EEG signal analysis for Seizure detection using Discrete Wavelet Transform and Random Forest

Suvadeep Bose, B.Tech, NIT Warangal, V Rama, Assistant Professor, NIT Warangal, and
Dr. C.B.Rama Rao, Associate Professor, NIT Warangal

***Abstract*—Epilepsy, a recurring disorder is symptomized by unprovoked seizures. Considered as one of the most common neurological disorders, Epilepsy affects people of all ages. Around 65 Million people across the world suffer from this disease. Manual diagnosis of EEG signals of long duration may be a source of error as well as a cumbersome task. Hence automation in Seizure Detection is essential for diagnosis of Epilepsy. Therefore, in this paper, an expert system for EEG signal classification has been proposed. The proposed system aims at classifying EEG signals into 3 classes- Normal , inter-ictal (EEG recordings of epileptic patients during non-seizure period) and ictal (EEG recordings of epileptic patients during seizure period). Discrete Wavelet Transform, a technique for analyzing signals in time-frequency domain has been quite successful in this respect for extraction of features from EEG signals. Further Envelope Analysis has been used on the DWT coefficients before feature extraction in order to elevate the performance of the system. Finally several Machine Learning techniques (E.g: MLP and Random Forest) have been used for classification of the EEG signals and the relative accuracy of these Classifiers with regard to this problem have been compared. The results from the experiments show that the Random Forest Classifier is the most effective classifier for this problem delivering an accuracy of 98%. The data that has been used for this work has been taken from the publicly available EEG database of University of Bonn which presents 100 single channel EEG records for each of the above classes.**

***Keywords*—EEG, Epilepsy, seizures, ictal, inter-ictal, DWT, MLP, Random Forest**

I. INTRODUCTION

Many people having epilepsy experience multiple types of seizure and it is also possible that they may also present other symptoms of having neurological problems.[1] A seizure is a sudden surge of electrical activity in the brain. The complex change in the chemical activity of our neurons causes a sudden change in electrical signal and it usually lasts for a short time. With the part or lobe of the brain from which seizure is originated, the nature of seizure may also vary. [20] The Seizures remain for a very short time (a few seconds to around a minute).Occurrence of seizures is also non-deterministic. Different types of seizures cause different changes in the person's behavior who is experiencing it which originate due to various reasons. Seizures are focal or generalized. Focal seizures localize the burst in electrical activity in a part of brain.[7] Therefore, one tends to have localized (focal) symptoms. Different functions are controlled by different parts of brain and so symptoms depend on which area of brain is affected. It should be kept in mind that experiencing a single seizure does not indicate epilepsy. Generally, every 1 in 20 people experience a seizure in some part of their lifetime but it may be the only seizure that they experienced. [13] But epilepsy causes multiple seizures. The occurrence rate of seizures in individuals with epilepsy is different. In few cases, there might be years between two seizures. On the other hand, in few cases the seizures occur every day. For some cases, the occurrence rate of seizures is somewhere in between these extremes. The reason for epilepsy may remain undetermined in many cases. Sometimes it may be caused due to a brain injury, excessive stress, heredity, etc. It is hard to ascertain whether a person is undergoing seizure.[18] During a seizure, the person undergoes different psychological and physiological changes. Tests such as brain scans, electroencephalogram (EEG - brainwave recordings) and blood tests might help in diagnosing Epilepsy. Out of these EEG test has been found to be most successful method for detecting epilepsy.[7] Electroencephalogram (EEG) signals which are considered as one of the prime Biomedical signals in the area of Biomedical research are the records of the electrical activity going on inside the brain as taken from the scalp with the help of electrodes.

The measured signal intensity for EEG signals is quite very small , measured in microvolts (V). The main frequency components of Human EEG waves are:[1]

- Delta: has a frequency of $<$ 3 Hz. It tends to be the highest in intensity and the wave with largest time period. It is common as the dominant rhythm in newborns up to 1 year and in stages 3 and 4 of sleep. It may occur focally with subcortical lesions and in general distribution with diffuse lesions, metabolic encephalopathy hydrocephalus or deep midline lesions.[1] It is usually most prominent frontally in adults (e.g. FIRDA - Frontal Intermittent Rhythmic Delta) and posteriorly in non adults e.g. OIRDA - Occipital Intermittent Rhythmic Delta).
- Theta: has frequency of 3.5- 7.5 Hz and is classified as "slow" activity. It is perfectly common in children $<$13 years and in sleep but not normal in awake adults.[1] It can be seen as a manifestation of focal subcortical lesions; it can also be seen in generalized distribution in diffuse disorders such as metabolic encephalopathy or some instances of hydrocephalus.
- Alpha: has a frequency of 7.5 to 13 Hz. Is usually best viewed in the posterior regions of the brain on each side, being higher in intensity on the dominant side. It is seen when closing the eyes and relaxing, and vanishes when opening the eyes or alerting by mechanisms like thinking, calculating, etc.[1] It is the major rhythm seen

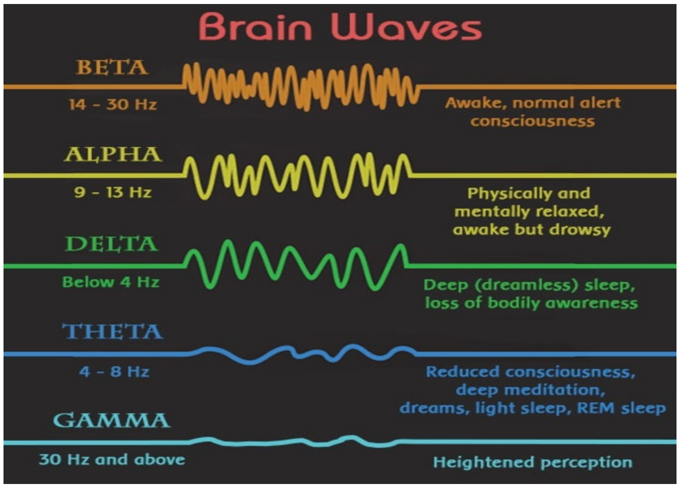


Figure 1: Types of brain waves [4]

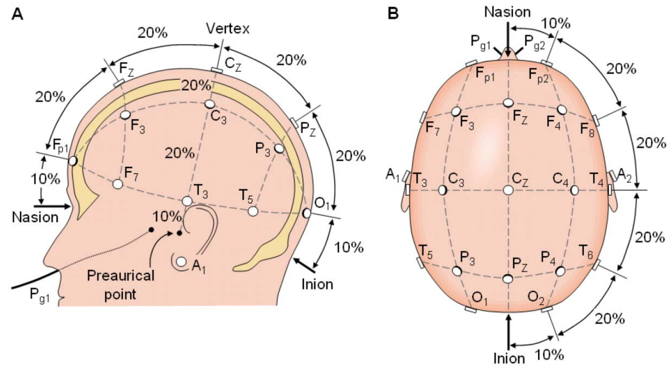


Figure 2: International 10-20 system: placement of electrodes on the scalp [3]

in normal relaxed adults. It is present for the most of the life particularly after the thirteenth year.

- Beta: beta activity is "fast" activity. It has a frequency of greater than 14Hz. It is usually viewed on both sides in symmetrical distribution and is most evident frontally.[1] It is accelerated by sedative-hypnotic drugs particularly the benzodiazepines and barbiturates. It may be not be present or less in areas of cortical damage. It is usually considered as a normal rhythm. It is the dominant rhythm in patients who are alert, anxious or have their eyes opened. The prominent brainwaves have been illustrated in Fig. 1.

The internationally standardized 10-20 system is usually used to record spontaneous EEG. In this system 21 electrodes are attached on the surface of the scalp.[10] Fig. 2 shows the electrode placement in International 10-20 system. An EEG test or electroencephalogram can look at the electrical activity of the brain and may help in determining whether a patient undergoing EEG test is experiencing a seizure.[19], [9] It may also help in predicting whether more seizures will occur. Certain patterns on the EEG are typical of epilepsy. If a person's brain waves show those type of patterns, the person is about twice as likely to develop epilepsy. There are different mental stages of an epileptic patient as can be detected from the change in their brain waves and hence are reflected on the EEG signals. These are: Ictal refers to the physiologic state or event of a seizure. The word originates from the Latin ictus, meaning a stroke or a blow. In (EEG), the recording during a seizure is called "ictal". The following definitions refer to the relationship of seizures with time. Pre-ictal refers the condition just before the seizure, cerebral pain, or stroke, however it has as of late been found that a few characteristics of this state, (for example, Migraine aura) indicate the beginning of the ictal state. Post-Ictal demonstrates the state soon after the ictal state. Interictal refers to the interim between seizures, that are characterizes an epileptic disorder. For the vast majority of the general population of epileptic patients, the interictal state exists for more than 99% of their lifetime. This Interictal period EEG recording is analyzed by many neurologists to detect epilepsy in patients.

Epileptic patients suffer from various challenges in their daily life. They may undergo various physical damages and/ or accidents at the time of the seizure if they are busy in certain tasks which may involve heavy machinery for instance driving. Therefore detection of seizure is very important for ascertaining its cause, origin, etc. so that proper treatment and medication can be offered to those epileptic patients suffering from epilepsy and reduce the chance of physical damage or accidents. The traditional way of detecting seizure in EEG signals is the manual skimming through 10s or even 100s of hours of recorded EEG signals along with the need for an expert for the task. This is really a strenous and time consuming job and also prone to human errors. Hence auto detection of seizure in EEG signals is necessary. Many such methods are available using different approaches for the same problem. In this project we made a system whch analyzes the EEG signals channelwise, and extracts features from them using Discrete Wavelet Transform(DWT) followed by feeding those features to a neural network for detection of seizure. The available EEG signals have been classified into 3 classes-normal , interictal and ictal. Various classifiers have been tested against the same features extracted from the EEG signals to judge their relative performace . Finally it has beeen found that The Random Forest Classifier gave the highest overall accuracy of 98% folllowed by MultiLayer Perceptron(MLP) which has an overall accuracy of 96.7%. The flow-chart for the proposed system has been shown in Fig. 3. The following sections of the paper have been organised as follows: Section II talks about the data used in the system and the details of the proposed methodology which includes feature extraction and classification techniques used in the present work. The Results and Discussions Section (Section III) shows the results obtained by the proposed system, the methods for evaluating the effectiveness of such results and compares them with some of the previous works centred on the same problem. Finally Section IV presents a conclusion of the work done.

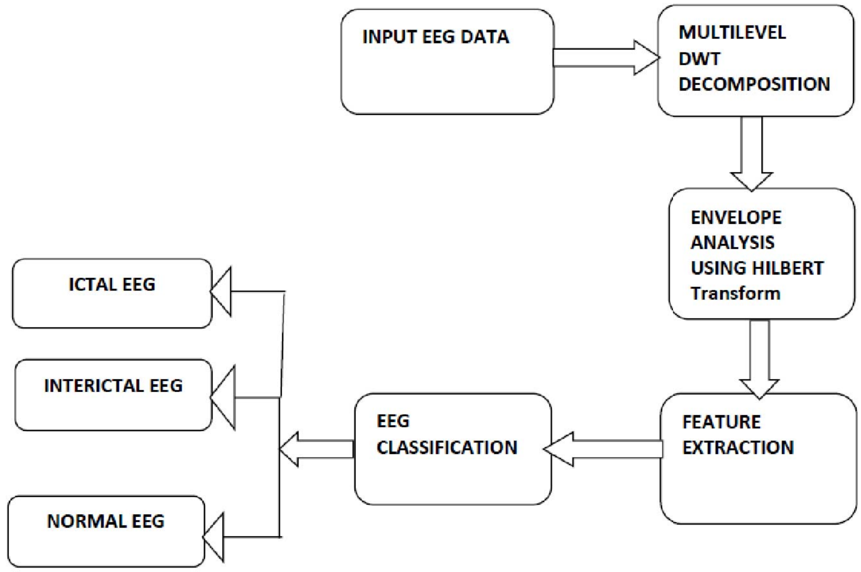


Figure 3: Flow-chart of proposed system

II. Data and Methodology

A. Data

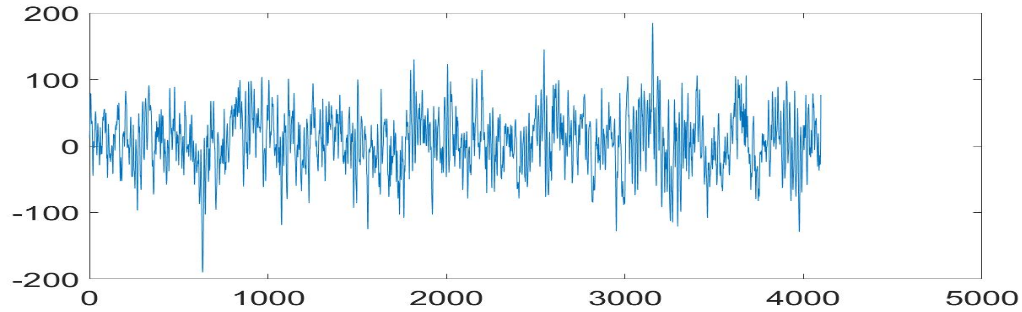
For this study the publicly available resource of EEG database for seizure patients created by University of Bonn has been used.[5], [2] The Database contains 0f 5 sets (Sets A to E) of EEG signals where each set consists of 100 single channel EEG records. Each of these records have been recorded for a duration of 23.6 seconds at a sampling rate of 173.61 Hz with 12 bit ADC resolution. Hence there are total 4097 samples in each of the records. Out of the continuous multi-channel EEG recordings those segments that had very less artifacts example:-muscle activity or eye movement where selected to be included in the database. In this work, in order to evaluate the performance of the proposed methodology, the EEG signals of three subsets namely, A, D, and E have been used. Data in set A consists of segments taken from EEG recordings from the test conducted upon 5 healthy volunteers using the standardized 10-20 electrode placement scheme. Volunteers were relaxed in an awake state with their eyes open. Data in sets D and E were taken from the EEG archive of pre-surgical diagnosis. For these sets EEG recordings of 5 patients were chosen who had achieved complete control over seizure after resection of one of the hippocampal formations which carries the evidence that it was correctly diagnosed to be the epileptogenic zone. While set D contains records taken during seizure free intervals (inter-ictal periods), the set E contains only seizure activities(Ictal period). These segments had been selected from all the recording sites showing ictal activity. The example waveforms for classes A, D and E are displayed below in the following figures (Fig 4 (a), (b) and (c)).

B. Feature Extraction

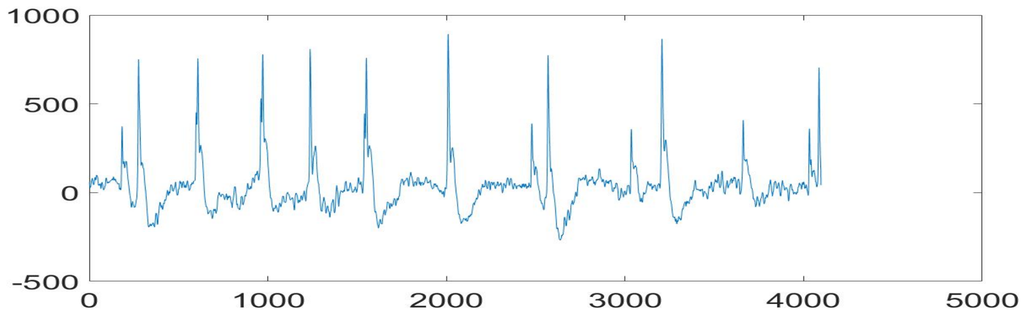
1) DWT Decomposition: The Wavelet Transforms are often used for Multi-Resolution Analysis(MRA) whereby a signal can be broken into a set of basis functions at different levels of resolutions. It is much similar to the Fourier Transform(or windowed Fourier Transform) where a signal is decomposed into sine and cosine terms. In Wavelet Transform , the signals are broken into scaled and time-shifted versions some basis function. Continuous Wavelet Transform(CWT) does this scaling arbitrarily with any arbitrary shift in time (i.e arbitrary wavelets, not necessarily orthogonal) of the function and the Forward transform used for finding its Coefficients is given by:

$$F(a,b) = \int_{-\infty}^{\infty} f(x)\psi^{*}_{a,b}(x)\ dx$$

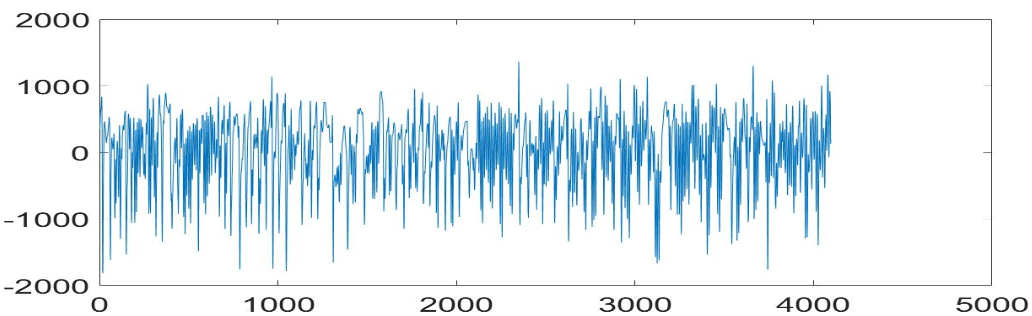
where * is the symbol for complex conjugate and ψ is the basis function. where a and b are scale and position parameters respectively. Computing wavelet coefficients at each conceivable scale is a decent lot of work. So just some discrete scales and positions are picked up, which is called Discrete Wavelet Transfrom(DWT) . Summarily, the DWT disintegrates the signal into mutually orthogonal basis functions called the Scaling functions, which are nothing but integral translates of the mother scaling function. The set of represent-able signals spanned by the scaling function at low scales are subsets of those spanned at higher scales. The subspace between that of the scaling functions of two adjacent scales is spanned by another set of basis functions called the Wavelet Functions. The full DWT for signal x(t) can be represented as a combination of a scaling function $\phi_{j,k}$ and a translated and scaled form of a mother wavelet function $\psi_{j,k}$.



(a) Set A (normal)



(b) Set D (inter-ictal)



(c) Set E (ictal)

Figure 4: Samples of EEG records from sets A, D and E

Using DWT the a discrete signal f(n) can be represented as:

$$f(n) = \frac{1}{\sqrt{M}} \sum_{k} W_{\phi}(j_0, k)\phi_{j_0,k}(n)$$

$$+\frac{1}{\sqrt{M}} \sum_{j=j_0}^{\infty} \sum_{k} W_{\psi}(j, k)\psi_{j,k}(n)$$

where $W_{\psi}(j, k)$ are the wavelet coefficients or detailed coefficients and $W_{\phi}(j_0, k)$ is the scaling coefficient or approximation coefficient. These coefficients are obtained using the forward transform by:

$$W_{\phi}(j_0, k) = \frac{1}{\sqrt{M}} \sum_{n} f(n)\phi_{j_0,k}(n)$$

$$W_{\psi}(j, k) = \frac{1}{\sqrt{M}} \sum_{n} f(n)\phi_{j,k}(n)$$

The Fast Wavelet Transform(FWT) proposed by Mallat (1988), exploits a special relationship among the DWT coefficients at adjacent scales. With the fast wavelet transform making use of filter banks, we can implement signal decomposition in an efficient way, as shown in Fig. 5. The signal can thus be decomposed into its subbands and the time variation

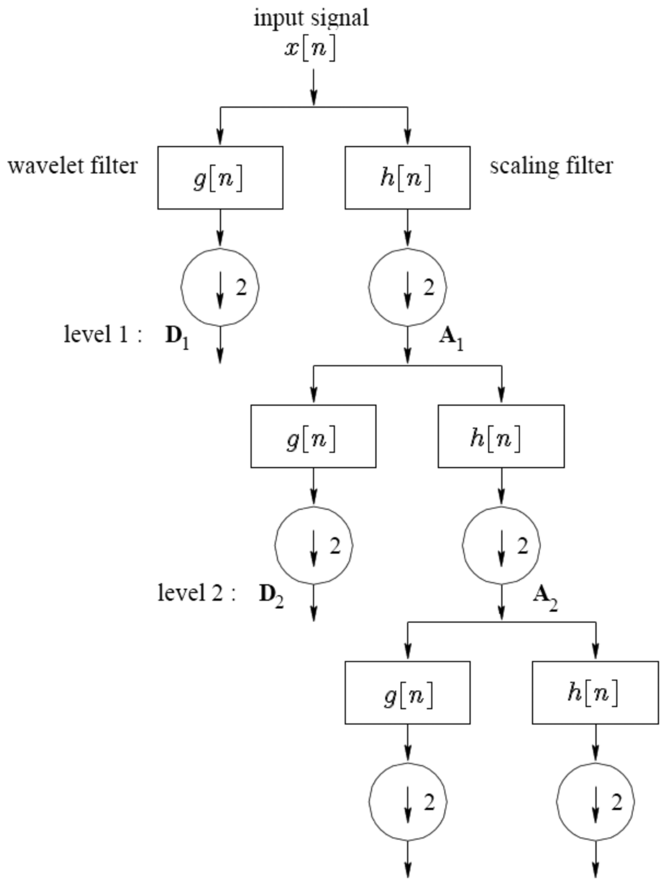
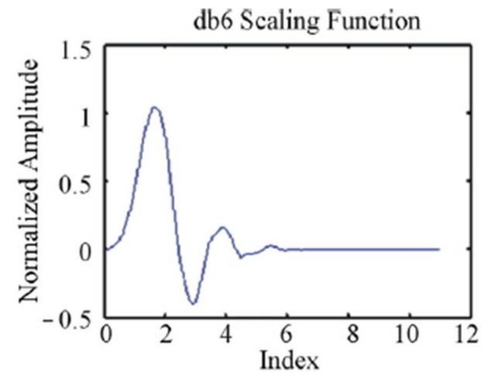
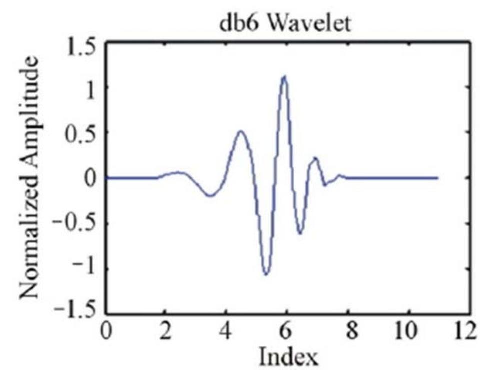


Figure 5: Block diagram of 2 level DWT decomposition [8]



(a) db6 scaling function



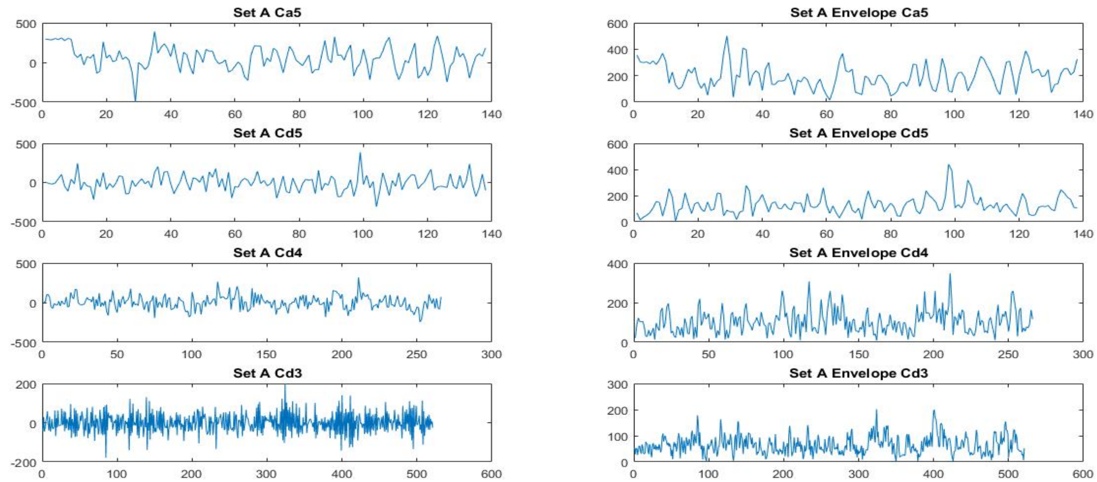
(b) db6 wavelet function

Figure 6: db6 scaling and wavelet function

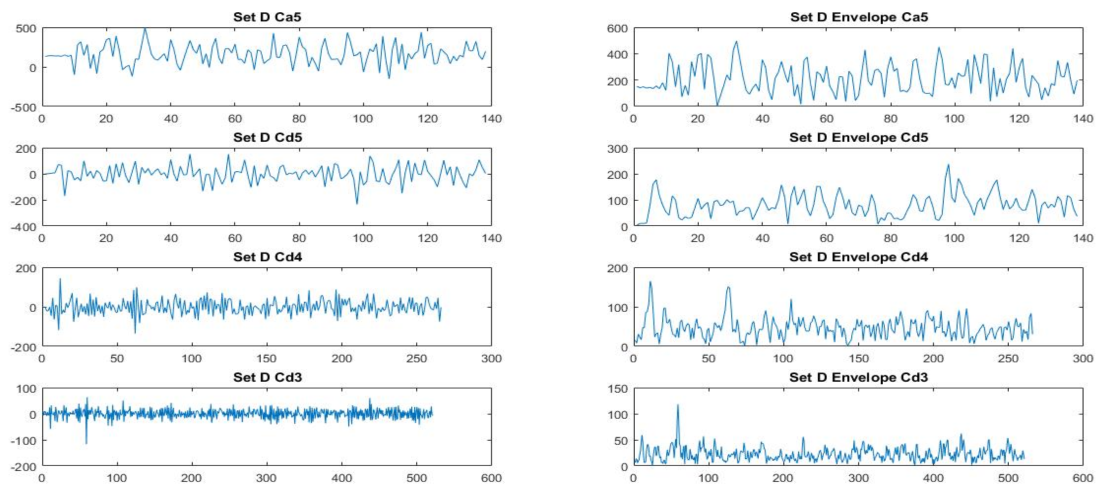
of each of these subbands can be analysed . Using a N level DWT decomposition , we get N+1 DWT coefficients(1 Approximation Coefficient and N Detailed Coefficients). Fig shows block diagram for 2 level FWT decomposition of a signal using Filters. The selection of wavelet function is very crucial for DWT decomposition of a signal. In this paper, we extract features by using db6 wavelet function as it is seen that Daubechies' wavelet is one of the most used wavelet function in the earlier works on this problem [11]. Fig.6(a) and (b) shows the scaling and wavelet functions for db6 respectively. The discrete Fourier change (DFT) reveals to you the frequency components of a signal, averaged over the whole duration of the signal. On the other hand, the Time Domain Analysis enlightens just regarding the changes in signal values with time and misses the elements that can be extricated from its frequency domain. It is the Discrete Wavelet Transform (DWT) that gives data about the frequency (basis) components and having the capacity to demonstrate what time these segments happen at. Henceforth complex and resource intensive DWT has emmerged as the favored technique for extraction of features from non-stationary bio-medical signals, for example, EEG Signals instead of basic time or frequency domain investigation. We have selected the Daubechies wavelet of order 6 (db6). The number of decomposition levels was set to 5. In other words, the EEG signals have been decomposed into the detailed coefficients D1-D5 and approximation coefficient c.

2) Hilbert Huang Transform for Envelope Analysis: The Hilbert Huang Transform (HHT) which can extricate the frequency components from conceivably nonlinear and non-stationary discontinuous signals like the EEG signal[21] , is a powerful tool which can portray the frequency components locally and adaptively for about any oscillating signal.[6]. Instead of a Fourier or wavelet based transforms, the Hilbert Transform can be utilized, with a specific end goal to process prompt frequencies and amplitudes and portray the signal all the more locally.

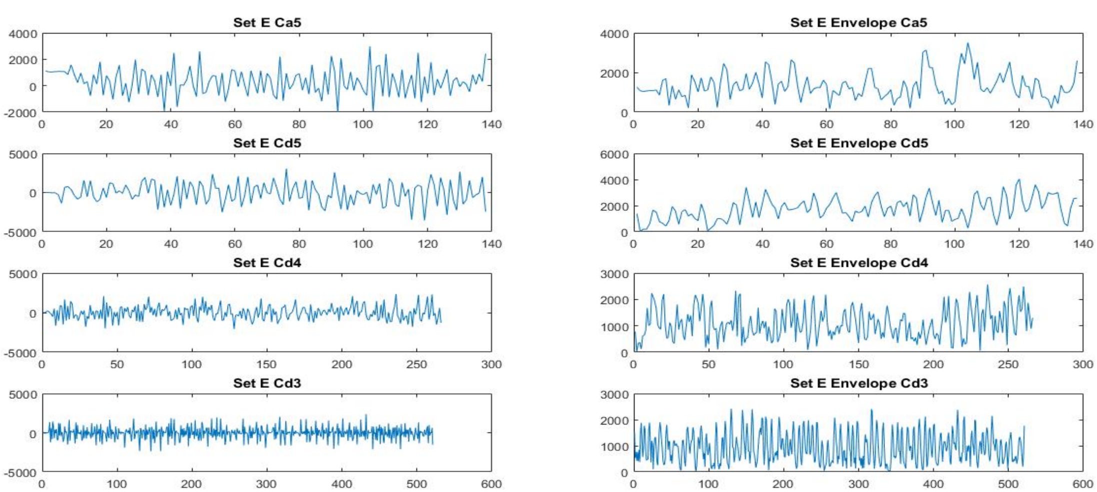
In this technique the envelope spectrum of a signal is made utilizing the popularly known Hilbert Transform (HT) to get clearer and smoother plots for the DWT coefficients of the EEG signals at various frequency bands. The formula for HT is given by: In this method the envelope spectrum of a signal is created using the popularly known Hilbert Transform (HT) to get clearer and smother graphs for the DWT coefficients of the EEG signals at different frequency bands.



(a) Set A (normal)



(b) Set D (inter-ictal)



(c) Set E (ictal)

Figure 7: Relevant DWT Coefficients and their envelopes of sample EEG records from sets A, D and E

The formula for HT is given by:

$$\bar{x}(t) = x(t) * \frac{1}{\pi\tau} = \int_{-\infty}^{\infty} \frac{x(\tau)}{t-\tau}\ d\tau$$

Where $\bar{x}(t)$ is denoted as HT of real signal x (t). And the envelope spectrum is calculated as:

$$S(t) = |\bar{x}(t)|$$

There is a general increase in classification accuracy due to envelope analysis of the DWT coefficients before performing feture extraction.

3) Features: Feature extraction is one of the most vital steps in training a classifier. While extracting features one should be careful about how each feature is going to affect the accuracy of the classifier. More features does not necessarily guarantee the robustness of a classifier ,on the other hand addition of more number of dependent or redundant features into the feature vector makes the classifier over-fitted to the training data and might yield poor accuracy for the test data. The data that has been taken from the University of Bonn Database for EEG has been sampled at 173.61 Hz. But the frequency components above 40 Hz is lack of use in epilepsy analysis and also the power line interference at 50-60 Hz is also unwanted in the analysis of the EEG signal, therefore the advisable envelope sub-bands (Ca5, Cd3-Cd5) are selected for feature acquisition. This in turn helps to reduce the dimensionality of the extracted feature vectors since we are not taking into account the unnecessary DWT coefficients. A total of 21 features has been used in the feature vector for each EEG record for training the Classifiers. The sets of features are given bellow:

- Maximum values of the 4 wavelet coefficients ((Ca5, Cd3-Cd5) and undecomposed EEG signal.
- Minimum values of the 4 wavelet coefficients ((Ca5, Cd3-Cd5) and undecomposed EEG signal.
- Mean values of the 4 wavelet coefficients ((Ca5, Cd3-Cd5) and undecomposed EEG signal.
- Energy values of the 4 wavelet coefficients ((Ca5, Cd3-Cd5) and undecomposed EEG signal
- Shannon entropy of the original EEG signal record The Shannon Entropy(H(X)) of a discrete random variable X with possible values x1, ..., xn and probability mass function P(X) may be defined as:

$$H(X) = E[I(X)] = E[-ln(P(X))]$$

Where E is the expectation and I is the total information contained in signal X. Therefore I(X) is also a random variable. Hence the entropy can be written as:

$$H(X) = \sum_{i=1}^{n} P(x_i)I(x_i) = -\sum_{i=1}^{n} P(x_i)log_b P(x_i)$$

Fig. 7 shows the relevant DWT coefficients and their envelopes for sample EEG signals taken from Sets A, D and E of the database respectively.

C. Classification

In the present work 2 types of Classifiers have been used and their results have been compared. These are: Multi-Layer Perceptron(MLP) and Random Forest. Out of these, Random Forest has been seen to give the highest accuracy. Following subsections present a brief discussion on these classification algorithms.

1) MLP: A multilayer perceptron (MLP) is a feedforward artificial neural network model that takes a vector of features from the given dataset as input and map them onto appropriate classes or outputs. In MLP there are several layers of nodes in a directed graph where all the nodes in a layer are connected to all the nodes in the next layer. All other nodes except the input nodes, have an activation function(E.g.: Sigmoid function).

The MLP has an input layer, an output layer and one or more hidden layers. The input layer consists of the same number of nodes as the size of the input feature vector. The output layer has the same number of nodes as the number of classes that the input data should be classified into. Each node in one layer is connected to all other nodes in the next layer through different weights. At the beginning of training phase the weights are initialized randomly. During the training phase the perceptron first processes the given data through feed forward propagation. The error in the output node j for the nth data point is referred to as

$$e_j(n) = d_j(n) - y_j(n)$$

where d is the correct lebelled value used for training purpose and y is the output value produced by the perceptron. We then make suitable adjustments in the weights of the nodes so that the cost function or the overall error in the output is minimized. The cost function may be defined as:

$$E(n) = \frac{1}{2}\sum_j e_j^2(n)$$

which is called the mean square error. Using gradient descent, we try to minimize the cost function and in turn change the weights of connections between each of the nodes. The change Δw_tin each weight is given by:

$$\Delta w_t = -\eta \frac{\partial E(w)^T}{\partial w}|_{w_t}$$

Where η is called the learning rate which should be carefully chosen so that the cost function converges to a global minimum. If the learning rate is too large the algorithm may never converge. On the other hand if kept very low it may take a long time for the algorithm to converge. Generally it is kept between 0 to 1. The MLP network used for this project has 3 layers: input, output and 1 hidden layer. The input layer contains 21 nodes as the size of the feature vector is 21, the output layer has 3 nodes since there are 3 classes for the problem and the hidden layer contains 15 nodes and the learning rate has been set as 0.3. Fig. 8 shows the block diagram of the MLP neural network.

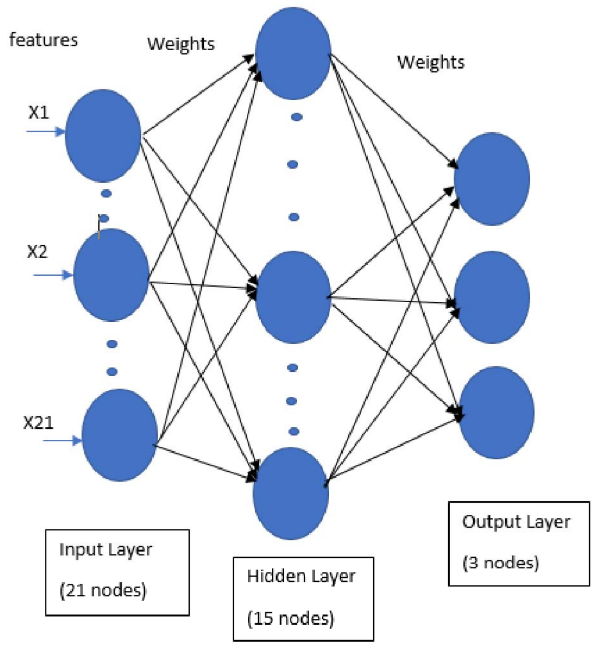


Figure 8: Block diagram for proposed MLP neural network architecture

2) Random Forest: Random forests is an ensemble learning method for classification, regression and other tasks, In which a number of decision tress are created at the training time and a class is given as output which is the mode of the classes. In order to create number of decision trees, data and variables are selected randomly from the available set of data and variables. To build a tree during training time a finite set of thresholds is used among which a threshold is selected for each node. While constructing a tree separation of classes is being done and probability of data point to be of any class is different for each node. The newly arrived data point go down in tree and it ends at leaf and the class with highest probability for that node shows the actual class of data point in that tree. Single random tree is not a good classifier but if we combine a number of random trees then it becomes a very good classifier. Let a dataset of N data points is X = x1, ..., xn with responses C= c1, ..., cn. Every input has some features. In order to create more randomness among these datapoints every time some data points are selected and from these data inputs some features are selected randomly in order to make decision tree. If one or a few features are very strong predictors for the target output, these features will be selected in many of the decision trees. Typically, for a classification problem with 'f' features, square root of 'f' features are used in each split. For $t \in \{1, ..., T\}$

1) Sample, with replacement, n training examples from X, C; call these Xt, Ct.
2) Train a decision or regression tree ft on Xt, Ct.

In the forest with T trees we have $t \in \{1, ..., T\}$.All these trees are trained independently. During test case each test point V is simultaneously pushed through all the trees starting at root until it reaches to corresponding leaves. For different trees datapoint will follow different path when it goes to leaf. The class probabilities for that point V is different in each tree. For the output prediction bagging method is used which leads to better model performance because it decreases the variance of the model, without increasing the bias. It can be made by

$$p(c|v) = \frac{1}{T}\sum_{t=1}^{T} p_t(c|v)$$

averaging the predictions from all individual decision trees Where 'c' represents the class of 'v'.

III. Results and Discussions

3 sets of data which are normal person's data with eyes open (SET A), EEG during interictal period of people suffering from epilepsy (SET D), and EEG record containing seizure recorded for the same patients (SET E) has been taken from the BONN University Database for Seizure detection and classification into the above three groups. Each of these data sets contain 100 records of EEG signals for 23.6 seconds' duration recorded at a sampling rate of 173.6 Hz. The feature extracted from these records were fed to various Machine Learning Classifiers to compare their performance on the given set of features. The best result was obtained from Random Forest Classifier followed by Multilayer Perceptron. In table I, we present a list of classifiers that were used initially and the resulting accuracy obtained by training them over the whole dataset using 5 fold cross validation. An n fold cross validation comprises of the following steps:

1) Break the whole dataset randomly into n equal parts.
2) Choose (n-1) of these sets as training set and the remaining 1 as test set.
3) Model the classifier based on this training set and calculate accuracy based on the test set.
4) Change the test set among the (n-1) other sets and repeat the training and testing for accuracy.
5) The accuracy of the trained model is finally given by the average of the accuracies obtained over the n tests (in this case n=5).

These results for these classifiers were obtained using MATLAB's Classifier learner app and Weka toolbox. Here accuracy is defined as the ratio of total number of correct predictions by the classifier to the total no. of instances.

Based on the accuracy obtained from initial cross validation results, we decided to further optimize the classifier parameters of MLP and Random Forest and investigate their performances in through further experiments.

For obtaining the final result as shown in Table II, 50% of the total dataset (150 records) were taken for training and the rest 50% (150 records) were taken for testing by choosing randomly from the whole dataset.

Although Accuracy is a measure of performance of a classifier which is both easy to calculate, and easy to understand, there are other metrics which reflect the performance of a classifier more appropriately. Hence, apart from accuracy,

Classifier Name	Set A Correctly detected	Set D Correctly detected	Set E Correctly detected	Accuracy
Complex Tree	96/100	89/100	96/100	93.7%
Medium Tree	96/100	89/100	96/100	93.7%
Linear Discriminant Analysis(LDA)	92/100	43/100	67/100	67.3%
Quadratic Discriminant Analysis(QDA)	97/100	63/100	91/100	83.7%
Linear SVM	95/100	95/100	95/100	95%
Medium Gaussian SVM	93/100	90/100	92/100	91.7%
K Nearest Neighbours(KNN)	96/100	73/100	93/100	87.3%
Multi-Layer Perceptron(MLP)	100/100	90/100	97/100	95.7%
Random Forest	100/100	90/100	97/100	95.7%

Table I: 5-fold Cross Validation result with different classifiers

Classifier Name	Set A Correctly detected	Set D Correctly detected	Set E Correctly detected	Accuracy
Multi-Layer Perceptron(MLP)	45/48	42/45	57/57	96.67%
Random Forest	48/48	42/45	57/57	98%

Table II: Test Set results for MLP and RAndom Forest Classifiers

the above 2 classifiers (MLP and Random Forest) have also been compared based on their sensitivity, specificity and ROC Area which are considered as more concrete measures of the performance of a classifier.

The metrics used for comparing the 2 classifiers are explained below:

- Accuracy = (1 – Error) = (TP + TN)/ (PP + NP) = P(C), the probability of a correct classification.
- Sensitivity of class j = TP/ (TP + FN), where class j has been taken as the positive class.
- Specificity of class j = TN/ (TN + FP), where class j has been taken as the positive class.
    - TP=True Positive
    - TN=True Negative
    - FP=False Positive
    - FN=False Negative
- ROC Area for class j: Receiver Operating Curve(ROC) is a plot between true positive and false positive rates for class j taken as the positive class at different threshold values. The area under the ROC curve is called the ROC Area. This is a significant measure of performance of a classifier. It lies between 0 and 1). An ideal classifier, which predicts all the instances correctly should have an ROC area of 1. More the ROC area, better is the classifier performance.

Table III shows the various statistical metrics mentioned above for MLP classifier and table IV shows the same for Random Forest Classifier.

Class	Sensitivity(%)	Specificity(%)	ROC Area
Normal EEG (Set A)	97.9	98	0.999
Interictal EEG (Set D)	91.1	99.04	0.977
Ictal EEG (Set E)	100	97.85	0.995

Table III: Statistical metrics for MLP Classifier

Table V presents a comparison of accuracy of some of the recent implementations of the EEG classification problem with our present work.

Class	Sensitivity(%)	Specificity(%)	ROC Area
Normal EEG (Set A)	100	100	1.00
Interictal EEG (Set D)	93.3	100	0.979
Ictal EEG (Set E)	100	96.78	0.993

Table IV: Statistical metrics for Random forest Classifier

Authors	Features	Classifiers	Accuracy
Hasan Ocsk [16]	Approximate Entropy(APen) & DWT	Classification on the basis of Apen values for different classes	96%
Ling Guo et al.[14]	Genetic Programming (GP)	KNN Classifier	93.5%
Ghosh Dastidar et al.[17]	9-parameter mixed band feature	PCA enhanced RBF neural network classifier	96.6%
Narendra Kumar et al.[15]	Optimised S-Transform	Back Propagation Artificial Neural Networks	93.33%
Enamul Kabir et al.[12]	Optimum Allocation Technique(OAT)	Logistic Model Trees(LMT)	95.3%
This work	DWT and Envelope Analysis(EA)	Multi-Layer Perceptron	96.7%
This work	DWT and Envelope Analysis(EA)	Random Forest	98%

Table V: Statistical metrics for Random forest Classifier

IV. Conclusions

EEG signal records were decomposed using multi-level DWT and further subjected to envelope analysis of the coefficients obtained from DWT. This preprocessing before feature extraction from the EEG signals particularly help in increasing the classification accuracy for detection of seizure or epilepsy. Finally the EEG records obtained from the database were classified into 3 classes: (i) healthy person's EEG, (ii) EEG records of people suffering from epilepsy during inter ictal period and (iii) the same patients' EEG records during seizure(ictal period). Among the various classifiers used for classifying this data Multi-Layer Perceptron (MLP) and Random Forest showed promising results in terms of accuracy, sensitivity and specificity when verified using 50% of the available EEG data as training set and the remaining 50% as the test set. Random forest gave an over-all accuracy of 98% followed by MLP (96% accuracy).

References

[1] Biomedical signal acquisition. https://www.medicine.mcgill.ca. Accessed: 2017-04-30.

[2] The bonn eeg database. http://epileptologie-bonn.de/cms/front$_content.php?idcat = 193lang = 3. Accessed : 2017 - 04 - 30$.

[3] Recognition of speech artifacts in eeg records. http://docplayer.com.br/18572674-Reconhecimento-de-artefactos-relativos-a-fala-em-registos-de-eeg.html. Accessed: 2017-06-20.

[4] the-mystery-of-gamma-waves. http://q4lt.com/the-mystery-of-gamma-waves. Accessed: 2017-06-20.

[5] Lehnertz K. Rieke C. Mormann F. David P. Andrzejak, R. and C. Elger. 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. v64. 061907.*

[6] Bradley Lee Barnhart. *The Hilbert-Huang Transform: theory, applications, development*. PhD thesis, University of Iowa, 2011.

[7] Dr Colin Tidy Dr Tim Kenny. Epilepsy with focal seizures. http://www.stillingtondoctors.co.uk. Accessed: 2017-04-30.

[8] David Monaghan Edmond Mitchell and Noel E. O'Connor. Classification of sporting activities using smartphone accelerometers. *Sensors*, 13(4), 2013.

[9] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. Ch. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley. PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation*, 101(23):e215–e220, 2000 (June 13). Circulation Electronic Pages: http://circ.ahajournals.org/content/101/23/e215.full PMID:1085218; doi: 10.1161/01.CIR.101.23.e215.

[10] Robert plonsey Jaakko malmivuo. Bio-electromagnetism : Principles and applications of bioelectric and biomagnetic fields. http://www.bem.fi/book/13/13.htm, 1995. Accessed: 2017-04-30.

[11] Y. Brain Inf. Kabir, E.Siuly Zhang. World academy of science, engineering and technology international journal of medical, health, biomedical, bioengineering and pharmaceutical engineering. 1(3), 2007.

[12] Y. Brain Inf. Kabir, E.Siuly Zhang. Epileptic seizure detection from eeg signals using logistic model trees. *Brain Informatics*, 3(2):93–100, June 2016.

[13] Dr Tim Kenny. Epilepsy with focal seizures. https://patient.info/in/health/epilepsy-with-focal-seizures, 2015. Accessed: 2017-04-30.

[14] Julián Dorado Cristian R. Munteanu Alejandro Pazos Ling Guo, Daniel Rivero. Automatic feature extraction using genetic programming: An application to epileptic eeg classification. *Expert Systems with Applications: An International Journal*, 38(8):10425–10436, August 2011.

[15] S.N. Sharma Narendra Kumar Ambulkar. Detection of epileptic seizure in eeg signals using window width optimized s-transform and artificial neural networks. In *IEEE Transactions on Biomedical Engineering-Bombay Section Symposium (IBSS)*, September 2015.

[16] Hasan Ocak. Automatic detection of epileptic seizures in eeg using discrete wavelet transform and approximate entropy. *Expert Systems with Applications: An International Journal*, 36(2):2027–2036, March 2009.

[17] Nahid Dadmehr Samanwoy Ghosh-Dastidar, Hojjat Adeli. Principal component analysis-enhanced cosine radial basis function neural network for robust epilepsy and seizure detection. *IEEE Transactions on Biomedical Engineering*, 55(2):512–518, February 2008.

[18] Amit M. Shelat. Seizures. https://medlineplus.gov/ency/article/003200.htm. Accessed: 2017-04-30.

[19] Ali Shoeb. *Application of Machine Learning to Epileptic Seizure Onset Detection and Treatment*. PhD thesis, Massachusetts Institute of Technology, September 2009.

[20] oseph I. Sirven Steven C. Schachter, Patricia O. Shafer. About epilepsy:the basics. http://www.epilepsy.com/learn/epilepsy-101/what-seizure, 2014. Accessed: 2017-04-30.

[21] M. Şeker and M. S. Özerdem. Classification of eeg data sets with hilbert transform. In *2016 24th Signal Processing and Communication Application Conference (SIU)*, pages 1949–1952, May 2016.