International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

AN ANALYSIS ON THE DETECTION OF EPILEPTIC SEIZURES USING ENSEMBLE MODEL

Akkshaya Sri J[a], Sanchez Innocencia D[a], Swetha Srinivasan[a], Dr. J.B.Jona[b], Mr.S.A.Gunasekaran[c]

*[a]*Msc. Decision and Computing Sciences, Department of Computing, Coimbatore Institute of Technology, Coimbatore , India*
*[b]*Associate Professor, Department Of Computer Applications, Coimbatore Institute of Technology, Coimbatore, India.*
*[c] *Assistant Professor, Department Of Computer Applications, Coimbatore Institute of Technology, Coimbatore, India.*

A B S T R A C T

Epilepsy is a neurological disorder condition of the brain and Encephalography (EEG) is one which is used commonly for its detection in the clinical approach. Identification of the brain signals manually is a tedious process and time-consuming, where it puts a heavy burden on the Neurosurgeons which may even affect their performance. In the Detection of Binary Epileptic scenarios, Several automatic techniques have been proposed to assist neurosurgeons to use the traditional approach. This project aims in detecting binary epileptic conditions like the seizure vs. non-seizure or normal vs. ictal. This system has been proposed based on the Deep Learning techniques, which is an ensemble of Pyramidal One-Dimensional Convolutional Neural Network (P-1D-CNN) models. To overcome the limitations of a small amount of data, Data Augmentation of Schemes were done for the P-1D-CNN Model. Further the Accuracy score has been calculated and predicted. The prediction determines whether the condition of the patient is Epileptic or not.

Keywords: Electroencephalogram (EEG), epilepsy, seizure, ictal, interictal, 1D-CNN

1. Introduction

Epilepsy is a neurological disorder affecting about fifty million people in the world. Electroencephalogram (EEG) is an effective and non-invasive technique commonly used for monitoring the brain activity and diagnosis of epilepsy. EEG readings are analyzed by neurologists to detect and categorize the patterns of the disease such as pre-ictal spikes and seizures. The visual examination is time-consuming and laborious; it takes many hours to examine a one day data recording of a patient, and also it requires the services of an expert. As such, the analysis of the recordings of patients puts a heavy burden on neurologists and reduces their efficiency. These limitations have motivated efforts to design and develop automated systems to assist neurologists in classifying epileptic and non-epileptic EEG brain signals.

From the machine learning (ML) point of view, recognition of epileptic and non-epileptic EEG signals is a challenging task. Usually, there is a small amount of epilepsy data available for training a classifier due to infrequently happening seizures. Further, the presence of noise and artifacts in the data creates difficulty in learning the brain patterns associated with normal, ictal, and non-ictal cases.

It is still a challenging problem due to three reasons, i) there does not exist a generalized model that can classify binary as well as ternary problem (i.e. normal vs. ictal vs. interictal), ii) less available labeled data, and ii) low accuracy. To help and aid neurologists, we need a generalized automatic system that can show good performance even with fewer training samples.

* *Corresponding author.* Tel.: ; fax: +0-000-000-0000.
E-mail address:

2. Literature Review

The detection of epilepsy conditions using EEG signals is solved using classification models. It starts with extracting unique features from the data and then classifying under different categories of the disorder. In the following section, we give an overview of the related classification models, which use different feature extraction and classification methods for classification of epileptic and non-epileptic EEG signals.

2.1 Detection using Support Vector Machines:

Nicolaou et al. (Nicolaou et al., 2012) extracted the permutation entropy feature from EEG signals. They employed a support vector machine (SVM) as a classifier and achieved an accuracy of 93.55% for A-E case on the University of Bonn dataset. However, maximum accuracy for other cases such as B-E, C-E, D-E, and ABCD-E is 86.1 %.

Gandhi et al. (Gandhi et al., 2011) has extracted the entropy, the standard deviation and the energy features from EEG signals using DWT. They used SVM and probabilistic neural network (PNN) as a classifier and reported the maximum accuracy of 95.44% for ABCD-E case.

Shoeb et al. (Shoeb, 2009) used an SVM classifier and adopted a patient-specific prediction methodology; the results indicate that a 96% accuracy was achieved. In most of the works, the common classifier used to distinguish between seizure and non-seizure events is the support vector machine (SVM).

2.2 Detection using LDA Model:

However, in (Khan et al., 2012) linear discriminant analysis (LDA) classifier was used for classification of five subjects consisting of sixty-five seizures. It achieved 91.8%, 83.6% and 100 % accuracy, sensitivity, and specificity, respectively.

Patel et al. (Patel et al., 2009) proposed a lowpower, real-time classification algorithm, for detecting seizures in ambulatory EEG. They compared Mahalanobis discriminant analysis (MDA), quadratic discriminant analysis (QDA), linear discriminant analysis (LDA) and SVM classifiers on thirteen (13) subjects. The results indicate that the LDA show the best results when it is trained and tested on a single patient. It gave 94.2% sensitivity, 77.9% specificity, and 87.7% overall accuracy. When generalized across all subjects, it gave 90.9% sensitivity, 59.5% specificity, and 76.5% overall accuracy.

2.3 Detection using Neural Network Models:

Swami et al. (Swami et al., 2016) extracted hand-crafted features such as Shannon entropy, standard deviation, and energy. They employed the general regression neural network (GRNN) classifier to classify these features and achieved maximum accuracy, i.e., 100% and 99.18% for A-E (non-seizure vs. seizure) and AB-E (normal vs. seizure) cases, respectively on University of Bonn dataset. However, maximum accuracy for other cases like B-E, C-E, D-E, CD-E, and ABCD-E is 98.4 %.

In another study, Guo et al. (Guo et al., 2010) achieved the accuracy of 97.77% for the ABCD-E case on the same dataset. They used artificial neural network classifier (ANN) to classify the line length features that were extracted by using discrete wavelet transform (DWT)

2.4 Detection using ensemble models:

Acharya et al. (Acharya et al., 2012) focused on using entropies for EEG seizure detection and seven different classifiers. The best-performing classifier was the Fuzzy Sugeno classifier, which achieved 99.4% sensitivity, 100% specificity, and 98.1% overall accuracy. The worst performing classifier was the Naive Bayes Classifier, which achieved 94.4% sensitivity, 97.8% specificity, and 88.1% accuracy.

The overview of the papers given above indicates that most of the feature extraction techniques are hand-crafted, which are not adapted to the data. Several authors have not used data augmentation techniques before training the model. To the best of our knowledge, so far no one has used the DL approach for epilepsy detection, perhaps the reason is the small amount of available data, which is not enough to train a deep model. In this study we apply data augmentation techniques to use the small amount of data in the best way possible. We employed the bagging and ensemble techniques for the model to identify and learn unique parameters to classify.

3. Dataset

The dataset which has been used for this project has been taken from the research team at the University of Bonn and it has been extensively used for the research on the Detection of Epilepsy. The brain EEG signals were recorded using standard 10-20 electrode placement systems. This dataset consists of five sets (A to E) of records in which each record consists of 100 one-channel instances. Set A and Set B consists of the EEG signals that have been recorded from five healthy volunteers in different states like when they were in a relaxed and awake state with their eyes opened (A) and eyes closed (B), respectively. The Sets C, D, and E were also recorded from five patients. EEG signals that have been detected in set D were taken from the Epileptogenic zone. The Set C dataset was recorded from the hippocampal formation of the opposite hemisphere of the brain. The Set C and Set D consists of the EEG signals that has been measured during the seizure-free intervals (interictal), whereas, the EEG signals in the Set E were recorded only during seizure activity (ictal). The categories in the dataset are depicted in Table 1.

Table 1: Categories in Dataset

A	B	C	D	E
Non - Epileptic	Non - Epileptic	Epileptic	Epileptic	Epileptic
Eyes Opened	Eyes Closed	Inter-ictal	Inter-ictal	Ictal

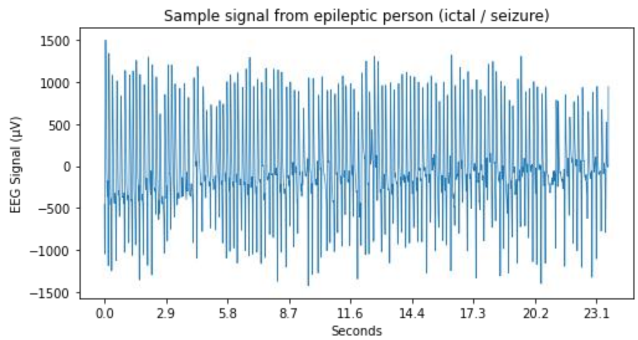


Fig 1: Sample EEG signal

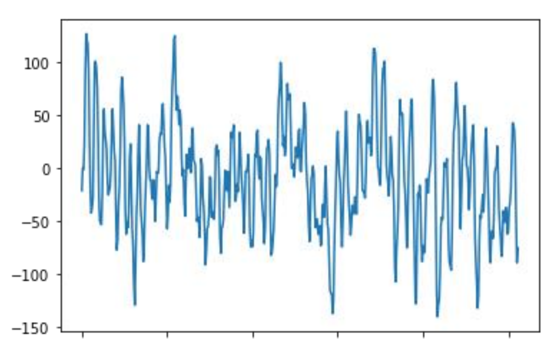
3.1 Data Augmentation

The number of instances that have been collected in this dataset are not sufficient to train the deep learning model. Acquiring a larger number of the EEG brain signals for this problem which is not practical and labeling them by the expert neurosurgeons is not an easy challenge. An Augmentation Scheme is needed which can help in increasing the amount of the data which is sufficient for training a deep CNN Model, which requires larger dataset for better generalization. The available EEG signal dataset is smaller which can learn the model but causes overfitting of the data. To overcome the problem of overfitting, Two Data Augmentation Schemes for training our model is proposed. Each data record in the dataset consists of 4097 samples. For generating many instances from one record, the sliding window approach is adopted. Based on the size and stride of the window size and stride, we propose two data augmentation schemes.

3.1.1 Scheme-1

The signals of the brain are divided into training and testing datasets, where it consists of about 90% and 10% of the total brain signals, respectively. The Data is initially augmented by a training set. Then further by choosing the size of the window which 512 and a stride of 64 (12.5% of 512 with an overlap of 87.5%) where each of the signals in the brain are of length 4097 which is given in the training set and is divided into 57 sub-signals and each of which is being treated and an independent signal instance. In a similar way, a total of about 5130 instances have been created for each category, which has been used for the P-1D-CNN Model. The n instances of the trained P-1D-CNN model which is used to form an ensemble.

In the case of testing, each brain signal is of length 4097 in the testing set which is divided into 4 sub-signals Sts, where each of length 1024; these Sub-signals are being considered as independent signal instances for testing. Each signal instance $Sts$of length 1024 is divided further into three sub-signals with a window of size 512 and 50% overlap. This gives rise to 3 independent signal instances $Sits$,i=1,2,3, each of size 512, which are passed to three trained P-ID-CNN models in the ensemble and majority vote is used as a fusion strategy to take the decision about the signal instance Sts. Each model in the ensemble plays the role of an expert, which analysis a local part of the signal instance Sts independently and the global decision is given by ensemble by fusing the local decisions.



Fig, 2: Scheme 1

3.1.2 Scheme-2:

The method in Scheme-2 is similar to Scheme-1. In this case, the size of the window is 512 with an overlap of 25% (i.e. stride of 128) for creating the training instances Str. For testing, each testing signal instance Sts of length 1024 is divided into three sub-signals with a window of size 512 and 75% overlap. This gives rise to 5 independent signal instances $Sits$,i=1,2,3,4,5, each of size 512, which are passed to five trained P-1D-CNN models in the ensemble and the rule of majority vote has been used as a strategy of fusion for taking the decision about the brain signal instance.

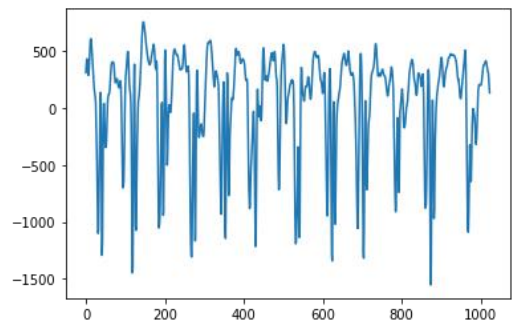


Fig. 3: Scheme 2

4. Proposed System

An electroencephalogram could be a 1D signal, as the algorithm tends to propose a pyramidic 1D-CNN (P-1D-CNN) model for detection of brain disease that involves so many fewer learnable parameters. Because the quantity of accessible knowledge is very small, thus for training the model P-1D-CNN, there is a requirement to propose two augmentation schemes. Exploiting the trained P-1D-CNN models as specialists, the algorithm tends to design a system as an ensemble of P-1D-CNN models that employs a majority vote strategy to fuse the local choices for detection of brain disease. The projected system takes an electroencephalogram signal, segments it with a fixed-size window, and passes every sub-signal to the corresponding P-1D-CNN model (Fig. 2) that processes it and provides the local minima value to the majority-voting module. In the end, the majority-voting module takes the final layer of the parameters with the minima (Fig. 1).

A general deep model desires an enormous quantity of information for training, except for brain disease detection problems as the quantity of information is prescribed. To tackle this issue, the algorithm tend to introduce knowledge augmentation schemes in Section 4, where every electroencephalogram signal is comparable to the brain disease or the traditional case is split into overlapping windows which is the sub signals and every window is treated as an freelance instance to coach P-1D-CNN model. Exploitation copies of the trained P-1D-CNN model, the algorithm tends to build an ensemble classifier, wherever every model plays the role of a skilled examining a particular part of the signal. For classification, keeping the known parameters in the augmentation approach, an input electroencephalogram signal is split into overlapping windows, that are passed to totally different P-1D-CNN models within the ensemble, as shown in Fig. 1, i.e. totally and completely different components of the signal are assigned to different specialist models for its analysis. When input analysis is done, every model provides a neighborhood decision; in conclusion, these choices are a coalesced exploitation majority vote for final calculation. The quantity of P-1D-CNN models within the ensemble depends on the quantity of windows. As an example, just in case an input electroencephalogram signal is split into n windows, the ensemble can comprise of n P-1D-CNN models. The core part of the system could be a P-1D-CNN model. It's a deep model that consists of convolutional, batch standardization, ReLU, absolutely connected and dropout layers.

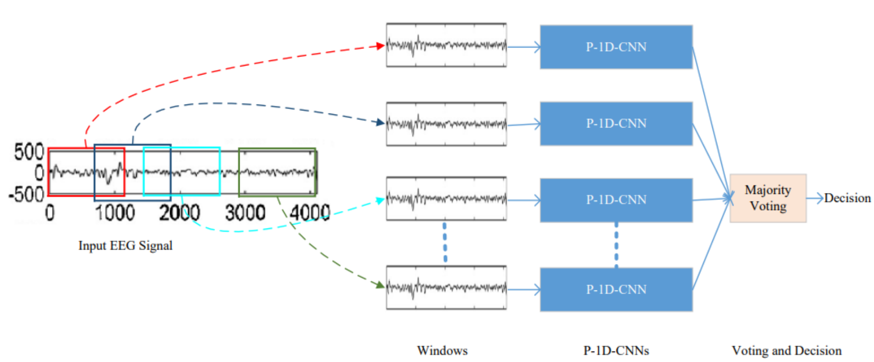


Fig. 4: Overall Structure ensemble EEG classification

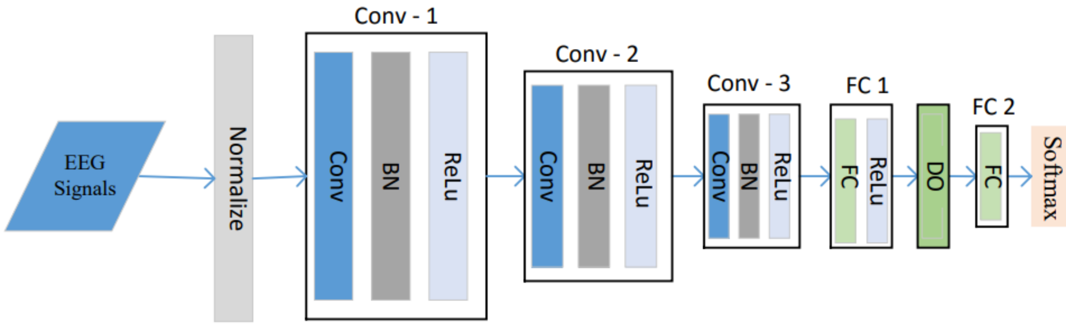


Fig. 5: Proposed Deep Pyramidal 1D-CNN Architecture (P-1D-CNN)

EEG signals could be a 1D time series; as those are the values that are used for analysis. The algorithm tend to propose a pyramidic 1D-CNN model that is likely to decide on P-1D-CNN and its generic design is shown in Fig. 2, its a structured end-to-end model unlike ancient CNN models which was not fully connected with structured layers. The traditional CNN models do not embrace any pooling layer; the redundant or any other parameter to reduce the error rate with the assistance of larger strides in convolution layers. Convolutional and completely connected layers learn a hierarchy of low to high-level pass on the weight parameters from the given signaling. The high-level parameters with the illustration are passed as input to the softmax function category parameters within the last layer to predict the several classes of the input electroencephalogram signal. A CNN model is often structured by adopting a rough to fine approach, wherever low-level layers have a little variety of kernels, and high-level layers contain an oversized variety of kernels. However this structure involves an enormous variety of learnable parameters i.e. its complexness is really high. Instead, the algorithm tends to adopt pyramid design wherever low-level layers have an oversized variety of kernels and better level layers contain a few layers of kernels. This structure considerably reduces the quantity of learnable parameters, avoiding the chance of overfitting. an oversized variety of kernels are taken in a very Conv1 layer, that are reduced by a continuing variety in Conv2 and Conv3 layers e.g. Models E and H which passes the value for Conv1, Conv2 and Conv3 layers with around 24, 16, and 8 kernels. The thought is that low-level layers extract an oversized variety of microstructures, that are composed of upper level layers into higher level options, that are of very low values with different parameters and sometimes gets discriminative, because of the network as it gets deeper.

The input signals are normalized with the value of mean 0 and the unit variance. This standardization helps in quicker convergence and tends to ignore the local minima. The normalized input is processed by 3 convolutional blocks, where every block consists of structured 3 layers: Convolutional layer ($Conv$), Batch Normal standardized layer (BN) and non-linear activation layer ($ReLU$). The output of $R.eLU$ layer within the third block is passed to a completely connected layer ($FC1$) which will also be followed by a $ReLU$ layer and another fully connected layer ($FC2$). So as to avoid overfitting, the algorithm tend to use dropout before $FC2$. The output of $FC2$ is passed to a softmax layer, which acts as a classifier and predicts the category of the signaling.

4.1 Training P-1D-CNN Model

Training of P-1D-CNN wants the calculated weight parameters or the kernels to learn from the data information. For learning these parameters, the algorithm tends to use the standard back-propagation technique with cross entropy loss which performs and generate a random gradient descent approach with Adam optimizer. Adam algorithm has around six hyper-parameters which includes learning rate (0.001), beta1 (0.9), beta2 (0.999), epsilon (0.00000001), use locking parameter (false) and name (Adam); the algorithm have a tendency to use default values of of these parameters mentioned above except learning rate, as it is set to a really tiny value of 0.00002 which does not make a bigger difference. Though batch standardization permits higher learning rate, a small and low learning rate is required to manage the oscillation of the network and to avoid any minima problem while using Adam optimizer. The ctrained accuracy for the model is 0.78 and therefore the Validation/Dev Accuracy is 0.76

4.2 Evaluation of System

During Evaluation of the model, a 10-fold Cross Validation method was adopted for ensuring that the system is tested over different variations in the dataset. The 100 brain signals for each of the classes are divided into 10 folds, each fold (10%) in return, is kept for testing where the remaining 9 folds (90% signal) is used for learning the proposed model. The Performance was evaluated using the performance metrics like the accuracy, specificity, sensitivity,f-measure, precision and then g-mean. Most of the state-of-the-art systems for epilepsy also employ these metrics, the adaptation of these metrics for evaluating our system helps in fair comparison with state-of-the-art systems. The formulas defining the metrics are given below.

$$Accuracy\ (Acc) = \frac{TP + TN}{Total\ Samples}$$

$$Specificity\ (Spe) = \frac{TN}{TN + FP}$$

$$Sensitivity\ (Sen) = \frac{TP}{FN + TP}$$

$$F - Measure\ (F\_M)) = \frac{2 * Precision * Sensitivity}{Precision + Sensitivity}$$

$$G - Mean\ (G\_M) = \sqrt{Specificity * Sensitivity}$$

where TP (true positives) is the number of abnormal cases (e.g. epileptic), which are predicted as abnormal, FN (false negatives) is the number of abnormal cases, which are predicted as normal, TN (true negatives) is the number of normal case that is predicted as normal and FP (false positives) is the number of normal cases that are identified as abnormal by the system.

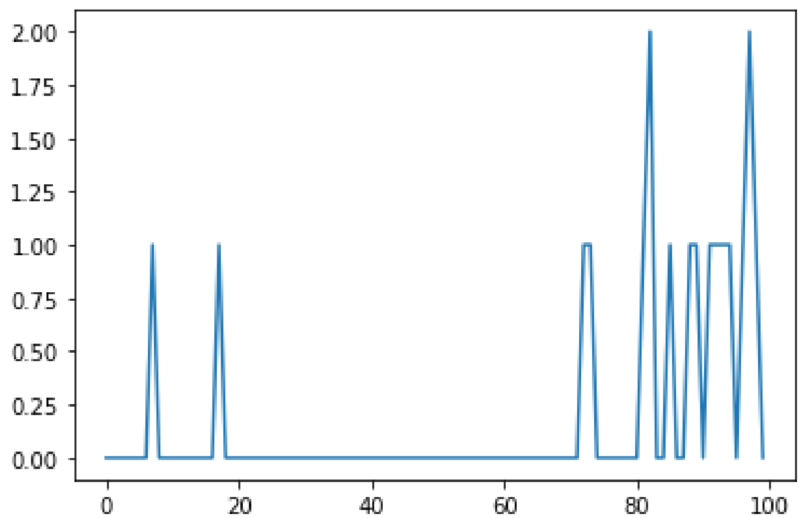


Fig 6: Predicted Graph

5. Conclusion

An automatic epilepsy detection system has been proposed, which deals with ternary detection problems of ictal vs. normal vs. interictal. This proposed system is based on deep learning, which is a state-of-the-art Machine Learning approach. In this system, a memory efficient and simple Pyramidal One-Dimensional Deep Convolutional Neural Network (P-1D-CNN) model has been introduced, which is an end-to-end model, and this involves a smaller number of learnable parameters. This system has been designed as an ensemble of P-1D-CNN models, which takes up an EEG signal as an input, and passes it to different P-1D-CNN models and finally fuses their decisions using majority vote. To overcome the issue of small datasets, two data augmentation schemes have been introduced for learning the P-1D-CNN model. Due to lesser parameters, the P-1D-CNN model is easy to train as well as easy to deploy on chips where memory is limited. It will assist neurosurgeons to detect epilepsy, and will greatly increase their efficiency and reduce their burden. This system will be useful for other similar classification problems based on EEG brain signals. Currently, the Epilepsy Detection methods detect seizures after their occurrence. In the future, investigation and its usefulness for detecting seizures prior to their occurrence, which is a challenging problem, can be done.

REFERENCES

Nicolaou, N., & Georgiou, J. (2012). Detection of epileptic electroencephalogram based on permutation entropy and support vector machines. Expert Systems with Applications, 39(1), 202-209.

Gandhi, T., Panigrahi, B. K., & Anand, S. (2011). A comparative study of wavelet families for EEG signal classification. Neurocomputing, 74(17), 3051-3057.

Shoeb, A. H. (2009). Application of machine learning to epileptic seizure onset detection and treatment. Massachusetts Institute of Technology.

Khan, Y. U., Rafiuddin, N., & Farooq, O. (2012). Automated seizure detection in scalp EEG using multiple wavelet scales. Paper presented at the Signal Processing, Computing and Control (ISPCC), 2012 IEEE International Conference

Patel, K., Chua, C.-P., Fau, S., & Bleakley, C. J. (2009). Low power real-time seizure detection for ambulatory EEG. Paper presented at the Pervasive Computing Technologies for Healthcare, 2009. PervasiveHealth 2009. 3rd International Conference

Acharya, U. R., Molinari, F., Sree, S. V., Chattopadhyay, S., Ng, K.-H., & Suri, J. S. (2012). Automated diagnosis of epileptic EEG using entropies. Biomedical Signal Processing and Control, 7(4), 401- 408.

Guo, L., Rivero, D., Dorado, J., Rabunal, J. R., & Pazos, A. (2010). Automatic epileptic seizure detection in EEGs based on line length feature and artificial neural networks. Journal of neuroscience methods, 191(1), 101-109.

Swami, P., Gandhi, T. K., Panigrahi, B. K., Tripathi, M., & Anand, S. (2016). A novel robust diagnostic model to detect seizures in electroencephalography. Expert Systems with Applications, 56, 116- 130