Biomed Eng Lett (2012) 2:147-157
DOI 10.1007/s13534-012-0066-7

ORIGINAL ARTICLE

Relative Wavelet Energy and Wavelet Entropy Based Epileptic Brain Signals Classification

Yatindra Kumar, Mohan Lal Dewal and Radhey Shyam Anand

Received: 29 April 2012 / Revised: 16 July 2012 / Accepted: 11 August 2012

Abstract
Purpose Manual analysis of EEG signals by an expert is very much time consuming due to the long length of EEG recordings. The suitable computerized analysis is essentially required to differentiate among the normal, interictal and ictal (epileptic) EEGs.
Methods In the present work the EEG signals are decomposed into different sub-bands using discrete wavelet transform (DWT) to obtain the detail and the approximation wavelet coefficients. The coefficients are used to calculate the quantitative values of relative wavelet energy and wavelet entropy from different data sets to select the features of EEG signals. The support vector machine (SVM), feed forward back- propagation neural network (FFBPNN), k-Nearest Neighbor Classifier (k-NN) and Decision tree classifier (DT) are used to classify the EEG signals.
Results It is revealed that the accuracy between normal subjects with eyes open condition (data set A) epileptic data set E using SVM is obtained as 96.25%. Classification accuracy between the normal subjects with eye closed condition and epileptic data set E is obtained as 83.75% using k-NN classifier. Similar accuracies while discriminating the interictal data set C versus ictal data set E, and interictal data set D versus ictal data set E are obtained as 97.5% and 97.5% respectively, using a FFBPNN. These accuracies are quite higher than the earlier results published. The results are discussed quite in detail towards the last sections of the present paper.
Conclusions Our experimental results demonstrate that the proposed method gives quite high statistical parameters for EEG classifications especially to classify the interictal data(C, D) and ictal data (E). These experiments indicate that the present method can be useful in analyzing and detecting the EEG signal associated with epilepsy.

Yatindra Kumar (✉)
Research Scholar, Department of Electrical Engineering, Indian Institute of Technology, Roorkee, Uttarakhand, India
Tel : +91-1332-286364 / Fax : +91-1332-273560
E-mail : ykr01dee@iitr.ernet.in

Mohan Lal Dewal and Radhey Shyam Anand
Department of Electrical Engineering, Indian Institute of Technology, Roorkee, Uttarakhand India

Keywords Electroencephalogram (EEG), Discrete wavelet transform (DWT), Wavelet entropy (WEN), Support vector machine (SVM), Artificial neural network (ANN), k-Nearest neighbor (k-NN)

INTRODUCTION

The neural activity of the human brain starts between the 17th and 23rd week of prenatal development. It is believed that from this early stage and throughout life electrical signals generated by the brain represent not only the brain function but also the status of the whole body. EEG is a non-invasive testing method which contains a lot of information about the state of a patient's health and different physiological states of the brain. It can be recorded over a long time span for monitoring neurological disorders like epileptic seizures which are not permanently present in the recordings. These EEG recordings are visually inspected by highly trained professionals for detecting epileptic seizures. This information is then used for clinical diagnosis and possible treatment plans [1].

Approximately 1% of the world's population suffers from epilepsy, a disorder of the normal brain function, characterized by the existence of abnormal synchronous discharges in large ensembles of neurons in brain structures. Epileptic seizures are manifestations of epilepsy, which are due to the sudden development of synchronous neuronal firing in the cerebral cortex and are recorded using the EEG. Epileptic seizures may occur in the brain locally (partial seizures), which are seen only in a few channels of the EEG recording, or involving the whole brain (generalized seizures), which are seen in every channel of the EEG recording. Clinical neurologists in daily practice commonly examine short recordings (usually 20-min recordings) of interictal periods. The most common forms of the interictal periods are the individual or isolated spikes, the sharp wave, and the spike-and-wave complex. These are perceived in the majority of patients with epilepsy. For this reason, interictal event detection plays a vital role in the diagnosis of epilepsy. However, during an isolated spike, the brain is not in a clinical seizure. A different EEG pattern is observed during the ictal period consisting of rhythmical waveforms for a wide variety of frequencies, polyspike activity, and low-amplitude desynchronization, as well as spike and wave complexes [2].

Research on epileptic EEG detection began in the 1970s and various methods addressing this problem have been presented. The early methods for processing of EEG signals are based on traditional Fourier transform and Fast Fourier transform (FFT) which have disadvantage of being highly sensitive to noise [3]. Discrete wavelet transform (DWT) is more convenient compared to the Fourier and Fast Fourier transforms because it has the advantages of time-frequency localization, multirate filtering, and scale-space analysis [3]. Moreover, DWT is appropriate for analysis of non-stationary signals. By means of DWT, each EEG signal is decomposed into its different sub-bands. These sub-bands may yield more accurate information than whole EEG about constituent neuronal activities [4].

EEG signals can be analyzed with nonlinear chaotic methods like entropies. Andrzejak et al. used correlation dimension (CD) to characterize the interictal EEG for seizures predictions and found that the CD values calculated from interictal EEG recordings are significantly lower for the epileptogenic zone as compared to other areas of the brain [5, 6]. Rosso et al. [7-9] has shown that the normalized wavelet entropy carries information about the degree of order/disorder associated with a multi-frequency signal response, and allows for the evaluation of the complexity behavior associated to the signal throughout the normalized wavelet complexity. Mirzaei et al. [10] has discussed the wavelet spectral entropy and spectral entropy for epileptic seizure detection. Nashash et al. [11] has discussed the Sub band wavelet Energy (SWE), which is used to analyze the recorded EEG during global hypoxic-ischemic injury and subsequent recovery stages. The SWE in each level of wavelet decomposition reflects the essential degree of spiking and bursting activity associated with the recovery phenomenon. Based on observations of the entropy variations obtained from different levels of injury, the normalized mean entropy can be used to segment the recovering EEG [11]. Guo et al. [12] analyzes the EEG signals using Relative wavelet energy (RWE) associated with different frequency bands and classify using ANNs. Kumar, et al. discussed the wavelet entropy to analyze the EEG signals for epileptic seizure detection using ANNs. [13].

Support vector machine (SVM) was first introduced in 1995 and has been used in many EEG signal classification problems. Nicolaou et al. [14] has employed a Permutation Entropy (PE) as a feature which drops during seizure interval for automated seizure detection using SVM. Hsu et al. [15] developed a method using the SVML and SVMRBF classifiers with nonlinear features for automatic seizure detection in EEG signals. Acharya et al. [16] have extracted the four entropy features from the EEG signals and fed to seven different classifiers and reported that the fuzzy classifier is able to differentiate with quite high accuracy [16]. Ubeyli [17] carried out a study for classification of EEG signals by combination of the model based methods and least square support vector machine and used the Burrg coefficients as features to achieve the high classification accuracy. Iscan et al. [18] proposed to combine the time and frequency feature approach for classification of healthy and epileptic EEG signals using different classifier including SVM.

Artificial neural networks (ANNs) have been widely applied to classify EEG signals over the last two decades [19-20]. A variety of different ANN based approaches are reported in the literature for epileptic seizure detection [21-23]. Kalayci et al. [24] used wavelet transform to capture characteristic features of the EEG signals and then combined with ANN to get satisfying classification result. Nigam et al. [25] described a method for automated detection of epileptic seizures from EEG signals using a multistage non-linear preprocessing filter for extracting two features: relative spike amplitude and spike occurrence frequency. These features are fed to a diagnostic artificial neural network.

In the present work a method is proposed based on the decomposition of EEG signals into different sub-bands using discrete wavelet transform. The RWE and WEN values are calculated from these sub-bands to form feature vectors. The proposed method is tested using clinical EEG data recorded from five healthy subjects and five epileptic patients during both ictal and interictal periods. Classification of EEG signals is done by SVM, FFBPNN, k-NN and Decision tree classifiers. The SVM gives the best classification accuracy (96.25%) based on these features vector for data set A versus data set E. The k-nearest neighbor classifier gives the 83.75% classification accuracy for data set B versus data set E. FFBPNN gives best classification accuracy (97.5%) for data set C, D individually, versus data set E. Where A and B are the data set of normal subject with eyes open and closed conditions, respectively. C, D and E are the data sets of interictal and ictal periods, respectively, for epileptic patents. The paper is organized as follows. In Section 2, the basics of DWT theory is described in brief. Section 3 discusses the dataset and the proposed methodology. In Section 4, the evaluation procedure and the experimental results are presented which are concluded in Section 5.

TECHNICAL BACKGROUND

Discrete wavelet transforms (DWT)

The continuous wavelet transform (CWT) [1] of a signal, $x(t)$, is the integral of the signal multiplied by scaled and shifted versions of a wavelet function ψ and is defined by,

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

Where a and b are called the scaling and shifting parameters, respectively. Calculation of wavelet coefficients at every possible scale is computationally very expensive task. Instead, if the scales and shifts are selected based on powers of two, the so-called dyadic scales and positions, then the wavelet analysis will be much more efficient. Such analysis is obtained from the DWT which is defined as,

$$DWT(j,k) = \frac{1}{\sqrt{|2^j|}}\int_{-\infty}^{\infty} x(t)\,\psi\left(\frac{t-2^j k}{2^j}\right)dt \tag{2}$$

Where a and b are replaced by 2^j and $2^j k$, respectively. Mallat et al. [4] developed an efficient way for implementing this scheme by passing the signal through a series of low-pass (LP) and high-pass (HP) filter pairs named as quadrature mirror filters. In the first step of the DWT, the signal is simultaneously passed through an LP and a HP filter with the cut-off frequency being the one fourth of the sampling frequency. The outputs from the low and high pass filters are referred to as approximation and detail coefficients of the first level, respectively. The output signals having half the frequency bandwidth of the original signals can be down-sampled by two according to Nyquist rule. The same procedure can be repeated for the first level approximation coefficients to get the second level coefficients. At each step of this decomposition process, the frequency resolution is doubled through filtering and the time resolution is halved through down sampling.

EEG DATA SETS AND PROPOSED METHOD

EEG data sets

Five sets (denoted by A-E) each containing 100 single channel EEG segments of 23.6-sec duration each, are obtained from University of Bonn Germany. These segments were selected and cut out from continuous multichannel EEG recordings after visual inspection for artifacts, which may be due to muscle activity or eye movements. Sets A and B consist of segments taken from surface EEG recordings that are carried out on five healthy volunteers using a standardized electrode placement scheme. Volunteers are relaxed in an awake state with eyes open (A) and eyes closed (B), respectively. Sets C, D, and E are recorded using intracranial electrodes exhibiting interictal and ictal epileptic activities. Segments in set D are recorded from within the epileptogenic zone, and those in set C from the hippocampal formation of the opposite hemisphere of the brain. While sets C and D contained only activity measured during seizure free intervals, set E only contained seizure activity.

All the EEG signals are recorded with the 128-channel amplifier system, using an average common reference [omitting electrodes containing pathological activity (C, D, and E) or strong eye movement artifacts (A and B)]. After converting these signals using 12 bit analog-to-digital converter, the data are written continuously onto the disk of a data acquisition computer system at a sampling rate of 173.61 Hz. Band-pass filter settings are 0.53−40 Hz (12 dB/oct.) [12].

Proposed method

In the present work, EEG data sets (A, B, C, D and E) are pre-processed through DWT to decompose into six sub-bands signals using fifth level decomposition. RWE and WEN values are calculated from these sub-band signals to form feature vector. These features are used as an input to SVM, ANN, k-NN and Decision tree classifiers to classify the EEG data sets into normal/interictal and ictal data sets. The block diagram of the proposed approach is shown in Fig. 1.

Preprocessing of EEG using DWT

Sampling frequency of the given EEG signals is 173.61 Hz. As per Nyquist sampling theorem, the maximum useful frequency that should be applied is half of the sampling frequency. Hence the frequency band that is decomposed by DWT is 0-86.81 Hz. The DWT that is applied to band-limited EEG is fifth-order bi-orthogonal wavelet transform.

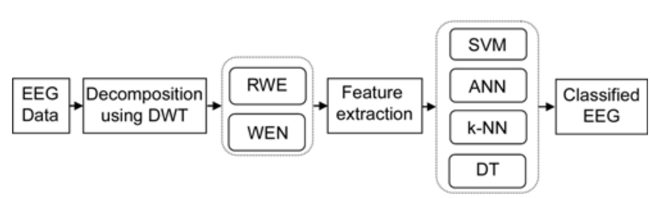

Fig. 1. Block diagram of proposed method.

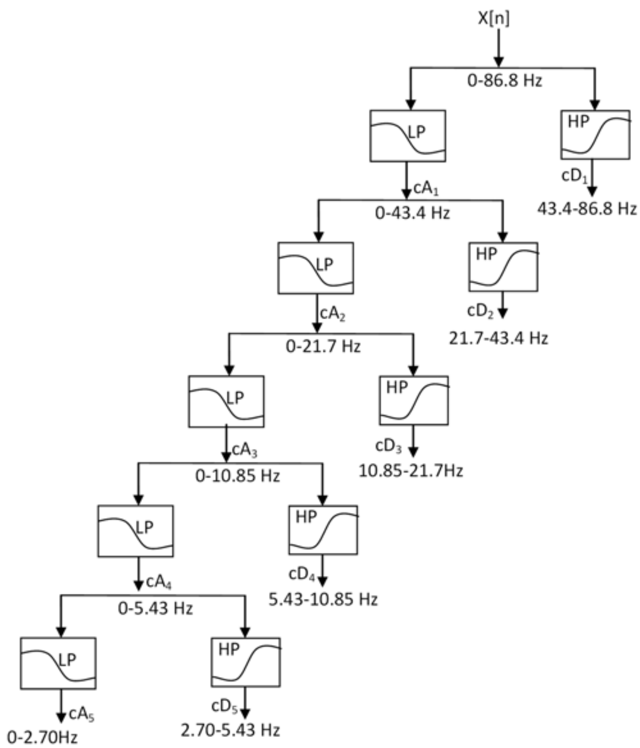

Fig. 2. Schematic of the EEG signal decomposition based on DWT.

In the first level decomposition, the whole EEG signal is passed through an LP and a HP filter, simultaneously. The cut off frequency of the filters is one fourth of the sampling frequency. The first level outputs are named cA1 and cD1 wavelets coefficients. According to wavelet analysis theory, the cA1 and cD1 wavelet coefficients are achieved by applying down sampling; where cA1 corresponds to 0-43.4 Hz band and cD1 corresponds to 43.4-86.8 Hz band. Time resolution is halved by down sampling and frequency resolution is doubled by filtering. In the second level of decomposition the cA1 coefficient that is obtained from the first level is decomposed to cA2 and cD2 wavelet coefficient sets that correspond to 0-21.7 Hz and 21.7-43.4 Hz, respectively. This process is repeated for five levels of decomposition. After that six coefficient sets cD1, cD2, cD3, cD4, cD5 and cA5 are achieved representing the full frequency band 0-86.8 Hz. In this process after each level of decomposition the number of coefficient are halved. Fig. 2 shows the schematic of this process.

Relative wavelet energy and wavelet entropy

The entropies of the wavelet coefficients on different scales are measured to reveal the information content carried by the coefficients. Wavelet entropy is a measure of the degree of order/disorder of the signal and it indicates the latent dynamical properties of the non-linear signals [7-9]. The given discrete signal $x(n)$ is transformed at instant k and scale j. It has a high frequency component wavelet coefficient $D_j(k)$ and a low-frequency component wavelet coefficient $A_j(k)$. The frequency bands of the information contained in signal component $D_j(k)$ and $A_j(k)$ are obtained as follows

$$\left.\begin{array}{l} D_j(k):[2^{-(j+1)}f_s,\ 2^{-j}f_s] \\ A_j(k):[0,2^{-(j+1)}f_s] \end{array}\right\} \quad \text{Where } j=1,2,\ldots,J \tag{3}$$

where f_s is the sampling frequency.

The energy at different decomposition level from (1,…,N) is the energy of wavelet detail coefficients $D_j(k)$ and wavelet approximate coefficient $A_j(k)$.

Energy at each decomposition level is defined as

$$E_j = \sum_k |D_j(k)|^2 \tag{4}$$

$$E_{N+1} = \sum_k |A_j(k)|^2 \tag{5}$$

Mean wavelet Energy of detail coefficients is defined as

$$\bar{E}_j = \frac{1}{N_j}\sum_k |D_j(k)|^2, \tag{6}$$

Where, N_j is the number of wavelet detail coefficients at j level.

Mean wavelet Energy of approximate coefficients,

$$\bar{E}_{N+1} = \frac{1}{N_J}\sum_k |A_j(k)|^2 \tag{7}$$

Where, N_J is the number of wavelet approximate coefficients at J level.

Then, the total energy of the signal, after wavelet decomposition, is obtained by

$$E_{total} = \sum_{j=1}^{N+1} \bar{E}_j \quad j = 1,\ldots,N+1 \tag{8}$$

Thus the relative wavelet energy (RWE) is defined as,

$$\rho_j = \frac{\bar{E}_j}{E_{total}} \tag{9}$$

$\sum_j \rho_j = 1$ and distribution ρ_j can be considered as time scale density. This provides information to characterize signal energy distribution at different frequency bands.

Wavelet entropy is defined as

$$WEN = -\sum \rho_j \log \rho_j \tag{10}$$

Classifications

Support vector machine (SVM)

Support Vector Machines are among the most prominent multiclass classification techniques for high dimensional feature vectors. The SVM is originally designed for binary classification. The SVMs find a hyper-plane in the feature space, which maximizes the distance between this hyper-plane and the nearest data point of each class. The architecture of the SVM depends on the regularization parameter and the type of the kernel function. The regularization parameter is employed to control the amount of allowed overlap between classes. There are various kernel functions including: linear, polynomial, radial basis function (RBF), and sigmoid. These functions have one or more free parameters referred to as hyper parameters. To train a SVM, the user must determine a suitable kernel function, optimum hyper parameters and proper regularization parameter. This goal is usually accomplished by cross-validation techniques.

Feed forward back propagation neural network

Neural network is an information processing system and it has been the choice of many researchers for the classification due to its special characteristics such as self learning, adaptability and robustness and massive parallelism [24, 27]. It consists of many computational neural units connected to each other. In Neural networks knowledge about the problem is distributed through the connection weights of links between neurons. The neural network has to be trained to adjust the connection weights and biases in order to produce the desired mapping. Neural networks are widely used in the biomedical area for modeling, data analysis, and diagnostic recognition. The training algorithm is an important part of the neural network model. A suitable training algorithm has a short training process, while achieving better accuracy. The

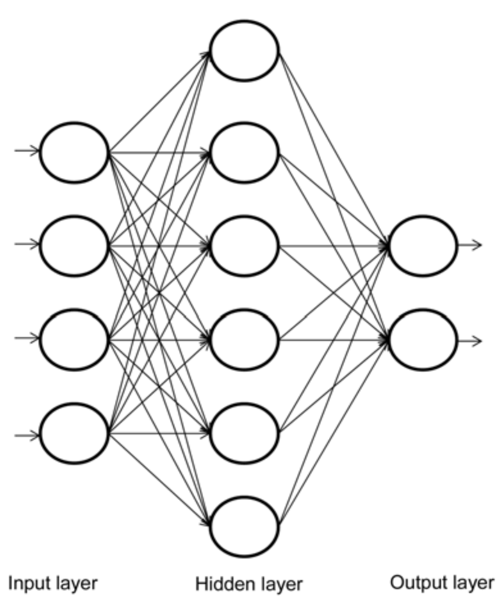

Fig. 3. Architecture of FFBPNN.

feed forward back propagation neural network (FFBPNN) is used to classify the EEGs. The function 'newff' creates a feed forward network and initializes the weight/bias of the network. The architecture of feed forward network are shown in Fig. 3

k-Nearest neighbor classifiers

The k-nearest neighbor (k-NN) algorithm is a non-parametric supervised classification algorithm. The k-NN algorithm is amongst the simplest algorithms for classifying objects in machine learning. All feature vectors extracted from the sub-training set are located in the feature space. A feature vector belonging to the test data is classified with the class of the majority of the k-nearest neighbors among all the located feature vectors. The performance of the nearest neighbor classifier depends on the distance function and the value of the neighborhood parameter k, which controls the volume of the neighborhood. The most common choice for the distance function is the Euclidean metric. The k-NN algorithm is very sensitive to the local distribution of feature vectors. However, this algorithm does not require a prior process of training, as it is not necessary to set the value of the parameters [38].

Decision tree classifier

A decision tree (DT) is a classifier articulated as a recursive partition of the instance space. The decision tree consists of nodes that form a rooted tree; this means it is a directed tree with a node called root that has no incoming edges. All other nodes have exactly one incoming edge. A node with outgoing edges is called an internal or test node. All other nodes are called leaves, or terminal or decision nodes. In a decision tree, each internal node splits the instance space into two or more sub-spaces according to a certain discrete function of the input attributes values. In the simplest and most frequent case, each test considers a single attribute, such that the instance space is partitioned according to the attribute's value. In the case of numeric attributes, the condition refers to a range. Each leaf is assigned to one class representing the most appropriate target value. Alternatively, the leaf may hold a probability vector indicating the probability of the target attribute having a certain value. Instances are classified by navigating them from the root of the tree down to a leaf, according to the outcome of the tests along the path [39].

Working of decision tree learning is as follows [30]:

- Select an attribute and formulate a logical test on attribute.
- Branch on each outcome of test, move subset of training data satisfying that outcome to the corresponding child node.

- Run recursively on each child node.
- Termination rule specifies when to declare a leaf node.

Definitions that used training of decision tree learning:

- Selection: used to partition training data.
- Termination condition: determines when to stop partitioning.
- Pruning algorithm: attempts to prevent over-fitting.

K-Fold cross-validation

Cross-validation is the statistical practice of partitioning a sample of data into subsets such that the analysis is initially performed on a single subset, while the other subset(s) are retained for subsequent use in confirming and validating the initial analysis. The initial subset of data is called the training set; the other subset(s) are called validation or testing sets [26]. In K-fold cross-validation, the original sample is partitioned into K sub samples. K-1 sub samples are used as training data and single sub sample is retained as the validation data for testing the model. The cross-validation process is then repeated K times, with each of the K sub samples used exactly once as the validation data. The K results from the folds then can be averaged to produce a single estimation. In this study, we have used default 5-fold scheme to achieve best performance accuracies.

Statistical parameters

The performance of neural network is evaluated by using three parameters, namely sensitivity, specificity and classification accuracy. The definitions of these parameters are as follows:

$$Sensitivity(SEN) = \frac{True\ Positives}{True\ Positives + False\ Negatives} \times 100\%$$

$$Specificity(SPE) = \frac{True\ Negatives}{True\ Negatives + False\ Positives} \times 100\%$$

$$Classification\ Accuracy(CA) = \frac{Correct\ Classified\ Patterns}{Total\ Patterns} \times 100\%$$

RESULTS AND DISCUSSION

All the 500 patterns of normal, interictal and epileptic (ictal) EEG data sets are decomposed into different sub-bands using DWT. The frequency ranges of these sub-bands are as follows: cA1 (0–43.4 Hz), cA2 (0–21.7 Hz), cA3 (0–10.85 Hz), cA4 (0–5.43 Hz), cA5 (0–2.70 Hz), cD1 (43.4–86.8 Hz), cD2 (21.7–43.4 Hz), cD3 (10.85–21.7 Hz), cD4 (5.43–10.85 Hz) and cD5 (2.70–5.43 Hz). The number of wavelet coefficients is halved after each level of decomposition which in turn reduces the calculation time appreciably. Figs. 4 and 5 illustrate a typical normal and epileptic EEG and its wavelet coefficients.

In this work, Energy ($\overline{E}_1, \overline{E}_2, ..., \overline{E}_6$) can be calculated using Eqs. (6) and (7) from the wavelet detail coefficients cD1-cD5 and wavelet approximate coefficients cA5. Thus, relative wavelet energy ($\rho_1, \rho_2, ..., \rho_6$) can be calculated according to Eq. (9) for different frequency bands from the Energy ($\overline{E}_1, \overline{E}_2, ..., \overline{E}_6$) and tabulated in Table 1. Table 1 shows that the RWE (ρ_6) value of 0-2.70 Hz sub-band for epileptic signal is lower than the ρ_6 of normal and interictal EEG signals. RWE (ρ_3, ρ_4, ρ_5) values of 2.70-21.7 Hz sub-band for data set E are higher than the data sets A, C, and D and lower than the data set B. RWE (ρ_1, ρ_2) values of 21.7-86.8 Hz sub-band for epileptic EEGs (data set E) are higher than the (ρ_3, ρ_4, ρ_5) values of the data sets C, and D and lower than that of the data sets A, and B. It can be, thus, concluded that epileptic signal have more rhythmicity than the normal and interictal EEGs. Based on the values of RWE as in Table 1, the following features are considered for

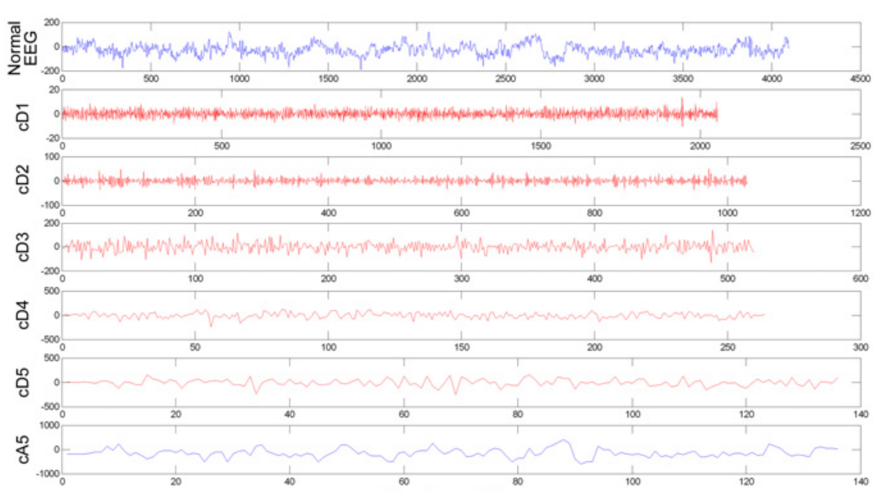

Fig. 4. Normal EEG (with eyes open) and its wavelet coefficients.

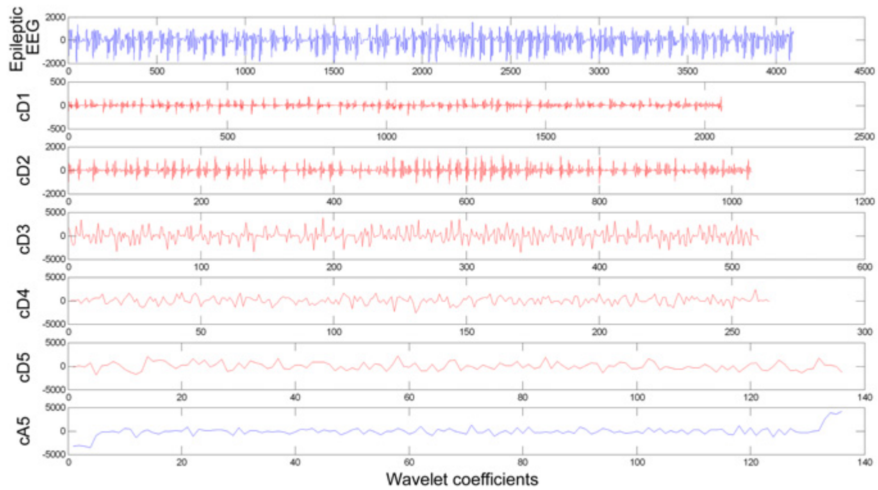

Fig. 5. Epileptic EEG and its wavelet coefficients.

Table 1. Average relative wavelet energy of different sub-bands.

RWE	Data Sets					Sub-band
	A	B	C	D	E	
ρ_1	0.000508	0.000469	0.000171	0.000149	0.000173	43.4–86.8
ρ_2	0.008810	0.007804	0.001454	0.001378	0.006409	21.7–43.4
ρ_3	0.058392	0.106870	0.012451	0.014129	0.079862	10.85–21.7
ρ_4	0.120826	0.250782	0.064089	0.088033	0.227105	5.43–10.85
ρ_5	0.145397	0.112384	0.197263	0.201341	0.365422	2.70–5.43
ρ_6	0.666068	0.521691	0.724571	0.694969	0.321029	0–2.70

Table 2. Average values of features for different data sets.

Data set	Features			
	x_1	x_2	x_3	x_4
A	0.333932451	0.33342492	0.666067549	0.961366222
B	0.478308702	0.477839624	0.5216913	1.103895635
C	0.275428614	0.275257464	0.724571386	0.758538
D	0.305030492	0.304881417	0.69497	0.80801
E	0.678971	0.678798	0.321029	1.148783637

classification;

$$x_1 = \rho_1 + \rho_2 + \rho_3 + \rho_4 + \rho_5$$
$$x_2 = \rho_2 + \rho_3 + \rho_4 + \rho_5$$
$$x_3 = \rho_6$$
$$x_4 = WEN \tag{11}$$

The average values of x_1, x_2, x_3 and x_4 features are given in Table 2. These features are used to differentiate among the normal/interictal and ictal EEGs using SVM, FFBPNN, k-NN and DT classifiers.

From Table 2, it is seen that the average values of x_1, x_2, and x_4 features for data set E are higher than the data sets A, B, C, and D. The average value of x_3 feature for the data set E is lower than the data sets A, B, C, and D. Therefore, these differences in average value features can also be utilized to form feature vectors and these feature vectors can be used to classify the EEGs.

Figs. 6a-b show that the x1values which are the signal energy distributions in the frequency band (2.7-86.8 Hz) for epileptic signals are higher than the x1 values of normal and interictal signals. These figures show significant difference between the ictal (epileptic) data based feature (x1-E), the normal subject data based features (x1-A and x1-B) and the

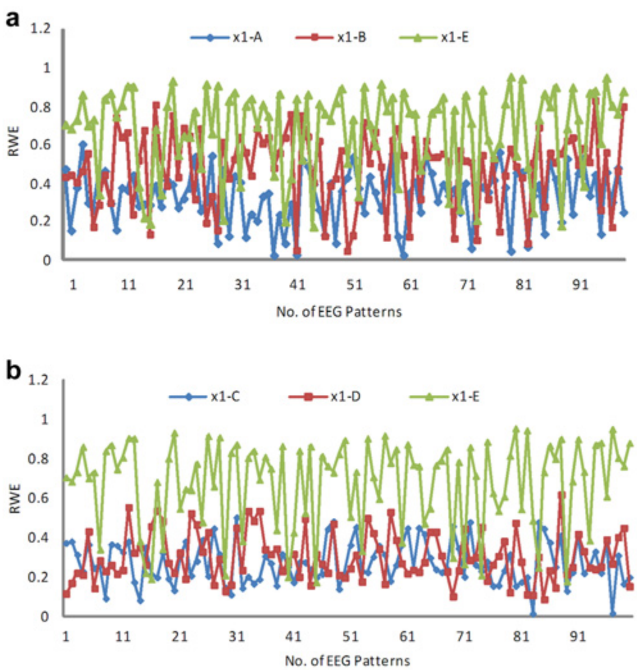

Fig. 6. (a) x1 values of data set A, B and E. (b) x1 value of data set C, D and E.

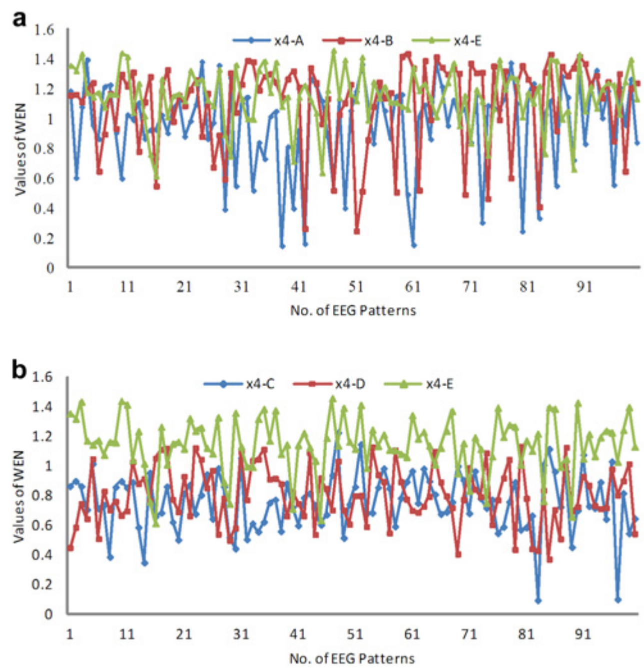

Fig. 8. (a) x4 values of data set A, B and E. (b) x4 values of data set C, D and E.

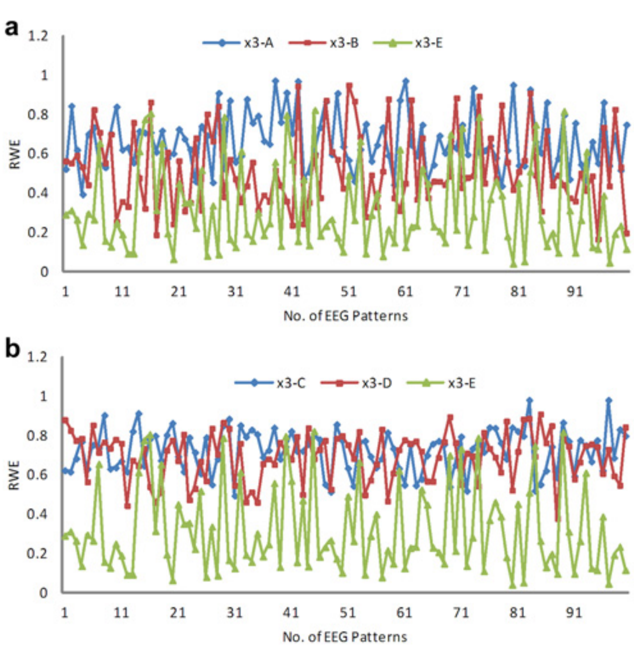

Fig. 7. (a) x3 values of data set A, B and E. (b) x3 values of data set C, D and E.

interictal data based features (x1-C and x1-D) for EEG classification. Figs. 7a-b show that the x3 values which are the signal energy distributions in a frequency band (0-2.7 Hz) for epileptic signals are lower than the values in case of normal and interictal signals. These figures also show significant difference between the x3 features of different data sets i.e. x3-A, x3-B, x3-C, x3-D and x3-E. Therefore, this feature can also be utilized in EEG classification. Feature (x_4) of data sets A, B, C and D versus E are plotted in Figs. 8a-b. These plots show that the most of the x_4 values of data set E are higher than the x_4 values of data set A, B, C and D values. This leads to conclude that the data set E is more ordered as compare to the data sets A, B, C and D. Significant overlap is found among the feature vectors of A. B, C, D and E data sets. This proves that the feature vectors are not linearly separable. The x_1, x_2, x_3 and x_4 feature vectors are applied to input of classifiers to classify the EEG.

The following cases have been taken up for study.

Case 1: Set A vs Set E
Case 2: Set B vs Set E
Case 3: Set C vs Set E
Case 4: Set D vs Set E
Case 5: Set A, C, D vs Set E
Case 6: Set A, B, C, D vs Set E

The SVM and FFBPNN are implemented by using MATLAB software version 7.8.0 (R2009a). The 60% data of the features are randomly selected for training and the rest of the 40% data of features for testing. The training features are used to train the SVM and FFBPNN, whereas the testing features are used to verify the accuracy and effectiveness of the trained SVM and FFBPNN for the given EEG classification problem. The 120, 240 and 300 different training patterns of features are randomly generated for cases 1-4, case-5 and case-6, respectively. The architecture of the

Table 3. Statistical parameters.

Cases	SVM (L)			SVM(rbf)			FFBPNN			k-NN			Decision Tree		
	%age SEN	%age SPE	%age CA	%age SEN	%age SPE	%age CA	%age SEN	%age SPE	%age CA	%age SEN	%age SPE	%age CA	%age SEN	%age SPE	%age CA
Case-1(A-E)	97.50	92	95.00	100	92.5	96.25	94.5	95.20	95.00	94.87	92.68	93.75	92.30	90.24	91.25
Case-2(B-E)	87.50	75	81.25	80	80	80.00	73.6	90.4	82.00	84.61	82.92	83.75	81.57	78.57	80.00
Case-3(C-E)	97.50	90	93.75	97.5	97.5	97.50	94.73	100	97.50	93.02	100	96.25	92.5	92.5	92.50
Case-4(D-E)	87.50	87.5	87.50	97.5	95	96.25	94.73	100	97.50	90.69	97.29	93.75	92.5	92.5	92.50
Case-5(ACD-E)	99.17	80	94.37	97.5	92.5	95.00	78.94	100	95.00	93.75	100	95.00	92.30	88.05	88.75
Case-6(ABCD-E)	99.38	62.5	93.00	94.37	77.5	91.00	63.15	100	93.00	93.41	87.87	92.50	82.15	90.11	89.00

FFBPNN used in this work is: 4 neurons in the input layer, 6 neurons in the hidden layer, and 2 neurons in the output layer. The performance of the proposed method is calculated using three statistical parameters which are tabulated in Table 3 for cases 1-6.

For the SVM training and classifying, the svmtrain and svmclassify functions from the MATLAB Bioinformatics Toolbox have been used. The SVM algorithm is used with Linear and Gaussian Radial basis kernel functions. Regularization parameter (box constraints) for linear kernel functions are set to its default value C =1. The SVM algorithm parameters with Gaussian Radial Basis kernel function (rbf) are set with its default value of scaling factor sigma of 1 and box constraints value C=10. Sixty percent data of the features are randomly selected for training and the rest of the 40% data of features for testing by SVM. This process is repeated for fifteen times and the maximum performance parameters are recorded and summarized in Table 3.

In this study the k-Nearest Neighbor (k-NN) is used as classifier and knnclassify function of MATLAB Bioinformatics Toolbox is used to classify the features data. The number of nearest neighbors' parameter (k) is varied from 1 to 10 for reducing the computation cost and getting the best performance parameters. The Euclidean distance is used to measure the similarity between two objects as another parameter. Majority rule with nearest point tie-break are used to decide how to classify the samples. The maximum statistical parameters using k-NN for cases (1-6) are summarized in Table 3.

The decision tree is implemented by using Statistical Toolbox of MATLAB software version 7.8.0 (R2009a). The functions treefit and treeval with default parameters are used to create the decision tree and predicting the response value. The maximum statistical parameters obtained using these functions for cases (1-6) are summarized in Table 3. The present work has mainly two disadvantages. First, the feature vector order chosen for this study is complex and quite difficult to estimate. Second, If the feature vectors are linearly separable than the linear classifiers will give better results using SVM (linear). If the feature vectors are not linearly separable or overlapped than the SVM (rbf) classifiers give better classification results and the linear classifiers fail for such features. Hence, the computational complexity is dependent on the nature of features vectors. These two factors increase the computational complexity of the method.

A comparison of classification accuracy obtained by the proposed method and by the other methods for EEG classification problems [12, 14, 18, 19, 25, and 27-37] is presented in Table 4. It represents the comparison between results obtained from the proposed method and the existing methods. Only the methods which have used the same data sets and same cases are included for the comparison. The classification accuracy of different cases 1-6 obtained using the proposed method is better than the reported results of some other researchers.

For case 1, the classification accuracy (96.25%) obtained using the present method is better than the results reported by different researchers [12, 14, 27, and 29] for the same data set, which were obtained using different methods like Subasi [27] works, in which discrete wavelet transform and mixture expert model are used to evaluate the classification accuracy.

For case 2, case 3 and case 4, the classification accuracies obtained using the present method are 83.75%, 97.5%, and 97.5% respectively, which are the best presented. As per the knowledge gathered, only N. Nicolaou [14] has raised these cases and reported 82.8%, 88.0% and 79.94% accuracy, respectively, which was obtained employing permutation entropy with SVM using the same data set.

For case 5, the classification accuracy obtained using the proposed method is 95.00 %. This case is also presented in Lingo work [34], the line length feature reflects the waveform dimensionality changes and is a measure sensitive to variation of the signal amplitude and frequency. Lingo reported 97.75% accuracy using the line length features based on wavelet transform multi-resolution decomposition and combines with an ANN to classify the EEG.

For case 6, the classification result obtained in our work is 93.0 %, and is less than the accuracy demonstrated in Tzallas [31] work, in which the authors employed the energy

Table 4. Comparison of classification accuracy.

Researchers	Year	Methods	Problem	CA (%)
Nigam and Graupe	2004	Non-linear preprocessing filter -diagnostic neural network	A-E	97.20
Srinivasan et al.	2005	Time and frequency- domain features -recurrent neural network	A-E	99.60
Kannathal et al.	2005	Entropy measures- adaptive neuro-fuzzy inference system	A-E	92.22
Polat and Günes¸	2007	Fast Fourier transform-decision tree	A-E	98.72
Subasi	2007	Discrete wavelet transform-mixture of expert model	A-E	95.00
Tzallas et al.	2007	Time–frequency analysis-Artificial neural network	A-E ABCD-E	100 97.73
Guo et al.	2009	Discrete wavelet transform-relative wavelet energy-MLPNN	A-E	95.20
Subasi et al.	2010	DWT- PCA, ICA, LDA and SVM	A-E	98.75(PCA) 99.50(ICA) 100(LDA)
Ling Guo et al.	2010	Multiwavelet transform-approximate entropy feature-MLPNN	A-E ABCD-E	99.85 98.27
Ling Guo et al.	2010	line length feature -ANN	A-E ACD-E ABCD-E	99.60 97.75 97.77
Elif Derya Ubeyli	2010	LS-SVM model-based methods coefficients	A-E	99.56
Umut orhan et al.	2011	EEG classification -K-means clustering-MLPNN	A-E ABCD-E	100 99.60
Ling Guo et al.	2011	GP-based feature extraction-KNN classifier	A-E	99.20
Zafer Iscan et al.	2011	Classification of EEG signal with combined time and frequency features	A-E	100
Deng Wang et al.	2011	Wavelet packet entropy- hierarchical EEG classification	A-E	99.44
Nicolaou et al.	2012	Permutation Entropy -SVM	A-E B-E C-E D-E	93.55 82.88 88.00 79.94
Present reporting	2012	RWE and WEN Based EEG Classification	A-E B-E C-E D-E ACD-E ABCD-E	**96.25** **83.75** **97.50** **97.50** **95.00** **93.00**

distribution features extracted from the time frequency plane to an ANN for classifying EEGs with 97.73% classification accuracy. Classification accuracies 99.73% is obtained by lingo et al. [34], which uses the line length features which are used as input to SVM.

CONCLUSIONS

The relative wavelet energy (RWE) and wavelet entropy (WEN) are calculated from the wavelet coefficients which are obtained after DWT decomposition of EEG data sets. The average value of wavelet entropy for data set E are more as compared to the data set (A, B, C, and D), which prove that the data set E is more ordered than the other data sets. Relative wavelet energy provides information about the relative energy associated with different frequency bands of the EEG signals. In this work, four features are extracted from the RWE and WEN to classify the EEG signals using SVM, FFBPNN, k-NN and DT. The proposed method gives the classification accuracy for case-1, i.e. normal subject (data set A) versus ictal set (data set E) which is 96.25% using SVM. The k-NN classifier classifies the case -2 with the best classification accuracy (83.75%). FFBPNN gives 97.5% classification accuracy for both, case-3 and case-4. These are the best reported results for the classification of interictal data versus ictal data. These accuracies demonstrate the improved performance of the proposed method. These findings indicate that the present method can be employed as a quantitative measure for monitoring the EEG and it may prove to be a useful tool in analyzing the EEG signal associated with epilepsy.

REFERENCES

[1] Ocak H. Automatic detection of epileptic seizures in EEG using discrete wavelet transform and approximate entropy. Expert Syst Appl. 2009; 36(2):2027-36.
[2] Tzallas AT, Tsipouras MG, Fortiadis DI. Epileptic seizure detection in EEG using time-frequency analysis. IEEE T Inform

Technol Biomed. 2009; 13(5):703-10.
[3] Adeli H, Dastidar SG, DadmehN. A wavelet-chaos methodology for analysis of EEGs and EEG sub-bands to detect seizure and epilepsy. IEEE T Bio-med Eng. 2007; 54(2):205-11.
[4] Mallat S. A theory for multi-resolution signal decomposition: the wavelet representation. IEEE T Pattern Anal. 1989; 11(7):674-93.
[5] Andrzejak RG, Lehnertz K, Rieke C. Indications of nonlinear deterministic and finite dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state. Phys Rev E. 2001; 64(6): 061907.
[6] Andrzejak RG, Widman G, Lehnertz K. The epileptic process as nonlinear deterministic dynamics in a stochastic environment: An evaluation on mesial temporal lobe epilepsy. Epilepsy Res. 2001; 44:129-40.
[7] Rosso OA, Martin MT, Plastino A. Brain electrical activity analysis using wavelet-based informational tools (II): Tsallis non- extensivity and complexity measures. Physica A. 2003; 320:497-11.
[8] Rosso OA, Blanco S, Yordanova J, Kolev V, Figliola A, Schurmann M, Basar E. Wavelet entropy: a new tool for analysis of short duration brain electrical signals. J Neurosci Meth. 2001; 105:65.
[9] Rosso OA, Martin MT, Figliola A, Keller K, Plastino A. EEG analysis using wavelet based information tools. J Neurosci Meth. 2006; 153:163-82.
[10] Mirzaei A, Ayatollahi A, Gifani P, Salehi L. EEG analysis based on wavelet spectral entropy for epileptic seizures detection. Conf Proc Biomed Eng Inform. 2010; 878-82.
[11] Al-Nashash HA, Paul JS, Thakor NV. Wavelet entropy method for EEG analysis: application to global brain injury. Conf Proc IEEE EMBS Neural Eng. 2003; 348-51.
[12] Guo L, Rivero D, Seoane J, Pazos A. Classification of EEG signals using relative wavelet energy and artificial neural networks. Proc first ACM/SIGEVO Summit Genet Evol Comput (GEC). 2009; 177-84.
[13] Kumar SP, Sriraam N, Benakop, PG, Jinaga BC. Entropies based detection of epileptic seizures with artificial neural network classifiers. J Expert Syst Appl. 2010; 37:3284-91.
[14] Nicolaou, N, Georgiou J. Detection of epileptic electroencephalogram based on permutation entropy and support vector machine. Expert Syst Appl. 2012; 39:202-9.
[15] Hsu KC, Yu SN. Detection of seizures in EEG using subband nonlinear parameters and genetic algorithm. Comput Biol Med. 2010; 40:823-30.
[16] Acharya UR, Molinari F, Sree SV, Chattopadhyay S. Automatic diagnosis of epileptic EEG using entropies. Biomed Signal Process Control. 2011; doi:10.1016/j.bpsc.2001.07.007.
[17] Ubeyli ED. Least square support vector machine employing model-based methods coefficients for analysis of EEG signals. Expert Syst Appl. 2010; 37:233-9.
[18] Iscan Z, Dokur Z, Demiralap T. Classification of electroencephalogram signals with combined time and frequency features. Expert Syst Appl. 2011; 38:10499-505.
[19] Foo SY, Stuart G, Harvey B, Baese AM. Neural network-based EKG pattern recognition. Eng Appl Artif Intel. 2002; 15:253-60.
[20] Kiymik MK, Akin M, Subasi A. Automatic recognition of alertness level by using wavelet transform and artificial neural network. J Neurosci Meth. 2004; 139:231-40.
[21] Kiymik MK, Subasi A, Ozcalik HR. Neural networks with periodogram and autoregressive spectral analysis methods in detection of epileptic seizure. J Med Syst. 2004; 28(6):511-22.
[22] Petrosian, A, Prokhorov D, Homan R, Dashei R, Wunsch D. Recurrent neural network based prediction of epileptic seizures in intra and extracranial EEG. Neurocomputing. 2000; 30:201-18.
[23] Subasi A. Automatic detection of epileptic seizure using dynamic fuzzy neural networks. Expert Syst Appl. 2006; 31:320-8.
[24] Kalayci T, Ozdamar O. Wavelet preprocessing for automated neural network detection of EEG spikes. IEEE Eng Med Biol. 1995; 14(2):160-6.
[25] Nigam V, Graupe D. A neural-network-based detection of epilepsy. Neurosci Res. 2004; 26(1):55-60.
[26] Gandhi T, Panigrahi BK, Anand S. A comparative study of wavelet families for EEG signals classification. Neurocomputing. 2011; 74:3051-7.
[27] Subasi A. EEG signal classification using wavelet feature extraction and a mixture of expert model. Expert Syst Appl. 2007; 32(4):1084-93.
[28] Srinivasan V, Eswaran, C, Sriraam, N. Artificial neural network based epileptic detection using time-domain and frequency-domain features. J Med Syst. 2005; 29(6):647-60.
[29] Kannathal N, Choo ML, Acharya UR, Sadasivan PK. Entropies for detection of epilepsy in EEG. Comput Meth Prog Bio. 2005; 80:187-94
[30] Polat K, Günes S. Classification of epileptic form EEG using a hybrid system based on decision tree classifier and fast Fourier transform. Appl Math Comput. 2007; 187(2):1017-26.
[31] Tzallas AT, Tsipouras MG, Fotiadis DI. Automatic seizure detection based on time-frequency analysis and artificial neural networks. Comput Intell Neurosci. 2007; doi:10.1155/2007/80510.
[32] Guo L, Riveero D, Pazaos A. Epileptic seizure detection using multiwavelet transform based approximate entropy and artificial neural networks. J Neurosci Meth. 2010; 193:156-63.
[33] Subasi A, Gursoy MI. EEG signal classification using PCA, ICA, LDA and support vector machine. Expert Syst Appl. 2010; 37:8659-66.
[34] Guo L, Rivero D, Dorado J, Rabunal, JR, Pazos A. Automatic epileptic seizure detection in EEG based on line length feature and artificial neural network. J Neurosci Meth. 2010; 191:101-9.
[35] Orhan U, Hekim M, Ozer M. EEG signals classification using the K means clustering and a multilayer perceptron neural network model. Expert Syst Appl. 2011; 38:13475-81.
[36] Guo L, Rivero D, Dorado J, Munteanu CR, Pazos A. Automatic feature extraction using genetic programming: An application to epileptic EEG classification. 2011; 38:10425-36.
[37] Wang D, Miao D, Xie C. Best basis-based wavelet packet entropy feature extraction and hierarchical EEG classification for epileptic detection. Expert Syst Appl. 2011; 38: 14314-20.
[38] K Temel, A Onder. A polynomial fitting and k-NN based approach for improving classification of motor imagery BCI data. Pattern Recogn Lett. 2010; 31:1207-15.
[39] L Rokach, O Maimon. Decision trees: Data mining and knowledge discovery handbook in Springer Science, Business media, Inc, 2005, pp. 165-92. http://www.ise.bgu.ac.il/faculty/liorr/hbchap9.pdf