Research Article

Classification of EEG Signal Using Wavelets and Machine Learning Techniques

Mohammed Ziaullah [a], Dr. Kalpana Vanjerkhede [b]

[a] Assistant Professor & Research Scholar, Dept. of ECE, SECAB I.E.T Vijayapur, Karnataka, India
[b] Professor, Dept of E&IE, PDACE, Kalaburagi, Karnataka, India

Abstract Brain is the crucial organ which performs various functions of body. Electroencephalogram (EEG) is an efficient modality which helps to acquire brain signals from scalp. Analysis of EEG signals helps in diagnosis and treatment of brain diseases like epilepsy seizure, and various problems associated with brain disorders. These signals are contaminated with unwanted artifacts and complicate to analyze. The diagnosis requires flawless analysis of EEG signals. Different methods are proposed to examine with high accuracy. In this paper, Discrete Wavelet Transform is used to pre-process the EEG signal and decompose into five frequency bands and then, the features like mean, standard deviation, RMS, entropy, energy and relative energy are computed. These features are evaluated by machine learning classifier such as Support Vector Machine (SVM), K-Nearest Neighbor (KNN) and Naive Bayes (NB). The results demonstrate the highest classification accuracy (99.5%), Specificity (100%), Sensitivity(100%) by SVM for normal versus epileptic subjects.

Keywords: EEG, epilepsy, DWT, Artifacts, Classifier

1. Introduction

Epilepsy is abnormal neural activity of the brain which has affected about fifty million people in the world. In 2015, as of 1,25000 deaths were recorded due to seizures. This diseases is more common in old age people. Epilepsy seizures are nearly undetectable over long period of time. The diagnosis is made by observing the seizure onset and underlying cause. An Electroencephalogram(EEG) can be used to assist the brain activity with suggestive risk of seizure. A EEG signal are read experts to distinguish and classify the different patterns brain seizures. The examination of such signals by visual method is time-consuming and laborious and also requires an expert to analyse. The efficiency of such analysis of the recordings of patients by neurologists get reduces. To overcome such problems of inefficiency in diagnosis of seizure, the computing systems are encouraged to develop effective model to classify and diagnosis the seizures with increased efficiency of detection.

The electroencephalogram (EEG) is an approach used to record the electrical potentials of brain activity. These brain signals are detected and measured by attaching electrodes over the scalp according to 10-20 international system. EEG signals has five major frequency , such as delta, theta, alpha, beta and gamma. EEG is a painless test which is carried by group of small sensors attached over the scalp. Then signals are recorded by the computer machine and analysed by doctor or expert. However, EEG signal contain useful information of brain, but always contaminated due to artifacts and makes the doctor analysis inefficient and mistaken in proper diagnosis. Hence EEG signal processing techniques are used to analyze these waveforms by intelligently selecting the method of diagnosis and treatment.

2. Literature Review

Epilepsy is a class of brain diseases with occurrence of about 1-2% of the world's population (Mormann, Andrzejak et al 2001). It occurs because of unexpected electrical storm and symptoms can vary widely. It is the behaviour of brain resulted from injury or trauma etc. In this state, a patient experiences sensations and may go under unconsciousness. The electroencephalogram (EEG) is used for monitoring these brain activities and has become an tool in epilepsy diagnosis.

Artifacts are basically caused by external sources and biological interferences in the brain .These artifacts present in EEG signal causes the brain signals to misinterpret the cause of brain disorder. Therefore, removal of artifacts has become a prominent role dung EEG analysis. Removing these artifacts may causes lose of important data because EEG signals carries most of the information below 100Hz frequency. EEG is described into rhythm activity and transient. The rhythm activity is divided into frequency bands such as delta wave (0.5-4H , theta wave (4-8Hz), alpha wave(8-13Hz) beta wave (13-30Hz) and >30Hz as gamma wave[2]. Therefore, to remove the artifacts by denoising the EEG signals, a range of techniques and methods have been proposed in literature this makes use of wavelet transform as a tool to find the EEG signal features efficiently.

Morlet and Grossman[2]. first introduced the term Wavelet based investigation of to overcome the limitations of Fourier analysis. A mother wavelet is used for investigation of transient signal with zero mean value, which has its energy concentrated in time well suited for time variance signals. The scaling functions and wavelet, are the basis functions in wavelet study. In signal and image processing applications, Haar basis functions have been effectively.

Asma et al. (2016) conducted the study using Discrete Wavelet Transform (DWT) & provided advantage over other different standard transformation. The analysis of signals with time and frequency localization can be done with multiresolution technique. Hence it is widely used in compression techniques. **Haslaile. Abdullah et al.(2013)** makes use of Double Density and Double Density Complex wavelet based image processing techniques for removal of noise with different sizes of window. The performance was compared and tested with RMSE. The minimum RMSE was obtained for threshold of 20. It was seen that the proposed methods were better than other denoising techniques. **Geeta Kaushik et al. (2012)** illustrate a mechanism for processing and analysis of biomedical signals. It was analyzed that wavelet based denoising are more effective and it can be accomplished by separate signal from noise by thresholding wavelet coefficients.

P.Ashok Babu et al.(2011) proposed, different methods for noise removal in EEG signals. Wavelet based PCA and adaptive thresholding is used to remove artifacts. In contrast to others various methods it was observed that , elapsed time for adaptive threshold is less compared to others**, Patil Suhas S. et al. (2012)** proposed, the wavelet transform can be used in denoising of EEG signal during performing different mental tasks. The proposed algorithm was validated by Colorado state EEG dataset.The experimental results are computed for the signal-to-noise ratio. **Priyanka Khatwani et al. (2013)** Proposed a method which uses wavelet method of denoising signal. The different of noise removal techniques were discussed in this paper. An EEG signal gets contaminated during signal recording hence it should be removed. The author concluded from the literature that the wavelet based denoising method is effective and gives better results.

Weidong Zhou et al. (2004) discussed the application of wavelets and ICA. The experimental results suggested that the integration of the above techniques can efficiently remove EMG noise and ECG artifacts in EEG signals and further increases the ratio of correctness**. Sartori et al (2010),** proposed a method which evaluate the electrical system failures using a non-invasive methodology. The analysis of noninvasive is done using magnetic signature recognition using wavelet technique. The effectiveness of this work was verified using a set of different faults, positioning of sensor and combined parameters**. Dwivedi et al. (2010),** presented a simple and effective technique in denoising PQ signal for enhanced detection. The coefficient of wavelets are correlated with its neighboring sub band and across sub bands. The proposed algorithm is tested power line disturbances and results in significantly better outputs in noise removal.

Lakshmi et al(2013) Proposed a method based on voltage regulation with dual purpose wavelet detection model. Therefore it is observed from literature survey that the use of Wavelets in EEG signal analysis gives efficiently and better results in determining the various frequency bands and also in obtaining the signals features. **Aayesha et al(2021)** proposed a framework which is used for classifying ictal and interactal classes of EEG signal. The work presented over here is automated model for diagnosis of seizure and provide alert for medical emergency. The experimental values show that classification made over CBHMIT and Bonn dataset gives fair results with FRNN and KNN

3. Data Collection and Methodology

The proposed classification of the EEG signal is done using following approach 1) pre-processing of EEG signals 2) feature extraction from decomposed signal, and 3) classification using classifier model.

In the present study, EEG data sets (A, B, C, D and E) is preprocessed by DWT to decompose into five sub band signals using seven level decomposition [12]. Next, useful features like Entropy, Mean, Standard deviation, Energy, RMS, and Relative power are derived from each sub-band. Finally, Extracted features are selected and given as input classifier for epilepsy classification.

3.1 Description of EEG Dataset

The publicly available EEG is taken from University of Bonn ([12]. This dataset include five sets labels such as A, B, C, D and E, with each having 100 single- EEG channel of time duration 23.6 sec having f =173.6 Hz sampling rate. In Each segment of data, N=4097 data points are positioned at the intervals of 1/173.61th of 1s.The sets A and B are acquired from scalp EEG recordings of five healthy volunteers subjects using a standard international electrode placement. The subjects were asked to relax in an awake state with eyes open and eyes closed for SET A and SET B respectively [18]. For Interictal and Ictal epileptic activities, the database sets C, D, and E are acquire from the epileptic subjects through intracranial electrodes. The summary of dataset is shown in table 1.

3.2 EEG Pre-processing

For the EEG pre-processing phase, DWT decomposition is considered as a tool for pre-processing to extract five physiological EEG bands like delta (0-4 Hz), theta (4-8 Hz), alpha (8-12 Hz), beta (13-30 Hz), and gamma (30-60 Hz). In the primary stage of the DWT, the signal is concurrently passed through an LP and HP filters. The approximation (A1) and detailed (D1) outputs from low and high pass filters are at the first level. Using the Nyquist rule, the output signal is down sampled by two which was holding half the bandwidth of original signal. The same process can be repetitively carriedout for the first level approximation and the detailed coefficients at the second level. In each step of decomposition, the frequency resolution is multiplied and down-sampled by splitting its time resolution. Since the sampling frequency used for EEG dataset is 173.61 Hz therefore, as per Nyquist rate, the utmost useable frequency should be half of the sampling frequency or 86.81 Hz[15].

EEG signal frequency are basically classified based on position of electrode on the scalp. In Signal processing, the normal and abnormal rhythm is determined by frequency units. Similarly , the EEG waveforms such as alpha, beta , gamma, theta and delta are also classified based on signal frequency. The continuous rhythm of brain waves is categorized by frequency bands and these frequency bands differ based on mental states of the brain. The below are the details of EEG frequency bands.

Delta wave(0.1-4Hz): the amplitude of Delta wave is highest but slow. This slowest wave is found in all the stages of sleep. It is normal in infants and abnormal in adults who are awake.

Theta wave ranges from (4-8Hz) which are associated to subconscious activity. These waves are observed in meditation and deep sleep.

Alpha wave (8-13Hz): These waves are seen in all age groups, generally in adults with closed eyes. It also represent the white matter of the brain.

Beta wave(13-30Hz): Beta wave are symmetrical on both side of frontal area of brain. These waves are low in amplitude.. This wave is observed when brain is engaged in mental activities.

Gamma wave(30-100Hz):gamma waves are associated with consciousness and perception. These waves occur in hyper alertness.

In order to determine these frequency bands a mechanism /tool knows as Discrete Wavelet transform (DWT) is used to find the occurrence of these bands and further features are obtained by decomposing using Daubechies Wavelet.

3.3 Feature Extraction

The below Fig 1 shows the basic methodology for determining the frequency bands of EEG signal. The input EEG signal is taken from Bonn database. The Dataset consist of both normal and epileptic seizure signal. A Sampling rate of 173.61 Hz. is used to sample the signal. MATLAB software is used for analyzing the input EEG signal. As wavelet transform is used for analyzing the signal both in time frequency domain, hence the input signal is decomposed to 7levels obtain the various filter coefficient and details vector values. As the EEG signal consist of many artifacts due to interferences, hence using this transformation we are able to obtain the various frequency bands of EEG signal and thus eliminating the unwanted signal frequencies. In the further analysis, we calculate the time frequency domain parameter of these EEG frequency bands like mean, standard deviation, entropy, energy, relative power and RMS.

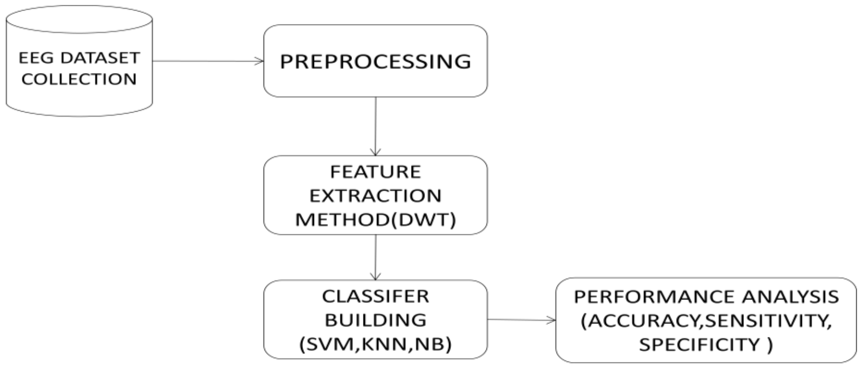


Fig 1: Proposed classification model

(1) The mean of wavelet filter coefficient in different sub band is calculated by

$$ui = \frac{\sum_{i=1}^{n} xi}{n} \qquad i=0,1,2,3,\ldots.n \tag{1}$$

(2) The standard deviation for each sub band is square root of mean give by

$$\sigma^2 = \frac{\sum_{i=1}^{n}(xi-x)^2}{n-1} \tag{2}$$

(3) The subband energy is obtained by

$$E = \sum_{n=0}^{N-1} |xi|^2 \qquad i=0,1,2,3,\ldots..n \tag{3}$$

(4) The Entropy for each sub-band is given by

$$Entropy\ (EN) = \sum_{i=1}^{N} xilog\ xi^2 \quad i=0,1,2,3,\ldots..n \tag{4}$$

(5) The Relative energy(RE) is calculated by

$$\rho i = \frac{Ei}{Etotal} \qquad i=0,1,2,3,\ldots..n \tag{5}$$

3.4 Classification

Several approaches and methods can be used for features extraction and classification of seizure. The differences between normal and abnormal cases are based on these extracted and selected features. The extracted features are given to a classifier model for seizure detection between normal and epileptic patient. In the proposed study we have used three classification model named as SVM, KNN & NB classifier to evaluate the performance proposed model.

Support Vector Machine [SVM]: Support vector Machine is most successful machine learning algorithm used in many engineering application. SVM is a classification tool which is used for classification of biomedical signals in detection of any abnormalities. SVM is considered as the best and common method for diagnosis. It handles high dimensional and non linear data very efficiently. The SVM classifier segregates the two classes i.e hyper plane or line. SVM algorithm is based on finding a hyperplane in N dimensional space which separates the data points.

K-Nearest Neighbor (KNN): is a non-parametric supervised learning method. This classifier type relies on distance for its classification. This technique assigns weights to contribution of the neighbor. The nearest neighbor is a new unknown variable which has to be classified is denoted by K. KNN is sensitive and east to implement..

Naive Bayes (NB): This classifier is a simplest class of classifier and also efficient statistical method. It is used in classification of text which include high dimensional training data set. Naive Baye classifier is a tool used as a classifiers developing models which assign a class label and represents feature values as vectors, the class labels are derived from finite set. Naïve Bayes models uses maximum likelihood method.

4. Results and Discussion

Extracting features of EEG signal is done in 2 steps. Initially, the decomposition of the input EEG signal is done using Discrete wavelet transform into five sub band signals with different frequency like alpha, beta, gamma, theta and delta. A appropriate decomposition level Db 7 is selected with proper wavelet function. The decomposition of the EEG signal is done with a sampling frequency of 173.6 Hz. These obtained wavelet filter coefficients further analyzed by selecting appropriate feature and computed using MATLAB (R2021). The below figure 2,3 and figure 3,4 shows the plot of Input EEG signal of Subject S001 and subject Z001 different EEG frequency bands respectively.

Fig 2: Input EEG signal of subject S001

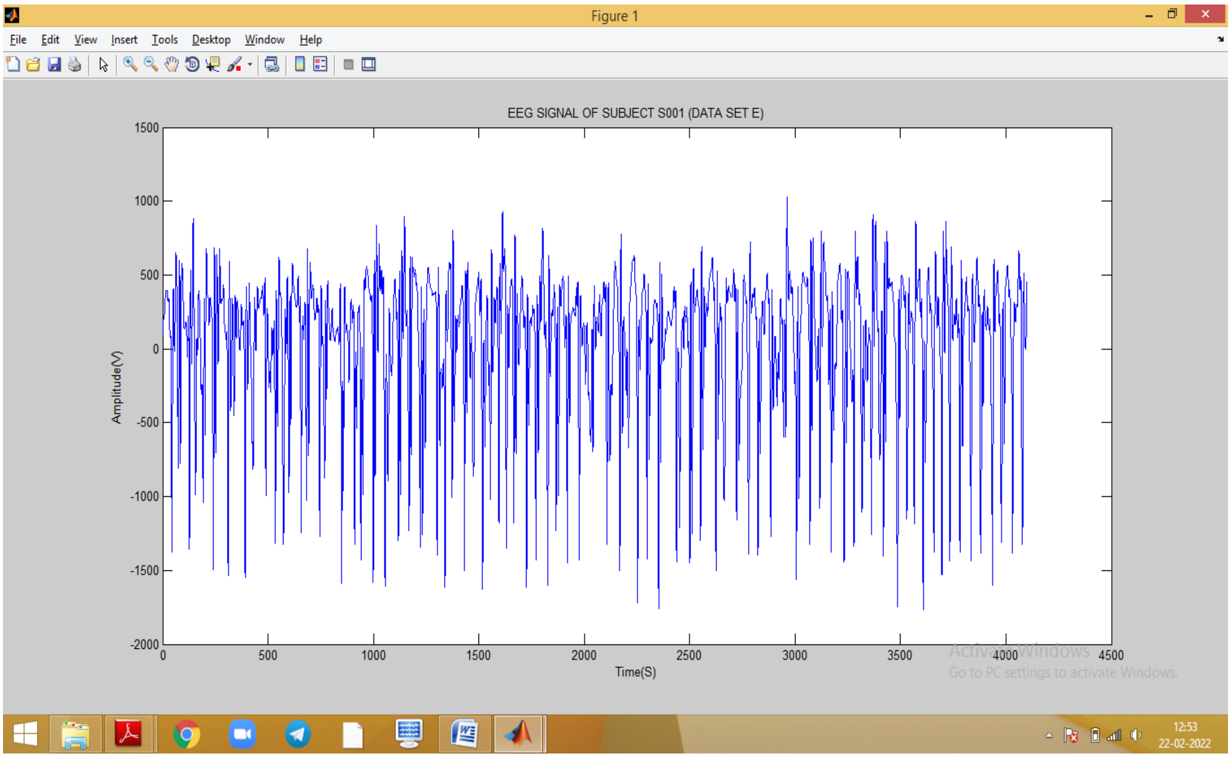


Fig 3: Obtained Frequency bands of subject S001

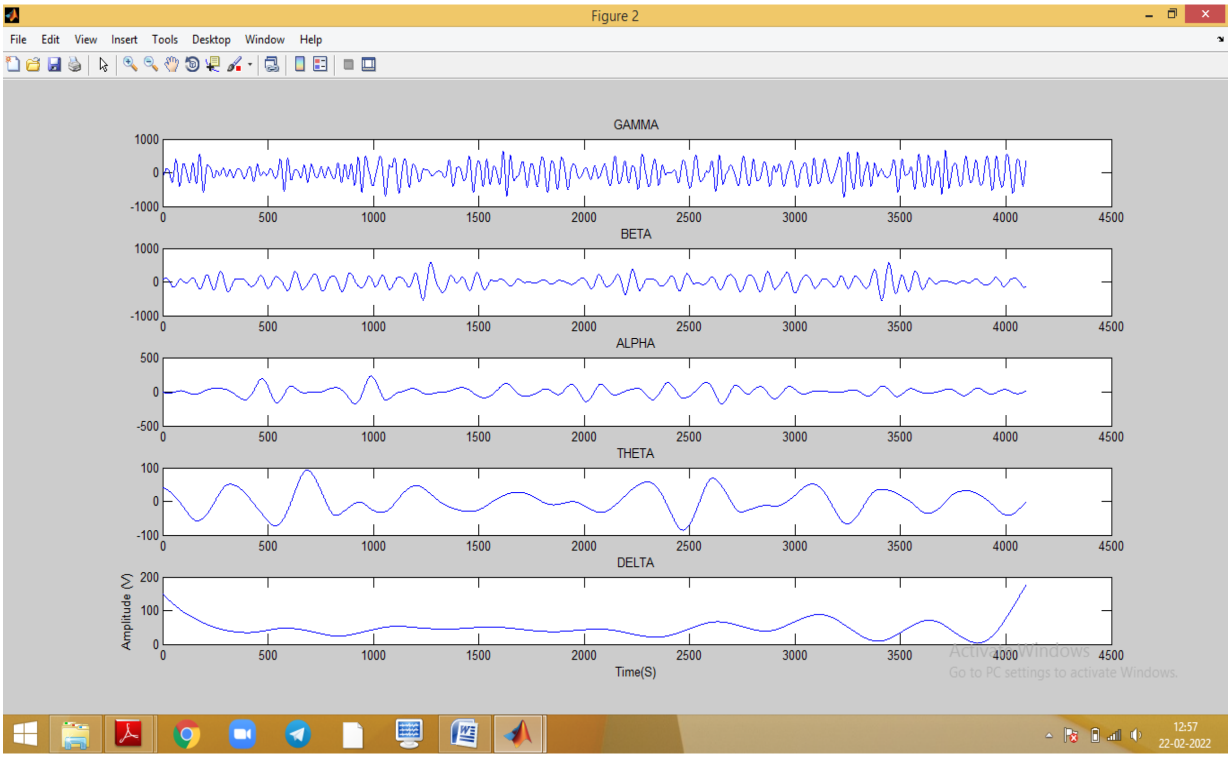


Fig 4: Input EEG signal of subject Z001

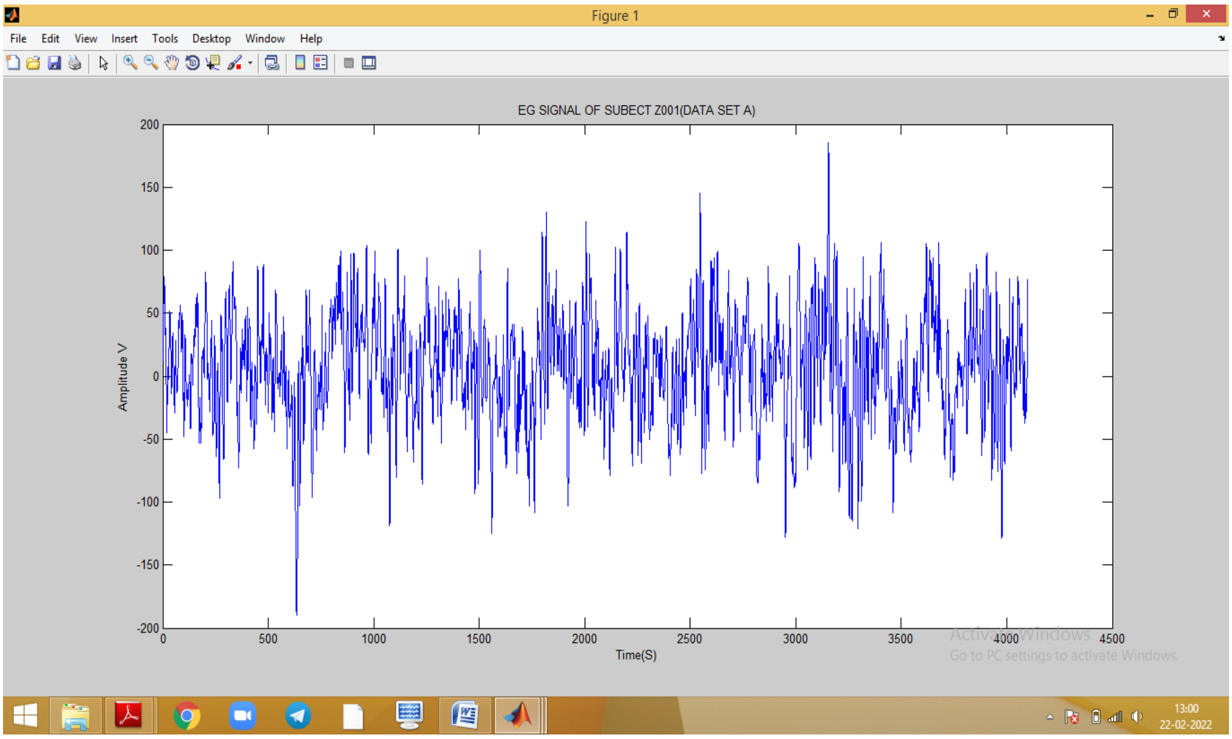
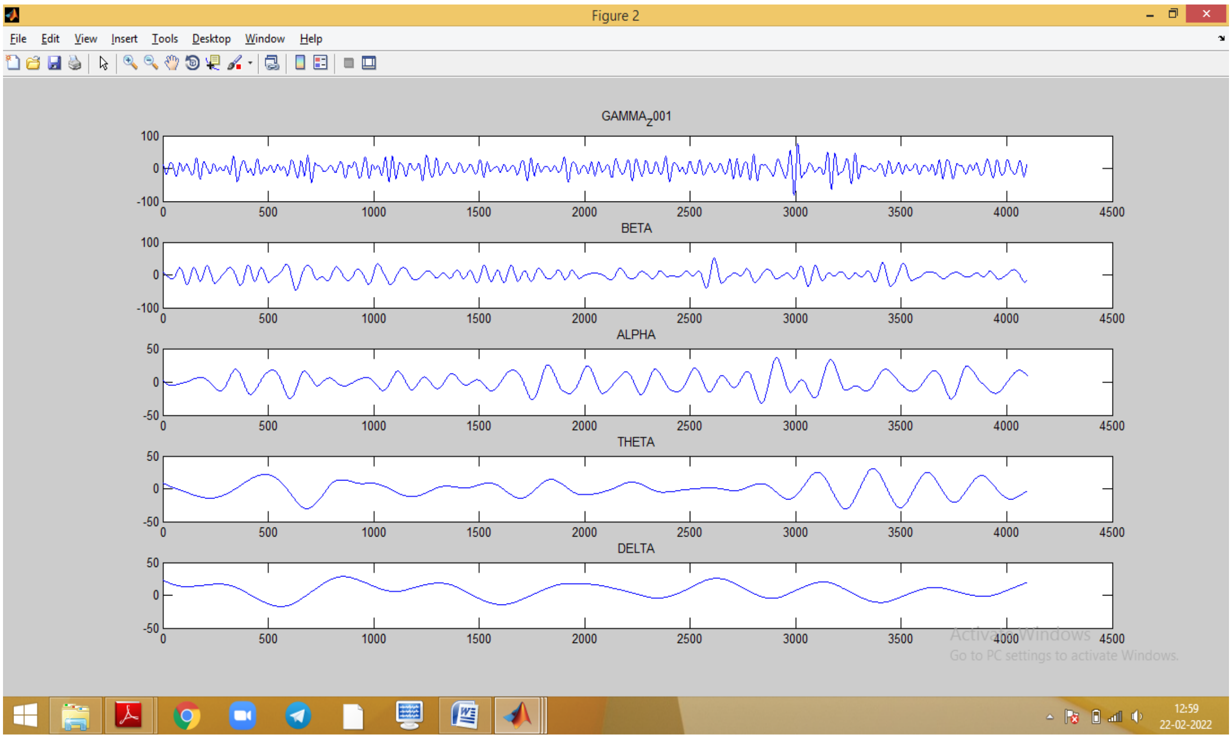


Fig 5: Obtained Frequency bands of subject Z001



In the next step, the input EEG signal is divided into five different sub-signals, and then statistical features values of each sub bands for both normal and epileptic signal like Mean, Standard Deviation, Energy, Entropy, Relative Powers and RMS values are obtained. The extracted features from each sub-bands of Dataset A and E are give as input to classifiers.

The Classifier performance in Seizure Detection is done based on specificity, sensitivity, and accuracy which are defined as[19]

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP+FN}} * 100 \quad (6)$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN+FP}} * 100 \quad (7)$$

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+FN+TN+FP}} * 100 \quad (8)$$

A sample test data from SET A and SET E is shown in table 4. These sample data shows the statistical parameter calculation of SET A and SET E for six different parameter. The classification description of test cases is shown in table 2 and the performance of the signal analysed is shown in table 3 which is tested with different classification model with respective to accuracy, specificity and sensitivity.

Table 1: A summary of clinical data used

Settings	SET A	SET E
Subjects	5 healthy	5 epileptic
Type of Electrode	surface	Intracranial
Placement of Electrode	International 10-20 system	Epileptogenic zone
Subject condition	Awake, eyes open	Seizure (Ictal)
Number of epochs/ Epoch duration (s)	100 23.6	100 23.6

Table 2: Classification description of test cases.

Test case	Case of Seizure	Classification Description
Case1	SET A VS SETE	Healthy Persons with eye open vs Epileptic patients during seizure activity

Table 3: The Performance Evaluation of Set A vs Set E using Different Machine Learning techniques

.Method	% in Accuracy	% Specificity	% Sensitivity
KNN	98.5	99	100
SVM	99.5	100	100
NB	98.3	98	98

5.Conclusion

Epilepsy seizure is the biological brain disorder which occurs by random changes in brain signals. The visual detection seizure in EEG wave is tedious and can be mistaken if the signal is of long duration in time. The seizures detection using supervised machine learning based methods for classification of EEG is studied from the literature. Therefore, in this proposed approach, Discreet Wavelet Transform is used as a tool to detect Epilepsy in EEG signal by decomposing input EEG signal into various sub bands to acquire the detailed and approximate wavelet coefficient and further different classifier like SVM ,KNN and NB is used and their attainment is calculated in terms of accuracy , sensitivity and specificity. It was seen that the finest classification accuracy is obtained for SVM with remarkable and convincing results than compared to other classification model.

Table 4: Sample test data

Patient Number	Statistical Parameters	Delta band	Theta band	Alpha band	Beta band	Gamma band
Patient 1 S001	Mean	48.42	-0.832	-0.366	-0.165	0.106
	Standard deviation	24.42	34.49	64.54	156.93	260.06
	Energy	1.025e+08	4.87e+06	1.706e+07	1.008e+08	2.77e+08
	Relative Power	0.089	0.036	0.126	0.748	2.054
	Entropy	0.451	1.552	1.125	1.001	1.055
	RMS	54.238	34.502	64.534	156.912	260.035
Patient 2 S002	Mean	40.796	-3.231	-0.249	-0.148	-0.0611
	Standard deviation	36.797	52.305	111.401	173.027	244.123
	Energy	1.236e+08	1.124e+07	5.083e+07	1.226e+08	2.441e+08
	Relative Power	0.062	0.057	0.257	0.622	1.238
	Entropy	0.529	1.066	1.084	1.044	1.021
	RMS	54.937	52.399	111.388	173.006	244.094
Patient 1 Z001	Mean	15.755	-3.449	0.141	0.211	0.136
	Standard deviation	38.417	52.079	51.013	106.510	221.215
	Energy	7.06e+06	1.11e+07	1.06e+07	4.64e+07	2.00e+08
	Relative Power	0.093	0.148	0.141	0.616	2.660
	Entropy	0.800	1.177	1.106	1.051	1.020
	RMS	41.518	52.187	51.007	106.498	221.188
Patient 2 Z002	Mean	-52.815	0.409	0.006	-0.063	0.030
	Standard deviation	10.704	10.744	13.855	16.083	16.378
	Energy	1.18e+07	4.73e+05	7.86e+07	1.059e+06	1.098e+06
	Relative Power	0.836	0.033	0.055	0.074	0.077
	Entropy	0	1.713	1.390	1.364	1.279
	RMS	53.889	10.750	13.854	16.081	16.376

References

[1] R. G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David,and C. E. Elger, (2001). "Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: dependence on recording region and brain state," Physical ReviewE, vol. 64,no. 6, Article ID061907, 8 pages,

[2] Mohammad Zuber, et al, (2017)" Denoising of EEG Signals For Analysis of Brain Disorders: A Review," International Research Journal of Engineering and Technology (IRJET), vol. 4, Issue: 07.

[3] Haslaile. Abdullah et al. (2013)., "Double Density Wavelet For EEG Signal Denoising," 2ndInternational Conference on Machine Learning and Computer Science, Kuala Lumpur,Malaysia, August 25-26.

[4] Geeta Kaushik et al. (2012)., "Biomedical Signal Analysis through Wavelets : A Review," International Journal Of Advanced Research in Computer Science And Sofware Engineering, vol. 2, issue 9, pp. 422-428, September,

[5] P.Ashok Babu, K.V.S.V.R Prasad (2011)., "Removal Of Ocular Artifacts From EEG Signals Using Adaptive Threshold PCA and Wavelet Transforms," International Conference on Communication Systems and Network Technologies, Katra, Jammu, June 3-5,

[6] Patil Suhas S.et al ,(2012)., "Wavelet Transform to Advance the Quality of EEG Signals in Biomedical Analysis," 3rdInternational Conference On Computing Communication Systems and Networking Technologies, Coimbatore, July 26- 28

[7] Priyanka Khatwani, Archana Tiwari (2013)., "A Survey on Different Noise Removal Techniques Of EEG Signals," International Journal of Advanced Research in Computer and Communication Engineering, vol. 2,issue 2, pp. 1091-95, February

[8] Weidong Zhou et al(2004)., "Removal of EMG and ECG Artifacts from EEG Based On Wavelet Transform and ICA," Proceedings Of the 26th Annual Conference Of The IEEE EMBS San Francisco, CA, USA, September 1-5,

[9]Sartori CAF, Sevegnani FX (2010). Fault classification and detection by wavelet-based magnetic signature recognition. IEEE Trans. Magn., 46(8):2880-2883.

[10]Dwivedi UD, Singh SN (2010). Enhanced detection of power-quality events using intra and interscale dependencies of wavelet coefficients. IEEE Transction. Power Deliv. 25(1):358-366.

[11] L. Guo, D. Rivero, J. Dorado, J. R. Rabuñal, and A. Pazos, "Automatic epileptic seizure detection in EEGs based on line length feature and artificial neural networks," Journal of Neuroscience Methods, vol. 191, no. 1, pp. 101–109, (2010).

[12] F. Pedregosa, G. Varoquaux, A. Gramfort et al., "Scikit-learn: machine learning in python," Journal of Machine Learning Research, vol. 12, no. 10, pp. 2825–2830, (2012)

[13] Lakshmi VIK, Xiaomin L, Usama Y, Ramakrishnan V, Narayan CK A Twofold Daubechies-wavelet-based module for fault detection and voltage regulation in SEIGs for distributed wind power generation. IEEE Trans. Ind.Electron.60(4):1638-1651. (2013).

[14] O. Faust, U. R. Acharya, H. Adeli, A.D. of Epileptology University of Bonn, Eeg time series data. URL http://www.meb.uni-bonn.de/epileptologie/science/physik/eegdata.htmlAccessedOct (2016)

[15] Asmaa Hamad, Essam H. Houssein, Aboul Ella Hassanien, Aly A. Fahmy. "Feature extraction of epilepsy EEG using discrete wavelet transform", 2016 12th International Computer Engineering Conference ICENCO), (2016)

[16] Aayesha Muhammad Bilal Qureshi & Muhammad Afzaal, Muhammad Shuaib Qureshi3 , Muhammad Fayaz "Machine learning-based EEG signals classification model for epileptic seizure detection", in Multimedia Tools and Applications, doi.org/10.1007/s11042-021-10597-6, Feb (2021). Online Springer.

[17] Tingxi Wen, Yu Du, Ting Pan, Chuanbo Huan and Zhongnan Zhang "A Deep Learning-Based Classification Method for Different Frequency EEG Data" in Computational and Mathematical Methods in Medicine Volume 2021, Article ID 1972662, 13 pages https://doi.org/10.1155/2021/1972662 (2021)

[18] Ibrahim Omerhodzic, Samir Avdakovic, Amir Nuhanovic, Kemal Dizdarevic, Kresimir Rotim. "Chapter 7 Energy Distribution of EEG Signal Components by Wavelet Transform", IntechOpen, (2012)

[19] Fathima, Thasneem, Yusuf U. Khan, M. Bedeeuzzaman, and Omar Farooq. "Discriminant Analysis for Epileptic Seizure Detection", 2011 International Conference on Devices and Communications (ICDeCom), (2011).