Asian Journal of Information Technology 13 (10): 639-645, 2014
ISSN: 1682-3915

An Improved Detection of Epilepsy EEG Using Multi Wavelet Transform and Modified Effective Extreme Learning Machine

A.S. Muthanantha Murugavel and S. Ramakrishnan
Department of Information Technology,
Dr. Mahalingam College of Engineering and Technology,
Pollachi, Tamil Nadu, India

Abstract: Epilepsy is the most widely recognized chronic neurological diseases and the most well-known neurological chronic disease of childhood. The Electroencephalogram (EEG) signal furnishes significant information neurologists contemplate in the investigation and analysis of epileptic seizures. EEG being the non-stationary signals, legitimate analysis is greatly needed to classify the normal, inter-ictal and ictal EEGs. This srudy presents a propose method is to classify the EEG signals regarding the existence or absence of seizure. Modified EELM is more accurate for automatic epileptic seizure detection. Multi wavelet transform contains both scaling and wavelet functions, simultaneously offers orthogonality and symmetry which is not possible for scalar wavelet transform. With these major properties, the multi wavelet transform is promising in signal processing applications. Multi wavelet based characteristic is utilized to differentiate the normal EEGs of healthy subjects, inter-ictal and ictal EEGs of epileptic patients. To acquire the detail and approximation wavelet coefficients, the EEG signals are decomposed into sub-groups. The proposed strategy decomposes with wavelet transform, reduce dimensionality by a set of features and have a classification with Modified Effective Extreme Learning Machine (Modified-EELM). The modified-EELM has a better accuracy than ELM. The Modified-EELM Classification algorithm with the sigmoid function has 98.01% testing accuracy, good performance, easy to implementing and consumes only 0.0008 sec of time. This is very less amount of time compared with other learning machines.

Key words: Multi wavelet transform, extreme learning machine, modified effective extreme learning machine, epilectic seizures, patients

INTRODUCTION

In epileptic patients, the nerve cells in the brain convey over the top electrical impulses that cause episodes called seizures. An epileptic seizure is a transient event of signs or symptoms because of irregular or synchronous neuronal action in the cerebrum. Epilepsy is a disorder of the brain described by a persisting predisposition to generate epileptic seizures. An epileptic seizure (Fisher *et al.*, 2005) is characterized as a transient symptom of synchronous neuronal activity in the brain. Epilepsy is depicted by more unwarranted seizures. Seizures might be central or generalized. In case of central seizures, just a specific segment of the brain is affected while the whole brain is affected in the case of generalized seizures. Epilepsy is assessed by utilizing the Electroencephalogram (EEG) test in which electrodes are set on the affected area of the brain and brain signals are recorded and examined (Ruiz *et al.*, 2011; Coyle *et al.*, 2010; Ince *et al.*, 2009). Reviewing of whole recorded EEG signals by a trained proficient is time-consuming. Research on seizure detection exhibited in different techniques. The Short Time Fourier Transform (STFT) analysis of EEG signals and its extracted features depend upon the pseudo wigner-file distribution (Mohseni *et al.*, 2006). So, those features were utilized as inputs to an Artificial Neural Network (ANN) for classifications.

The wavelet transform used to capture features of EEG signals and after that joined together with ANN to get a classification result (Kalayci and Ozdamar, 1995). It portrayed a method for automated detection of epileptic seizures from EEG signals utilizing a multistage non-linear pre-processing filter for concentrating on relative spike amplitude and spike occurrence frequency (Nigam and Graupe, 2004). The EEGs were decomposed with wavelet transform into distinctive sub-groups and some statistical information extracted from the wavelet coefficients (Jahankhani *et al.*, 2006). The two popular and successful EEG features, Root Mean Square (RMS) and mean average value (Oskoei and Hu, 2007; Naik *et al.*, 2010) as per representative features. Multi-Wavelet Transform (MWT) is a hot research topic in the wavelet field which is from an extension of scalar wavelet transform. MWT is not only maintaining the better time and frequency

Corresponding Author: A.S. Muthanantha Murugavel, Department of Information Technology, Dr. Mahalingam College of Engineering and Technology, Pollachi, Tamil Nadu, India

domain localization properties which scalar wavelet possesses but also overcomes the scalar wavelet shortcomings. MWT is orthogonality, symmetry and smoothness are the main desired properties of the MWT which are most useful for both signal analysis and processing. To improve the compression ratio, MWT can transfer high-frequency energy to the low-frequency energy. MWT provides perfect reconstruction and good performance at the boundaries. With these properties, MWT has a vital role in signal processing applications. There have been different classifiers, both for linear and nonlinear. Artificial Neural Network (ANN), Support Vector Machine (SVM), Linear Discriminant Analysis (LDA), hidden markov classifier and Z-scale base Discriminant Analysis (ZDA). The main weakness of these techniques is that they do not work well for nonlinear problems. The present srudy falls into the classification of ANN. It is realized that ANN methodologies are usually difficult in manual tuning control parameters. Back Propagation Neural Networks (BPNNs) and Support Vector Machines (SVMs) have been demonstrated to enhance the classification accuracy compared with linear techniques. Support Vector Machines (SVM) are the most Succeeded and Utilized algorithms to model the EEG characteristics. However, execution time for SVMs training process is more. To accelerate the training process without compromising, a new Classification algorithm referred to as Extreme Learning Machine (ELM) algorithm. The reason is that ELM (Huang *et al*., 2006) arbitrarily chooses the input weights and bias of the SLFNs as opposed to tuning. ELM is proposed by Huang which is a Single-hidden-Layer Feed-forward Network (SLFN) and which could be utilized within regression and classification issues.

The Modified Effective Extreme Learning Machine (Modified-EELM) is a straightforward and effective algorithm for SLFNs to conquer the short comings of ELM. There are a few interesting characteristics of the Modified-EELM algorithm in comparison with the ELM algorithm. The Modified-EELM is faster than ELM. The main difference between Modified-EELM and ELM lie in the determination of input weights and biases. The ELM algorithm selects them randomly which consumes little time. The Modified-EELM algorithm chooses the input weights and bias appropriately which consumes less time compared with the training time of output weights. Also, the Modified-EELM algorithm by making legitimate selection of input weights and bias of the networks neglects the risk of yielding singular or not full column rank hidden layer output network H. It considers utilization of a faster method which can calculate the Moore-penrose's inverse of H. And also, the great characteristic of the Modified-EELM is that it has a longer prediction term much accuracy than the ELM algorithm. Additionally, Modified-EELM is better robust than ELM. However, modified-EELM remains steady and has a great performance. The proposed method concentrated on multi wavelet transform to generate sub band coefficients, feature extraction is to reduce the dimensionality and modified-EELM is proposed for classifying the EEG data into normal/non-seizure and epileptic. At first, the EEG data is decomposed by MWT to generate sub-band signals and then the feature is extracted for each sub-band signal to form a feature vector. And finally, the modified-EELM is to classify the EEG into normal/non-seizure and seizure. The block diagram of the proposed approach is shown in Fig. 1.

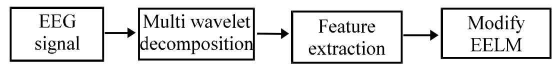

Fig. 1: Block diagram of proposed method

MATERIALS AND METHODS

EEG data collection: International 10-20 system is a globally distinguished method to describe and apply the electrodes in proper location to scalp for an EEG test. It is dependent upon the relationship between the electrode location and the underlying area of cerebral cortex, shown in Fig. 2. The numbers 10 and 20 shows the distance between adjacent electrodes are either 10 or 20% of total front-back or right-left separation of the skull.

Each site has a letter to recognize the lobe and a number to distinguish the hemisphere area. A zero (z) refers to an electrode put on the midline. Even numbers (2, 4, 6, 8) refers to electrode positions on the right hemisphere and odd numbers (1, 3, 5, 7) refers to those on the left hemisphere. The letter codes A, Pg and Fp recognizes the earlobes, nasopharyngeal and frontal polar sites separately. The letters F stands for frontal, T stands for fleeting, C stands for focal, P stands for parietal and O stand for occipital projections. But, there exists no central lobe; the "C" letter is utilized for distinguishing purposes. Two anatomical landmarks are utilized for the positioning of EEG electrodes:

- The Nasion which is depressed area between eyes, means just above the bridge of the nose
- The Inion which is the skull lowest point from the back of the head

The point when recording a more EEG details with additional electrodes, extra electrodes are included using 10% division which fills in halfway between those of the existing 10-20 system. The Modified Combinatorial Nomenclature (MCN) is a new electrode naming system.

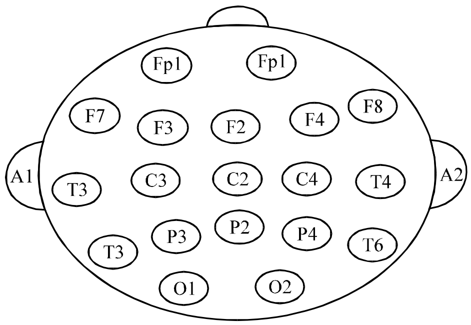

Fig. 2: EEG electrodes placement

MCN System utilizes 1, 3, 5, 7, 9 for the left hemisphere which represent that 10, 20, 30 and 40% of Inion to Nasion distance separately. These additional letter codes permits the naming of intermediate electrode sites. The patient's EEG recordings from the Freiburg database held between 2 and 6 seizures and 50 min of pre-ictal information for most seizures and approximately 24 h of EEG-recordings without seizure action and spreading over the full wake-sleep cycle. Depending upon patient's total number of seizures, it can be separated pre-ictal samples before the last 1 or 2 seizures and for testing data 33% of the inter-ictal samples are required. The remaining samples are training data. The publicly available EEG information is recorded by method of a 128-channel 12 bit EEG system with 173.5 samples/sec and total 500 segments are grouped into five sets (A-E). The duration for each segment is 23.6 sec. All sets are recorded from EEG after purifying artifacts caused by eye and muscle movements.

Multi wavelet transform: Multi wavelets are very similar to wavelets. But, it has some important differences. Specifically, multi wavelets have two or more scaling and wavelet functions. The set of scaling functions can be written using the vector notation $\phi(t) = [\phi_1(t), \phi_2(t), ..., \phi_r(t)]^T$, where $\phi(t)$ is called the multi-scaling function. And the multi wavelet function is defined from the set of wavelet functions as $\psi(t) = [\psi_1(t), \psi_2(t), ..., \psi_r(t)]^T$. While r can be arbitrarily large, the multi wavelets studied to date are primarily for r = 2. When r = 1; $\psi(t)$ is a scalar wavelet. The multi wavelet two-scale equations for scalar wavelets:

$$\phi(t)=\sqrt{2}\sum_{k=-\infty}^{\infty} L_k\phi(2t-k) \tag{1}$$

$$\psi(t)=\sqrt{2}\sum_{k=-\infty}^{\infty} H_k\phi(2t-k) \tag{2}$$

Here, $\{L_k\}$ and $\{H_k\}$ are the matrix filters, i.e., $\{L_k\}$ and $\{H_k\}$ are r×r matrices for every integer value k. The matrix elements in these filters give more degrees of freedom than a traditional scalar wavelet. Similar to the scalar wavelet transform, the Multi Wavelet Transform (MWT) is based on the idea of multi resolution analysis. During a single level decomposition using scalar wavelet transform, the input data is decomposed into two sub-bands which represent the output of low-pass and high-pass filters individually. MWT (Martin and Bell, 2001; Jafari-Khouzani and Soltanian-Zadeh, 2003) decomposition produces two low-pass sub-bands and two high-pass sub-bands for every level of decomposition. In MWT, each multi filter bank contains two channels, so that there will be two scaling coefficients and two wavelet coefficients for each level decomposition. Since, multiple iterations over the low-pass information is desired, the scaling coefficients for the two channels are stored together. Similarly, the wavelet coefficients for the two channels are stored together. Successive iterations are performed on the two sets of scaling coefficients from the previous stage. The MWT Decomposition algorithm is written as:

$$C_{j-1,k} = \sum_n L_n C_{j,2k+n} \tag{3}$$

$$d_{j-1,k} = \sum_n H_n d_{j,2k+n} \tag{4}$$

The relations 3 and 4 are illustrated by multi wavelet filter bank. The low-pass filter and high-pass filter consist of coefficients corresponding to Eq. 1 and 2 and $[c_{1,0,k}, c_{2,0,k}]^T$ is the input signal vector which is the initial expansion coefficients of the given multi wavelet system. After the input signal vector is filtered through L and H matrix filters and also four output streams are generated. So, each of streams is then down sampled by a factor of two. Then, $[c_{1,-1,k}, c_{2,-1,k}]^T$ represents the low-pass sub-band signal vector which is a multi scaling coefficient of the input signal vector. And also, $[d_{1,-1,k}, d_{2,-1,k}]^T$ represents the high-pass sub-band signal vector which is the multi wavelet coefficient of the input signal vector. For next decomposition, the same method is repeated on the first level multi scaling coefficients vector $[c_{1,-1,k}, c_{2,-1,k}]^T$ to get the second level multi scaling and multi wavelet coefficients. The MWT analysis is decomposed signals into frequency bands. The appropriate wavelet and the number of decomposition levels are very important in the analysis of EEG signals through multi wavelet transform. The EEG frequency bands cover the 0-60 Hz frequency scope. The original EEG signals were then subjected to a level four decomposition utilizing order 4 wavelet transform. After the first level of decomposition, the EEG signal has decomposed into higher resolution d1 and lower resolution components a1. MWT's each sub-signal represents the original signal information in different frequency sub-bands since it reaches to good results in the classification of EEG signals.

Feature extraction: The multi wavelet coefficients provide a compact representation that shows the energy distribution of the signal in time and frequency. It is expected that the coefficients of the seizure frequency spectrum ranges from 0.5-30 Hz. The wavelet coefficient subsets (cd1-cd4, ca4) and the reconstructed EMG signals (D1-D4, A4) were extracted from their features. The popular and successful feature is MAV for dimensionality reduction. Equation 5 shows the definition of MAV feature is defined as:

$$MAV = \frac{1}{N}\sum_{n=1}^{N}|x_n| \tag{5}$$

where, x_n represents the nth sample of the EMG signal (S) or the wavelet coefficients subsets (cd1-cd4, cA4) or the reconstructed EMG signals (D1-D4, A4) in a window segment and N denotes the length of EMG signal window segment (N = 256). The comparison of class separability in each type is discussed to find the suitable EMG subset. For feature extraction, the MWT coefficients provide a compact representation that shows the energy distribution of EEG signal in both time domain and frequency domain. Therefore, the computed detail and approximation coefficients of the EEG signals were used as feature vectors to represent the signals. The most significant step in EEG signal recognition to reduce the dimensionality, feature extraction is a highly effective approach. The recognition performance depends on the features which are used to characterize the primary EEG signals. In the characteristic extraction stage, EEG time series signals are analyzed by inspecting frequency bands related with different conscious states or mental activities.

Modified Effective Extreme Learning Machine algorithm: Extreme Learning Machine (ELM) is a proficient training algorithm to manage Single-hidden-Layer Feed forward neural Networks (SLFNs) by Huang *et al*. (2006). ELM randomly generates the input weights and the bias of hidden nodes while analytically decides the output weights of SFLN.

From Fig. 3, a standard Single-hidden Feed-forward Neural network (SLFN) with N hidden nodes could be represented as mathematical description of SLFN and activation function g can be defined as:

$$G_N(X) = \sum_{i=1}^{N}\beta_i g(w_i.X + b_i) \tag{6}$$

here, the input vector is $X = (x_1, x_2, x_3, ..., x_d)^T \in \mathbb{R}^d$, $W = (w_{i1}, w_{i2}, w_{i3}, ..., w_{id})$ is the input weight vector connecting the ith hidden nodes and the input nodes b_i is

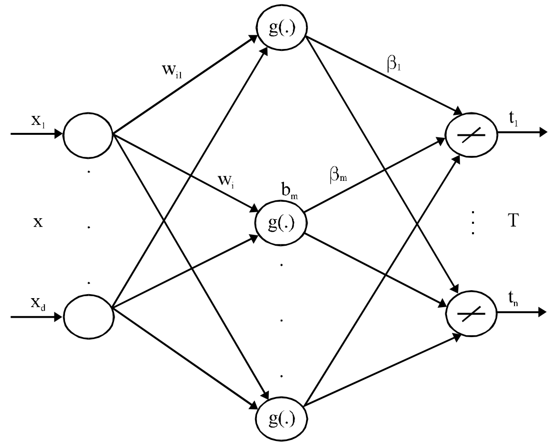

Fig. 3: Simple ELM architecture

the bias of ith hidden node and the output weight vector connecting with output node is $\beta = (\beta_1, \beta_2, \beta_3, ..., \beta_N)^T$. An SLFN with N hidden-layer nodes approximating both input vector x and corresponding observation value ti have n training data with zero error means for set of distinct training data $\{(X_i^*, t_i^*)\}_{i=1}^{n} \subset \mathbb{R}^d \times \mathbb{R}$ such that:

$$\sum_{i=1}^{N}\beta_i g(w_i.X + b_i) = t_i \quad j = 1,2,...,n \tag{7}$$

Equation 7 can be written compactly as:

$$H\beta = T \tag{8}$$

By Huang *et al*. (2006), H is called the hidden layer output matrix of the neural network concept. Where:

$$H = \begin{bmatrix} g(W_1.X_1 + b_1) & \cdots, & g(W_N.X_1 + b_N) \\ \vdots & \ddots & \vdots \\ g(W_1.X_n + b_1) & \cdots & g(W_N.X_n + b_N) \end{bmatrix}_{n\times N} \tag{9}$$

$$\beta = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_N \end{bmatrix}_{N\times 1} \text{ and } T = \begin{bmatrix} t_1 \\ \vdots \\ t_n \end{bmatrix}_{n\times 1} \tag{10}$$

ELM designed as SLFN with N hidden neurons can learn N distinct samples with zero error. According to Huang's *et al*. (2004) ELM, the hidden layer parameters w_i and b_i of SLFN need not to be tuned during the process of learning and can be allocated with random values. Then, Eq. 7 becomes a linear system and the output weight can be assessed by:

$$\hat{\beta} = H^{\dagger}T \tag{11}$$

Here, $H^{\dagger}$ is the Moore-Penrose generalized inverse of the hidden layer output matrix H. ELM randomly chooses the input weights w_i and bias b_i of the SLFNs oppose to tuning. Thus, ELM is an Extremely Fast Learning algorithm. But this major advantage makes the algorithm less effective. The random choice of input weights and bias produce an unexpected result. And the result is that the hidden layer output matrix H is not full column rank or singular by following two difficulties:

- It develops training error of the samples which to some extent lowers the prediction accuracy
- ELM cannot utilize the orthogonal projection technique to calculate Moore-Penrose generalized inverse of H due to the singularity of H. To calculate the Moore-Penrose generalized inverse, it prefers Singular Value Decomposition (SVD) which costs more time and lower the effectiveness of ELM

To improve the learning rate and robustness of the networks, the proposed algorithm Modified-EELM makes a legitimate selection of the input weights and bias before calculating the output weights which decides the full column rank of H. The difference between Modified-EELM and ELM algorithm lies in the training of input weights and bias. And time spent in the first phase of Modified-EELM is very short compared to the second step. So, modified-EELM is actually faster than ELM. And Modified-EELM algorithm also possesses the qualities of the ELM algorithm including easy implementing, good generalization performance. Moreover, the new algorithm improved the effectiveness of learning: the full column rank property of the matrix makes the Orthogonal Projection Method, a fast algorithm for solving generalized inverse, available. So, it is called effective extreme learning machine.

The Modified-EELM is a simple and effective algorithm for SLFNs in an attempt to overcome the short comings of ELM. There are several interesting features of the Modified-EELM algorithm in comparison with the ELM algorithm:

- The learning speed of Modified-EELM is generally faster than ELM. The main difference between Modified-EELM and ELM lie in the selection of input weights and biases. The ELM algorithm chooses them randomly which consumes little time. The Modified-EELM algorithm selects the input weights and bias properly which also consumes short time compared with the training time of output weights
- The Modified-EELM algorithm by making proper selection of input weights and bias of the networks avoids the risk of yielding singular or not full column rank hidden layer output matrix H. This allows for use of a faster method which can calculate the Moore-Penrose generalized inverse of H much more rapidly
- An impressive feature of the Modified-EELM is that it has a longer prediction term with acceptable accuracy than the ELM algorithm. Also, Modified-EELM has better robustness property than ELM. In particular, in the regression, the performance of ELM is sometimes poor. But Modified-EELM remains steady and has a good performance

Proposed algorithm: Set of distinct training data set, $N=\{(X_i^*, t_i^*)\}_{i=1}^{n}$ where X is input vector and t is corresponding observation value. For input-output relationships of biological neurons, choose activation function of radial basis function $g(x) = 1/1+\exp(-\rho x)$, where ρ is a parameter that determines its slope. This function has an S-shape that introduces a nonlinear property in the network. The S-shape gives the functions the property that they act linearly around the origin while large inputs are squeezed into the interval zero to one. It is the function that endows the model with non-linear behavior and biological plausibility. However, this form assumes that the variance of the states is fixed because the sigmoid encodes the density on neuronal states. And also choose the hidden node number n_0.

Suppose there are n samples X, t, where $X = (x_1, x_2, x_3, ..., x_d)T\in\mathbb{R}^d$ is the input vector, $\beta = (\beta_1, \beta_2, \beta_3, ..., \beta_N)^T$ is the output weight vector connecting the hidden nodes with output node, $W = (w_{i1}, w_{i2}, w_{i3}, ..., w_{id})$ is the input weight connecting with the ith hidden nodes and the input nodes and is the bias of the ith hidden node.

Step1: Denote the sorted data by (X_i, t_i) $(i = 1, 2, 3, ..., n)$ and $X_i = (x_{i1}, x_{i2}, x_{i3}, ..., x_{id})$ $(i = 1, 2, ..., n)$. Select weights w_i and bias $b_i (i = 1, 2, 3, ..., n_0)$. Sort $w.x_1^*, w.x_2^*, ..., w.x_{n_0}^*$ in order that $w.x_{i_1}^* < w.x_{i_2}^* < ... w.x_{i_{n_0}}^*$ $(i_j \neq i_k$ for $j \neq k$, $j, k = 1, 2, 3, ..., n_0$ and $i_j = 1, 2, 3, ..., n_0)$ are satisfied and then correspondingly change the order of the forward n_0 samples (X_i^*, t_i^*) $(i = 1, 2, 3, ..., n_0)$. For $j = 1, 2, ..., d$:

$$w_j^1 = \frac{1}{\max_{i=1,2,3,...,n_0}\{|X_{ij}|\}} > 0,$$

$$X_{ij}^1 = w_j^1 X_{ij} \in [-1, 1],$$

$$Y_{ij}^1 = \left|x_{i+1,j}^1 - x_{ij}^1\right| (i = 1, 2, ..., n_0 - 1),$$

$$\delta = \log_{10} d + \log_{10} 2,$$

$$n_j = \left[-\log_{10}(i = 1, 2, ..., n^{\min})\{y_{ij}^1\}\right] + \delta,$$

$$w_j^2 = w_j^1 10 \sum\nolimits_{p=1}^{j} n_p \text{ and set } W = (w_1^2, w_1^2, ..., w_d^2)$$

For exact characterizing the process of time series through resulting network and obtaining the best classification result, the average distance need to be found in the interval of minimum and maximum function:

$$M = \frac{(\max\{g(x)\}+\min\{g(x)\})}{2},$$
$$K_i = \frac{2M}{\min_{j=i,i+1}\{W.X_j\text{-}W.X_{j-1}\}}(i = 2,3,n_0 - 1),$$
$$W_i = K_i W(i = 2,3,...,n_0 - 1),$$
$$W_1 = W_2, W_{n_0} = W_{n_0-1}, b_i = x_0 - W_i$$
$$X_i(i = 1,2,...,n_0)$$

Step 2: Calculate the output weights $\beta = (\beta_1, \beta_2, \beta_3, ..., \beta_{n_0})^T(i = 1, 2, 3, ..., n_0)$, $T = (t_1, t_2, t_3, ..., t_n)^T$

$$H = \begin{bmatrix} g(W_1.X_1 + b_1) & \cdots & g(W_{n_0}.X_1 + b_{n_0}) \\ \vdots & \ddots & \vdots \\ g(W_1.X_{n_0} + b_1) & \cdots & g(W_{n_0}.X_{n_0} + b_{n_0}) \\ g(W_1.X_{n_0+1} + b_1) & \cdots & g(W_{n_0}.X_{n_0+1} + b_{n_0}) \\ \vdots & \ddots & \vdots \\ g(W_1.X_n + b_1) & \cdots & g(W_{n_0}.X_n + b_{n_0}) \end{bmatrix}_{n\times n_0}$$

Then:

$$\beta = H^{\dagger}T = (H^TH)^{-1}H^TT$$

Where, $H^{\dagger}$ is the Moore-Penrose generalized inverse of H. The Modified-EELM obtains minimum square training error, the best generalization ability on unseen data samples, remains steady and has a good performance.

RESULTS AND DISCUSSION

In this study, a novel method Modified-EELM is proposed for epileptic seizure detection in EEGs. This proposed method discriminates between normal brain states and inter-ictal brain states of epileptic subjects effectively. This method combines multi wavelet transform and its feature extraction with Modified-EELM for classifying the EEGs. This method is evaluated through two different classification problems originated from medical diagnosis. The Modified-EELM is used in order to classify EEG signal and some important features of Modified-EELM has better robustness, good performance and consumes short time compared with the training time of output weights. For signal processing, Modified-EELM achieves higher classification accuracy than that of ELM (Huang *et al*., 2004, 2006). These learning machines is used perform by accuracy, sensitivity and specificity.

The accuracy, sensitivity and specificity of Modified-EELM, ELM, ANN and SVM learning machines are shown in the Table 1. The accuracy, sensitivity and specificity of learning machine parameters are determined in Eq. 12-14, respectively:

Table 1: EEG classification results with different machine learning methods

Learning machine	Accuracy (%)	Sensitivity (%)	Specificity (%)	Processing time (sec)
Modified-Eelm	98.01	96.24	98.80	0.0008
Elm	96.85	95.40	98.29	0.0011
Svm	87.85	85.24	90.46	1.8200
Ann	95.14	93.81	96.47	21.7200

Table 2: Performances of the modified-EELM for different activation functions

Activation function	Training accuracy (%)	raining process time (sec)	Testing accuracy (%)	Testing process time (sec)	Average RMSE
Sigmoid	98.17	0.2850	98.01	0.0008	0.138
Radial basis	97.95	0.0608	95.12	0.0288	0.164

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \tag{12}$$

$$\text{Sensitivity} = \frac{TP}{TP+FN} \tag{13}$$

$$\text{Specificity} = \frac{TN}{TN+FP} \tag{14}$$

Where:

TP = Patients with the characteristic and Tested Positive
FN = Patients with the characteristic and Tested Negative
FP = Patients without the characteristic and Tested Positive
TN = Patients without the characteristic and Tested Negative

In Table 1, the classification results show that the Modified-EELM with multi-wavelet features has the highest classification accuracy as compared to other classifiers. Classification accuracy achieves 98.01% for epileptic EEG signals while using sigmoid activation function.

The Modified-EELM performance depends upon the activation function and the number of nodes in the hidden layer. In Table 2, the output of hidden layers: sigmoid and radial basis activation functions train and testing network results.

CONCLUSION

The proposed method has effectively classified EEG data sets with the emphasis for epileptic seizure detection.

It is concentrated on multi wavelet transform, its feature extraction is to reduce the dimensionality and modified-EELM is proposed for classifying the EEG data into normal/non-seizure and epileptic. Compared with other classification methods, the proposed method has better robustness, easy implementing and good generalization performance and also significant improvement in the testing accuracy can be achieved by smoothing their raw outputs. The proposed approach demonstrates its potential for real-time seizure detection. In experiment results, researchers will compare the testing accuracy of the proposed method with some other popular methods with different classifiers.

REFERENCES

Coyle, D., T.M. McGinnity and G. Prasad, 2010. Improving the separability of multiple EEG features for a BCI by neural-time-series-prediction-preprocessing. Biosignal Proces. Control, 5: 196-204.

Fisher, R., W. van Emde Boas, W. Blume, C. Elger, P. Genton, P. Lee and J. Engel Jr., 2005. Epileptic seizures and epilepsy: Definitions proposed by the International League Against Epilepsy (ILAE) and the International Bureau for Epilepsy (IBE). Epilepsia, 46: 470-472.

Huang, G.B., Q.Y. Zhu and C.K. Siew, 2004. Extreme learning machine: A new learning scheme of feed forward neural networks. Proceedings of the International Joint Conference on Neural Networks, June 25-29, 2004, Budapest, Hungary, pp: 985-990.

Huang, G.B., Q.Y. Zhu and C.K. Siew, 2006. Extreme learning machine: Theory and applications. Neurocomputing, 70: 489-501.

Ince, N.F., F. Goksu, A.H. Tewfik and S. Arica, 2009. Adapting subject specific motor imagery EEG patterns in space-time-frequency for a brain computer interface. Biomed. Signal Proces. Control, 4: 236-246.

Jafari-Khouzani, K. and H. Soltanian-Zadeh, 2003. Multi-wavelet grading of pathological images of prostate. IEEE Trans. Biomed. Eng., 50: 697-704.

Jahankhani, P., V. Kodogiannis and K. Revett, 2006. EEG signal classification using wavelet feature extraction and neural networks. Proceedings of the IEEE John Vincent Atanasoff 2006 International Symposium on Modern Computing, October 3-6, 2006, Sofia, pp: 120-124.

Kalayci, T. and O. Ozdamar, 1995. Wavelet preprocessing for automated neural network detection of EEG spikes. IEEE Eng. Med. Biol., 14: 160-166.

Martin, M. and A. Bell, 2001. New image compression techniques using multiwavelets and multiwavelet packets. IEEE Trans. Image Process., 10: 500-510.

Mohseni, H., A. Maghsoudi, M.H. Kadbi and J. Hashemi and A. Ashourvan, 2006. Automatc detection of epileptic seizure using time-frequency distributions. Proceedings of the IET 3rd International Conference on Advances in Medical, Signal and Information Processing, July 17-19, 2006, Glasgow, Scotland, UK., pp: 1-4.

Naik, G.R., D.K. Kum and S.P. Arjunan, 2010. Pattern classification of Myo-electrical signal during different maximum voluntary contractions: A study using BSS techniques. Meas. Sci. Rev., Vol. 10. 10.2478/v10048-010-0001-y.

Nigam, V.P. and D. Graupe, 2004. A neural-network-based detection of epilepsy. Neurol. Res., 26: 55-60.

Oskoei, M.A. and H. Hu, 2007. Myoelectric control systems: A survey. Biomed. Signal Process. Control, 23: 275-294.

Ruiz, R.A.S., R. Ranta and V. Louis-Dorr, 2011. EEG montage analysis in the blind source separation framework. Biosignal Proces. Control, 6: 77-84.