How to Cite:
Sundaram, C. K., Gayathri, S., & Soundarya, P. (2022). Discrete wavelet transform based on EEG signal analysis for diagnosing neurological disorder. *International Journal of Health Sciences*, *6*(S1), 9556–9566. https://doi.org/10.53730/ijhs.v6nS1.7213

Discrete wavelet transform based on EEG signal analysis for diagnosing neurological disorder

Dr. C. Kalyana Sundaram, M.E.,Phd.,AP(Sl.G.)
Department of Electronics and communication engineering, Mepco Schlenk Engineering College
Email: kalyan@mepcoeng.ac.in

Gayathri. S
Department of Electronics and communication engineering, Mepco Schlenk Engineering College
Email: gayushanmugam1316@gmail.com

Soundarya. P
Department of Electronics and communication engineering, Mepco Schlenk Engineering College
Email: soundarya1024@gmail.com

***Abstract*---**Electro Encephalo Gram (EEG) signal analysis is critical since it is a reliable approach for detecting neurological brain diseases. In this work, the artifacts in the EEG dataset are removed using the Independent Component Analysis (ICA) technique. The EEG dataset was then filtered with a band-pass filter to eliminate noise. In this work using a Discrete Wavelet Transform (DWT) to deconstruct the filtered data, the EEG signal features are recovered. The features are also supplied into four separate classifiers. Five statistical techniques are utilised to extract characteristics from EEG sub bands: Local Binary Pattern (LBP), Standard Deviation (SD), Variance, Kurtosis, and Shannon Entropy (SE). To classify the features related to their classes Linear Discriminant Analysis (LDA), Support Vector Machine (SVM), K-Nearest Neighbour (KNN), and Artificial Neural Network (ANN) are four classifiers that use the features.The overall classification accuracy approaches 90.5% using SVM, 99% using ANN, 87.5% using LDA and 97% using KNN respectively. In this work ANN gives better performance accuracy than other classifers.

***Keywords*---**Artificial Neural Network, Discrete Wavelet Transform, Electro Encephalo Gram, Entropy, Epilepsy, K-Nearest Neighbour, Linear Discriminant Analysis, Support Vector Machine.

Manuscript submitted: 9 March 2022, Manuscript revised: 18 April 2022, Accepted for publication: 1 May 2022

Introduction

Electroencephalography (EEG) data reflects the electrical activity of brain behaviours. Electrical activity or abnormalities of neurons in the human brain are reflected in EEG data. EEG signal analysis- based signal processing methods are an essential clinical tool for monitoring and diagnosing neurological brain disorders like epilepsy. Brain diseases such as epileptic disorders are defined by such processes in the human brain. The major portion of brain disorder diagnoses are now performed manually by neurologists or qualified doctors using EEG patterns. The human brain is the body's most complex organ, providing a lot of knowledge regarding limbic activities and neurological disorders. In order to build and enhance an effective diagnosis system, several types of studies are now being conducted in this area.

Methodology

This section describes the suggested feature extraction and classification techniques, as well as their validation using MATLAB software tools. The EEG data is read first, and then the ICA method is used to remove ocular artefacts from the recorded signals. The EEG data will be segmented into 50-second time frames once the artefacts have been removed. After that, we use a band-pass filter to remove the noise from the framing output. LBP, SD, variance, kurtosis, and entropy are used in combination with DWT to extract features from EEG signals, generate feature vectors, and increase classification accuracy. Classifiers included LDA, SVM, KNN, and ANN. All potential combinations of the recommended approaches were implemented and validated.

Block Diagram

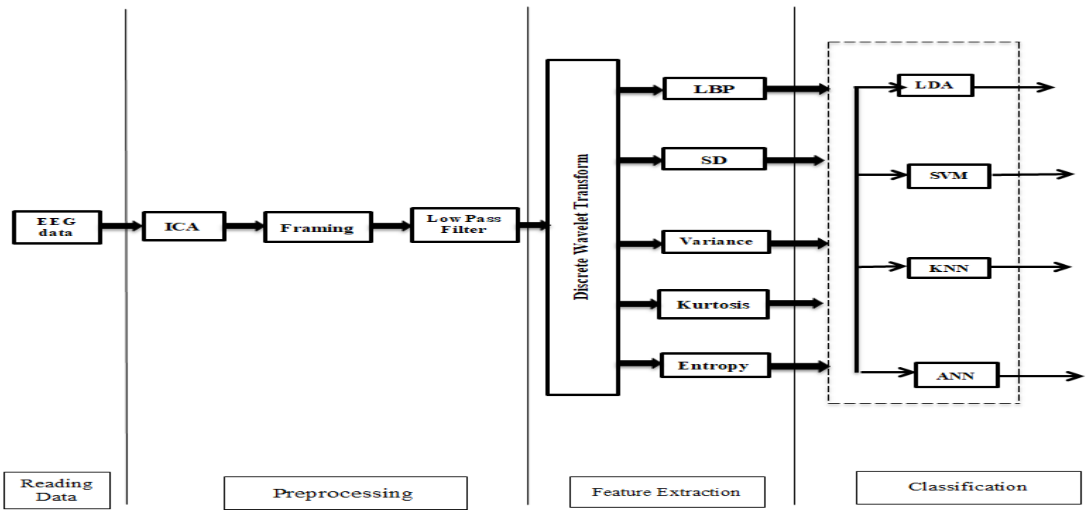

Figure.1.1. Block diagram

Dataset Descriptions

Our approaches were implemented and verified using epilepsy EEG data set. The EEG signal is received from CHB-MIT Scalp EEG database. The EEG signal in this database contains signals for about 1 hour or 2 hours or even 4 hours long including noise and seizures. Epileptologists have annotated the presence of seizure and non-seizure in seconds which is useful for further work. Each and every signal in the database has 23 channels(Channel I: FP1-F7, Channel II: F7-T7, Channel III: T7-P7, Channel IV: P7- O1, Channel V: FP1-F3, Channel VI: F3-C3, Channel VII: C3-P3, Channel VIII: P3-O1, Channel IX: FP2- F4, Channel X: F4-C4, Channel XI: C4-P4, Channel XII: P4-O2, Channel XIII: FP2-F8, Channel XIV: F8- T8, Channel XV: T8-P8, Channel XVI: P8-O2, Channel XVII: FZ-CZ, Channel XVIII: CZ-PZ, Channel XIX: P7-T7, Channel XX: T7-FT9, Channel XXI: FT9-FT10, Channel XXII: FT10-T8, Channel XXIII: T8-
P8). The EEG signal is received from CHB-MIT scalp EEG dataset.

Preprocessing

The EEG signal acquisition process, there were artifacts, noise, and interferences from numerous sources such as the magnetic field of electronic equipment, mobile waves, and so on. In this work, the ICA method and adaptive filtering were used as the first step in the preprocessing stage to eliminate ocular artifacts. EEG data from four electrodes around the eyes were used as reference signals to reduce eye blink artifacts. To guarantee that each segment has the same amount of data, the EEG dataset is segmented into equal segments of distinct wavelengths. The disturbances and interferences created during EEG signal recording were eliminated by filtering the EEG segments after they were segmented.

To increase classification accuracy, the filtering strategy aims to eliminate all noise and interference while boosting the signal to noise ratio. Filtering techniques such as finite impulse response (FIR) and infinite impulse response (IIR) filters were utilised. The elliptic band pass filter outperforms other types of filters in terms of experimental findings because it uses less memory, needs less calculation, and has a shorter delay time. This figure represents channel 1 and channel 2 for an EEG dataset

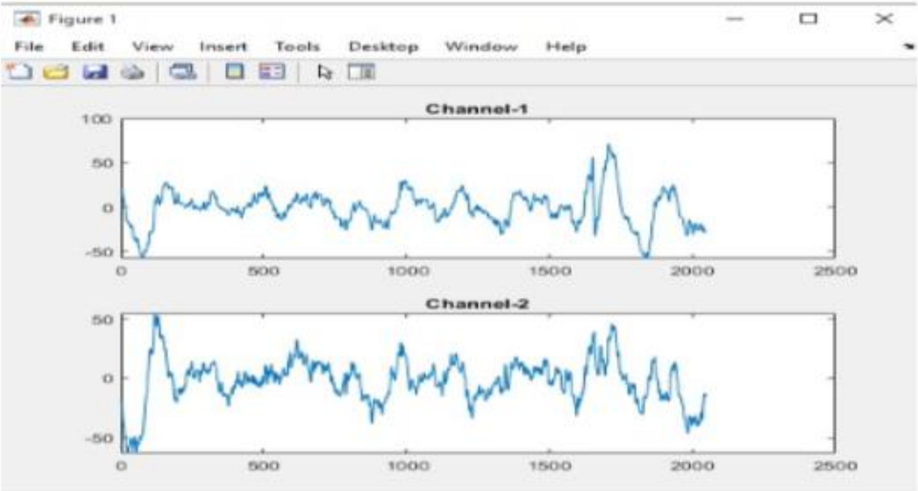

Figure 1.2 Channel 1 and Channel 2

Here the CHB-MIT Scalp dataset is executed and its shown Channel 1 and Channel 2 After removing eye blink artifacts using ICA technique and the results are given below:

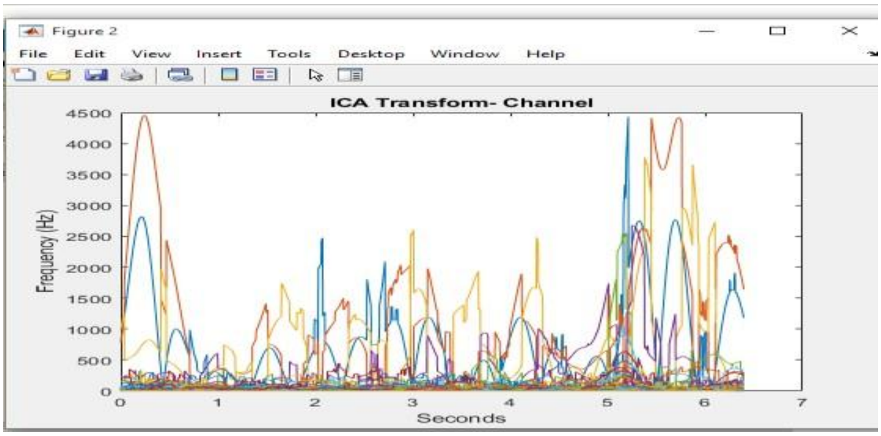

Figure 1.3 Independent Component Analysis Transform

In this preprocessing method ICA technique is used to remove the eye blink artifact from an recorded EEG signal. Here around 23 channels of EEG dataset is obtained. In order to remove the noises and interferences, FIR filter is used and the results were given below:

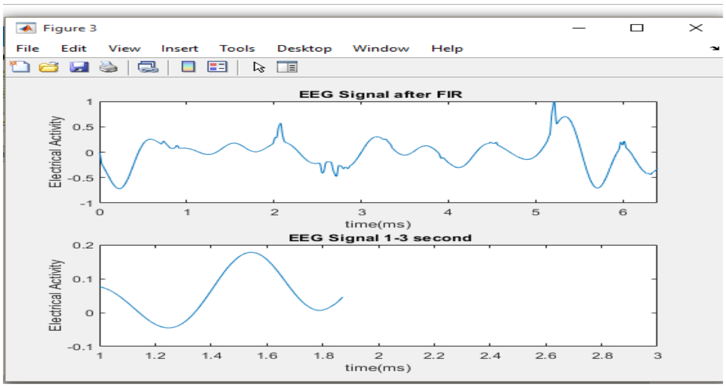

Figure 1.4 FIR Filter

FIR Filter is used to remove the noise from the signal. In FIR filter Beta frequency bands are used and cutoff frequency 100Hz.
Magnitude Response of an FIR filter is given below:

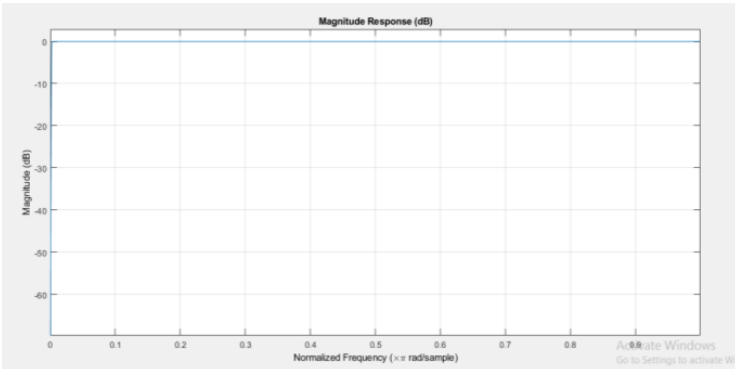

Figure 1.5 Magnitude Response

The Magnitude Response is a real even function of the frequency. Noises and interferences are removed with the help of IIR Filter and the filtered output is given below:

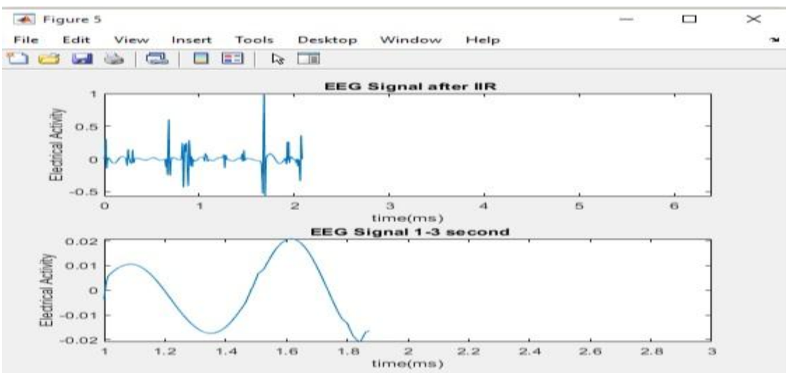

Figure 1.6 IIR Filter

In IIR Filter sampling frequency 1kHz and the order of the filter is 4.

Feature Extraction

When dealing with a limited number of values that describe the EEG signal's characteristics, massive time series of EEG signals are acquired and split. These values are called features, and they are gathered into a vector called the feature vector. As a result, feature-extraction methods are defined as signal-to-feature-vector conversion procedures. A variety of feature extraction approaches are used to extract features. This study made use of DWT, the most popular and commonly used technique. In this work, we recommend creating feature vectors using DWT based on LBP, SD, variance, kurtosis, and entropy. The STFT is useless with non-stationary signals like EEG. STFT has a fixed resolution across all frequencies. Varying frequencies with different resolutions are evaluated multi-resolution is used in the wavelet transform method. Furthermore, the wavelet transform can provide a fewer number of characteristics for the signal being analysed, potentially eliminating the dimensionality issue. On the other hand, wavelet transformations look at signal properties in terms of time and frequency. Domain by utilising a single function to decompose a signal into many functions. The following equation shows the wavelet transform, equ(1)

$$\int ( ) \quad ( ) \qquad (1)$$

The scaling and shifting parameters are x and y, respectively, while the wavelet space is S.

Local Binary Pattern

Local binary patterns (LBP) are a type of visual descriptor used for categorization in computer vision. The LBP is a valuable tool for texture feature extraction, and it is computed using the formula provided in equ (2). The operator may also be utilised with a variety of neighbourhood sizes. By employing circular neighbourhoods and bilinearly interpolating pixel values, any radius and number of pixels in the neighbourhood may be obtained.

$$\text{LBP}=\log\left(\frac{1}{N}\sum_{n=1}^{N}|s(n)|^{2}\right) \qquad (2)$$

S(n) is a discrete signal, where n = 1, 2,.. N, where N is the number of signal samples.

Standard Deviation

The standard deviation is determined for a window of many frame level feature vectors during the feature extraction procedure. The same data for means and standard deviations are collected separately at the clip level, then layered to form a single feature vector per clip. The Standard Deviation for signal is calculated by equ (3).

$$=\sqrt{(\sum \quad (( )))} \qquad (3)$$

Where mean of signal samples.

Variance

The square of the signal's standard deviation is the variance of a discrete-time signal. The divergence of a signal from its mean value is measured by its variance.

$$= \sum \; (( ))) \qquad (4)$$

Kurtosis

Kurtosis is a statistical term for describing the features of a signal. It essentially determines the "peakedness" of a random signal. Three-sigma peaks, or peaks that are three times the signal's RMS value, are more common in signals with a greater kurtosis value. It can be calculated by using the formula given in equ (5). There are no units in Kurtosis.

$$\text{Kurtosis} = \frac{\sum \; (\quad \bar{\bar{\ }})}{\sum \; (\quad \bar{\ })} \qquad (5)$$

Entropy

Entropy is a quantitative EEG device that acquires a single-lead frontal EEG signal using a 3-electrode sensor on the patient's forehead. The ElectroEncephaloGram (EEG) signals' "spectral entropy," which is a measure of how uniform the power spectrum is, and it is calculated by the system using equ (6).

$$( ) \sum \qquad ( ) ( ) \qquad (6)$$

The table represents the results of Feature extraction and it is given below:

Table 1.1.
Feature Extraction output for chb01

Mean	3.5577e-06
Variance	0.02409
Standard Deviation	0.15521
Kurtosis	24.8643
Entropy	33.4822
LBP	3.7259

Classification

The characteristics are fed into four different classifiers: LDA, SVM, KNN, and Artificial Neural Networks (ANNs).

LDA

LDA and regression analysis are all methods for expressing one dependent variable as a linear combination of other features or data. LDA works when measurements on independent variables are continuous values for each observation. When dealing with categorical independent variables, discriminant correspondence analysis is the analogous approach. The mean vectors of the feature vectors and their covariance matrices for each class are required by LDA, which is a generalisation of Fisher's linear discriminant.

Support Vector Machine

Support-vector machines are supervised learning models that use related learning techniques to analyse data for classification and regression analysis. An SVM training approach generates a model that assigns new instances to one of two categories using a non- probabilistic binary linear classifier.

K-Nearest Neighbour

The K-Nearest Neighbor approach is used to classify and predict data by dividing it into groups depending on the distance between data points. The K-Nearest Neighbor technique assumes that data points near together are similar, hence the data point to be classified will be grouped with the cluster closest to it. The K-Nearest Neighbour method is one of the most fundamental Machine Learning algorithms. It is based on the Supervised Learning approach.

Artificial Neural Network

Artificial neural networks (ANNs), sometimes known as neural networks (NNs), are computer systems that mimic the organic neural networks seen in animal brains. Artificial neurons are a collection of linked units or nodes in an artificial neural network (ANN) that loosely mimic the neurons in a biological brain. Each connection, like synapses in the human brain, may transmit a signal to other neurons.

By using Neural network training Tool the performance is given below:

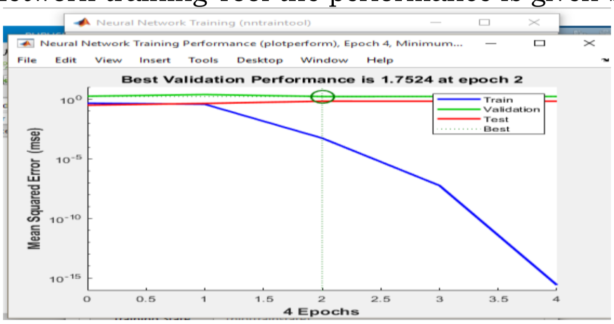

Figure 1.7 Neural Network Training Performance

By using Neural network training tool the Regression plot is given below:

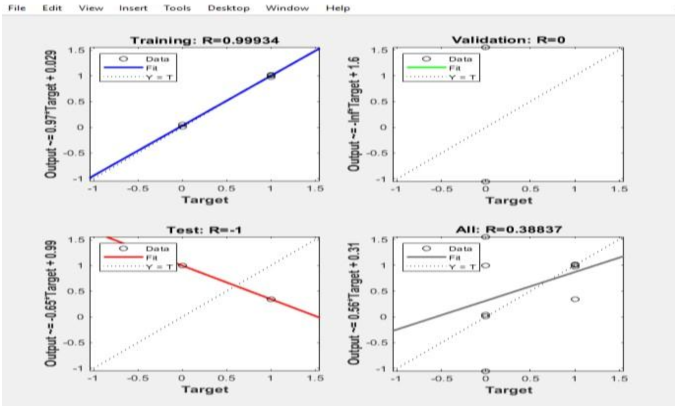

Figure 1.8 Regression plot table for classification

Table 1.2
Classification analysis for chb01

Parameter	LDA	SVM	KNN	ANN
Accuracy	83.5	92	97	99
Sensitivity	100	100	100	100
Specificity	75.19	86.2	94.33	98
Precision	67	84	94	98
Recall	100	100	100	100

Literature Survey

[1]. Nigam and Graupe proposed an EEG-based computer-aided diagnosis for epilepsy using a multistage nonlinear preprocessing filter and an artificial neural network (ANN). The approach they provide has a 97.2 percent accuracy rate (2015). [2]. Kannathal et al. looked at a variety of entropy approaches and recommended that entropy values be used to distinguish between neurotypical and epileptic EEGs. They employed a 92.2 percent accurate adaptive neuro-fuzzy inference method for classification (2017). [3]. Sadati et al employed an adaptable neural fuzzy network to diagnose epilepsy. To extract features, the energy of discrete wavelet transform (DWT) sub-bands was employed. Their recommended strategy, on the other hand, was accurate to the tune of 85.9%. (2014). [4].Ocak provided a method for feature extraction with DWT that used estimated entropy and reached an accuracy of over 96 percent with DWT and 73 percent without it.(2012). [5]. Instead of only categorising only sets A and E, Nunes et al. examined different feature extraction and classification methods using the whole dataset (Sets A, B, C, D, and E) given by Bonn University. Wavelet coecients as feature extractors and ideal route forest as a classifier produced the best results, with an average accuracy of 89.2 percent. (2014). [6]. Subasi used the wavelet transform to extract features and an expert model to classify them. The proposed method's total accuracy was 94.5 percent. (2007). [7].Sheikhani et al. employed In their research, they employed the short-time Fourier transform (STFT) approach to extract EEG signal features and then used k-nearest neighbours (KNN) as a classifier. The total accuracy of this method is up to 82.4 percent.(2015).[8]. Fan

et colleagues coupled spectral elements of EEG data with therapist judgements of behavioural links, enjoyment, discouragement, boredom, and difficulty to develop a series of classification models. To achieve an overall classification accuracy of 75 to 85 percent, they looked at the outcomes of seven classification methods, including Bayes network, naive Bayes, multilayer perceptron, SVM, KNN, decision tree classifier (J48), and random forest. (2016).

Conclusion

EEG signal-analysis techniques have improved in recent years, and it is an important tool for diagnosing neurological brain disorders such as epilepsy. The CHB MIT Scalp database provided the datasets used in this study. To eliminate eye blink artifacts from an EEG dataset, the ICA approach is utilized. After framing and filtering the EEG data to remove noise and interference, the EEG characteristics are extracted from the filtered signal using DWT to break it down into its sub-bands. Five statistical techniques are examined in order to extract features from EEG sub-bands. Four separate classifiers are used as inputs to sort the features into their relevant categories. The outcomes of our proposed approaches the four types of classifiers are employed—LDA, SVM, KNN, and ANN. The overall classification accuracy approaches 90.5000% using SVM and 99% using ANN and 97% using KNN. The combination of DWT + (LBP + Entropy) + ANN achieved the highest accuracies among other three classifiers.

Future Work

We will test and evaluate our suggested methodologies with more datasets in future research. Our investigation will also involve another neurological condition.

References

[1] Noachtar, S.; Rémi, J., 2009, The role of EEG in epilepsy: A critical review. Epilepsy Behav., 15, 22–33. [2]. Ocak, H., 2009, Automatic detection of epileptic seizures in EEG using discrete wavelet transform and approximate entropy. Expert Syst. Appl. 36.

[2] Nunes, T.M.; Coelho, A.L.; Lima, C.A.; Papa, J.P.; De Albuquerque, V.H.C., 2014, EEG signal classification for epilepsy diagnosis via optimum path forest—A systematic assessment. Neurocomputing, 136, 103–12.

[3] Subasi, A., . 2007 EEG signal classification using wavelet feature extraction and a mixture of expert model. Expert Syst. Appl, 32, 1084–1093.

[4] Subasi, A.; Gursoy, M.I., 2010, EEG signal classification using PCA, ICA, LDA and support vector machines. Expert Syst. Appl.,37, 8659–8666.

[5] Morabito, F.C.,2000, Independent component analysis and feature extraction for NDT data. Mater. Eval.,58, 85–92.

[6] Yong, Y.; Hurley, N.; Silvestre, G.,2005, Single-trial EEG classification for brain-computer interface using wavelet decomposition. In Proceedings of the European Signal Processing Conference, Antalya, Turkey, 4–8. [8].

[7] Patel, K.; Chua, C.P.; Fau, S.; Bleakley, C.J., 2009 Low power real-time seizure detection for ambulatory EEG. In Proceedings of the 3rd International

Conference on Pervasive Computing Technologies for Healthcare,London, UK, 1–3.

[8] Yuan, Q.; Zhou, W.; Liu, Y.; Wang, J., 2012, Epileptic seizure detection with linear and nonlinear features. Epilepsy Behav.,24, 415–421.

[9] El Safty, S.; El-Zonkoly, A., 2009 Applying wavelet entropy principle in fault classification. Int. J. Electr. Power Energy Syst., 31, 604–607.

[10] Shoeb, A., 2009 Application of Machine Learning to Epileptic, Seizure Onset Detection and Treatment. Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, MA, USA.

[11] Sadati, N.; Mohseni, H.R.; Maghsoudi, A., 2006. Epileptic seizure detection using neural fuzzy networks. In Proceedings of the IEEE International Conference on Fuzzy Systems, Vancouver, BC, Canada, 16–21.