ELSEVIER

Journal of Neuroscience Methods xxx (2005) xxx–xxx

JOURNAL OF NEUROSCIENCE METHODS

www.elsevier.com/locate/jneumeth

Adaptive neuro-fuzzy inference system for classification of EEG signals using wavelet coefficients

İnan Güler [a,*], Elif Derya Übeyli [b]

[a] Department of Electronics and Computer Education, Faculty of Technical Education, Gazi University, 06500 Teknikokullar, Ankara, Turkey
[b] Department of Electrical and Electronics Engineering, Faculty of Engineering, TOBB Ekonomi ve Teknoloji Üniversitesi, 06530 Söğütözü, Ankara, Turkey

Received 12 January 2005; received in revised form 12 April 2005; accepted 15 April 2005

Abstract

This paper describes the application of adaptive neuro-fuzzy inference system (ANFIS) model for classification of electroencephalogram (EEG) signals. Decision making was performed in two stages: feature extraction using the wavelet transform (WT) and the ANFIS trained with the backpropagation gradient descent method in combination with the least squares method. Five types of EEG signals were used as input patterns of the five ANFIS classifiers. To improve diagnostic accuracy, the sixth ANFIS classifier (combining ANFIS) was trained using the outputs of the five ANFIS classifiers as input data. The proposed ANFIS model combined the neural network adaptive capabilities and the fuzzy logic qualitative approach. Some conclusions concerning the saliency of features on classification of the EEG signals were obtained through analysis of the ANFIS. The performance of the ANFIS model was evaluated in terms of training performance and classification accuracies and the results confirmed that the proposed ANFIS model has potential in classifying the EEG signals.

Keywords: Adaptive neuro-fuzzy inference system (ANFIS); Fuzzy logic; Wavelet transform; Electroencephalogram (EEG) signals

1. Introduction

The electroencephalogram (EEG) signal is widely used clinically to investigate brain disorders. The study of the brain electrical activity, through the electroencephalographic records, is one of the most important tools for the diagnosis of neurological diseases (Adeli et al., 2003; Hazarika et al., 1997; Rosso et al., 2004). Large amounts of data are generated by EEG monitoring systems for electroencephalographic changes, and their complete visual analysis is not routinely possible. Computers have long been proposed to solve this problem and thus, automated systems to recognize electroencephalographic changes have been under study for several years (Glover et al., 1989; Gabor and Seyal, 1992; Webber et al., 1993; Nigam and Graupe, 2004). There is a strong demand for the development of such automated devices, due to the increased use of prolonged and long-term video EEG recordings for proper evaluation and treatment of neurological diseases and prevention of the possibility of the analyst missing (or misreading) information (Webber et al., 1993).

Abnormalities in the EEG in serious psychiatric disorders are at times too subtle to be detected using conventional techniques. Such techniques work by transforming the mostly qualitative diagnostic criteria into a more objective quantitative signal feature classification problem. The techniques have been used to address this problem such as the analysis of EEG signals for detection of electroencephalographic changes using the autocorrelation function, frequency domain features, time frequency analysis, and wavelet transform (WT) (Adeli et al., 2003; Hazarika et al., 1997; Rosso et al., 2004; Glover et al., 1989). The results of the studies in the literature have demonstrated that the WT is the most promising method to extract features from the EEG signals (Adeli et al., 2003; Hazarika et al., 1997; Rosso et al., 2004). In this respect, in the present study the WT was used for feature extraction from the EEG signals.

* Corresponding author. Tel.: +90 312 212 3976; fax: +90 312 212 0059.
E-mail address: iguler@gazi.edu.tr (İ. Güler).

0165-0270/$ – see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.jneumeth.2005.04.013

The 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 (Daubechies, 1990; Soltani, 2002; Unser and Aldroubi, 1996). Thus, the EEG signal, consisting of many data points, can be compressed into a few parameters. These parameters characterize the behavior of the EEG signal. This feature of using a smaller number of parameters to represent the EEG signal is particularly important for recognition and diagnostic purposes (Adeli et al., 2003; Hazarika et al., 1997; Rosso et al., 2004).

Artificial neural networks (ANNs) have been used as computational tools for pattern classification including diagnosis of diseases because of the belief that they have greater predictive power than signal analysis techniques (Baxt, 1990; Miller et al., 1992; Güler and Übeyli, 2003; Übeyli and Güler, 2003). However, fuzzy set theory plays an important role in dealing with uncertainty when making decisions in medical applications. Therefore, fuzzy sets have attracted the growing attention and interest in modern information technology, production technique, decision making, pattern recognition, diagnostics, data analysis, etc. (Dubois and Prade, 1998; Kuncheva and Steimann, 1999; Nauck and Kruse, 1999). Neuro-fuzzy systems are fuzzy systems, which use ANNs theory in order to determine their properties (fuzzy sets and fuzzy rules) by processing data samples. Neuro-fuzzy systems harness the power of the two paradigms: fuzzy logic and ANNs, by utilizing the mathematical properties of ANNs in tuning rule-based fuzzy systems that approximate the way humans process information. A specific approach in neuro-fuzzy development is the adaptive neuro-fuzzy inference system (ANFIS), which has shown significant results in modeling nonlinear functions. In ANFIS, the membership function parameters are extracted from a data set that describes the system behavior. The ANFIS learns features in the data set and adjusts the system parameters according to a given error criterion (Jang, 1992, 1993). Successful implementations of ANFIS in biomedical engineering have been reported, for classification (Usher et al., 1999; Belal et al., 2002; Güler and Übeyli, 2004; Übeyli and Güler, 2005a, 2005b) and data analysis (Virant-Klun and Virant, 1999).

In this study, a new approach based on ANFIS was presented for classification of the EEG signals. The proposed technique involved training the five ANFIS classifiers to classify the five classes of EEG signals when wavelet coefficients defining the behavior of the EEG signals were used as inputs. The ANFIS classifiers were trained with the backpropagation gradient descent method in combination with the least squares method. We used the data described in reference (Andrzejak et al., 2001), which is publicly available. In our applications, a detailed classification between set A (healthy volunteer, eyes open), set B (healthy volunteer, eyes closed), set C (seizure-free intervals of five patients from hippocampal formation of opposite hemisphere), set D (seizure-free intervals of five patients from epileptogenic zone), and set E (epileptic seizure segments) was performed. Each of the ANFIS classifier was trained so that they are likely to be more accurate for one class of EEG signals than the other classes. The predictions of the five ANFIS classifiers were combined by the sixth ANFIS classifier. The correct classification rates and convergence rates of the proposed ANFIS model were examined and then performance of the ANFIS model was reported. Finally, some conclusions were drawn concerning the saliency of features (inputs of the ANFIS classifiers) on classification of the EEG signals.

2. Materials and method

Decision making was performed in two stages: feature extraction using the WT (20 extracted features as ANFIS inputs) and classification using the ANFIS classifiers trained with the backpropagation gradient descent method in combination with the least squares method. In this section, we restrict ourselves to only a short description of the used datasets and refer to reference (Andrzejak et al., 2001) for further details. The complete dataset consists of five sets (denoted A–E), each containing 100 single-channel EEG signals of 23.6 s. Each signal has been selected after visual inspection for artifacts and has passed a weak stationarity criterion. Sets A and B have been taken from surface EEG recordings of five healthy volunteers with eyes open and closed, respectively. Signals in two sets have been measured in seizure-free intervals from five patients in the epileptogenic zone (D) and from the hippocampal formation of the opposite hemisphere of the brain (C). Set E contains seizure activity, selected from all recording sites exhibiting ictal activity. Sets A and B have been recorded extracranially, whereas sets C, D, and E have been recorded intracranially. Apart from the different recording electrodes, the recording parameters were fixed.

2.1. Feature extraction using wavelet transform

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 Fig. 1 (Daubechies, 1990; Soltani, 2002; Unser and Aldroubi, 1996).

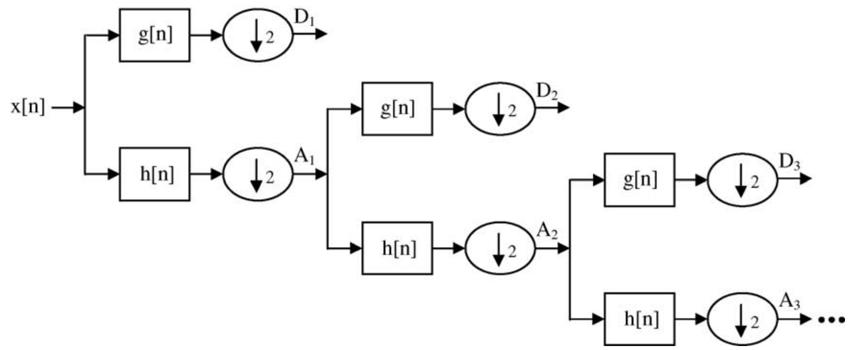

Fig. 1. Subband decomposition of discrete wavelet transform implementation; $g[n]$ is the high-pass filter, $h[n]$ is the low-pass filter.

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

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

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

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

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

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

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

$$\begin{aligned} h_{i+1}(k) &= [h]_{\uparrow 2^i} h_i(k) \\ g_{i+1}(k) &= [g]_{\uparrow 2^i} h_i(k) \end{aligned} \tag{4}$$

where the subscript $[\,]_{\uparrow m}$ 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

$$\begin{aligned} \varphi_{i,l}(k) &= 2^{i/2} h_i(k - 2^i l) \\ \psi_{i,l}(k) &= 2^{i/2} g_i(k - 2^i l) \end{aligned} \tag{5}$$

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

$$\begin{aligned} a_{(i)}(l) &= x(k)\varphi_{i,l}(k) \\ d_{(i)}(l) &= x(k)\psi_{i,l}(k) \end{aligned} \tag{6}$$

where $a_{(i)}(l)$ and $d_{(i)}(l)$ are the approximation coefficients and the detail coefficients at resolution i, respectively (Daubechies, 1990; Soltani, 2002; Unser and Aldroubi, 1996).

2.2. Adaptive neuro-fuzzy inference system (ANFIS)

2.2.1. Architecture of ANFIS

The ANFIS is a fuzzy Sugeno model put in the framework of adaptive systems to facilitate learning and adaptation (Jang, 1992, 1993). Such framework makes the ANFIS modeling more systematic and less reliant on expert knowledge. To present the ANFIS architecture, two fuzzy if-then rules based on a first order Sugeno model are considered:

Rule 1: If (x is A_1) and (y is B_1) then ($f_1 = p_1x + q_1y + r_1$)
Rule 2: If (x is A_2) and (y is B_2) then ($f_2 = p_2x + q_2y + r_2$)

where x and y are the inputs, A_i and B_i are the fuzzy sets, f_i are the outputs within the fuzzy region specified by the fuzzy rule, p_i, q_i and r_i are the design parameters that are determined during the training process. The ANFIS architecture to implement these two rules is shown in Fig. 2, in which a circle indicates a fixed node, whereas a square indicates an adaptive node.

In the first layer, all the nodes are adaptive nodes. The outputs of layer 1 are the fuzzy membership grade of the inputs, which are given by:

$$O_i^1 = \mu_{A_i}(x) \quad i = 1, 2 \tag{7}$$

$$O_i^1 = \mu_{B_{i-2}}(y) \quad i = 3, 4 \tag{8}$$

where $\mu_{A_i}(x)$, $\mu_{B_{i-2}}(y)$ can adopt any fuzzy membership function. For example, if the bell shaped membership

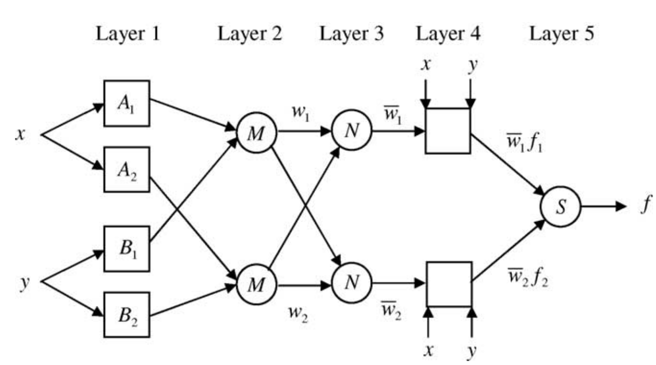

Fig. 2. ANFIS architecture.

function is employed, $\mu_{A_i}(x)$ is given by:

$$\mu_{A_i}(x) = \frac{1}{1 + \left\{ \left(\frac{x - c_i}{a_i} \right)^2 \right\}^{b_i}} \tag{9}$$

where a_i, b_i and c_i are the parameters of the membership function, governing the bell shaped functions accordingly.

In the second layer, the nodes are fixed nodes. They are labeled with *M*, indicating that they perform as a simple multiplier. The outputs of this layer can be represented as:

$$O_i^2 = w_i = \mu_{A_i}(x)\mu_{B_i}(y) \quad i = 1, 2 \tag{10}$$

which are the so-called firing strengths of the rules.

In the third layer, the nodes are also fixed nodes. They are labeled with *N*, indicating that they play a normalization role to the firing strengths from the previous layer.

The outputs of this layer can be represented as:

$$O_i^3 = \bar{w}_i = \frac{w_i}{w_1 + w_2} \quad i = 1, 2 \tag{11}$$

which are the so-called normalized firing strengths.

In the fourth layer, the nodes are adaptive nodes. The output of each node in this layer is simply the product of the normalized firing strength and a first order polynomial (for a first order Sugeno model). Thus, the outputs of this layer are given by:

$$O_i^4 = \bar{w}_i f_i = \bar{w}_i(p_i x + q_i y + r_i) \quad i = 1, 2 \tag{12}$$

In the fifth layer, there is only one single fixed node labeled with *S*. This node performs the summation of all incoming signals. Hence, the overall output of the model is given by:

$$O_i^5 = \sum_{i=1}^{2} \bar{w}_i f_i = \frac{\sum_{i=1}^{2} w_i f_i}{w_1 + w_2} \tag{13}$$

It can be observed that there are two adaptive layers in this ANFIS architecture, namely the first layer and the fourth layer. In the first layer, there are three modifiable parameters $\{a_i, b_i, c_i\}$, which are related to the input membership functions. These parameters are the so-called premise parameters. In the fourth layer, there are also three modifiable parameters $\{p_i, q_i, r_i\}$, pertaining to the first order polynomial. These parameters are so-called consequent parameters (Jang, 1992, 1993).

2.2.2. Learning algorithm of ANFIS

The task of the learning algorithm for this architecture is to tune all the modifiable parameters, namely $\{a_i, b_i, c_i\}$ and $\{p_i, q_i, r_i\}$, to make the ANFIS output match the training data. When the premise parameters a_i, b_i and c_i of the membership function are fixed, the output of the ANFIS model can be written as:

$$f = \frac{w_1}{w_1 + w_2} f_1 + \frac{w_2}{w_1 + w_2} f_2 \tag{14}$$

Substituting Eq. (11) into Eq. (14) yields:

$$f = \bar{w}_1 f_1 + \bar{w}_2 f_2 \tag{15}$$

Substituting the fuzzy if-then rules into Eq. (15), it becomes:

$$f = \bar{w}_1(p_1 x + q_1 y + r_1) + \bar{w}_2(p_2 x + q_2 y + r_2) \tag{16}$$

After rearrangement, the output can be expressed as:

$$\begin{aligned} f = {} & (\bar{w}_1 x)p_1 + (\bar{w}_1 y)q_1 + (\bar{w}_1)r_1 \\ & + (\bar{w}_2 x)p_2 + (\bar{w}_2 y)q_2 + (\bar{w}_2)r_2 \end{aligned} \tag{17}$$

which is a linear combination of the modifiable consequent parameters p_1, q_1, r_1, p_2, q_2 and r_2. The least squares method can be used to identify the optimal values of these parameters easily. When the premise parameters are not fixed, the search space becomes larger and the convergence of the training becomes slower. A hybrid algorithm combining the least squares method and the gradient descent method is adopted to solve this problem. The hybrid algorithm is composed of a forward pass and a backward pass. The least squares method (forward pass) is used to optimize the consequent parameters with the premise parameters fixed. Once the optimal consequent parameters are found, the backward pass starts immediately. The gradient descent method (backward pass) is used to adjust optimally the premise parameters corresponding to the fuzzy sets in the input domain. The output of the ANFIS is calculated by employing the consequent parameters found in the forward pass. The output error is used to adapt the premise parameters by means of a standard backpropagation algorithm. It has been proven that this hybrid algorithm is highly efficient in training the ANFIS (Jang, 1992, 1993).

3. Results and discussion

The EEG 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. Spectral analysis of the EEG signals was performed using the DWT (described in Section 2.1). 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 EEG 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 EEG signals. Therefore, the wavelet coefficients were computed using the db2 in the present study. The wavelet coefficients were computed using the MATLAB software package.

Selection of the neural network 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. Input selection has two meanings: (1) which components of a pattern, or (2) which set of inputs best represent a given pattern. The computed 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 EEG signals were used as the feature vectors representing the signals. A rectangular window, which was formed by 256 discrete data, was selected so that the EEG signal considered to be stationary in that interval. For each EEG segment, 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 EEG segment. In order to reduce the dimensionality of the feature vectors, statistics over the set of the wavelet coefficients were used. The following statistical features were used to represent the time-frequency distribution of the EEG signals:

(1) Maximum of the wavelet coefficients in each subband.
(2) Minimum of the wavelet coefficients in each subband.
(3) Mean of the wavelet coefficients in each subband.
(4) Standard deviation of the wavelet coefficients in each subband.

These feature vectors, which were calculated for the D_1–D_4 and A_4 frequency bands, were used in classifying the EEG signals. Table 1 presents the extracted features of five exemplary records from five classes. From Table 1, one can see that the extracted features of the five classes of EEG signals are different from each other. Therefore, we decided that they can serve as useful parameters in classifying the EEG signals.

The five ANFIS classifers were trained with the backpropagation gradient descent method in combination with the least squares method when 20 features (dimension of the extracted feature vectors) representing the EEG signals were used as inputs. Samples with target outputs sets A, B, C, D, and E were given the binary target values of (0, 0, 0, 0, 1), (0, 0, 0, 1, 0), (0, 0, 1, 0, 0), (0, 1, 0, 0, 0), and (1, 0, 0, 0, 0), respectively. To improve classification accuracy, the sixth ANFIS classifier (combining ANFIS) was trained using the outputs of the five ANFIS classifiers as input data. The fuzzy rule architecture of the ANFIS classifiers were designed by using a generalized bell shaped membership function defined in Eq. (9). Each ANFIS classifier was implemented by using the MATLAB software package (MATLAB version 7.0 with fuzzy logic toolbox). The data sets (sets A, B, C, D, and E) were divided into two separate data sets—the training data set and the testing data set. The adequate functioning of the ANFIS depends on the sizes of the training set and test set. In this study, the 100 time series of 4096 samples for each class windowed by a rectangular window composed of 256 discrete data and then training and test sets of the five ANFIS classifiers were formed by 8000 vectors (1600 vectors from each class) of 20 dimensions (dimension of the extracted feature

Table 1
The extracted features of five exemplary records from five classes

Dataset	Extracted features	Subbands				
		D_1	D_2	D_3	D_4	A_4
Set A	Maximum	12.0394	31.3064	75.7695	120.0146	192.6771
	Minimum	−12.0140	−42.0737	−92.3744	−105.3666	−172.4994
	Mean	−0.2611	0.1775	1.6022	2.1703	34.4130
	Standard deviation	4.9689	14.8416	41.1865	60.3469	96.4623
Set B	Maximum	14.1446	46.9284	102.2603	242.5219	302.9787
	Minimum	−14.7570	−51.4840	−139.1860	−157.7330	−208.8999
	Mean	0.4727	0.0892	−7.3273	−2.7413	24.0453
	Standard deviation	6.0482	17.9383	60.0547	88.4955	146.4562
Set C	Maximum	6.4079	17.1955	49.5235	142.3749	231.6009
	Minimum	−7.3739	−21.1106	−42.6390	−182.4811	−269.4633
	Mean	0.0668	−0.1359	2.2645	−12.3407	−39.0668
	Standard deviation	2.8001	9.5142	25.9131	95.0770	153.3921
Set D	Maximum	26.0292	117.9646	32.3480	88.2469	320.4451
	Minimum	−20.6820	−82.1600	−61.5424	−89.1512	−175.7673
	Mean	−0.1935	0.1121	−2.2112	−2.6360	94.1584
	Standard deviation	4.3874	19.2455	20.1756	43.6354	126.3576
Set E	Maximum	258.0806	644.3659	1524.4000	1420.1000	1639.2000
	Minimum	−325.4508	−1074.6000	−1508.9000	−1107.0000	−1917.6000
	Mean	−0.1337	0.1052	65.5614	−77.2298	281.4010
	Standard deviation	75.1448	303.6744	716.0870	614.2615	1138.5000

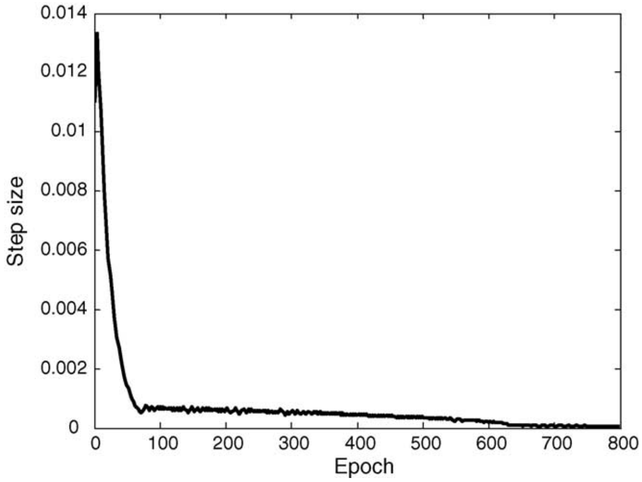

Fig. 3. Adaptation of parameter steps of each ANFIS.

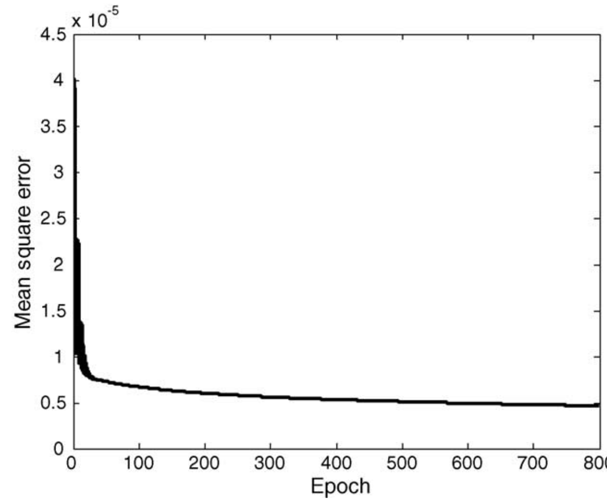

Fig. 4. The curve of network error convergence of each ANFIS.

vectors). The 4000 vectors (800 vectors from each class) of 20 dimensions were used for training and the 4000 vectors (800 vectors from each class) of 20 dimensions were used for testing.

The training data set was used to train the ANFIS model, whereas the testing data set was used to verify the accuracy and the effectiveness of the trained ANFIS model for classification of the five classes of EEG signals. Each ANFIS used 4000 training data in 800 training epochs and the step size for parameter adaptation had an initial value of 0.011. The steps of parameter adaptation of each ANFIS are shown in Fig. 3. The step size is decreased (by multiplying it with the component of the training option corresponding to the step size decrease rate) if the error measure undergoes two consecutive combinations of an increase followed by a decrease. The step size is increased (by multiplying it with the increase rate) if the error measure undergoes four consecutive decreases. At the end of 800 training epochs, the network error (mean square error) convergence curve of each ANFIS was derived as shown in Fig. 4. From the curve, the final convergence value is 4.6721×10^{-6}.

In a real world domain, just like the one used in the present study, all of the features used in the descriptions of instances may have different levels of saliency. Feature saliency provides a means for choosing the features, which are best for classification. Therefore, in the present study changes of the final (after training) membership functions with respect to the initial (before training) membership functions of the input parameters were examined. Membership function of each input parameter was divided into three regions, namely, small, medium, and large. The examination of initial and final membership functions indicates that there are considerable changes in the final membership functions of the wavelet coefficients. Figs. 5 and 6 show the initial and final membership functions of the first and seventh inputs (input 1: maximum values of the wavelet coefficients in D_1 subband and input 7: mean values of the wavelet coefficients in

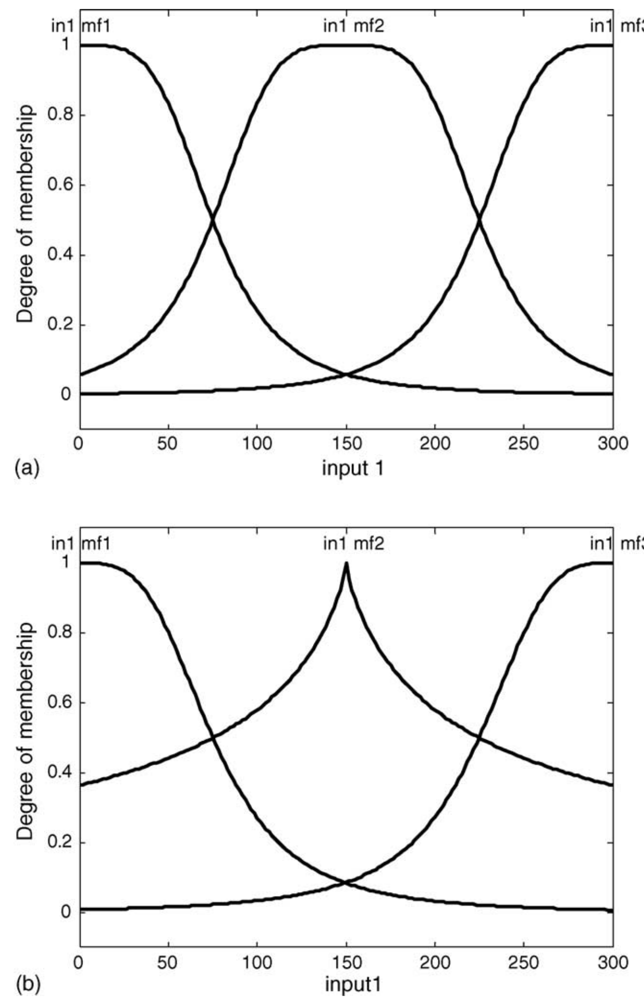

Fig. 5. (a) Initial and (b) final generalized bell shaped membership function of input 1 (maximum values of the wavelet coefficients in D_1 subband).

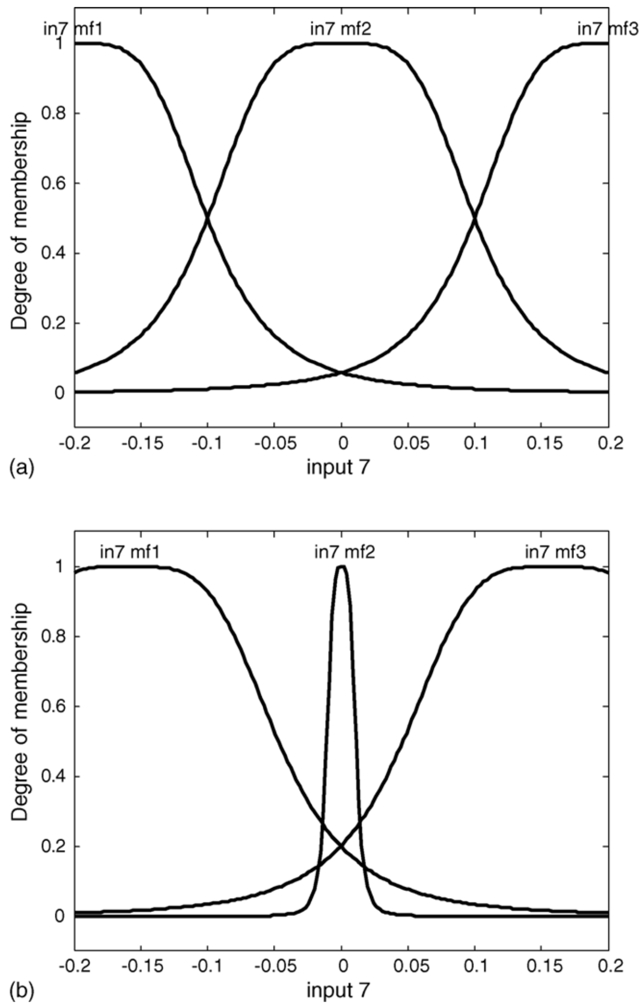

Fig. 6. (a) Initial and (b) final generalized bell shaped membership function of input 7 (mean values of the wavelet coefficients in D_2 subband).

D_2 subband) using the generalized bell shaped membership function, respectively. Based on the analysis of membership functions of each input features, it can be mentioned that all of the 20 inputs have saliency on the EEG signals classification.

After training, 4000 testing data were used to validate the accuracy of the ANFIS model for classification of the EEG signals. In classification, the aim is to assign the input patterns to one of several classes, usually represented by outputs restricted to lie in the range from 0 to 1, so that they represent the probability of class membership. While the classification is carried out, a specific pattern is assigned to a specific class according to the characteristic features selected for it. In this application, there were five classes: set A (healthy volunteer, eyes open), set B (healthy volunteer, eyes closed), set C (seizure-free intervals of five patients from hippocampal formation of opposite hemisphere), set D (seizure-free intervals of five patients from epileptogenic zone), and set E (epileptic seizure segments). Classification results of the ANFIS model were displayed by a confusion matrix. In a confusion matrix, each cell contains the raw number of exemplars classified for the corresponding combination of desired and actual network outputs. The confusion matrix showing the classification results of the ANFIS model is given below.

Confusion matrix

Output/desired	Set A	Set B	Set C	Set D	Set E
Set A	789	9	0	0	0
Set B	9	791	0	0	0
Set C	2	0	793	10	5
Set D	0	0	6	788	9
Set E	0	0	1	2	786

According to the confusion matrix, nine EEG segments from set A were classified incorrectly by the ANFIS model as segments from set B, two segments from set A were classified as segments from set C, nine segments from set B were classified as segments from set A, six segments from set C were classified as segments from set D, one segment from set C was classified as a segment from set E, 10 segments from set D were classified as segments from set C, two segments from set D were classified as segments from set E, nine segments from set E were classified as segments from set D, and five segments from set E were classified as segments from set C.

The test performance of the classifiers can be determined by the computation of sensitivity, specificity and total classification accuracy. The sensitivity, specificity and total classification accuracy are defined as:

Sensitivity: number of true positive decisions/number of actually positive cases.
Specificity: number of true negative decisions/number of actually negative cases.
Total classification accuracy: number of correct decisions/total number of cases.

The values of the statistical parameters (sensitivity, specificity and total classification accuracy) are given in Table 2. As it is seen from Table 2, the ANFIS classified sets A, B, C, D, and E with the accuracy of 98.63, 98.88, 99.13, 98.50, and 98.25%, respectively. All of the sets were classified with the accuracy of 98.68% (total classification accuracy). Nigam and Graupe (2004) described a method for automated detection of epileptic seizures from EEG signals using a multistage

Table 2
The values of statistical parameters

EEG datasets	Statistical parameters		
	Sensitivity (%)	Specificity (%)	Total classification accuracy (%)
Set A	98.63	99.72	98.68
Set B	98.88	99.72	
Set C	99.13	99.46	
Set D	98.50	99.53	
Set E	98.25	99.91	

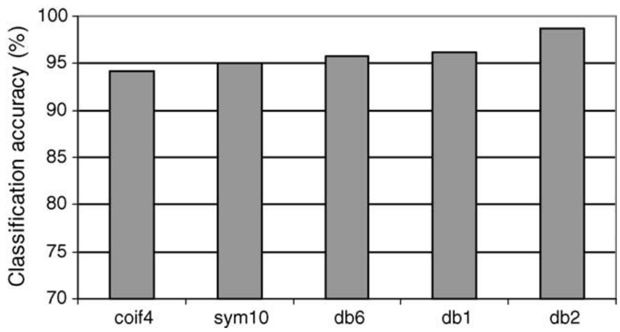

Fig. 7. Total classification accuracy obtained for different wavelets when the EEG signals were classified using the proposed ANFIS.

nonlinear pre-processing filter in combination with a diagnostic ANN. In order to train their network, they used two sets of the same EEG datafiles—sets A and E (Andrzejak et al., 2001) and the total classification accuracy of their model was 97.2%. Thus, the accuracy rates of the ANFIS model presented for this application were found to be higher than that of the ANN model presented by Nigam and Graupe (2004).

The classification accuracy, which is defined as the percentage ratio of the number of segments correctly classified to the total number of segments considered for classification, depends on the type of wavelet chosen for the application. The db2 was used and found to yield good results in classifying the EEG signals. In order to investigate the effect of other wavelets on classifications accuracy, tests were carried out using other wavelets also. Apart from the db2, Coiflet of order 4 (coif4), Symmlet of order 10 (sym10), Daubechies of order 1 (db1), and Daubechies of order 6 (db6) were also tried. Total classification accuracy obtained for each wavelet when the EEG signals were classified using the proposed ANFIS model, is shown in Fig. 7. One can see that the Daubechies wavelet offers better accuracy than the others, and the db2 is marginally better than the db1 and db6. Hence, the db2 wavelet was chosen for this application.

4. Conclusion

Fuzzy set theory plays an important role in dealing with uncertainty when making decisions in medical applications. Using fuzzy logic enabled us to use the uncertainty in the classifier design and consequently to increase the credibility of the system output. This paper presented a new application of ANFIS for classification of the EEG signals. The five ANFIS classifiers were used to classify five classes of EEG signals when the wavelet coefficients of the EEG signals were used as inputs. The predictions of the five ANFIS classifiers were combined by the sixth ANFIS classifier. The presented ANFIS model combined the neural network adaptive capabilities and the fuzzy logic qualitative approach. Some conclusions concerning the saliency of features on classification of the EEG signals were obtained through analysis of the ANFIS. The classification results and statistical measures were used for evaluating the ANFIS. The total classification accuracy of the ANFIS model was 98.68%. We therefore have concluded that the proposed ANFIS model can be used in classifying the EEG signals by taking into consideration the misclassification rates.

Acknowledgements

This study has been supported by Scientific Research Project of Gazi University (Project No: 07/2004-04).

References

Adeli H, Zhou Z, Dadmehr N. Analysis of EEG records in an epileptic patient using wavelet transform. J Neurosci Methods 2003;123(1):69–87.

Andrzejak RG, Lehnertz K, Mormann F, Rieke C, David P, Elger CE. Indications of nonlinear deterministic and finite-dimensional structures in time seires of brain electrical activity: dependence on recording region and brain state. Phys Rev E 2001;64:061907.

Baxt WG. Use of an artificial neural network for data analysis in clinical decision making: the diagnosis of acute coronary occlusion. Neural Comput 1990;2:480–9.

Belal SY, Taktak AFG, Nevill AJ, Spencer SA, Roden D, Bevan S. Automatic detection of distorted plethysmogram pulses in neonates and paediatric patients using an adaptive-network-based fuzzy inference system. Artif Intell Med 2002;24:149–65.

Daubechies I. The wavelet transform, time-frequency localization and signal analysis. IEEE T Inform Theory 1990;36(5):961–1005.

Dubois D, Prade H. An introduction to fuzzy systems. Clin Chim Acta 1998;270:3–29.

Gabor AJ, Seyal M. Automated interictal EEG spike detection using artificial neural networks. Electroencephalogr Clin Neurophysiol 1992;83(5):271–80.

Glover Jr JR, Raghaven N, Ktonas PY, Frost Jr JD. Context-based automated detection of epileptogenic sharp transients in the EEG: elimination of false positives. IEEE Trans Biomed Eng 1989;36(5):519–27.

Güler İ, Übeyli ED. Detection of ophthalmic artery stenosis by least-mean squares backpropagation neural network. Comput Biol Med 2003;33(4):333–43.

Güler İ, Übeyli ED. Application of adaptive neuro-fuzzy inference system for detection of electrocardiographic changes in patients with partial epilepsy using feature extraction. Expert Syst Appl 2004;27(3):323–30.

Hazarika N, Chen JZ, Tsoi AC, Sergejew A. Classification of EEG signals using the wavelet transform. Signal Process 1997;59(1):61–72.

Jang J-SR. Self-learning fuzzy controllers based on temporal backpropagation. IEEE Trans Neural Netw 1992;3(5):714–23.

Jang J-SR. ANFIS: Adaptive-network-based fuzzy inference system. IEEE Trans Syst Man Cybern 1993;23(3):665–85.

Kuncheva LI, Steimann F. Fuzzy diagnosis. Artif Intell Med 1999;16:121–8.

Miller AS, Blott BH, Hames TK. Review of neural network applications in medical imaging and signal processing. Med Biol Eng Comput 1992;30:449–64.

Nauck D, Kruse R. Obtaining interpretable fuzzy classification rules from medical data. Artif Intell Med 1999;16:149–69.

Nigam VP, Graupe D. A neural-network-based detection of epilepsy. Neurol Res 2004;26(1):55–60.

Rosso OA, Figliola A, Creso J, Serrano E. Analysis of wavelet-filtered tonic-clonic electroencephalogram recordings. Med Biol Eng Comput 2004;42(4):516–23.

Soltani S. On the use of the wavelet decomposition for time series prediction. Neurocomputing 2002;48:267–77.

Übeyli ED, Güler İ. Neural network analysis of internal carotid arterial Doppler signals: Predictions of stenosis and occlusion. Expert Syst Appl 2003;25(1):1–13.

Übeyli ED, Güler İ. Automatic detection of erythemato-squamous diseases using adaptive neuro-fuzzy inference systems. Comput Biol Med 2005a;35(5):421–33.

Übeyli ED, Güler İ. Adaptive neuro-fuzzy inference systems for analysis of internal carotid arterial Doppler signals. Comput Biol Med (2005b), in press.

Unser M, Aldroubi A. A review of wavelets in biomedical applications. Proc IEEE 1996;84(4):626–38.

Usher J, Campbell D, Vohra J, Cameron J. A fuzzy logic-controlled classifier for use in implantable cardioverter defibrillators. Pace-Pacing Clin Electrophysiol 1999;22:183–6.

Virant-Klun I, Virant J. Fuzzy logic alternative for analysis in the biomedical sciences. Comput Biomed Res 1999;32:305–21.

Webber WRS, Litt B, Lesser RP, Fisher RS, Bankman I. Automatic EEG spike detection: what should the computer imitate? Electroencephalogr Clin Neurophysiol 1993;87(6):364–73.