A Machine Learning Application for Epileptic Seizure Detection

Ayappan Anugraha[1], Elangovan Vinotha[2], Rangarajan Anusha[3], Sadagopan Giridhar[4], K. Narasimhan[5]
[1,2,3,4] School of Electrical & Electronics Engineering,
[5]Assistant Professor, Department of ECE
SASTRA University, Thanjavur- 613401, Tamilnadu, India
[5]knr@ece.sastra.edu

***Abstract*-Electroencephalography can be treated as an electrophysiological method that can be used to monitor the electrical activity of the brain. Having EEG signal as an aid, there are innumerable diseases that can be detected. Epilepsy is one such disease that can be easily encountered with the abnormalities in the EEG signal. Epilepsy is a condition that affects many people, rendering it the most common neurological disorder after stroke. However it is still difficult to detect some subtle but critical changes in an EEG signal. In this paper we are designing an automated system (classifier) that classifies the recorded EEG signal into Normal, Interictal and Ictal cases. The automation is achieved by extracting various features that include statistical data in the transformed domain using Wavelet and Hilbert techniques. Also the approximate entropy of the sub-bands are included. Now having known the range of the values, each of the features is given a rank. These features are used to enhance the differences between the three cases. Classifier performance is evaluated in terms of its accuracy, specificity and sensitivity. This automated classifier can classify the EEG signal into the desirable cases and has found out its way in biomedical applications by simply repudiating the conventional methods.**

Keywords: EEG; epilepsy; Wavelet transform; Hilbert transform, Approximate Entropy; Feature extraction; Classifier; accuracy; specificity; sensitivity

I. INTRODUCTION

Machine learning is a type of Artificial Intelligence that provides computers the ability to learn without being explicitly programmed. Also Machine leaning focuses on computer programs that can actually changed when exposed to new data. Machine learning algorithms are often categorized into Supervised and Unsupervised learning. Supervised Algorithm can apply what has been learnt in the past to new data. Unsupervised algorithm can draw inferences from data sets. Machine learning in both supervised and unsupervised phase is going to hold the planet in few years. As a scientific endeavor Machine Learning grew out of the quest for Artificial Intelligence. In the biomedical arena, there is always a room for advancements. Having detecting one frightful disease automatically has its own advantages in the medical field. This however produces almost accurate results that can rule out the conventional method in no time. Epilepsy is one of the most common neurological disorders that accounts for 0.6-0.8% of the world's population. Epilepsy means the same thing as "seizure disorders". Epilepsy is characterized by unpredictable seizures and can cause other health problems.

Although the symptoms of a seizure affect any part of the body, the electrical events that produce the symptoms occur in the brain. Electrical events in the brain are recorded through an electroencephalogram (EEG) [6, 7]. Ictal refers to a physiological state or event such as a seizure, stroke or headache. In Electroencephalography, the recording during a seizure is said to be ictal. Interictal refers to a period between seizures that are characteristic of an epileptic disorder. Interictal EEG discharges are those abnormal waveforms not associated with seizure symptoms. The recorded EEG signals are complex non-linear, non-stationary and random in nature. So there are various methods to detect epilepsy using EEG signals. So here, we classify the recorded EEG signal [9], as ictal, interictal or normal through a machine learning approach. Generally, in the technique of machine learning there are various steps involved. They are pre-processing, feature extraction and classification [17,19]. The main purpose of feature extraction is to find some characteristics which are distinct for each of the case. So, feature extraction is vital in any machine learning application [15].

In the pre-processing step, the artifacts from the EEG signal are removed. The features that are extracted [20] includes statistical data of sub-bands of discrete wavelet transform [16,17], statistical data of Hilbert envelope for various wavelet sub-bands and approximate entropy [13] for various wavelet sub-bands. Classification is done using the k-nearest neighbor classifier which has a reasonably higher accuracy for the detection of epilepsy.

II. MATERIALS

EEG time series data are obtained from the open database available at the University of Bonn (department of epileptology) [11]. The obtained time series data were pre-processed.(i.e.)free of artifacts. Set Z is the EEG of healthy subjects with eyes closed. Set O is the data obtained from normal subjects with eyes open. Set S is the EEG of epileptic patients during seizure(ictal). Set F and Set N are the EEG of epileptic patients in the period between seizures(interictal). Each set contains 100 samples of EEG data and each sample contains 4096 values for 23.6 s.

978-1-5090-5595-1/17/$31.00 ©2017 IEEE

III. METHODS

A. *Approximate Entropy:*

Approximate entropy [2, 3] is one of the best features for the automatic detection of epilepsy. It is a technique that is used to quantify the amount of regularity and unpredictability over time series data. The presence of repeated patterns in a time series makes it more predictable and has a lower value. Absence of such patterns leads to a higher value of approximate entropy. In order to calculate[4][16] approximate entropy for a signal we need to choose two input parameters namely sub-sequence length 'm' and similarity coefficient 'r'. The proper selection of these values will aim in the classification of the three cases. Mathematically this can be calculated by the formula

$$apen = \ln(\frac{c_m(r)}{c_{m+1}(r)}$$

where $C_m(r)$ is the mean fraction of pattern of length 'm' and $Cm_{+1}(r)$ the mean fraction of pattern length of 'm+1'. Given a sequence S consisting of N instantaneous measurements approximate entropy can be calculated using the above mentioned formula. Advantages of approximate entropy are it can work for smaller dataset and can also be applied in real time. The disadvantage is that it is highly prone to noise and hence un processed or raw data will not yield a good result.[20]

B. *Discrete Wavelet Transform:*

DWT[16][17] found its application in biomedical engineering arena, especially in epileptic detection due to the fact that it provides features related to transient nature of the signal. Fourier transform of a signal contains the frequency content of the signal over the fixed analysis window. While DWT, has a variable window size by which accurate frequency information can be gathered. Discrete wavelet transform is one of the transform techniques which convert a time series data into its frequency domain. Wavelet method is the best for non-stationary signal analysis Fig.1 shows the DWT plots of sub-bands. Hence to detect epilepsy through an EEG discrete wavelet transform is used. Mathematically,

$$DWT(j,k) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{|2^j|}} \varphi\left(\frac{t-2^j k}{2^j}\right) dt$$

Absolute value of the transformed signal is considered for further statistical analysis. The advantage of Hilbert transform is that it is useful in calculating the instantaneous attributes of a time series data in real time. So Hilbert transform is one of the methods for feature extraction in this paper.

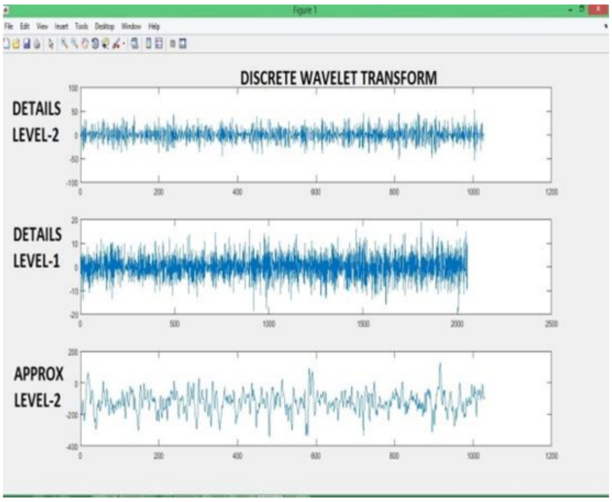

Fig.1 shows the DWT of various wavelets

In the analysis of DWT, suitable wavelet and decomposition levels play a vital role. In this paper, approximation and detail coefficients at levels 1 and 2 are calculated and statistical analysis is made. And we have used Rbio 3.3 wavelet transform for the whole of our study.

C. *Hilbert Envelope*

Extraction of envelope [13] in the field of signal processing has become prevalent in the recent years. It plays a significant role in the analysis of EEG signals. Here, Hilbert transform is applied on all the sub-bands after wavelet analysis. Mathematically, Hilbert envelope of a signal can be calculated from the below mentioned formula.

$$\bar{x}(t) = x(t) * \frac{1}{\pi t} = \int_{-\infty}^{\infty} \frac{x(\tau)}{t-\tau} d$$

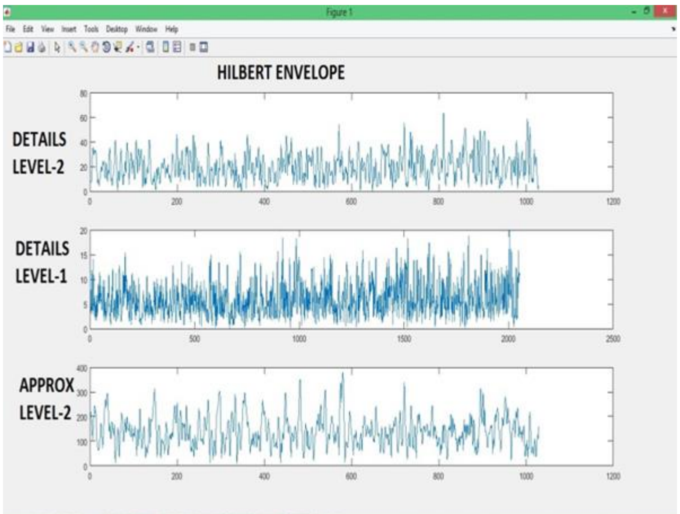

Fig.2 shows the Hilbert envelope of the various wavelets

IV. FEATURE EXTRACTION

Feature extraction [20] is an essential part of machine learning process. Classification accuracy is obtained by the selection of proper features. Here, 21 features were chosen to classify the EEG signal. Firstly, 2clevel decomposition is performed using wavelet transform and various features are extracted. These features include

- The approximate entropy for A2, D1, D2 wavelet sub-bands.
- The statistical data such as Mean, Maximum, Standard deviation for A2, D2, and D1 wavelet sub-bands.
- The above statistical features of the Hilbert envelope curves in each sub-band (A2, D1, D2).

The various features were extracted using the Wave menu command in MATLAB. The features were chosen based on the value obtained considering a significant difference between them for all three cases. Hence the features were then tabulated and used for further work[15] Fig.3. shows the Feature extraction using wave menu.

V. CLASSIFICATION

The classification process involves categorising the EEG signal into three different cases namely Normal, Interictal and Ictal. Having prepared the database for the above extracted features, we used the CLASSIFICATION LEARNER APPLICATION (MATLAB 2016b) to classify the same. Various classifiers like KNN, SVM, Linear classifier were used for the same database and key results like Accuracy, Scatter plot, ROC curve [18] and Confusion Matrix were then analysed. Among the above mentioned classifiers the one with the best accuracy was chosen. Added to this, an user interface was created using MATLAB-GUI. The classifier was trained and then KNN classifier was used to obtain the desired result with an accuracy of 98%. The various Experimental results like the Sensitivity, Specificity and Accuracy for the various classifiers are tabulated below. Analysing the below obtained table, we have chosen KNN classifier for our work and hence accurate results were obtained. Table.1. shows the sensitivity, specificity and accuracy obtained for various classifiers.

TABLE.1. Performance of classifiers

Classifier	**Sensitivity**	**Specificity**	**Accuracy**
KNN	100%	99%	99.7%
SVM	95.5%	96%	95.3%
LINEAR	80%	79.7%	80.7%

VI. EXPERIMENTAL ANALYSIS

In this paper, we have used Approximate entropy along with DWT statistics and statistics of Hilbert envelope to identify the class of the EEG signal. Table.2. shows the results of various papers along with the accuracy of the proposed method.

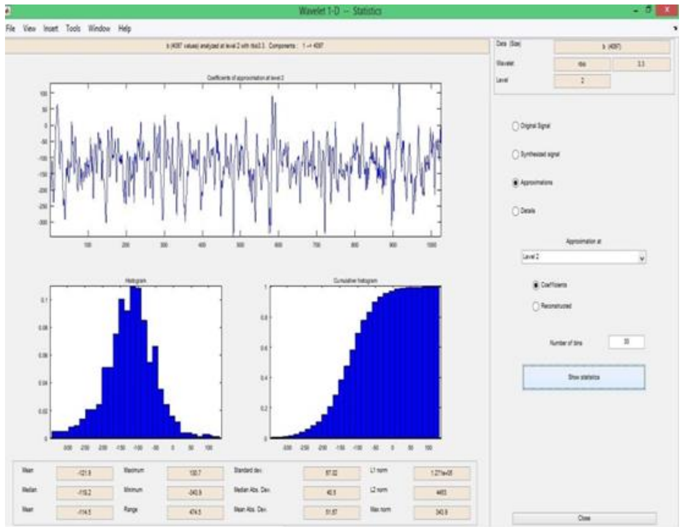

Fig.3 shows the wave menu for extracting features

Fig.4. shows the block diagram of the proposed method.

TABLE.2. Comparison with various methods

Author name	**Method used**	**Classifier**	**Accuracy**
Kannathal et al [1]	Entropy method	ANFIS	92.2
Mingyang Li et al [13]	Hilbert envelope	NNE	98.78
Acharya et al [21]	Recurrence quantification analysis	SVM	95.6
Acharya et al [2]	Entropy Method	FUZZY	98.10
Orhan et al [8]	DWT	ANN	96.67
Proposed method	DWT, Approximate Entropy, Hilbert Envelope	KNN	99.7

VII. PROPOSED METHOD

In the proposed paper, the pre-processed data is directly used for feature extraction. Features like approximate Entropy of the wavelet sub-bands, Hilbert Envelope of the sub-bands and wavelet statistical features are considered and the machine is trained with the same. The training is performed by importing the selected features to the Classification Learner Application and the accuracy is obtained using various Classifiers.

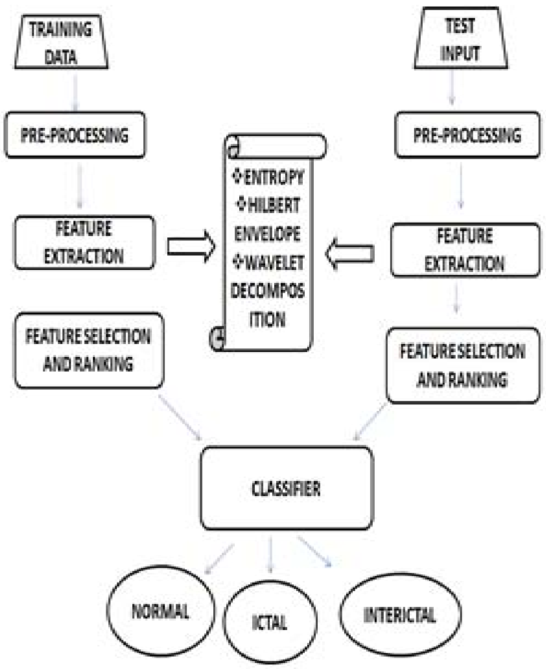

Fig.4 shows the block diagram for the proposed method

We have used Sets S, N and half the values of Z for training while the remaining values of Z , sets O and F for testing. The Classifiers like K-Nearest Neighbour, Support Vector Machine and Linear Classifier are used. The features are ranked and the differences between the three cases are enhanced. Parameters like sensitivity, specificity and accuracy determine the performance of the Classifiers. Added to this Scatter plot and confusion matrix are considered to evaluate the working.

VIII. RESULTS

In the MATLAB, we have three panels, each containing push buttons. Once the LOAD DATA is clicked, the user can choose an existing data from the system. Then the chosen data is displayed in a text box. Then SHOW EEG displays the EEG plot of the respective data. Then GET FEATURES should be clicked to obtain the values of all the features.

Then moving to the next panel GET CLASS gives the class of the previously chosen data. The classifier is trained by storing data bases of training data before hand. So once the class of the chosen data is obtained, the user can now be aware if the corresponding data is epileptic or not. Finally, CLEAR button can be used to clear all the fields and the GUI is now ready to be used again.

IX. CONCLUSION

Epilepsy is a threat to human population which affects a significant part of the globe. This dangerous neurological disorder can be taken care of with proper early detection and treatment. Hence the work was aimed at designing an automated classifier. KNN classifier was used for the same. For real time application, a graphical user interface design was our goal. This was achieved using the MATLAB- GUI. The future of this project work is to make this work purely real time by implementing this in various hospitals by linking this to the internet with web services

REFERENCES

[1]. N.Kannathal, Min LimChoo,U.Rajendra Acharya, P.K. Sadasivan, Entropies for detection of epilepsy in EEG(2005).
[2]. U.Rajendra Acharya,H.Fujitha,Vidhya K. Sudarshan, Shreya Bhat, Joel E.W.Koh, Application of entropies for automated diagnosis of epilepsy using EEG signals(2015).
[3]. S.M.Pincus,Approximate entropy as a measure of system complexity, Proc.Nall.Acl.Sci(1991).
[4]. S.M.Pincus,Approximate entropy(ApEn) as a complexity measure,Chao 5(1)(1995).
[5]. A.J.E.Seely,P.T.Macklem,Complex Systems and the technology of variability analysis,Grit Care 8(6)(2004).
[6]. N.V.Thakor,S.Tong,Advances in quantitative electroencephalogram analysis methods,Ann.Rev.6(2004).
[7]. J.Wackermann,Beyond mapping:estimating complexity of multichannel EEG recordings,Acta Neurobiol.Exp.56(1996).
[8]. Orhan U., Hekim M., Ozer M., EEG signals classification using the K-means clustering and a multilayer perception neural network model Expert system with Application,38(2011).
[9]. Agarwal,R.,Gotman J., Flanagan,D., Rosenblatt,B,Automatic EEG analysis during long-term monitoring in the ICU,Electroencephelography and clinical Neurophysiology(1998).
[10]. Elmas C., Yapay Sinir Aglari,Seckin Yayinicilik,Ankara(2003).
[11]. EEG time series data are available under http://epileptologiebonn.de/cms/front_content.php?idcat=193&lang=3,01.12.2014
[12]. https://www.slideshare.net/syedirshad murtatza/ilae- classification-of-seizure-by-murtaza ,29.01.2015
[13]. Mingyang Li,Wanhong Chen, Classification of epilepsy EEG signals using DWT based envelope analysis and neural network Envelope.
[14]. Srinivasan,V.,Eswaran,C.,Sriram,N., artificial nueral network based epileptic detection using time domain and frequency domain features .Med.Syst.,29(2008).
[15]. Tzallas,A.T.,Tsipouras,M.G.,Fotiadis,D.I.,2007.Automatic seizure detection based on time-frequency analysis and artificial neural networks.comput.intell.nuerosci.,80510.
[16]. Kumar,Y.,Dewal,M.L.,Anand,R.S.,2014.epileptic seizures detection in EEG using DWT-based ApEn and artificial neural network. Signal and Video Processing.8.,1323-1334.
[17]. H.Ocak,Automatic detection of epileptic seizures in EEG using discrete wavelet transform and approximate entropy, Expert Syst.Appl.36 (2) (2009) 2027-2036.
[18]. ROC-The Area Under an ROC Curve. https://gim.unmc.edu/dxtests/roc3.htm (date accessed 02.11.2017).
[19]. Temko,A.,Nadeu,C.,Marnane,W.,Boylan,G.B.,Light body, G.,2011.EEG signal description with spectral-envelope-based speech recognition features for detection of neonatal seizures. IEEE Transactions on Information Technology in Biomedicine.15,839-847.
[20]. Li,S.F.,Zhou,W.D.,Qi,y.,Geng,S.J.,Cai,D.M.,2013.Feature extraction and recognition of ictal EEG using EMD and SVM.Computers in Biology&Medicine.43,807-816.
[21]. U.R. Acharya, S. VinithaSree, S. Chattopadhyay, W. Yu, A.P.C. Alvin, Application of recurrence quantification analysis for the automated identification of epileptic EEG signals, Int. J. Neural Syst. 21 (3) (2011) 199–211.