $1 \leq i \leq H$ and H is the number of neurons in hidden layer (80 in our case).

The RBF learning is done by:

- Initialization of the centers of the hidden layer using SOM (Self Organizing Map) networks (Lagerholm *et al*., 2000; Risk *et al*., 1997)
- Calculation of the synaptic weights between the hidden and output layer, using gradient descent algorithm

BPNN Bayesian classification: The MLP (Ceylan and Ozbay, 2007; Prasad and Sahamb, 2003) is one of the neural networks the most used and the most efficient in the classification problems.

- The first layer is the input layer who receives as the input data, 17 parameters characterizing the ECG beat
- The second layers, also called the hidden layers, have respectively 32 neurons. To determine the optimal number of hidden neurons, we realized several trainings for different architectures. A validation data base enabled us to choose among all these architectures that which gives in general the best rates
- The output layer has six neurons, which is equal to the number of ECG beat types to be classified and are regarded as membership probability of classes.

Dempster Shafer fusion: The (MLP) neural network and (RBF) neural network classifiers are very successful techniques in pattern recognition and using in combination, these two classifiers provide an excellent stage in a data-fusion framework.

In the other hand, Dempster-Shafer (D-S) theory (Shafer, 1976) is a mathematical theory of evidence and plausibility reasoning. It provides means of representing and combining measures of evidence. Major advantages of this theory are the ability to discriminate between ignorance and uncertainty, the ability to easily represent evidence at different levels of abstraction and the possibility to combine evidence from different sources. The theory may be summarized as follows:

Let Ω be a finite set of n mutually exclusive atomic hypotheses $\Omega = \{\theta_1, \ldots, \theta_n\}$ also referred to as a frame of discernment (FOD) representing the universe of discourse and let 2^Ω the set of all the subsets of Ω including itself and a null set $\emptyset$. Each subset is called a focal element.

A basic probability assignment or mass function m over a frame of discernment Ω is a function m: $2^\Omega \rightarrow [0, 1]$ that satisfies the following two conditions:

$$m(\phi) = 0 \text{ and } \sum_{A \subset \Omega} m(A) = 1 \tag{10}$$

The mass m(A) specifies the degree of belief that is assigned to the set $A \subset \Omega$. With m being a basic probability assignment the beliefs function: $2\Omega \rightarrow [0, 1]$ is defined as follows:

$$\text{Bel}(A) = \sum_{B \subset A} m(B) \tag{11}$$

If m is a basic probability assignment the plausibility functions: $2\Omega \rightarrow [0, 1]$ is defined as:

$$\text{Pl}(A) = \sum_{A \cap B = \phi} m(B) \tag{12}$$

Two basic probability assignments m_1 and m_2 from two independent sources can be combined via Dempster's combination rule, which is called orthogonal sum and defined as:

$$m_{1,2}(C) = m_1 \oplus m_2 = \frac{\sum_{A \cap B} m_1(A) \times m_2(B)}{K} \qquad \forall C \neq 0 \tag{13}$$

where, K is a measure of the conflict between the two sources. The conflict K is defined as:

$$K = 1 - \sum_{A \cap B = \phi} m_1(A) \times m_2(B)$$
$$K = \sum_{A \cap B \neq \phi} m_1(A) \times m_2(B) \tag{14}$$

The orthogonal sum exists only if $K \neq 0$ and the result $m_{1,2}$ is then a basic probability assignment. Otherwise the two sources are said to be totally contradictory.

Please note that, in the context of this study, the primary classifiers MLP, RBF form a set of evidence to compute the belief functions.

EXPERIMENTAL RESULTS

Preparing the data base: The ECG records used in this study, were obtained from the MIT-BIH Arrhythmia database. This database has been used in a number of studies, it contains records obtained from 48 subjects and sampled at 360 Hz. They have approximately 30 min in length. For each record, there is a corresponding annotation file, created by qualified cardiologists, that identifies the category of each beat.

We have used 23 ECG records of this database and we have considered six ECG types, including the normal beat (N), Left Bundle Branch Bloc (LBBB), Right Bundle Branch Block (RBBB), Premature Ventricular Contraction (PVC), Atrial Premature Contraction (APC) and the Paced Beat (PB).

Table 2: Origins and numbers of ECG samples used in the study

Type	Record	Training	Testing
N	103,113,115,123,220,234	600*6	600*6
LBBB	109,111,207,214	600*4	600*4
RBBB	118,124,212,231	600*4	600*4
PVC	119	200	200
	221	150	150
	200,233	400*2	400*2
APB	209	150	150
	222	100	100
	232	600	600
PB	107,217	600*2	600*2
Total		11600	11600

We have created from this database to data sets, one for training the two classifiers and the second for the testing phase. The origins of ECG beats are shown in Table 2.

RESULTS AND DISCUSSION

The test performance of our system can be determined by the computation of the total classification accuracy, the specificity and the sensitivity; which are defined as:

$$\text{Accuracy} = \frac{\text{No. of correct decisions}}{\text{No. of total beats specificity}}$$
$$= \frac{\text{No. of correct decision normal beats}}{\text{No. of total normal beatssensitivity}}$$
$$= \frac{\text{No. of correct decision abnormal beats}}{\text{No. of total abnormal beats}}$$

Two other performance indices for the classification system are false positive and false negative. False negative is defined as misclassification of the desired data to other categories, which is also called target missing. False positive is defined as the misclassification of the classified data that does not belong to this class, which is also called false alarm. A good classification system should have both lowered false negative and false positive.

Classification results using BPNN, RBFNN and fusion are shown in Table 3-5, respectively. We considered that a beat is correctly classified by both networks and after fusion, only if the posteriori probability of a class is higher than 0.75. The diagonal elements in each table are the number of correctly classified beats of specific ECG type using the proposed method.

The results show well recognition power throughout all categories; in general the global rate is more than 98%. This rate was slightly decreased after fusion, this is logical, because for few beats, the two networks are not in agreement.

Table 3: Confusion matrix and overall classification accuracy using BPNN classifier: 98.99%

	N	LBBB	RBBB	PVC	APB	PB	Reject
N	3593	0	0	1	6	0	0
LBBB	3	2367	6	20	4	0	0
RBBB	2	18	2367	8	5	0	0
PVC	0	13	5	1130	2	0	0
APB	2	3	6	1	840	0	0
Learning	100	99.87	98.95	99.30	99.41	100	0
Test	99.80	98.62	98.62	98.26	98.82	100	0

Table 4: Confusion matrix and overall classification accuracy using RBFNN classifier: 99.14%

	N	LBBB	RBBB	PVC	APB	PB	Reject
N	3596	1	0	0	3	0	0
LBBB	2	2378	4	14	2	0	0
RBBB	0	13	2380	4	3	0	0
PVC	0	12	3	1135	1	0	0
APB	1	2	10	2	835	0	0
PB	0	0	0	0	0	1200	0
Learning	100	99.91	99.37	99.56	99.64	100	0
Test	99.88	99.08	99.16	98.69	98.23	100	0

Table 5: Confusion matrix and overall classification accuracy using fusion: 98.87%

	N	LBBB	RBBB	PVC	APB	PB	Reject
N	3588	0	0	2	2	0	9
LBBB	2	2375	5	7	0	0	11
RBBB	0	5	2380	2	1	0	12
PVC	0	8	2	1129	1	0	10
APB	1	2	3	1	833	0	10
Learning	99.94	99.58	99.16	98.86	98.23	100	0
Test	99.66	98.95	99.16	98.17	98	100	0

We also note that the rejection rate is higher after fusion; in fact, when the two classifiers are in disagreement they can consider beat as rejected (it is better to say I do not know that making a wrong diagnosis).

Consequently, it is observed that the rates of false negative for class LBBB, RBBB, PVC and APB obtained by the BPNN and RBFNN classifier, are higher comparatively with those obtained after fusion.

To examine more precisely the effect of fusion in ECG beat classification, the misclassification numbers beats represented as false negative and false positive are illustrated in Fig. 6a and b, respectively.

In Fig. 6a and b, the false negative and false positive rates are zeros in the PB category, because they are very different from the other beat types.

Influence of training set size: The number of beats which are used to training must be very small in comparison to the total number of beats to be classified. If this is not the case, the classifier will perform very well on the given set of data, but very poorly on new data.

Table 6 shows the classification results on test data, when using different training data size. The training data size was halved by evenly selecting the training signals from the training set in previous test. The result shows

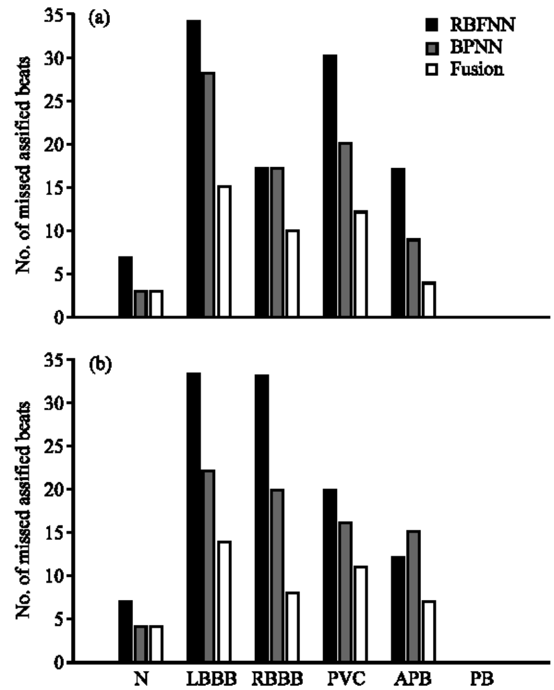

Fig. 6: (a) False positive and (b) false negative

that the performance of our system, slightly degraded with the smallest data size. Even with smallest data size, i.e., 725 training data, degradations of the specificity and the sensitivities of the LBBB and RBBB are less than 0, 3%. The most degradation due to the decrease in training data size occurs in PVC, APB and PB, who it same does not exceed 0.55% after fusion. These results prove the effectiveness and robustness of our classifier which generalize very well on new data.

Many methods for automatic detection and classification of various arrhythmias have recently been presented in literature and it's interesting to compare our method with these methods. Some representative ECG beat recognition systems are chosen for this comparison: a modified mixture of experts network structure for ECG beats classification with diverse features (MME) (Guler and Ubeyli, 2005); ECG beat classification using neuro-fuzzy network (neuro-fuzzy) (Engin, 2004). The ECG rhythm classification using artificial neural networks (MLP-LVQ) (Oien *et al*., 1996). Comparison of different wavelet subband features in the classification of ECG Beats using Probabilistic Neural Network (PNN) (Yu and Chen, 2006) and real-time discrimination of ventricular tachyarrhythmia with Fourier-transform neural network (MLP-Fourier) (Minami *et al*., 1999). Table 6 compares the accuracy of these systems. Since, different numbers of beat types were exploited in different systems, the averaged classification accuracy was calculated for comparison.

Tabel 6: Effects of different training data size

Networks	Training data size	Specificity	Sensitivity (%) LBBB	RBBB	PVC	APB	PB	Classification accuracy (%)
BPNN	11600	99.80	98.62	98.62	98.26	98.82	100	99.02
RBFNN		99.88	99.08	99.16	98.69	98.23	100	99.17
Fusion		99.66	98.95	99.16	98.17	98.00	100	98.99
BPNN	5800	99.75	98.62	98.58	98.08	98.23	100	98.87
RBFNN		99.83	99.00	99.12	98.52	97.76	100	99.04
Fusion		99.61	98.87	99.08	97.82	97.64	100	98.75
BPNN	2900	99.75	98.62	98.58	98.08	98.23	100	98.84
RBFNN		99.77	99.00	99.12	98.52	97.29	100	98.95
Fusion		99.55	98.75	99.08	97.82	97.64	100	98.72
BPNN	1450	99.72	98.50	98.50	98.08	98.11	100	98.78
RBFNN		99.77	99.00	99.12	98.52	97.29	100	98.95
Fusion		99.55	98.70	99.08	97.65	97.52	100	98.67
BPNN	725	99 .69	98.45	98.50	97.91	98.11	99.66	98.71
RBFNN		99.72	99.00	99.08	98.40	97.05	99.83	98.84
Fusion		99.55	98.66	99.00	97.65	97.52	99.50	98.54

Table 7: Comparison of our method with other systems

Method	No. of beat types	Accuracy (%)
Proposed	6	98.99
MME	5	97.78
Neuro-fuzzy	4	98.00
MLP-LVQ	2	96.80
PNN	6	99.65
MLP-Fourier	3	98.00

The result shows that our proposed method is comparatively with other systems, however, the major difference between the proposed system and others is that the obtained rates are not only high, but homogeneous throughout all ECG beats type; moreover, false alarms are reduced, since the system does not give a response in the ambiguity case (it is better to say, I don't know to give a false answer) (Table 7). In the other hand, the outputs of our system are given in term of the class probabilities and not like only as truth or false (0 or 1), so it is possible to well analyze the response of the system for taking the best decisions. Our classifiers generalizes very well even when it is entrained with a small set; indeed classifiers are effective when the test base is very large compared to the learning set.

CONCLUSIONS

In this study, we have proposed an ECG classification scheme based on wavelet transform and fusion of two Bayesian neural networks, to aid diagnosis of five common cardiac arrhythmias: Premature Ventricular Contraction (PVC) ventricular, Atrial Premature Contraction (APC), Right Bundle Branch Block (RBBB) Left Bundle Branch Bloc (LBBB) and the Paced Beat (PB), in addition to normal beat (N). The signals were first decomposed into subbands components with five levels of DWT. Three categories of statistical features corresponding to the three last decomposed signals and the last approximation, also the instantaneous RR interval and statistical parameters of the original signal are exploited as input of our RBFNN and BPNN classifier.

The two modules work in parallel, each one determines the beat class in term of posterior probability and they are regarded as two different sources each one having an opinion to give. The two module outputs are then fused by using the Dempster Shafer rule, to reduce at the maximum the misclassification.

Present system was validated on real ECG records taken from the MIT-BIH database. This study proves that the proposed method is an excellent model for the computer-aided diagnosis of heart diseases based on ECG signals. However, there are some important issues we want to explore in future study, including the effect of noises on the classifier, the comparison of the RBFNN and BPNN to the most popular neural network classifier and detailing the comparison of the proposed method to the other methods published in the literatures.

REFERENCES

Ceylan, R. and Y. Ozbay, 2007. Comparison of FCM, PCA and WT techniques for classification ECG, arrhythmias using artificial neural network. Expert Syst. Appl., 33: 286-295.

Challis, R.E. and R.I. Kitney, 1990. Biomedical signal processing's (in four parts). Part1: Time domain mathods. Med. Biol. Eng. Comput., 28: 509-524.

Chen, S.W., 2007. Complexity-measure-based sequential hypothesis testing for real-time detection of lethal cardiac arrhythmias. EURASIP J. Adv. Signal Process., 2007: 1-8.

Cid-Sueiro, J., J.I. Arribas, S. Urban-Muioz and A.R. Figueiras-Vidal, 1999. Cost functions to estimate a posteriori probabilities in multiclass problems. IEEE Trans. Neural Network, 10: 645-655.

De Chazel, P. and R.B. Reilly, 2003. Automatic classification of ECG beats using waveform shape and heart beat interval features. Proceedings of the IEEE International Conference on Acoustic, Speech and Signal Prcessing, (IcASSP'03), Hong Kong, China, pp: 269-272.

Engin, M., 2004. ECG beat classification using neuro-fuzzy network. Pattern Recognit Lett., 25: 1715-1722.

Guler, I. and E.D. Ubeyli, 2005. A modified mixture of experts network structure for ECG beats classification with diverse features. Eng. Appl. Artificial Intell., 18: 845-856.

Hu, Y.H., S. Palreddy and W. Tompkins, 1997. A patient adaptable ECG beat classifier using a mixture of experts approach. IEEE Trans. Biomed. Eng., 44: 891-900.

Jekova, L., G. Bortolan and L. Christov, 2008. Assessment and comparison of different methods for heartbeat classification. Med. Eng. Phys., 30: 248-257.

Khadra, L., A. Al-Fahoum and S. Binajjaj, 2005. A quantitative analysis approach for cardiac arrhythmia classification using higher order spectral techniques. IEEE Trans. Biomed. Eng., 52: 1840-1845.

Lagerholm, M., C. Peterson, G. Braccini, L. Edenbrandt and L. Sornmo, 2000. Clustering ECG complexes using Hermite functions and self-organizing maps. IEEE Trans. Biomed. Eng., 47: 838-848.

Minami, K., H. Nakajima and T. Toyoshima, 1999. Real-time discrimination of ventricular tachyarrhythmia with Fourier-transform neural network. IEEE Trans. Biomed. Eng., 46: 179-185.

Moraes, J., M.O. Seixas, F.N. Vilani and E.V. Costa, 2002. A real time QRS complex classification method using Mahalanobis distance. Comput. Cardiol., 29: 201-204.

Oien, G.E., N.A. Berteslsen, T. Eftestol and J.H. Husoy, 1996. ECG rhythm classification using artificial neural networks. Proceedings of the IEEE Digital Signal Processing Workshop, Sept. 1-4 Stavanger College, Ullandhaug, pp: 415-517.

Osowski, S. and T.H. Linh, 2001. ECG beat recognition using fuzzy hybrid neural network. IEEE Trans. Biomed. Eng., 48: 1265-1271.

Prasad, G.K. and J.S. Sahamb, 2003. Classification of ECG arrhythmias using multi-resolution analysis and neural networks. IEEE. Trans. Biomed. Eng., 1: 227-231.

Risk, M.R., J.F. Sobh and J.P. Saul, 1997. Beat detection and classification of ECG using self organizing maps. Proceedings of the 19th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Oct. 30-Nov. 2, Chicago, IL. USA., pp: 89-91.

Shafer, G., 1976. A Mathematical Theory of Evidence. 1st Edn., Princeton University Press, Princeton, New Jersey, USA.

Ubeyli, E.D., 2007. ECG beats classification using multiclass support vector machines with error correcting output codes. Digit Signal Prog., 17: 675-684.

Vargas, F., D. Lettnin, M.C.F. de Castro and M. Macarthy, 2002. Electrocardiogram pattern recognition by means of MLP network and PCA: A case study on equal amount of input signal types. Proceedings of the 7th Brazilian Symposium on Neural Networks, Nov. 11-14, IEEE Computer Society, Washington, DC, USA., pp: 200-205.

Yu, S.N. and Y.H. Chen, 2006. Comparison of different wavelet subband features in the classification of ECG beats using probabilistic neural network. Proc. 28th IEEE Conf. Eng. Med. Biol. Soc., 1: 1398-1401.

Yu, S.N. and Y.H. Chen, 2009. Noise-tolerant electrocardiogram beat classification based on higher order statistics of subband components. Artif. Intell. Med., 46: 165-178.