IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received October 23, 2019, accepted November 18, 2019, date of publication November 29, 2019,
date of current version December 26, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2956865

Epileptic Seizure Detection With Permutation Fuzzy Entropy Using Robust Machine Learning Techniques

WAQAR HUSSAIN[1], BIN WANG[1], YAN NIU[1], YUAN GAO[1], XIN WANG[1], JIE SUN[1], QIONGHUI ZHAN[1], RUI CAO[2], ZHOU MENGNI[3], MUHAMMAD SHAHID IQBAL[4], AND JIE XIANG[1]

[1]College of Information and Computer, Taiyuan University of Technology, Taiyuan 030000, China
[2]College of Software, Taiyuan University of Technology, Taiyuan 030000, China
[3]Graduate School of Interdisciplinary Science and Engineering in Health Systems, Okayama University, Okayama 700-8530, Japan
[4]School of Computer Science and Technology, Anhui University, Hefei 23000, China

Corresponding author: Jie Xiang (xiangjie@tyut.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61687023623, Grant 61503272, Grant 61305142, Grant 61741212, and Grant 61373101, in part by the Shanxi Province Key Research and Development Project under Grant 201803D421047, in part by the Natural Science Foundation of Shanxi Province under Grant 2015021090 and Grant 201601D202042, in part by the China Postdoctoral Science Foundation under Grant 2016M601287, and in part by the Shanxi Provincial Foundation for Returned Scholars, China, under Grant 2016-037.

ABSTRACT The automatic and accurate determination of the epileptogenic area can assist doctors in presurgical evaluation by providing higher security and quality of life. Visual inspection of electroencephalogram (EEG) signals is expensive, time-consuming and prone to errors. Several numbers of automated seizure detection frameworks were proposed to replace the traditional methods and to assist neurophysiologists in identifying epileptic seizures accurately. However, these systems lagged in achieving high performance due to the anti-noise ability of feature extraction techniques, while EEG signals are highly susceptible to noise during acquisition. The present study put forwards a new entropy index Permutation Fuzzy Entropy (PFEN), which may delineate between ictal and interictal state of epileptic seizure using different machine learning classifiers. 10-fold cross-validation has been used to avoid the over-fitting of the classification model to achieve unbiased, stable, and reliable performance. The proposed index correctly distinguishes ictal and interictal states with an average accuracy of 98.72%, sensitivity of 98.82% and a specificity of 98.63%, across 21 patients with six epileptic seizure origins. The proposed system manifests the fact that lower PFEN characterizes the EEG during seizure state than in the Interictal seizure state. The study also helps us to investigate the more profound enactment of different classifiers in term of their distance metrics, learning rate, distance, weights, multiple scales, etc. rather than the conventional methods in the literature. Compared to other state of art entropy-based feature extraction methods, PFEN showed its potential to be a promising non-linear feature for achieving high accuracy and efficiency in seizure detection. It also show's its feasibility towards the development of a real-time EEG-based brain monitoring system for epileptic seizure detection.

INDEX TERMS Classification, electroencephalogram (EEG), machine learning, permutation fuzzy entropy (PFEN), seizure detection.

I. INTRODUCTION

Despite the availability of drug and surgical treatment options, epilepsy manifests 1% of the world population [1] as mental and neurological disorder. Epilepsy stood fourth most common neurological syndrome after migraine, spike, and Alzheimer's disease with approximately 2.4 Million people newly diagnosed annually in the world. Epilepsy is an acute, chronic, and recurring neurological disorder hallmarked by frequent unpredictable seizures. Epileptic seizure transpires owing to the abrupt malfunctioning and synchronization of neurons, thereby imitating the excessive and hypersynchronous neuronal activity in the brain [2]. Epileptic seizures do not strike randomly, instead, they emerge from slow

The associate editor coordinating the review of this manuscript and approving it for publication was Quan Zou.

pre- ictal variations in brain excitability, which advance over long timescales and predispose the brain to seizure [3]. Currently, the epileptic disorder is primarily diagnosed by non-invasive electroencephalogram (EEG), whose data is obtained by monitoring brain electrical activity oscillations over time, generated by neural synapses. The international 10-20 system is used to map electrode positions over the scalp. In the past, when epilepsy was diagnosed, patient needing surgery undergo long-term pre-surgical valuations for locating the epileptogenic hubs in the brain. Through the assessment phase, a massive amount of EEG recording is acquired and later visually examined by neurologists for recognizing epileptic seizure information. This information was used to confine epileptogenic centers and to establish which area of the brain needs to be resected during surgery. However, this method was found tiresome, time taking, and high-cost job, particularly when having a large number of patients, making analysis unfeasible. Expert opinion's diversity over a specific pattern perceived in EEG verification increase the diagnostic difficulty over a particular brain disease. For a similar EEG segment, several experts might suggest different diagnoses. Even the same expert might recommend different diagnosis on different evaluations for the same EEG segment. Along with other issues, EEG might contain abnormal patterns that are hard to be visualized by amateur medical specialists. As a result, it is primarily needed to develop automated epileptic seizure detection technique which can eradicate the afore- mentioned issues with the highest level of accuracy, robust- ness, and reliability [4]. Mounting attention has recently been given towards the development of computer-aided diagnostic techniques based on signal processing and machine learning techniques with the aim of swift and accurate classification of seizures and epilepsy, based on non-stationary and non- linear EEG signals [5]. These techniques assistances neuro- physiologists towards accurate and fast detection of presence or absence of disorder [6]. Several highly accurate and robust automatic systems with efficient EEG signal processing have been proposed by researchers to classify seizure, and non- seizure intervals by extracting linear and non-linear features from intracranially recorded EEG signals analysis. The hu- man brain contains highly interconnected nerve cells which makes the brain a non-linear system. Song *et al.* [7] advocate that non-linear EEG analysis of the epileptic zone might deliver strong seizure detection report as epileptogenic zones possess convincing identification of non-linear determinism. Therefore, it will be more appropriate to study brain dynamics using non-linear theory like entropy which reflect the disorder of the dynamic system and help in discriminating stages. Entropy reflects predictability and randomness, with high entropy values refer to less system order and more randomness. The advantage of the entropy-based system is that it needs fewer data to get significant results. Entropy has previously been used for quantitative analysis of EEG signal in various brain diseases like the cognitive task, sleep disease, and other states [8]. Several entropy measures have previously been used for epileptic seizure detection.

Guo *et al.* [9] achieved classification accuracy of about 99% from the epileptic classification of epilepsy using Ap- proximate entropy (ApEn). Ocak [10] implemented epilepsy classifications grounded on Approximate Entropy using wavelet, having 94.3% accuracy. Guo *et al.* [9] calculated Approximate entropy for epilepsy classification achieving an accuracy as high as 99%. Ocak [11] developed a model for automatic detection using approximate entropy feature extraction and got an accuracy of more than 96%. Giannakakis *et al.* [12] proposed a technique for detection of absence epileptic seizures employing Approximate entropy and achieved an accuracy of 90.12%. Hussain *et al.* [13] use sample entropy (SampEn) and approximate entropy to extract nonlinear features based on the k-d (dimension) tree algorithmic approach. Song *et al.* [7] augmented SampEn for epileptic seizure classification with an accuracy of up to 99%. Song and Liò [14] achieved an accuracy of 86% using SampEn to analyze epileptic EEG in conjunction with the extreme learning machine (ELM) for classification. Shen *et al.* [15] achieved an accuracy of 91.18% for epilepsy detection by combining the SampEn feature and support vector machine (SVM) classifier. Song and Zhang [16] deployed sample entropy (SampEn) as a method for feature extraction to detect epileptic seizures with an accuracy of 95%. Kumar *et al.* [17] performed the epileptic signal classification by computing wavelet entropy (WE) to have 94.5% accuracy. Bedeeuzzaman *et al.* [18] used WE to classify preictal and interictal data with 100% sensitivity value. Using Permutation Entropy (PE) and support vector machines (SVM), Nicolaou *et al.* [19] was the pioneer towards performing an epileptic seizure classification who achieved a classification accuracy of 94.38%. Mateos *et al.* [20] developed a framework based on Permutation Entropy to manifest different stages EEG towards the treatment of a chronic epileptic patient. Li textitet al. [21] predict absence seizure using Permutation entropy in genetic absence epilepsy of rats. Xiang *et al.* [22] obtained classification accuracy of 97.16% using SampEn and 98.31% using Fuzzy entropy (FuzzyEn) on epileptic EEG signals. Kumar *et al.* [23] used Fuzzy-ApEn by decomposition in subband for the complexity measure of EEG signal using Support Vector Machine and achieve an accuracy of 97.38%. Distribution entropy (DistEn) was used by Li *et al.* [24] for discrimination normal and interictal seizure over non- overlapping segments specific to EEG signal.

Although, number of entropy variants have been used for detecting epileptic seizure from EEG signal and promising results were accomplished from espousal of entropy variants. It signifies that entropy-based approaches are promising towards EEG analysis of epilepsy. Unfortunately, most of these entropy variants are sensitive to the underlying noise during EEG acquisition, which compromised their classification accuracy. Therefore, it is highly needed to develop

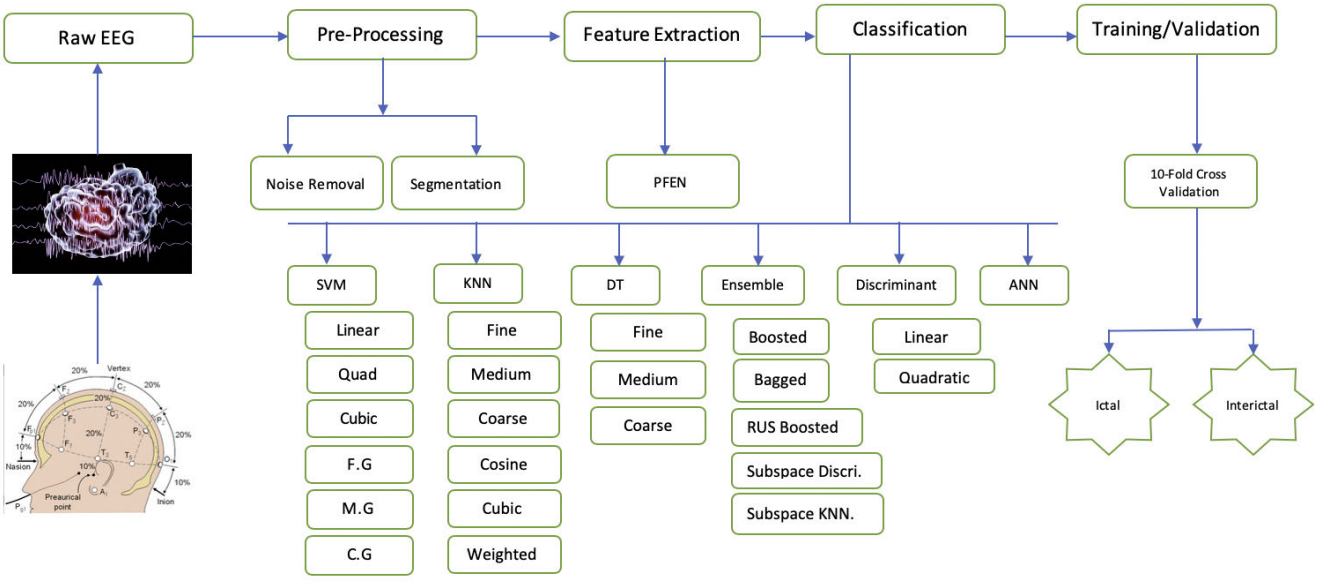


FIGURE 1. A proposed framework to automatically discriminate ictal and interictal seizure from multi-channel EEG signal data.

new entropy variant which can deal with the noisy EEG signals and can perform accurate classification. Also, in past classifiers were just used with the few options and parameters or used with their default values rather than using by fine-tuning their parameter or by using an appropriate variant of the classifier based on their weight, distance or other metrics and their actual ability was undermined. Therefore, the current study aims at validating the use of nonlinear index entropy, as a diagnostic measure to discriminate individual subject's ictal and interictal stages of epilepsy by studying his clinical EEG dataset. We used PFEN characterizing excellent detection property of Fuzzy entropy along with anti-noise improvement ability of Permutation Entropy using the most robust machine learning techniques that outperformed the existing classification techniques. In comparison to Permutation entropy, our index gains much better improvement in results which demonstrate that PFEN analysis of the brain might be a promising prospect towards EEG-based evaluation for classification of seizure. The acquired high detection accuracy and low computational burden signify tremendous aptitudes of the proposed technique for real-time discovery of epileptic seizure.

II. DESCRIPTION OF EEG DATA

Our methodology has been trained, tested, and validated on long-term invasive EEG from 21 patients suffering from medically intractable focal epilepsy, recorded at Epilepsy Center of the University Hospital of Freiburg, Germany. The iEEG dataset obtained from the publicly available Freiburg Seizure Prediction EEG database containing ictal files (the time when seizure onsets) and interictal files (time between seizures), which is now available through the EPILEPSIAE project. The ictal data contains more epileptic component and it would better suit to discriminate between seizure and non-seizure brain activity. There were 2-5 recorded seizure intervals for every patient lasting from few seconds to few minutes. To achieve a better signal-to-noise ratio (SNR) and low artifacts, EEG recordings were made during pre-surgical epilepsy monitoring using grid, strip, and depth electrodes positioned on the cortex of the patients or implanted in their brain. The available data for all 21 patients with six intracranial EEG channels were selected by certified epileptologists. Out of six channels, three focal channels were chosen from inside the epileptogenic zones exhibiting the earliest sign of seizure activity while the remaining three extra-focal electrodes were selected from remote locations which were not at all involved in seizure activity during seizure propagation. The EEG recordings were sampled at a frequency of 256 Hz or 512 Hz and a 16-bit A/D converter annotated by certified epileptologists for seizure onset and epileptiform activities. A total of 87 seizure were analyzed, including 509 hours of interictal and 199 hours of both pre-ictal and ictal EEG data. As the database is heavily dominated by interictal segments, we used an equal number of instances from both ictal and interictal recordings of all 21 patients to make a balance between both the classes. Ictal segments among the seizure onset and seizure end were selected which only contains seizure activity of every patient. For interictal segments, same number of random segments from different interictal sessions of each patient was chosen.

III. SYSTEM ARCHITECTURE

The overall System architecture and epileptic classification process from EEG acquisition to the resulting output are shown as a block diagram in Fig 1.

A. PREPROCESSING

EEG signals usually suffer from complex and low-frequency noises which are not originating from the brain and are considered as disturbances in brain-signal measurements. These artifacts need to preprocess EEG signals and remove noise. To reduce the high-frequency noise and low-frequency artifacts on all the intracranial EEG segments, we apply the bandpass FIR filter between 0.5Hz to 48Hz. The notch filter was applied to eliminate the 50Hz power-line interference. The long-term EEG of each patient was initially divided into 5s non-overlapping windowing to be used as input to the classifier. Each EEG segment represents a feature vector contains time series recording of 6 EEG channels.

B. PERMUTATION FUZZY ENTROPY (PFEN)

Permutation Fuzzy Entropy (PFEN) is a novel index that inherits the seizure detection capabilities from fuzzy entropy and uses anti-noise capabilities of permutation entropy [25]. PFEN employs fuzzy exponential function with soft and continuous boundaries, thus closer the neighbor vector points are, more the similar they are. Firstly, it performs permutation signifying on the EEG signal then it computes the fuzzy entropy over EEG signal [26]. PFEN is calculated by using the following steps:

(1) Assume a time series [X(i): $1 \leq i \leq L$], where L is the length of series X. These time series are used to construct a matrix.

$$\begin{bmatrix} X(1) & x(1+\tau) & \dots & \dots & x(1+(pm-1)\tau \\ X(2) & x(2+\tau) & \dots & \dots & x(2+(pm-1)\tau \\ X(3) & x(3+\tau) & \dots & \dots & x(3+(pm-1)\tau \\ & \dots\dots\dots & & & \dots\dots\dots \\ X(j) & x(j+\tau) & \dots & \dots & x(j+(pm-1)\tau \\ & \dots\dots\dots & & & \dots\dots\dots \\ X(k) & x(k+\tau) & \dots & \dots & x(k+(pm-1)\tau \end{bmatrix} \quad (1)$$

where j = 1,2,3,,,,,,,,,,,,,,k

Where τ and pm are the embedding time delay and the permuted dimension, respectively. K = L - (pm - 1) τ. Each row of matrix is regarded as a reconstruction component. So, there will be K reconstruction components in above matrix.

(2) Arrange elements in ascending order based on values. Thus, we have a new time series constructed from the original time series with the values between 1 and pm!

$$\{U(i) : 1 \leq i \leq 1-(pm-1)*\tau\} \quad (2)$$

(3) Arrange elements in order to reconstruct U (phase-space reconstruction). By considering the length of U is N. The created m-dimensional vector is

$$Y_i^m = \{u(i), u(i+1), \dots\dots\dots\dots\dots u(i+m-1)\} - u_0(i) \quad (3)$$

where $i = 1, 2, \dots, N - m + 1$, m < N - 2, and u0(i) is the average value, which is defined in Eq. (3).

$$u_o(i)\frac{1}{m}\sum_{j=0}^{m-1} u(i+j) \quad (4)$$

(4) The distance d_{ij}^m between vectors Y_i^m and Y_j^m is defined as the largest difference between corresponding elements.

$$d_{ij}^m = d[Y_i^m, Y_j^m] = \max_{k\in(0,m-1)} \{|u(i+k) - u_0(i) - (u(j+k)) - u_0(j)|\} \quad (5)$$

where (i, j = 1 ∼ N-m, j ≠ i) and the degree of similarity d_{ij}^m between vectors Y_i^m and Y_j^m is defined using a fuzzy membership function $\mu(d_{ij}^m, n, r)$.

$$D_{ij}^m = \mu\left(d_{ij}^m, n, r\right) = \exp\left[\frac{-(d_{ij}^m)^n}{r}\right] \quad (6)$$

in this above expression, the fuzzy function $\mu(d_{ij}^m, n, r)$ is an exponential function. Width and gradient of the exponential function are represented by n and r respectively. Define function

$$\emptyset^n(n,r) = \frac{1}{N-m}\sum_{j=1}^{N-m}\left[\frac{1}{N-m-1}\right]\sum_{j=1,j\neq i}^{N-m} D_{ij}^m \quad (7)$$

Increase the reconstruction dimension from m to m+1 and repeat step (3) to

Generate a group of m + 1- dimensional vectors. Define function

$$\emptyset^{n+1}(n,r) = \frac{1}{N-m}\sum_{i=1}^{N-m}\left[\frac{1}{N-m-1}\right]\sum_{j=1,j\neq i}^{N-m} D_{ij}^{m+1} \quad (8)$$

the FuzzyEn of a given series U is defined by Eq. (9)

$$FuzzyEn(m,n,r) = \lim_{N\to\infty}\left[\ln\emptyset^m(n,r) - \ln\emptyset^{m+1}(n,r)\right] \quad (9)$$

when the length N of series U is finite, the estimated value of the corresponding FuzzyEn is Shown in Eq. (9)

$$FuzzyEn(m,n,r,N) = \left[\ln\emptyset^m(n,r) - \ln\emptyset^{m+1}(n,r)\right] \quad (10)$$

where m is the dimensions of the phase space, and r is the similarity tolerance.

The above steps from (1) to (3) signify the original time series X(i) to U (i). Steps (4) to (10) obtain the PFuzzyEn of the original time series X(i) by calculating the FuzzyEn of U (i).

There are four parameters that must be chosen wisely for the calculation of PFuzzyEn i.e. value of m, pm, r and τ. The value of pm should be in the range of 3 to 7 because the smaller value of pm reduces the time complexity of the algorithm under the premise that it is sensitive to changes in the system's transient feature. But if pm is too small the reconstructed sequence contains too few states. Therefore, we have set pm to 4. The time delay function τ was set to 1 in permutation signifying process, the dimension of phase space m was used with 2 and similarity tolerance r was set to 0.25 times the standard deviation of original time series.

C. FEATURE EXTRACTION

To minimize the computational time and complexity of the proposed framework, most significant and prominent feature need to be extracted from every patient's ictal and interictal files for training the classification model. We derived the PFEN as an innovative feature from both groups of EEG signals. Fig 2 demonstrates PFEN value for patient 12. That ictal values of PFEN are found lower than the interictal phase value, signifying that high regularity exists in ictal phase than in interictal EEG time series and much randomness arises in EEG with non-seizure existence than the inception of seizure.

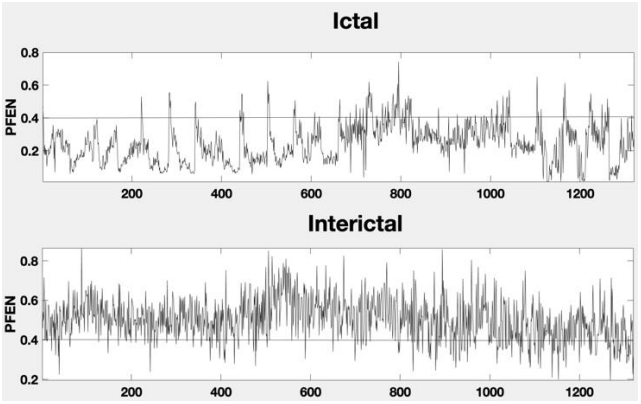


FIGURE 2. Difference in value of Ictal and Interictal phase after extracting Permutation Fuzzy Entropy.

IV. CLASSIFICATION PROCESS

After computing permutation fuzzy entropy as a discriminating feature, the next step was to effectively classify the unseen EEG segments into either ictal or interictal class so as to avoid the tedious and time-consuming screening procedure. The classification was conducted individually for every patient to classify seizure in a patient-specific manner. In all patients, seizure instances were found tremendously small than interictal instance, which leads to unbalance data and could results in poor classification. To avoid large interictal class dominating the ictal class, we used all the instance of ictal class and randomly choose same number of EEG segments from interictal class for all patient. To avoid the problem of over-fitting in classification, we used 10-fold cross validation technique which is assumed to be the best method for validating the accuracy of classifier. The process is repeated for all the classifiers and their performances are shown in Table I–III.

A. SUPPORT VECTOR MACHINE (SVM)

In machine learning, Support Vector Machine (SVM) is a well-known robust supervised learning method based on finite sample theory that maximizes the accuracy of results by avoiding over-fitting of data. Recently SVM is used extensively for binary and multiclass problem in classification and regression majorly in the fields of medical diagnostic area, machine learning, recognition, biometrics etc. and highly suitable for non-linear and high–dimensional data. For classifying different groups of patterns, it creates the hyper-plane in high dimensional space to give largest minimum distance to the training samples called margin in SVM theory. Vectors over margins are known as support vectors which formalize crucial component of training data in classification problem.

Technique for higher classification is to locate optimal hyperplane by maximizing the distance between margin and support vector and by minimizing the classification error. The optimum margin is obtained for the maximized hyperplane. SVM uses Kernel function which has the responsibility of transformation of higher dimension space. Kernel function used by SVM can be linear, radial base function (RBF), polynomial or sigmoid kernel. SVM with RBF and Gaussian kernel function has cost and sigma training parameters. Cost function controls the overfitting of model whereas sigma function controls the degree of non-linearity of the model. Linear kernel has one or several hyperplanes and uses only one parameter i.e. C which is used as a constraint of soft margin representing cost violation constraint association with wrong sided data points over the decision surface. SVM can be used for both linear and non-linear separable data. The non-linear SVM lead to more flexible decision boundary and may have high accuracy, which map input space to higher dimension feature space. Dot product is taken between the input space and some kernel functions. For non-linear mapping polynomial and Radial base functions (RBF) are mostly used. The mathematical expressions of different Machine learning kernel are expressed as below.

SVM sigmoid Kernel.

$$K\,(a_i, b_i) = \tanh(a_i^t, b_i + 1) \tag{11}$$

SVM Polynomial Kernel.

$$K\,(a_i, b_i) = \tanh(a_i, b_i + 1)^n \tag{12}$$

SVM Gaussian (RBF) kernel.

$$K\,(a_i, b_i) = \exp(-\frac{||a_i - b_i||^2}{2\sigma^2}). \tag{13}$$

SVM Fine Gaussian (RBF) kernel

$$K\,(a_i, b_i) = \exp(-\frac{a_i - b_i' ||a_i - b_i||}{\sigma^2}) \tag{14}$$

Here a, b represents the vectors in input space while σ representing width of RBF and n defines the polynomial kernel order.

B. DECISION TREE (DT)

Decision Tree (DT) recursively check the similarities in dataset and classify attribute into best distinctive classes by splitting data and expanding leaf nodes until some termination criterion is met. The choice of splitting is based on comparing impurity of leaf nodes and also on type of impurity being used. Further the size of decision tree is reduced by tree-pruning step to avoid overfitting. Pruning helps to improve the generalization capability of DT. The processing time of decision tree depends on the height of the tree.

TABLE 1. Patient-specific classification accuracy of SVM and ANN (mean ± standard deviation) of all the 21 cases.

Subject	Parameter	SVM						ANN
		Linear SVM	Quadratic SVM	Cubic SVM	Fine Gaussian	Medium Gaussian	Coarse Gaussian	
	Accuracy	86.1±0.61	94.33±0.95	89.03±1.5	85.1±9.04	92.73±2.71	88.93±1.15	98.15±2.14
Pat 1	Sensitivity	87.06±1.5	94.25±3.34	86.45±6.93	86.45±6.93	94.32±4.16	88.71±0.07	98.32±0.1
	Specificity	87.42±2.37	94.29±3.31	81.73±2.12	81.73±2.12	92.68±1.17	89.27±1.03	98.63±0.49
	Accuracy	94.4±3.77	93.73±3.4	97.07±0.58	89.23±8.73	96.67±2.95	96.57±2.31	96.65±2.26
Pat 2	Sensitivity	93.63±0.75	92.66±2.19	88.23±1.27	88.23±1.27	96.32±4.27	94.01±1.29	98.79±0.58
	Specificity	94.9±3.75	94.91±2.97	93.54±0.41	93.54±0.41	98.3±0	94.94±2.91	94.5±0.61
	Accuracy	83.1±1.56	91.2±0.82	89.9±1.15	88.3±5.46	89.03±7.84	79.63±0.64	96.28±2.49
Pat 3	Sensitivity	81.94±0.21	90.96±1.31	96.79±0.02	96.79±0.02	96.06±0.62	95.67±0.06	96.92±0.62
	Specificity	87.67±4.39	90.63±0.58	89.66±3.59	89.66±3.59	92.39±0	63.4±0.63	97.55±0.47
	Accuracy	91.2±1.22	92.83±1.27	91.63±2.43	91.77±2.8	92.4±0.52	91.77±1.1	95.58±1.13
Pat 4	Sensitivity	92.59±4.58	89.12±1.42	92.83±1.02	92.83±1.02	88.09±1.37	86.26±1.28	94.7±1.99
	Specificity	96.21±1.39	96.67±0.98	86.17±0.98	86.17±0.98	97.25±1.3	96.85±1.28	96.97±1.93
	Accuracy	81.3±1.48	89.77±1.77	88.73±1.5	90.47±2.29	88.87±0.93	86.33±1.27	94.73±0.67
Pat 5	Sensitivity	81.44±2.65	95.49±2.24	87.77±0.46	87.77±0.46	96.2±1.99	96.12±1.8	94.65±0.21
	Specificity	79.93±0.75	81.5±3.46	89.5±2.29	89.5±2.29	81.08±5.13	74.64±3.26	94.58±0.29
	Accuracy	99.6±0	99.47±0.12	99.33±0.06	97.87±1.25	99.4±0	99.47±0.06	99.85±0.17
Pat 6	Sensitivity	99.67±0.14	99.42±0.14	94.75±0.43	94.75±0.43	99.5±0	99.75±0	99.89±0.1
	Specificity	99.4±0.26	99.22±0.06	99.42±0.14	99.42±0.14	99.17±0.14	99.33±0.29	99.89±0.1
	Accuracy	93.23±0.98	98.27±1.01	98.03±0.81	98.67±0.76	97.33±0.5	96.77±0.06	99.65±0.44
Pat 7	Sensitivity	94.57±1.78	97.35±2.34	96.1±0.86	96.1±0.86	97.73±2.23	96.68±1.94	99.69±0.27
	Specificity	96±4.85	99.13±0.57	99.38±0.24	99.38±0.24	99.37±0.41	98.26±1.13	99.69±0.27
	Accuracy	92.87±1.29	91.9±0.1	82.47±0.38	84.57±7.47	88.6±0.4	88.83±0.21	99.5±0.71
Pat 8	Sensitivity	95.83±0.98	92.57±2.88	90.53±1.96	90.53±1.96	94.08±2.88	87.43±0.46	97.14±4.09
	Specificity	92.8±1.13	86.77±1.07	58.58±1.7	58.58±1.7	81.65±0.21	88.85±1.99	99.67±0.29
	Accuracy	95.97±0.06	96.6±0	96.3±0.46	96.33±0.06	96.5±0	96.47±0.06	97.35±0.98
Pat 9	Sensitivity	96.09±0.08	95.27±1.33	96.29±0.27	96.29±0.27	96.65±0.04	96.53±0.06	97.03±0.2
	Specificity	96.4±0.43	96.53±0.2	96.2±0.17	96.2±0.17	96.24±0.1	96.18±0.1	97.03±0.2
	Accuracy	83.4±1.04	91.3±0.92	90.27±0.31	88.83±1.29	86.9±0.53	86.43±0.23	96.4±0.96
Pat 10	Sensitivity	96.51±1.98	95.64±2.15	84.6±1.56	84.6±1.56	96.84±1.7	97.75±1.15	96.72±1.02
	Specificity	68.65±1.58	83.62±1.97	89.43±0.29	89.43±0.29	80.47±1.15	71.13±2.14	97.05±1.6
	Accuracy	93.17±0.06	95.37±0.06	94.67±0.7	95.37±0.06	95.17±0.59	94.57±0.06	99.68±0.65
Pat 11	Sensitivity	97.6±0.17	97.86±0.27	98.67±0.06	98.67±0.06	98.63±0.12	98.92±0.03	99.17±0.91
	Specificity	94.73±5.14	90.37±1.93	91.15±1.7	91.15±1.7	92.13±0	90.01±0.09	99.53±0.4
	Accuracy	98±0	98.33±0.29	98.1±0.46	98.17±0.4	97.8±0.17	97.2±1.14	99.1±0.62
Pat 12	Sensitivity	98.02±0.28	98.71±0.13	96.57±3.52	96.57±3.52	98.25±0.13	98.33±0.26	99.42±0.5
	Specificity	97.55±0.3	97.8±0.13	97.86±0.28	97.86±0.28	97.78±0.14	96.58±0.39	99.42±0.5
	Accuracy	95.83±0.12	96.73±0.29	96.27±0.23	95.9±0.5	95.9±1.21	94.5±0	98.48±0.15
Pat 13	Sensitivity	95.31±0.08	97.15±0.28	98.28±0.56	98.28±0.56	96.48±0.31	93.41±0.08	97.54±1.28
	Specificity	95.86±0.05	97±0.27	96.15±0.18	96.15±0.18	96.26±0.36	95.6±0	98.27±0.03
	Accuracy	100±0	100±0	100±0	99.33±1.15	100±0	100±0	100±0

TABLE 1. (*Continued.*) Patient-specific classification accuracy of SVM and ANN (mean ± standard deviation) of all the 21 cases.

Pat 14	Sensitivity	100±0	100±0	98.41±1.38	98.41±1.38	100±0	100±0	100±0
	Specificity	100±0	100±0	98.4±2.77	98.4±2.77	100±0	100±0	100±0
	Accuracy	93.27±0.06	95.47±0.06	94.4±0.17	95.53±0.06	94.27±0.06	94.3±0	98.73±1.66
Pat 15	Sensitivity	93.63±1.62	95.47±1.67	94.8±2.03	94.8±2.03	93.86±3.74	94.8±1.73	98.08±1.75
	Specificity	91.48±0.5	91.5±1.73	89.87±0.64	89.87±0.64	90.61±1.93	89.81±1.73	98.45±1.34
	Accuracy	99.47±0.12	99.47±0.12	98.6±0	97.03±4.45	99.8±0	99.73±0.12	100±0
Pat 16	Sensitivity	99.58±0	99.31±0.24	83.6±0.95	83.6±0.95	99.58±0	99.59±0.01	100±0
	Specificity	99.31±0.24	99.31±0.24	99.73±0.23	99.73±0.23	99.59±0.01	100±0	100±0
	Accuracy	90.23±0.12	97.63±0.06	96.87±0.23	97.77±0.12	96.67±0.06	96.57±0.06	98.63±0.49
Pat 17	Sensitivity	65.84±48.71	98.29±0.13	98.13±0.12	98.13±0.12	98.29±0.07	97.83±0.06	98.68±0.16
	Specificity	86.91±0.08	97.21±0.01	98.96±1.62	98.96±1.62	97.02±0.06	95.35±0	98.77±0.03
	Accuracy	92.4±1.39	95.77±0.98	95.3±1.39	95.83±0.06	92.77±0.06	92.8±0	100±0
Pat 18	Sensitivity	96.84±0.42	97.5±1.99	97.99±1.57	97.99±1.57	98.38±1	97.8±0	98.53±1.27
	Specificity	88.35±0.57	95.52±2.83	89.88±0.31	89.88±0.31	90.79±1.89	84.8±1.73	100±0
	Accuracy	81.6±0.87	90.67±0.29	90.53±0.81	93.83±1.15	86.83±8.95	75.9±0.52	95.3±1.51
Pat 19	Sensitivity	95.37±1.1	96.33±1.15	95±0	95±0	97±0	100±0	95.2±0.17
	Specificity	79.67±12.7	86.33±1.15	95.67±0.58	95.67±0.58	86.33±0.58	52±0	93.53±3.06
	Accuracy	90.1±0.69	96.33±0.58	94.83±0.58	95.8±0.87	95.17±1.27	93.03±0.58	99.65±0.41
Pat 20	Sensitivity	87.53±0.06	95.08±0.19	95.98±0.06	95.98±0.06	94.35±0.07	88.59±0.79	97.97±1.78
	Specificity	91.76±0.28	96.78±0.14	94.61±0.13	94.61±0.13	97.32±0	97.02±0.73	100±0
	Accuracy	95.8±0	96.8±0.35	97.2±0.69	96.4±1.04	97.2±0.35	97.4±0.35	99.85±0.3
Pat 21	Sensitivity	95.19±0.01	96.37±0.06	87.58±1.36	87.58±1.36	96.37±0.06	94.8±0.69	99.91±0.08
	Specificity	96.37±0.06	97.17±0.75	97.4±0.35	97.4±0.35	98.77±0.06	99.2±0.69	99.9±0.09
	Accuracy	91.95±0.74	95.33±0.64	94.26±0.69	93.91±2.33	94.29±1.39	92.72±0.47	98.26±0.84
Average	Sensitivity	92.58±3.2	95.94±1.21	93.3±1.26	93.3±1.26	96.52±1.18	95.38±0.56	98.02±0.81
	Specificity	91.49±1.94	93.92±1.16	92.06±0.99	92.06±0.99	93.58±0.7	89.2±0.96	98.26±0.56

Mathematically, decision tree uses following expression for constructions.

$$A' = \{A_1, A_2, A_3, , , , , , , , , , , , , , , A_n\}^t \quad (15)$$

$$A_i = \{a_1, a_2, a_3, , , , , , , , , , , , , , , a_n\}. \quad (16)$$

$$B = \{B_1, B_2, B_3, , , , , , , , , , , , , , , B_n\} \quad (17)$$

where n denotes the number of observations, n denotes the number of independent variables, t represents the transpose, B represents the m-dimensional vector predicted from A′. A′ representing the ith n-dimensional component autonomous variables a1, a2, a3..... An DT are used to predict observations of A′ Large number of DT can be constructed from A′ having different accuracy values but still optimally best DT is challenging due to large dimensional search space. Generally, a trade off exists between the accuracy and its complexity. The optimized DT will be constructed as a result of local optimal decision regarding feature parameters used in partitioning.

C. ENSEMBLE CLASSIFIER

These are the machine learning paradigm comprising a number of individual sets of trained classifiers whose predictions are combined together when classifying some new instances. These classifiers are successively used in a variety of prediction applications like predicting signal peptide, subcellular location prediction and in making protein subcellular location predictions. According to [28] there are several applications where the combined classification approaches give more accurate and more efficient classification results than the individual classifier results. Individual classifiers are diverse and can make several errors during the classification process; however, in combined classification approaches the error produced by one classifier can be compensated by other classifier. In this way the errors can be reduced by combining several classifiers. In ensemble classifiers individual classification decisions are combined by weighted or unweighted voting to classify new samples. Ensemble methods differs to each other in term of their inter classifier relationship, combining methods, diversity generation

TABLE 2. Patient specific classification accuracy of KNN and discriminant (mean ± standard deviation) of all the 21 cases.

Subject	Parameter	KNN						Discriminant	
		Fine KNN	Medium KNN	Coarse KNN	Cosine KNN	Cubic KNN	Weighted KNN	Linear	Quadratic
Pat 1	Accuracy	88.63±1.67	91.2±2.25	51.83±1.15	88.03±1.15	94±2.25	92.73±1.15	88.03±1.15	90.83±1.15
	Sensitivity	96.78±2.14	99.41±1.03	100±0	88.73±0.05	99.41±1.03	98.15±0.06	90.58±0.03	88.54±3.48
	Specificity	79.51±1.15	83.36±1.15	0±0	87.91±1.2	89.13±1.15	93.61±7.79	86±1.21	89.13±1.15
Pat 2	Accuracy	89.4±2.05	93.07±2.05	83.27±3.19	92.47±2.67	95.03±1.31	95.83±2.57	93.4±2.29	97.8±2.6
	Sensitivity	87.9±5.72	95.54±2.26	97.91±1.38	88.58±5.16	94.26±3.4	94.53±3.86	87.28±2.92	87.37±10.33
	Specificity	90.83±4.04	86.7±8.61	56.8±5.2	88.19±7.96	90.88±4.13	89.2±7.95	95.94±2.85	96.53±2.3
Pat 3	Accuracy	89±1.05	86.67±1.05	74.43±0.82	84.67±1.1	86.83±0.29	90.47±0.29	83.6±0.35	83.8±1.6
	Sensitivity	92.13±0.64	94.61±0.01	94.23±0.64	79.52±0.03	94.6±0	94.6±0	80.99±0.61	85.5±0.52
	Specificity	86.93±0.03	80.79±0.62	53.61±0.6	82.97±4.39	84.4±4.42	87.6±1.56	87.27±1.27	85.8±0
Pat 4	Accuracy	90.23±0.49	91.53±0.38	87.33±0.36	91.27±1.36	91.53±0.9	92.07±0.74	89.4±1.04	91.57±0.99
	Sensitivity	86.12±3.46	89.09±0.51	88.06±2.15	86.42±2.64	87.49±0.72	88.93±1.78	80.86±0.91	89.65±1.15
	Specificity	85.46±7.03	93.83±0.53	86.63±0.4	94.25±3.9	94.62±0.72	90.09±5.44	97.86±1.15	93.38±0.88
Pat 5	Accuracy	86.5±0.85	87.7±0.42	82.17±0.81	86.33±2.12	87.67±1.86	89.1±2.23	82.1±1.61	90.17±1.48
	Sensitivity	88±3.12	97.64±1.45	97.84±0.9	84.53±4.91	95.9±2.94	95.79±2.76	83.63±2.57	93.48±3.4
	Specificity	80.89±3.09	74.53±5.39	62.73±5.95	76.91±8.86	78.24±1.8	91.75±5.68	75.74±4.38	83.51±2.37
Pat 6	Accuracy	98.5±0.49	98.7±0.21	95.4±2	99.03±0.12	98.83±0.21	98.7±0.17	99.47±0.06	99.37±0.06
	Sensitivity	99.17±0.14	99.75±0	99.75±0	99.25±0	99.75±0	99.75±0	99.75±0	99±0
	Specificity	97.83±0.14	97.92±0.14	90.59±0.16	98.83±0.14	97.83±0.29	84.07±11.85	99.08±0.14	99.58±0.14
Pat 7	Accuracy	97.4±2.11	98.43±1.04	97.33±0.87	96.2±0.7	98.33±0.72	98.47±0.84	92.93±1.1	94.33±0.64
	Sensitivity	98.18±1.83	98.08±2.18	97.95±1.86	93.92±0.74	97.98±2	97.94±2.06	96.34±1.83	96.11±2.24
	Specificity	96.62±2.07	98.63±0.84	96.7±0.69	98.18±0.05	98.99±0.5	98.56±0.97	88.67±0.29	90.28±2.92
Pat 8	Accuracy	79.83±1.22	85.03±0.21	57.7±0.17	82.8±0.26	85±0.26	85.27±0.21	94.27±0.06	93.47±0.06
	Sensitivity	72.87±2.88	83.47±2.89	16.87±1.1	90.53±1.96	81.47±4.62	81.47±4.62	90.94±0.31	92.57±2.89
	Specificity	76±6.41	82.18±1.11	99.93±0.12	79.58±6.04	82.55±3.54	87.97±18.59	92.94±3.78	95.32±3.61
Pat 9	Accuracy	94.37±0.46	96.23±0.29	95.97±0.1	96.07±0.06	96±0	96±0	95.7±0	95.53±0.06
	Sensitivity	95.2±0.26	97.37±0.06	96.53±0.06	96.25±0.1	96.82±0.27	96.71±0.08	96.13±0	96.5±0
	Specificity	93.43±0.46	94.87±0.06	95.07±0	95.16±0.47	96.22±1.15	88.13±5.95	95.35±0.13	92.1±2.08
Pat 10	Accuracy	85.4±46.49	88.2±1.14	81.07±0.74	89±0.52	91.3±5.85	89.45±1.06	84.33±0.67	89.97±0.58
	Sensitivity	87.93±1.44	98.47±1.15	99.13±0.12	92.67±1.33	97.07±2.1	95.21±1.13	91.4±1.13	94.09±1.14
	Specificity	79.47±1.15	75.37±1.15	59.1±1.56	81.13±2.48	77.03±1.41	91.63±9.12	76.33±5.95	91.37±8.89
Pat 11	Accuracy	93.27±0	94.6±0.4	93.57±0.46	94.1±0	94.7±0	95±0	92.97±0.06	93.3±0.52
	Sensitivity	94.09±0.09	98.87±0.06	98.9±0	96.49±0.27	98.79±0.1	98.69±0.08	96.47±0.12	94.57±0.06
	Specificity	92.4±0	90.46±0.07	88.1±0.05	90.7±0.52	91.05±0.51	82.33±7.77	89.4±0	92.72±1.23
Pat 12	Accuracy	97.5±0.92	98.27±0.12	98.07±0.6	98.13±0.12	98.13±0.06	97.73±0.64	98.07±0.12	98.33±0.12
	Sensitivity	97.42±1.84	98.76±0.28	98.4±0.39	98.49±0.24	98.78±0.27	98.86±0.39	96.73±1.44	98.53±0.12
	Specificity	97.27±0.79	95.71±0.34	96.06±1.84	97.69±0.39	97.5±0.39	91.23±0.2	94.97±3.99	95.79±2.07
Pat 13	Accuracy	95.97±0.2	96.07±0.23	91.03±0.7	96±0	95.87±0.06	96.07±0.06	95.2±0	96.27±0.12
	Sensitivity	96.52±0.28	97.19±0.35	92.03±0.12	96±0.09	96.99±0.53	96.35±0.57	93.67±1.33	97.63±0
	Specificity	94.47±0.27	94.67±0.09	90.03±0.02	95.15±0.91	94.84±0.73	97.15±0.99	96.27±0.92	97.43±2.02
	Accuracy	100±0	100±0	94.2±0	100±0	100±0	100±0	100±0	100±0

TABLE 2. (*Continued.*) Patient specific classification accuracy of KNN and discriminant (mean ± standard deviation) of all the 21 cases.

Pat 14	Sensitivity	100±0	100±0	97.6±0	100±0	100±0	100±0	100±0	100±0
	Specificity	100±0	100±0	84.95±0.05	100±0	100±0	100±0	100±0	100±0
	Accuracy	92.2±3.64	93.23±2.02	93.6±0.69	93.57±0.15	93.7±0.3	95±0	92.83±0.58	93.23±0.55
Pat 15	Sensitivity	95.24±1.93	97.16±1.06	97.04±1.73	93.02±1.63	97.04±1.73	96.82±1.35	91.47±1.93	92.33±2.23
	Specificity	89.63±2.03	87.17±1.63	86.94±1.83	88.37±3.42	87.43±1.19	88.36±2.81	91.59±1.54	90.56±1.54
	Accuracy	99.17±0	99.2±0	99.13±0.29	99.4±0	99.17±0.75	99.4±0	99.4±0	99.6±0
Pat 16	Sensitivity	99.17±0	99.58±0	99.72±0.24	99.03±0.24	99.58±0	99.58±0	99.58±0	99.65±0.13
	Specificity	99.17±0	99.31±0.47	95.47±2.13	99.45±0.48	99.58±0	99.44±0.24	99.17±0	99.86±0.24
	Accuracy	96.9±0.29	97.73±0.4	94.37±0	92.8±0.17	97.73±0.12	97.53±0.46	84.5±0.17	96.8±1.04
Pat 17	Sensitivity	96.93±0.33	99.3±0	98.5±0.01	90.95±0.61	99.18±0.19	99.11±0.06	82.22±0.43	94.48±0.55
	Specificity	96.51±0.01	95.78±0.13	90.37±0.22	95.58±0.4	96.27±0.4	96.46±0.06	87.34±0	97.33±0.01
	Accuracy	93.93±2.46	94.93±1.15	86.37±0.87	96.13±1.62	95.27±2.02	95.63±0.75	91.23±0.75	95.87±1.15
Pat 18	Sensitivity	91.17±1.15	96.29±2.41	94.54±0.77	97.29±1.55	95.8±2.84	95.07±1.58	95.53±1.1	97.99±1.57
	Specificity	95.73±1.53	94.06±1.14	78.61±1.15	95.01±1.55	94.77±1.15	94.77±1.15	86.41±0.68	93.79±0.7
	Accuracy	90.33±0	80.6±0	63.2±0.58	81.1±0.87	80.77±1.15	84.6±0.87	80.1±0	90.33±0.58
Pat 19	Sensitivity	98.33±0.58	100±0	100±0	91.33±2.31	100±0	100±0	93.67±0.58	96.67±0.58
	Specificity	82.67±0.58	60.67±1.15	25.33±0.58	62.33±5.77	65.67±2.31	66.67±0.58	66±0	86.33±0.58
	Accuracy	94.2±0.75	96±0.75	93.97±0.75	95.2±0.52	96.13±0.4	96.17±0.46	89.6±0.52	91.27±0.64
Pat 20	Sensitivity	93.88±0.07	93.93±0.12	88.78±0.31	95.49±0.16	93.48±0.03	93.48±0.03	83.08±0.08	85.24±0.13
	Specificity	94.18±0	97.2±0.17	98.6±0	96.34±2.18	97.59±0.27	97.59±0.27	95.11±0	96.6±0
	Accuracy	94.4±2.08	97.2±2.08	91.6±1.73	96.8±0.35	97±0	98.2±0	97.2±0.35	97.2±0.35
Pat 21	Sensitivity	94.33±0.58	97.5±0	96.37±0.06	96.33±1.15	97.57±0.06	97.4±0.35	96.77±0.64	97.98±0.66
	Specificity	94.33±0.58	98.37±0.75	89.17±2.14	98.37±0.75	98.77±2.14	99.57±0.75	97.57±0.06	97.17±0.75
	Accuracy	92.72±3.15	93.55±0.69	85.98±0.85	92.81±0.66	93.95±0.88	94.45±0.6	91.63±0.52	94.24±0.68
Average	Sensitivity	93.4±0.55	96.76±0.43	92.86±0.16	93.09±0.9	96.28±0.72	96.12±0.59	91.77±0.86	94.18±1.48
	Specificity	90.63±1.12	91.25±0.07	77.37±0.94	90.58±1.55	91.11±0.37	91.25±0.07	90.43±1.35	93.55±1.59

or by ensemble size. Performance of an ensemble greatly depends on the diversity and accuracy among base learner of ensemble [29]. In our study we have used boosted, bagged, subspace discriminant, subspace KNN and RUSboosted Tree ensemble.

D. K-NEAREST NEIGHBOR (KNN)

KNN is a non-parametric method based on lazy learning, used in the field of machine learning for pattern recognition, classification and regression etc. using similarity measure commonly the distance functions. Here weights of neighbor contribute more than weights of distant contributors. In KNN model is not built immediately rather than all training samples are saved and are awaited until all the new observations are classified. These characteristics are totally opposite of eager learning, which build classifier model before the classification of new observations. These models are simple to implement and suits to the applications where data is dynamic and needs to be changed and updated frequently. The most crucial parameter is this algorithm is selection of parameter K. For larger dataset, the value of K is kept high whereas for smaller dataset this parameter value should be kept small. The most commonly used distance in KNN is Euclidean distance but we have also used cosine, cubic and weighted distance in this study.

Euclidean Similarity

$$\sqrt{\sum_{i=1}^{k}(x_i - y_i)^2} \tag{18}$$

Cosine Similarity

$$Cos\,\varphi == \frac{a.b}{||a||\,|b|} \tag{19}$$

where and b are the attribute vectors.

Weighted KNN

$$F(x) = \frac{\sum_i^k W_i f(x_i)}{\sum_i^k W_i} \tag{20}$$

$$\left(W_i = \frac{1}{d(x_d, x_i)^2}\right). \tag{21}$$

TABLE 3. Patient specific classification accuracy of ensemble and tree (mean ± standard deviation) of all the 21 cases.

Subject	Parameter	Ensemble					Tree		
		Boosted Tree	Bagged Tree	Subspace Discriminant	Subspace KNN	RUSBoosted Tree	Fine	Medium	Coarse
Sub 1	Accuracy	51.17±1.15	88.3±1.04	86.13±1.15	85.07±1.02	54±3.08	85.53±0.64	85.53±0.64	84.23±2.14
	Sensitivity	100±0	89.97±2.25	90.6±0.06	92.49±0.06	99.41±1.03	84.31±1.02	84.02±1.25	90.62±0.04
	Specificity	0±0	82.71±0.04	81.43±1.15	78.87±0.05	9.33±2.83	84±2.26	85.33±1.11	73.05±0.09
Sub 2	Accuracy	65.13±8.4	95.27±2.83	91.33±2.25	96.53±3.93	74.77±17.85	90.23±2.05	90.23±2.05	88.97±3.19
	Sensitivity	64.87±17.44	92.41±3.48	89.86±2.35	94.25±3.38	63.87±15.7	88.94±0.58	90.05±7.71	87.57±5.13
	Specificity	55.53±0.46	93.06±2.88	91.43±3.52	90.83±5.77	55.75±1.65	89.61±1.73	90.13±1.27	86.33±4.57
Sub 3	Accuracy	50.3±0	91.53±0.81	78.2±0.82	89.17±1.44	51.37±1.85	91.37±1.05	91.37±1.05	89±0.82
	Sensitivity	100±0	92.1±1.21	77.41±1.86	95.33±0.64	100±0	90.68±1.25	90.32±1.53	92.8±1.22
	Specificity	0±0	89.13±0	78.26±0	83.33±0.64	0±0	92.04±2.52	87.64±1.29	86.9±0
Sub 4	Accuracy	90.93±0.71	90.67±0.78	86±1.25	87.93±1.5	87.73±0.38	86.93±0.49	87.27±0.38	85.1±0.36
	Sensitivity	88.69±0.68	89.83±0.13	84.05±5.63	89.3±5.72	87.44±4.75	86.21±0.02	85.67±0.8	83.96±2.53
	Specificity	91±2.42	90.62±1.25	91.02±0.9	84.48±1.52	88.15±0.77	86.5±0.51	85.63±1.26	85.99±1.89
Sub 5	Accuracy	87.13±1.63	88.63±1.45	80.3±1.91	87.57±0.98	86.7±1.18	82.97±0.85	85.47±0.42	87.07±0.81
	Sensitivity	91.84±1.34	90.5±3.64	87.35±3.56	93.17±2.1	90.03±3.68	80.67±1.33	88.23±4.14	89.76±3.15
	Specificity	81.41±2.59	87.53±1.5	70.72±2.63	79.03±2.31	78.72±2.58	82.96±1.42	77.11±2.88	77.83±4.13
Sub 6	Accuracy	49.9±0	99.3±0	98.9±0	98.43±0.29	49.9±0	98.03±0.49	98.37±0.21	96.67±2
	Sensitivity	89.72±0.04	98.83±0.14	99.5±0	98.5±0	89.72±0.04	98.75±0	98.63±0.18	97.58±0.14
	Specificity	90±0	99.67±0.14	98.08±0.14	98.57±0.75	90±0	98.67±0.14	98.83±0.29	99.08±0.14
Sub 7	Accuracy	97.6±0.95	98.03±0.99	92.1±1.15	97.2±1.4	97.2±0.95	96.63±2.11	97.3±1.04	97.07±0.87
	Sensitivity	97.18±2.64	97.33±2.37	95.88±1.43	96.7±1.73	97.07±2.31	97.27±2.14	96.47±3.02	95.51±3.45
	Specificity	98.27±0.81	98.98±0.4	87.71±0.79	97.73±0.92	97.49±0.5	95.4±3.46	97.69±0.32	98.03±0.92
Sub8	Accuracy	74.33±0.12	84.33±0.12	91.93±0.12	82.77±0.15	76.5±0.52	75.1±1.22	74.27±0.21	82.3±0.17
	Sensitivity	78.32±4.45	88.53±1.5	83.48±2.88	77.31±5.32	81.97±2.89	74.4±2.42	73.7±2.97	83.99±6.39
	Specificity	67.3±12.64	92.03±4.56	89.23±0	78.09±4.65	68.82±0.71	77.05±0.22	72.94±3.78	89.23±14.32
Sub 9	Accuracy	96.47±0.15	96.5±0.1	95.7±0.1	95.47±0.15	95.43±0.29	95.03±0.46	95.43±0.29	95.9±0.1
	Sensitivity	96.33±0.06	95.87±0.15	96.05±0.13	94.97±0.12	95.07±0.32	94.5±0.17	94.8±0.14	95.87±0.14
	Specificity	94.77±1.33	96.31±0.6	94.99±0.18	95.65±0.21	96.11±0.62	94.37±0.06	95.89±0.42	96.19±0.1
Sub 10	Accuracy	90.2±1.08	89.33±0.23	79.8±0.52	87.6±0.17	88.5±0	58.97±46.49	87.13±1.14	90.03±0.74
	Sensitivity	92.3±1.39	91.1±0.61	82.83±1.44	88.98±1.14	91.43±0.32	88±8.31	90.3±0.71	90.78±1.73
	Specificity	85.6±1.56	85.9±8.4	76.9±6.93	83.39±1.57	78.3±2.94	85.18±2.57	85.13±0.75	88.32±2.53
Sub 11	Accuracy	95.07±0.06	95.47±0.06	93±0	94.87±0.06	94.5±0.17	93±0	94.33±0.4	93.27±0.46
	Sensitivity	98±0.17	97.47±0.12	97.08±0.05	98.5±0.17	97.55±0.09	93.6±0	97.7±0	97.36±0.12
	Specificity	92.57±0.32	96±2.42	88.9±0	90.9±0.7	91.31±0.05	93.2±0.69	90.9±0.86	90.29±0.2
Sub 12	Accuracy	54.8±8.31	97.7±0.17	97.4±0.52	97.67±0.06	54.8±8.31	97.33±0.92	96.87±0.12	97.43±0.6
	Sensitivity	33.23±57.56	97.87±0.27	97.87±0.27	98.48±0.53	99.92±0.13	98.17±0.81	97.37±0.47	97.51±0.38
	Specificity	90.3±16.8	97.72±0	96.21±1.05	97.04±0.39	23.64±40.94	96.28±0.66	97.04±0.43	92.61±3.83
Sub 13	Accuracy	96.67±0.23	96.47±0.06	93.93±0.12	95.93±0.12	95.7±0.44	95.1±0.2	95.57±0.23	94.1±0.7
	Sensitivity	96.31±0.18	94.83±3.49	92.45±2.81	96.21±0.27	95.03±1.76	94.86±0.14	95.83±0.32	94.25±2.39
	Specificity	95.33±0.67	96.73±0.2	96.5±0.17	95.59±0.25	94.3±1.91	94.4±0.87	95.62±0.19	95.94±0.36
Sub 14	Accuracy	48.4±0	100±0	99.47±0.23	99.87±0.23	48.2±0	100±0	100±0	100±0
	Sensitivity	59.15±0.73	100±0	99.47±0.46	100±0	58.72±0.02	100±0	100±0	100±0
	Specificity	38.33±1.44	100±0	100±0	100±0	38.06±0.05	100±0	100±0	100±0
	Accuracy	92.6±1.23	94.3±0.35	91.53±1.15	93.97±0.49	93.03±0.64	85.3±3.64	91.07±2.02	92.9±0.69

TABLE 3. (*Continued.*) Patient specific classification accuracy of ensemble and tree (mean ± standard deviation) of all the 21 cases.

Sub 15	Sensitivity	94.06±1.94	94.92±1.03	89.58±1.63	95.71±1.24	94.91±1.93	96.47±3.96	95.3±2.36	93.02±1.93
	Specificity	89.43±1.36	90.04±1.83	90.71±1.81	88.07±1.41	87.04±2.3	87.22±2.63	91.17±1.27	90.15±1.73
	Accuracy	49.6±0	99.13±0.12	99.2±0	99.27±0.23	49.6±0	98.8±0	98.8±0	98.63±0.29
Sub 16	Sensitivity	79.22±0.19	99.17±0	99.31±0.24	99.31±0.24	79.22±0.19	98.34±0	98.34±0	98.07±0.24
	Specificity	39.89±34.74	98.76±0	98.76±0	98.76±0	39.87±34.76	98.9±0.24	98.34±0.72	98.9±0.24
	Accuracy	97.33±0.12	97.53±0.12	87±0.17	97.47±0.12	96.83±0.29	96.53±0.29	96.97±0.4	96.9±0
Sub 17	Sensitivity	98.45±0.27	97.98±0.14	86.17±0	98.57±0.13	97.97±0.57	97.13±0.54	97.62±0.74	96.4±0.01
	Specificity	96.35±0.06	96.82±0.33	88.19±0.06	96.7±0.06	96.1±0.35	96.46±0.48	96.23±0.29	97.24±0.07
	Accuracy	50.87±1.15	94.9±1.39	92.27±1.15	96.3±0.52	57.57±0.58	93.1±2.46	94.07±1.15	94.9±0.87
Sub 18	Sensitivity	33.33±57.74	94.06±4.96	97.89±1.4	97±1.13	99.51±0.42	93.37±2.37	92.61±5.03	97.5±0.73
	Specificity	100±0	97±2.67	87.01±0.43	93.04±1.96	56.39±41.04	93.31±1.15	93.16±1.25	91.11±1.15
	Accuracy	50.53±0.58	90.5±0.87	80.1±0	91.83±0.29	52.03±3.18	86.1±0	86.1±0	86.93±0.58
Sub 19	Sensitivity	100±0	92±0	96.33±1.15	98±0	100±0	85.1±0.09	85.07±0.1	89.67±0.58
	Specificity	0±0	90±0	66±1.73	84±0	9.33±0.58	85.43±0.5	66±17.32	85.67±0.58
	Accuracy	95.6±0.35	96.23±0.58	88.7±0.69	95.63±0.58	95.37±0.12	94.03±0.75	94.63±0.75	93.93±0.75
Sub 20	Sensitivity	93.5±0.52	95.24±0.43	79.05±0.47	94.27±0.2	92.81±0.25	93.07±0.29	92.52±0.17	91.7±1.42
	Specificity	96.94±0.13	96.5±0.4	97.36±0.14	95.61±0.47	96.12±0.33	94.4±0	96.47±0.13	96.47±0.14
	Accuracy	48.2±0	94.8±0.35	96.8±0.35	95.8±1.04	50.2±1.73	90.4±2.08	90.4±2.08	90.8±1.73
Sub 21	Sensitivity	67.5±0	96.77±0.64	96.37±0.06	96.77±0.64	67.5±0	90.7±2.93	90.35±5.16	89.13±4.21
	Specificity	28.9±0	93.5±0.69	89.7±11.6	96.37±0.06	28.9±0	95.77±3.7	89.93±1.36	89.93±1.36
	Accuracy	72.99±1.25	94.23±0.59	90.47±0.65	93.64±0.7	73.81±1.98	90.02±3.15	91.96±0.69	92.2±0.85
Average	Sensitivity	83.43±7.02	94.61±1.26	91.36±1.33	94.94±1.18	89.48±1.73	91.64±1.35	92.14±1.75	93±1.71
	Specificity	68.19±3.68	93.76±1.35	88.53±1.58	90.76±1.13	63.03±6.42	91.48±1.23	90.06±1.77	90.92±1.83

TABLE 4. Performance evaluation results to distinguish healthy ictal subjects from interictal subjects using quadratic kernel SVM classifier and tuning KS with different values.

KS	Accuracy (%)	Sensitivity (%)	Specificity (%)
1	95.17	95.12	93.91
2	95.69	95.60	94.25
3	95.27	93.60	92.15
4	95.00	93.53	88.21
5	94.53	93.17	87.93
10	92.19	92.19	89.24
15	90.30	90.28	86.61
20	87.53	87.52	81.43
25	86.40	86.40	79.95
30	85.22	85.23	78.76
35	83.42	83.54	76.42
40	81.56	81.54	74.04
45	80.22	80.19	72.77
50	78.49	78.47	71.52
Average	88.64	88.31	83.37

E. ARTIFICIAL NEURAL NETWORK (ANN)

ANN gained huge popularity in the fields of Pattern classification problem due to its characteristics like self-learning, robustness, adaptability and parallelism used in solving complex, large-scale and nonlinear problems [30]. It constructs connectionist classifiers by creating a mathematical model inspired by biological neural structure. ANN is characterized by highly interconnected processing elements called neurons. Training of ANN depends on the gradient descent algorithms. In our case training was done by scale gradient descent which belongs to conjugate gradient method class. It shows superfast performance by avoiding time-consuming line search per learning iteration compared to other training methods [32]. The training samples were divided into 80% training, 10% validation and remaining 10% for testing samples. The network was built up with 6 six input nodes, 10 hidden nodes and 1 output node.

F. EVALUATION PARAMETERS

Previously the gold standard for performance evaluation of seizure- detection algorithms remains visual annotation of seizures by EEG experts. However, to evaluate the performance of proposed EEG signal classification models to classify an unseen data instance as either ictal or interictal, we used accuracy, sensitivity, specificity and Area under Curve (AUC) parameters.

$$Sensitivity = \frac{TP}{TP + FN} * 100 \tag{22}$$

$$Specificity = \frac{TN}{TN + FP} * 100. \quad (23)$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} * 100 \quad (24)$$

where True Positive (TP) = Correctly identified as Ictal False Positive (FP) = Incorrectly identified as Ictal True Negative (TN) = Correctly identified as Interictal False Negative (FN) = Incorrectly identified as Interictal.

V. RESULTS

The Freiburg database is one of the most comprehensive, long-term datasets, which addressed a variety of classification problems based on its recordings [38]. Since the onset and the offset of seizure episodes are known, the epochs between the seizure onset and offset were marked as ''ictal'' and the rest of the epochs as ''Interictal,'' forming the corresponding classes. To judge the performances on a broader parameter, patient-specific classification performance of 21 patients in terms of Accuracy, sensitivity, specificity, and AUC have been evaluated using the most robust machine learning classifiers. The classifiers were trained using all 21 subjects, and therefore, classification is generalized across the whole population contained in the Freiburg database rather than a single individual. For proper analysis of the problem, efficient use of feature extraction techniques and use of optimal classifier is most importantly needed. In this study, we used PFEN for feature extraction from both the ictal and interictal seizure states of all the patients, keeping the dynamic and non-linear variations of the brain. All the classifiers were tuned with different parameters to view their performance using various metrics like weight, distance, learning rate, etc. that will help to investigation their performance in research. In the past, these classifiers were used with the few options and default parameters, and their actual ability was under- mined. The higher performance in classification depends on the selection of different kernel and other metrics. To minimize the impact of over-fitting, standard 10-fold cross- validation approach was used for evaluation of model to find the optimal classification results. To get an in-depth investigation of classifiers performance, we deployed SVM with linear, quadratic, cubic, fine gaussian, medium gaussian, and coarse gaussian. The Decision tree was implemented with a fine tree, medium tree, and coarse tree. KNN was used using fine, medium, coarse, cosine, and cubic KNN. The Ensemble was implemented using the boosted tree, bagged tree, space discriminant KNN, subspace discriminant KNN and RUS- Boosted tree. There was linear and quadratic discriminant. Artificial neural network was implemented with two-layered feedforward networks that was trained using scaled conjugate gradient backpropagation containing ten hidden layers. The performance evaluation for different classifiers is shown in Tables I–III. SVM and ANN (Table I), KNN and Discriminant (Table II) while Ensemble and Decision Tree (Table III) using their default values. The average accuracy of these classifiers was 98.26 ± 0.84% (ANN), 91.95 ± 0.74% (Lin- ear SVM), 95.33 ± 0.64% (Quadratic SVM), 94.26 ± 0.69% (Cubic SVM), 93.91 ± 2.33% (Fine Gaussian), 94.29 ± 1.39% (Medium Gaussian SVM), 92.72 ± 0.47% (Coarse Gaussian SVM), 92.72 ± 3.15% (Fine KNN), 93.55 ± 0.69% (Medium KNN), 85.98 ± 0.85% (Coarse KNN), 92.81 ± 0.66% (Cosine KNN), 93.95 ± 0.88% (Cubic KNN) 94.53 ± 0.6% (Weighted KNN) 91.64 ± 0.52% (Linear Discriminant) 94.24 ± 0.68% (Quadratic Discriminant), 72.99 ± 1.25% (Boosted Tree), 94.24 ± 0.59% (Bagged Tree) 90.47 ± 0.65% (Subspace Discriminant) 93.63 ± 0.7% (Subspace KNN) 73.81 ± 1.98% (RUSBoosted Tree) 91.2 ± 3.15% (Fine Tree) 91.96 ± 0.69% (Medium Tree) and 92.2 ± 0.85% (Coarse Tree). The best classification accuracy of 98.26 ± 0.84% was achieved by using ANN, while the lowest accuracy of 72.99 ± 1.25% was found using Boosted Tree Ensemble.

Using SVM, the highest accuracy was achieved for quadratic Gaussian, followed by medium SVM, and the lowest classification accuracy was achieved for linear SVM. Using KNN, the most promising accuracy was achieved by Weighted KNN followed by Cubic KNN, and most low results were obtained by Coarse KNN. Using Discriminant, Quadratic discriminant made the highest accuracy while the Linear discriminant was the second one. The highest discrimination using ensemble was found for Bagged tree ensemble and the lowest discrimination was found with the boosted tree. For Decision Tree, the highest accuracy was achieved by Coarse tree followed by the medium and fine tree. The average highest values of these classifiers are shown in Table VIII and Fig 8. Most of the results of classification are very high, which again justifies the higher discriminative ability of PFEN based feature extraction. We have further evaluated the performance of SVM Quadratic kernel in more depth by varying the kernel scales (KS) and fixing box constraint level (BCL) at 1. The performance was evaluated by changing the values of Kernel scale (KS), which can be seen in Table IV. The overall highest accuracy of 95.69% was observed with KS = 2. The accuracy was trending down with the increase in the value of KS, and we got 78.49% accuracy at KS 50, which can be shown in Fig 3. Similar behavior was observed in case of sensitivity and specificity. It demonstrates that the value of KS at different levels is significant in dis- criminating the subjects for determination of epileptic seizure stages. Thus, these parameters help us in getting more in- sight towards getting improved performances. In Table V, we evaluated the performance of KNN by using altered weight metrics and fixed neighbor of 10. We have used Cosine, City Block, Chebyshev, Minkowski, Euclidean, Co-relation, Spearman, Hamming, and Jackard distance weights. The highest performance accuracy of 93.53% was observed for KNN with Minkowski followed by Euclidean with 93.50% accuracy, followed by City Block with 93.49% accuracy. The lowest accuracy was observed with Jackard and hamming distance metrics. The detailed comparison of the classification performance shown in Fig 4.

TABLE 5. Performance evaluation results to distinguish ictal subjects from interictal subjects using K-nearest neighbor classifier with varying distance metrics and selection criterion of the number of neighbors.

	Parameter	Cosine	City Block	Chebyshev	Minkowski	Euclidean	Co-relation	Spearman	Hamming	Jackard
	Accuracy	88.7	89.6	90.6	93.4	92.5	88.7	87.7	49.5	49.5
Sub 1	Sensitivity	90.6	98.2	98.2	98.2	98.11	94.4	92.5	0	0
	Specificity	86.6	80.8	82.7	88.6	86.69	82.7	82.76	100	100
	Accuracy	94	96.9	93.5	94.5	96.9	91.1	88.2	50	50
Sub 2	Sensitivity	90.61	99.2	99.2	99	99.2	84.84	84.8	0	0
	Specificity	97.3	94.5	87.7	90.6	94.5	97.34	91.6	100	100
	Accuracy	85.9	87.6	84.9	88.1	86.5	80.5	77.3	49.5	49.5
Sub 3	Sensitivity	82.7	95.6	92.4	93.5	94.6	75.2	75.2	0	0
	Specificity	89.13	79.3	77.17	82.6	78.26	85.86	79.3	100	100
	Accuracy	92.1	92	91.4	91.9	91.9	92.4	90.8	50	50
Sub 4	Sensitivity	86.09	89.9	89.9	89.7	89.5	86.96	84.58	0	0
	Specificity	98.19	94.08	92.7	94.08	94.3	97.7	97.11	100	100
	Accuracy	82.4	86.2	85.5	86.2	85.7	81.9	77.7	50	50
Sub 5	Sensitivity	86.8	98.8	98.4	99.3	97.98	87.26	75.2	0	0
	Specificity	81.9	73.4	72.5	72.9	73.42	76.1	80.12	100	100
	Accuracy	99	98.6	98.5	98.8	98.9	99.3	99.1	50	50
Sub 6	Sensitivity	99.25	99.75	99.5	99.75	99.75	99.3	99.3	0	0
	Specificity	98.75	97.5	97.5	97.75	98	99.3	99.2	100	100
	Accuracy	96.8	98.5	98.5	98.6	98.6	89.1	88.1	50	50
Sub 7	Sensitivity	94.4	99.13	99.13	99.2	99.2	92.5	89.9	0	0
	Specificity	99.13	97.93	97.8	98.04	98.04	85.7	86.2	100	100
	Accuracy	82.3	82.3	81.6	83.9	82.4	82.4	85.4	50	50
Sub 8	Sensitivity	93.9	93.9	83.3	83.8	93.9	93.9	87.8	0	0
	Specificity	67.6	67.6	79.9	83	67.6	67.6	83	100	100
	Accuracy	95.9	96.2	95.6	95.9	96.1	65.5	64.4	50	50
Sub 9	Sensitivity	95.9	97.2	96.13	96.8	97.2	71.8	58.5	0	0
	Specificity	95.9	95.2	95	94.9	95	59.15	70.2	100	100
Sub 10	Accuracy	89.1	88.6	88	88.6	88.6	83.9	79	50	50
	Sensitivity	94.9	99.8	99.3	99.8	99.8	95.2	82.8	0	0
	Specificity	83.2	77.45	76.7	77.45	77.45	72.5	78.2	100	100
	Accuracy	93.5	94.8	94.6	94.6	94.5	70.1	65.3	50	50
Sub 11	Sensitivity	96.3	98.8	98.9	98.9	98.8	74.3	56.7	0	0
	Specificity	90.6	90.6	90.14	90.14	90.2	65.8	74	100	100
	Accuracy	97.5	97.7	98.2	98.2	97.7	77.5	77.7	50	50
Sub 12	Sensitivity	98.18	98.6	99.09	99.09	98.6	74.09	74.5	0	0
	Specificity	96.8	96.8	97.27	97.27	96.8	80.9	80.9	100	100
	Accuracy	95.9	95.7	94.9	95.7	95.9	96.1	94.4	50	50
Sub 13	Sensitivity	95.73	97.15	96.99	97.15	95.73	95.89	93.04	0	0
	Specificity	96.05	94.15	92.89	94.15	96.03	96.2	95.7	100	100
	Accuracy	100	100	100	100	100	81.3	82.9	50	50
Sub 14	Sensitivity	100	100	100	100	100	79.3	76.2	0	0
	Specificity	100	100	100	100	100	83.3	89.6	100	100
	Accuracy	92.4	93.3	92.9	93	93.2	89.5	87.5	50	50

TABLE 5. (*Continued.*) Performance evaluation results to distinguish ictal subjects from interictal subjects using K-nearest neighbor classifier with varying distance metrics and selection criterion of the number of neighbors.

Sub 15	Sensitivity	94.08	98.2	98.38	98.2	96.2	92.7	89.3	0	0
	Specificity	90.8	88.4	87.3	87.8	86.23	86.34	85.8	100	100
	Accuracy	99.2	99.4	99	99.2	99.2	94.4	93.2	50	50
Sub 16	Sensitivity	98.34	99.58	99.58	99.2	98.34	96.7	95.86	0	0
	Specificity	100	99.17	98.34	99.2	100	92.14	90.5	100	100
	Accuracy	92.8	97.9	97.3	97.6	97.7	74.4	71.6	50	50
Sub 17	Sensitivity	90.47	99.3	99.3	99.4	99.3	79.09	65.9	0	0
	Specificity	95.12	96.4	95.35	95.8	96.05	69.8	77.2	100	100
	Accuracy	96	94.5	94.1	94.5	93.8	83.2	81.3	50	50
Sub 18	Sensitivity	97.08	95.6	96.35	96.35	94.89	84.67	82.48	0	0
	Specificity	94.8	93.38	91.91	92.64	92.64	69.8	80.14	100	100
	Accuracy	81.1	79.6	77.6	79.1	80.1	72.1	69.2	49.8	49.8
Sub 19	Sensitivity	92.07	99.1	99	99	99	79.2	75.2	0	0
	Specificity	70	40	56	59	61	65	63	100	100
	Accuracy	94.9	95.6	95	95.5	95.6	90.4	88.9	49.9	49.9
Sub 20	Sensitivity	95.2	93.7	92.3	93.35	93.7	91.3	93.2	100	100
	Specificity	94.5	97.4	97.8	97.66	97.4	89.5	84.5	0	0
	Accuracy	97	98.2	96.4	97.6	97.6	84.3	84.3	50	50
Sub 21	Sensitivity	96.38	97.59	95.18	96.38	96.38	85.5	85.5	0	0
	Specificity	97.59	98.8	97.16	98.8	98.8	83.13	83.13	100	100
Average	Accuracy	92.69	93.49	92.77	93.57	93.50	84.20	82.57	49.4	49.4
	Sensitivity	93.76	97.58	96.69	96.96	97.15	86.39	81.83	4.76	4.76
	Specificity	91.62	88.23	88.75	90.11	89.45	81.23	83.44	95.24	95.24

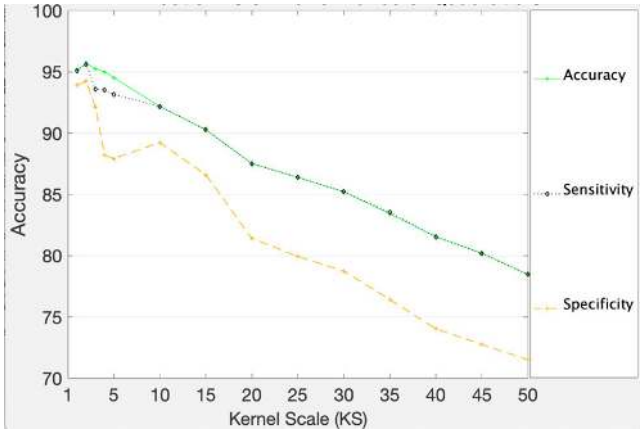


FIGURE 3. Effect of varying Kernel Scale (KS) on the performance of SVM.

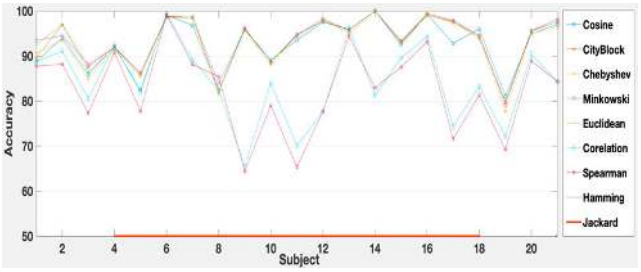


FIGURE 4. Effect of varying Distance Metrics on performance of KNN.

A. SCATTER PLOT

To visually show the classification performance of the different classifier, we used the scatter plot of PFEN against the ictal and interictal stage of subjects in Fig 5. We selected patient 14 with one variant of each classifier for the sake of illustration. We chose SVM with the quadratic kernel, KNN with Cosine, Discriminant with linear, subspace discriminant Ensemble, Fine DT and ANN. Here the brown dots denote interictal whereas blue dots represent the ictal class. The cross dots indicate the errors in classification.

B. RECEIVER OPERATING CURVE (ROC)

Receiver Operating Curve (ROC) is a plot against true positive rate (sensitivity) and false positive rate (specificity) at various cutoff values to visualize the behavior of the classification system. Ictal values are classified as 1, whereas interictal values are represented as 0. Sensitivity is plotted on y-axis

TABLE 6. Area under curve (AUC) for patient specific values of SVM, tree, discriminant and ANN for all the 21 cases.

Subject	SVM						Tree			Discriminant		ANN
	Linear SVM	Quadratic SVM	Cubic SVM	Fine Gaussian	Medium Gaussian	Coarse Gaussian	Fine Tree	Medium Tree	Coarse Tree	Linear	Quadratic	
Sub1	0.95	0.96	0.94	0.92	0.96	0.96	0.89	0.89	0.83	0.94	0.95	0.99
Sub2	0.99	1.0	1.0	0.97	1.0	0.99	0.90	0.90	0.93	0.98	1.0	1.0
Sub3	0.90	0.96	0.96	0.94	0.97	0.93	0.93	0.93	0.90	0.91	0.92	0.99
Sub4	0.97	0.99	0.97	0.98	0.99	0.98	0.91	0.94	0.91	0.97	0.99	1.0
Sub5	0.91	0.98	0.97	0.97	1.0	0.95	0.89	0.89	0.89	0.91	0.98	0.98
Sub6	1.0	1.0	1.0	1.0	1.0	1.0	0.99	0.99	0.99	1.0	1.0	1.0
Sub7	0.93	0.98	0.97	0.97	0.97	0.97	0.95	0.95	0.94	0.93	0.96	1.0
Sub8	0.98	0.94	0.91	0.85	0.95	0.95	0.77	0.77	0.81	0.98	0.99	1.0
Sub9	0.99	0.99	0.99	0.98	0.98	0.99	0.96	0.97	0.97	0.99	0.99	0.98
Sub10	0.88	0.98	0.96	0.95	0.97	0.90	0.90	0.93	0.92	0.88	0.97	0.98
Sub11	0.94	0.98	0.98	0.98	0.98	0.97	0.95	0.96	0.94	0.95	0.97	1.0
Sub12	0.99	1.0	1.0	0.99	1.0	1.0	0.98	0.97	0.97	0.99	1.0	0.99
Sub13	0.99	0.99	0.99	0.99	1.0	0.98	0.96	0.95	0.95	0.99	0.99	0.99
Sub14	1.0	1.0	1.0	1.0	1.0	1.0	0.1.0	1.0	1.0	1.0	1.0	1.0
Sub15	0.98	1.0	0.98	0.98	1.0	0.99	0.95	0.95	0.94	0.98	0.99	1.0
Sub16	1.0	1.0	1.0	1.0	1.0	1.0	0.99	0.99	0.99	1.0	1.0	1.0
Sub17	0.89	0.99	0.97	0.99	0.99	0.99	0.97	0.97	0.96	0.89	0.99	1.0
Sub18	0.96	0.99	0.98	0.97	0.98	0.93	0.94	0.94	0.93	0.97	0.99	1.0
Sub19	0.80	0.96	0.91	0.98	0.97	0.80	0.87	0.87	0.86	0.79	0.97	0.99
Sub20	0.95	0.99	0.97	0.99	0.99	0.97	0.96	0.95	0.95	0.97	0.95	1.0
Sub21	1.0	1.0	0.99	0.98	1.0	1.0	0.94	0.94	0.93	1.0	1.0	1.0
Average	0.953	0.985	0.973	0.97	0.986	0.964	0.933	0.936	0.929	0.953	0.981	0.994

whereas specificity is plotted on the x-axis. The area under the curve (AUC) is used to the get better discrimination ability of the system, and its value ranges between 0 and 1. Higher the value of AUC, higher will be the system discrimination ability and vice versa [1]. The AUC values obtained were found higher which matches to the higher accuracy values of our classifiers as visualized in Fig 6. The AUC values of different classifier are shown in Tables VI & VII. Table VI contains AUC values of SVM, Tree, Discriminant and ANN classifiers while, Table VII contains AUC values of KNN and Ensemble classifiers.

Furthermore, the capability of PFEN entropy in discriminating seizure states was conducted using student t-test on all 21 independent patient values, which were normally distributed. Lower the value of p; higher will be the power of discrimination between different states, i.e., ictal and interictal [1]. This means lesser p-value has a greater discriminating capability. The average p-value was found less than 0.05 which infers that PFEN has better discriminative power of seizure detection.

VI. DISCUSSION

Objective of this study was to find the potential of PFEN towards discriminating ictal and interictal stage based on individual subjects using robust machine learning algorithms. We utilize only the ictal EEG segments between the seizure onset and seizure end, containing only the seizure event and missed the non-ictal recording providing in the database. Thus, it is more suitable to precisely capture the seizure activity. According to [38], human brain exhibits a variation in the chaotic electrical physiological act from interictal to ictal or seizure state. As the nerve cells inside the epileptogenic zone turn isolated, they grow vacant which might incline to an epileptic seizure. Through ictal state, a huge amount of nerve cells on the cerebral cortex sharply jump discharging in a tremendously ordered recurring pattern. We observed that there is a sharp fall in PFEN values after the seizure start, which can be due to a sudden increase in magnitude either slightly before or slightly after the seizure onset, which reveals that brain gets affected by the occurrence of such seizures so that it becomes momentarily

TABLE 7. Area under curve (AUC) for patient specific values of KNN and discriminant for all the 21 cases.

Subject	KNN						Ensemble				
	Fine KNN	Medium KNN	Coarse KNN	Cosine KNN	Cubic KNN	Weighted KNN	Boosted Tree	Bagged Tree	Subspace Discrim-inant	Subspace KNN	RUSBoosted Tree
Sub1	0.90	0.95	0.49	0.93	0.96	0.96	0	0.93	0.94	0.91	0.92
Sub2	0.90	1.0	0.93	0.98	0.99	1.0	0.93	1.0	0.97	1.0	0.85
Sub3	0.89	0.95	0.84	0.92	0.96	0.95	0	0.96	0.85	0.95	0
Sub4	0.92	0.98	0.96	0.98	0.98	0.99	0.98	0.99	0.95	0.97	0.94
Sub5	0.88	0.96	0.96	0.96	0.96	0.97	0.95	0.96	0.90	0.94	0.89
Sub6	0.99	1.0	1.0	1.0	1.0	1.0	0	0.1.0	1.0	1.0	0
Sub7	0.97	0.97	0.96	0.96	0.97	0.97	0.97	0.92	0.97	0.97	0.97
Sub8	0.80	0.90	0.79	0.91	0.90	0.90	0.88	0.93	0.97	0.87	0.82
Sub9	0.94	0.98	0.99	0.98	0.98	0.98	0.98	0.98	0.99	0.98	0.96
Sub10	0.85	0.96	0.94	0.94	0.96	0.97	0.97	0.98	0.89	0.94	0.94
Sub11	0.93	0.97	0.98	0.97	0.97	0.98	0.98	0.98	0.93	0.97	0.96
Sub12	1.0	1.0	0.99	0.99	1.0	0.99	0.98	0.99	0.99	1.0	0.98
Sub13	0.96	1.0	0.99	0.98	0.99	0.99	0.99	0.99	0.91	0.99	0.96
Sub14	1.0	1.0	0.98	1.0	1.0	1.0	0	1.0	0.99	1.0	0
Sub15	0.93	0.99	0.99	0.98	0.99	0.99	0.99	0.99	0.98	0.98	0.94
Sub16	1.0	1.0	1.0	1.0	1.0	1.0	0	1.0	1.0	1.0	0
Sub17	0.97	0.99	0.99	0.97	0.99	0.99	0.98	0.99	0.90	0.99	0.96
Sub18	0.93	0.98	0.91	0.98	0.98	0.98	0	0.97	0.93	0.97	0.96
Sub19	0.90	0.95	0.83	0.88	0.94	0.96	0	0.97	0.79	0.97	0.95
Sub20	0.94	0.99	0.99	0.98	0.99	0.99	0.99	0.99	0.96	0.99	0.96
Sub21	0.95	1.0	0.99	0.98	1.0	1.0	0	0.99	1.0	0.99	0
Average	0.929	0.977	0.928	0.966	0.976	0.979	0.599	0.975	0.945	0.969	0.714

TABLE 8. Comparison of highest performance achieved by different classifier's.

Classifier	Accuracy (%)	Sensitivity (%)	Specificity (%)
ANN	98.26	98.02	98.26
SVM	95.33	95.94	93.92
KNN	94.53	96.12	91.25
Ensemble	94.23	94.61	93.76
Discriminant	94.24	94.18	93.55
DT	92.2	93	90.92

irregular before becoming regular during the seizure. This type of recurring pattern regularly begins and finishes impulsively without any external consequences, which fallouts in lessening values of PFEN. We found that interictal subjects have higher complexity than that of ictal subjects. The author in [19], [31], [32] while analyzing human epileptic EEG using permutation entropy (PE), experience's the lower value of PE during seizure stage and higher PE values in the normal stage. Our findings of PFEN are also in agreement to their studies where we find lower PFEN values in ictal state and higher PFEN values in interictal state, shown in Fig 7. This infers that the brain activity during the seizure period has a more regular pattern than the normal interictal state having a more repetitive pattern with a similar pattern. This feature of PFEN helps us in discriminating EEG signal states with high accuracy. A substantial difference in the values of

TABLE 9. Comparison of our techniques with other techniques using Freiburg Data.

Author	Year	Dataset	Feature	Classifier	Patients	Acc (%)	Sen (%)	Spe (%)
Giannakakis	2013	University General Hospital of Herklion	ApEn	CSB	8	90.12	97.33	83.91
Xiang et al	2015	CHB-MIT	FuzzyEn	SVM	18	98.31	98.27	98.36
Ling et al	2010	Bonn	ApEn	ANN	5	98.27	95.5	99
Song and lio	2010	Bonn	SampEn	ELM	5	95	98	91
Nicolaou et al	2012	Bonn	PE	SVM	5	94	94	93
Kumar et al	2014	Bonn	FuzzyEn	ANN	5	97.38	98.1	94.4
Subasi et al	2018	Bonn	DWT	PSO-SVM	5	99.38	99.5	99.25
Xie et al	2012	Freiburg	Wavelet based functional linear mode	KNN	4	100	100	100
Patel et al	2009	Freiburg	Power	LDA	21	87.7	94.2	77.9
Park et al	2011	Freiburg	Spectral Power	SVM	18	-	97.5	
Williamson et al.	2011	Freiburg	Spatio-Temporal Correlation	SVM	21	-	90.80	
Yuan et al.	2012	Freiburg	Fractal Intercept	ELM	21	94.9	93.85	94.89
Geng et al.	2016	Freiburg	WT	WNN	20	98.9	96.72	98.9
Yuan et al.	2017	Freiburg	WPT	WELM	19	95.74	92.22	-
Patnaik et al.	2008	Freiburg	WT	BPNN	21	91.29	99.19	-
Alickovic et al.	2018	Freiburg	WPD	KNN	21	99.5	99.91	99.16
Our Approach	2019	Freiburg	PFuzzyEn	NN	21	98.26±0.84	98.02±0.81	98.26±0.56
				SVM	21	95.33±.64	95.94±1.21	93.92±1.16
				KNN	21	94.45±0.6	96.12±0.59	91.25±0.37
				Ensemble	21	94.23±0.59	94.61±1.26	93.76±1.35
				LD	21	94.24±0.68	94.18±1.48	93.55±1.59
				DT	21	92.2±0.85	93±1.71	90.92±1.83

ictal and interictal seizure can be seen after using PFEN, which signifies the legitimacy of PFuzzyEn based feature extraction. The study result itemized above shows that feature extraction grounded over PFuzzyEn method is not only capable of lessening feature dimension of the EEG data, but likewise capable of imitating the characteristic discrepancy of original epileptic EEG signals efficiently, which is major reasons of achieving high accuracy. High-frequency artifacts can increase the chances of misclassification [33]. Thus, the Notch filter was used to detain the artifacts. Also, the smoothing used in post-processing can process the imprecise boundaries between preictal and interictal EEGs, which

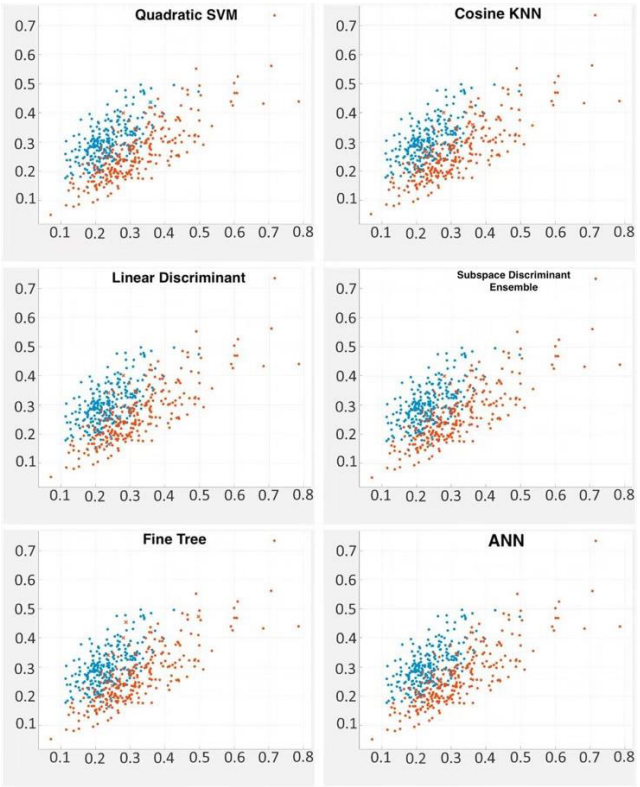


FIGURE 5. Patient 16 Scatter plot using different classifier for PFEN for different classifiers. The brown color represents Ictal while blue color denotes Interictal points. The dot point represents correctly classified point whereas cross point represents incorrectly classified data point. (a) Quadratic SVM (b) Cosine KNN (c) Linear Discriminant (d) Subspace Discriminant Ensemble (e) Fine Decision Tree (f) Artificial Neural Network.

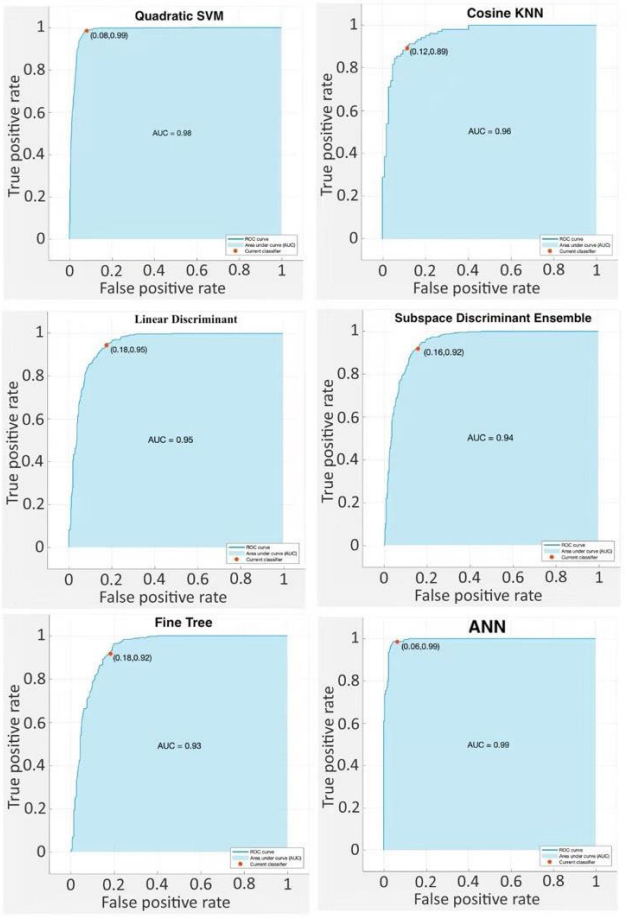


FIGURE 6. Area Under Curve to discriminate Ictal subjects and interictal subjects using SVM, KNN, LD, FT, SSE and ANN.

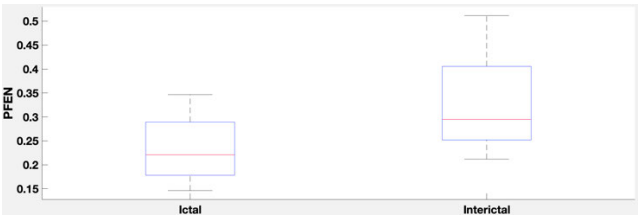


FIGURE 7. Boxplot of extracted ictal and interictal value of PFuzzyEn using EEG signal with values of m = 2 and r = 0.25 ∗ std. Lower PFuzzyEn value were observed during the ictal state and higher PFuzzyEn values were found in normal or interictal stage.

makes the model capable of classification more accurate. Before analyses, we tried to minimize all the possible effects of the skull by applying band-pass filter on all the raw EEGs (cut-off frequencies: 0.53–48 Hz). All the classifiers show excellent results with the proposed entropy; however, it was more desirable to find the capability of the state of art robust machine learning classifiers in discriminating ictal and interictal states. We compared the classification performance used in our framework and found that the ANN achieved a comparable classification accuracy and a much faster computation speed than other states of art models. The ANN gives the classification accuracy of 98.26 ± 0.84% for discrimination and diagnostic purpose than the traditional approaches. It also requires the minimum time to train and test unseen EEG pattern in comparison to SVM, KNN and other classifiers which suits to the rapid real-time detection of epileptic EEG. The higher accuracy using SVM kernel was 95.33 ± 0.64% with quadratic SVM, 94.65 ± 0.6% with KNN the best result was using Weighted KNN having Euclidean distance, Fine tree with 100 splits using Gini's diversity index produce 92.2 ± 0.85%, bagged tree ensemble with decision tree learner has 92.23 ± 0.59% and quadratic discriminant with 94.24 ± 0.68% accuracy. The comparison of the proposed method with the previously reported methods [33]–[39], which also uses the same database is shown in Table IX. In contrast with other works adopted on said database, our technique produced better overall results. Different methods have been proposed in the literature for seizure detection on the same dataset used in this study. The maximum accuracy reported on the said dataset is reported by Xie and Krishnan [39] produced 100% accuracy, but they have used just four patients, so it's quite difficult to extract some conclusions from their proposed model. But in our case, our methodology was tested on data from 21 patients, which achieves a very higher value of accuracy, sensitivity, specificity, and AUC. For instance, Patel *et al.* [34] report 94% for sensitivity, 77.9% for specificity, and 87.7% for

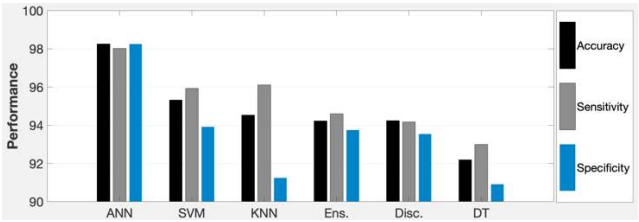


FIGURE 8. Comparative analysis of different classifiers in discriminating ictal and interictal states using PFEN.

overall accuracy. Yuan *et al.* [40] report 91.72% for sensitivity, 94.89% for specificity, and 94.9% accuracy. Park *et al.* [35] used SVM and reported a mean seizure prediction sensitivity of 90.8% using 19 patients. Patel *et al.* [34] performed binary classification using SVM on 18 patients and find the sensitivity of 93.8% to 97.5%. Geng *et al.* [33] performed epileptic seizure detection using the improved wavelet neural network over 20 patients and achieved a recognition accuracy of 98.9%, 96.72% sensitivity and 98.91% specificity. Patnaik *et al.* [38] uses feed- forward propagation ANN for classification of epileptic seizure and obtained sensitivity of 91.29% and specificity of 99.19%. Alickovic *et al.* [41] decomposed EEG signal using wavelet packet decomposition along with different classifiers to achieve the classification accuracy of 99.5%. Subasi *et al.* [42] decomposed EEG signals into time-frequency sub-bands using DWT with optimized parameters of SVM and GA to achieve classification accuracy of 99.38% over Bonn dataset. From all the analysis results and discussion above, it can be said that the performance along with the robustness of any framework could be enhanced by the ability of feature extraction technique, i.e., entropy and by choosing optimized parameter value of classifiers. From the simulations, we conclude that due to obtained highly promising results, our proposed method, combined with simple classifier's such as artificial neural network can be handy for the detection of epilepsy and seizure detection problems.

VII. CONCLUSION

Epileptic seizures encompass huge portions of the cerebral cortex, thus effective means of switching state from interictal to ictal is exceptionally complex. Even in the same subject, the involved cortical regions and the time consumed in the development remain changed through every seizure onset. Therefore, a swift and proficient detection technique capable of discriminating interictal and ictal state is highly desired. In this work, PFEN has been explored to see its potentials toward discriminating the transitional change from interictal to the ictal state of the human brain in a complex time series extracted from multichannel EEGs. The classification model uses PFEN and robust machine learning algorithms with advance parametric tactics to get superior accuracy and more profound knowledge to detect seizures for each specific patient. The results of all subjects were compared for all classifiers to check their performance. The artificial neural network gives the most efficient and optimized results for discrimination and diagnostic purpose than the traditional approaches. The support vector machine was found second best by optimizing their kernel and kernel scales. The classification results varied significantly across subjects, signifying abnormal activities in the brain and the potential advantage of patient-specific seizure classification methods. The study proposed a novel approach for automatic epileptic seizure detection with a very high identification accuracy in classification with a low computational budget that can be further used to construct real-time epileptic seizure detection system. These findings indicate that PFEN has much better anti-noise performance than the Permutation entropy and fuzzy entropy. Also, it has much better seizure detection ability than the current state of art entropy variants. PFuzzyEn is not sensitive to noise, so it can be extended to analyze EEG signals of other diseases and electromyography (EMG) and electrocardiogram (ECG).

VIII. LIMITATION

Freiburg database is mixed with scalp and intracranial recordings, so it might not be perfect for testing diverse classification algorithms. Possibly amplitudes of intracranial recordings, are higher partly due to the different locations of electrodes and the filtering mechanism of the skull. Also, this method uses stored offline data to assesses algorithms; therefore, it's not a suitable marker of the system's ability, because the data for training and testing algorithms are pre-processed and filtered with significant features extracted. It's a substantial concern, in our future work we will be looking towards implementing the real-time signals, using advances in the big data and deep learning techniques.

ACKNOWLEDGMENT

(Waqar Hussain and Bin Wang are co-first authors.)

REFERENCES

[1] M. N. Tibdewal, H. R. Dey, M. Manjunatha, A. Ray, and M. Malokar, "Multiple entropies performance measure for detection and localization of multi-channel epileptic EEG," *Biomed. Signal Process. Control*, vol. 38, pp. 158–167, Sep. 2017.
[2] L. Hussain, "Detecting epileptic seizure with different feature extracting strategies using robust machine learning classification techniques by applying advance parameter optimization approach," *Cogn. Neurodyn.*, vol. 12, no. 3, pp. 271–294, Jun. 2018.
[3] M. P. Richardson and J. G. R. Jefferys, "Introduction—Epilepsy research UK workshop 2010 on 'preictal phenomena,'" *Epilepsy Res.*, vol. 97, no. 3, p. 229, Dec. 2011.
[4] L. Wang, W. Xue, Y. Li, M. Luo, J. Huang, W. Cui, and C. Huang, "Automatic epileptic seizure detection in EEG signals using multi-domain feature extraction and nonlinear analysis," *Entropy*, vol. 19, no. 6, p. 222, Jun. 2017.
[5] S. Lahmiri and A. Shmuel, "Accurate classification of seizure and seizure-free intervals of intracranial EEG signals from epileptic patients," *IEEE Trans. Instrum. Meas.*, vol. 68, no. 3, pp. 791–796, Mar. 2019.
[6] M. Sharma, R. B. Pachori, and U. R. Acharya, "A new approach to characterize epileptic seizures using analytic time-frequency flexible wavelet transform and fractal dimension," *Pattern Recognit. Lett.*, vol. 94, pp. 172–179, Jul. 2017.
[7] Y. Song, J. Crowcroft, and J. Zhang, "Automatic epileptic seizure detection in EEGs based on optimized sample entropy and extreme learning machine," *J. Neurosci. Methods*, vol. 210, no. 2, pp. 132–146, Sep. 2012.

[8] Y. Niu, R. Cao, H. Wang, C. Li, M. Zhou, Y. Guo, B. Wang, P. Yan, and J. Xiang, "Permutation fuzzy entropy—An index for the analysis of epileptic electroencephalogram," *J. Med. Imag. Health Inform.*, vol. 9, no. 3, pp. 637–645, Mar. 2019.

[9] L. Guo, D. Rivero, and A. Pazos, "Epileptic seizure detection using multiwavelet transform based approximate entropy and artificial neural networks," *J. Neurosci. Methods*, vol. 193, no. 1, pp. 156–163, Oct. 2010.

[10] H. Ocak, "Optimal classification of epileptic seizures in EEG using wavelet analysis and genetic algorithm," *Signal Process.*, vol. 88, no. 7, pp. 1858–1867, Jul. 2008.

[11] H. Ocak, "Automatic detection of epileptic seizures in EEG using discrete wavelet transform and approximate entropy," *Expert Syst. Appl.*, vol. 36, no. 2, pp. 2027–2036, Mar. 2009.

[12] G. Giannakakis, V. Sakkalis, M. Pediaditis, C. Farmaki, P. Vorgia, and M. Tsiknakis, "An approach to absence epileptic seizures detection using approximate entropy," in *Proc. 35th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2013, pp. 413–416.

[13] L. Hussain, W. Aziz, S. Saeed, S. A. Shah, M. S. A. Nadeem, I. A. Awan, A. Abbas, A. Majid, and S. Zaki H. Kazmi, "Quantifying the dynamics of electroencephalographic (EEG) signals to distinguish alcoholic and non-alcoholic subjects using an MSE based Kd tree algorithm," *Biomed. Eng./Biomedizinische Technik*, vol. 63, no. 4, pp. 481–490, Jul. 2018.

[14] Y. Song and P. Liò, "A new approach for epileptic seizure detection: Sample entropy based feature extraction and extreme learning machine," *J. Biomed. Sci. Eng.*, vol. 3, no. 6, p. 556, Jun. 2010.

[15] C.-P. Shen, S.-T. Liu, W.-Z. Zhou, F.-S. Lin, A. Y.-Y. Lam, H.-Y. Sung, W. Chen, J.-W. Lin, M.-J. Chiu, M.-K. Pan, J.-H. Kao, J.-M. Wu, F. Lai, "A physiology-based seizure detection system for multichannel EEG," *PLoS ONE*, vol.8, no. 6, Art. no. 065862, Jun. 2013.

[16] Y. Song and J. Zhang, "Discriminating preictal and interictal brain states in intracranial EEG by sample entropy and extreme learning machine," *J. Neurosci. Methods*, vol. 257, pp. 45–54, Jan. 2016.

[17] S. P. Kumar, N. Sriraam, P. G. Benakop, and B. C. Jinaga, "Entropies based detection of epileptic seizures with artificial neural network classifiers," *Expert Syst. Appl.*, vol. 37, no. 4, pp. 3284–3291, Apr. 2010.

[18] M. Bedeeuzzaman, T. Fathima, Y. U. Khan, and O. Farooq, "Mean absolute deviation and wavelet entropy for seizure prediction," *J. Med. Imag. Health Inform.*, vol. 2, no. 3, pp. 238–243, Sep. 2012.

[19] N. Nicolaou and J. Georgiou, "Detection of epileptic electroencephalogram based on permutation entropy and support vector machines," *Expert Syst. Appl.*, vol. 39, no. 1, pp. 202–209, Jan. 2012.

[20] D. Mateos, J. Diaz, and P. Lamberti, "Permutation entropy applied to the characterization of the clinical evolution of epileptic patients under pharmacologicaltreatment," *Entropy*, vol. 16, no. 11, pp. 5668–5676, Oct. 2014.

[21] X. Li, G. Ouyang, and D. A. Richards, "Predictability analysis of absence seizures with permutation entropy," *Epilepsy Res.*, vol. 77, no. 1, pp. 70–74, Oct. 2007.

[22] J. Xiang, C. Li, H. Li, R. Cao, B. Wang, X. Han, and J. Chen, "The detection of epileptic seizure signals based on fuzzy entropy," *J. Neurosci. Methods*, vol. 243, pp. 18–25, Mar. 2015.

[23] Y. Kumar, M. L. Dewal, and R. S. Anand, "Epileptic seizure detection using DWT based fuzzy approximate entropy and support vector machine," *Neurocomputing*, vol. 133, pp. 271–279, Jun. 2014.

[24] P. Li, C. Yan, C. Karmakar, and C. Liu, "Distribution entropy analysis of epileptic EEG signals," in *Proc. 37th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug. 2015, pp. 4170–4173.

[25] W. Chen, Z. Wang, H. Xie, and W. Yu, "Characterization of surface EMG signal based on fuzzy entropy," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 15, no. 2, pp. 266–272, Jun. 2007.

[26] H. Waqar, J. Xiang, M. Zhou, T. Hu, B. Ahmed, S. H. Shapor, M. S. Iqbal, and M. Raheel, "Towards classifying epileptic seizures using entropy variants," in *Proc. IEEE 5th Int. Conf. Big Data Comput. Service Appl. (BigDataService)*, Apr. 2019, pp. 296–300.

[27] M. Hayat and A. Khan, "Discriminating outer membrane proteins with fuzzy K-nearest neighbor algorithms based on the general form of Chou's PseAAC," *Protein Peptide Lett.*, vol. 19, no. 4, pp. 411–421, Feb. 2012.

[28] L. A. Cacha, S. Parida, S. Dehuri, S.-B. Cho, and R. R. Poznanski, "A fuzzy integral method based on the ensemble of neural networks to analyze fMRI data for cognitive state classification across multiple subjects," *J. Integrative Neurosci.*, vol. 15, no. 4, pp. 593–606, Dec. 2016.

[29] J. T. Oliva and J. L. G. Rosa, "Classification for EEG report generation and epilepsy detection," *Neurocomputing*, vol. 335, pp. 81–95, Mar. 2019.

[30] J. Orozco, C. A. R. García, "Detecting pathologies from infant cry applying scaled conjugate gradient neural networks," in *Proc. Eur. Symp. Artif. Neural Netw.*, Bruges, Belgium, Apr. 2003, pp. 349–354.

[31] A. A. Bruzzo, B. Gesierich, M. Santi, C. A. Tassinari, N. Birbaumer, and G. Rubboli, "Permutation entropy to detect vigilance changes and preictal states from scalp EEG in epileptic patients. A preliminary study," *Neurol. Sci.*, vol. 29, no. 1, pp. 3–9, Feb. 2008.

[32] Y. Cao, W.-W. Tung, J. B. Gao, V. A. Protopopescu, and L. M. Hively, "Detecting dynamical changes in time series using the permutation entropy," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 70, no. 4, Oct. 2004, Art. no. 046217.

[33] D. Geng, W. Zhou, Y. Zhang, and S. Geng, "Epileptic seizure detection based on improved wavelet neural networks in long-term intracranial EEG," *Biocybern. Biomed. Eng.*, vol. 36, no. 2, pp. 375–384, Jan. 2016.

[34] K. Patel, C.-P. Chua, S. Fau, and C. J. Bleakley, "Low power real-time seizure detection for ambulatory EEG," in *Proc. 3rd Int. Conf. Pervasive Comput. Technol. Healthcare*, Apr. 2009, pp. 1–7.

[35] Y. Park, L. Luo, K. K. Parhi, and T. Netoff, "Seizure prediction with spectral power of EEG using cost-sensitive support vector machines," *Epilepsia*, vol. 52, no. 10, pp. 1761–1770, Oct. 2011.

[36] J. R. Williamson, D. W. Bliss, and D. W. Browne, "Epileptic seizure prediction using the spatiotemporal correlation structure of intracranial EEG," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2011, pp. 665–668.

[37] Y. Wang, Z. Li, L. Feng, C. Zheng, and W. Zhang, "Automatic detection of epilepsy and seizure using multiclass sparse extreme learning machine classification," *Comput. Math. Methods Med.*, vol. 2017, Jun. 2017, Art. no. 6849360.

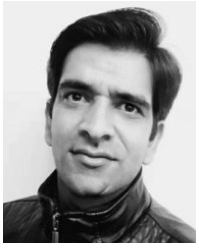
[38] L. M. Patnaik and O. K. Manyam, "Epileptic EEG detection using neural networks and post-classification," *Comput. Methods Programs Biomed.*, vol. 91, no. 2, pp. 100–109, Aug. 2008.

[39] S. Xie and S. Krishnan, "Wavelet-based sparse functional linear model with applications to EEGs seizure detection and epilepsy diagnosis," *Med. Biol. Eng., Comput.*, vol. 51, nos. 1–2, pp. 49–60, 2013.

[40] Q. Yuan, W. Zhou, Y. Liu, and J. Wang, "Epileptic seizure detection with linear and nonlinear features," *Epilepsy Behav.*, vol. 24, no. 4, pp. 415–421, Aug. 2012.

[41] E. Alickovic, J. Kevric, and A. Subasi, "Performance evaluation of empirical mode decomposition, discrete wavelet transform, and wavelet packed decomposition for automated epileptic seizure detection and prediction," *Biomed. Signal Process. Control*, vol. 39, pp. 94–102, Jan. 2018.

[42] A. Subasi, J. Kevric, and M. A. Canbaz, "Epileptic seizure detection using hybrid machine learning methods," *Neural Comput. Appl.*, vol. 31, no. 1, pp. 317–325, Jan. 2019.



WAQAR HUSSAIN received the B.S. degree in computer science from the University of Azad Jammu and Kashmir, Muzaffarabad, Pakistan, in 2006, and the M.S. degree from the Department of Computer Science, Muhammad Ali Jinnah University Islamabad, Pakistan, in 2011. He is currently pursuing the Ph.D. degree with the College of Information and Computer, Taiyuan University of Technology, Taiyuan, China. His major is computer application technology. His research interests include biomedical signal processing, deep learning, machine learning, and pattern learning.

BIN WANG received the Ph.D. degree from Okayama University, Japan, in 2013. He is currently an Associate Professor with the School of Information and Computer Science and Technology, Taiyuan University of Technology. He is mainly engaged in the study of brain cognitive function and the diagnosis of mental diseases.

YAN NIU is currently pursuing the Ph.D. degree with the College of Information and Computer, Taiyuan University of Technology, Taiyuan, China. Her major is computer application technology. She has published more than 17 articles. She is also certified by a senior software engineer. Her research interests include the exploration of nonlinear dynamics methods, study of functional magnetic resonance signal complexity, and early diagnosis of Alzheimer's disease.

YUAN GAO is currently pursuing the M.S. degree with the College of Information and Computer Science, Taiyuan University of Technology, Taiyuan, China. She is currently conducting a research in brain neuroscience. She has published two articles. Her research interests include brain network analysis, deep learning, and complexity analysis.

XIN WANG is currently pursuing the bachelor's degree with the College of Information and Computer, Taiyuan University of Technology. He is currently conducting a research in brain neuroscience. He has published one article. His research interests include intelligent information processing, brain informatics, and machine learning.

JIE SUN received the B.S. degree from the Taiyuan University of Technology, where she is currently pursuing the Ph.D. degree. Her major is computer science and technology research which directed at intelligent information processing, brain non-linear analysis, and machine learning. Her research interests include get high accuracy for Alzheimer's disease and known brain complexity changes during aging. She received the computer three-level and certified by a middle software engineer.

QIONGHUI ZHAN received the B.S. degree in computer science and technology from Hubei Normal University, Huangshi, China, in 2013. She is currently pursuing the M.S. degree in information and computer science with the Taiyuan University of Technology, Taiyuan, China. She has published an article in Cerebral Cortex on February 2019. Her research interests include medical image processing, brain networks, neurocognitive science, image processing, cognitive neural networks, neurodegenerative disease, diffusion tensor imaging, complex networks, and machine learning.

RUI CAO received the Ph.D. degree in computer application technology from the Taiyuan University of Technology. He is currently an Associate Professor of the Software College, Taiyuan University of Technology. He is mainly engaged in the research of brain science and intelligent information processing.

ZHOU MENGNI is currently pursuing the Ph.D. degree with the Graduate School of Interdisciplinary Science and Engineering in Health Systems, Okayama University. She is currently conducting a research in cognitive neuroscience. She has published five articles. Her research interests include EEG signal analysis based on deep learning, complexity analysis, and brain network analysis.

MUHAMMAD SHAHID IQBAL is currently pursuing the Ph.D. degree with the School of Computer Science and Application Technology, Anhui University, Hefei, China. His research interests include computer vision, image processing, machine learning, deep learning, bioinformatics, automated software engineering, and software testing.

JIE XIANG received the M.S. and Ph.D. degrees in computer science and technology from the Taiyuan University of Technology. She is currently a Professor and the Doctoral Supervisor with the College of Information and Computer, Taiyuan University of Technology. She is mainly engaged in brain science and intelligent computing, intelligent information processing, and big data management and analysis.