

# Hybrid Feature Extraction Technique for Multi-Classification of Ictal and Non-Ictal EEG Epilepsy Signals

Sani Saminu<sup>1,3\*</sup>, Guizhi Xu<sup>1</sup>, Shuai Zhang<sup>1</sup>, Abd El Kader Isselmou<sup>1</sup>, Adamu Halilu Jabire<sup>2</sup>, Ibrahim Abdullahi Karaye<sup>1</sup> and Isah Salim Ahmad<sup>1</sup>

<sup>1</sup>State Key Laboratory of Reliability and Intelligence of Electrical Equipment, Hebei University of Technology Tianjin 300130, China

<sup>2</sup>School of Electronics and Information Engineering, Hebei University of Technology Tianjin 300130, China

<sup>3</sup>Biomedical Engineering Department, University of Ilorin, Ilorin-Nigeria

\*Corresponding author: sansam4k@gmail.com

**Abstract:** These Electroencephalography (EEG) signals is an effective tool for identification, monitoring, and treatment of epilepsy, but EEG signals need highly experienced personnel to interpret it correctly due to its complexity, even for an expert it is monotonous and usually consume much time. Therefore, the automatic computer-aided device (CAD) needs to be developed to overcome those challenges associated with epilepsy interpretation and diagnosis. The system efficiency relies largely on the quality of features supply as input to classifiers. This paper presents an efficient feature extraction technique to develop a CAD system that can detect and classify normal, interictal and ictal epilepsy signals correctly with high accuracy. Our approach employs time-frequency features, statistical features and nonlinear features combined as hybrid features to train and test the classifier. Machine learning classifiers of multi-class support vector machine (mSVM) and feed-forward neural network (FFNN) with fivefold cross-validation are used to classifies normal, interictal and ictal with our proposed features. Our system was tested using a publicly available database with three classes each of 100 single channels EEG signals of 4096 samples point each. Based on sensitivity, specificity, and accuracy, our proposed approach of multiclass classification shows a good performance with 96.7%, 98.3% and 100% of sensitivity, specificity, and accuracy respectively.

**Keywords:** EEG, Epilepsy, CAD System, DWT, Entropy.

©2020 Penerbit UTM Press. All rights reserved

*Article History: received 6 June 2020; accepted 13 August 2020; published 29 August 2020.*

## 1. INTRODUCTION

In recent years, EEG signal processing receives much attention by researchers, one of the fields that have been growing fast is EEG epilepsy detection and classification. Epilepsy is one of the chronic neurological diseases which happen due to unprovoked seizures because of outrageous and uncontrolled brain cell activities. It is very difficult to predict its occurrence and as such, it may almost occur at any time. Some noticeable signs and symptoms of epilepsy include loss of consciousness, jerking, loss of memory and uncontrollable motions that results in serious injury and even death. To monitor and alert a patient and doctors of its occurrence is a serious concern by health care practitioners, therefore proper monitoring system is quite desirable to ease the doctor's work and improve the patient life [1-4].

With the recent technological revolution in smart devices and low power portable devices, the need for easy and low-cost devices to be used in developing countries is paramount as a large number of epilepsy patients are from these countries [5]. The accuracy and efficiency of these devices depend highly on the type and quality of features extracted from epilepsy signals as well as on the classifier's efficiency. EEG is believed to be a prime signal

that contains information about the electrical activity of the brain and is widely used to evaluate and asses different parts of brain functions and activities including epileptogenic zones [6,7]. Usually, EEG recording is traditionally recorded on a chart and computer monitors [8], these recordings contain a time-domain representation which does not manifest some useful information about the signal. This may result in misinterpretation of EEG signals which may lead to missing epilepsy signs [7,9,10]. Several studies/systems have been carried out and developed to detect, understand, and classify seizures nature and types.

To overcome those challenges, some of the earlier studies show that EEG signals are nonstationary and exhibit nonlinearity [11] which cannot be well detected by time and frequency domain analysis like Fourier transform [12,13] and short-time Fourier transform [14] to extract features from EEG epilepsy signals. As such, this work addressed some challenges associated with traditional feature extraction techniques by combining time-frequency features with statistical and nonlinear features to provide efficient feature extraction that is suitable for mobile applications and integrated smart facilities in our health care centers.

The remaining of this paper is as follows: related work on EEG epilepsy detection and classification was discussed in section 2. Section 3 addressed background studies related to our proposed approach. Section 4 highlights our system design and architecture while section 5 presents our experimental results and discussion of the results. Finally, conclusion and future direction and improvement were provided in section 6.

## 2. RELATED WORK

Although Last decade has witnessed a lot of attention by researchers on the epilepsy detection and classification with various feature extraction developed as well as improvement and development of numerous types of classifiers to achieve high accuracy in discriminating between different types of epilepsy signals rhythms such as normal pattern, abnormal spikes during the seizures (ictal) and before the seizures (interictal). Gotman and Gloor [15] present one of the pioneering work in automatic detection of seizures in which the EEG signal segments are compared based on the continuous temporal analysis. The time-domain approach has been presented by [16-18] as a means of extracting EEG characteristics as features for detection and classification, but the nonstationary of EEG cannot be well captured by time-domain approaches. Frequency domain approach was reported by [12,13,19,20], time-frequency domain features were reported by [21-24] which they used wavelet transform with approximate entropy for feature extraction and employ artificial neural network for classification [25-30]. Other researchers that used time domain, time-frequency approach for feature extraction includes [30-37].

Recently, some researches focus on nonlinear techniques to capture the nonlinearity nature of the EEG epilepsy signal. Different types of entropy have been investigated such as approximate entropy [38-40], sample entropy [41]. Empirical mode decomposition is used to extract statistical features [42]. [43] used the Empirical Mode Decomposition (EMD) method to extract features of epilepsy EEG signals and used a modified fuzzy clustering algorithm as a classifier to detect epilepsy signals.

There are some drawbacks with some reported work in literature such as either utilized a single domain such as time domain or frequency domain that cannot extract some hidden EEG signals dynamics, others used only binary classification to classify normal and seizure states while some works reported low accuracy. This work aims to address some of the mentioned drawbacks by utilizing the efficacy of time-frequency features combined with statistical and nonlinear features to extract significant and hidden information such as nonlinearity and nonstationarity of the EEG epilepsy signal. Also, we used multi-class SVM and feed-forward neural network to classify normal, interictal and ictal states. Discrete wavelet transform was used to decompose the signal into decomposition levels, statistical features and nonlinear features were extracted from DWT coefficients, the hybrid features are then fed to SVM and FFNN for classification.

## 3. MATERIALS AND METHODS

### 3.1 Discrete Wavelet Transform

Discrete wavelet transform is a quite popular and convenient approach for analyzing biomedical signals including EEG due to its ability to represent a signal in a time-frequency representation. DWT analyzed signals that are nonstationary like EEG signals conveniently. The approach in DWT is to decomposes the Epilepsy EEG signal into signal sub-bands or decomposition levels, thereafter, significant features are calculated and extracted from the decomposition levels selected. Selection of appropriate decomposition levels, type of mother wavelet to be used and identifying type of features from selected sub-bands are some of the challenging issues in using wavelet for classification of normal from seizure periods [44]. Wavelet function can be described as follows:

$$\Psi_{s,\tau} = \frac{1}{\sqrt{s}} \Psi\left(\frac{t-\tau}{s}\right) \quad (1)$$

Where  $\tau$  is called shift parameter while,  $s$  is called scale parameter

Wavelet transform:

$$\gamma(s, \tau) = \int f(t) \Psi_{s,\tau}^*(t) dt \quad (2)$$

Inverse wavelet transform:

$$f(t) = \iint \gamma(s, \tau) \Psi_{s,\tau}(t) d\tau ds \quad (3)$$

Simple Haar wavelet threshold is used in initializing the process of feature extraction. Haar wavelet function can be described as follows [44]:

$$\Psi(t) = \begin{cases} 1, & 0 \leq t < \frac{1}{2}, \\ -1, & \frac{1}{2} \leq t < 1, \\ 0, & \text{otherwise.} \end{cases} \quad (4)$$

To realize the DWT, a series of successive filters were designed by [45], they are made of low and high pass quadrature mirror filters with two downsamplers by 2. Figure 1. shows a successive of low-pass represented by  $g[n]$  and high-pass filter represented by  $h[n]$  for three-level of DWT decomposition. At every level of decomposition, there are two coefficients produced by the filters. Approximations coefficients and detail coefficients produced by low pass and high pass filters respectively, the process of filtering and decimation is continuous up to the last selected level of decomposition [46].

### 3.2 Support Vector Machine

SVM is a machine learning classifier highly suitable for binary classification with feature vectors of high dimension. It is very suitable and popularly used in biomedical signal processing and applications due to its capability of dealing with a high number of predictors and also high accuracy. The distance of optimal hyperplane obtained by SVM from feature space of high dimension and that of each class closest data sample is maximized by SVM [47]. It depends on its regularization parameter which controls the level of overlap between the class and kernel functions which is used for mapping of training data to a feature space of higher dimensional from an input space [48,49]. Figure 2. depicts an example of a 2-dimensional separable classification problem by denoting

the maximum margin and optimal hyperplane. The support vectors are those data points on the margin line [50].

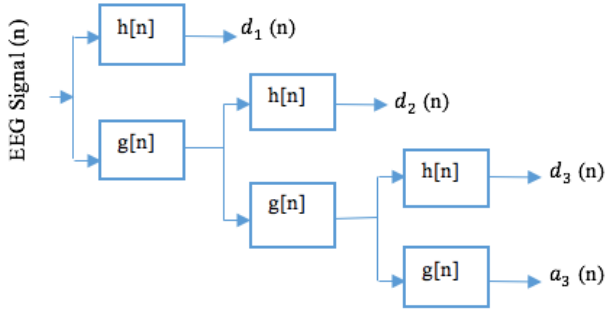


Figure 1. Three level Wavelet decomposition tree

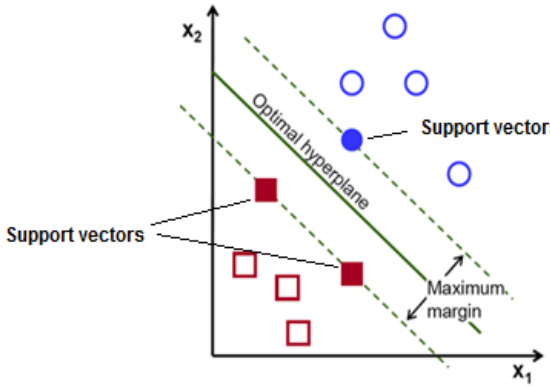


Figure 2. An example of a separable problem in a 2 dimensional space

Some of the common kernel functions used in SVM are linear kernel function with the equation as follows:

$$K(X, Y) = X^T Y. \quad (5)$$

Another type of kernel function is polynomial with degree  $d$  as follows

$$k(X_i, X_j) = (X_i, X_j)^d \quad (6)$$

Where  $d(d \geq 1)$  is the number of polynomials. If the number of polynomials  $d = 2$  or  $d = 3$ , then the function is called quadratic kernel function.

### 3.2.1 Multi Class Support Vector Machine (mSVMs)

Unlike binary classification using SVM, the multi-class classification is a complex as the output is more than one group each of mutually exclusive. Some of the methods developed to solve mSVMs problems include One-Against-All (OAA), Binary Tree (BT), Directed Acyclic Graph (DAG), and One-Against-One (OAO).

OAA SVM problem with  $M$ -class ( $M > 2$ ) was tested in this study. In this method, binary classifiers of number  $M$  are constructed with one for each class. The hyper plane is constructed by  $m^{th}$  classifier between class  $m$  and the remaining  $M - 1$  classes [51].

SVMs optimization problem can be formulated with slack variables  $\xi_i$  as follows:

Find  $w$ ,  $b$ , and  $\xi \geq 0$  such that

$$\Phi(w) = w^T w + C \sum_{i=1}^N \xi_i \text{ is minimized}$$

Where  $w$  is weight vector,  $b$  is a bias term

$$\text{For all } (x_i, y_i), i = 1, \dots, N: s. t. y_i(w^T x_i + b) \geq 1 - \xi_i \quad (7)$$

Where  $C$  is regularization term to control over fitting

Equation (7) can be generalized by minimizing

$$\Phi(w, \xi) = \frac{1}{2} \sum_{m=1}^M (w_m^T w_m) + C \sum_{i=1}^N \sum_{m \neq y_i} \xi_i^m, \quad (8)$$

$$s. t. (w_{y_i}^T x_i) + b_{y_i} \geq (w_m^T x_i) + b_m + 2 - \xi_i^m, \quad (8)$$

$$\xi_i^m \geq 0, \text{ for } i = 1, \dots, N: \quad m \in \{1, \dots, M\} \setminus \{y_i\}$$

The decision boundary can be written as follows:

$$f(x) = \text{argmax}_m [(w_m^T w_m) + b_m] \text{ for } m=1, \dots, M$$

The final classifier is of the form (9) after finding the saddle point of the Lagrangian

$$f(x) = \sum \alpha_i x_i^T x + b_m. \quad (9)$$

Regardless of the number of  $M$  classes, there are only  $N$  coefficients and the number of non-zero  $\alpha$  can be directly reduced by regularization

### 3.3 Artificial Neural Network

The Artificial neural network used a predefined algorithm to mimic learning input-output behavior. Its structures are inspired by the human nervous system and widely used for system modeling and function approximation. Multi-layer feedforward neural network with backpropagation learning is the simplest and most common ANN structure as shown in Figure 3.

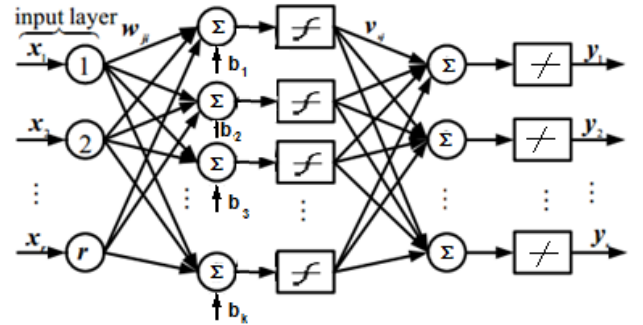


Figure 3. Multi-layer feedforward neural network

The ANN process is to choose the structure of the system, number of layers, and the number of neurons in each layer, training samples are fed to ANN and weights are determined based on the learning algorithm. Steepest descent direction is used by a backpropagation algorithm to adjust its weight and this direction is where the performance function rapidly decreased.

## 4. SYSTEM ARCHITECTURE

Our proposed system overall working principle is depicted in Figure 4. The EEG epilepsy raw data after recording from scalp surface or intracranially is preprocessed to remove the unwanted artifacts and noises as well as baseline corrections and other processing such as dimensionality reduction among others. After preprocessing, the preprocessed signal is passed to feature extraction stage where domain based techniques, statistical, nonlinear and other feature extraction techniques have been applied to extract a meaningful and

relevant features that are supplied to classifiers for classification.

Different types of classifiers are available for classification of epileptic EEG signals into seizure and non-seizure conditions. Finally, the analysis and performance evaluation of classification results is conducted to ascertain the efficiency or otherwise of the system so that monitoring and other diagnosis and development of CAD system would be feasible.

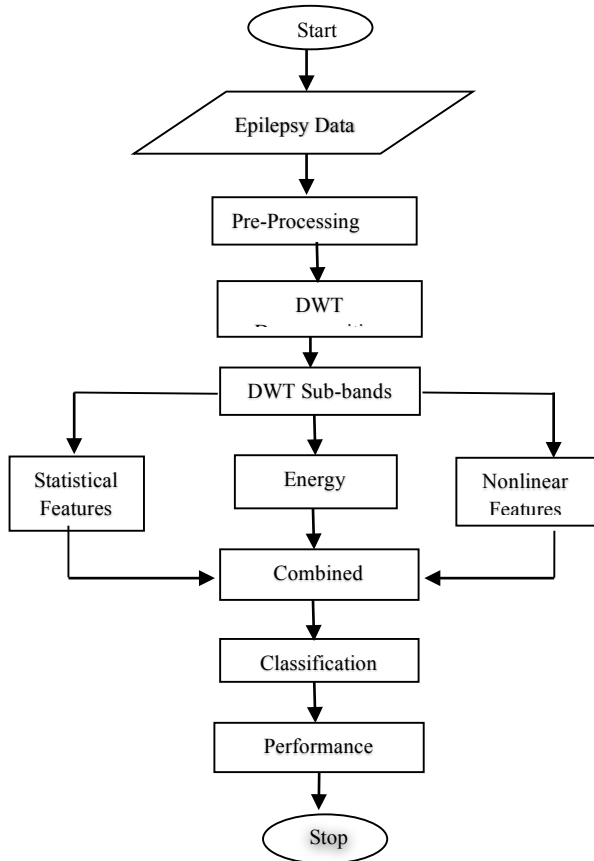


Figure 4. Block diagram of the proposed technique

#### 4.1 EEG Datasets

The experiment for this study was conducted using a well and widely acceptable university of Bonn database which was available online by the Department of Epileptology of Bonn University, Germany. Detail explanation and description of this dataset were given by Dr. Ralph Andrzejak of Epilepsy Center of the Bonn University Hospital of Freiburg [52]. During the signal acquisition, the data were recorded based on the standard 10-20 electrode placement system. The EEG signals are of 23.6s duration with 173.6Hz as sampling frequency which was digitized using a 12 bit A/D converter. This dataset is using five different conditions and then divided into five classes/groups each of 100 single channels. The groups are Z, O, N, F, and S or A, B, C, D, and E classes. This work selected three different classes to test a multi-class classification and compared with binary one, the classes are set Z or A which is recorded from the surface of the scalp of five healthy subjects, set F or C dataset is recorded from five epileptic patients at the time of non-seizure activity (interictal) intracranially during the period of

acquisition and set S or E is recorded during epileptic seizure activity of five pre-surgical epileptic patients (seizure) using the intracranial recording. More information about the dataset can be found at [53] and the description of the epileptic EEG dataset is shown in table 1 while figure 5 shows the example of raw epileptic EEG signal for S and N classes.

Table 1. Description of EEG dataset

Dataset Class	Subject Condition	Epileptogenic Foci
Z	Health	Surface Scalp
F	Seizure-free	Intracranial
S	Intervals Seizure	Intracranial

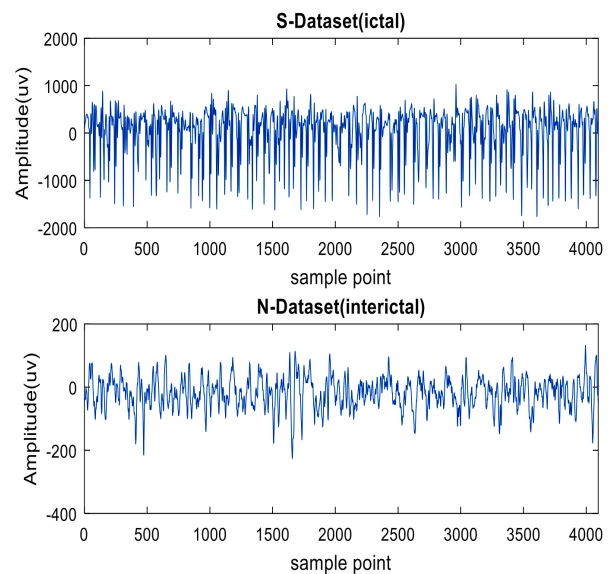


Figure 5. Example of EEG signals from set C and set

#### 4.1 Pre-processing

This stage involves preprocessing the EEG epileptic signal before extracting meaningful features that can be used for classification. It is very essential to denoise the signal by removing artifacts and noise that contaminate the original signal during recording and processing, some of these artifacts include muscle artifacts, environmental artifacts, power line interference, etc. apart from artifacts elimination by EEGLAB software, baseline correction, digital filtering, and signal re-referencing have been performed using the same software.

It is very essential to extract meaningful information with a smaller number of parameters to improve the accuracy of computer-aided device (CAD) system especially smart internet of things (IoT) devices. Therefore, instead of using raw EEG epileptic signals with full-length data which contain a lot of redundancy and large dimensionality that results in a poor signal input representation to classifiers and render the classification tedious and complex. To overcome these issues, time-frequency features have been deployed to decompose the signal into sub-level and thereafter statistical and nonlinear features were extracted which gives a smaller number of parameters by reducing its dimension and upheld the system accuracy.

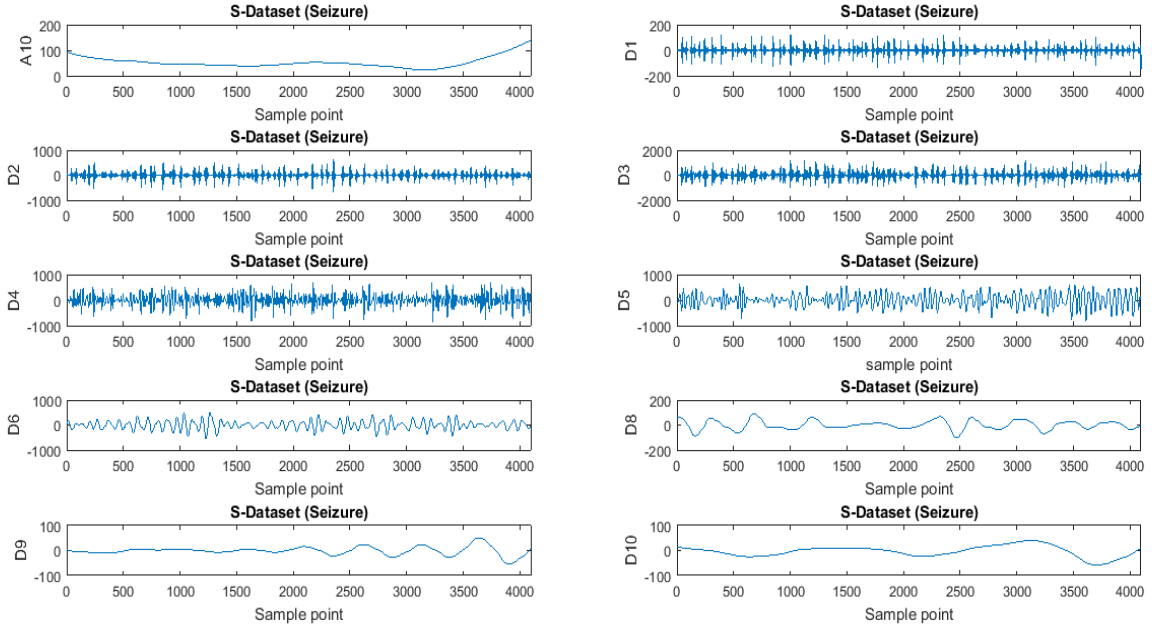


Figure 6. Decomposition levels of DWT

### 4.3 Feature Extraction and Classification

#### 4.3.1 Time-Frequency Features

Time-frequency features are suitable features for biomedical signals to cater to their nonlinearity and nonstationarity. In this paper, DWT has been used as a technique for extracting time-frequency features by decomposing the signal into many sub-bands. To ensure that all the signal levels have been properly captured, Daubechies mother wavelet was chosen as a mother wavelet because it has been discovered in the literature to be very effective in dealing with EEG signal. Also, level 10 decomposition was selected. Figure 6 depicts the decomposition of seizure class into sub-bands.

#### 4.3.2 Statistical Features

At each sub-band of DWT decomposition, some statistical parameters have been selected to further reduce the dimensionality of the signal. Details and approximation coefficients of  $d_1 - d_8$  combined with  $a_8$  respectively were calculated and prepared for further analysis or as input to the classifier. The standard statistical parameters used in this study are: mean, median, mode, maximum, minimum and standard deviation.

#### 4.3.3 Nonlinear Features

Approximate entropy and sample entropy have been further calculated from the DWT statistical features to capture the nonlinearity of EEG epilepsy signal to make it more robust and significance to classifiers.

For approximate entropy, the template vector of size  $m$  is defined as a windowed signal

$$u[i] = [x[i] \ x[i + 1] \ \dots \ x[i + 1 - m]]^T, \quad (10)$$

The self-similarity of the template vector  $u[i]$  is defined as

$$C_i^m(r) = \frac{1}{N-m+1} \sum_{j=0}^{N-m} \Theta(r - \|u[i] - u[j]\|_\infty), \quad (11)$$

Where,  $r$  is the tolerance and  $\Theta(x)$  is the Heaviside step function, i.e.

$$\Theta(x) = \begin{cases} 1, & x \geq 0, \\ 0, & \text{otherwise.} \end{cases} \quad (12)$$

$C_i$  would be high when  $u[i]$  and  $u[j]$  sequences are very near if  $X$  is mostly self-similar. The self-similarity indices are aggregated by Approximate entropy over all shifted possibilities of template vectors given the tolerance and template length.

The Approximate entropy is given as

$$ApEn(X, m, r) = \frac{1}{N-m+1} \sum_{i=0}^{N-m} \log C_i^m(r) - \frac{1}{N-m} \sum_{i=0}^{N-m-1} \log C_i^{m+1}(r). \quad (13)$$

For the sample entropy, it's difference with approximate entropy is that the self-similarity of template vectors for all pairs  $u[i]$  and  $u[j]$  with tolerance  $r$  is defined as

$$\phi^m(r) = \sum_{j=0, j \neq i}^{N-m} \sum_{i=0}^{N-m} \Theta(r - \|u[i] - u[j]\|_\infty). \quad (14)$$

If the signals are self-similar,  $\phi^m(r)$  is high. The SampEn is given as

$$SampEn(X, m, r) = \log \phi^m(r) - \log \phi^{m+1}(r). \quad (15)$$

#### 4.3.4 Energy Features

To further capture the EEG Epilepsy signals significant characteristics and information, energy of the signal was also calculated as features to be used for classification.

The above features are combined in one file and serve as our features for training and testing mSVM and ANN classifiers.

#### 4.4 K-fold Cross-Validation

Cross-validation is a method of testing the validity of machine learning algorithm in which features are divided into two subsets one for training data and the other one for test data. If the data is separated into k-subsets the cross-validation is called k-fold cross-validation. 5-fold cross-validation have been employed in this study as shown in Figure 7.

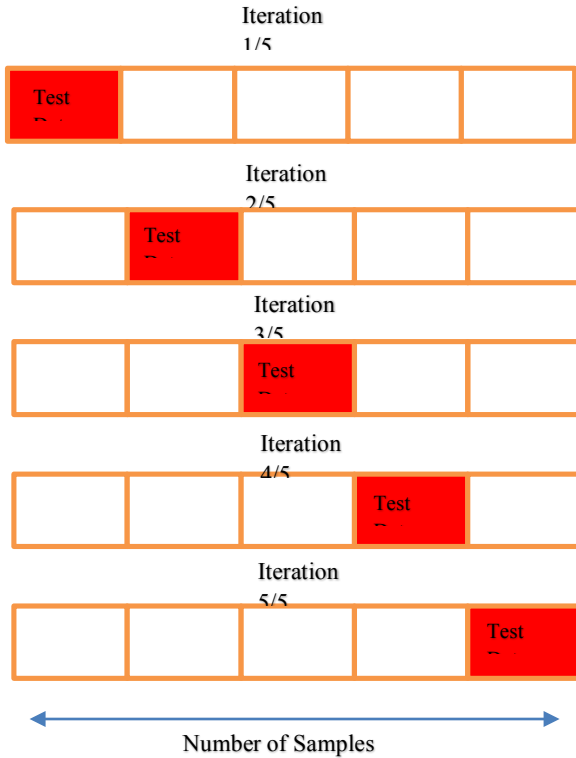


Figure 7. Illustration of 5-fold cross-validation

The features were divided into 5-fold groups as subsets with one subset as test data and the remaining four subsets as training data so that the iteration would be five times.

#### 4.5 Classification

At this stage, ANN and SVM were used for classification of pre-ictal and ictal (normal and seizure) classes. To optimized the classifier's performance and to select efficient and robust classifier for smart IoT devices, different types of SVM classifier as well as different number of neurons were selected and tested.

### 5. EXPERIMENTAL RESULTS AND DISCUSSION

In this paper, EEG epilepsy signal from University of Bonn, Germany database were used for training and testing. 300 EEG records from two classes of normal, pre-ictal and ictal were downloaded as .txt files and distributed randomly for classification of normal and seizure beats. 300 EEG records are fed to artificial neural network and SVM, approximately 70% is used for training and the remaining is used for validation and testing. The classification performance is considered in terms of sensitivity, specificity and accuracy. The used measures are defined as

$$Sensitivity = \frac{TP}{TP + FN} * 100\%$$

$$Specificity = \frac{TN}{TN + FP} * 100\%$$

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} * 100\%$$

where TP is True Positive (correctly identified), FP is False Positive (incorrectly identified), FN is False Negative (incorrectly rejected), TN is True Negative (correctly rejected).

To test the robustness of classifiers as well as to advance from binary to multiclass classification which is much challenging task as the EEG signals is recorded from different subjects at different brain location during ictal, interictal and normal conditions for real implementation in our smart devices to be used by individuals and health care units. Various classification between different groups have been performed as follows:

- Binary classification between class S (ictal) and class Z (normal)
- Binary classification between class S (ictal) and class F (interictal)
- Binary classification between class Z (normal) and class F (interictal)
- Classification between class S (ictal), class F (interictal), and class Z (normal) have been performed as multiclass classification

The detail performance analysis of different group is shown in table and figures

#### 5.1 Performance of S vs Z Group

The performance of seizure condition with normal condition for SVM classifier with different kernel function is shown in table 1, with linear, quadratic, and polynomial function achieved 100% for accuracy, sensitivity, and specificity. Also, various SVM methods have been experimented to optimize the SVM classifier performance as depicted in Figure 8. FFNN classifiers using the proposed scheme with different number of neurons is shown in Figure 9 with almost all the number of neurons tested in this study achieved 100% accuracy, sensitivity, and specificity.

Table 2. Different Kernel functions of SVM for S vs Z

Kernel Function	Sensitivity(%)	Specificity(%)	Accuracy(%)
Linear	100	100	100
Quadratic	100	100	100
Polynomial	100	100	100
Rbf	78.8	78.7	63.3

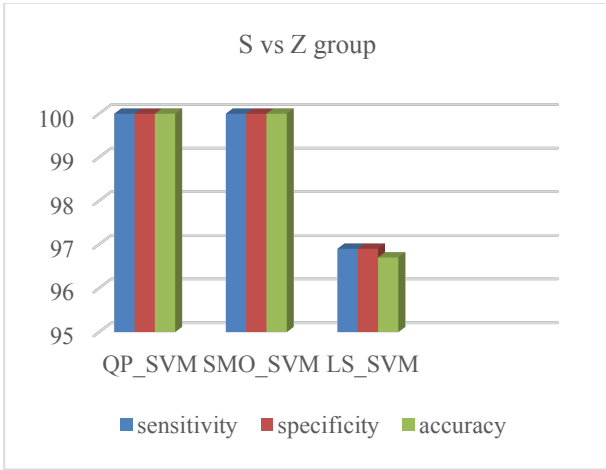


Figure 8. Performance of SVM methods for S vs Z classes

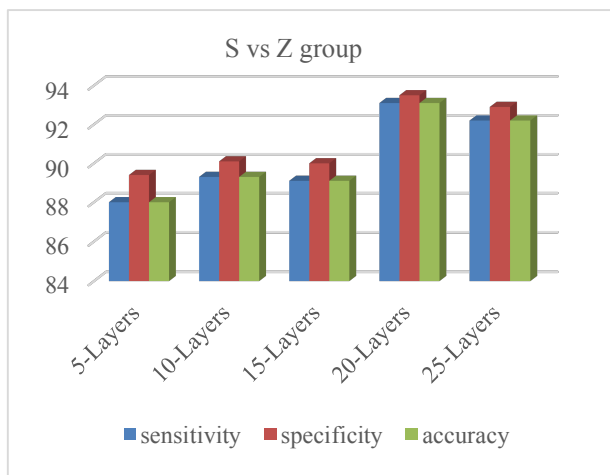


Figure 9. Performance of FFNN layers for S vs Z classes

### 5.2 Performance of S vs F Group

Our proposed approach was tested on ictal and interictal groups with SVM and FFNN classifiers. Table 2 shows the performance of different SVM kernel functions with polynomial function gives the highest accuracy, sensitivity, and specificity of 96.7%, while Figure 10 shows different SVM methods. Furthermore, the performance of FFNN classifier is shown in Figure 11 with 25 number of neurons gives highest accuracy, sensitivity, and specificity of 95.66%, 95.72%, and 95.70% respectively.

### 5.3 Performance of Z vs F Group

To further validate our suggested technique using different epilepsy conditions, group Z of normal signal was compared with interictal group. SVM classifier with linear kernel function outperformed other functions as it attains 100% accuracy, sensitivity, and specificity. While for SVM methods, QP\_SVM and SMO\_SVM scored 100% accuracy, sensitivity, and specificity as shown in table 3 and Figure 12. For FFNN classifier, Figure 13 shows the performance analysis for various number of neurons with 10 number of neurons gives the highest accuracy, sensitivity, and specificity of 77.02%, 96.99%, and 97.20% respectively.

Table 3. Different Kernel functions of SVM for S vs F

Kernel Function	Sensitivity(%)	Specificity(%)	Accuracy(%)
Linear	95.1	95.1	95.0
Quadratic	90.0	90.0	90.0
Polynomial	96.7	96.7	96.7
Rbf	60.0	71.6	71.6

Table 4. Different Kernel functions of SVM for Z vs F

Kernel Function	Sensitivity(%)	Specificity(%)	Accuracy(%)
Linear	100	100	100
Quadratic	88.2	88.2	86.7
Polynomial	89.5	89.5	86.7
rbf	76.3	76.3	55.0

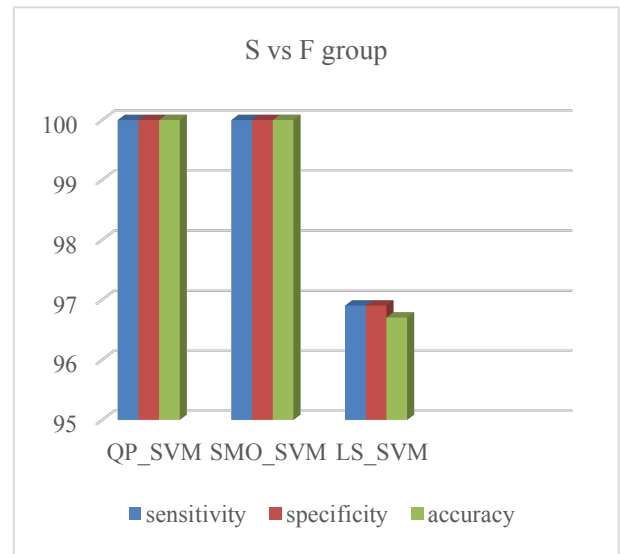


Figure 10. Performance of SVM methods for S vs F classes

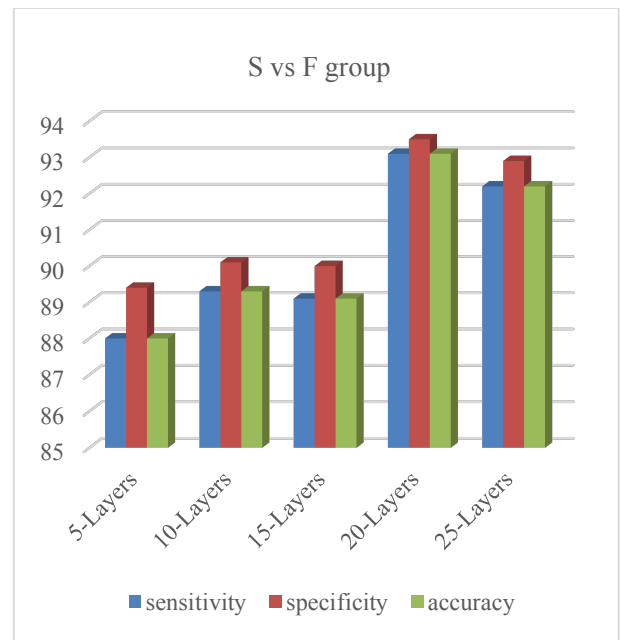


Figure 11. Performance of FFNN layers for S vs F classes

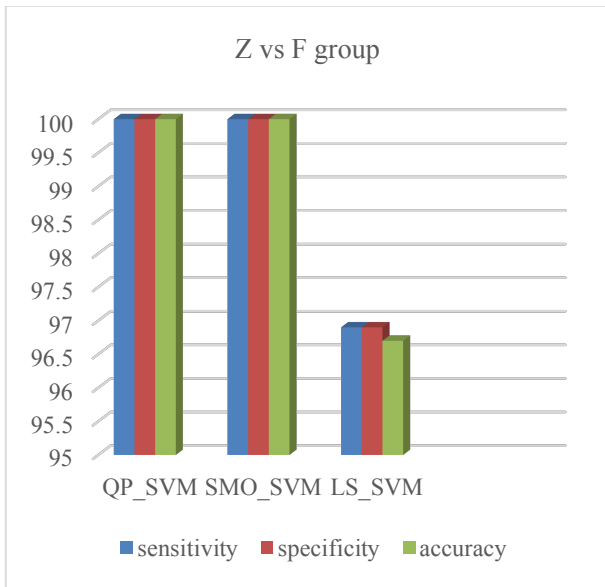


Figure 12. Performance of SVM methods for Z vs F classes

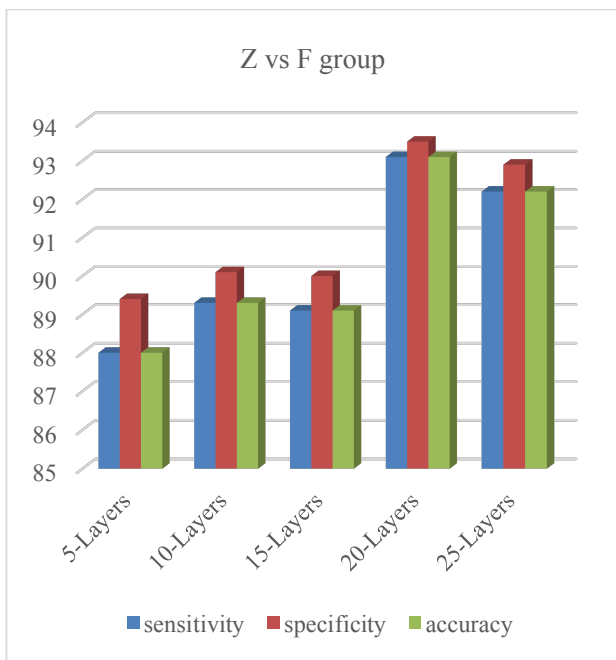


Figure 13. Performance of FFNN layers for Z vs F classes

#### 5.4 Performance of Multiclass Classification

The method proposed in this study have performed very well for binary classification, therefore, to further investigate the robustness of this technique we performed multiclass classification on our epilepsy signals to develop and implement a real clinical application device. The epilepsy signals of ictal, interictal, and normal conditions have been classified using SVM and FNNN chosen classifiers. Multiclass SVM classification have accuracy, sensitivity, and specificity of 100%, 98.33% and 96.67% respectively.

We compare the performance between the chosen classifiers SVM and FFNN to find the best classifier for our epilepsy signals. Best kernel function has been compared with best FFNN system in each classification category as shown in table 5.

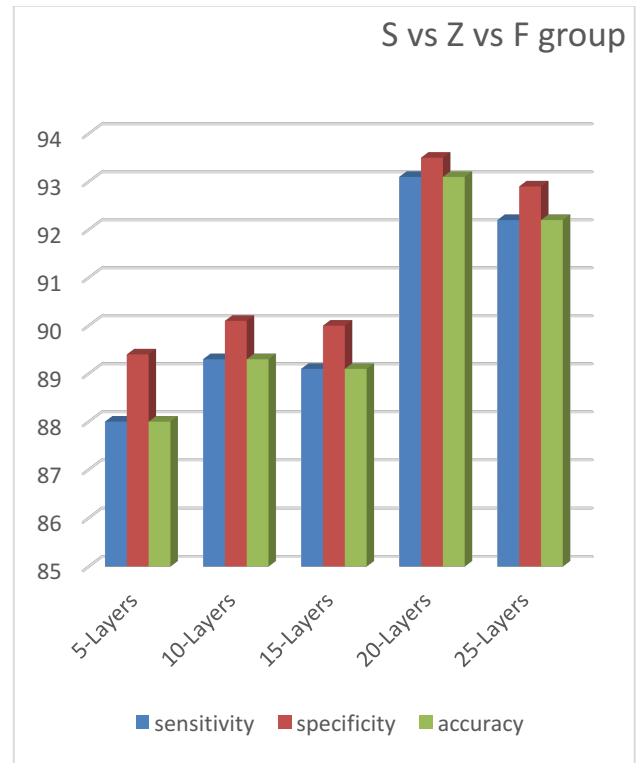


Figure 14. Performance of FFNN layers for S vs Z vs F classes

## 6. CONCLUSION

Due to the increase in the need to develop a suitable CAD system that to be incorporated in today's smart devices for the use in clinical environment and individual for quick, easy, and proper monitoring of epileptic seizure occurrence and condition, this work proposed a suitable feature extraction technique that combine the time-frequency, nonlinear and statistical feature to extract a significant features to characterize the ictal, interictal and normal condition using Bonn University database. Our features were combined and fed into FFNN and SVM classifiers for classification, binary and multiclass classification have been experimented to validate and evaluate our approach which shows a significant improvement in clinical diagnosis to detect various epileptic conditions effectively and accurately as compared to other works in literature as shown in table 6.

This work can be further improved in the future by using large amount of dataset and classes for the system to characterize more epileptic seizure conditions. Also, optimization techniques would be incorporated to optimize both the feature extraction techniques as well as classifiers. Deep learning approach investigation is under way to eliminate the use of manual feature extraction by conventional feature extraction approaches and classifiers.

## REFERENCES

- [1] Berger, H.: 'Über das elektroencephalogramm des menschen', *Eur. Arch. Psychiatry Clin. Neurosci.* 1929, 87, pp. 527–570.
- [2] Fisher, R.S., Boas, W.E., Blume, W., et al.: 'Epileptic seizures and epilepsy: definitions proposed by the international league against epilepsy (ILAE) and the international bureau for epilepsy (IBE)', *Epilepsia*, 2005, 46, (4), pp. 470-472.



Table 5. Comparison between SVM and FFNN

No	Case	Classifier	Sensitivity(%)	Specificity(%)	Accuracy(%)
1	S vs Z	SVM	100	100	100
		FFNN	100	100	100
2	S vs F	SVM	96.7	96.7	96.7
		FFNN	95.7	95.7	95.7
3	Z vs F	SVM	100	100	100
		FFNN	97.0	97.2	97.0
4	S vs F vs Z	SVM	96.7	98.3	100
		FFNN	93.1	93.1	93.5

Table 6. Comparison of proposed method with some recent methods in literature

Authors and year	Method	Sensitivity(%)	Specificity(%)	Accuracy(%)
Alcin, et al, 2016, [54]	Time-frequency image, Filter vector, Extreme learning machine	-	-	96.4
	Kroskov entropy and LS_SVM			
Patidar et al, 2017 [55]	Sample entropy and SVM	97.0	99.5	97.8
Rizal et al, 2018 [41]		-	-	97.7
Fasil et al, 2019 [30]	Time domain exponential Energy features			
Mahjoub et al, 2019 [35]	Linear and nonlinear features, raw data	-	-	99.5
		99.33	96.56	98.78
This study	Time-frequency, Statistical, nonlinear, entropy and Energy	96.7	98.3	100

- [3] Berg, A.T., Berkovic, S.F., Brodie, M.J., et al.: 'Revised terminology and concepts for organization of seizures and epilepsies: report of the ILAE Commission on Classification and Terminology 2005–2009', *Epilepsia* 51 (4) 2010, pp. 676–685.
- [4] Siuly, S., Zhang, Y.: 'Medical big data: neurological diseases diagnosis through medical data analysis', *Data Sci. Eng.* 2016, 1, pp. 54–64.
- [5] 'World Health Organization', <http://www.who.int/news-room/fact-sheets/detail/epilepsy>, Accessed 17 February 2020.
- [6] Gotman, J.: 'Automatic seizure detection: improvement and evaluation', *Electroencephalography and Clinical Neurophysiology*, 1990, 76, (4), pp. 317–324.
- [7] Siuly, S., Li, Y., Zhang, Y.: 'EEG Signal Analysis and Classification: Techniques and Applications', Springer, 2017.
- [8] Gavvala, J., Abend, N., LaRoche, S., et al.: 'Continuous EEG monitoring: a survey of neurophysiologists and neurointensivists', *Epilepsia*, 2015, 55, (11), pp. 1864–1871.
- [9] Zhu, G., Li, Y., Wen, P.P.: 'Epileptic seizure detection in EEGs signals using a fast weighted horizontal visibility algorithm', *Comput. Methods Programs Biomed.* 2014, 115, (2), pp. 64–75.
- [10] Wang, X., Gong G., Li N., Ma Y. 'A Survey of the BCI and Its Application Prospect. In: Zhang L., Song X., Wu Y. (eds) *Theory, Methodology, Tools and Applications for Modeling and Simulation of Complex Systems*', AsiaSim, SCS AutumnSim., Communications in Computer and Information Science, Springer, Singapore, 2016, 646. pp. 102–111.
- [11] Liu, A., Hahn, J. S., Heldt, G. P., et al.: 'Detection of neonatal seizures through computerized EEG analysis', *Electroencephalogr. Clin. Neurophysiol.* 1992, 82, (1), pp. 30–37.
- [12] Kohtoh, S., Taguchi, Y., Matsumoto, N., et al.: 'Algorithm for sleep scoring in experimental animals based on fast Fourier transform power spectrum analysis of the electroencephalogram', *Sleep Biol. Rhythms*, 2008, 6, (3), pp. 163–171.
- [13] Murugappan, M., Murugappan, S., Balaganapathy and Gerard, C.: 'Wireless EEG signals based Neuromarketing system using Fast Fourier Transform (FFT)', *IEEE 10th International Colloquium on Signal Processing and its Applications*, Kuala Lumpur, 2014, pp. 25–30.

- [14] Guerrero-Mosquera, C., Trigueros, A. M., Franco, J. I.: 'New feature extraction approach for epileptic EEG signal detection using time-frequency distributions', *Med. Biol. Eng. Comput.*, 2010, 48, (4), pp. 321–330.
- [15] Gotman, J., Gloor, P.: 'Automatic recognition and quantification of interictal epileptic activity in the human scalp EEG', *Electroencephalogr Clin Neurophysiol*, 1976, 41, (5), pp. 513–529.
- [16] Gotman, J.: 'Automatic recognition of epileptic seizures in the EEG', *Electroencephalogr Clin Neurophysiol*, 1982, 54, (5), pp. 530–540.
- [17] Meier, R., Dittrich, H., Schulze-Bonhage, A.: 'Detecting epileptic seizures in long-term human EEG: a new approach to automatic online and real-time detection and classification of polymorphic seizure patterns', *J. Clin. Neurophysiol*, 2008, 25, (3), pp. 119–131.
- [18] Subasi, A., Gursoy, M.I.: 'EEG signal classification using PCA, ICA, LDA and support vector machines', *Expert Syst. Appl.*, 2010, 37, (12), pp. 8659–8666.
- [19] Minasyan, G. R., Chatten, J. B., Chatten, M. J.: 'Patient specific early seizure detection from scalp EEG', *J. Clin. Neurophysiol*, 2010, 27, (3), pp. 163–178.
- [20] Polat, K., and Günes, S.: 'Classification of epileptiform EEG using a hybrid system based on decision tree classifier and fast Fourier transform', *Appl. Math. Comput.*, 2007, 187, (2), pp. 1017–1026.
- [21] Subasi, A., Alkan, A., Koklukaya, E., et al: 'Wavelet neural network classification of EEG signals by using AR model with MLE preprocessing', *Neural Netw.*, 2005, 18, (7), pp. 985–997.
- [22] Zhang, Y., Liu, B., Ji, X., Huang, D.: 'Classification of EEG signals based on autoregressive model and wavelet packet decomposition', *Neural Process. Lett.* 2016, 45, pp. 365–378.
- [23] Lekshmi, S., Selvam, V., Pallikonda R. M.: 'EEG signal classification using Principal Component Analysis and Wavelet Transform with Neural Network', *International Conference on Communication and Signal Processing, Melmaruvathur*, 2014, pp. 687-690.
- [24] Gajic, D., Djurovic, Z., Di Gennaro, S., Gustafsson, F.: 'Classification of EEG signals for detection of epileptic seizures based on wavelets and statistical pattern recognition', *Biomed. Eng. Appl. Basis Commun.*, 2014, 26, (2), 1450021.
- [25] Yatindra, K., Dewal, M.L., Anand, R.S.: 'Epileptic seizure detection using DWT based fuzzy approximate entropy and support vector machine', *Neurocomputing*, 133, 2014, pp.271-279.
- [26] Rajaguru, H., Tsangavel, V.: 'Wavelets and morphological operators based classification of epileptic risk level', *Mathematical Problems in Engineering*, 2014, 2014, (1), pp. 1-13.
- [27] Karumari, B.K., Vlachos, L., Hu, R., et al.: 'Classification of preictal and interictal period periods based on EEG frequency features in Epilepsy,' 32<sup>nd</sup> Southern Biomedical Engineering Conference, 2016, Shreveport, LA, pp. 9-10.
- [28] Jaiswal, A.K., Banka, H., 'Epileptic seizure detection in EEG using machine learning techniques', *Australas Phys Eng Sci Med*, 2018, 41, pp. 81-94.
- [29] Zhang, T., Chen, W., Li, M.: 'Fuzzy distribution entropy and its application in automated seizure detection technique', *Biomedical Signal Processing and Control*, 2018, 39, pp. 360-377.
- [30] Fasil, O.K., Rajesh, R.: 'Time domain exponential energy for epileptic EEG signal classification', *NeuroScience Letters*, 2019, 694, pp. 1-8.
- [31] Tzallas, A. T., Tsipouras, M. G., and Fotiadis, D. I.: 'Epileptic seizure detection in EEGs using time-frequency analysis', *IEEE Trans. Inform. Technol. Biomed.*, 2009, 13, (5), pp. 703–710.
- [32] Alam, S. M. S., and Bhuiyan, M. I. H.: 'Detection of seizure and epilepsy using higher order statistics in the EMD domain', *IEEE J. Biomed. Health Inform.* 2013, 17, (2), pp. 312–318.
- [33] Kumar, Y., Dewal, M. L., and Anand, R. S.: 'Epileptic seizures detection in EEG using DWT-based ApEn and artificial neural network', *Signal Image Video Process.*, 2014, 8, pp. 1323–1334.
- [34] Karumari, B.K., Vlachos, L., Hu, R., et al: 'Classification of preictal and interictal period periods based on EEG frequency features in Epilepsy', 32<sup>nd</sup> Southern Biomedical Engineering Conference, 2016,
- [35] Mahjoub, C., Le Bouquin Jeannès R., Lajnef T., et al.: 'Epileptic seizure detection on EEG signals using machine learning techniques and advanced preprocessing methods', *Biomed Tech (Berl)*. 2020, 65, (1), pp. 33-50.
- [36] Chua, K., Chandran, V., Acharya, U. R., and Lim, C. M.: 'Application of higher order spectra to identify epileptic EEG', pp. *J. Med. Syst.*, 2011, 35, (6), pp.1563–1571.
- [37] Selesnick, I.W.: 'Wavelet transform with tunable Q-factor', *Ieee Trans. Signal Process.*, 2011, 59, (8), pp. 3560–3575.
- [38] Ocak, H.: 'Automatic detection of epileptic seizures in EEG using discrete wavelet transform and approximate entropy', *Expert Syst. Appl.*, 2009, 36, (2), pp. 2027–2036.
- [39] Guo, L., Rivero, D., and Pazos, A.: 'Epileptic seizure detection using multiwavelet transform based approximate entropy and artificial neural networks', *J. Neurosci. Methods*, 193, (1), pp. 156–163.
- [40] Kumar, Y., Dewal, M.L., Anand, R.S.: 'Epileptic seizure detection using DWT based fuzzy approximate entropy and support vector machine', *Neurocomputing*, 2014, 133, (10), pp. 271–279.
- [41] Rizal, A., Hadiyoso, S.: 'Sample Entropy on Multidistance Signal Level Difference for Epileptic EEG Classification', *Hindawi the Scientific World Journal*, 2018, Article ID 8463256, 6 pages.
- [42] Hassan, A.R, Bhuiyan, M.I.H.: 'Automatic sleep scoring using statistical features in the EMD domain and ensemble methods', *Biocybern Biomed Eng*, 2016, 36, (1), pp. 248–255
- [43] Priyanka, J., Rupali, K.: 'EEG signal classification using modified fuzzy clustering algorithm', *International Journal of Computer Science and Information Technologies*, 2015, 6, (3), pp. 2031-2034.

- [44] Kumar, P., and Agnihotri, D.: 'Biosignal Denoising via wavelet thresholds', IETE Journal of Research, 2010, 56, (3), pp. 132–138.
- [45] Mallat, S.: 'A theory for multi-resolution signal decomposition: the wavelet representation', IEEE Trans. Pattern Anal. Mach. Intell., 1989, 11 (7), pp. 674–693.
- [46] Saminu, S., Özkurt, N.: 'Stationary Wavelet Transform and Entropy-Based Features for ECG Beat Classification', International Journal of Research Studies in Science, Engineering and Technology, 2015, 2, (7), pp. 23-32.
- [47] Cortes, C., Vapnik, V.: 'Support-vector networks', Mach. Learn., 1995, 20, pp. 273–297.
- [48] Dagher, I., 2008. 'Quadratic kernel-free non-linear support vector machine', J. Glob. Optim.