



# MACHINE LEARNING BASED EEG SIGNAL CLASSIFICATION

Anjali Kadwe  
Electronics and  
Communication Department  
Sapkal College of Engineering  
Nashik

**Abstract**—Epilepsy is a neurological disorder which is characterized by transient and unexpected electrical disturbance of the brain. The electroencephalogram (EEG) is a commonly used signal for detection of epileptic seizures. The proposed method is based on the classification of EEG signal with the less number of sample and more accurately by using the Matlab software. The Bonn University Data set use in this project provide classification of EEG signal by using the latest transform method. The project consist of Extraction of the data from text file, Frequency domain low pass filtering And Feature extraction by three most recent transform such as Coiflet Transform, Stationary Wavelet Transform (SWT) and Walsh Hadamard Transform (WHT). This transformed signal is the classified by KNN ensemble classification. This project provide an overall classification accuracy of 99%.

**Keywords**—Epilepsy, extraction of data from text file, frequency domain low pass filtering, feature extraction, Classification of EEG signal.

## I. INTRODUCTION

Epilepsy is a common brain disorder that, according to an estimate of the World Health Organization, affects almost 60 million people around the world. Approximately one in every 100 persons will experience a seizure at some time in their life. Epilepsy is characterized by the recurrent and sudden incidence of epileptic seizures which can lead to dangerous and possibly life-threatening situations. The seizures are the result of a transient and unexpected electrical disturbance of the brain and excessive neuronal discharge that is evident in the electroencephalogram (EEG) signal representative of the electrical activity of the brain. Consequently, the EEG signal has been the most utilized signal in clinical assessments of the state of the brain and detection of epileptic seizures, and is very important for a proper diagnosis of epilepsy. The detection of epileptic seizures by visual scanning of a patient's EEG data usually collected over a few days is a tedious and time-consuming process. In addition, it requires an expert to analyze the entire length of the EEG recordings, in order to detect epileptic activity.

The major challenge in the classification of MI EEG signals arises due to the fact that the brain signals that are recorded are very small in amplitude. Therefore, events such as eye blink, eye movement, muscular movements, teeth grinding

and heart rhythm interfere with the EEG signal resulting in a signal having low signal to noise ratio (SNR). This prevents the decoding system to correctly decode the user thoughts.

Various techniques have been proposed by the scientific community aiming to improve the temporal filtering methods, spatial filtering, feature extraction and feature selection techniques, dimensionality reduction techniques and classification algorithms. Several feature extraction techniques such as power spectral density (PSD), common spatial pattern (CSP), statistical features, self organizing maps (SOM), correlation, spectral coherence and information entropy have been studied. Classifiers such as support vector machine (SVM), k-nearest neighbors (KNN), random forest (RF), etc. have been explored for classification of MI-EEG signals.

In this paper we are classify the EEG signal with the less number of samples and with more accuracy by using the latest feature extraction method to an epileptic and non epileptic EEG signal.

## II. LITERATURE SURVEY

Many automated EEG signal classification and seizure detection systems, using different approaches, have emerged in recent years. Among such studies, Gotman presented a computerized system for detecting a variety of seizures, while Qu and Gotman proposed the use of the nearest-neighbor classifier on EEG features extracted in both time and frequency domains to detect the onset of epileptic seizures. Gigola et al. applied a method based on the evolution of accumulated energy using wavelet analysis for the prediction of epileptic seizure onset from intracranial epileptic EEG recordings, while Adeli et al., Guler et al. and Ubeyli et al. discussed the potential of nonlinear time series analysis in seizure detection. Artificial neural network-based detection systems for diagnosis of epilepsy have been proposed by several researchers. The method put forward by Weng and Khorasani uses the features proposed by Gotman and Wang, namely, average EEG amplitude, average EEG duration, variation coefficient, dominant frequency and average power spectrum, as inputs to an adaptive structured neural network. The method proposed by Pradhan et al. uses a raw EEG signal as an input to a learning vector quantization network. Nigam and Graupe proposed a new neural network model called LAMSTAR (large memory storage and retrieval) network and two time-domain



attributes of EEG; namely, relative spike amplitude and spike rhythmicity have been used as inputs for the purpose of detecting seizures. The algorithm proposed by Kiymiket al. uses a back propagation neural network with periodogram and auto regressive (AR) features as inputs for automated detection of epileptic seizures. Ghosh Dastidar et al. discussed a classification methodology based on wavelet analysis and both radial basis function and Levenberg-Marquardt back propagation neural network. Srinivasan et al. presented an algorithm based on approximate entropy as an input to an artificial neural network classifier, while Subasi used wavelet analysis and mixture of experts, in addition to the artificial neural network, to classify EEG signals and detect seizures.

Due to quite a low understanding of the mechanisms underlying the problem, most existing methods suffer from low accuracy, a high rate of false alarms and missed detections. In addition, due to a lack of reliable standardized data, most reported EEG analysis-based algorithms are performed on a small number of datasets, which often demonstrate good accuracy for selected EEG segments but are not robust enough to adjust to EEG variations commonly encountered in a hospital setting.

In this research, however, a larger number of EEG data sets, which belong to two subject groups, were used: a) healthy subjects (normal EEG), b) epileptic subjects during a seizure-free interval (interictal EEG). The EEG signal classification and seizure detection problem was modeled as a three-group classification problem that could be of great clinical significance. An automated system able to accurately differentiate between normal and interictal EEG signals can be used to diagnose epilepsy, while a system that can accurately differentiate between interictal and ictal EEG signals can be used to detect seizures in a clinical setting. Therefore, the classification algorithm must be able to classify all two groups accurately and at the same time be robust with respect to EEG signal variations across various mental states and subjects. The improvement of the classification accuracy is mainly based on the design of both an appropriate feature space, by identifying combinations of all extracted features that increase the inter-class separation, and classifiers that can accurately classify all two groups of EEG signals based on the selected and reduced feature space. Real EEG recordings were applied to test algorithm performance and the results indicated that the algorithm has a potential to be applied within an automatic epilepsy diagnosis system.

### III. PROPOSED WORK

#### A. Extraction of the data from text file:

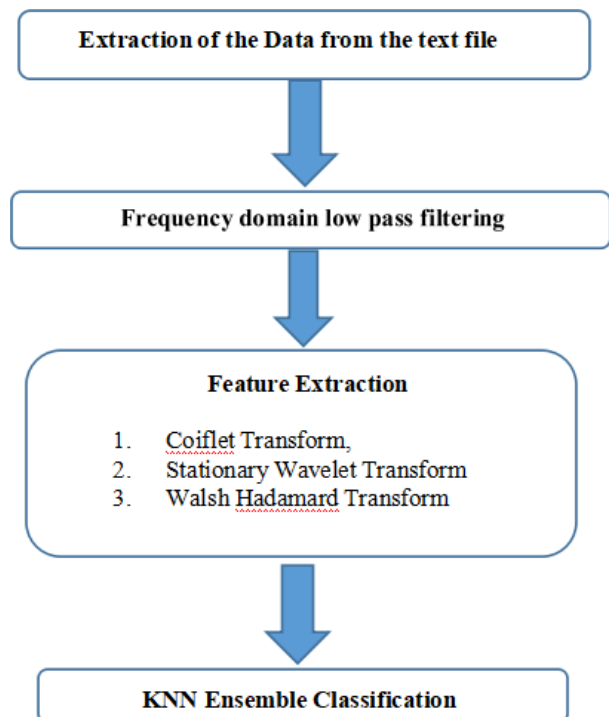
In this project we have used the A and E dataset from the Bonn University for detection and classification of the EEG signal and all this project is carried out with the help of MATLAB software so extraction of the data from the text file is the first step in the process as all the flow work of the process is shown in the fig 1 as follows:

#### B. Frequency domain low pass filtering:

The frequency domain processing techniques are based on modifying the Fourier transform of an image. The basic idea in using this technique is to enhance the image by manipulating the transform coefficient of the image, such as Discrete Fourier Transform (DFT), Discrete Wavelet Transform (DWT), and Discrete Cosine Transform (DCT). This methods advantages includes low complexity of computations, ease of viewing and manipulating the frequency composition of the image and the easy applicability of special transformed domain properties.

#### C. Feature Extraction

Feature extraction is either time domain or frequency domain for EEG signal. In most of the cases we have used different features in bio potential signals, this is because of the characteristics of the EEG signal. However, there are different frequency band of EEG signals, such as alpha beta, delta, and gamma.



#### 1. Coiflate Transform:

In the proposed method, the Coiflet wavelet was selected as the wavelet basis function. This wavelet exists under the name of Coiflets, but it is indeed constructed by I. Daubechies at the request of R. Coifman.

#### 2. Walsh Hadamard Transform (WHT):

Linear image transforms [25] such as: Discrete Cosine Transform (DCT), Discrete Fourier Transform (DFT), Walsh Hadamard Transform (WHT) and Karhunen-Loeve transform (KLT) have been used in various image/video processing applications due to their energy compaction, entropy, flexibility, robustness and performance. Through analyzing or classifying the image frequencies or coefficients of the linear image



transforms, edges of the image can be detected. The kernels or basis of the image transforms used for extracting edges can also be used to compute a set of energy measures which could characterize the local texture properties of a given region in an image. Hence, the kernels can extract the local texture property.

#### B. KNN ensemble classification:

Ensemble learning is an effective and increasingly adopted technique that combines multiple learning algorithms to improve overall performance

#### IV. ADVANTAGES

1. It required very less number of samples to get the decision.

#### V. DISADVANTAGES

1. The data is noisy since it is in mv, microvolt.
2. Very few public data sets are available for experimentation.

#### VI. CONCLUSION

This paper presented an EEG data classification algorithm to epileptic and non-epileptic by using existing dataset which, based on a large number of features extracted after wavelet transform. We have studied the existing literature and found that very few approaches can successfully classify the EEG signal with less number of samples. It is a challenging problem to be solved using machine learning techniques in coming months an efficient algorithm for EEG signal classification can be developed. Therefore, the conclusion is that the proposed algorithm can be used to classify EEG signals with 99% accuracy.

#### VII. REFERENCES

- [1] Hosseini M-P, Nazem-Zadeh M-R, Pompili D, Jafari-Khouzani K, Elisevich K, Soltanian-Zadeh H. "Comparative performance evaluation of automated segmentation methods of hippocampus from magnetic resonance images of temporal lobe epilepsy patients." *Med Phys* 2016;43:538–53.
- [2] Hosseini M-P, Hajisami A, Pompili D. Real-time epileptic seizure detection from EEG signals via random subspace ensemble learning. In 2016 IEEE international conference on autonomic computing (ICAC). IEEE; 2016. p. 209–18.
- [3] Kuzum D, Takano H, Shim E, Reed JC, Juul H, Richardson AG, de Vries J, Bink H, Dichter MA, Lucas TH, Coulter DA, Cubukcu E, Litt B. Transparent and flexible low noise graphene electrodes for simultaneous electrophysiology and neuroimaging. *Nat Commun* 2014;5.
- [4] Jayapandian C, Chen C-H, Dabir A, Lhatoo S, Zhang G-Q, Sahoo SS. Domain ontology as

- conceptual model for big data management: application in biomedical informatics. In: International conference on conceptual modeling. Springer; 2014. p. 144–57.
- [5] Pagán J, Risco-Martín JL, Moya JM, Ayala JL. Modeling methodology for the accurate and prompt prediction of symptomatic events in chronic diseases. *J Biomed Inform* 2016;62:136–47.
- [6] Ombao H, Lindquist M, Thompson W, Aston J. Handbook of neuroimaging data analysis; 2016.
- [7] Cui L, Sahoo SS, Lhatoo SD, Garg G, Rai P, Bozorgi A, Zhang G-Q. Complex epilepsy phenotype extraction from narrative clinical discharge summaries. *J Biomed Inform* 2014;51:272–9.
- [8] Combi C, Pozzani G, Pozzi G. Telemedicine for developing countries: a survey and some design issues. *Appl Clin Inform* 2016;7:1025–50.
- [9] Zhou W, Liu Y, Yuan Q, Li X. Epileptic seizure detection using lacunarity and Bayesian linear discriminant analysis in intracranial EEG. *IEEE Trans Biomed Eng* 2013;60:3375–81.
- [10] Khan YU, Gotman J. Wavelet based automatic seizure detection in intracerebral electroencephalogram. *Clin Neurophysiol* 2003;114:898–908.
- [11] Ghosh-Dastidar S, Adeli H, Dadmehr N. Mixed-band wavelet-chaos-neural network methodology for epilepsy and epileptic seizure detection. *IEEE Trans Biomed Eng* 2007;54:1545–51.
- [12] Conradsen I, Beniczky S, Hoppe K, Wolf P, Sorensen HB. Automated algorithm for generalized tonic-clonic epileptic seizure onset detection based on SEMG zero-crossing rate. *IEEE Trans Biomed Eng* 2012;59:579–85.
- [13] Hopfeng R, Kasper BS, Graf W, Gollwitzer S, Kreiselmeyer G, Stefan H, Hamer H. Automatic seizure detection in long-term scalp EEG using an adaptive thresholding technique: a validation study for clinical routine. *Clin Neurophysiol* 2014;125:1346–52.
- [14] Pachori RB, Patidar S. Epileptic seizure classification in EEG signals using second-order difference plot of intrinsic mode functions. *Comput Methods Programs Biomed* 2014;113:494–502.
- [15] Hassan AR, Siuly S, Zhang Y. Epileptic seizure detection in EEG signals using tunable-Q factor wavelet transform and bootstrap aggregating. *Comput Methods Programs Biomed* 2016;137:247–59.

