Research Article

Copy Right@ Mohamed M Dessouky

Rapid Seizure Classification Using Feature Extraction and Channel Selection

Athar A Ein Shoka¹, Mohamed M Dessouky^{1*}, AS El-Sherbeny² and Ayman El-Sayed¹

¹Department of Electronic Engineering, Computer Science and Engineering, Menoufia University, Egypt

*Corresponding author: Mohamed M Dessouky, Department of Electronic Engineering, Computer Science and Engineering, Menoufia University, Menouf, Egypt

To Cite This Article: Mohamed M Dessouky, Rapid Seizure Classification Using Feature Extraction and Channel Selection. 2020 - 7(3). AJBSR. MS.ID.001149. DOI: 10.34297/AJBSR.2020.07.001149.

Received:

☐ January 28, 2020; Published:
☐ February 18, 2020

Abstract

The Seizure is an abnormal electrical activity in the brain; it can be diagnosed by a neurologist and could be classified using recorded data. Medical data, such as EEG signal usually contain many features and attributes that are not important for the classification process. Dimension reduction is an important step to reduce irrelevant information. Features extraction is one algorithm for dimension reduction step. Another one is the channel selection algorithm. These algorithms speed up the process of classification and improve accuracy. This paper proposes an approach based on extracting EEG features, channel selection to reduce the computation capacity, and trained model used for classification. Variance parameter is used for channels selection, by taking the maximum three ones. Eleven features are extracted from the selected channels and averaged to be the input for the classifier. Six classifiers are used to select the most accurate one. Ensemble classifier was the more accurate one to classify all seizure cases correctly as it is got 100% sensitivity for continuous testing and 97.6% for the random testing set.

Index Terms: Seizure, Epilepsy, Electroencephalography (EEG), Features extraction, Channel selection, Seizure Classification

Background

Epilepsy is the most common neurological condition affecting humans of all ages. Due to the unpredictable characteristics of epilepsy, this makes it an understood neurological disease. There are an anticipated 50 million human beings with epilepsy in the world, up to 75% live in corrupt countries with little or no get entry to scientific services or treatment. Epilepsy is described by recurrent seizures related to sudden sporadic neuronal releases in the cerebrum and can result in other problems, even death [1,2].

Electroencephalography (EEG) is one of the primary modalities regularly utilized for remote epileptic seizure recognition. It becomes an inexpensive and noninvasive stage to investigate the inconspicuous qualities of the disease. The Seizure is the characterizing property of epilepsy, which reflects abnormal periods of activity in the EEG [2,3]. Features extraction is a critical step in epilepsy detection. Its importance lies in building an accurate model for classification. It is profitable to restrict the number of information in a classifier to have models with less computation [4]. There are

many techniques used for features extraction like Auto-Regressive (AR) [5], Principle Component Analysis (PCA) [6], Empirical Mode Decomposition (EMD) [7], and Statistical features technique [8]. A statistical feature had been used for extracting features in several algorithms [9,10].

Full-channel EEG signals recorded using electrodes. Their number varies from 18 to 23 electrodes on the scalp is neither wearable nor computationally effective. The process of channel selection is used to eliminate calculations, particularly in real-time applications. With the channel selection, it aims to decrease the number of channels not including distinguishing information [11,12].

Several algorithms used the concept of channel selection, such as; combining the advantages of both feature enhancement and channel selection to progress the detector performance [13]. Another algorithm compares electrode montage reduction by using only nine electrodes instead of using all electrodes [14]. Different algorithm selected EEG channels to eliminate power consumption

²Department of Electronic Engineering, Industrial Electronics and Control Engineering, Menoufia University, Egypt

in the detection process without affecting accuracy [15], and variance, the difference invariance, entropy, random selection and extra focal channels, and doctor's choice are also used and resulted from a valid range [16].

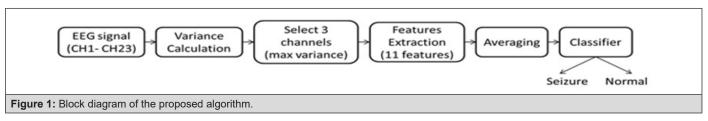
Classification is the step of identifying groups or classes based on similarities between them. This step is essential in this proposed approach to differentiate between seizure itself –ictal- and normal non-ictal periods. Classification involves two main steps: First step, the dataset information or concepts are grouped into two classes (seizure and normal) to learn the model. The second step, the model of the previous step is used for classification [17].

Several algorithms are used for this task such as Artificial Neural Network (ANN) [18], Support Vector Machine (SVM) [17,19], Linear Discriminant Analysis (LDA) [20,21], K-nearest neighbors (KNN) [22,23], decision tree [24], ensemble [25] and logistic regression [26]. This paper is arranged as follow: section II describes

the proposed algorithm. The EEG dataset is presented in Section III. Section IV lists the performance metrics evaluation. The results are presented in section V. Section VI shows the conclusion, and the paper ends with the references.

Proposed Approach

The main target of this work is to detect the epileptic seizures from EEG recording automatically. EEG records the ictal and non-ictal cases, the proposed approach detects every seizure states and it acceptable to classify the normal state to be a seizure in limited instances, and this makes the assurance of avoiding any complications while the opposite case is not fair. The proposed approach fundamentally relies upon consequently recognizing epilepsy from a brief period EEG recording. The proposed method is work as indicated by the following four stages: channel selection, feature extraction, averaging and classification as shown in (Figure 1).



The EEG signal is composed of 23 channels generated from electrodes, which attached on the scalp. These channels make calculations complex and increase the load on the system. Due to these limitations, the channels selection step is extensively essential. Once gutters are selected, this is the time to extract features. At this point, the averaging step is taking place for feature reduction. Finally, the features will be utilized to prepare the classifier. The prepared classifier will be used to test new cases to evaluate the performance of the ready classifier.

Variance Channel Selection

The channel selection step is intended to choose the most affected channels by seizure. Only one feature is used to be calculated for all channels, according to this feature a channel would be selected. Then the other features would be calculated for only the selected ones.

The simple method for selecting channels for features extraction and classification is the variance of EEG signal amplitude:

$$V_{ict}(c) = \frac{1}{k} \sum_{i=1}^{k} (X_c(i) - \mu_c) 2$$
 (1)

Where c is the channel, \mathbf{X}_c is training seizure data, μ_c is the mean of training seizure data, \mathbf{k} is the number of samples of training seizure data. The channels selection based on the highest values of $V_{ict}(c)$.

$$chosen channel = max(V_{ict}(c))$$
 (2)

Feature Extraction

In this step, a group of features of the EGG signal are extracted. These features are most influential in the form of the signal. From their synthesis, a classification model would be trained to classify the case as a seizure or not. They are extracted in ten seconds from EEG recording; these features are Standard Deviation, Mean, Variance, Median, Kurtosis, Skewness, Entropy, Moment, Power, Maximum and Minimum of the EEG signal.

Averaging

Each channel produces 11 features, so the input for each model would be 11×3 for each case, which would affect the real-time calculations and may prolong the classification time. Averaging the values of extracted features would reduce the number of input nodes of the model, and then eliminate the processing load and classification time.

Classification

Extracted averaged features would be used in this step to build the classification model. Different six models are created in this work to use the best model as a classifier. The used classifiers models are Support Vector Machine (SVM) which has the following characteristics; speed is medium, memory usage is also medium, and interpretability is easy. K-Nearest Neighbors (KNN); medium

for speed and memory usage and hard for interpretability. Decision Tree; speed is fast, memory usage is small, and interpretability is easy. Ensemble and Linear Discriminant; quick, short and easy. Logistic Regression; fast, medium and smooth.

Dataset Description

This work is performed using the CHB-MIT EEG dataset, CHB-MIT scalp EEG database became gathered on December 2010 at the children's hospital of Boston. The dataset incorporates 23 cases for 22 patients. Patients in the dataset contain five adult males with ages from three to 22 years old and 17 female with a while from 1.5 to 19 years old. The dataset has a sampling rate of 256 samples per second with 16-bit resolution [27].

Evaluation Metric Parameters

After finishing the training of the different classifier, each classifier is tested using 30 samples (containing both seizures and normal). The results are evaluated using accuracy, Sensitivity (True Positive Ratio - TPR), Specificity (True Negative Ratio - TNR), False Positive Ratio (FPR) (Fall-out), False Negative Ratio(FNR) (Miss-Rate), positive predictivity ($P_{\rm pr}$), and F1 score [9,28].

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \times 100\%$$
 (3)

$$sensitivity = \frac{TP}{TP + FN} \times 100 \tag{4}$$

$$specificity = \frac{TN}{TN + FP} \times 100\%$$
 (5)

$$fall - Out = \frac{FP}{TN + FP} \tag{6}$$

$$Miss - Rate = \frac{FN}{TP + FN} \tag{7}$$

$$positive \ predicticity = \frac{TP}{TP + FP} \tag{8}$$

$$F_1 = 2 \times \frac{p_{pr} \times sensitivity}{p_{pr} + sensitivity} \tag{9}$$

Experimental Results

In this work, the model goes under two phases, one is the training phase, and the other is the test phase. In the training phase, the proposed model is trained using massive data (250 samples). The training phase is applied to the data after channel selection step, and the results are estimated for these cases.

In the test phase, the proposed model is evaluated in two cases. The first case, the test is carried out on random dataset (80 samples / each sample is 10 Sec period -2560 points for each channel), and the other is carried out on continuous dataset (82 samples). The continuous data is taken from patient number 3. The obtained results are shown in the tables below.

The proposed algorithm primarily goes through the channel selection stage, this selection is performed using the variance calculation of all channels, and this is illustrated in (Table 1).

Table 1 shows an example for one sample from patient 1, at hour 21 (from 387 to 397 sec). Then the algorithm selects the three channels with maximum variance, then extract all other features (11 features) for these selected channels only as shown in (Table 2).

Table 1: Variance	Calculation for	Channel Selection.
-------------------	-----------------	--------------------

Table 1. Valiano	Table 1. Variance Calculation for Chainer Selection.										
CHn	VAR	CHn	VAR	CHn	VAR						
CH1	1.44E+05	СН9	7.23E+04	CH17	7.22E+04						
CH2	2.14E+05	CH10	4.09E+04	CH18	6.64E+04						
СНЗ	9.76E+04	CH11	2.97E+04	CH19	9.76E+04						
CH4	8.92E+04	CH12	4.10E+04	CH20	7.76E+04						
CH5	1.37E+05	CH13	8.32E+04	CH21	1.56E+05						
СН6	1.51E+05	CH14	4.79E+04	CH22	2.02E+04						
CH7	4.16E+04	CH15	8.25E+04	CH23	8.25E+04						
CH8	7.67E+04	CH16	3.62E+04								

Table 2: Selected Channels with Extracted Features.

Table 21 Colocted Charmole Will Extracted 1 Catalog.											
STD	MEAN	MAX	MIN	VAR	MED	SKW	ENT	KRT	МОМ	POW	CHn
463.1	1.61	1095	-1385	2.14E+05	17	-0.11	4.7	2.35	1.08E+11	1.41E+12	CH2
395.41	9.95	1279	-1295	1.56E+05	-25	0.26	4.47	3	7.33E+10	1.03E+12	CH21
388.61	4.13	941	-1293	1.51E+05	42	-0.4	4.84	3.01	6.86E+10	9.90E+11	СН6

Then the features extracted of the selected channels are averaged to minimize the calculations as in (Table 3). These steps are repeated for all input data (250 samples), to prepare the data for

training the model. Some examples of the training samples are shown in (Table 4).

Table 3: Averaged Values of Extracted Features.										
STD	MEAN	MAX	MIN	VAR	MED	SKW	ENT	KRT	мом	POW
415.71	5.23	1105	-1324.33	1.74E+05	11.33	-0.08	4.67	2.79	8.34E+10	1.14E+12

Table 4: Sam	Table 4: Samples of Results of Features Extraction in the Training Phase										
STD	MEAN	MAX	MIN	VAR	MED	SKW	ENT	KRT	мом	POW	state
415.7081	5.229598	1105	-1324.33	173944.2	11.33333	-0.08272	4.669139	2.788487	8.34E+10	1.14E+12	s
467.4201	-9.18222	1786.667	-1309	218865.1	-20.3333	0.362146	4.292831	3.875785	1.96E+11	1.44E+12	s
366.6224	12.58714	1046.333	-1296.67	134524.1	17	-0.21187	4.648266	3.686893	6.70E+10	8.83E+11	s
241.3414	-1.70558	1694.333	-967	59063.96	-2.33333	1.505989	3.675731	15.32168	6.66E+10	3.87E+11	n
254.11	-3.04178	1351.333	-1346.33	65207.83	1.666667	-0.10194	4.134625	7.827703	3.14E+10	4.28E+11	N
303.9874	0.370949	1477.667	-1665	92876.37	1.333333	0.2845	4.067168	7.142509	6.06E+10	6.09E+11	n
311.1377	-0.80893	1654.667	-1582	97136.49	1.666667	-0.33156	3.97244	8.298503	8.07E+10	6.37E+11	n
347.6363	-3.59677	955	-1061	123772.3	5	0.019804	4.49983	3.803735	6.33E+10	8.12E+11	s
277.2208	0.704282	841.6667	-798.667	76893.41	-1.66667	0.175046	4.582805	3.129974	1.87E+10	5.04E+11	s

Before starting the model training, an important data observation step should be performed, the cases are plotted with any selected parameter, the mean is selected, and the result is as shown in (Figure 2).

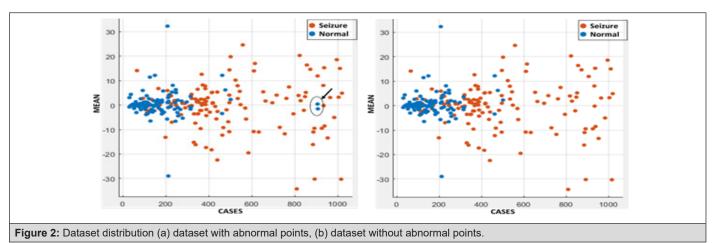


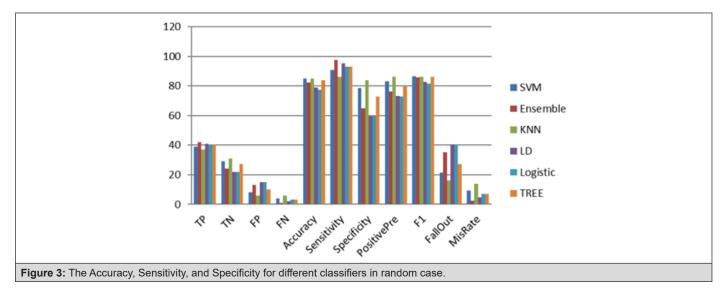
Table 5: Training Result Accuracy Before and After Filtering. Accuracy [un-filtered data] Accuracy [filtered data] Model SVM 81.3 83.2 Ensemble 80.1 82.8 KNN 83.3 84.8 LD 81.3 82.8 80.9 Logistic 84.8 TREE 78.1 82.4

As it is clear from Figure 2, there are two extremely abnormal points noticed, which affects the accuracy of the model at this stage as in (Table 5). These points are inpatient 4, 12 hours 5, 32 at 510

to 520. These points are filtered from the training data, and now the data was ready for the training phase.

Table 6: The Results for Different Classifiers in Test Phase in Random Case.									
	SVM	Ensemble	KNN	LD	Logistic	TREE			
TP	39	42	37	41	40	40			
TN	29	24	31	22	22	27			
FP	8	13	6	15	15	10			

FN	4	1	6	2	3	3
Accuracy	85	82.5	85	78.75	77.5	83.75
Sensitivity	90.69767	97.67442	86.04651	95.34884	93.02326	93.02326
Specificity	78.37838	64.86486	83.78378	59.45946	59.45946	72.97297
PositivePre	82.97872	76.36364	86.04651	73.21429	72.72727	80
F1	86.66667	85.71429	86.04651	82.82828	81.63265	86.02151
FallOut	21.62162	35.13514	16.21622	40.54054	40.54054	27.02703
MisRate	9.302326	2.325581	13.95349	4.651163	6.976744	6.976744



After model training, it should be evaluated by other data to ensure the accuracy and performance of the model. In this phase, the trained model is tested using two cases; random and continuous testing sets. Then the evaluation metrics are calculated for the six classifiers. The random testing data set is built from different patients at not connected EEG recording. The performance of this

test is shown in the following (Table 6) (Figure 3) shows the result in graphical form.

The other testing data set are continuous dataset from one patient with connected EEG recording. Also, the performance result is mentioned in the next (Table 7) with graphical form as in (Figure 4).

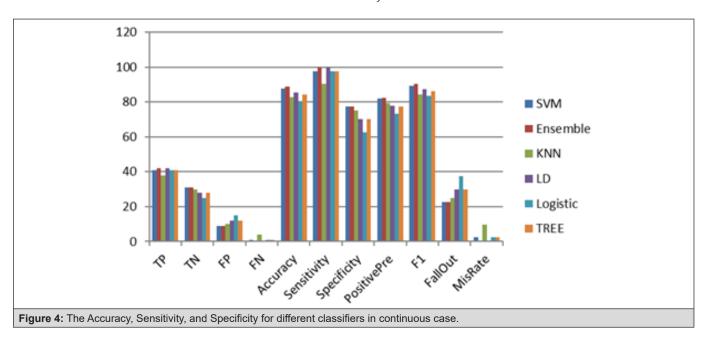
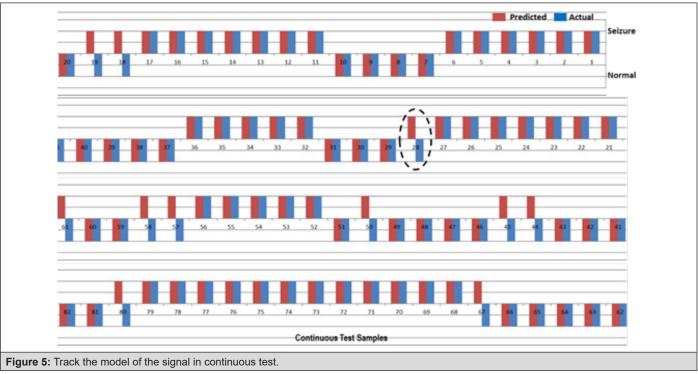


Table 7: The Results for Different Classifiers in Test Phase in Continuous Case.									
	SVM	Ensemble	KNN	LD	Logistic	TREE			
TP	39	42	37	41	40	40			
TN	29	24	31	22	22	27			
FP	8	13	6	15	15	10			
FN	4	1	6	2	3	3			
Accuracy	85	82.5	85	78.75	77.5	83.75			
Sensitivity	90.69767	97.67442	86.04651	95.34884	93.02326	93.02326			
Specificity	78.37838	64.86486	83.78378	59.45946	59.45946	72.97297			
PositivePre	82.97872	76.36364	86.04651	73.21429	72.72727	80			
F1	86.66667	85.71429	86.04651	82.82828	81.63265	86.02151			
FallOut	21.62162	35.13514	16.21622	40.54054	40.54054	27.02703			
MisRate	9.302326	2.325581	13.95349	4.651163	6.976744	6.976744			

Result Discussion

From (Figure 5), it is clear that the proposed model successfully detected all cases of seizure outshout any missing, while there are limited normal instances are classified as a seizure. It is evident from the quantitative results that are in (Table 7). The model Sen-

sitivity is 100% which means the same indication as well as the MisRate equals to zero, which means the detection of all positive –seizure cases- are well classified. From that, it would be taken into consideration that sensitivity is the primary criterion for evaluating the model. The higher the sensitivity of the model is the more accurate model.



Accuracy parameter is also essential, but it is not the better indication of the model performance to detect the soigure cases, between

tion of the model performance to detect the seizure cases, however in training or testing phase. For example KNN model gives an accuracy higher than all other models in training phase as shown in (Table 5&6). But, in its sensitivity is lower than ensemble model, these parameters are also evaluated by MisRate in the same (Table 6) which is equals to 13.9% for KNN and 2.3% for ensemble. This final value indicates that ensemble model classifies all seizure

cases except only one example –FN- only while KNN miss 6 cases, all these results makes the sensitivity parameter is the primary parameter of model evaluation.

Conclusion

Epilepsy patient suffers from a lot of convulsions and complications, so, it is strongly needed a model that works to determine the seizure as fast as possible to avoid these symptoms. This paper presents a proposed algorithm based on extracted EEG features, channel selection to reduce the computation capacity, and trained model for classification. Variance parameter is calculated for all channels and then used to select the channels by taking the maximum three ones. Eleven features are extracted from the selected channels and averaged to be the input for the classifier in both training and testing steps. Six classifiers are used in this work; Support Vector Machine, Linear Discriminant Analysis, K-nearest neighbors, decision tree, ensemble and logistic regression. The classifiers are tested using two sets of data; random data and continuous data. The results showed that the ensemble classifier was the more accurate to classify all seizure cases correctly as it is got 100% sensitivity for continuous testing data set and also the maximum sensitivity with 97.6% for the random testing set.

Funding

There was no funding organization for this study.

Compliance with Ethical Standards

Conflict of Interest

Author Athar Ein-shoka declares that she has no conflict of interest. Author Mohamed Moawad declares that he has no conflict of interest. Author Ahmed Salem declares that he has no conflict of interest. Author Ayman El-sayed declares that he has no conflict of interest.

Ethical Approval

In this study, cases of epilepsy patients are treated through an offline database from the physionet site.

This article does not contain any studies with human participants or animals performed by any of the authors.

Informed Consent

Informed consent was obtained from all individual participants included in the study.

References

- 1. Y Yuan, G Xun, K Jia, A Zhang (2019) A Multi-View Deep Learning Framework for EEG Seizure Detection. IEEE Journal of Biomedical and Health Informatics 23(1): 83-94.
- H Rajaei, M Cabrerizo, P Janwattanapong, Pinzon-Ardila, S Gonzalez-Arias, et al. (2016) Connectivity maps of different types of epileptogenic patterns. Conf Proc IEEE Eng Med Biol Soc Pp. 1018-1021.
- K Abualsaud, A Mohamed, T Khattab, E Yaacoub, M Hasna, et al. (2018) Classification for Imperfect EEG Epileptic Seizure in IoT applications: A Comparative Study. International Wireless Communications & Mobile Computing Conference (IWCMC) Pp. 364-369.
- F Wang, Q Meng, Y Chen, Y Zhao (2014) Feature extraction method for epileptic seizure detection based on cluster coefficient distribution of complex network. WSEAS Trans Comput 13: 351-360.
- Y Zhang, X Ji, B Liu, D Huang, F Xie, et al. (2017) Combined feature extraction method for classification of EEG signals. Neural Comput Appl 28(11): 3153-3161.

- E M Imah, Arif Widodo (2017) A Widodo A comparative study of machine learning algorithms for epileptic seizure classification on EEG signals. International Conference on Advanced Computer Science and Information Systems (ICACSIS) Pp. 401-408.
- F Riaz, A Hassan, S Rehman, IK Niazi, K Dremstrup, et al. (2016) EMD-Based Temporal and Spectral Features for the Classification of EEG Signals Using Supervised Learning. IEEE Trans Neural Syst Rehabil Eng 24(1): 28-35.
- N Rafiuddin, Y U Khan, O Farooq (2011) Feature extraction and classification of EEG for automatic seizure detection. International Conference on Multimedia, Signal Processing and Communication Technologies Pp. 184-187.
- A K Manisha Chandani, Arun kumar (2017) Classification of EEG Physiological Signal for the Detection of Epileptic Seizure by Using DWT Feature Extraction and Neural Network. Int J Neurol Phys Ther 3(5): 38-42
- 10. A Hamad, E H Houssein, A E Hassanien, A A Fahmy (2016) Feature extraction of epilepsy EEG using discrete wavelet transform. 12th International Computer Engineering Conference (ICENCO) Pp. 190-195.
- 11. J Birjandtalab, M Baran Pouyan, D Cogan, M Nourani, J Harvey, et al. (2017) Automated seizure detection using limited-channel EEG and non-linear dimension reduction. Comput Biol Med 82(1): 49-58.
- 12. G Chen, W Xie, T D Bui, A Krzyżak (2017) Automatic Epileptic Seizure Detection in EEG Using Nonsubsampled Wavelet-Fourier Features. J Med Biol Eng 37(1): 123-131.
- 13. M Qaraqe, M Ismail, Q Abbasi E Serpedin (2014) Channel Selection and Feature Enhancement for Improved Epileptic Seizure Onset Detector. International Conference on Wireless Mobile Communication and Healthcare (MOBIHEALTH) Pp. 258-262.
- 14. H Tekgul, B F D Bourgeois, K Gauvreau, A M Bergin (2005) Electroencephalography in neonatal seizures: Comparison of a reduced and a full 10/20 montage. Pediatr Neurol 32(3): 155-161.
- 15. S Faul, W Marnane (2012) Dynamic, location-based channel selection for power consumption reduction in EEG analysis. Comput Methods Programs Biomed 108(3): 1206-1215.
- 16. J Duun Henriksen, T W Kjaer, R E Madsen, L S Remvig, C E Thomsen, et al. (2012) Channel selection for automatic seizure detection. Clin Neurophysiol 123(1): 84-92.
- 17. R S Selvakumari, M Mahalakshmi (2019) RETRACTED ARTICLE: Epileptic seizure detection by analyzing high dimensional phase space via Poincaré section. Multidimens Syst Signal Process 30(2): 1029.
- 18. M Kaur, G Singh (2017) Classification of Seizure Prone EEG Signal Using Amplitude and Frequency Based Parameters of Intrinsic Mode Functions. J Med Biol Eng 37(4): 540-553.
- 19. C Xiao, S Wang, L Iasemidis, S Wong, W A Chaovalitwongse, et al. (2017) An Adaptive Pattern Learning Framework to Personalize Online Seizure Prediction. IEEE Trans Big Data Pp. 1-13.
- 20. JW Lin, Wei Chen, Chia-Ping Shen, Ming-Jang Chiu, Yi-Hui Kao, et al. (2018) Visualization and Sonification of Long-Term Epilepsy Electroencephalogram Monitoring. J Med Biol Eng 38(6): 943-952.
- 21. E Tessy, PPM Shanir, S Manafuddin (2016) Time domain analysis of epileptic EEG for seizure detection. International Conference on Next Generation Intelligent Systems (ICNGIS) Pp. 1-4.
- 22. UI Awan, UH Rajput, G Syed, R Iqbal, I Sabat, et al. (2016) Effective Classification of EEG Signals Using K-Nearest Neighbor Algorithm. International Conference on Frontiers of Information Technology (FIT) Pp. 120-124.
- 23. A K Jaiswal, H Banka (2017) Local Transformed Features for Epileptic Seizure Detection in EEG Signal. J Med Biol Eng 38(2): 222-235.

- 24. H Rajaguru, Sunil Kumar Prabhakar (2017) Sparse PCA and Soft Decision Tree Classifiers for Epilepsy Classification from EEG Signals. Int Conf Electron Commun Aerosp Technol ICECA Pp. 581-584.
- 25. S Bhattacharyya, A Konar, D N Tibarewala, A Khasnobish, R Janarthanan, et al. (2014) Performance analysis of ensemble methods for multi-class classification of motor imagery EEG signal. International Conference on Control, Instrumentation, Energy and Communication (CIEC): 712-716.
- 26. A Subasi, E Erc (2005) Classification of EEG signals using neural network and logistic regression. Comput Methods Programs Biomed 78(2): 87-99.
- 27. AL Goldberger, Amaral LA, Glass L, Hausdorff JM, Ivanov PC, et al. (2000) PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. Circulation 101(23): 215-20.
- 28. A Garcés Correa, L L Orosco, P Diez, E Laciar Leber (2019) Adaptive Filtering for Epileptic Event Detection in the EEG. J Med Biol Eng 39: 912-918.