Biomedical Signal Processing and Control 77 (2022) 103820

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

An EEG based real-time epilepsy seizure detection approach using discrete wavelet transform and machine learning methods

Mingkan Shen [a,1], Peng Wen [a], Bo Song [a], Yan Li [b]

[a] School of Mechanical and Electrical Engineering, University of Southern Queensland, Toowoomba, QLD 4350, Australia
[b] School of Science, University of Southern Queensland, Toowoomba, QLD 4350, Australia

ARTICLE INFO

Keywords:
EEG
Real-time seizure detection
Discrete wavelet transform
Support vector machine
RUSBoosted tree Ensemble

ABSTRACT

Epilepsy is one of the most common complex brain disorders which is a chronic non-communicable disease caused by paroxysmal abnormal super-synchronous electrical activity of brain neurons. This paper proposed an electroencephalogram (EEG) based real-time approach to detect epilepsy seizures. Discrete wavelet transform and eight eigenvalues' algorithms are applied to extract features in different sub-frequency bands. Then support vector machine is employed for three-classes classification of health control, seizure free and seizure active, and finally RUSBoosted tree Ensemble method is used for real-time seizure onset detection. The proposed algorithm is evaluated using two public datasets: one short-term dataset named UB and one long-term dataset named CHB-MIT. The results show that the algorithm achieves 97% accuracy and 96.67% sensitivity in the three-classes classification of health control, seizure-free and seizure-active groups in UB dataset, and 96.38% accuracy, 96.15% sensitivity, 3.24% false positive rate for the real time seizure onset detection in CHB-MIT Dataset.

1. Introduction

Epilepsy is a chronic non-communicable disease caused by the abnormal synchronous electrical activity of brain neurons [1,2]. It is also one of the most common neurological diseases in the world, and affects approximately 50 million people [2,3]. Due to the differences in the starting region and propagation mode of abnormal electrical activity in the brain, the clinical manifestations of epilepsy are diversified and complicated [4]. Repeated seizures can cause persistent negative effects on patients' mental and cognitive functions, and bring life-threatening risks [5]. Therefore, research on the diagnosis and treatment of epilepsy has very important clinical significance. Automatic identification of epilepsy seizures from electroencephalogram (EEG) signals and its real-time implementation can provide an objective reference basis for the diagnosis and in time evaluation of epilepsy, thereby reducing the workload of doctors and improving the efficiency of treatment [6]. Majority of the recent papers have set the ultimate objective of developing automated EEG monitoring system to detect epileptic seizures. Bhattacharyya et al. highlighted a real-time seizure detection through empirical Wavelet transform method [7]. Disruptive EEG networks for epileptic seizures in real-time application reported by Bomela et al. [8]. Harmonic Wavelet packet transform with relevance vector machine method were proposed by Vidyaratne et al. [2].

EEG is a microvolt level electrical signal generated by synchronized neuronal activity in the brain collected by electrodes placed at a specific position on the scalp [9,10]. EEG abnormalities in epileptic seizures are mainly manifested as spike waves and sharp waves [11]. Therefore, using feature extraction methods to find the eigenvalues which can divide the normal waves and spike or sharp waves in different seizure state should be thoroughly investigated. So far many methods in time, frequency, and time–frequency domains have been developed such as discrete wavelet transform (DWT), empirical mode decomposition (EMD), Q-wavelet transformation, Hilbert-Huang transform (HHT), mean amplitude spectrum (MAS), etc [7,12–15]. Another important progress of epilepsy seizure detection is the development of machine learning based classification methods. The main objective of machine learning methods is to overcome the robustness of EEG individuals in epilepsy detection. Specifically, support vector machine (SVM), linear discriminant analysis (LDA), naive Bayes, logistic regression (LR), random forest were used to classify the different seizure states in previous studies [12,16–19]. Currently, automatic epilepsy detection can be divided into two types: offline seizure detection and real-time seizure detection. The purpose of offline seizure detection is to identify epileptic seizure signals as accurately as possible from EEG signal [20]. The

E-mail addresses: Mingkan.Shen@usq.edu.au (M. Shen), Paul.Wen@usq.edu.au (P. Wen), Bo.Song@usq.edu.au (B. Song), Yan.Li@usq.edu.au (Y. Li).
[1] 0000-0002-6770-4366

https://doi.org/10.1016/j.bspc.2022.103820
Received 27 October 2021; Received in revised form 29 March 2022; Accepted 16 May 2022
Available online 25 May 2022

purpose of real-time seizure detection is to identify seizures onsite with the shortest possible delay when the patient has a seizure during continuous EEG monitoring [21].

Two public datasets are available in EEG seizure detection. Dataset UB is a short-term dataset from the University of Bonn, which is used to do seizure event detection through the classifications of two classes (seizure free and seizure active) and three classes (health control, seizure free and seizure active). In the studies of two-classes classification, fractional linear prediction and SVM were used by Joshi et al. and achieved 95.33% accuracy [22]. Fast Fourier transform with k-nearest neighbor (k-NN) model proposed by Ghaderyan et al. could result in 98.72% accuracy [23]. In addition, The Dual-tree complex wavelet and the nearest neighbor (NN) model was reported to have 95.5% accuracy by Chen et al. [24]. Meanwhile, in the three classes classification studies, Acharya et al. discussed four entropy parameters (approximate entropy, sample entropy, two phase entropies) combined with fuzzy classifier, and achieved 98.1 % accuracy [25]. Omidvar et al. used the DB4-DWT method based on the artificial neural network and SVM models, and got 98.7% accuracy [12]. Currently, feature extraction based on machine learning classification is one of the most researched approaches in seizure event detection using EEG signal. Tunable-Q wavelet transform based multiscale entropy measure proposed by Bhattacharyya, A., et al. is used to classifier 3-classes between normal, seizure-free and seizure EEG signals and achieve 98.6% accuracy results [26]. Gupta, V. and R.B. Pachori stated that Fourier-Bessel series expansion (FBSE) and weighted multiscale Renyi permutation entropy (WMRPE) for EEG rhythms and get 97.3% accuracy results [27]. Empirical wavelet transform (EWT) with FBSE method highlighted by Anuragi, A., et al. can also achieved 97.7% accuracy classification [28].

Dataset CHB-MIT is a long-term dataset from Boston Children's Hospital, which is used by many researchers to do the real-time automatic seizure detection. Samiee et al. used multivariate textural features with gray-level co-occurrence matrix (GLCM) in SVM and reported 70.19% sensitivity in the real-time seizure detection [29]. As a contrast, time delay embedding method proposed by Zabihi et al. was reported to have 89.01% sensitivity [30]. In particularly, graph theory analysis, function connectivity analysis and effective connectivity analysis were used in the seizure detection [16,31,32]. Bomela et al. constructed the network connectivity using Fourier transform to detect the seizure onset and reported 93.6 % sensitivity and a false positive rate of 0.16 per hour (FP/h) result [8]. A stactked 1D-CNN model is presented via Wang, X., et al. to detect seizure onset automatically and achieved 88.14% accuracy and 0.38% false positive (FP) result [33]. Orthogonal matching pursuit with DWT as pre-processing progress with non-linear features and SVM classifier can also detect the seizure onset in same dataset, Zarei, A. and B.M. Asl, used this method reported 96.81% sensitivity and 2.74% FP result [34]. Li, C., et al. proposed EMD, common spatial pattern and SVM model get 97.34% sensitivity, 2.5% FP output as well [35].

In this study, the proposed real time EEG based seizure detection method includes four major steps using the aforementioned both datasets (UB and CHB-MIT) in two experiments. Specifically, in the first experiment using Dataset UB, DWT analysis with DB4 mother wavelet was used to decompose the raw EEG signal data. After feature extraction and selection, 12 eigenvalues were evaluated as the input of the SVM model to classify health control, seizure-free and seizure-active subjects. Based on the first experiment, real-time seizure detection was implemented using Dataset CHB-MIT. Similarly, DB16 DWT analysis with 7 eigenvalues were fed into the SVM and RUSBoosted tree Ensemble model to obtain the final results. All the experiments in this study were carried out in a Dell workstation with dual Intel Xeon E5-2697V3 CPUs using MATLAB 2019b. The main contributions and innovations of this study are: (1) DB4-DWT and DB16-DWT were proposed to extract approximate and details of signals and remove redundant information. (2) Improved the robustness of EEG based epilepsy detection via machine learning methods. (3) Proposed a method that can achieve 97% accuracy and 96.67% sensitivity in 3-class classification (health control, seizure free and seizure active) using Dataset UB, and 96.38% accuracy, 96.15% sensitivity and 3.24% false positive rate in the real-time seizure detection using Dataset CHB-MIT. (4) Implemented an automatic seizure detection approach in real-time way.

The first section of the paper provided a brief introduction of the work. Section II described the details of the short-term dataset (Dataset UB) and long-term dataset (Dataset CHB-MIT). The pre-processing, feature extraction, classification and real-time application are also introduced in this section. Section III reported the work in our experiments and results obtained using the proposed method. Comparisons of previous work using the same datasets were conducted and evaluated in Section IV. Section V concluded the paper.

2. Methodology

The proposed methodology utilized DWT for the data pre-processing, and calculated nine eigenvalues via entropy-based and statistical measures to extract features. SVM and RUSBoosted tree Enemble methods were used to train and test Dataset UB and Dataset CHB-MIT. The framework of the proposed method is described as follow Fig. 1:

2.1. Datasets

Dataset UB is collected from the University of Bonn which consists of the sets F, N, O, S and Z [36]. Each set contains 100 single channel segments with 23.6 s duration in 173.61 Hz sample rate. Sets Z and O were collected from 5 healthy control subjects via surface EEG standard 10–20 system caps, which Set Z is open eyes EEG data and Set O is closed eyes EEG data [37]. Sets N and Sets F were collected from 5 epileptic patients through seizure-free state via intracranial EEG (iEEG) signals. Sets S involves 5 epileptic patients in seizure active state by iEEG signals as well [36,37]. According to the Dataset UB has three different classes that health control, seizure free and seizure active respectively, thus, the first experiment in this study is to do the 3 classes' classification detection.

Dataset CHB-MIT is collected from Boston Children's Hospital with 23 subjects (5 males between ages 3 to 22 years and 17 females of ages between 1.5 and 19 years) [38]. All CHB-MIT database was sampled at 256 Hz, and it collected in 23 bipolar channels by scalp EEG standard 10–20 system caps. Here, Channel FT10 - T8 is used to detect the seizure onset in real-time applications. This study used 16 patents from the CHB-MIT database, and excluded patients that had seizures characterized by amplitude depression [15]. Furthermore, case'Chb24′ was not used in the real-time seizure detection due to the insufficient specific real-time information. The details of the CHB-MIT used in this study is shown in Table 1.

2.2. Pre-processing

One of the most challenging parts in epilepsy seizure detection is to detect the sharp waves and spike waves. However, not all features in all frequency bands are significantly different among seizure active, seizure free and health control states. Thus, DWT is proposed to decompose an EEG signal in different frequency sub-bands into frequency components, and the formula of DWT is shown as follow [39,40]:

$$C(a,b) = \frac{1}{\sqrt{a}} \int \overline{\psi}(\frac{t-b}{a})x(t)dt \tag{1}$$

where ψ is the analyzing wavelet method, 'a' and 'b' are the parameters of time dilation and time translation, respectively.

After DWT decomposition, two style coefficients the detail coefficients and approximation coefficients of each sub-bands are calculated in Eqs. (2) and (3).

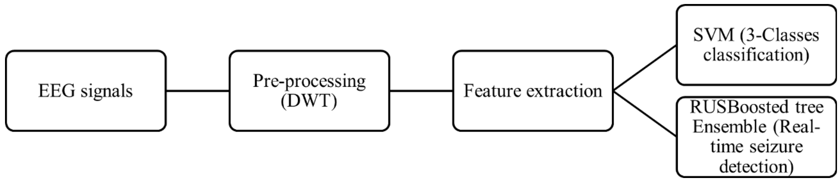

Fig. 1. The framework of DWT and machine learning methods for seizure detection.

Table 1
Long term EEG data (Dataset CHB-MIT).

Patient	EEG used (h)	Number of seizures	Seizure duration (s)
Chb01	25	7	40,27,40,51,90,93,101
Chb02	16	3	82,81,9
Chb03	36	7	52,65,69,52,47,64,53
Chb04	25	4	49,111,102,116
Chb05	14	5	115,110,96,120,117
Chb07	28	3	86,96,143
Chb08	16	5	171,190,134,160,264
Chb09	34	4	64,79,71,62
Chb10	20	7	35,70,65,58,76,89,54
Chb11	12	3	22,32,752
Chb17	15	3	90,115,88
Chb18	18	6	50,30,68,55,68,46
Chb19	14	3	78,77,81
Chb20	15	8	29,30,39,38,35,49,35,39
Chb22	15	3	58,74,72
Chb23	14	7	113,20,47,71,62,27,84

$$A_j(n) = \sum_{l=-\infty}^{+\infty} g(l-2n)A_{(j-1)}(l), j = 1, 2, ..., J \tag{2}$$

$$D_j(n) = \sum_{l=-\infty}^{+\infty} h(l-2n)A_{(j-1)}(l), j = 1, 2, ..., J \tag{3}$$

where $A_j(n)$ and $D_j(n)$ are the approximation coefficients and detail coefficients at level j respectively.

The sample rate of Dataset UB is 173.61 Hz, 5-level DWT with 'DB4′ wavelet technique is used to decompose the data.D_1,D_2,D_3, D_4 and A_4 coefficients are used to represent the EEG sub-bands described in Table 2. The pre-processing method used in Dataset CHB-MIT was slightly changed to accommodate the different sampling rates. 6-level DWT with 'DB16′ wavelet method was applied to decompose the components and the details are shown as Table 3.

2.3. Feature extraction

Nine features were calculated to assess the time-domain signal correspond to sub-band j to find the difference in different seizure states. The features include the standard deviation (SD), mean band power (BP), Shannon entropy (SE), log-energy entropy (LE), fuzzy entropy (FE), maximum, kurtosis, and median.

The SD can describe the degree of dispersion of the signal, is defined as.

Table 2
5-level DWT decomposition (DB4) in Dataset UB.

Sub-band j	Decomposed signal	Frequency band (Hz)
1	D1	43.4–86.8
2	D2	21.7–43.4
3	D3	10.8–21.7
4	D4	5.4–10.8
4	A4	0–5.4

Table 3
6-level DWT decomposition (DB16) in Dataset CHB-MIT.

Sub-band j	Decomposed signal	Frequency band (Hz)
1	D1	64–128
2	D2	32–64
3	D3	16–32
4	D4	8–16
5	D5	4–8
5	A5	0–4

$$SD = \sqrt{\frac{1}{N}\sum_{n=0}^{N}(S_n - \mu)^2} \tag{4}$$

where μ is the mean value of the EEG segments.

$$\mu = \frac{1}{N}\sum_{i=1}^{N} S_n \tag{5}$$

The BP of the signal is calculated to assess the power of time-domain amplitude.

$$BP = \frac{1}{N}\sum_{i=1}^{N} S_n^2 \tag{6}$$

Entropy is the parameter to define the confusion of the signal. In particular, SE, LE and FE are calculated to detect seizures [41,42].

$$SE = -\sum_{i=1}^{N} p_i \log_2 p_i \tag{7}$$

$$LE = -\sum_{i=1}^{N} (\log_2 p_i)^2 \tag{8}$$

$$FE(m, n, r, N) = \ln\phi^m(n, r) - \ln\phi^{m+1}(n, r) \tag{9}$$

where p_i is the probability of occurrence in the EEG segments, and.

$$\phi(n, r) = \frac{1}{N-m}\sum_{i=1}^{N-m}\left[\frac{1}{N-m-1}\sum_{j=1, j\neq i}^{N-m} D_{ij}^m\right] \tag{10}$$

where m is the embedding dimension, r is the threshold which equals to $0.15 \times$ SD and n is the fuzzy power. Empirically, values like $m = 2, n = 2$ produced the best performance.

Kurtosis is a measure of the peak of the probability distribution of a random variable. High kurtosis means that the increase in variance is caused by extreme differences in low frequencies that are greater than or less than the average.

$$Max = \max(S_n) \tag{11}$$

$$M = median(S_n) \tag{12}$$

$$K = \frac{1}{N-1}\sum_{i=1}^{N}\frac{(S_i - \mu)^4}{\sigma^4} \tag{13}$$

where μ is the mean value and σ is the SD of the EEG segments.

There are nine features for each sub-band's signal, and total 45 eigenvalues for the 5 sub-bands from Dataset UB. Similarly, nine features for each sub-band's signal, total 54 eigenvalues from Dataset CHB-MIT.

2.4. Classification

In Dataset UB, SVM was applied in the 3-classe classifications (health control, seizure free and seizure active). The leaving one out training method and RUSBoosted tree Ensemble were used to detect seizure onset based on Dataset CHB-MIT.

2.4.1. Three -class classification for Dataset UB

Polynomial function of SVM helps the model to improve the accuracy in classification. Two key parameters of SVM model, γ and c, were selected, where γ is the inverse of the radius of influence of samples selected by the model as support vectors, and c parameter trades off correct classification of training examples against maximisation of the decision function's margin. Here, $\gamma = 0.1$, and $c = 1$.

In Dataset UB, 80% data (segment 1 to segment 80, all 400 segments) of each set (Set Z, O, N, F, S) is used to training in SVM model while the remaining 20% (segment 81 to segment 100, all 100 segments) were used to test the performance of the proposed method.

2.4.2. Leaving one out experiment for Dataset CHB-MIT

Sixteen patients' data from dataset CHB-MIT (detail shown in Table 1) was used in this study. In leaving one out training method, one subject data is used for testing and the other 15 subjects were used for training. As a result, 16 models have been trained. The same SVM model constructed in Dataset UB was re-used in Dataset CHB-MIT. But in Dataset CHB-MIT, it is a two-class classification problem in seizure free and seizure active, thus the RUSBoosted tree Ensemble model function with 5-fold cross validation in MATLAB Classification Learner toolbox was applied to conduct the seizure detection and comparisons with the SVM model.

In this study, the EEG data was segmented into 30 s epoch with 256 Hz sample rate, which resulted 7680 sampling points in each epoch. In all 16 subjects, the EEG raw data of 30 min before seizure epochs and 30 min after seizure epochs for each subject data were used to train.

2.5. Real time implementation

The real-time application is implemented and evaluated in Dataset CHB-MIT, as the Dataset UB is not a continuous real-time data. In particular, a 30-second sliding window was developed and data within the moving window was considered as the input data, and the sliding window overlap was selected as 1 s in the real-time detection. The features of$BP-D_5$,$SD-D_4$,$SD-D_5$,$SE-D_4$,$SE-D_5$,$LE-D_4$, $LE-D_5$ with the EEG raw data of Channel FT10 - T8 for case 'Chb01′ first seizure is shown in Fig. 2.

Obviously, the eigenvalues have changed significantly before and after the seizure onset (the 2996 s). Thus, these eigenvalues were selected to train in the machine learning model. Using leaving one out training method, 16 models have been trained, the left one subject as the test data is used to test by corresponding models.

3. Experiments and results

Accuracy and sensitivity were used to evaluate the 3-class classification for Dataset UB, while accuracy, sensitivity, false positive rate and seizure onset detection delay were used to evaluate the proposed method for Dataset CHB-MIT.

Accuracy is a direct parameter in method evaluation which is define as follow:

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \tag{14}$$

where, TP is the true positive, TN is the true negative, FP is the false positive and FN is the false negative.

Sensitivity is another parameter for evaluation which is defined as:

$$Sen = \frac{TP}{TP + FN} \tag{15}$$

However, the sensitivity of Dataset UB and Dataset CHB-MIT is different because this parameter is focus on how much seizure times has

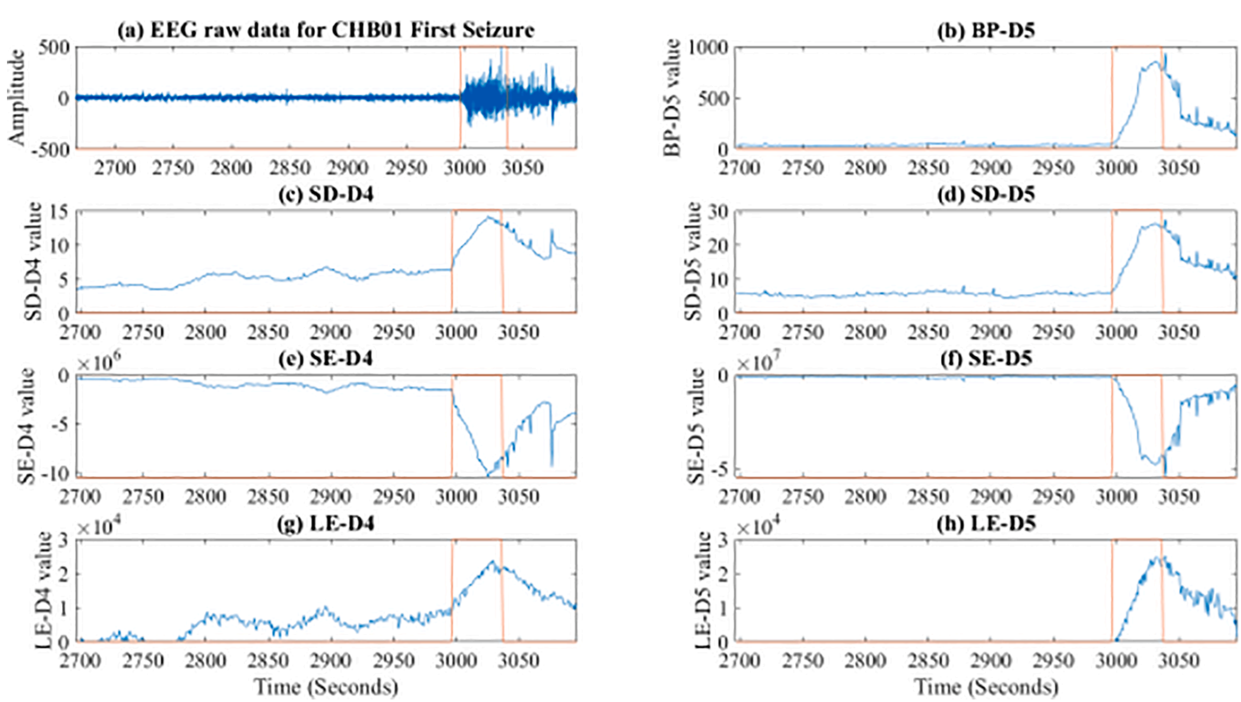

Fig. 2. (a) Raw EEG signal of Chb01, FT10 – T8 Channel, (b) BP of D5 level, (c) SD of D4 level, (d) SD of D5 level, (e) SE of D4 level, (f) SE of D5 level, (g) LE of D4 level, (h) LE of D5 level, orange line is the seizure states labelled by the doctor from the 2996 s to 3036 s.

been detected in real-time detection. The sensitivity of this part is calculated as:

$$Sen_{real-time} = \frac{TP}{NS} \tag{16}$$

where *NS* means the number of the seizures.

3.1. Three-class classification results using Dataset UB

There are 45 eigenvalues for these EEG signals. However, the performance of those features was quite different. Empirically, 12 eigenvalues were selected to train the model, and their details are shown in Table 4(a). The last 20% data (segment 81 to segment 100 for each Set, all 100 segments) were used to test, and the classification results of 97% accuracy, 96.67% sensitivity were received for Dataset UB.

3.2. Real-time seizure onset detection results using Dataset CHB-MIT

About 7 eigenvalues from 53 eigenvalues were selected in the real-time implementation using Dataset CHB-MIT, and the details of features selection is shown in Table 4(b). SVM and RUSBoosted tree Ensemble machine method were applied to evaluate the model using leaving one training method. As a result, RUSBoosted tree Ensemble achieved results of 96.15% sensitivity, 96.38% accuracy and 3.24% false positive rate (Tables 5 and 6).

4. Discussion

We used DWT to decompose EEG raw signal into different frequency bands, after that nine eigenvalues are calculated in each sub-band. However, not every sub-band's features are obviously different between seizure active state and seizure free state. Thus, we studied the level of decompositions, mother wavelet selection, and sliding window size selection to get the best performance. The work and results are presented below:

4.1. Level of decompositions and mother wavelet selection

To extract features from EEG raw signal, we compared several eigenvalues in 3 different states in Dataset UB, and before and after seizure onset in Dataset CHB-MIT. After computing all 9 eigenvalues of different seizure state, 5 levels for Dataset UB and 6 levels for Dataset CHB-MIT are selected as they show a significant difference between different seizure states. The performance of BP in different levels is shown in Fig. 3, which clearly shows that the eigenvalues of BP in D_4 and D_5 have a better performance. Therefore, the parameter of the level of decompositions was selected as 6.

There is no standard method to choose the best wavelet. Omidvar et al. reported a method using 5-level DB4 DWT to the seizure classification using the Set Z, Set F and Set S of Dataset UB, and they got 98.7 % accuracy [12]. Following the same mother wavelet technique DB4, this study achieved a satisfying result as well using Dataset UB. However, in another real-time experiment of Dataset CHB-MIT, DB4 is not the best mother wavelet because of different sample rate of two datasets and different level of decompositions. After increasing the DB value from 4 to 6, 8, …20 while calculate 7 eigenvalues, we found DB16 mother wavelet can distinguish seizure active and seizure free states more accurately.

Table 4
Selected features for Dataset UB and CHB-MIT.

(a) Features selection for Dataset A		(b) Features selection for Dataset B	
Decomposition level	Features	Decomposition level	Features
D3	SD	D5	BP
A4	SD	D4	SD
D3	Mean	D5	SD
D4	Mean	D4	SE
D2	LE	D5	SE
D3	LE	D4	LE
D3	FE	D5	LE
D2	Max		
D3	Max		
D3	Median		
D2	Kurtosis		
D3	Kurtosis		

Table 5
Real time detection using Dataset CHB-MIT and SVM method.

Patient	NS	TP	FP (%)	Sen (%)	Delay (s)	Acc (%)
Chb01	7	7	0.0	100	20.7	99.60
Chb02	3	3	0.0	100	19.0	99.76
Chb03	7	4	0.2	57.1	25.8	99.30
Chb04	4	3	0.1	75	60.0	99.62
Chb05	5	5	16.3	100	16.4	83.59
Chb07	3	3	0.2	100	18.7	99.64
Chb08	5	5	0.1	100	19.6	98.60
Chb09	4	3	11.4	75	−171.0	88.40
Chb10	7	7	0.7	100	23.6	98.83
Chb11	3	1	0.1	33.3	91.0	98.16
Chb17	3	2	0.0	66.7	43.0	99.53
Chb18	6	3	0.0	50	30.3	99.57
Chb19	3	3	0.0	100	46.0	99.66
Chb20	8	0	0.0	0	–	99.41
Chb22	3	3	0.0	100	39.7	99.66
Chb23	7	6	0.7	85.7	29.0	98.60
Total	78	58				
Mean			1.86		20.79	97.62
Sen		74.36				

Table 6
Real time detection using Dataset CHB-MIT and RUSBoosted tree Ensemble method.

Patient	NS	TP	FP (%)	Sen (%)	Delay (s)	Acc (%)
Chb01	7	7	0.0	100	3.1	99.70
Chb02	3	3	6.1	100	−52.7	93.86
Chb03	7	7	2.3	100	18.6	97.32
Chb04	4	4	8.3	100	38.5	91.49
Chb05	5	5	7.0	100	−4.2	92.75
Chb07	3	3	4.6	100	4.3	95.41
Chb08	5	5	3.0	100	1.8	96.50
Chb09	4	4	0.0	100	13.6	98.72
Chb10	7	7	5.6	100	18.3	94.09
Chb11	3	1	0.6	33.3	19.0	98.14
Chb17	3	3	0.5	100	22.7	99.27
Chb18	6	5	1.2	83.3	27.0	98.42
Chb19	3	3	0.4	100	15.7	99.48
Chb20	8	8	3.4	100	27.8	96.10
Chb22	3	3	0.2	100	14.7	99.68
Chb23	7	7	8.6	100	−1.5	91.13
Total	78	75				
Mean			3.24		10.42	96.38
Sen		96.15				

'*NS*' is the number of seizures, '*TP*' is true positive, '*FP*' is false positive rate, '*Sen*' is sensitivity, and '*Acc*' is accuracy. 'delay' represents the difference between the detected seizure onset time and the doctor's marker. Delay is negative if detected the seizure active signal early.

4.2. Length of the sliding window

The sliding window size selection also directly affects the results. If the sliding window size is too small, the eigenvalues do not change before and after seizure. If the size is too big, the results may cause big delay in seizure onset detection. The performance of eigenvalue LE are provided as the results of two sliding window size (10 s and 30 s). According to the performance described in Fig. 4, the sliding window size was selected as 30 s in this study for seizure onset detection in Dataset

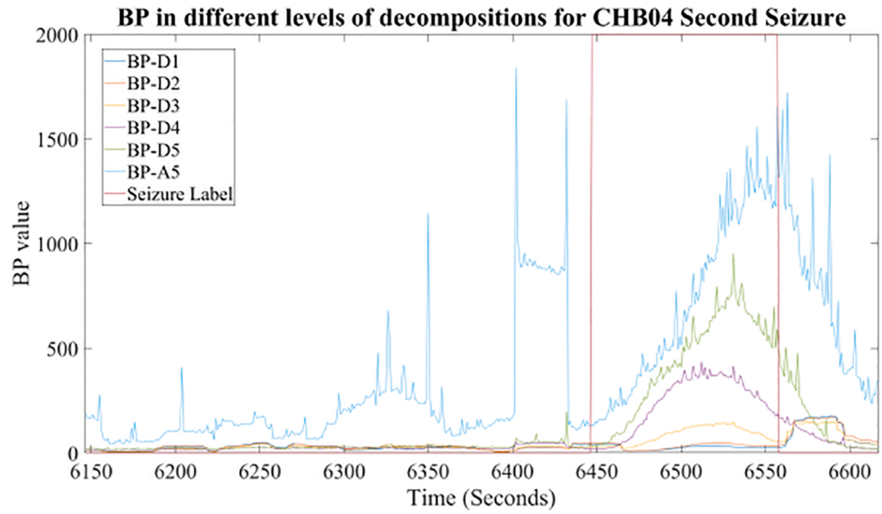

Fig. 3. BP in different levels of decompositions for Chb04 Second Seizure, seizure label line is the seizure state labelled by the doctor from the 6446 s to 6557 s.

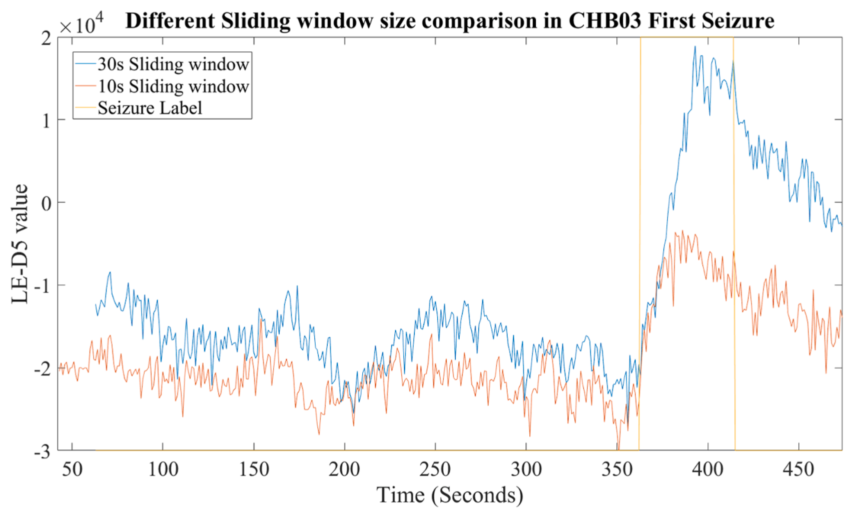

Fig. 4. LE-D5 in different sliding window size for Chb03 First Seizure, seizure label line is the seizure state labelled by the doctor from the 362 s to 414 s.

CHB-MIT because there are more significant differences in eigenvalue LE before and after seizure onset.

4.3. Training data and training method selection

Through the real-time sliding window analysis, 12 eigenvalues of Dataset UB and 7 eigenvalues of Dataset CHB-MIT were selected as the input of SVM and RUSBoosted tree Ensemble technique of the machine learning methods respectively, according to their performance in different seizure states and before and after seizure onset. The raw EEG data was segmented into 30 s epochs as the input for training. This work could avoid the confusing data, because in epilepsy onset detection the raw EEG signal wave will begin to change drastically in pre-seizure state which always appear before seizure onset several seconds. In addition, it significantly reduced the cost of training time as well.

Leaving one out training method based on SVM and RUSBoosted tree Ensemble models was used to analysis Dataset CHB-MIT. Leaving one out training method can be evaluated more objectively and compared with the random subject allocations for training and test. In addition, in seizure detection, the proportion of seizure free and seizure active differs significantly. Thus, the sensitivity of results is a very important parameter to assess the proposed method, and here RUSBoosted tree Ensemble model can provide a better result than SVM model (detail in Tables 5 and 6).

4.4. DWT and other methods in seizure detection

Tunable-Q wavelet transform (TQWT) as an advanced wavelet transform method has been used in epileptic EEG signals analysis. Bhattacharyya et al. proposed the TQWT with the entropy measure and achieved 98.6% accuracy results in Dataset UB [26]. Three parameters need to be defined in this experiment, which are the quality factor ‘Q’, the redundancy parameter ‘r’ and the number of decomposition levels ‘J’. To compare the performance of TQWT and DWT in this experiment, parameter ‘J’ was selected as $J = 5$ in Dataset UB and $J = 6$ in Dataset CHB-MIT to replace the proposed DWT using the same eigenvalues and process. Thus, these TQWT were extracted with the same amounts of sub-bands as the DWT in this work. The redundancy parameter ‘r’ will be sufficient if $r \geq 3$. For $r \approx 1$, the wavelet will resemble the ‘sinc’ wavelet. When $r \geq 3$, the passband of the level-J frequency response will not have a ‘flat top’ (where the frequency response is equal to a constant over a sub-interval of its passband). The parameter ‘r’ was selected as $r = 3$ in this experiment. The specified Q-factor should be chosen from the range of $Q \geq 1$. Setting $Q = 1$ leads to a wavelet transform for which the wavelet resembles the second derivative of a Gaussian. Higher values of Q lead to more oscillatory wavelets. In this experiment, we select $Q =$

1,2,3,4 with r = 3, J = 5 in Dataset UB and Q = 1,2,3,4 with r = 3, J = 6 in Dataset CHB-MIT. Comparing the results with different values of quality factor 'Q', the best performance was observed when Q = 2. To compare the TQWT with the proposed DWT in this case, we selected the parameters of TQWT as Q = 2, r = 3, J = 5 for dataset UB and Q = 2, r = 3, J = 6 for dataset CHB-MIT to ensure the same decomposition levels in DWT. Using the same features and machine learning models, the TQWT achieved 79% accuracy in the three classes classification of health control, seizure free and seizure active groups in Dataset UB, and 74.36% sensitivity in Dataset CHB-MIT during the real time simulation.

This study focused more on the features' performance in each decomposition level for seizure detection than the wavelet methods. EEG is often described in terms of rhythmic activity, so that different frequency bands in DWT correspond to various EEG rhythms. As shown in Table 3, Dataset CHB-MIT, each decomposition level can correspond to the defined frequency band such as delta band (0–4 Hz), theta band (4–8 Hz), alpha–beta band (8–16 Hz), low-gamma band (16–32 Hz), high gamma band (64–128 Hz). We found the features in theta band and alpha–beta band had more significant difference before and after seizure onset. In Dataset UB, because of the different sample rate from the Dataset CHB-MIT, we divided the dataset in five frequency bands as shown in Table 2. In addition, it still can correspond to various EEG rhythms just a little difference which are delta band (0–5.4 Hz), theta band (5.4–10.8 Hz), alpha–beta band (10.8–21.7 Hz), low gamma band (21.7–43.4 Hz) and high gamma band (43.4–86.8 Hz) respectively. Therefore, using DWT in EEG analysis can provide more insights from the clinic detection perspective and we can focus on the specific frequency bands rather than the whole frequency band to reduce the computational load during the calculation.

4.5. Performance comparison with Dataset UB

Table 7 summarizes the performance of the proposed method and other peer works in the three-class classification using Dataset UB. Comparing with the previous works, this study used the whole data of 5 Sets and achieved a promising result (97% accuracy and 96.67% sensitivity). In this study, 80% data (segment 1 to segment 80, all 400 segments) of each set (Set Z, O, N, F, S) is used to training and the remaining 20% (segment 81 to segment 100, all 100 segments) were used to test the performance of the proposed method.

4.6. Performance comparison with previous work using Dataset CHB-MIT

A similar method is used to implement the real-time detection in Dataset CHB-MIT. Table 8 summarizes the results of the proposed real-time method and previous offline works in seizure detection using Dataset CHB-MIT. The proposed method achieved 96.38% accuracy, 96.15% sensitivity and 3.24% false positive rate in real-time seizure onset detection.

5. Conclusion

This study proposed an EEG based real-time epilepsy seizure detection approach using DWT, SVM and RUSBoosted tree Ensemble models of machine learning, and evaluated its performance by comparison. Using the 12 eigenvalues in corresponding decomposition levels extracted by DB4-DWT, and SVM models, our study achieved a 97% accuracy and 96.67% sensitivity in the three classes classification (health control, seizure-free and seizure-active) of Dataset UB. Experiments show the proposed method can classify the seizure free, health control and the seizure active states. In addition, our study also implemented the real-time seizure detection using DB16-DWT in seven eigenvalues with RUSBoosted tree Ensemble method, and obtained a 96.38% accuracy, 96.15% sensitivity and 3.24% false positive rate in Dataset CHB-MIT. The proposed method is also suitable for real-time seizure detection.

Table 7
Comparison of the proposed method and previous works using Dataset UB.

References	Feature extraction	Classifier	Datasets	Acc (%)
Kumar et al. (2014) [44]	DWT + fuzzy approximate entropy	SVM	Z, F, S	95.67
Acharya et al. (2012) [25]	Four entropy parameters	Fuzzy classifier	Z, F, S	98.1
Kaya and Ertuğrul (2018) [45]	One-dimensional ternary patterns	Random Forest	Z, F, S	95.7
Zhang et al. (2018) [46]	Generalized Stockwell Transform, singular value decomposition	Random Forest	Z, F, S	99
Omidvar et al. (2021) [12]	DWT + 11 features	ANN, SVM	Z, F, S	98.7
Bhattacharyya, A., et al. (2017) [26]	Tunable-Q wavelet transform, K-NN entropy	SVM	Z, F, N, O, S	98.6
Gupta, V. and R.B. Pachori (2019) [27]	Fourier-Bessel series expansion (Morelet wavelet)	Least squares SVM	Z, F, S	97.3
Anuragi, A., et al. (2022) [28]	Fourier-Bessel series expansion, EWT	Ensemble classifiers	Z, F, N, O, S	97.7
Proposed method	DWT + 12 features	SVM	Z, F, N, O, S	97

Table 8
Comparison of the proposed method and previous works using Dataset CHB-MIT.

Reference	Sen (%)	FP (%)	Delay (s)
Ahammad et al. (2014) [20]	98.50	14.4	1.76
Samiee et al. (2015) [29]	70.19	2.26	Not reported
Zabihi et al. (2015) [30]	88.27	6.79	Not reported
Bhattacharyya and Pachori (2017) [7]	97.91	0.43	Not reported
Fan and Chou (2018) [47]	97	8.61	6–7
Bomela et al. (2020) [8]	93.6	0.16 per hour	10.06
Wang, X., et al. (2021) [33]	88.14	0.38	Not reported
Li, C., et al. (2021) [35]	97.34	2.5	Not reported
Zarei, A. and B.M. Asl (2021) [34]	96.81	2.74	Not reported
Proposed method	96.15	3.24	10.42

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] A.S. Zandi, M. Javidan, G.A. Dumont, R. Tafreshi, Automated real-time epileptic seizure detection in scalp EEG recordings using an algorithm based on wavelet packet transform, IEEE Trans. Biomed. Eng. 57 (7) (2010) 1639–1651.
[2] L.S. Vidyaratne, K.M. Iftekharuddin, Real-time epileptic seizure detection using EEG, IEEE Trans. Neural Syst. Rehabil. Eng. 25 (11) (2017) 2146–2156.
[3] R.B. Yaffe, P. Borger, P. Megevand, D.M. Groppe, M.A. Kramer, C.J. Chu, S. Santaniello, C. Meisel, A.D. Mehta, S.V. Sarma, Physiology of functional and effective networks in epilepsy, Clin. Neurophysiol. 126 (2) (2015) 227–236.
[4] J. Xiang, C. Li, H. Li, R. Cao, B. Wang, X. Han, J. Chen, The detection of epileptic seizure signals based on fuzzy entropy, J. Neurosci. Methods 243 (2015) 18–25.
[5] D.C. Bergen, Do seizures harm the brain? Epilepsy Curr. 6 (4) (2006) 117–118.
[6] R. Rosch, T. Baldeweg, F. Moeller, G. Baier, Network dynamics in the healthy and epileptic developing brain, Network Neurosci. 2 (1) (2018) 41–59.
[7] A. Bhattacharyya, R.B. Pachori, A multivariate approach for patient-specific EEG seizure detection using empirical wavelet transform, IEEE Trans. Biomed. Eng. 64 (9) (2017) 2003–2015.
[8] W. Bomela, S. Wang, C.-A. Chou, J.-S. Li, Real-time inference and Detection of Disruptive eeG networks for epileptic Seizures, Sci. Rep. 10 (1) (2020).
[9] Y.i. Liang, C. Chen, F. Li, D. Yao, P. Xu, L. Yu, J.T. Cao, Altered Functional Connectivity after Epileptic Seizure Revealed by Scalp EEG, Neural Plasticity 2020 (2020) 1–8.
[10] H. Yu, L. Zhu, L. Cai, J. Wang, C. Liu, N. Shi, J. Liu, Variation of functional brain connectivity in epileptic seizures: an EEG analysis with cross-frequency phase synchronization, Cogn. Neurodyn. 14 (1) (2020) 35–49.

[11] J. Wu, T. Zhou, T. Li, Detecting epileptic seizures in EEG signals with complementary ensemble empirical mode decomposition and extreme gradient boosting, Entropy 22 (2) (2020) 140.
[12] M. Omidvar, A. Zahedi, H. Bakhshi, EEG signal processing for epilepsy seizure detection using 5-level Db4 discrete wavelet transform, GA-based feature selection and ANN/SVM classifiers, J. Ambient Intell. Hum. Comput. 12 (11) (2021) 10395–10403.
[13] K. Jindal, R. Upadhyay, H.S. Singh, Application of tunable-Q wavelet transform based nonlinear features in epileptic seizure detection, Analog Integr. Circ. Sig. Process 100 (2) (2019) 437–452.
[14] R.J. Oweis, E.W. Abdulhay, Seizure classification in EEG signals utilizing Hilbert-Huang transform, Biomed. Eng. Online 10 (1) (2011) 1–15.
[15] W. Hu, et al., Mean amplitude spectrum based epileptic state classification for seizure prediction using convolutional neural networks, J. Ambient Intell. Hum. Comput. (2019) 1–11.
[16] B. Akbarian, A. Erfanian, A framework for seizure detection using effective connectivity, graph theory, and multi-level modular network, Biomed. Signal Process. Control 59 (2020) 101878.
[17] Y. Gao, Z. Zhao, Y. Chen, G. Mahara, J. Huang, Z. Lin, J. Zhang, Automatic epileptic seizure classification in multichannel EEG time series with linear discriminant analysis, Technol. Health Care 28 (1) (2020) 23–33.
[18] Y.L. Niriayo, A. Mamo, T.D. Kassa, S.W. Asgedom, T.M. Atey, K. Gidey, G. T. Demoz, S. Ibrahim, Treatment outcome and associated factors among patients with epilepsy, Sci. Rep. 8 (1) (2018).
[19] C. Donos, M. Dümpelmann, A. Schulze-Bonhage, Early seizure detection algorithm based on intracranial EEG and random forest classification, Int. J. Neural Syst. 25 (05) (2015) 1550023.
[20] N. Ahammad, T. Fathima, P. Joseph, Detection of epileptic seizure event and onset using EEG, BioMed. Res. Int. 2014 (2014) 1–7.
[21] M. Qaraqe, M. Ismail, E. Serpedin, H. Zulfi, Epileptic seizure onset detection based on EEG and ECG data fusion, Epilepsy Behav. 58 (2016) 48–60.
[22] V. Joshi, R.B. Pachori, A. Vijesh, Classification of ictal and seizure-free EEG signals using fractional linear prediction, Biomed. Signal Process. Control 9 (2014) 1–5.
[23] P. Ghaderyan, A. Abbasi, M.H. Sedaaghi, An efficient seizure prediction method using KNN-based undersampling and linear frequency measures, J. Neurosci. Methods 232 (2014) 134–142.
[24] G. Chen, Automatic EEG seizure detection using dual-tree complex wavelet-Fourier features, Expert Syst. Appl. 41 (5) (2014) 2391–2394.
[25] U.R. Acharya, F. Molinari, S.V. Sree, S. Chattopadhyay, K.-H. Ng, J.S. Suri, Automated diagnosis of epileptic EEG using entropies, Biomed. Signal Process. Control 7 (4) (2012) 401–408.
[26] A. Bhattacharyya, et al., Tunable-Q wavelet transform based multiscale entropy measure for automated classification of epileptic EEG signals, Appl. Sci. 7(4) (2017) 385.
[27] V. Gupta, R.B. Pachori, Epileptic seizure identification using entropy of FBSE based EEG rhythms, Biomed. Signal Process. Control 53 (2019) 101569.
[28] A. Anuragi, D.S. Sisodia, R.B. Pachori, Automated FBSE-EWT based learning framework for detection of epileptic seizures using time-segmented EEG signals, Comput. Biol. Med. 136 (2021) 104708.
[29] K. Samiee, S. Kiranyaz, M. Gabbouj, T. Saramäki, Long-term epileptic EEG classification via 2D mapping and textural features, Expert Syst. Appl. 42 (20) (2015) 7175–7185.
[30] M. Zabihi, S. Kiranyaz, A.B. Rad, A.K. Katsaggelos, M. Gabbouj, T. Ince, Analysis of high-dimensional phase space via Poincaré section for patient-specific seizure detection, IEEE Trans. Neural Syst. Rehabil. Eng. 24 (3) (2016) 386–398.
[31] G. Wang, D. Ren, K. Li, D. Wang, M. Wang, X. Yan, EEG-based detection of epileptic seizures through the use of a directed transfer function method, IEEE Access 6 (2018) 47189–47198.
[32] H. Rajaei, M. Adjouadi, M. Cabrerizo, P. Janwattanapong, A. Pinzon, S. Gonzales-Arias, A. Barreto, J. Andrian, N. Rishe, I. Yaylali, Dynamics and Distant Effects of Frontal/Temporal Epileptogenic Focus Using Functional Connectivity Maps, IEEE Trans. Biomed. Eng. 67 (2) (2020) 632–643.
[33] X. Wang, X. Wang, W. Liu, Z. Chang, T. Kärkkäinen, F. Cong, One dimensional convolutional neural networks for seizure onset detection using long-term scalp and intracranial EEG, Neurocomputing 459 (2021) 212–222.
[34] A. Zarei, B.M. Asl, Automatic seizure detection using orthogonal matching pursuit, discrete wavelet transform, and entropy based features of EEG signals, Comput. Biol. Med. 131 (2021) 104250.
[35] C. Li, W. Zhou, G. Liu, Y. Zhang, M. Geng, Z. Liu, S. Wang, W. Shang, Seizure Onset Detection Using Empirical Mode Decomposition and Common Spatial Pattern, IEEE Trans. Neural Syst. Rehabil. Eng. 29 (2021) 458–467.
[36] R.G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, C.E. Elger, Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state, Phys. Rev. E 64 (6) (2001) 061907.
[37] U. Herwig, P. Satrapi, C. Schönfeldt-Lecuona, Using the international 10–20 EEG system for positioning of transcranial magnetic stimulation, Brain Topogr. 16 (2) (2003) 95–99.
[38] A.L. Goldberger, L.A.N. Amaral, L. Glass, J.M. Hausdorff, P.C. Ivanov, R.G. Mark, J. E. Mietus, G.B. Moody, C.-K. Peng, H.E. Stanley, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals, Circulation 101 (23) (2000).
[39] M. Stephane, A wavelet tour of signal processing, Elsevier, 1999.
[40] M. Vetterli, C. Herley, Wavelets and filter banks: Theory and design, IEEE Trans. Sign. Process. 40(ARTICLE) (1992) 2207–2232.
[41] C.E. Shannon, A mathematical theory of communication, Bell Syst. Techn. J. 27 (3) (1948) 379–423.
[42] A. Ishikawa, H. Mieno, The fuzzy entropy concept and its application, Fuzzy Sets Syst. 2 (2) (1979) 113–123.
[44] Y. Kumar, M. Dewal, R. Anand, Epileptic seizure detection using DWT based fuzzy approximate entropy and support vector machine, Neurocomputing 133 (2014) 271–279.
[45] Y. Kaya, Ö.F. Ertuğrul, A stable feature extraction method in classification epileptic EEG signals, Australas. Phys. Eng. Sci. Med. 41 (3) (2018) 721–730.
[46] Y. Zhang, et al., Integration of 24 feature types to accurately detect and predict seizures using scalp EEG signals, Sensors 18 (5) (2018) 1372.
[47] M. Fan, C.-A. Chou, Detecting abnormal pattern of epileptic seizures via temporal synchronization of EEG signals, IEEE Trans. Biomed. Eng. 66 (3) (2018) 601–608.