International Journal of Engineering and Techniques - Volume 3 Issue 6, Nov - Dec 2017

RESEARCH ARTICLE OPEN ACCESS

Fractal Dimension and Higher Order Statistics Based Features for Classification of Different Epileptic States

Kaushik Das[1], Asif Ahmed[2]

[1,2](Electronics and Communication Engineering, North-Eastern Hill University, Shillong)

Abstract:

Here we have presented a method for the classification of different types of electroencephalogram (EEG) signals in the empirical mode decomposition (EMD) domain. Here we have used a EEG dataset which is available online, in the dataset out of five subsets we have considered three subsets forming normal, interictal and ictal states. Here we have used some of the statistical moments like variance, skewness and kurtosis and we have also used fractal dimension and sample entropy on the intrinsic mode functions (IMFs) which are obtained by doing EMD on the main EEG signals. All the obtained features are feed to a support vector machine for classification of normal and ictal states as well as interictal and ictal states. The mentioned method gives a classification accuracy of 100% in almost all the cases for classification of these states.

Keywords — **Electroencephalogram (EEG), Support vector machine (SVM), Epileptic seizure, Empirical mode decomposition (EMD)**

I. INTRODUCTION

The human brain is a highly complex system and consists of billons of neurons and these are responsible for maintaining brains electrical charge. EEG is an electrophysiological monitoring method that contains the information about the human brain activity. The EEG signal can be obtained by placing the sensors on the scalp or using the intracranial electrodes. The EEG signals can be used for various purposes like emotion recognition [1], brain–computer interfaces [2], etc. Seizure is an neurological disorder of the brain and this occurs when the human brain cell fires electrical impulses three times more than that of normal one, a pattern of repeated seizure is called epilepsy. Epilepsy can be caused by alcohol, drugs, brain injury and can also be genetically. During seizure period some of the symptoms like rapid eye blinking, shaking of body, clenching of teeth are observed [3]. There are lots of method developed in the literature for EEG signal analysis and classification are based on time domain, frequency domain and time-frequency domain. The spikes detection methods for EEG signal analysis have been proposed in [4], time domain and frequency domain features along with artificial neural network (ANN) have been used for normal and epileptic EEG signal analysis [5]. EMD is a technique now researchers are using widely, The measure namely, area of analytic IMF in the complex plane has been used in order to discriminate between seizure and normal EEG signals [6]. Computation of amplitude modulation and frequency modulation bandwidths have been used for the classification of seizure and normal EEG signals [7].

Here firstly we will perform EMD on the EEG signals to get the IMFs, after that we will perform higher order statistical moments like variance, skewness and kurtosis as well as some of the features like fractal dimension and sample entropy on the IMFs and these will be used as a input to the support vector machine for classification. The paper is organised as follows: The EEG dataset in section II, methods in section III, which includes EMD, fractal dimension, sample entropy, higher order statistical moments and support vector machine. The results and discussion is in section IV and finally section V concludes the paper.

II. DATASET

In this paper the dataset we have used is available online in [7],[8]. The dataset consists of five subsets (namely Z, O, N, F, and S), each containing 100 single-channel EEG signals, each of having 23.6 s duration. These signals have been taken out from continuous multichannel EEG recording after visual inspection of artefacts. The subsets Z and O have been recorded extracranially, whereas the subsets N, F, and S have been recorded intracranial. The subsets Z and O have been acquired from surface EEG recordings of five healthy volunteers with eyes open and closed, respectively. Subsets N and F consist of EEG recorded during seizure free intervals from epileptogenic zone and hippocampal formation of the opposite hemisphere, respectively. subset E contains signals with seizure activity. The signals are recorded in a digital format at a sampling rate of 173.61 Hz. Out of five subsets we have taken three subsets named Z, F and S forming normal, interictal and ictal states respectively. Here we have considered two cases first case is consisting of normal and ictal states and second case is consisting of interictal and ictal states.

III. METHODS

A. *EMPIRICAL MODE DECOMPOSITION (EMD)*

It is a signal processing technique that decomposes a signal so called intrinsic mode functions. The EMD method does not require any condition about whether a signal is stationary or not. The main aim of the EMD method is to decompose a signal *x(t)* into a numbers of intrinsic mode functions (IMFs). Each IMF satisfies two basic conditions: 1) the number of maxima-minima and the number of zero crossings must be the same or differ at most by one; 2) at any instant, the mean value of the envelope formed by the local maxima and the envelope formed by the local minima is zero.

The EMD algorithm of the signal $x(t)$ can be explained as follows [7],[9].

1) Detect the extrema (maxima and minima) of the dataset *x(t).*

2) Formation of upper and lower envelopes $e_m(t)$ and $e_l(t)$ respectively, by connecting the maxima and minima with cubic spline interpolation.

3) Calculate the local mean as $a(t)=\frac{e_m(t)+e_l(t)}{2}$

4) Extract the detail $h_1(t)=x(t)-a(t)$

5) Decide whether $h_1(t)$ is an IMF or not by checking the two basic conditions as described above.

6) Repeat steps (1) to (4) and end when an IMF $h_1(t)$ is obtained.

As soon as the first IMF is derived, define $c_1(t)=h_1(t)$, which is the smallest temporal scale in $x(t)$. To determine the rest of the IMFs, generate the residue $r_1(t)=x(t)-c_1(t)$. The residue can be treated as the new signal and we can repeat the above explained process until the final residue is a constant or a function from which no more IMFs can be derived. At the end of the decomposition, the original signal $x(t)$ is represented as

$$x(t)=\sum_{m=1}^{M}c_m(t)+r_M(t) \quad (1)$$

where M is the number of IMFs, $c_m(t)$ is the mth IMF and $r_M(t)$ is the final residue. Each IMF in (1) is assumed to yield a meaningful local frequency, and different IMFs do not exhibit the same frequency at the same time. Then, (1) can be written as

$$x(t)=\sum_{m=1}^{M}A_m(t)\cos\left[\varphi_m(t)\right] \quad (2)$$

The plot of subsets Z, F and S are shown in figure 1. and their empirical mode decomposition are shown in fig. 2, 3 and in 4.

B. *FRACTAL DIMENSION (FD)*

The term "fractal dimension" is given by Mandelbrot, on the basis of fractal geometry. It has been found that fractal dimension is a promising parameter in distinguishing the non linear and non stationary property of EEG signal [16]. In this paper, the most used Higuchi's algorithm [17] is used for the fractal dimension calculation of EEG sequences.

Let us consider a discrete time sequence X containing N data points, X = {*x(1), x(2),* . . . ,

$x(N)\}$, the reformed time series x_r^n with n discrete time interval between points is given as [17]:

$$x_r^n = \{x(r) + x(r+n) + x(r+2n) + ... + x(r + \left\lfloor \frac{N-r}{n} \right\rfloor n)\} \quad (3)$$

Here, symbol $\lfloor y \rfloor$ represents the greatest integer which is less than or equal to y; $r = 1, 2, . . . , n$ and it means the initial time value.

The average length $J_m(n)$ of each subsequence x_r^n is defined as

$$J_m(n) = \frac{\sum_{i=1}^{\lfloor b \rfloor} |x(r+in) - x(r-(i-1)n)| (N-1)}{\lfloor b \rfloor n} \quad (4)$$

where b is computed as

$$b = (N-r)/n \quad (5)$$

Therefore, the fractal dimension of any time sequence X can be calculated as

$$FD = \log(J_m(n)) / \log(1/n) \quad (6)$$

In this paper we have taken n=5.

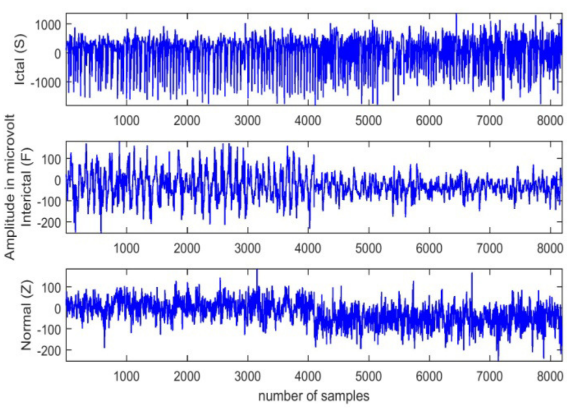

Fig. 1. Plot of two normal (Z), interictal (F) and ictal (S) EEG signals.

C. SAMPLE ENTROPY (SaEn)

Sample entropy is an alteration of the approximate entropy used for measuring the complexity of a time-series data. Sample Entropy is given by the negative natural logarithm of the conditional probability means any two sequences which are same for m points will remain same at the next point where r is identified as similarity criterion and m is the length of data segment [11],[13]. Compared to Approximate entropy, Sample Entropy is more immune to noise and does not depend on the data series length. A lower value of Sample Entropy means an increased matching in the time series data. Sample Entropy can be given by the following equation

$$SaEn(m,r,k) = -\ln[B^{m+1}(r) / B^{m(r)}] \quad (7)$$

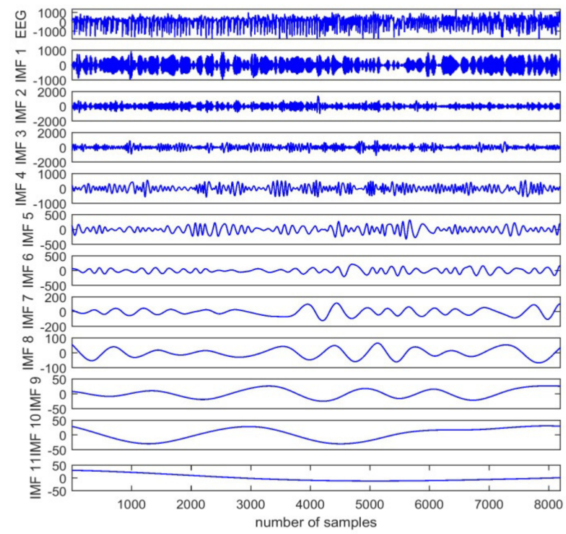

Fig. 2. Empirical Mode Decomposition of two ictal EEG signals.

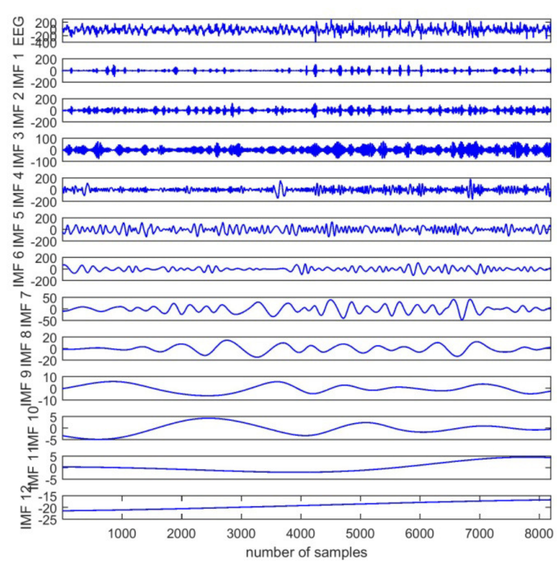

Fig. 3. Empirical Mode Decomposition of two interictal EEG signals.

where $B^{m(r)}$ is the probability of matching two sequences for m points and $B^{m+1}(r)$ is the probability that two sequence will match in the next point i.e. for m+1 points.

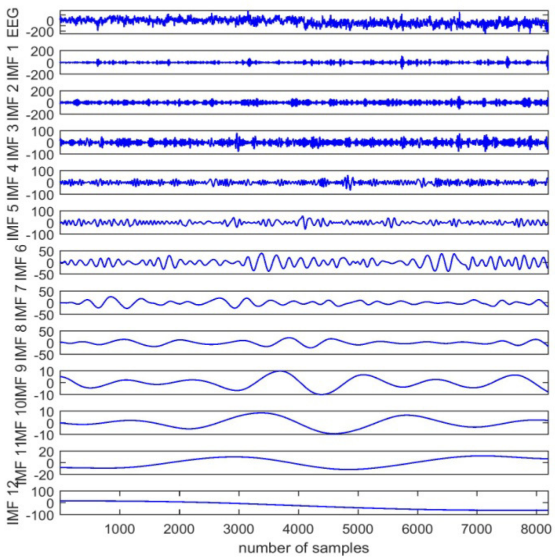

Fig. 4. Empirical Mode Decomposition of two normal EEG signals.

D. HIGHER ORDER STATISTICS OF IMFS

Here we have used higher order statistical moments like variance, skewness and kurtosis. Generally variance is a measure of how far a set of numbers are spread out from their mean, Skewness tells us about asymmetry of a probability distribution function and Kurtosis tells us about the thickness or weight of the tails of a distribution [14].

Let us consider a random variable *X* has the mean, μ = A(*X*), then the variance *Var*(*X*) of *X* is given by:

$$Var(X) = A[(X-\mu)^2] \quad (8)$$

Skewness is given by the following formula:

$$\delta_1 = \frac{\mu_3}{\sigma^3} \quad (9)$$

where σ is the standard deviation and μ_3 is a moment of third order around the mean.

Kurtosis is given by the following formula:

$$\delta_2 = \frac{\mu_4}{\sigma^4} \quad (10)$$

where μ_4 is a moment of fourth order around the mean.

E. SUPPORT VECTOR MACHINE (SVM)

In order to classify the normal and ictal as well as interictal and ictal EEG signals, all the parameter values obtained are given as a input to a support vector machine. Suppose we have two classes and an unknown feature vector which is to be classified, then our goal is to design a hyperplane that classifies all the training vectors in two classes, but we may have different hyperplane so best choice will be the hyperplane that leaves the maximum margin for both classes [12]. The decision function of a SVM is given by

$$l(x) = \text{sgn}\left(\sum_{j=1}^{k} \alpha_i r_i \Psi(x, x_i) + b\right) \quad (11)$$

where Ψ defines the kernel type, k is the number of support vectors, x_i is the input data, r_i is the target class of training dataset and b is the bias term. In this paper we have used radial basis function kernel.

IV. RESULTS AND DISCUSSION

Here we have calculated variance, skewness, kurtosis, fractal dimension and sample entropy for different IMFs obtained as a result of EMD and before that we have segmented each IMF into 8 parts. Here we have considered two cases first one is consisting of normal and ictal EEG signals and the second one is consisting of interictal and ictal EEG signals. The class of discrimination ability of above mentioned parameters for both the cases are performed using Kruskal-Wallis statistical test and the *p*-values obtained as a result of this test are shown in table I and II.

TABLE I
p-values obtained as a result of Kruskal-Wallis statistical test for normal and ictal signal

IMF	SaEn	Kurtosis	Skewness	Variance	FD
IMF1	8.6e-19	2.1e-38	0.0606	1.5e-77	4.3e-39
IMF2	2.7e-15	3.9e-25	0.0783	1.7e-78	1.1e-45
IMF3	0.0811	9.0e-05	0.0349	4.0e-80	1.7e-39
IMF4	0.0342	0.0156	0.0140	9.1e-80	3.9e-51

The calculation of the sample entropy is done by keeping pattern length parameter *m* as 1 and tolerance parameter *r* as .2, after that we have designed a classifier in order to classify the ictal and normal signal as well as interictal and ictal signal and to get the classification accuracy we need the training and the test data, for that we have selected 60% of the dataset for training and remaining data are for testing and the results are shown in table III.

TABLE II
p-values obtained as a result of Kruskal-Wallis statistical test for interictal and ictal signal

IMF	SaEn	Kurtosis	Skewness	Variance	FD
IMF1	7.5e-35	1.3e-61	0.0357	4.1e-79	5.0e-28
IMF2	1.8e-21	1.6e-67	0.0378	8.0e-78	2.8e-54
IMF3	0.0811	2.3e-37	0.0973	1.9e-77	1.2e-17
IMF4	0.2642	4.6e-04	0.4603	5.9e-60	4.8e-12

TABLE III
Classification performances of different cases of EEG signals for the first four IMFs

IMF	Ictal (S) vs Normal (Z) Accuracy(%)	Ictal (S) vs Int. ictal (F) Accuracy(%)
IMF1	100%	100%
IMF2	100%	100%
IMF3	100%	100%
IMF4	100%	99.88%

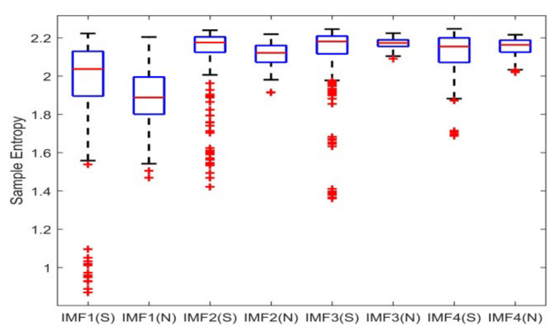

Fig. 5. Box-plot for the comparison of normal and ictal signals using sample entropy values.

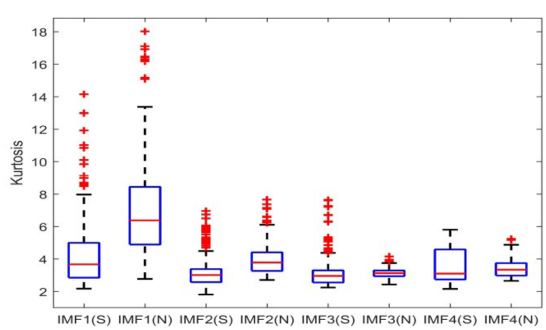

Fig. 6. Box-plot for the comparison of normal and ictal signals using kurtosis values.

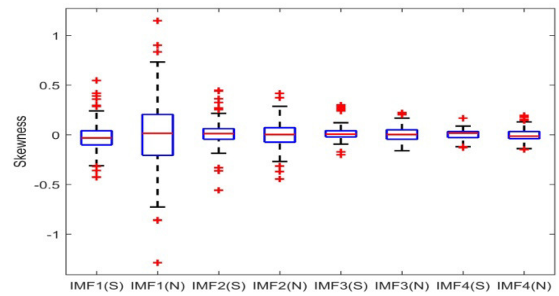

Fig. 7. Box-plot for the comparison of normal and ictal signals using skewness values.

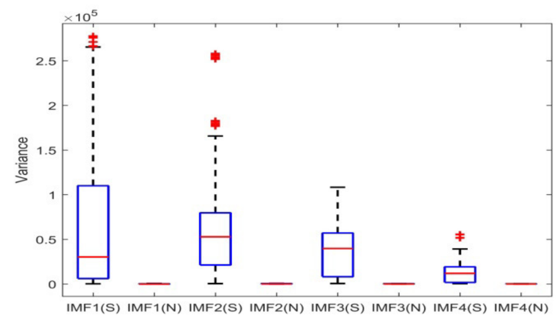

Fig. 8. Box-plot for the comparison of normal and ictal signals using variance values.

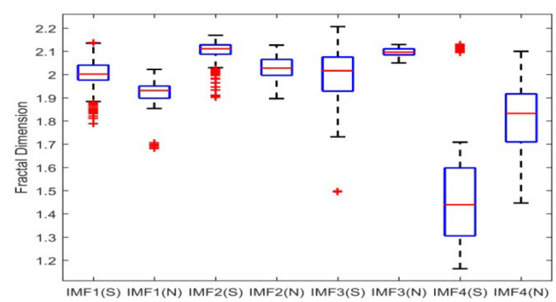

Fig. 9. Box-plot for the comparison of normal and ictal signals using fractal dimension values.

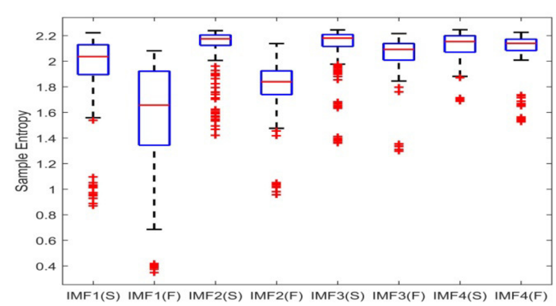

Fig. 10. Box-plot for the comparison of interictal and ictal signals using sample entropy values.

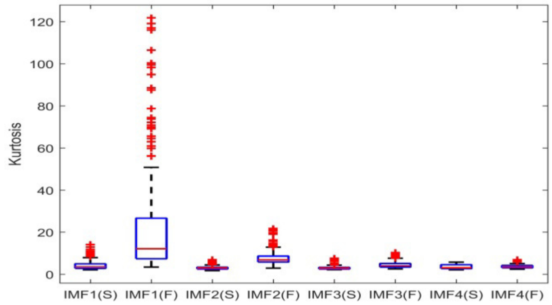

Fig. 11. Box-plot for the comparison of interictal and ictal signals using kurtosis values.

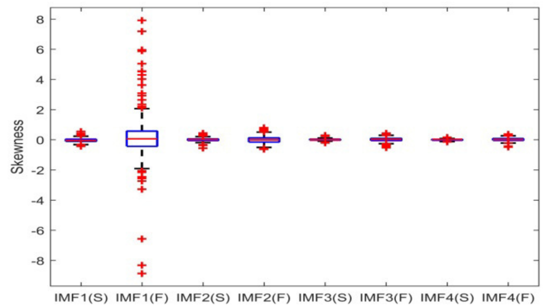

Fig. 12. Box-plot for the comparison of interictal and ictal signals using skewness values.

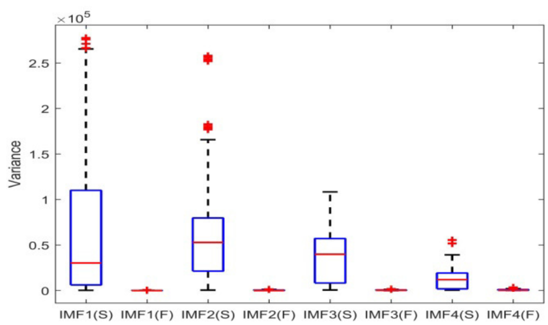

Fig. 13. Box-plot for the comparison of interictal and ictal signals using variance values.

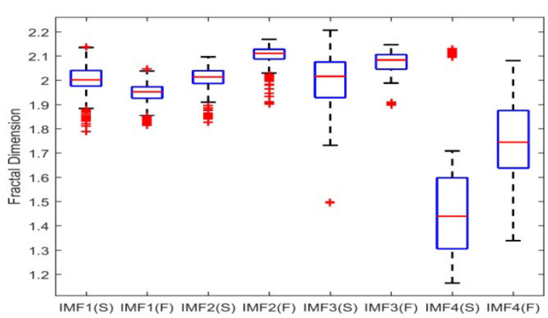

Fig. 14. Box-plot for the comparison of interictal and ictal signals using fractal dimension values.

V. CONCLUSION

In this paper, at the beginning we have applied empirical mode decomposition on the EEG signals to obtain IMFs, the parameters extracted from the IMFs have been used in order to distinguish between normal (Z) and ictal (S) as well as interictal (F) and ictal (S). Here we have calculated above mentioned parameters up to fourth IMF, and got an average accuracy of 100% in order to discriminate between normal and ictal, also got an average accuracy of 99.97% in order to discriminate between interictal and ictal EEG signals.

REFERENCES

[1] P. C. Petrantonakis and L. J. Hadjileontiadis, "Emotion recognition from EEG using higher order crossings," *IEEE Trans. Inf. Technol. Biomed.*, vol. 14, no. 2, pp. 186–197, Mar. 2010.

[2] Y. Li et al., "An EEG-based BCI system for 2-D cursor control by combining mu/beta rhythm and p300 potential," *IEEE Trans. Biomed. Eng.,* vol. 57, no. 10, pp. 2495–2505, Oct. 2010.

[3] Ling Guo, Daniel Rivero, Julian Dorado, Juan R. Rabunal, Alejandro Pazos, "Automatic Epileptic Seizure Detection in EEGs based on Line Length Feature and Artificial Neural Network," *Journal of Neuroscience Methods*, Vol. 191, 2010, pp. 101-109.

[4] Adeli, H., Zhou, Z., Dadmehr, N., "Analysis of EEG records in an epileptic patient using wavelet transform," *J. Neurosci. Methods*,vol . 123, pp.69-87, 2003.

[5] Srinivasan, V., Eswaran, C., Sriraam, N., "Artificial neural network based epileptic detection using time-domain and frequency-domain features," *J. Med. Syst.,* vol. 29, pp. 647-660, 2005.

[6] R. B. Pachori and V. Bajaj, "Analysis of normal and epileptic seizure EEG signals using empirical mode decomposition," *Comput. Methods Progr. Biomed.*, vol. 104, no. 3, pp. 373–381, 2011.

[7] V. Bajaj and R. Pachori, "Classification of seizure and non-seizure EEG signals using empirical mode decomposition," *IEEE Trans. Inf. Technol. Biomed.*, vol. 16, no. 6, pp. 1135–1142, Nov. 2012.

[8] EEG Time Series Download Page 2012 [Online]. Available: http://epileptologie-bonn.de/cms/front_content.php?idcat=193&lang3

[9] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N. C.Yen, C. C.Tung, andH.H.Liu, "The empirical mode decomposition and Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proc. R. Soc. Lond. A,* vol. 454, pp. 903–995, 1998.

[10] S. M. Pincus, "Approximate entropy as a measure of system complexity," *Proceedings of the National Academy of Sciences,* vol. 88, pp. 2297-2301, 1991.

[11] J. S. Richman, D. E. Lake, and J. R. Moorman, "Sample Entropy," vol. 384, pp. 172-184, 2004.

[12] C.-W. Hsu, C.-C. Chang, and C.-J. Lin, A Practical Guide to Support Vector Classification. [Online]. Available: http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf

[13] Vijith V S, Jisu Elsa Jacob, et al., "Epileptic seizure detection using nonlinear analysis of EEG," *IEEE Int. Conf. on Inventive Computation Technol.,* vol. 3, pp. 1-6, 2016.

[14] Mohamed Bedeeuzzaman.V, Omar Farooq, Yusuf Uzzaman Khan, "Automatic Seizure Detection Using Higher Order Moments," *IEEE Int. Conf. on Recent Trends in Information, Telecommunication and Computing*, pp. 159-163, 2010.

[15] George B. Moody, Approximate Entropy. [Online]. Available: http://physionet.caregroup.harvard.edu/physiotools/ApEn/

[16] B. B. Mandelbrot, *The Fractal Geometry of Nature*. San Francisco, CA, USA: Freeman, 1982.

[17] T. Higuchi, "Approach to an irregular time series on the basis of the fractal theory," *Phys. D*, vol. 31, no. 2, pp. 277–283, Jun. 1988.