Journal of Biological Engineering Research and Review, 2021; 8 (2): 147-152
ISSN: 2349-3232 Online
Available online at www.biologicalengineering.in/Archive

Research Article

Detecting and Classifying Epileptic Seizures by Higher-Order Spectral Cumulants

Deepthi K, Shama B.N.

Department of Electronics and Communication Systems, St. Joseph Engineering College, Mangaluru

***Corresponding author. E-mail:** deepthirao950@gmail.com (**Deepthi K**)

ARTICLE INFO:

Article History:
Received: 12/08/2021
Revised: 05/10/2021
Accepted: 25/11/2021
Published: 31/12/2021

Keywords: Epilepsy, *Seizure, Electroencephalography, higher-order spectral, PCA*

Citation: Deepthi K, Shama B.N. Detecting and Classifying Epileptic Seizures by HOS Cumulants. Journal of Biological Engineering Research and Review. 2021, 8 (2): 147-152.

Abstract: Epilepsy is brain illness and characterized by brain's excessive weird activity. Long-term Electroencephalography (EEG) recordings of an epileptic patient contain a vast amount of EEG data. The requirement of testing the EEG's entire length for the epileptic prediction is a demanding process. If well-handled and diagnosed, 70% of epilepsy individuals can stay free from seizures. Two different classes of EEG data are taken for analysis purposes. The higher-order spectral (HOS) estimates are obtained by computing the cumulant after the fourth-order Butterworth filter and the Infinite Impulse Response (IIR) notch filter preprocessing processes. Principal Component Analysis (PCA) are used for extracting the characteristic and those characteristics are classified using the Decision Tree algorithm. MATLAB software is used as a tool to implement this proposed methodology. The training and testing the data during the classification is carried using the 10-fold cross validation. The final result indicates that the decision tree algorithm precisely classifies normal and seizure EEG signals with an accuracy of 98 percent and an average AUC of 0.9828. Classifying and detecting the two classes viz, normal and epilepsy from EEG signals is the extreme focus. It is clear that the feature chosen using principal component analysis to separate the two classes of EEG after computing the higher-order spectra of the EEG data was successful. Non-linear features taken from EEG segments are known to produce classifier results with high classification accuracies of more than 97 percent.

INTRODUCTION

The brain is a very composite part of our body. It comprises roughly 100 billion neurons or nerve cells. Excitation of these signals generates signals to the brain's different parts [13]. For other parts of the body, external stimulus and the process of physiological control transmit these signals. Hence EEG signals provide wealthy information about the electrical activity of the brain. Currently, for clinical and research purposes, EEG signals are chiefly used to detect the activity of various actions inside the brain. EEG signal generates a huge amount of data which is tricky to analyze by observation. They are having low amplitude because of the skull's composition. From the EEG signal, abnormalities recognition is done using computers. Sub-bands of EEG signal are gamma (30-40 Hz), beta (13-30 Hz), alpha (8-13 Hz), theta (4-8 Hz), and delta (0.5-4 Hz) [17]. The scalp's surface is arranged by electrodes for electrical activity recording. This arrangement is suggested by the International Federation of societies for electroencephalography and clinical neurophysiology [6].

Epilepsy is a neurological disorder categorized by frequent seizures as a result of brain anomalous electrical discharges. It is a continual neuron-disordered related situation that influences about fifty million people throughout the world, which makes Epilepsy a usual neurological illness globally in step with the report of World Health Organization (WHO), June 2019 [6]. Around 1-2% of inhabitants in the world are affected by this state. The sudden and feasibly unpredictable nature of seizures is one of the mainly disabling aspects of Epilepsy. Possibilities of epilepsy healing are possible by the Seizure occurrence prediction technique. Treatment concepts could move from preventative strategies (e.g., Long-term prescription with antiepileptic drugs) in the direction of an EEG- triggered on-demand therapy or other simulation in an attempt to reset brain dynamics to a state that will no longer build up into a seizure. The uncertainty of the onset of seizures is one of the utmost important reasons for morbidity and pressure in patients with epilepsy [19]. EEG is a controlling method that uses more than one electrode, which is placed alongside the brain scalp to degree electric attention of mind generated with the neocortex nerve cells [6]. The enlightening of epileptic seizures by visual scanning of a patient's EEG data usually collected over a few days is a droning and sustained process.

An epilepsy detection system is required for the following reasons: (1) Detecting tiny changes in EEG signals correctly through the manual process is a complex task; (2) As this process is dreary, the physician or clinician is unfeasible to monitor EEG signals continuously; (3) Having a skilled person who can diagnose the signals about a probable seizure is very difficult; (4) By the manual method, differentiating the normal and epileptic seizure signals is complex; and (5) The delicate changes in the amplitudes and intervals of EEG signals for a different dataset for more exactness are analysed through an automated or semi-automated system.

METHODS AND MODELS

Block Diagram

EEG data belonging from two different classes are taken for analysis purposes. The cumulant is computed and higher-order spectral (HOS) estimates are obtained. FIG I shows the block diagram. Features of the signals are extracted using Principal Component Analysis (PCA) and classified using the Decision Tree algorithm.

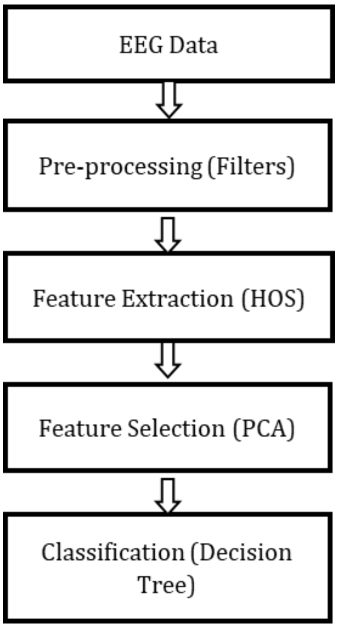


Fig. 1: Block Diagram

EEG Data and Pre-processing

Bonn University data is used for the revise of seizure event recognition. With 23.6 seconds duration segments of 100 channels, the entire dataset has 5 sets. Each channel TXT-file consists of 4096 samples of one EEG time series in ASCII code [12].

Table 1: overview of Bonn dataset

Set	Patients	Setup	Phase
A	Healthy	Surface EEG	Open eyes
B	Healthy	Surface EEG	Closed eyes
C	Epilepsy	Intracranial EEG	Interictal
D	Epilepsy	Intracranial EEG	Interictal
E	Epilepsy	Intracranial EEG	Seizure

The Pre-processing method is done as the obtained data is a raw signal. It is the primary processing of signal to formulate it for the prime processing or added analysis. Each sample is 23.6 seconds long, with data sampled at 173.61Hz and band-filtered between 0.5 to 50Hz using a fourth-order Butterworth filter.

The pre-processing is the way to filter delta (0.54Hz), theta (4-8Hz), alpha (8-13Hz), beta (13- 30 Hz) and, gamma (30-40Hz) frequency ranges. Infinite Impulse Response (IIR) notch filter is performed on the Butterworth filtered (0.5-50Hz) signal to remove a single frequency 50Hz noise signal. Pre-processing results in signals that are clear and of good quality used for further feature extraction methods.

Computation of HOS Cumulants

From the present ones, new features creation is possible by reducing the features using feature extraction.

Higher-order application has an intensifying interest in the past few years to an extensive range of signal processing and system hypothesis problems. This information is very valuable in problems where either non-Gaussian, non-minimum phase, or nonlinearities are considerable and should be accounted for. Deterministic signals are defined by moment and random processes by cumulant of HOS [20]. The autocorrelation function or sequence of a fixed process, x (n), is defined by:

$$R_{xx}(m) = E\{x(n)x(n-m)\} \quad (1)$$

Where E {·} denotes the ensemble expectation operator.

The power spectrum is officially defined as the Fourier Transform (FT) of the autocorrelation sequence.

$$P_{xx}(f) = \sum_{-\infty}^{\infty} R_{xx}(m)\exp(-j2\pi fm) \quad (2)$$

Where f denotes the frequency. An equivalent definition is given by

$$P_{xx}(f) = E\{X(f)X*(f)\} \quad (3)$$

Autocorrelation normal generalization is the moments, and their nonlinear combinations are cumulants.

The first-order cumulant of a static method is the mean,

$$C_{1x}: = E\{x\ (t)\} \quad (4)$$

The second and third-order cumulants of a zero-mean stationary process are defined by

$$C_{2x}(k) = E\{x(n)x(n+k)\} \quad (5)$$

$$C_{3x}(k.l) = E\{x(n)x(n+k)x(n+l)\} \quad (6)$$

In practice, a finite amount of data must obtain consistent estimates of cumulants.

$$C_{xyz}(k,l) = \frac{1}{N}\sum_{N_1}^{N_2} x(n)\, y(n+k)z(n+l) \quad (7)$$

Principal Component Analysis

From the raw data, the best features are chosen to abolish EEG redundancy for extracting features. The process of translation from a set of correlated variables to a set of uncorrelated variables through orthogonal alteration is mathematically termed PCA [9]. After normalizing, correlation matrix decay is performed in the PCA. PCA is the multivariate examination based on eigenvector [1].

Steps to implement PCA:
Step 1: Normalize the data
Step 2: Covariance matrix computation
Step 3: Determine the eigenvalues and eigenvectors
Step 4: Principal components selection
Step 5: Feature vector formation
Step 6: Principal Component's configuration

PCA can deliver the user lower-dimensional data from high-dimensional one.

Decision tree

In the classification stage, all the selected features will be specified to a classifier. A Decision Tree is the most influential and popular implement for classification and prediction [1]. A Decision Tree is a flowchart-like tree arrangement, where each inner node denotes a quality test, each branch represents a result of the test, and every terminal node holds a class label. Subsets were formed based on a quality value test by splitting the origin set.

This method is reiterated on each resultant subset in a recursive manner called recursive partitioning.

k-fold cross-validation
The method of cross-validation with 10-fold is applied during classification to train and test the data. In this process, the complete dataset is divided into 10 non-overlapping subsets such that a nearly equal number of data from each class belongs to each fold. One of the ten folds is used for testing and the remaining nine subsets were together used for training the classifier for each fold of classification.

Classifier performance in the detection of Seizure Event
Classifiers are trained (learned) on a fixed training multiset. A learned classifier has to be tested on a dissimilar test set experimentally. The classifier executes on different data in the run mode that on which it has learned. Then the label is predicted based on the classification tree and true testing data.

Confusion Matrix

The Confusion matrix is one of the most instinctive and at ease metrics used for discovering the accuracy and correctness of the model. The confusion matrix is two-dimensional ("Actual" and "Predicted") table and has groups of "classes" in both dimensions. Actual labels or classifications are rows and Predicted ones are columns.

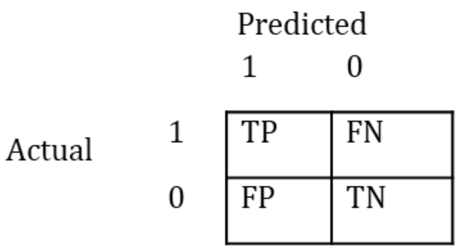


Fig. 2: Confusion matrix diagrammatic representation

Figure 2 shows the TP, TN, FP, and FN metrics of the confusion matrix.

Receiver Operating Curve (ROC) and Area Under the Curve (AUC)

AUC-ROC graphs are to figure the classifiers and to picture their performance. It is a 1-specificity versus sensitivity curve.

RESULTS AND DISCUSSION

The proposed methodology on the two-class EEG data is implemented in MATLAB.

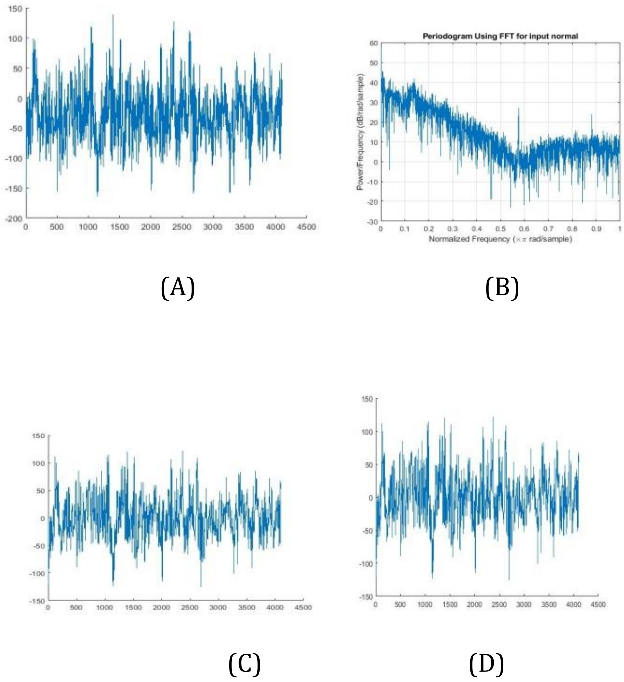


Fig. 3 Normal EEG (A), its periodogram (b), Butterworth filtered (C), and notch filtered (d)

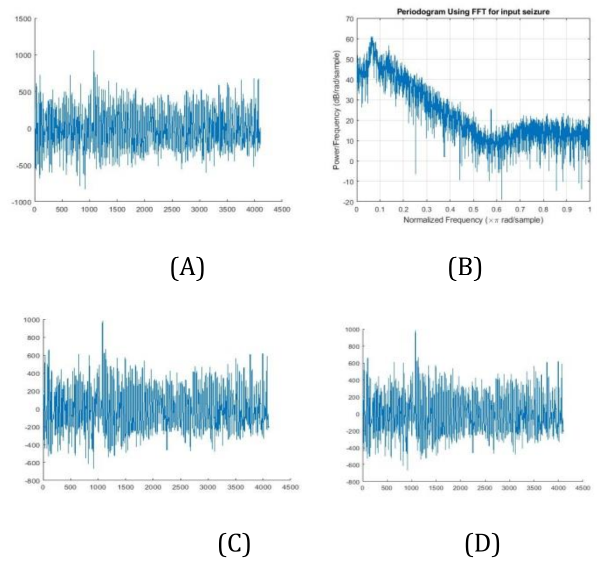
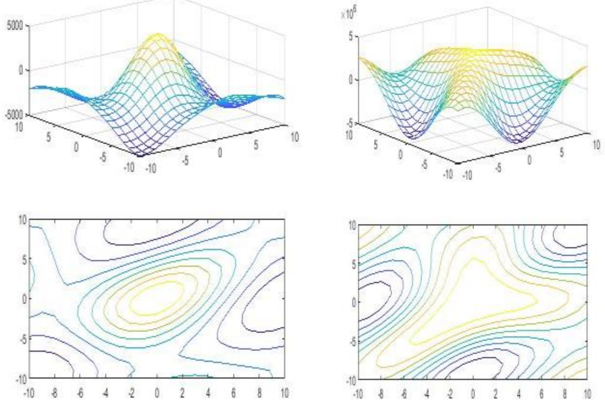


Fig. 4: Seizure EEG (A), its periodogram (b), Butterworth filtered (C), and notch filtered (d)

FIG III and IV show the normal input and Seizure EEG signal and their periodogram. A periodogram is a graphical data analysis method for revising frequency-domain models of an equispaced sequence.

To eliminate noise Fourth order Butterworth filters between 0.5- 50Hz are applied to the input seizure and normal EEG and the frequency ranges alpha, beta, theta, delta, and gamma are obtained, and then the IIR notch filter is performed on the Butterworth filtered signal. The Butterworth and Notch filtered normal and Seizure signal is also shown in FIG III and IV.

The input signal frequency is 173.61 Hz. The filtered signals magnitude levels are decreased and its frequency ranges are between 0.5 to 50 Hz. The power line interference at 50Hz frequency is suppressed from the filtered signal.



(A) (B)

Fig. 5: A typical bi-spectrum of normal (A) and Seizure (B) EEG. (3-D and contour plot)

Time series filtered EEG data is segmented into records of 512 samples each, with 50% overlap. Unbiased estimates of the third-order cumulant are obtained from each segment and then averaged.

The contour plot reviews the basic symmetry of third-order cumulants. FIG V is the cumulant and 3D plot.

The features are selected using PCA from the cumulant calculated.

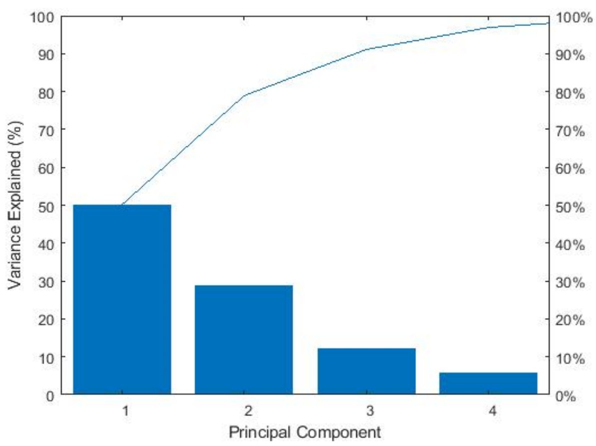
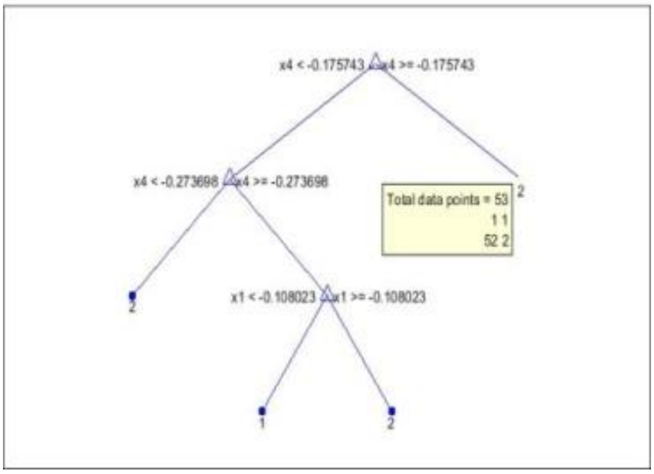


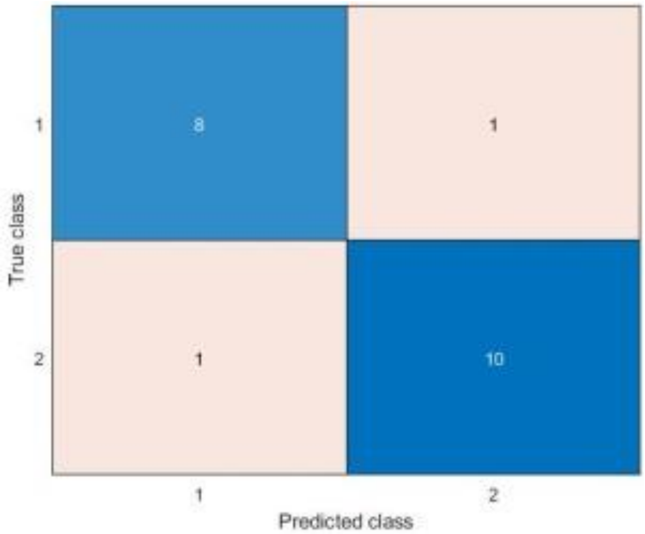
Fig. 6: Principal component's variance explanation graph

FIG VI is the graphical depiction showing the principal components and each principal component variance explanation. It is evident that the first 5 principal components elucidate 99 percent of features.

In the classification stage, the first five principal components which explain 99% of variance are given to a classifier.



(A)



(B)

Fig. 7: Decision tree (a) and Confusion matrix (b)

FIG VII (A) is the classification Decision tree. Here the number of nodes is seven. In first or a root node it takes the variable x4 and makes the decision such that it makes the branches and split into two nodes, called leaf nodes. If data or variables are not present then this process of spitting stops and based on the classifier decision different classes are assigned with their data. True or real class is the rows and the predicted class is the columns of the confusion matrix. FIG VII (B) includes all TP, TN, FP, and FN. Here 8 out of 9 data from class 1 data are classified correctly and the remaining 1 data is classified as FN that is class 2. 1 data of class 2 out of 11 is erroneously classified as class 1 that is FP and the other 10 are classified correctly as class 2.

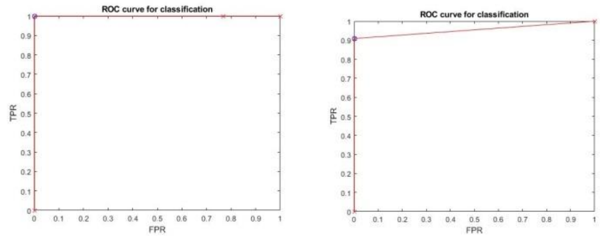


Fig. 8: ROC curves for decision tree

ROC curve and average AUC of the Decision tree is shown in FIG VIII.

Table: 2 classification results using decision tree

Classifier	Decision Tree
Accuracy	98
Precision	97.8889
Sensitivity	97.9798
Specificity	98.1818
F Score	97.8864
AUC	0.9828

The outcome implies that the decision tree algorithm precisely classifies normal and seizure EEG signals with an accuracy of 98 percent and an average AUC of 0.9828. By Table II, the algorithm has a precision, sensitivity, specificity, and F Score of about 97.8889, 97.9798, 98.1818, and 97.8864 percent respectively.

The test error of a predictive model or the cross-validation error is estimated as 0.0111 whereas the in-sample classification error or their substitution loss is 0.0056. 50% of the data are from class1 and the other 50% are from class 2 as calculated from the 50% Prevalence rate. So, there are an equal number of observations present in both classes.

CONCLUSION

EEG signals can be efficiently used to study the mental statuses and ailments associated with the brain. The EEG signals are nonlinear and their visual analyses are tedious. Here the extreme focus is on detecting and classifying the two classes viz, normal and epilepsy from EEG signals. It is evident that the feature selected using Principal component analysis after calculating the higher-order spectra of the EEG data effectively distinguished two classes of EEG. It is realized that the practice of nonlinear features extracted from EEG segments in classifier outcomes in high classification accuracies of more than 97%.

Though the proposed methodology affords adequate results, more studies on multiclass classification can be done by using better nonlinear features, different databases, and robust classifiers. This process can also be done for the remaining three sets of the same dataset. Deep learning concepts like Neural Networks, Convolution Neural Networks, and Artificial Neural Networks may also play important roles in this process.

In the proximate future, the following concerns need to be addressed for precise seizure detection and prediction.

1. Future may have more channels in the technologies that capture the EEG: It needs modern techniques which can exploit inter-channel connection for better detection and prediction.

2. Captured EEG signals have other signals interventions: EEG signals may have interferences from other signals produced from movable electronics devices. It contains diverse line noise and artifacts. Those noise characteristics are inspected by dissimilar systems to remove noise.

3. Wireless signals and wired signals: EEG signals can be captured through wired and wireless procedures. Investigating the required characteristics and unwanted noise in those signals is important.

Manipulating another classification algorithm

Using other classification algorithms like k-nearest neighbor (k-NN), support vector machine (SVM), Naïve Bayes (NB), Radial Basis Function (RBF), Artificial neural network and comparing the performance of these algorithms and thereby applying an improved algorithm if obtainable.

FUNDING

N/A

CONFLICT OF INTEREST

The author declared no potential conflicts of interest concerning the research, authorship, and publication of this article.

DECLARATION OF COMPETING INTEREST

The author declared no financial competing interests.

REFERENCES

1. Bingbing Wang, Zeju Xu, Tong Luo, et al., "EEG-based Closed-Loop Neurofeedback for Attention Monitoring and Training in Young adults", Journal of Healthcare Engineering, vol. 2021, Article ID 5535810, 13 pages, 2021. https://doi.org/10.1155/2021/5535810.
2. Wu Wen, "Sleep Quality Detection Based on EEG Signals Using Transfer Support Vector Machine Algorithm", Front. Neurosci., 23 April 2021.
3. Saminu, S.; Xu, G.; Shuai, Z.; et al., " A recent InvestigATION ON Detection and Classification of Epileptic Seizure Techniques Using EEG Signal, Brain Sci,2021, 11, 668. https://doi.org/10.3390/brainsci 11050668.
4. Si, Y., "Machine learning applications for electroencephalograph signals in epilepsy: a quick review", Acta Epileptologica, 2, 5 (2020); doi:10.1186/s42492-020-00014-0.
5. Mohammad Khubeb Siddiqui, Ruben Morales-Menendez, Xiaodi Huang, et al., "A review of epileptic seizure detection using machine learning classifiers", Brain Informatics, 7(1), 1-18, 2020.
6. Xuyang Zhao, Jordi Sole-Casals, Binghua li, et al., "Classification of Epileptic IEEG Signals by CNN and Data Augmentation", ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and, Signal Processing (ICASSP), 926-930, 2020.

7. Jefferson Tales Olivia, Joao Luis Garcia Rosa, "Classification for EEG report generation and epilepsy detection", Neurocomputing, 335, 81-95, 2019.
8. Poonam Sheoran, J.S. Saini, "Time-time analysis of electroencephalogram signals for epileptic seizure detection", International Journal of Biomedical Engineering and Technology, Vol.30 No.2, pp.113 – 131, 2019.DOI: 10.1504/IJBET.2019.100554.
9. Sreelekha Panda, Satyasis Mishra, Manas Ranjan Senapathi, "Detection and Classification Methods for EEG Epileptic Seizures", International Journal of Advanced Trends in Computer Science and Engineering, 8(6), 2925-2934, December 2019;doi:10.30534/ijatcse/2019/40862019.
10. Xiashuang Wang, Guanghong Gong, and Ni Li, "Automated Recognition of Epileptic EEG States Using a Combination of Symlet Wavelet Processing", Sensors, 19(2), 219,2019; doi:10.3390/s19020219.
11. Alexander Brenner, Ekaterina Kutafina, Stephan M Jonas, "Automatic recognition of Epileptiform EEG abnormalities", Studies in health technology and informatics, 247, 171-175, 2018.
12. Yang Y, Zhou M, Niu Y, et al., "Epileptic Seizure Prediction Based on Permutation Entropy", Frontiers computational neuroscience, 12,55, 2018.
13. Mengni Zhou, Chwng Tian, Rui Cao, et al., "Epileptic seizure detection based on EEG signals and CNN", Frontiers in neuroinformatics, 12, 95, 2018.
14. A. Ghosh, A. Sarkar, T. Das, et al., "Pre-ictal epileptic seizure prediction based on ECG signal analysis," 2nd International Conference for Convergence in Technology (I2CT), Mumbai, pp. 920–925, 2017.
15. U. Rajendra Acharya, Shu Lih Oh, Yuki Hagiwara, et al., "Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals", Computers in Biology and Medicine, September 2017; DOI: 10.1016/j.compbiomed.2017.09.017.

About Authors

Deepthi K. is pursuing her second year M.Tech in Department of Electronics and Communication systems, St. Joseph Engineering College, Mangaluru.

Shama B.N. is working as Assistant Professor in Department of Electronics and Communication Systems, St. Joseph Engineering College, Mangaluru