



Selection of features for patient-independent detection of seizure events using scalp EEG signals

Shuhan Yang^a, Bo Li^d, Yinda Zhang^a, Meiyu Duan^a, Shuai Liu^a, Yexian Zhang^a, Xin Feng^{a,b}, Renbo Tan^c, Lan Huang^a, Fengfeng Zhou^{a,*}

^a BioKnow Health Informatics Lab, College of Computer Science and Technology, Key Laboratory of Symbolic Computation and Knowledge Engineering of Ministry of Education, Jilin University, Changchun, 130012, Jilin, China

^b Jilin Institute of Chemical Technology, Jilin, 132000, Jilin, China

^c Cancer Systems Biology Center, The China-Japan Union Hospital, Jilin University, Changchun, 130033, China

^d BioKnow Health Informatics Lab, College of Software, Key Laboratory of Symbolic Computation and Knowledge Engineering of Ministry of Education, Jilin University, Changchun, 130012, Jilin, China

ARTICLE INFO

Keywords:

EEG
Epileptic seizure detection
Time-domain feature extraction
MinMaxHist
Nonlinear features
Feature selection
XGBoost

ABSTRACT

Epilepsy involves brain abnormalities that may cause sudden seizures or other uncontrollable body activities. Epilepsy may have substantial impacts on the patient's quality of life, and its detection heavily relies on tedious and time-consuming manual curation by experienced clinicians, based on EEG signals. Most existing EEG-based seizure detection algorithms are patient-dependent and train a detection model for each patient. A new patient can only be monitored effectively after several episodes of epileptic seizures. This study investigates the patient-independent detection of seizure events using the open dataset CHB-MIT Scalp EEG. First, a novel feature extraction algorithm called MinMaxHist is proposed to measure the topological patterns of the EEG signals. Following this, MinMaxHist and several other feature extraction algorithms are applied to parameterize the EEG signals. Next, a comprehensive series of feature screening and classification optimization experiments are conducted, and finally, an optimized EEG-based seizure detection model is presented that can achieve overall values for accuracy, sensitivity, specificity, Matthews correlation coefficient, and Kappa of 0.8627, 0.8032, 0.9222, 0.7504 and 0.7254, respectively, with only 30 features. The classification accuracy of the method with MinMaxHist features was 0.0464 higher than that without MinMaxHist features. Compared with existing methods, the proposed algorithm achieved higher accuracy and sensitivity, as shown in the experimental results.

1. Introduction

Chronic lesion epilepsy induces transient brain dysfunction via the abnormal discharge of brain neurons [1,2]. Epilepsy is recognized as the second most common neurological disorder, ranked just after stroke [3]. More than 20% of epilepsy patients suffer from seizures, which are usually refractory to medication [4]. Epileptic seizures occur recurrently and unpredictably, and substantially impact the quality of life of the patient [5,6]. Various modern imaging technologies have been applied to diagnose epilepsy, including magnetic resonance imaging (MRI) [7,8], computed tomography (CT) [9], positron emission tomography (PET) [10,11] and electroencephalogram (EEG) [12].

EEG technology records electrophysiological signals to describe brain activity, and has been widely used to detect epilepsy [5,12].

Although EEG has characteristics of low cost, non-invasiveness, and high signal resolution [13,14], clinical annotation of EEG signals still relies heavily on human screening [15]. The visual inspection of EEG signals is recommended by guidelines, but is tedious and time-consuming [16].

Machine learning algorithms have become the research focus of EEG-based epileptic seizure detection problems [17–19], and several studies have investigated classification based on EEG signals [20,21]. Epileptic seizure detection is formulated as a classification problem of ictal and interictal EEG signals; various feature extraction algorithms have been utilized to generate features from a fixed-length window of EEG signals, including the time domain [22], frequency domain [23], and non-linear characteristics [24,25]. After the feature extraction step, a variety of feature selection [26,27] and classification algorithms [5,28] may be utilized to generate seizure detection models. Deep learning algorithms

* Corresponding author.

E-mail addresses: fzzhou@jlu.edu.cn, FengfengZhou@gmail.com (F. Zhou).

<https://doi.org/10.1016/j.complbiomed.2020.103671>

Received 2 January 2020; Received in revised form 20 February 2020; Accepted 20 February 2020

Available online 21 February 2020

0010-4825/© 2020 Elsevier Ltd. All rights reserved.

Table 1

EEG signal samples retrieved from each patient. Each row gives the total number of EEG signal samples for a given patient. Columns marked “NumSeizure” and “NumNonIctal” give the numbers of seizure and non-seizure regions recorded for this patient, while the columns marked “NumIctal” and “NumInterIctal” give the numbers of sampling regions retrieved from the EEG signals of this patient. The column marked “SamplePair” shows the number of pairs of one positive (ictal) and one negative (interictal) region retrieved as modeling data for this patient.

Patient	NumSeizure	NumIctal	NumNonIctal	NumInterIctal	SamplePair
chb01	7	70	34	204	70
chb02	3	27	33	198	27
chb03	7	62	31	186	62
chb04	4	62	39	234	62
chb05	5	92	34	204	92
chb06	10	22	11	66	22
chb07	3	53	16	96	53
chb08	5	151	15	90	90
chb09	4	44	15	90	44
chb10	7	70	18	108	70
chb11	3	133	32	192	133
chb13	10	68	4	24	24
chb14	8	24	19	114	24
chb15	20	323	25	150	150
chb17	3	48	17	102	48
chb18	6	51	29	174	51
chb19	3	38	26	156	38
chb20	8	45	22	132	45
chb21	4	32	29	174	32
chb22	3	33	28	168	33
chb23	7	67	6	36	36
chb24	16	79	9	54	54
Total	146	1594	492	2952	1260

such as CapsNet and a deep belief network may also be utilized to characterize the phenotypic patterns within EEG signals [29,30].

Most studies calculate variables from EEG signals using feature extraction algorithms, and formulate EEG-based seizure detection as a machine learning problem using these extracted features [19,31,32]. Gabor et al. trained a self-organizing map (SOM) neural network to detect seizures in 24 long-term EEG recordings using time-frequency features, and achieved an accuracy of 90%, thus demonstrating the possibility of automated seizure detection [33]. Time-frequency features are widely used to describe EEG signals due to their non-stationary characteristics. Various forms of wavelet transformations have been utilized to extract time-frequency features from EEG signals. Saab and Gotman used wavelet decomposition to extract features from scalp EEG signals and achieved a sensitivity of 76.0% for the detection of seizures [34]. Discrete wavelet transformation (DWT) is a powerful tool for analyzing time-frequency domain data, and can give a more flexible representation of signals. An extreme learning machine (ELM) was used to train a seizure detection model in Ref. [35]. In addition to time-frequency features, non-linear features can also be extracted from EEG signals. Kannathal et al. investigated the seizure detection problem by integrating four entropy estimators: Shannon spectral entropy, Renyi's entropy, Kolmogorov-Sinai entropy and approximate entropy. Their entropy-based model achieved a seizure classification accuracy of 90% [36]. Another study demonstrated that the non-linear features such as correlation dimension (CD), largest Lyapunov exponent (LLE), Hurst exponent (H), and entropy were accurate when used for seizure detection [37].

This study proposes a novel feature extraction algorithm in the time domain for the patient-independent seizure detection (pidSeizure) problem, and demonstrates that the integrated modeling of various feature sources can achieve satisfactory detection performance. Although much progress has been achieved in the epileptic seizure detection problem, most prior studies have focused on training a separate detection model for each patient. In this study, we hypothesize that a pidSeizure model might better serve epileptic seizure patients. Our experimental data also suggest that the extracted features can be further selected for better classification performance. The final detection model utilizes both feature extraction and feature selection to deliver accurate seizure detection performance.

2. Materials and methods

2.1. Problem setting

This study investigated the patient-independent detection problem of epileptic seizures using scalp EEG signals, due to the following two challenges identified in existing studies.

Previous studies have usually used patient-specific modeling [21,24,38–40], i.e., they have trained a separate model for each patient. Although this approach can achieve high detection accuracy, clinicians must wait for several epileptic seizure events before a new patient can be monitored. In addition, the rich array of information from other patients is not fully utilized in optimizing these patient-specific models.

Many existing studies have screened intracranial EEG signals for patterns representing epileptic seizures [21,38–41]. An intracranial EEG (iEEG) signal has a high signal/noise ratio, which facilitates the generation of highly accurate seizure detection models [42]. However, an iEEG monitors electrophysiological signals by placing sensing electrodes directly on the exposed surface of the brain, making it difficult to deploy outside of a clinical environment. Scalp EEG (sEEG) is a non-invasive technology for detecting electrophysiological signals from the surface of the head, and has a much higher noise level compared with iEEG [42–44]. The detection of seizures from sEEG signals therefore remains a challenge.

2.2. Description of dataset

The CHB-MIT Scalp EEG database from PhysioNet was used in this study [45]. This pediatric scalp EEG dataset was generated at Boston Children's Hospital after the withdrawal of anti-seizure medications. A detailed description of these samples can be found on the PhysioNet website.

All pediatric patients were monitored at a sampling rate of 256 Hz with 16-bit quantization. Since most of the recordings used 23 EEG channels, this study only selected samples with 23 channels. These channels were placed following the International Federation of Clinical Neurophysiology 10–20 placement system, and the recordings were analyzed accordingly, (i.e., FP1–F7, F7–T7, T7–P7, P7–O1, FP1–F3, F3–C3, C3–P3, P3–O1, FZ–CZ, CZ–PZ, FP2–F4, F4–C4, C4–P4, P4–O2,

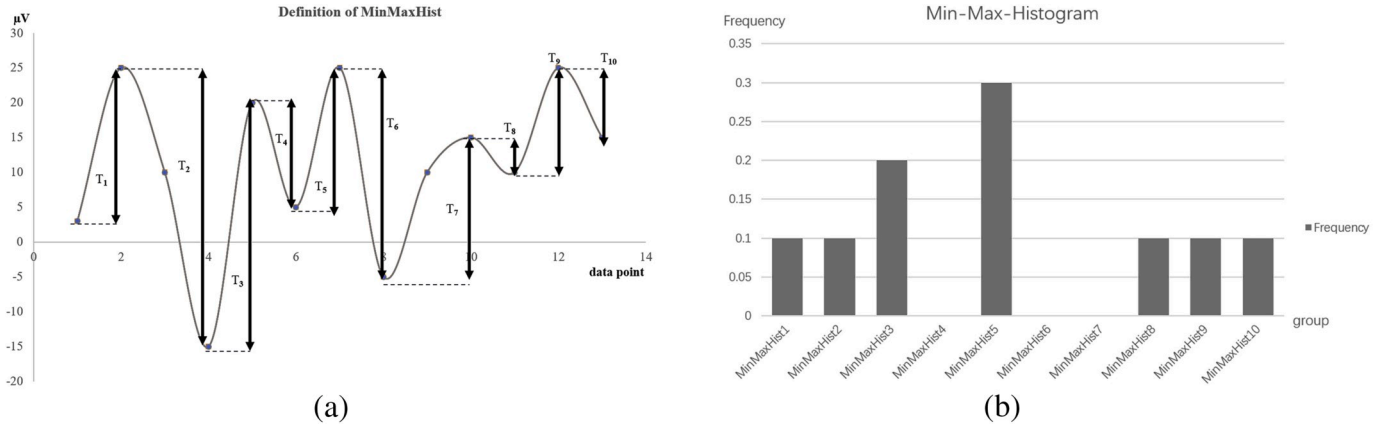


Fig. 1. Illustration of the MinMaxHist algorithm. (a) Schematic line chart of an EEG signal, where T_i ($i = 1, 2, \dots, n$) represents the absolute value of the difference between a pair of neighboring peak and valley values. (b) Min-Max-Histogram, where the maximum and minimum values in the array T represent the left and right limits of this histogram. The number of features returned by this algorithm is WinNum.

FP2-F8, F8-T8, T8-P8, P8-O2, P7-T7, T7-FT9, FT9-FT10, FT10-T8, and T8-P8, as in Ref. [46]). The last channel, T8-P8, was removed, since there were two of these channels, meaning that each EEG signal used in this study contained 22 channels. The interictal regions of the record for patient chb12 were very short (average value 415 s), as were the ictal regions of the record for patient chb16 (average value 8.6 s), and these two records were also excluded from further analysis.

2.3. Preprocessing of scalp EEG signals

The EEG recordings for these patients were of various lengths and showed different epileptic seizure times, so a preprocessing step was conducted to ensure that each sample contained the same number of features. First, a zero-phase 4th-order Butterworth filter was used to process the 22 channels of each sample. The band filter was applied between 0.5 and 32 Hz, since the ictal EEG signals were observed to have significant abnormalities within this frequency range [34,47]. Following this, 6-s [38] non-overlapping sliding windows were applied to the epileptic seizure regions, which were taken as positive samples. There were 1594 positive samples in total, retrieved from the 22 patients, as shown in Table 1. The interictal regions were usually much longer than the ictal ones, and to ensure a balance between positive and negative samples, six non-overlapping windows with 6-s lengths were randomly retrieved as negative samples from the interictal regions, excluding the initial 10-min windows. Interictal regions shorter than 10 min were not used for retrieving negative samples. Each sample contained 22 channels, giving $6 \text{ s} \times 256 \text{ Hz} \times 22 \text{ channels} = 33,792$ voltage values.

2.4. Feature extraction: Min-Max histogram (MinMaxHist)

We observed that ictal EEG signals contained more fluctuations than interictal regions, and therefore proposed our novel feature extraction algorithm, the Min-Max Histogram (MinMaxHist). Spikes and sharp waveforms were repeatedly observed in the EEG signals at seizure onset [48]. The purpose of the proposed MinMaxHist algorithm was to quantitatively describe the waveform characteristics of these spikes and sharp waves by describing the scales of change in the EEG signals in the feature extraction step and to distinguish seizure events from interictal events based on these types of features.

The aim of the MinMaxHist algorithm is to create a histogram describing a given EEG signal, as shown in Fig. 1. First, we constructed a one-dimensional array $T = \langle T_1, T_2, \dots, T_n \rangle$, where T_i was the absolute value of the difference between a pair of neighboring peak and valley values, as shown in Fig. 1(a). Then, Max_T and Min_T were defined as the maximum, and minimum values in array T , representing the right- and left-hand limits of the lateral axis in the histogram, respectively. We then

set the parameter $\text{WinNum} = 10$, to describe the number of histogram groups, in order to ensure that the new data distribution would not be too scattered or too centralized. Next, we calculated the bin size as $(\text{Max}_T - \text{Min}_T) / \text{WinNum}$ to represent the width of each interval in the histogram. As shown in Fig. 1(b), the lateral axis was evenly divided into 10 bins, with ranges $[\text{Min}_T, \text{Min}_T + \text{bin}]$, $[\text{Min}_T + \text{bin}, \text{Min}_T + 2 \times \text{bin}]$, ..., $[\text{Min}_T + 9 \times \text{bin}, \text{Max}_T]$. The values for the vertical axis were the data frequency of array T in each interval. The detailed algorithm was formulated as in Fig. 1. Hence, 10 MinMaxHist features were extracted from a given EEG signal, and were denoted as MinMaxHist1 , MinMaxHist2 , ..., MinMaxHist10 .

2.5. Other time-domain features

Several time-domain features were summarized from the raw EEG signals. The minimum, maximum, average, standard deviation and variance of a given window of EEG signal X , where $X = \langle X_1, X_2, \dots, X_N \rangle$, were extracted as the features minX , maxX , meanX , stdX and varX .

Another measurement totalVariance was defined as $\frac{\sum_{i=1}^N |X_{i+1} - X_i|}{(\text{maxX} - \text{minX}) \times (N-1)}$.

The skewness of a data distribution X was defined as the expected value of $E \left[\left(\frac{X - \text{meanX}}{\text{stdX}} \right)^3 \right]$ [49]. This feature was abbreviated as Skewness.

Kurtosis is a shape descriptor of a probability distribution (abbreviated as Kurs) and was defined as the expected value of $E \left[\left(\frac{X - \text{meanX}}{\text{stdX}} \right)^4 \right]$ [46].

The magnitude of a time-series signal X can be measured using the root mean square value [50], i.e.,

$$\text{RMS}(X) = \sqrt{\frac{\sum_{i=1}^N X_i^2}{N}} \quad (1)$$

The largest peak value of X was defined as $\text{Peak}(X) = \max(|\text{maxX}|, |\text{minX}|)$, since the voltages of the EEG signal X may be positive or negative values. The peak-to-average power ratio (PARS) was also defined as $\text{PARS}(X) = [\text{Peak}(X) / \text{RMS}(X)]^2$ [51].

2.6. Nonlinear features of EEG signals

EEG signals have also been described by nonlinear features in many existing studies, since they are inherently non-stationary, low-energy and transient [52]. Nonlinear descriptors can robustly capture the hidden patterns within such noisy signals [53,54].

Entropy is a representative descriptor of the complexity of a time-series variable, and increases with an increase in this complexity [55].

The nonlinear dynamic analysis wavelet entropy (abbreviated as WaveletEntropy) was used to describe the EEG signals in this study, and the Daubechies wavelet (db4) was used to decompose the EEG signals into six frequency sub-bands [56–58].

Richman et al. proposed a similar measurement to describe the data regularity of the EEG signals, abbreviated as SampleEntropy [59]. The definition and parameter choices for SampleEntropy may be found in Refs. [60,61].

Permutation entropy has demonstrated a very good anti-noise capability and a low computational complexity [62,63], and was therefore used in this study (abbreviated as PermutEntropy) to describe the complexity of the EEG signals.

The Hurst Exponent (abbreviated as HurstExp) was proposed in 1951 to quantize the long-term correlations in a time-series variable [64]. The HurstExp feature is usually calculated by the R/S method [65,66], and a value in the range [0.5, 1] demonstrates the existence of long-range dependence.

Non-static and non-stationary time-series variables can also be described by fractal patterns, e.g., using detrended fluctuation analysis (DFA) [67] and the Petrosian fractal dimension (PFD) [68]. DFA can robustly describe the fractal characteristics and self-similarity in a non-stationary time-series variable [67], while the fractal dimension describes the self-similarity within a time-series variable and demonstrates a better temporal resolution than Fourier analysis [68–70]. The Petrosian version of fractal dimension (PFD) was utilized in this study, as it has shorter time requirements [69,70].

The Hjorth parameters were proposed by Bo Hjorth in 1970 to describe EEG signals [71], and we utilized the Hjorth mobility and Hjorth complexity to detect seizures. The Hjorth mobility is defined as

$$HjorthMobility(X(t)) = \sqrt{\text{var}(dX(t) / dt) / \text{var}(X(t))} \quad (2)$$

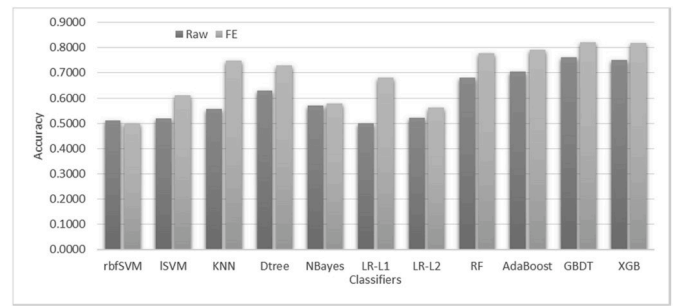
where $\text{var}(X(t))$ is the variance of the variable $X(t)$, while the Hjorth complexity is defined as

$$HjorthComplexity(X(t)) = HjorthMobility(dX(t) / dt) / HjorthMobility(X(t)) \quad (3)$$

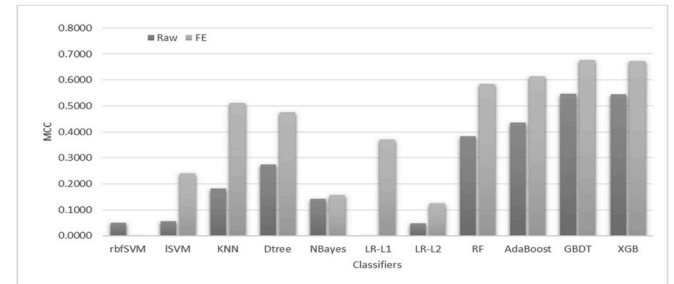
2.7. Feature selection algorithms

Various feature selection algorithms were utilized to find a feature subset that outperformed the list of 638 original features extracted from the EEG signals. Filters are the simplest form of feature selection and usually have the lowest time complexity [72]. The student's t-test (Ttest) is a widely-used statistical test for evaluating the difference in a feature between two groups of samples [5,73,74]. The statistical p-value of a Ttest evaluates the null hypothesis in which the distributions of a given feature in two groups of features are identical to each other. Only features with $p \leq 0.05$ are selected for further analysis. The rank sum test (Utest) evaluates whether a feature was collected from two populations with the same distribution, and does not assume that this feature follows a normal distribution [75,76]. The same p-value threshold of 0.05 is applied to select features using a Utest. In this study, we ranked the features in descending order based on their accuracy, calculated using a 5-fold cross-validation strategy and two classifiers: a support vector machine (SVM) with a linear kernel [77] and the extreme gradient boosting algorithm (XGB) [78]. The top 50% of the features were kept for further analysis. These two filter feature selection strategies were denoted as Acc-SVM and Acc-XGB, respectively.

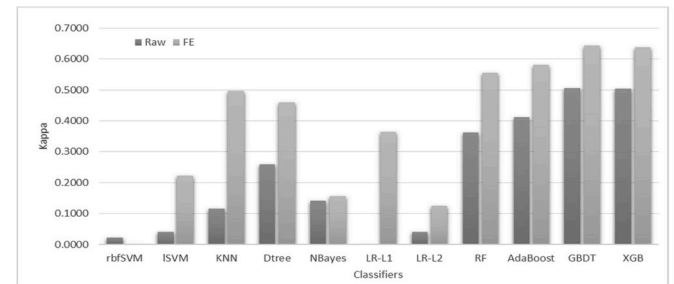
Recursive feature elimination (RFE) is a popular feature selection framework [79] that iteratively trains a classification model over a feature set using the embedded classifier and removes the feature with the lowest feature coefficient returned by the trained model. The feature subset with the best 5-fold cross-validation accuracy is returned for further analysis. If multiple feature subsets achieve the same best accuracy, the one with the fewest features is returned. Any classifier may



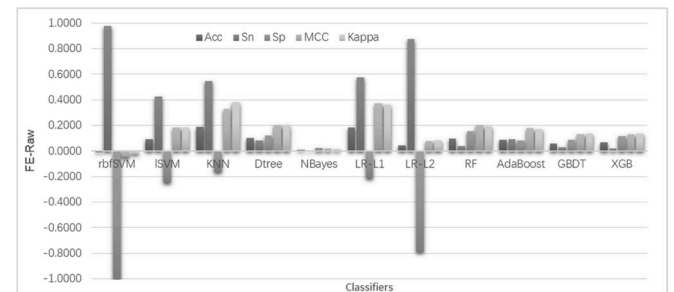
(a)



(b)



(c)



(d)

Fig. 2. Seizure detection performance of 10 classifiers using raw EEG data (Raw) and extracted features (FE). Values were calculated using 5-fold cross-validation for (a) accuracy, (b) MCC and (c) Kappa for each classifier on the two datasets (Raw and FE). (d) Difference in detection performance between the FE and raw data. Five performance measurements (Acc, Sn, Sp, MCC and Kappa) were evaluated.

be embedded in RFE as long as it returns the feature coefficients after the classification model is trained. The most popular embedded classifier is an SVM [80].

Another feature elimination framework BackFS has also been utilized to remove features in an exhaustive way [5,81]. BackFS does not assume that the coefficients from the trained model are positively

Table 2

Summary of the features extracted from the EEG signals. Two families of features were used in this study, from the time domain and the nonlinear domain. The columns marked “Abbreviation” and “Description” give abbreviations and basic descriptions of these extracted features, while the column marked “Number” gives the number of features extracted for each row.

Family	Abbreviation	Description	Number
Time domain	MinMaxHist1-MinMaxHist10	Waveform	10
	maxX, minX, meanX, stdX and varX	Maximum, minimum, mean, standard deviation and variance of a given variable X	5
	totalVariation, Skewness, Kurs, RMS, and peak	Total variation, skewness, kurtosis, root mean square, and peak	5
	PARS	peak-to-average power ratio	1
Non-linear domain	WaveletEntropy, Sample Entropy and PermutEntropy	Wavelet, sample and permutation entropy measurements	3
	HurstExp	Hurst exponent	1
	DFA	Detrended fluctuation analysis	1
	PFD	Petrosian fractal dimension	1
	HjorthMobility and HjorthComplexity	Hjorth parameters: mobility and complexity	2

correlated with the importance of each feature. In each round of the iteration, BackFS removes a feature if this achieves the smallest decrease in the accuracy of the embedded classifier. Any classifier can be embedded, and BackFS is used as a wrapper framework.

2.8. Classification algorithms

Ten popular classifiers were utilized in this study. SVM is a fast and effective classifier in many cases of binary classification problems, and calculates a hyperplane to maximize the margins between two groups of samples within the hyperspace, using a given kernel function [77]. In

many supervised learning algorithms. Another popular boosting algorithm is the gradient-boosted decision tree (GBDT), which can be used to solve both classification and regression problems [90–92]. Chen and Guestrin recently proposed a scalable tree boosting system (XGB), which has a very fast running speed and is a scalable end-to-end boosting system [78].

2.9. Performance measurements

In this study, we used a 5-fold cross-validation strategy to evaluate how well each model performed for a binary classification problem, using the measures of accuracy, sensitivity, and specificity. Let P and N be the numbers of positive and negative samples, respectively. The performance measurement accuracy (Acc) is defined as the percentage of correctly detected samples, i.e., $\text{Acc} = (\text{TP} + \text{TN}) / (P + N)$, where TP (true positive) and TN (true negative) are the numbers of correctly detected positive and negative samples, respectively.

Sensitivity (S_n) and specificity (S_p) are defined as the ratios of correctly detected positive and negative samples, respectively. Hence, sensitivity $S_n = \text{TP} / P$, and specificity $S_p = \text{TN} / N$. The numbers of incorrectly detected positive and negative samples were defined as FP (false positive) and FN (false negative), respectively.

The Matthews correlation coefficient (MCC) and Kappa are also used to evaluate the performance of our approach. The range of MCC is between zero and one, with one representing the best classification result. MCC is defined as

$$\text{MCC} = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP) \times (TP + FN) \times (TN + FP) \times (TN + FN)}} \quad (4)$$

The Kappa metric (also called Cohen’s kappa) is a coefficient for the qualitative (categorical) samples [93,94] that takes values in the range [0, 1]. Kappa is defined as

$$\text{Kappa} = (\text{Acc} - p_e) / (1 - p_e) \quad (5)$$

$$p_e = [(\text{TP} \times \text{FP}) + (\text{TP} \times \text{FN}) + (\text{TN} \times \text{FP}) + (\text{TN} \times \text{FN})] / (\text{TP} + \text{TN} + \text{FP} + \text{FN})^2 \quad (6)$$

this study, we utilized two popular kernel functions for SVM, the Gaussian and linear kernels. These two models are abbreviated as rbfSVM and lSVM, respectively.

Two simple classifiers were evaluated, the k-nearest neighbors (KNN) and naive Bayes (NBayes) approaches. KNN utilizes the essence of a classification problem (i.e., the intra-class similarity) and assigns a class label to a sample based on majority voting by the sample’s k nearest neighbors [82], where the default value of k is 5. NBayes assumes inter-feature independence and calculates the likelihood of a sample being in each class, based on Bayes’ theorem [83]. This study used a Gaussian kernel NBayes approach. A logistic regression (LR) binary classifier was evaluated in this study; both L1 and L2-regularizations were used as loss functions to avoid overfitting, and are denoted as LR-L1 and LR-L2, respectively [84].

Tree-based learning algorithms form another popular family of classifiers. The classification and regression tree (CART) is a widely used classifier with low model complexity and fast running speed [85]. A CART-based decision tree classifier was evaluated, referred to here as the DTree algorithm.

Boosting is an integration strategy that can create a strong classifier by summarizing the results of multiple weak classifiers [86]. Adaptive boosting (AdaBoost) was proposed by Freund and Schapire [87] in 1997 to generate a boosting classifier in an adaptive manner [88,89]. The AdaBoost algorithm is a meta-algorithm and can be integrated with

3. Results and discussion

3.1. Extraction of features from raw EEG data

Equal numbers of positive and negative samples were collected from each patient for further analysis, as shown in Table 1. Hence, the number of sample pairs collected from each patient was the minimum of the numbers of positive and negative samples for this patient. A total of 1260 pairs of samples were collected from all patients.

Two families of features were extracted from each channel of a given EEG sample, i.e., the time domain and the nonlinear domain, as shown in Table 2. We calculated 21 time-domain features from each 6 s EEG sample. The MinMaxHist feature extraction algorithm was used to generate 10 features (MinMaxHist1 to MinMaxHist10) from a given EEG sample. Five basic statistics metrics were also calculated: the maximum (maxX), minimum (minX), mean (meanX), standard deviation (stdX) and variance (varX) for this sample, giving a one-dimensional vector with $6 \text{ s} \times 256 \text{ Hz} = 1536$ data points. Five further time-domain metrics were also defined to describe the EEG signals: the total variance

Accuracy	FE	Ttest	Utest	Acc-SVM	Acc-XGB	AvgC
rbfSVM	0.5000	0.5180	0.5000	0.5000	0.5000	0.5036
ISVM	0.6110	0.7430	0.6090	0.5972	0.6294	0.6379
KNN	0.7480	0.7390	0.7480	0.7476	0.7520	0.7469
Dtree	0.7300	0.7460	0.7250	0.7440	0.7532	0.7396
NBayes	0.5790	0.5770	0.5790	0.5790	0.5782	0.5784
LR-L1	0.6820	0.7140	0.6370	0.6873	0.6996	0.6840
LR-L2	0.5620	0.5750	0.5620	0.5611	0.5623	0.5645
RF	0.7770	0.7950	0.7910	0.7857	0.7571	0.7812
AdaBoost	0.7910	0.7850	0.7880	0.7948	0.8004	0.7918
GBDT	0.8210	0.8200	0.8200	0.8171	0.8194	0.8195
XGB	0.8190	0.8210	0.8200	0.8206	0.8194	0.8200
AvgFS	0.6927	0.7121	0.6890	0.6940	0.6974	

(a)

MCC	FE	Ttest	Utest	Acc-SVM	Acc-XGB	AvgC
rbfSVM	0	0.1124	0	0	0	0.0225
ISVM	0.2417	0.5244	0.2435	0.2072	0.2719	0.2977
KNN	0.5119	0.4991	0.5119	0.5104	0.5191	0.5105
Dtree	0.4762	0.5083	0.4641	0.5032	0.5242	0.4952
NBayes	0.158	0.1533	0.158	0.1586	0.1573	0.1570
LR-L1	0.3705	0.4349	0.2783	0.3837	0.418	0.3771
LR-L2	0.1261	0.1498	0.126	0.124	0.1265	0.1305
RF	0.5849	0.6266	0.6198	0.5968	0.5486	0.5953
AdaBoost	0.6138	0.6056	0.6052	0.6244	0.6332	0.6164
GBDT	0.6761	0.677	0.6716	0.6671	0.6739	0.6731
XGB	0.6723	0.6779	0.6741	0.6744	0.6732	0.6744
AvgFS	0.4029	0.4518	0.3957	0.4045	0.4133	

(b)

Kappa	FE	Ttest	Utest	Acc-SVM	Acc-XGB	AvgC
rbfSVM	0	0.0365	0	0	0	0.0073
ISVM	0.2222	0.4857	0.2183	0.1944	0.2587	0.2759
KNN	0.4968	0.4778	0.4968	0.4952	0.504	0.4941
Dtree	0.4603	0.4921	0.4508	0.4881	0.5063	0.4795
NBayes	0.1571	0.1532	0.1571	0.1579	0.1563	0.1563
LR-L1	0.3635	0.4278	0.2746	0.3746	0.3992	0.3679
LR-L2	0.1246	0.1492	0.1246	0.1222	0.1246	0.1290
RF	0.5548	0.5905	0.5825	0.5714	0.5143	0.5627
AdaBoost	0.5817	0.5706	0.5754	0.5897	0.6008	0.5836
GBDT	0.6429	0.6405	0.6397	0.6341	0.6389	0.6392
XGB	0.6381	0.6413	0.6405	0.6413	0.6389	0.6400
AvgFS	0.3856	0.4241	0.3782	0.3881	0.3947	

(c)

Fig. 3. Heatmap of the performance of 10 classifiers on five feature sets. The FE feature set was the original list of all 638 features extracted from an EEG sample. The other four feature sets were selected using four filters: Ttest, Utest, Acc-SVM and Acc-XGB. The (a) accuracy, (b) MCC and (c) Kappa were calculated using each of the 10 classifiers with a 5-fold cross-validation strategy. The column marked AvgC gives the average value for each classifier, while the column marked AvgFS gives the average value for each feature set.

(totalVariance), skewness (Skewness), kurtosis (Kurs), root mean square (RMS) and peak (Peak). The peak-to-average power ratio (PARS) was also calculated as a feature.

We also calculated eight nonlinear features to describe the EEG signals, as shown in Table 2. Three types of entropy were calculated: the wavelet entropy (WaveletEntropy), sample entropy (SampleEntropy), and permutation entropy (PermutEntropy). Three other widely used nonlinear features were also calculated: the Hurst exponent (HurstExp), detrended fluctuation analysis (DFA), and Petrosian fractal dimension (PFD). The Hjorth parameters of mobility (HjorthMobility) and complexity (HjorthComplexity) were also used due to their discriminative capabilities.

3.2. Detection of seizures using raw EEG data and extracted features

Overall, the extracted features achieved better seizure detection performance than the raw EEG data, as shown in Fig. 2. The best values of Acc, MCC, and Kappa for seizure detection were achieved by the classifiers on the extracted features (dataset FE), as shown in Fig. 2(a–c). The GBDT and XGB boosting classifiers were the best performing. GBDT achieved the best values of Acc (0.8210) and MCC (0.6761), while the second-best classifier achieved values of 0.8190 for Acc and 0.6723 for MCC. The third best classifier, AdaBoost, achieved values of around 0.80 for Acc and 0.6138 for MCC. The other classifiers did not achieve values of Acc larger than 0.80 or MCC larger than 0.6. These results demonstrate that the boosting classifiers gave good performance for this problem.

We also demonstrated that this step achieved better detection performance by utilizing the features extracted from the EEG signals. Fig. 2 (d) displays the results of dataset FE minus the results of raw EEG data,

in order to show the differences between the two sets of results. The experimental results showed that except for rbfSVM, the classifiers achieved better seizure detection performance, with average improved values using the extracted features of 0.0918, 0.1819 and 0.1859 for Acc, MCC and Kappa, respectively, compared to using the raw EEG data. The largest improvement in Acc (0.1900) was achieved by the KNN classifier. The rbfSVM classifier performed in the same way as a random classifier, with a value of Acc of around 0.5000 using both raw EEG data and the extracted features, meaning that the small improvement in accuracy for rbfSVM between the dataset FE and raw EEG may be simply due to chance.

It was therefore essential to extract features from the EEG signals in the time and nonlinear domains in order to achieve better seizure detection performance.

The DFA and Hurst exponent features describe the general patterns in seizure EEG signals [95,96], and have been widely used to detect epileptic seizures based on EEG data [96,97]. In this study, we evaluated the classification performance of these two feature types alone and compared it with all feature types. The XGB classifier achieved the best classification performance (Acc = 0.7901, MCC = 0.6057, Kappa = 0.5802) using these two feature types, i.e., DFA and HurstExp. When all types of features were used, the XGB classifier again achieved the best classification model, with Acc = 0.8190, MCC = 0.6723 and Kappa = 0.6381. Both the experimental data and the literature suggest that these two general features (DFA and HurstExp) can contribute to the problem of epileptic seizure detection using EEG signals.

3.3. Finding seizure-associated features using filters

Due to the inherently noisy nature of the data, it is always

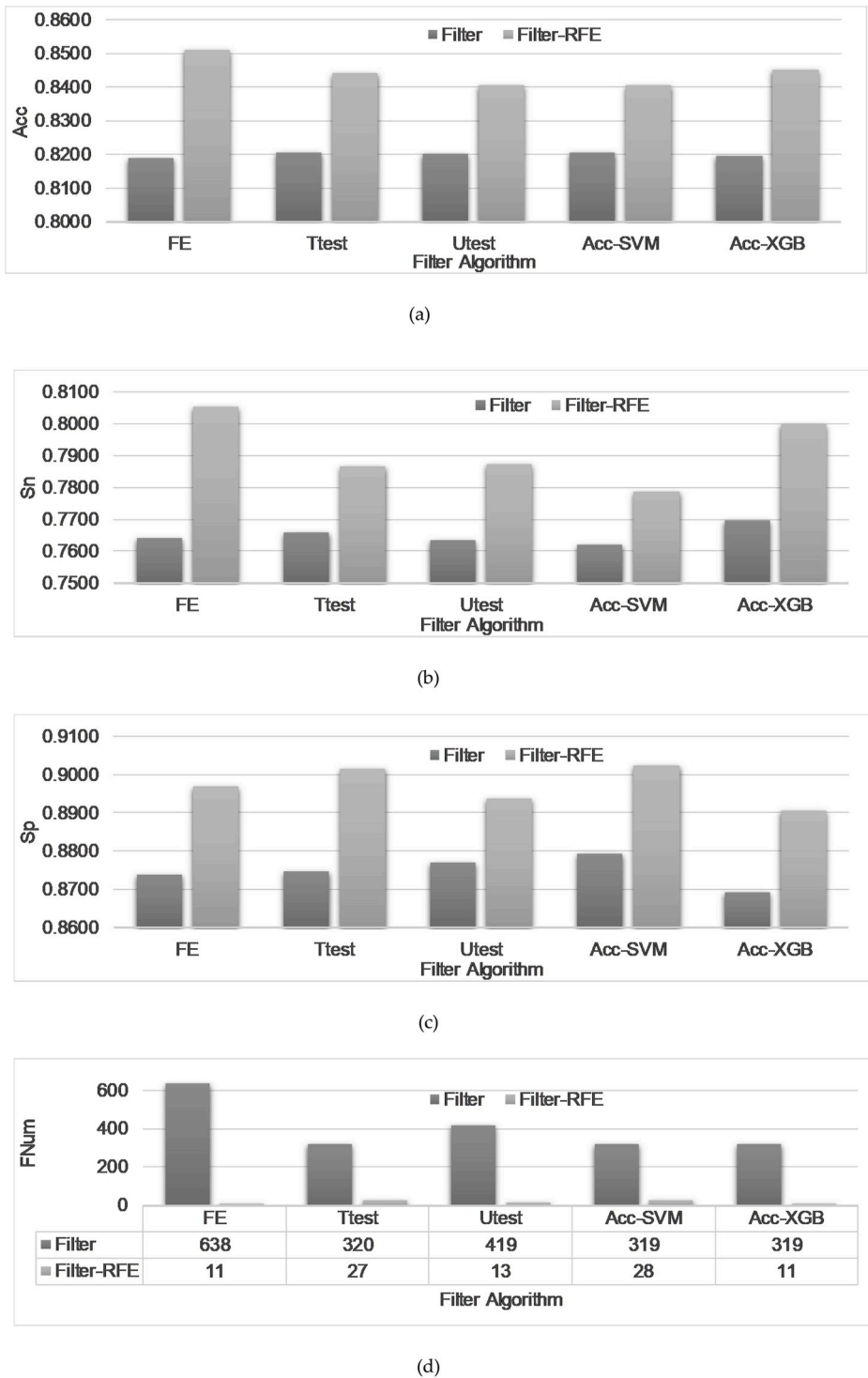


Fig. 4. Comparison of the detection performance of filters and filter-embedded RFE algorithms. The FE feature set contained a list of all 638 features extracted from the EEG signals. The other four feature sets were first selected using four filters: Ttest, Utest, Acc-SVM and Acc-XGB. The column marked Filter gives the seizure detection performance for the original and filter-selected features, while the column marked FilterRFE gives the detection performance for features that have been further screened using a filter-embedded RFE strategy. A filter was embedded for the FE-based RFE process, and the detection performance was evaluated based on (a) Acc, (b) Sn and (c) Sp, calculated using 5-fold cross-validation and the best classifier (XGB) identified in the previous section. The numbers of features are shown in (d).

challenging to accurately identify phenotype-associated patterns in EEG signals [43,44,52]. Some EEG-extracted features may be removed in order to refine EEG-based detection models for various biomedical issues, e.g., treatment responses for depressive disorders [98].

First, we evaluated four filter algorithms for finding seizure-associated features. Each sample contained 22 channels, and 29 features were calculated for each of these channels, meaning that the original FE feature set for each sample in this study had $22 \times 29 = 638$ features. The Ttest and Utest filters were used to select 320 and 419 features, respectively, with $p \leq 0.05$, while Acc-SVM and Acc-XGB were used to select 319 features. All four filters assumed that the features

were independent of each other and the features were ranked based on the discriminative power of each feature individually [57,72,99]. The features were scaled to a [0, 1] normal distribution.

Fig. 3 demonstrates that the use of these four simple filters can improve the seizure detection model for nine of the 10 classifiers. The largest improvement of 0.1320 in the value of accuracy was achieved by the LSVM model with the Ttest-selected features. No improvement was achieved by the four filter algorithms for the NBayes classifier. The GBDT boosting classifier needed all 638 features to achieve its best classification accuracy. It was also interesting to observe that the three top-performing models were trained with the ensemble classifiers, i.e.,

Acc	FE	Ttest	U test	Acc-SVM	Acc-XGB
SVM -RFE	0.8008	0.7849	0.7214	0.7766	0.7579
LR-L1-RFE	0.7960	0.7937	0.7202	0.7790	0.7556
LR-L2-RFE	0.7885	0.7956	0.7206	0.7778	0.7651
AdaBoost-RFE	0.8417	0.8298	0.8242	0.8313	0.8341
GBDT-RFE	0.8512	0.8365	0.8361	0.8361	0.8401
XGB-RFE	0.8512	0.8440	0.8405	0.8405	0.8452

(a)

Sn	FE	Ttest	U test	Acc-SVM	Acc-XGB
SVM -RFE	0.7579	0.7563	0.6421	0.7341	0.6873
LR-L1-RFE	0.7746	0.7524	0.6429	0.7278	0.6929
LR-L2-RFE	0.7524	0.7627	0.6444	0.7286	0.7190
AdaBoost-RFE	0.8111	0.7984	0.7960	0.8016	0.7960
GBDT-RFE	0.7992	0.7730	0.7865	0.7786	0.7889
XGB-RFE	0.8056	0.7865	0.7873	0.7786	0.8000

(b)

Sp	FE	Ttest	U test	Acc-SVM	Acc-XGB
SVM -RFE	0.8437	0.8135	0.8008	0.8190	0.8286
LR-L1-RFE	0.8175	0.8349	0.7976	0.8302	0.8183
LR-L2-RFE	0.8246	0.8286	0.7968	0.8270	0.8111
AdaBoost-RFE	0.8722	0.8611	0.8524	0.8611	0.8722
GBDT-RFE	0.9032	0.9000	0.8857	0.8937	0.8913
XGB-RFE	0.8968	0.9016	0.8937	0.9024	0.8905

(c)

FN um	FE	Ttest	U test	Acc-SVM	Acc-XGB
SVM -RFE	93	84	122	63	189
LR-L1-RFE	25	56	47	52	43
LR-L2-RFE	181	79	67	61	25
AdaBoost-RFE	35	36	22	17	34
GBDT-RFE	19	40	39	57	27
XGB-RFE	11	27	13	28	11

(d)

Fig. 5. Evaluation of the RFE strategy using various embedded classifiers. The detection performance was evaluated based on (a) Acc, (b) Sn, (c) Sp and (d) FNum (feature number). The six embedded classifiers were LSVM, LR-L1, LR-L2, AdaBoost, GBDT and XGB.

XGB (Acc = 0.8210, MCC = 0.6779 and Kappa = 0.6413), GBDT (Acc = 0.8210, MCC = 0.6761 and Kappa = 0.6429) and AdaBoost (Acc = 0.8004, MCC = 0.6332 and Kappa = 0.6008). Although both XGB and GBDT achieved the best value of 0.8210 for Acc, XGB (320 features) achieved a better MCC of 0.6779 and used only about half of the features compared with GBDT (638 features). The principle of Occam's Razor suggests that a simple model with a similar detection performance should be preferred over a complex one [100,101]. Hence, the experimental data suggest that a feature selection step can be utilized to further improve the seizure detection model based on the features extracted from an EEG.

3.4. Finding a better feature list using filter-embedded RFEs

This section evaluates how the features can be refined by filter-embedded RFE strategies, as shown in Fig. 4. The best classifier identified in the previous section, XGB, was then utilized to evaluate a given feature subset. The RFE strategy eliminates one feature per iteration of

Table 3

List of the 30 features selected by BackFS. The columns marked "No." and "Channel" give the IDs and channels for the features, while the feature names are given in the column marked "FeatureName".

No.	Channel	FeatureName	No.	Channel	FeatureName
1	P7_O1	MinMaxHist8	16	C3_P3	MinMaxHist2
2	F4_C4	MinMaxHist2	17	T7_FT9	MinMaxHist7
3	T8_P8	MinMaxHist4	18	P7_O1	MinMaxHist2
4	P7_O1	MinMaxHist3	19	P7_T7	MinMaxHist3
5	FT10_T8	MinMaxHist2	20	T7_FT9	MinMaxHist2
6	FT9_FT10	MinMaxHist4	21	T8_P8	RMS
7	FP1_F7	MinMaxHist5	22	T7_FT9	HjorthMobility
8	F7_T7	MinMaxHist2	23	T7_P7	HjorthMobility
9	T7_P7	MinMaxHist3	24	FT9_FT10	MinMaxHist5
10	FP2_F4	MinMaxHist1	25	T7_P7	meanX
11	FP2_F8	MinMaxHist2	26	T7_FT9	PFD
12	C4_P4	MinMaxHist1	27	P3_O1	DFA
13	P8_O2	MinMaxHist2	28	FP2_F4	PermutEntropy
14	F3_C3	MinMaxHist2	29	FT9_FT10	totalVariation
15	P8_O2	MinMaxHist7	30	FT9_FT10	RMS

model training.

FE-RFE achieved the best seizure detection accuracy (Acc = 0.8512), as shown in Fig. 4(a). The RFE strategy improved the results for all four filter algorithms, with a minimum improvement in Acc of 0.0198. It was also interesting to observe that the detection performance Acc/Sn/Sp was improved by the RFE strategy in all cases, as shown in Fig. 4(b) and (c). The RFE strategy also reduced the number of features by at least 91%, as shown in Fig. 4(d), thus demonstrating its power to improve features extracted from an EEG for better classification performance.

3.5. Selection of features using classifier-embedded RFEs

Six classifiers were evaluated in terms of their performance as the embedded classifier in the RFE framework, as shown in Fig. 5. The three classifiers LSVM, LR-L1 and LR-L2 performed the worst, with an average accuracy of only about 0.7690, while the other three classifiers achieved values of at least 0.8300. XGB and GBDT achieved the best detection accuracy of 0.8512 on the FE features. A larger difference between Sn and Sp was seen for GBDT (Sn = 0.7992 and Sp = 0.9032), while XGB achieved values of 0.8056 and 0.8968 for Sn and Sp, respectively. XGB was identified as the best embedded classifier, since it used only 11 features to achieve the same seizure detection accuracy as GBDT, which used 19.

3.6. Removal of redundant features using BackFS

The union of the 11 detected features and MinMaxHist features was refined using the BackFS strategy, as shown in Table 3. Thirty features were chosen to achieve the overall best values of Acc, MCC and Kappa of 0.8627, 0.7504 and 0.7254, respectively, using the XGB classifier as shown in Fig. 6. The values of Acc, MCC and Kappa for MinMaxHist features were 0.0464, 0.084 and 0.0929 higher, respectively, than for classification without the MinMaxHist features. The final results for Acc, MCC and Kappa showed that the proposed method achieved a more accurate classification. From Fig. 6, it can be seen that the best results were obtained when XGB was used as the classifier for the input feature matrix. The second best classifier, GBDT, was also an ensemble classifier. This observation suggests that ensemble classifiers may be good choices for modeling EEG-based classification problems, as shown in Fig. 6. The NBayes classifier achieved imbalanced values of Sn = 0.1540 and Sp = 0.9381, suggesting that the features extracted from the EEG did not facilitate the calculation of the Bayesian probabilities.

The final list of 30 features consisted of 21 MinMaxHist features and nine other features, as shown in Table 3. The nine other features were shared with the best subset of 11 features selected in the previous section. The RMS for channel T7_P7 and PermutEntropy for channel P7_O1

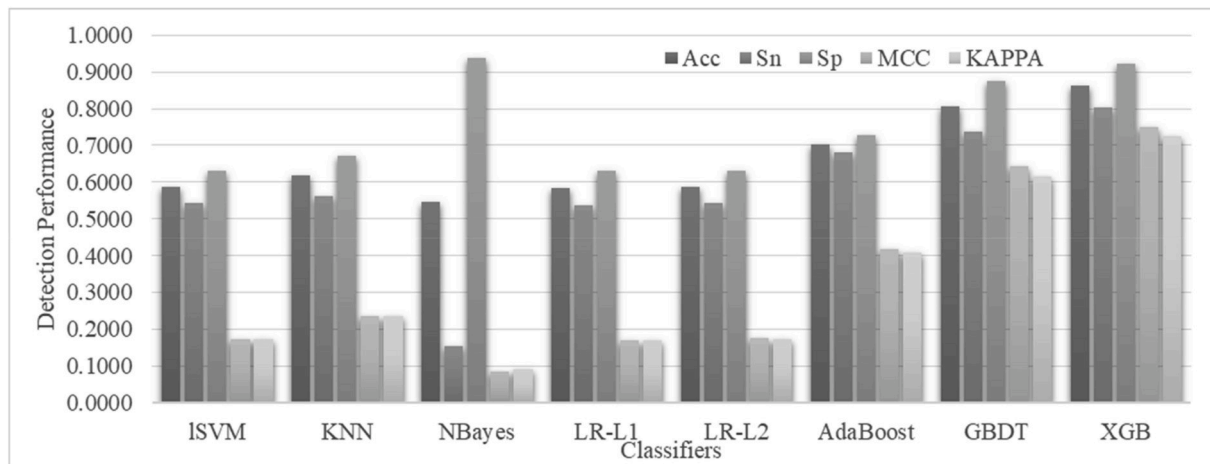


Fig. 6. Seizure detection performance of eight classifiers on the 30 features selected by BackFS. The horizontal axis lists the classifiers, while the vertical axis shows the detection performance measurements.

were not selected. This demonstrated that the MinMaxHist features extracted from the EEG signals exhibited the seizure-specific EEG patterns, and had potential applications in the EEG-based detection of other diseases.

3.7. Performance of patient-specific seizure detection

Many existing EEG-based seizure detection studies have focused on patient-specific models [38], in which a separate model is trained for each patient using EEG data from that patient. This approach is expected to give a higher seizure detection accuracy than the patient-independent model investigated in this study. We therefore carried out a comparative experiment to investigate whether the optimized feature selection procedure developed in this study can help to improve the classic patient-specific approach. Fig. 7 illustrates that the procedure developed in this study improved the accuracy of detection in the patient-specific approach, with an average improvement of 0.0233 in the value for Acc, as shown in Fig. 7(a). The values of Sn, Sp, MCC and Kappa were improved by 0.0346, 0.0141, 0.0509 and 0.0546, respectively, as shown in Fig. 7(c–f). A total of 638 features were extracted from the EEG signals, and our feature selection procedure selected at most six of these for patient-specific seizure detection models, for all patients except chb11. Our procedure selected 51 features, and achieved a seizure detection accuracy that was only 0.0071 lower than that of the “Raw” model, as shown in Fig. 7(b). The average values of Acc, Sn, Sp, MCC and Kappa were 0.9828, 0.9893, 0.9764, 0.9674 and 0.9655, respectively. The value of Sn (0.9893) obtained in our study was higher than the value of Sn (0.96) reported in the prior study [38].

3.8. Performance of semi-patient-specific seizure detection

Haoqu and Gotman proposed a semi-patient-specific approach in their study, with results that were better than non-specific but worse than patient-specific models [102]. Their experiment used semi-patient-specific information including a seizure from one patient and interictal EEG signals from many other patients to train a classifier for the seizure patient. The classifier was trained in this way for each patient. We used the algorithm proposed in this work to carry out a semi-patient-specific experiment and obtained a value of 0.8849 for the accuracy of detection for the onset of epileptic seizures. This result was higher than the non-patient-specific model and lower than the patient-specific model. This conclusion was consistent with that of Haoqu and Gotman, and our method achieved a higher accuracy. The non-specific method used EEG signals from all patients to train each classifier. Although this method did not perform as well as the other two

specific methods, it has a wider range of applications.

3.9. Comparison with existing studies

Various machine learning algorithms for automatic seizure detection have been proposed by other researchers. We carried out a comprehensive comparison between the model proposed in this study and existing models, as shown in Table 4.

The open dataset CHB-MIT has been widely used to evaluate seizure detection algorithms. Vidyaratne et al. [38] used the feature extraction algorithm harmonic wavelet packet transform (HWPT) and a classifier relevance vector machine (RVM) to achieve a seizure detection sensitivity of 0.96. Raghu et al. [103] used DWT-based sigmoid entropy in the time and frequency domains to extract features, and applied the SVM classifier, which achieved a sensitivity of 0.9421. Kaleem et al. [104] detected seizures based on signal-derived empirical mode decomposition (EMD)-based dictionary approach, and obtained average values for accuracy, sensitivity and specificity of 0.929, 0.943 and 0.915, respectively. Xiang et al. [105] used fuzzy entropy to extract features and achieved average values for accuracy, sensitivity and specificity of 0.9831, 0.9827 and 0.9836, respectively. In comparison, a model trained with the proposed method for each patient can achieve average values of accuracy, sensitivity and specificity of 0.9828, 0.9893 and 0.9764, respectively. These experimental results suggest that the proposed method is superior. Although the accuracy reported by Bhattacharyya et al. [106] was slightly better than in our study, their work was patient-dependent, and cannot be easily applied to new patients.

From a comparison with patient-independent studies that used the CHB-MIT dataset, we found that our proposed method achieved better results, as shown in Table 4. Our method required on 6 s to analyze and process the raw data, whereas in Ref. [46], the reported sampling time was 60 s. Although the values for the sensitivity and specificity of these authors' results were 0.93 and 0.94, the SMOTE over-sampling algorithm was used for data processing, and the sampling time was almost 10 times as long as ours. Before the SMOTE method was used, the sensitivity and specificity of the experimental results reported in Ref. [46] were both 0.88. In comparison, if the sampling time was extended to 60 s in our method, the values of sensitivity and specificity would reach 0.8896 and 0.9310. In Ref. [107], the author used a neural network classifier to analyze the dataset, and obtained a sensitivity value of 0.85; however, the sampling time reported in this paper was 30 s. With an equivalent sampling time, the sensitivity and specificity of our algorithm reached 0.8967 and 0.9393, respectively, i.e. our method gives superior results.

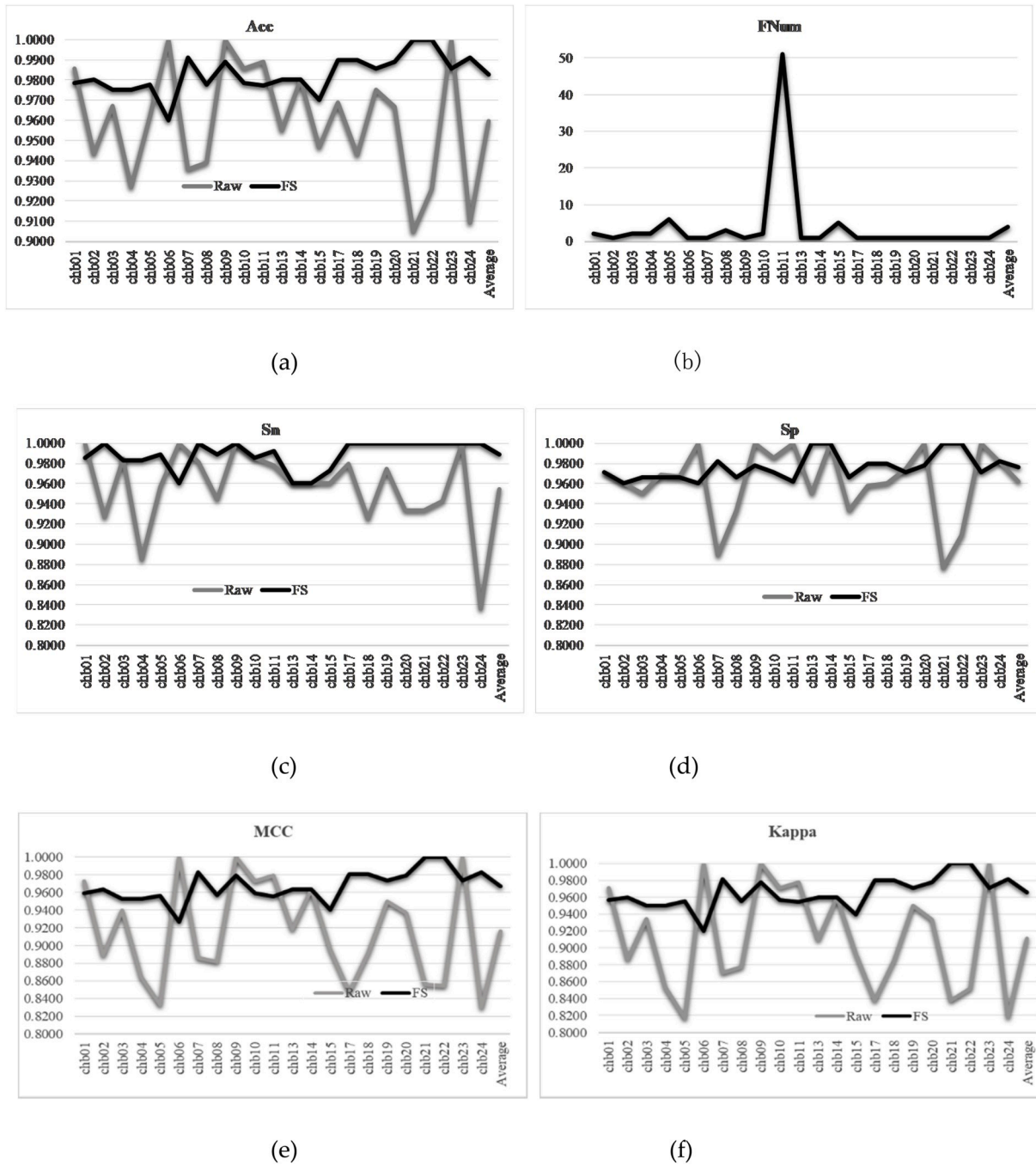


Fig. 7. Improving the classic patient-specific seizure detection model. An optimized feature selection procedure was applied to the classic model of patient-specific seizure detection. The detection performance was measured based on (a) Acc, (b) FNum (number of features), (c) Sn, (d) Sp, (e) MCC and (f) Kappa. The data series “Raw” refers to patient-specific seizure detection using all the extracted features, while the data series “FS” refers to detection models with features selected using the feature selection procedure optimized in the previous sections.

3.10. Evaluation of the proposed method on a new corpus

We also evaluated the proposed method on an independent dataset. The THU EEG corpus was released by the American Clinical Neurophysiology Society, and contains many types of EEG signals for use in various studies [108]. We selected the THU EEG Seizure corpus, a subset of the THU EEG Corpus with files that are known to contain seizure events. There were 929 EEG files in this corpus, 217 of which were samples of epileptic seizures. The EEG signals in this corpus were collected from 43 real epileptic patients, and had not been processed. The corpus contained male and female patients of various ages, for

whom the causes of the disease were different. There are both eye movements and artifacts in the EEG signals, and we can therefore treat this dataset as a real clinical dataset. We evaluated our method on this corpus, and achieved values for seizure prediction of Acc, Sn, and Sp of 0.8021, 0.80, and 0.8088, respectively. Although the performance of our method in this case was slightly lower than that for the CHB-MIT dataset, this corpus contained more patients and more causes of epilepsy, the patients had a larger age range, and the EEG signals contained more noise and artifacts. This independent corpus was therefore similar to a realistic clinical dataset, and the proposed method demonstrated a reasonable seizure prediction performance.

Table 4

Comparison with other existing studies. The detection performance was measured based on Acc, Sn, Sp, MCC and Kappa. “-” indicates that the value is not mentioned in the literature. The column marked “Category” gives the type of approach used in each paper.

Study	Category	Acc	Sn	Sp	MCC	Kappa
Vidyaratne et al.	patient-dependent	-	0.9600	-	-	-
Raghu et al.	patient-dependent	-	0.9421	-	-	0.94
Kaleem et al.	patient-dependent	0.9291	0.9427	0.9155	-	-
Xiang et al.	patient-dependent	0.9831	0.9827	0.9836	-	-
Bhattacharyya et al.	patient-dependent	0.9941	0.9791	0.9957	-	-
Fergus et al.	patient-independent	-	0.8800	0.8800	-	-
Thodoroff et al.	patient-independent	-	0.8500	-	-	-
Current study	patient-dependent	0.9828	0.9893	0.9764	0.9674	0.9656
Current study (6 s)	patient-independent	0.8627	0.8032	0.9222	0.7504	0.7254
Current study (30 s)	patient-independent	0.9180	0.8967	0.9393	0.8428	0.8356
Current study (60 s)	patient-independent	0.9103	0.8896	0.9310	0.8275	0.8207

4. Conclusions

Epileptic seizure is a nervous system disorder that does not directly cause lethal damage to patients; however, seizures can substantially impact the quality of life of patients, and can give rise to major mental pressures. The accurate detection of seizures may provide vital information to help both patients and their family members to maintain their daily routines and quality of life.

This study focused on the patient-independent seizure detection problem, which is more challenging to solve than the classic patient-specific seizure detection problem. Many existing studies have focused on patient-specific seizure detection, which requires dozens of seizures to be recorded for a specific patient before a seizure detection model for that patient can be well trained. Our patient-independent seizure detection model can be easily deployed for a given patient without requiring recordings of previous seizures.

By integrating the proposed feature extraction algorithm (Min-MaxHist) with several existing algorithms, we generated a seizure detection model that achieved classification values of 0.8627, 0.7504 and 0.7254 for Acc, MCC and Kappa, respectively. This model also used only 30 features, and we did not consider other possible features. It is possible that the integration of more feature extraction methods will improve the experimental results, and further combinations of features may also be tried in the future. The optimized feature selection procedure in this study was also shown to improve the classic patient-specific seizure detection model.

Data availability

The dataset used in this experiment was downloaded on May 4, 2018, and can be obtained from <https://www.physionet.org/content/chbmit/1.0.0/>.

Declaration of competing interest

None Declared.

Acknowledgements

This work was supported by the Jilin Provincial Key Laboratory of Big Data Intelligent Computing (20180622002JC), the Education Department of Jilin Province (JJKH20180145KJ), and the startup grant of the Jilin University. This work was also partially supported by the Bioknow MedAI Institute (BMCP-2018-001), the High Performance Computing Center of Jilin University, and by the Fundamental Research Funds for the Central Universities, JLU. The insightful comments from the three anonymous reviewers were much appreciated.

References

- [1] A.V. Bumanglag, R.S. Sloviter, No latency to dentate granule cell epileptogenesis in experimental temporal lobe epilepsy with hippocampal sclerosis, *Epilepsia* 59 (2018) 2019–2034.
- [2] C. Tang, Y. Gu, H. Wang, H. Wu, Y. Wang, Y. Meng, Z. Han, Y. Gu, W. Ma, Z. Jiang, Y. Song, M. Na, D. Lu, Z. Lin, Targeting of microRNA-21-5p protects against seizure damage in a kainic acid-induced status epilepticus model via PTEN-mTOR, *Epilepsy Res.* 144 (2018) 34–42.
- [3] J. Corsini, L. Shoker, S. Sanei, G. Alarcon, Epileptic seizure predictability from scalp EEG incorporating constrained blind source separation, *IEEE (Inst. Electr. Electron. Eng.) Trans. Biomed. Eng.* 53 (2006) 790–799.
- [4] A. Shueb, H. Edwards, J. Connolly, B. Bourgeois, S.T. Treves, J. Guttag, Patient-specific seizure onset detection, *Epilepsy Behav.* : E&B 5 (2004) 483–498.
- [5] Y. Zhang, S. Yang, Y. Liu, Y. Zhang, B. Han, F. Zhou, Integration of 24 feature types to accurately detect and predict seizures using scalp EEG signals, *Sensors* 18 (2018).
- [6] G.L. Birbeck, R.D. Hays, X. Cui, B.G. Vickrey, Seizure reduction and quality of life improvements in people with epilepsy, *Epilepsia* 43 (2002) 535–538.
- [7] X. Zhao, Z. Zhou, W. Zhu, H. Xiang, Role of conventional magnetic resonance imaging in the screening of epilepsy with structural abnormalities: a pictorial essay, *Am. J. Nucl. Med. Mol. Imaging* 7 (2017) 126–137.
- [8] G.P. Winston, The role of magnetic resonance imaging techniques in the diagnosis, surgical treatment and biological understanding of epilepsy, *Quant. Imag. Med. Surg.* 5 (2015) 186–187.
- [9] P. Fu, L. Wei, F. Zhang, J. Gao, J. Jing, W. Wu, H. Liu, Added value of NeuroGam software analysis in single photon emission computed tomography localization diagnosis of epilepsy in interictal stage, *Med. Sci. Mon. Int. Med. J. Exp. Clin. Res.* 24 (2018) 1494–1501.
- [10] W.T. Kerr, S.T. Nguyen, A.Y. Cho, E.P. Lau, D.H. Silverman, P.K. Douglas, N. M. Reddy, A. Anderson, J. Braman, N. Salamon, J.M. Stern, M.S. Cohen, Computer-aided diagnosis and localization of lateralized temporal lobe epilepsy using interictal FDG-PET, *Front. Neurol.* 4 (2013) 31.
- [11] A. Kumar, F. Semah, H.T. Chugani, W.H. Theodore, Epilepsy diagnosis: positron emission tomography, *Handb. Clin. Neurol.* 107 (2012) 409–424.
- [12] C. Minardi, R. Minacapelli, P. Valastro, F. Vasile, S. Pitino, P. Pavone, M. Astuto, P. Murabito, Epilepsy in children: from diagnosis to treatment with focus on emergency, *J. Clin. Med.* 8 (2019).
- [13] J. Nizard, J.P. Lefaucheur, M. Helbert, E. de Chauvigny, J.P. Nguyen, Non-invasive stimulation therapies for the treatment of refractory pain, *Discov. Med.* 14 (2012) 21–31.
- [14] L. Yang, C. Wilke, B. Brinkmann, G.A. Worrell, B. He, Dynamic imaging of ictal oscillations using non-invasive high-resolution EEG, *Neuroimage* 56 (2011) 1908–1917.
- [15] S. Koelstra, I. Patras, Fusion of facial expressions and EEG for implicit affective tagging, *Image Vis Comput.* 31 (2013) 164–174.
- [16] S.R. Sinha, L. Sullivan, D. Sabau, D. San-Juan, K.E. Dombrowski, J.J. Halford, A. J. Hani, F.W. Drislane, M.M. Stecker, American clinical neurophysiology society guideline 1: minimum technical requirements for performing clinical electroencephalography, *J. Clin. Neurophysiol.* 33 (2016) 303–307.
- [17] A. Emami, N. Kunii, T. Matsuo, T. Shinokaki, K. Kawai, H. Takahashi, Autoencoding of long-term scalp electroencephalogram to detect epileptic seizure for diagnosis support system, *Comput. Biol. Med.* 110 (2019) 227–233.
- [18] N.S. Frolov, V.V. Grubov, V.A. Maksimenko, A. Luttjohann, V.V. Makarov, A. N. Pavlov, E. Sitnikova, A.N. Pisarchik, J. Kurths, A.E. Hramov, Statistical properties and predictability of extreme epileptic events, *Sci. Rep.* 9 (2019) 7243.
- [19] R. San-Segundo, M. Gil-Martin, L.F. D'Haro-Enriquez, J.M. Pardo, Classification of epileptic EEG recordings using signal transforms and convolutional neural networks, *Comput. Biol. Med.* 109 (2019) 148–158.
- [20] X. Fan, N. Gaspard, B. Legros, F. Lucchetti, R. Ercek, A. Nonclercq, Automated epileptic seizure detection based on break of excitation/inhibition balance, *Comput. Biol. Med.* 107 (2019) 30–38.
- [21] M. Zhou, C. Tian, R. Cao, B. Wang, Y. Niu, T. Hu, H. Guo, J. Xiang, Epileptic seizure detection based on EEG signals and CNN, *Front. Neuroinf.* 12 (2018) 95.
- [22] E.M. Ventouras, E.A. Monoyiou, P.Y. Ktonas, T. Paparrigopoulos, D.G. Dikeos, N. K. Uzunoglu, C.R. Soldatos, Sleep spindle detection using artificial neural

- networks trained with filtered time-domain EEG: a feasibility study, *Comput. Methods Progr. Biomed.* 78 (2005) 191–207.
- [23] N. Gursel Ozmen, L. Gumusel, Y. Yang, A biologically inspired approach to frequency domain feature extraction for EEG classification, *Comput. Math Methods Med.* 2018 (2018) 9890132.
- [24] J. Birjandtalab, M. Baran Pouyan, D. Cogan, M. Nourani, J. Harvey, Automated seizure detection using limited-channel EEG and non-linear dimension reduction, *Comput. Biol. Med.* 82 (2017) 49–58.
- [25] S. Geng, W. Zhou, Q. Yuan, D. Cai, Y. Zeng, EEG non-linear feature extraction using correlation dimension and Hurst exponent, *Neurol. Res.* 33 (2011) 908–912.
- [26] A.H. Ansari, P.J. Cherian, A. Caicedo, G. Naulaers, M. De Vos, S. Van Huffel, Neonatal seizure detection using deep convolutional neural networks, *Int. J. Neural Syst.* 29 (2019) 1850011.
- [27] Y. Li, W. Cui, M. Luo, K. Li, L. Wang, Epileptic seizure detection based on time-frequency images of EEG signals using Gaussian mixture model and gray level Co-occurrence matrix features, *Int. J. Neural Syst.* 28 (2018) 1850003.
- [28] Q. Yuan, W. Zhou, L. Zhang, F. Zhang, F. Xu, Y. Leng, D. Wei, M. Chen, Epileptic seizure detection based on imbalanced classification and wavelet packet transform, *Seizure* 50 (2017) 99–108.
- [29] H. Chao, L. Dong, Y. Liu, B. Lu, Emotion recognition from multiband EEG signals using CapsNet, *Sensors* 19 (2019).
- [30] H. Chao, H. Zhi, L. Dong, Y. Liu, Recognition of emotions using multichannel EEG data and DBN-GC-based ensemble deep learning framework, *Comput. Intell. Neurosci.* 2018 (2018) 9750904.
- [31] A. Sharmila, P. Geethanjali, A review on the pattern detection methods for epilepsy seizure detection from EEG signals, *Biomed. Tech.* (2019).
- [32] N. Padfield, J. Zabalza, H. Zhao, V. Masero, J. Ren, EEG-based brain-computer interfaces using motor-imagery: techniques and challenges, *Sensors* 19 (2019).
- [33] A.J. Gabor, R.R. Leach, F.U. Dowl, Automated seizure detection using a self-organizing neural network, *Electroencephalogr. Clin. Neurophysiol.* 99 (1996) 257–266.
- [34] M.E. Saab, J. Gotman, A system to detect the onset of epileptic seizures in scalp EEG, *Clinical neurophysiology, Off. J. Int. Feder. Clin. Neurophysiol.* 116 (2005) 427–442.
- [35] J.L. Song, W.F. Hu, R. Zhang, Automated detection of epileptic EEGs using a novel fusion feature and extreme learning machine, *Neurocomputing* 175 (2016) 383–391.
- [36] N. Kannathal, M.L. Choo, U.R. Acharya, P.K. Sadasivan, Entropies for detection of epilepsy in EEG, *Comput. Methods Progr. Biomed.* 80 (2005) 187–194.
- [37] N. Kannathal, U.R. Acharya, C.M. Lim, P.K. Sadasivan, Characterization of EEG—a comparative study, *Comput. Methods Progr. Biomed.* 80 (2005) 17–23.
- [38] L.S. Vidyaratne, K.M. Iftikharuddin, Real-time epileptic seizure detection using EEG, *IEEE Trans. Neural Syst. Rehabil. Eng.* 25 (2017) 2146–2156, a publication of the IEEE Engineering in Medicine and Biology Society.
- [39] S.N. Baldassano, B.H. Brinkmann, H. Ung, T. Blevins, E.C. Conrad, K. Leyde, M. J. Cook, A.N. Khambhati, J.B. Wagenaar, G.A. Worrell, B. Litt, Crowdsourcing seizure detection: algorithm development and validation on human implanted device recordings, *Brain* 140 (2017) 1680–1691.
- [40] G.R. Minasyan, J.B. Chatten, M.J. Chatten, R.N. Harner, Patient-specific early seizure detection from scalp electroencephalogram, *J. Clin. Neurophysiol.* 27 (2010) 163–178, official publication of the American Electroencephalographic Society.
- [41] R. Hussein, H. Palangi, R. Ward, Z. Wang, Epileptic Seizure Detection: A Deep Learning Approach, 2018.
- [42] H. Stefan, S. Rampp, R.C. Knowlton, Magnetoencephalography adds to the surgical evaluation process, *Epilepsy Behav.* : E&B 20 (2011) 172–177.
- [43] P. Holler, E. Trinka, Y. Holler, High-frequency oscillations in the scalp electroencephalogram: mission impossible without computational intelligence, *Comput. Intell. Neurosci.* 2018 (2018) 1638097.
- [44] C. Yeon, D. Kim, K. Kim, E. Chung, Visual evoked potential recordings in mice using a dry non-invasive multi-channel scalp EEG sensor, *J Vis Exp* (2018).
- [45] A.L. Goldberger, L.A. Amaral, L. Glass, J.M. Hausdorff, P.C. Ivanov, R.G. Mark, J. E. Mietus, G.B. Moody, C.K. Peng, H.E. Stanley, PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals, *Circulation* 101 (2000) E215–E220.
- [46] P. Fergus, D. Hignett, A. Hussain, D. Al-Jumeily, K. Abdel-Aziz, Automatic epileptic seizure detection using scalp EEG and advanced artificial intelligence techniques, *BioMed Res. Int.* 2015 (2015) 986736.
- [47] J. Gotman, Automatic recognition of epileptic seizures in the EEG, *Electroencephalogr. Clin. Neurophysiol.* 54 (1982) 530–540.
- [48] W.T. Blume, G.B. Young, J.F. Lemieux, EEG morphology of partial epileptic seizures, *Electroencephalogr. Clin. Neurophysiol.* 57 (1984) 295–302.
- [49] I. Ginzburg, Prediction of the moments in advection-diffusion lattice Boltzmann method. I. Truncation dispersion, skewness, and kurtosis, *Phys. Rev. E* 95 (2017), 013304.
- [50] J.A. Blanco, A.C. Vanleer, T.K. Calibo, S.L. Firebaugh, Single-trial cognitive stress classification using portable wireless electroencephalography, *Sensors* 19 (2019).
- [51] J. Bai, Y. Li, Y. Yi, W. Cheng, H. Du, PAPR reduction based on tone reservation scheme for DCO-OFDM indoor visible light communications, *Optic Express* 25 (2017) 24630–24638.
- [52] A. Gramfort, D. Strohmeier, J. Hauelsen, M.S. Hamalainen, M. Kowalski, Time-frequency mixed-norm estimates: sparse M/EEG imaging with non-stationary source activations, *Neuroimage* 70 (2013) 410–422.
- [53] U.R. Acharya, O. Faust, N. Kannathal, T. Chua, S. Laxminarayan, Non-linear analysis of EEG signals at various sleep stages, *Comput. Methods Progr. Biomed.* 80 (2005) 37–45.
- [54] L. Gupta, A.M. Upadhye, Non-linear alignment of neural net outputs for partial shape classification, *Pattern Recogn.* 24 (1991) 943–948.
- [55] Y. Chen, Y. Li, R. Narayan, A. Subramanian, X. Xie, Gene expression inference with deep learning, *Bioinformatics* 32 (2016) 1832–1839.
- [56] H. Sharma, R. Raj, M. Juneja, EEG signal based classification before and after combined Yoga and Sudarshan Kriya, *Neurosci. Lett.* 707 (2019) 134300.
- [57] F. Han, B. Liu, J. Zhu, B. Zhang, Algorithm design for edge detection of high-speed moving target image under noisy environment, *Sensors* 19 (2019).
- [58] T. Zoughi, R. Boostani, M. Deypir, A wavelet-based estimating depth of anesthesia, *Eng. Appl. Artif. Intell.* 25 (2012) 1710–1722.
- [59] J.S. Richman, J.R. Moorman, Physiological time-series analysis using approximate entropy and sample entropy, *Am. J. Physiol. Heart Circ. Physiol.* 278 (2000) H2039–H2049.
- [60] Y. Song, P. Li[†], A new approach for epileptic seizure detection: sample entropy based feature extraction and extreme learning machine, 03, *J. Biomed. Sci. Eng.* (2010) 12, 06.
- [61] U.R. Acharya, S.V. Sree, A.P. Alvin, R. Yanti, J.S. Suri, Application of non-linear and wavelet based features for the automated identification of epileptic EEG signals, *Int. J. Neural Syst.* 22 (2012) 1250002.
- [62] X. Chen, N.-D. Jin, A. Zhao, Z.-K. Gao, L.-S. Zhai, B. Sun, The experimental signals analysis for bubbly oil-in-water flow using multi-scale weighted-permutation entropy, *Phys. Stat. Mech. Appl.* 417 (2015) 230–244.
- [63] A.A. Bruzzo, B. Gesierich, M. Santi, C.A. Tassinari, N. Birbaumer, G. Rubboli, Permutation entropy to detect vigilance changes and preictal states from scalp EEG in epileptic patients. A preliminary study, *Neurol. Sci.* 29 (2008) 3–9, official journal of the Italian Neurological Society and of the Italian Society of Clinical Neurophysiology.
- [64] H.E. Hurst, Long-term storage capacity of reservoirs, *Trans. Am. Soc. Civ. Eng.* 116 (1951) 770–799.
- [65] B.A. Carreras, B.P. van Milligen, M.A. Pedrosa, R. Balbin, C. Hidalgo, D. E. Newman, E. Sanchez, M. Frances, I. Garcia-Cortes, J. Bleuel, M. Endler, C. Riccardi, S. Davies, G.F. Matthews, E. Martinez, V. Antoni, A. Latten, T. Klinger, Self-similarity of the plasma edge fluctuations (vol 5, pg 3632, 1998), *Phys. Plasmas* 10 (2003), 4591–4591.
- [66] M. Nurujjaman, R. Narayanan, A.S. Iyengar, Comparative study of nonlinear properties of EEG signals of normal persons and epileptic patients, *Nonlinear Biomed. Phys.* 3 (2009) 6.
- [67] K. Ivanova, M. Ausloos, Application of the detrended fluctuation analysis (DFA) method for describing cloud breaking, *Physica A* 274 (1999) 349–354.
- [68] T. Higuchi, Approach to an irregular time series on the basis of the fractal theory, *Phys. Nonlinear Phenom.* 31 (1988) 277–283.
- [69] A. Petrosian, Kolmogorov complexity of finite sequences and recognition of different preictal EEG patterns, in: *Proceedings Eighth IEEE Symposium on Computer-Based Medical Systems*, IEEE, 1995, pp. 212–217.
- [70] R. Esteller, G. Vachtsevanos, J. Echaz, B. Litt, A comparison of waveform fractal dimension algorithms, *IEEE Trans. Circuits Syst. I: Fund. Theor. Appl.* 48 (2001) 177–183.
- [71] B. Hjorth, EEG analysis based on time domain properties, *Electroencephalogr. Clin. Neurophysiol.* 29 (1970) 306–310.
- [72] M. Alirezanejad, R. Enayatifar, H. Motameni, H. Nematzadeh, Heuristic filter feature selection methods for medical datasets, *Genomics* (2019).
- [73] A. Asif, M. Majid, S.M. Anwar, Human stress classification using EEG signals in response to music tracks, *Comput. Biol. Med.* 107 (2019) 182–196.
- [74] A. Ahmadi, S. Davoudi, M.R. Daliri, Computer Aided Diagnosis System for multiple sclerosis disease based on phase to amplitude coupling in covert visual attention, *Comput. Methods Progr. Biomed.* 169 (2019) 9–18.
- [75] S. Khoshnoud, M.A. Nazari, M. Shamsi, Functional brain dynamic analysis of ADHD and control children using nonlinear dynamical features of EEG signals, *J. Integr. Neurosci.* 17 (2018) 11–17.
- [76] D. Endres, S. Maier, B. Feige, N.A. Posielski, K. Nickel, D. Ebert, A. Riedel, A. Philipsen, E. Perlov, L. Tebartz van Elst, Altered intermittent rhythmic delta and theta activity in the electroencephalographies of high functioning adult patients with autism spectrum disorder, *Front. Hum. Neurosci.* 11 (2017) 66.
- [77] M.-W. Huang, C.-W. Chen, W.-C. Lin, S.-W. Ke, C.-F. Tsai, SVM and SVM ensembles in breast cancer prediction, *PLoS One* 12 (2017) e0161501–e0161501.
- [78] T. Chen, C. Guestrin, XGBoost: a scalable tree boosting system, in: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ACM, San Francisco, California, USA, 2016, pp. 785–794.
- [79] I. Guyon, J. Weston, S. Barnhill, V. Vapnik, Gene selection for cancer classification using support vector machines, *Mach. Learn.* 46 (2002) 389–422.
- [80] S. Sahrani, D. Albashish, A. Abdullah, N.A. Shukor, S. Hayati Md Pauzi, Absolute cosine-based SVM-RFE feature selection method for prostate histopathological grading, *Artif. Intell. Med.* 87 (2018) 78–90.
- [81] X. Feng, R. Zhang, M. Liu, Q. Liu, F. Li, Z. Yan, F. Zhou, An accurate regression of developmental stages for breast cancer based on transcriptomic biomarkers, *Biomarkers Med.* 13 (2019) 5–15.
- [82] S. Zhang, X. Li, M. Zong, X. Zhu, R. Wang, Efficient kNN classification with different numbers of nearest neighbors, *IEEE Trans. Neural Network. Learn. Syst.* 29 (2018) 1774–1785.
- [83] J. Hu, Automated detection of driver fatigue based on AdaBoost classifier with EEG signals, *Front. Comput. Neurosci.* 11 (2017) 72.
- [84] S.K. Palei, S.K. Das, Logistic regression model for prediction of roof fall risks in bord and pillar workings in coal mines: an approach, *Saf. Sci.* 47 (2009) 88–96.

- [85] L. Rutkowski, M. Jaworski, L. Pietruczuk, P. Duda, The CART decision tree for mining data streams, *Inf. Sci.* 266 (2014) 1–15.
- [86] Z.-H. Zhou, *Ensemble Methods: Foundations and Algorithms*, Chapman and Hall/CRC, 2012.
- [87] Y. Freund, R.E. Schapire, A decision-theoretic generalization of on-line learning and an application to boosting, *J. Comput. Syst. Sci.* 55 (1997) 119–139.
- [88] M. Maniruzzaman, M. Jahanur Rahman, B. Ahammed, M.M. Abedin, H.S. Suri, M. Biswas, A. El-Baz, P. Bangeas, G. Tsoulfas, J.S. Suri, Statistical characterization and classification of colon microarray gene expression data using multiple machine learning paradigms, *Comput. Methods Progr. Biomed.* 176 (2019) 173–193.
- [89] K. Li, G. Zhou, J. Zhai, F. Li, M. Shao, Improved PSO_AdaBoost ensemble algorithm for imbalanced data, *Sensors* 19 (2019).
- [90] C. Jia, Q. Yang, Q. Zou, NucPosPred: predicting species-specific genomic nucleosome positioning via four different modes of general PseKNC, *J. Theor. Biol.* 450 (2018) 15–21.
- [91] J.H. Friedman, Greedy function approximation: a gradient boosting machine, *Ann. Stat.* 29 (2001) 1189–1232.
- [92] J.H. Friedman, Stochastic gradient boosting, *Comput. Stat. Data Anal.* 38 (2002) 367–378.
- [93] N.S. Malan, S. Sharma, Feature selection using regularized neighbourhood component analysis to enhance the classification performance of motor imagery signals, *Comput. Biol. Med.* 107 (2019) 118–126.
- [94] M. Rawashdeh, S. Lewis, M. Zaitoun, P. Brennan, Breast lesion shape and margin evaluation: BI-RADS based metrics understate radiologists' actual levels of agreement, *Comput. Biol. Med.* 96 (2018) 294–298.
- [95] A. Gupta, P. Singh, M. Karlekar, A novel signal modeling approach for classification of seizure and seizure-free EEG signals, *IEEE Trans. Neural Syst. Rehabil. Eng.* 26 (2018) 925–935.
- [96] Z. Zhang, T. Wen, W. Huang, M. Wang, C. Li, Automatic epileptic seizure detection in EEGs using MF-DFA, SVM based on cloud computing, *J. X Ray Sci. Technol.* 25 (2017) 261–272.
- [97] H. Namazi, V.V. Kulish, J. Hussaini, J. Hussaini, A. Delaviz, F. Delaviz, S. Habibi, S. Ramezanpoor, A signal processing based analysis and prediction of seizure onset in patients with epilepsy, *Oncotarget* 7 (2016) 342–350.
- [98] F. Hasanzadeh, M. Mohebbi, R. Rostami, Prediction of rTMS treatment response in major depressive disorder using machine learning techniques and nonlinear features of EEG signal, *J. Affect. Disord.* 256 (2019) 132–142.
- [99] R.Q. Ge, M.L. Zhou, Y.X. Luo, Q.H. Meng, G.Q. Mai, D.L. Ma, G.Q. Wang, F. F. Zhou, McTwo: a two-step feature selection algorithm based on maximal information coefficient, *BMC Bioinf.* 17 (2016).
- [100] A. Blumer, A. Ehrenfeucht, D. Haussler, M.K. Warmuth, Occam's razor, *Read. Mach. Learn.* 24 (1987) 377–380.
- [101] M. Shah, M. Marchand, J. Corbeil, Feature selection with conjunctions of decision stumps and learning from microarray data, *IEEE Trans. Pattern Anal. Mach. Intell.* 34 (2012) 174–186.
- [102] H. Qu, J. Gotman, A patient-specific algorithm for the detection of seizure onset in long-term EEG monitoring: possible use as a warning device, *IEEE Trans. Biomed. Eng.* 44 (1997) 115–122.
- [103] S. Raghu, N. Sriraam, Y. Temel, S.V. Rao, A.S. Hegde, P.L. Kubben, Performance evaluation of DWT based sigmoid entropy in time and frequency domains for automated detection of epileptic seizures using SVM classifier, *Comput. Biol. Med.* 110 (2019) 127–143.
- [104] M. Kaleem, D. Gurve, A. Guergachi, S. Krishnan, Patient-specific seizure detection in long-term EEG using signal-derived empirical mode decomposition (EMD)-based dictionary approach, *J. Neural. Eng.* 15 (2018).
- [105] 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.
- [106] A. Bhattacharyya, R.B. Pachori, A multivariate approach for patient-specific EEG seizure detection using empirical wavelet transform, *IEEE (Inst. Electr. Electron. Eng.) Trans. Biomed. Eng.* 64 (2017) 2003–2015.
- [107] P. Thodoroff, J. Pineau, A. Lim, Learning robust features using deep learning for automatic seizure detection, in: D.-V. Finale, F. Jim, K. David, W. Byron, W. Jenna (Eds.), *Proceedings of the 1st Machine Learning for Healthcare Conference*, PMLR, *Proceedings of Machine Learning Research*, 2016, pp. 178–190.
- [108] A. Harati, S. Lopez, I. Obeid, J. Picone, M. Jacobson, S. Tobochnik, The TUH EEG CORPUS: a big data resource for automated EEG interpretation, in: *2014 IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*, IEEE, 2014, pp. 1–5.