ISTANBUL UNIVERSITY –
JOURNAL OF ELECTRICAL & ELECTRONICS ENGINEERING

YEAR : 2007 (393 - 402)
VOLUME : 7
NUMBER : 2

ANALYSIS OF ECG SIGNALS BY DIVERSE AND COMPOSITE FEATURES

Elif Derya ÜBEYLİ

Department of Electrical and Electronics Engineering, Faculty of Engineering,
TOBB Ekonomi ve Teknoloji Üniversitesi,
06530 Söğütözü, Ankara, Turkey

E-mail : edubeyli@etu.edu.tr

ABSTRACT

In this study, the automated diagnostic systems employing diverse and composite features for electrocardiogram (ECG) signals were analyzed and their accuracies were determined. In pattern recognition applications, diverse features are extracted from raw data which needs recognizing. Combining multiple classifiers with diverse features are viewed as a general problem in various application areas of pattern recognition. Because of the importance of making the right decision, classification procedures classifying the ECG signals with high accuracy were analyzed. The classification accuracies of multilayer perceptron neural network, combined neural network, and mixture of experts trained on composite features and modified mixture of experts trained on diverse features were compared. The inputs of these automated diagnostic systems composed of diverse or composite features and were chosen according to the network structures. The conclusions of this study demonstrated that the modified mixture of experts trained on diverse features achieved accuracy rates which were higher than that of the other automated diagnostic systems trained on composite features.

Keywords: : *Diverse features, Composite features, Electrocardiogram (ECG) signals, Automated diagnostic systems*

1. INTRODUCTION

Electrocardiography is an important tool in diagnosing the condition of the heart. The electrocardiogram (ECG) is the record of variation of bioelectric potential with respect to time as the human heart beats. It provides valuable information about the functional aspects of the heart and cardiovascular system. Early detection of heart diseases/abnormalities can prolong life and enhance the quality of living through appropriate treatment. Therefore, numerous research and work analyzing the ECG signals have been reported [1-5]. 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. However, biosignals being nonstationary signals, this reflection may occur at random in the time scale. In this situation, the disease symptoms may not show up 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

Received Date : 21.01.2007
Accepted Date: 05.12.2007

over several hours. Thus, the volume of the data being enormous, the study is tedious and time consuming. Naturally, the possibility of the analyst missing (or misreading) vital information is high. Therefore, computer-based analysis and classification of diseases can be very helpful in diagnostics [1-5].

Various methodologies of automated diagnosis have been adopted, however the entire process can generally be subdivided into a number of disjoint processing modules: beat detection, feature extraction/selection, and classification. The initial pre-processing module of beat detection aims to locate each cardiac cycle in each of the recording leads and insert reference markers indicating the beginning and end of each interwave component. The algorithm is designed with two main objectives: firstly, the detector should provide reliable detection of each cardiac cycle in all recording leads and secondly, the temporal location of the reference points should be described accurately. The accuracy of detection of each cardiac cycle is of great importance since it contributes significantly to the overall classification result. The markers are subsequently processed by the feature extraction module, where measurements are produced for wave amplitudes and durations. The collective term for the measurements produced is commonly referred to as the input feature vector, which is considered to describe the morphology of the current recorded signal. The module of feature selection is an optional stage, whereby the feature vector is reduced in size including only, from the classification viewpoint, what may be considered as the most relevant features required for discrimination. The classification module is the final stage in automated diagnosis. It examines the input feature vector and based on its algorithmic nature, produces a suggestive hypothesis [6,7].

The wavelet transform (WT) can be applied to extract the wavelet coefficients of discrete time signals. This procedure makes use of multirate signal processing techniques. The proposed scheme is the subband coding or multiresolution signal analysis. The multiresolution feature of the WT allows the decomposition of a signal into a number of scales, each scale representing a particular coarseness of the signal under study. The WT provides very general techniques which can be applied to many tasks in signal processing. One very important application is the ability to compute and manipulate data in compressed parameters which are often called features [8]. Thus, the ECG signal, consisting of many data points, can be compressed into a few parameters. These parameters characterize the behavior of the ECG signal. This feature of using a smaller number of parameters to represent the ECG signal is particularly important for recognition and diagnostic purposes [1,4,5].

Eigenvector methods are used for estimating frequencies and powers of signals from noise-corrupted measurements. These methods are based on an eigen-decomposition of the correlation matrix of the noise–corrupted signal. Even when the signal-to-noise ratio (SNR) is low, the eigenvector methods produce frequency spectra of high resolution. These methods are best suited to signals that can be assumed to be composed of several specific sinusoids buried in noise. Hence, to gain some noise immunity it is reasonable to retain only the principal eigenvector components in the estimation of the autocorrelation matrix. Using the frequency estimations provided by any one of these methods, the power levels of the signal can be determined from the power matrix [9,10]. In this study, three eigenvector methods (Pisarenko, multiple signal classification – MUSIC, and Minimum-Norm) were selected to generate the power spectral density (PSD) estimates.

As in traditional pattern recognition systems (Figure 1), the present model consists of three main modules: a feature extractor that generates a feature vector from the ECG signals, feature selection that composes diverse and composite features (wavelet coefficients and power levels of the PSDs obtained by the eigenvector methods), and feature classifiers that output the class based on the diverse and composite features (multilayer perceptron neural network – MLPNN, combined neural network – CNN, mixture of experts – ME and modified mixture of experts – MME). A significant contribution of the present work was the composition of diverse and composite features by the usage of discrete wavelet transform (DWT) and eigenvector methods which were used to train novel classifier (MME trained on diverse features) for the ECG signals. The ECG signals (normal beat, congestive heart failure beat, ventricular tachyarrhythmia beat, atrial fibrillation beat) from the Physiobank database [11] were used to train and test the classifiers. The present study was conducted with the purpose of answering the question of whether the automated diagnostic systems with diverse features (MME) or composite features (MLPNN, CNN and ME) improve the capability of classification of the ECG signals. To evaluate performance of the classifiers, the classification accuracies, the central processing unit (CPU) times of training of the classifiers (diverse or composite feature vectors used as inputs) were compared.

2. SPECTRAL ANALYSIS USING DISCRETE WAVELET TRANSFORM

The ECG signals are considered as representative signals of cardiac physiology, useful in diagnosing cardiac disorders. The most complete way to display this information is to perform spectral analysis. The ECG signal, consisting of many data points, can be compressed into a few parameters by the WT. These parameters characterize the behavior of the ECG signal and they can be used for recognition and diagnostic purposes. The WT can be thought of 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 decomposition of a signal into a number of scales, each scale representing a particular coarseness of the signal under study. The procedure of multiresolution decomposition of a signal $x[n]$ is schematically shown in Figure 2. Each stage of this scheme consists of two digital filters and two downsamplers by 2. The first filter, $g[\cdot]$ is the discrete mother wavelet, high-pass in nature, and the second, $h[\cdot]$ is its mirror version, low-pass in nature. The downsampled outputs of first high-pass and low-pass filters provide the detail, D_1 and the approximation, A_1, respectively. The first approximation, A_1 is further decomposed and this process is continued as shown in Figure 2.

All wavelet transforms can be specified in terms of a low-pass filter h, which satisfies the standard quadrature mirror filter condition:

$$H(z)H(z^{-1}) + H(-z)H(-z^{-1}) = 1, \quad (1)$$

where $H(z)$ denotes the z-transform of the filter h. Its complementary high-pass filter can be defined as

$$G(z) = zH(-z^{-1}). \quad (2)$$

A sequence of filters with increasing length (indexed by i) can be obtained:

$$H_{i+1}(z) = H(z^{2^i})H_i(z)$$
$$G_{i+1}(z) = G(z^{2^i})H_i(z), \quad i = 0, \ldots, I-1 \quad (3)$$

with the initial condition $H_0(z) = 1$. It is expressed as a two-scale relation in time domain

$$h_{i+1}(k) = [h]_{\uparrow 2^i} * h_i(k)$$
$$g_{i+1}(k) = [g]_{\uparrow 2^i} * h_i(k), \quad (4)$$

where the subscript $[\cdot]_{\uparrow m}$ indicates the up-sampling by a factor of m and k is the equally sampled discrete time.

The normalized wavelet and scale basis functions $\varphi_{i,l}(k)$, $\psi_{i,l}(k)$ can be defined as

$$\varphi_{i,l}(k) = 2^{i/2} h_i(k - 2^i l)$$
$$\psi_{i,l}(k) = 2^{i/2} g_i(k - 2^i l), \quad (5)$$

where the factor $2^{i/2}$ is an inner product normalization, i and l are the scale parameter and the translation parameter, respectively. The discrete wavelet transform (DWT) decomposition can be described as

$$a_{(i)}(l) = x(k) * \varphi_{i,l}(k)$$
$$d_{(i)}(l) = x(k) * \psi_{i,l}(k), \quad (6)$$

where $a_{(i)}(l)$ and $d_i(l)$ are the approximation coefficients and the detail coefficients at resolution i, respectively [8].

3. SPECTRAL ANALYSIS USING EIGENVECTOR METHODS

The Pisarenko method is particularly useful for estimating PSD which contains sharp peaks at the expected frequencies. The polynomial $A(f)$ which contains zeros on the unit circle can then be used to estimate the PSD.

$$A(f) = \sum_{k=0}^{m} a_k e^{-j2\pi fk} \quad (7)$$

where $A(f)$ represents the desired polynomial, a_k represents coefficients of the desired polynomial, and m represents the order of the eigenfilter, $A(f)$. From the eigenvector corresponding to the minimum eigenvalue, the

Pisarenko method determines the signal PSD from the desired polynomial:

$$P_{PISARENKO}(f) = \frac{1}{|A(f)|^2}\,. \qquad (8)$$

The MUSIC method is also a noise subspace frequency estimator and eliminates the effects of spurious zeros by using the averaged spectra of all of the eigenvectors corresponding to the noise subspace. The resultant PSD is determined from

$$P_{MUSIC}(f) = \frac{1}{\frac{1}{K}\sum_{i=0}^{K-1}|A_i(f)|^2} \qquad (9)$$

where K represents the dimension of noise subspace, $A_i(f)$ represents the desired polynomial that corresponds to all the eigenvectors of the noise subspace.

In addition to the Pisarenko and MUSIC methods, the Minimum-Norm method was investigated. In order to differentiate spurious zeros from real zeros, the Minimum-Norm method forces spurious zeros inside the unit circle and calculates a desired noise subspace vector a from either the noise or signal subspace eigenvectors. Thus, while the Pisarenko method uses only the noise subspace eigenvector corresponding to the minimum eigenvalue, the Minimum-Norm method uses a linear combination of all noise subspace eigenvectors. The Minimum-Norm PSD can be estimated as follows:

$$P_{MIN}(f, K) = \frac{1}{|A(f)|^2} \qquad (10)$$

where K represents the dimension of the noise subspace [9,10].

4. BRIEF REVIEW OF IMPLEMENTED CLASSIFIERS

Several problems occurring with the usage of a composite feature are given in the following:

- Its dimension is higher than that of any component feature and it is well known that high-dimension vectors will not only increase computational complexity but will also produce implementation problems and accuracy problems.
- It is difficult to combine several features due to their diversified forms, e.g., they may be continuous variables, binary values, discrete labels, structural primitives.
- The component features are usually not independent.

In general, therefore, the use of a composite feature does not provide a significantly improved performance. However, the combination of multiple classifiers is a good solution for the problem involving a variety of features [12-14].

There have recently been widespread interests in the use of multiple models for pattern classification and regression in statistics and neural network communities. The basic idea underlying these methods is the application of a so-called divide-and-conquer principle that is often used to tackle a complex problem by dividing it into simpler problems whose solutions can be combined to yield a final solution. Utilizing this principle, Jacobs et al. [15] proposed a modular neural network architecture called ME. The ME models the conditional probability density of the target output by mixing the outputs from a set of local experts, each of which separately derives a conditional probability density of the target output. The outputs of expert networks are combined by a gating network simultaneously trained in order to stochastically select the expert that is performing the best at solving the problem [16,17]. Based on the probabilistic model, learning in the ME architecture is treated as a maximum likelihood problem. Jordan and Jacobs [18] have proposed an expectation-maximization (EM) algorithm for adjusting the parameters of the architecture. In this framework a number of relatively small expert networks can be used together with a gating network designed to divide the global classification task into simpler subtasks [12,13,19]. Although the ME architecture has been successfully applied to several supervised learning tasks, it can only use a composite feature for classification with diverse features, since both gating and expert networks need to receive the same input. A MME network structure was proposed by Chen [12] for the effective use of diverse features representing the signals under study. Another network structure used in discrimination of the signals is the CNN, which combines the predictions of several models trained on composite features. The general framework for prediction using an ensemble of models consists of two levels and is often referred to as stacked generalization [5,20].

5. COMPUTATION OF DIVERSE FEATURES

Spectral analysis of the ECG signals was performed using the DWT as described in Section 2. The ECG signals can be considered as a superposition of different structures occuring on different time scales at different times. One purpose of wavelet analysis is to separate and sort these underlying structures of different time scales. Selection of appropriate wavelet and the number of decomposition levels is very important in analysis of signals using the WT. The number of decomposition levels is chosen based on the dominant frequency components of the signal. The levels are chosen such that those parts of the signal that correlate well with the frequencies required for classification of the signal are retained in the wavelet coefficients. In the present study, the number of decomposition levels was chosen to be 4. Thus, the ECG signals were decomposed into the details $D_1 - D_4$ and one final approximation, A_4. Usually, tests are performed with different types of wavelets and the one which gives maximum efficiency is selected for the particular application. 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 computed discrete wavelet coefficients provide a compact representation that shows the energy distribution of the signal in time and frequency. Therefore, the computed detail and approximation wavelet coefficients of the ECG signals were used as the feature vectors representing the signals. A rectangular window, which was formed by 256 discrete data, was selected so that it contained a single ECG beat. For each ECG beat, the detail wavelet coefficients ($d^k, \ k = 1,2,3,4$) at the first, second, third and fourth levels ($129 + 66 + 34 + 18$ coefficients) and the approximation wavelet coefficients (a^4) at the fourth level (18 coefficients) were computed. Then 265 wavelet coefficients were obtained for each ECG beat.

The Pisarenko, MUSIC, and Minimum-Norm methods were employed to obtain PSDs of the ECG signals. Using the frequency estimations provided by any one of these methods, the power levels of the signal can be determined from the power matrix. In the Pisarenko method, the eigenvector associated with the minimum eigenvalue of the estimated autocorrelation matrix is used to calculate the PSD. This method may produce spurious zeros and has a relatively poor statistical accuracy. In all cases, the Pisarenko PSD showed extra peaks as compared to the PSDs obtained from the MUSIC or Minimum-Norm methods. The MUSIC method eliminates these spurious zeros by averaging the spectra from all of the eigenvectors corresponding to noise subspace. The MUSIC method is the most widely studied, computationally simple, high-resolution eigenvector method. The MUSIC method can be considered as an appropriate method for spectral analysis of the ECG signals. The Minimum-Norm method treats the problem of spurious zeros by forcing them inside the unit circle. For each beat the 129 points of the logarithm of the power levels of the PSDs were computed.

Feature selection plays an important role in classifying systems such as neural networks. The feature selection process performed on a set of predetermined features. Features are selected based on either 1) best representation of a given class of signals, or 2) best distinction between classes. High-dimension of feature vectors increased computational complexity and the neural networks trained on these feature vectors produced lower accuracy. In order to reduce the dimensionality of the extracted diverse feature vectors, statistics over the set of the wavelet coefficients and power levels of the PSDs were used. The following statistical features were used in reducing the dimensionality of the extracted diverse feature vectors representing the ECG signals:

Maximum of the power levels of the PSDs obtained by the eigenvector methods, maximum of the wavelet coefficients in each subband.
Minimum of the power levels of the PSDs obtained by the eigenvector methods, minimum of the wavelet coefficients in each subband.
Mean of the power levels of the PSDs obtained by the eigenvector methods, mean of the wavelet coefficients in each subband.
Standard deviation of the power levels of the PSDs obtained by the eigenvector methods, standard deviation of the wavelet coefficients in each subband.

Tables 1 and 2 present the extracted features of four exemplary records from four classes. From Tables 1 and 2, one can see that the extracted diverse features of the four classes of ECG beats are different from each other. This result indicated that they can serve as useful parameters in classifying the ECG signals. The diverse feature vectors were computed by the usage of the MATLAB software package.

6. APPLICATION OF AUTOMATED DIAGNOSTICS SYSTEMS TO ECG SIGNALS

The adequate functioning of neural networks depends on the sizes of the training set and test set. Training and test sets were formed by 720 vectors (180 vectors from each class) of 32 dimensions (dimension of the extracted feature vectors). The 360 vectors (90 vectors from each class) of 32 dimensions were used for training and the 360 vectors (90 vectors from each class) of 32 dimensions were used for testing. The term vector is used for defining the extracted features of the samples of an ECG beat.

MME classifier with 12 expert networks was adopted as the structure of the MME classifier with four diverse feature vectors and the ensemble of expert networks in this structure was divided into four groups (three expert networks in each group). For the purpose of different classifiers comparison, ME classifier configured with 4 expert networks was implemented to deal with the same classification problem and they were based on a composite feature set (32 inputs). Both the gating and expert networks in the MME and ME classifiers were MLPNNs with a single hidden layer. Four sets of neural networks were trained for the first level models in the CNN since there were four possible outcomes of the diagnosis of the ECG beats. Networks in each set were trained so that they are likely to be more accurate for one type of beat than the other beat. The network architecture was the MLPNN with a single hidden layer. Each network had 32 input neurons, equal to the dimension of composite feature vector. The number of hidden neurons was 30 and the number of output was 4. Samples with target outputs were given the binary target values of (0,0,0,1), (0,0,1,0), (0,1,0,0), (1,0,0,0). Second level neural network was trained to combine the predictions of the first level networks. The second level network had 16 inputs which correspond to the outputs of the four groups of the first level networks. The targets for the second level network were the same as the targets of the original data. The number of outputs was four and the number of hidden neurons was chosen to be 30. In order to compare performance of the different classifiers, for the same classification problems the MLPNN, which is the most commonly used feedforward neural networks, was also implemented. The single hidden layered (25 hidden neurons) MLPNN was used to classify the ECG beats based on a composite feature vector (32 inputs). Different experiments were performed during implementation of these classifiers and the number of hidden neurons was determined by taking into consideration the classification accuracies. In the hidden layers and the output layers, the activation function was the sigmoidal function. Table 3 defines the network parameters of the classifiers implemented in this research.

In order to compare the classifiers used for classification of the ECG beats, the total classification accuracies on the test sets and the CPU times of training (for Pentium 4, 3.00 GHz) of the four classifiers are presented in Table 4. From the classification results presented in Table 4, one can see that the MME classifier trained on the four diverse feature vectors produce considerably better performance than that of the ME, CNN and MLPNN classifiers trained on the composite feature vector.

7. CONCLUSION

The MMEs used for classification of the ECG signals were trained, cross validated and tested with the extracted diverse feature vectors. For comparison different classifiers, the ME, CNN and MLPNN classifiers were implemented to deal with the same classification and the ME, CNN and MLPNN classifiers were used to handle ECG signals classification based on a composite features. The classification accuracies and the CPU times of training showed that the MME classifiers trained on the four diverse feature vectors produce considerably better performance than that of the ME, CNN and MLPNN classifiers trained on the composite features. The results of the present study demonstrated that the MME can be used in classification of the ECG signals by taking into consideration the misclassification rates.

REFERENCES

[1] Saxena, S.C., Kumar, V., Hamde, S.T., "Feature extraction from ECG signals using wavelet transforms for disease diagnostics", *International Journal of Systems Science*, Vol: 33, No: 13, pp. 1073-1085, 2002.

[2] Foo, S.Y., Stuart, G., Harvey, B., Meyer-Baese, A., "Neural network-based EKG pattern recognition", *Engineering Applications of Artificial Intelligence*, Vol: 15, pp. 253-260, 2002.

[3] Maglaveras, N., Stamkopoulos, T., Diamantaras, K., Pappas, C., Strintzis, M., "ECG pattern recognition and classification using non-linear transformations and neural networks: A review", *International Journal of Medical Informatics*, Vol: 52, pp. 191-208, 1998.

[4] Sternickel, K., "Automatic pattern recognition in ECG time series", *Computer Methods and Programs in Biomedicine*, Vol: 68, pp. 109-115, 2002.

[5] Güler, İ., Übeyli, E.D., "ECG beat classifier designed by combined neural network model", Pattern Recognition, Vol: 38, No: 2, pp. 199-208, 2005.

[6] Kwak, N., Choi, C-H., "Input feature selection for classification problems", *IEEE Transactions on Neural Networks*, Vol: 3, No: 1, pp. 143-159, 2002.

[7] Übeyli, E.D., Güler, İ., "Feature extraction from Doppler ultrasound signals for automated diagnostic systems", *Computers in Biology and Medicine*, Vol: 35, No: 9, pp. 735-764, 2005.

[8] Daubechies, I., "The wavelet transform, time-frequency localization and signal analysis", *IEEE Transactions on Information Theory*, Vol: 36, No: 5, pp. 961-1005, 1990.

[9] Akay, M., Semmlow, J.L., Welkowitz, W., Bauer, M.D., Kostis, J.B., "Noninvasive detection of coronary stenoses before and after angioplasty using eigenvector methods", *IEEE Transactions on Biomedical Engineering*, Vol: 37, No: 11, pp. 1095-1104, 1990.

[10] Übeyli, E.D., Güler, İ., "Comparison of eigenvector methods with classical and model-based methods in analysis of internal carotid arterial Doppler signals", *Computers in Biology and Medicine*, Vol: 33, No: 6, pp. 473-493, 2003.

[11] Goldberger, A.L., Amaral, L.A.N., Glass, L., Hausdorff, J.M., Ivanov, P.Ch., Mark, R.G., Mietus, J.E., Moody, G.B., Peng, C.K., Stanley, H.E. Physiobank, Physiotoolkit, and Physionet: Components of a New Research Resource for Complex Physiologic Signals, Circulation 101(23), e215-e220 [Circulation Electronic Pages; http://circ.ahajournals.org/cgi/content/full/101/23/e215]; 2000 (June 13).

[12] Chen, K., "A connectionist method for pattern classification with diverse features", *Pattern Recognition Letters*, Vol: 19, No: 7, pp. 545-558, 1998.

[13] Xu, L., Krzyzak, A., Suen, C.Y., "Methods of combining multiple classifiers and their applications to handwriting recognition", *IEEE Transactions on Systems, Man, and Cybernetics*, Vol: 22, No: 3, pp. 418-435, 1992.

[14] Chen, K., Wang, L., Chi, H., "Methods of combining multiple classifiers with different features and their applications to text-independent speaker identification", *International Journal of Pattern Recognition and Artificial Intelligence*, Vol: 11, No: 3, pp. 417-445, 1997.

[15] Jacobs, R.A., Jordan, M.I., Nowlan, S.J., Hinton, G.E., "Adaptive mixtures of local experts", *Neural Computation*, Vol: 3, No: 1, pp. 79-87, 1991.

[16] Chen, K., Xu, L., Chi, H., "Improved learning algorithms for mixture of experts in multiclass classification", *Neural Networks*, Vol: 12, No: 9, pp. 1229-1252, 1999.

[17] Hong, X., Harris, C.J., "A mixture of experts network structure construction algorithm for modelling and control", *Applied Intelligence*, Vol: 16, No: 1, pp. 59-69, 2002.

[18] Jordan, M.I., Jacobs, R.A., "Hierarchical mixture of experts and the EM algorithm", *Neural Computation*, Vol: 6, No: 2, pp. 181-214, 1994.

[19] Güler, İ., Übeyli, E.D., "A mixture of experts network structure for modelling Doppler ultrasound blood flow signals", *Computers in Biology and Medicine*, Vol: 35, No: 7, pp. 565-582, 2005.

[20] Wolpert, D.H., "Stacked generalization", *Neural Networks*, Vol: 5, pp. 241-259, 1992.

Elif Derya ÜBEYLİ graduated from Çukurova University in 1996. She took her M.S. degree in 1998, all in electronic engineering. She took her Ph.D. degree from Gazi University, electronics and computer technology. She is an Associate Professor at TOBB Economics and Technology University, Department of Electrical and Electronics Engineering. Her interest areas are biomedical signal processing, neural networks, and artificial intelligence. She has written more than 75 articles on biomedical engineering.

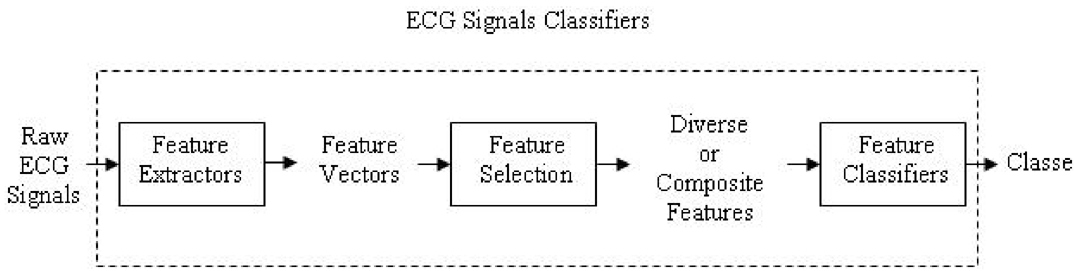

Figure 1. *General structure of the developed ECG signals classifiers*

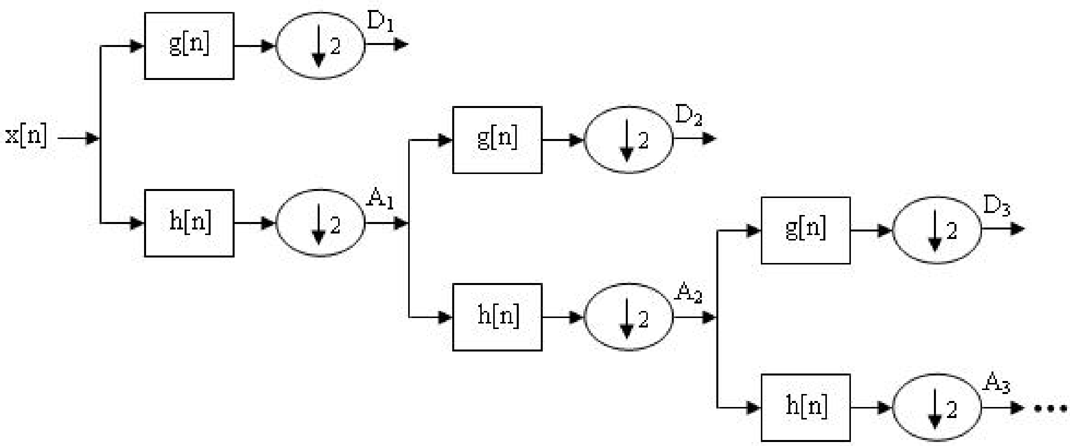

Figure 2. *Subband decomposition of discrete wavelet transform implementation;* $g[n]$ *is the high-pass filter,* $h[n]$ *is the low-pass filter*

Table 1. *The wavelet coefficients of four exemplary records from four classes*

ECG beat types	Extracted Features	Wavelet Coefficients				
		Subbands				
		D_1	D_2	D_3	D_4	A_4
Normal beat	Maximum	0.2062	1.5757	0.2792	0.9683	0.5843
	Minimum	-0.1814	-0.3593	-0.3977	-0.3625	-0.5159
	Mean	-0.0003	0.0429	-0.0269	0.1097	-0.0941
	Standard deviation	0.0436	0.3174	0.1546	0.3707	0.3116
Congestive heart failure beat	Maximum	0.1316	0.2344	1.3364	1.3463	1.1550
	Minimum	-0.1119	-0.1635	-1.0327	-1.9773	-1.2350
	Mean	-0.0003	-0.0066	-0.0056	-0.0248	-0.3698
	Standard deviation	0.0259	0.0604	0.3995	0.7862	0.4014
Ventricular tachyarrhythmia beat	Maximum	0.1568	0.4554	2.1134	2.5063	4.1980
	Minimum	-0.0839	-0.3181	-0.8983	-1.4226	-0.5930
	Mean	-0.0001	0.0002	0.0344	0.0838	0.9075
	Standard deviation	0.0232	0.0919	0.4845	0.8299	1.1458
Atrial fibrillation beat	Maximum	0.0665	0.4417	0.3574	1.3044	-0.9396
	Minimum	-0.0564	-0.1832	-0.3312	-0.3328	-2.0488
	Mean	-0.0002	0.0037	-0.0058	0.0774	-1.5942
	Standard deviation	0.0173	0.0849	0.1238	0.4051	0.2892

Table 2. *The power levels of the PSDs obtained by the eigenvector methods of four exemplary records from four classes*

ECG beat types	Extracted Features	Pisarenko PSD values	MUSIC PSD values	Minimum-Norm PSD values
Normal beat	Maximum	-8.3262	-6.1429	-5.9119
	Minimum	-63.3942	-48.5747	-45.7063
	Mean	-29.0350	-25.5815	-24.9883
	Standard deviation	15.5270	14.6241	13.9427
Congestive heart failure beat	Maximum	15.4373	14.8120	11.8117
	Minimum	-58.5668	-54.1374	-52.5681
	Mean	-34.2724	-33.0135	-32.4818
	Standard deviation	21.1599	19.0780	18.4921
Ventricular tachyarrhythmia beat	Maximum	8.9680	6.5634	5.3926
	Minimum	-73.2121	-66.8833	-65.4340
	Mean	-44.5406	-43.3792	-42.9259
	Standard deviation	28.8429	25.2899	24.1945
Atrial fibrillation beat	Maximum	19.4951	23.0237	22.6390
	Minimum	-62.0746	-54.3199	-52.6743
	Mean	-39.9064	-36.4000	-36.0503
	Standard deviation	21.4749	19.9989	19.1709

Table 3. *Network parameters of the classifiers*

Classifiers (features)	Network parameters
MME (diverse features)	8·25·4[a], 8·25·4[a], 8·25·4[a], 8·25·4[a], 8·25·4[b], 8·25·4[b], 8·25·4[b], 8·25·4[b], 600[c]
ME (composite feature)	32·25·4[a], 32·25·4[d] 800[c]
CNN (composite feature)	32·30·16[e], 16·30·4[f] 1700[c]
MLPNN (composite feature)	32·25·4[g] 2500[c]

[a]Design of expert networks: Number of input · hidden · output neurons, respectively.
[b]Design of gating networks in gate-bank: Number of input · hidden · output neurons, respectively.
[c]Number of training epochs.
[d]Design of gating network: Number of input · hidden · output neurons, respectively.
[e]Design of first level network: Number of input · hidden · output neurons, respectively.
[f]Design of second level network: Number of input · hidden · output neurons, respectively.
[g]Design of neural network: Number of input · hidden · output neurons, respectively.

Table 4. *The total classification accuracies and the CPU times of training of the classifiers*

Classifiers (features)	Total classification accuracy (%)	CPU time (min:s)
MME (diverse features)	97.78	8:43
ME (composite feature)	96.11	10:17
CNN (composite feature)	95.28	14:25
MLPNN (composite feature)	91.67	15:24