Journal of Neural Engineering

PAPER

A robust seizure detection and prediction method with feature selection and spatio-temporal casual neural network model

To cite this article: Yuanming Zhang *et al* 2023 *J. Neural Eng.* **20** 056036

IOP Publishing J. Neural Eng. 20 (2023) 056036 https://doi.org/10.1088/1741-2552/acfff5

Journal of Neural Engineering

RECEIVED
6 January 2023

REVISED
27 September 2023

ACCEPTED FOR PUBLICATION
4 October 2023

PUBLISHED
18 October 2023

PAPER

A robust seizure detection and prediction method with feature selection and spatio-temporal casual neural network model

Yuanming Zhang[1], **Xin Li**[1], **Shuang Wang, Haibin Shen and Kejie Huang***

Zhejiang University, 38 Zheda Road, Hangzhou, People's Republic of China

[1] The authors contributed equally.

* Author to whom any correspondence should be addressed.

E-mail: huangkejie@zju.edu.cn

Keywords: epilepsy, electroencephalogram, feature selection, channel selection, temporal causal neural network

Abstract

Objective. Epilepsy is a fairly common condition that affects the brain and causes frequent seizures. The sudden and recurring epilepsy brings a series of safety hazards to patients, which seriously affects the quality of their life. Therefore, real-time diagnosis of electroencephalogram (EEG) in epilepsy patients is of great significance. However, the conventional methods take in a tremendous amount of features to train the models, resulting in high computation cost and low portability. Our objective is to propose an efficient, light and robust seizure detecting and predicting algorithm. *Approach.* The algorithm is based on an interpretative feature selection method and spatial-temporal causal neural network (STCNN). The feature selection method eliminates the interference factors between different features and reduces the model size and training difficulties. The STCNN model takes both temporal and spatial information to accurately and dynamically track and diagnose the changing of the features. Considering the differences between medical application scenarios and patients, leave-one-out cross validation (LOOCV) and cross-patient validation (CPV) methods are used to conduct experiments on the dataset collected at the Children's Hospital Boston (CHB-MIT), Siena and Kaggle competition datasets. *Main results.* In LOOCV-based method, the detection accuracy and prediction sensitivity have been improved. A significant improvement is also achieved in the CPV-based method. *Significance.* The experimental results show that our proposed algorithm exhibits superior performance and robustness in seizure detection and prediction, which indicates it has higher capability to deal with different and complicated clinical situations.

1. Introduction

Epilepsy is a neurological disorder characterized by recurrent and sudden seizures. A seizure is an uncontrolled electrical disturbance in the brain, which may cause jerky muscle movements, loss of consciousness, or confusion. According to the World Health Organization, there are approximately 60 million people with epilepsy worldwide, making epilepsy one of the most common neurological disorders. Among them, about 30% of the patients cannot relieve their seizures by drug treatment. Therefore, an early seizure detection or prediction method could assist timely interventions to greatly reduce the risk of the epilepsy.

Electroencephalogram (EEG) is a measurement of the continuous electrical activity of the brain. EEG signals have been widely used in biomedical applications for the diagnosis, treatment, and monitoring of brain diseases. Deep learning is one of the most powerful techniques to analyze EEG signals to detect and predict seizures. However, conventional deep learning-based methods often ignore that EEG features have different levels of correlations with the epileptic seizure. They simply utilize numerous features to train their models, which may lead to large neural network model size and high computational cost. The mussy features increase the training difficulties and deteriorate the inference accuracy. As a result, the conventional models lack interpretability and also suffer from low accuracy in cross-patient experiments. Therefore, they are not suitable for wearable seizure detection devices to detect or predict seizures. In view of the above, great efforts should be made to build up a lightweight and robust neural network model for cross-patient applications.

To address the above problems, this paper proposes a seizure detection and prediction algorithm based on feature selection and spatial-temporal information. Firstly, a feature selection algorithm based on dynamic time warping (DTW) is used to screen the optimal feature space. Secondly, a spatial-temporal casual neural network (STCNN) model is proposed to extract spatial and temporal correlation information. From the perspective of spatial information, a channel self-selection module is designed to independently select key channels for each patient to reduce the confounding effect of irrelevant channels; from the perspective of temporal information, residual temporal convolutional network (Res-TCN) is used to learn the timing correlation of EEG signals.

Considering the differences between medical application scenarios, both leave-one-out cross validation (LOOCV) and cross-patient validation (CPV) methods are used to conduct experiments. LOOCV uses the data of the first N seizures as the training set and the rest of the seizure data as the test set. LOOCV could achieve high detection and prediction accuracy, which has been widely used in epileptic seizure diagnosis. However, LOOCV requires a lot of expert work to label the seizures. In contrast, CPV experiment only needs to label a part of the patients as the training set, and the trained model could be used directly on the other unlabeled patients. Therefore, it is more suitable for early seizure detection and prediction and has greater medical application value. The experimental results show that our proposed method outperforms the conventional methods in various application scenarios. For example, it improves the accuracy of the CPV prediction by more than 10% on the CHB-MIT dataset.

The rest of this paper is structured as follows. Section 2 introduces the related work of feature extraction, feature selection, detection and prediction models, and validation methods. Section 3 elaborates on our proposed methods, including the DTW-based feature selection algorithm and the STCNN model. The introduction of the datasets is given in section 4. Section 5 presents the experimental results, including the experimental setting, LOOCV experiments, CPV experiments, ablations, and comparisons with current techniques. Finally, conclusions are drawn in section 6.

2. Related work

2.1. Feature extraction

Time-domain and frequency-domain EEG features have been widely used in epilepsy diagnose. Time-domain features extract the shape and intensity information of EEG, which are the most straightforward features for epilepsy diagnose. References [1–3] extracted basic time-domain features such as mean, skewness, kurtosis, coefficient of variation, and square root amplitude. Besides, hurst exponent and Higuchi fractal dimension are used in [4, 5]. Moreover, entropy features such as approximate entropy, sample entropy, permutation entropy [6–9] also have been utilized in the field of epilepsy. Frequency-domain features can reflect the frequency components of the signal more intuitively, such as autoregressive model [10], fast Fourier transform [11], and mean amplitude spectrum [12]. However, EEG is a non-stationary signal, and a single frequency domain analysis or time domain analysis will lose the other dimension of information. Therefore, time–frequency analysis applied to non-stationary signals is also widely used in the field of EEG feature extraction. To obtain time–frequency variation information, short-time Fourier transform (STFT) [13], wavelet transform (WT) [14], empirical mode decomposition (EMD) [12], and power spectral density [15] are also widely used in EEG feature extraction.

2.2. Feature selection

Filtering, embedding, and wrapping are the commonly used feature selection algorithms. Among them, filtering and embedding algorithms have been widely used in the feature selection for EEG signals. For example, Zhang and Parhi [16] adds L1 regularization to the model for sparse feature selection. However, its weak scalability prevents it from defining the size of the feature subset. Hassan *et al* [17] utilizes the mutual information between a single feature and a label to rank features, and [17–19] select a filtering algorithm based on information metric and distance metric. Unfortunately, all of these methods fail to consider the feature redundancy.

2.3. Detection and prediction models

The traditional machine learning methods for seizure detection and prediction include K-nearest neighbor [20], support vector machine (SVM) [21–24], decision tree (DT) model [25], linear discriminant analysis (LDA) [26], and a combined model of multiple weak classifiers [27]. For example, [22, 23] use SVM as the seizure detection model. SVM is used to achieve the seizure prediction in [24, 28]. Alotaiby *et al* [26] uses LDA model to achieve the seizure prediction. What is more, some other machine learning methods like RUBoost model [29] also have application in the seizure detection and prediction. Recently, deep learning models have been used to improve the detection and prediction performance. In terms of the seizure detection, [30] uses one-dimension convolutional neural network (1D CNN) and [31] uses 2D CNN to achieve seizure detection. References [32, 33] employ graph convolutional networks (GCNs) to extract graph structure information between channels. But those works ignore the temporal correlation of sequences when processing EEG signals. References [34, 35] adopt a bidirectional long-short-term memory (Bi-LSTM) network to extract the time correlations. Li *et al* [36] uses a unified temporal-spectral squeeze-and-excitation network (SENet). Zhang *et al* [37] achieve the seizure detection based on the bidirectional gated recurrent unit network. Moreover, CNN and recurrent neural network (RNN) combined structure are used in [38, 39], where CNN further generalizes deep features, and RNN learns time series information. In terms of seizure prediction, [40–42] use CNN or three-dimension (3D) CNN as the seizure prediction model. Sun *et al* [43] adopts the dual-input CNN model to conduct the prediction task. References [12, 44] adopt the transfer learning method to transfer the pre-trained model to the EEG domain. However, those methods rarely consider the model's robustness. They do not achieve high performance on the CPV, which indicates those models are not robust for the complicated medical application scenarios.

3. Proposed method

To decrease the computational cost and improve the performance of the model in both LOOCV-based and CPV-based experiments, we propose a method that extracts the key features and utilizes the spatial and temporal information of EEG signals. As shown in figure 1, our proposed method consists of two parts. The first part is the DTW-based feature selection algorithm to select the best feature subset based on the importance. The second part is the STCNN seizure detection and prediction algorithm. The algorithm designs a channel self-selection module to extract spatial information by independently selecting key channels for each patient to reduce the confounding effect of irrelevant channels. Meanwhile, Res-TCN is used to learn the timing correlation of EEG signals to fully utilize the temporal information. In addition, the algorithm designs streaming computing and K-of-N module to improve the reliability of the results. Moreover, the adaptive focal loss function is used to solve the learning difficulty due to the imbalance samples.

3.1. Multiple features extraction

After cutting the data into 2 s segments, we extract a total of 134 dimension features from the time domain, frequency domain, and time-frequency domain. Some features have more than one dimension. It is worth noting that the raw data is filtered and normalized before extracting features.

As shown in table 1, 30 dimension time domain features are extracted at first which cover the basic statistical properties, entropy properties, energy properties, stochastic correlation properties, and Hjorth series parameters of the signal.Secondly, a total of 31 dimension frequency domain features are extracted. We convert the original EEG signal to the frequency domain to analyze the frequency components and get the frequency domain features. At the end, a total of 73 dimension time–frequency domain features are extracted. We adopt STFT, WT and other time spectrum analysis algorithms to extract transient and discontinuous waveform changes in time series signals.

3.2. DTW-based feature selection algorithm

The whole feature set contains not only key features relative to the seizure, but also many redundant and interferential features. To select an optimal feature subset, this paper proposes a DTW based feature selection algorithm. The algorithm mainly includes four parts: a Savitzky–Golay (S-G) smoothing module for extracting feature trends, a dynamic time-warping metric module for calculating feature distances, a de-redundant feature module, and a feature subset search for finding the best subset module.

3.2.1. SG smoothing module

To obtain the tendency profile of the feature, the S-G filter is used to smooth the feature curve and remove the mixed outlier data. The S-G filter adopts the least square method to fit parameters, which can retain more useful information than the simple moving average method.

The window length is crucial to feature selection performance. If the window length is too short, the feature curve still fluctuates in a short period of time. On the contrary, if the window length is too long, the S-G filter will lose a lot of useful information while removing noise fluctuations. An experiment is conducted to explore the optimum window length. The result is shown in figure 2. It obtains the highest accuracy when the ratio of the window length to the original EEG signal length is 1/10. Therefore, the window length selected in the feature extraction in this paper is 1/10 of the sequence length.

3.2.2. DTW metrics

The patient's seizure state changes from the interictal period to the preictal period over time, and finally enters the ictal period. The DTW algorithm is used to measure the consistency of the change trend of the feature over time with the patient's state changing. The DTW algorithm uses dynamic programming to capture the similarity in shape of two time series and calculates the distance between the contour curve of the feature and the state curve.

Both decreasing and increasing trends of the features are calculated, as shown in figure 3. When calculating the upward trend, the state quantity gradually increases with the different seizure states. The interictal, pre-ictal, and ictal periods are represented by 0, 1, and 2, respectively. When the downward trend is calculated, the state quantity gradually decreases with the seizure state, with 2 for the interictal period, 1 for the preictal period, and 0 for the ictal period. The sequences are normalized before being fed into the

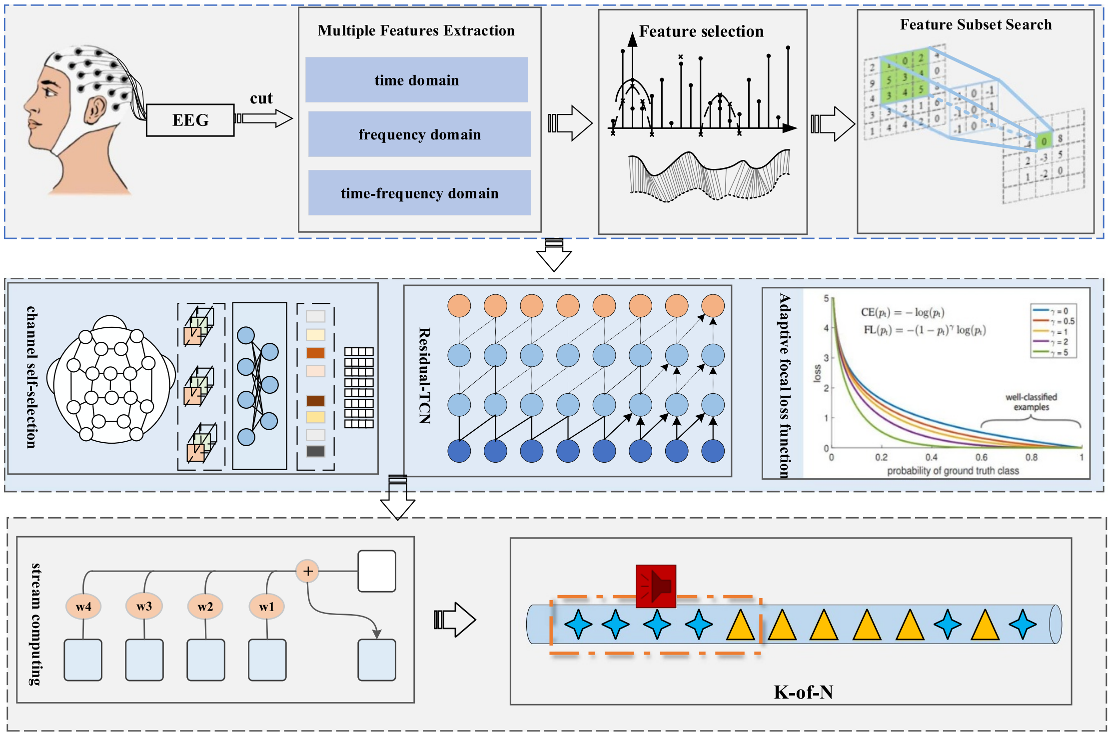

Figure 1. Illustration of the feature selection and temporal causal neural network model.

Table 1. Experimental results of different feature selection algorithms on seizure three-classification task.

Domain	Features	Details
Time domain	Basic statistical properties	Mean, variance, rectified mean, standard deviation, peak-to-peak, skewness, kurtosis, rms, rms amplitude, rectified mean, shape factor, crest factor, impulse factor, margin factor, zero-crossing times, coefficient of variation, line length (LL), detrended fluctuation analysis
	Entropy properties	Approximate entropy, sample entropy, singular value entropy, fuzzy entropy, permutation entropy
	Energy properties	Energy, nonlinear energy
	Stochastic correlation properties	Hurst exponent, Higuchi fractal dimension, Hjorth series parameters
Frequency domain	Mean amplitude spectrum	The α, β, γ, θ, and low-δ rhythm wave bands are divided into sub-bands and the average amplitude spectrum of each sub-band is calculated.
	Sub-band energy	Overall energy, spectral entropy in frequency domain, intensity-weighted average frequency, intensity-weighted bandwidth, spectral edge frequency, peak frequency
Time frequency domain	Discrete wavelet transform	Mean, variance, skewness, kurtosis, energy, line length and Higuchi fractal dimension of each order coefficient after DWT decomposition
	Empirical mode decomposition	Mean, variance, energy, coefficient of variation and information entropy features of the IMF after EMD decomposition of the signal

DTW algorithm. If the distance between the feature changing curve and the uptrend state curve is smaller, it demonstrates the feature changing is uptrend, and vice versa. Therefore, the smaller distance is adopted to participate in next operations.

3.2.3. De-redundant feature module

In the process of feature selection, not only the separability of data but also the redundancy of features should be considered. Therefore, after obtaining the ranking of the consistencies between the features and

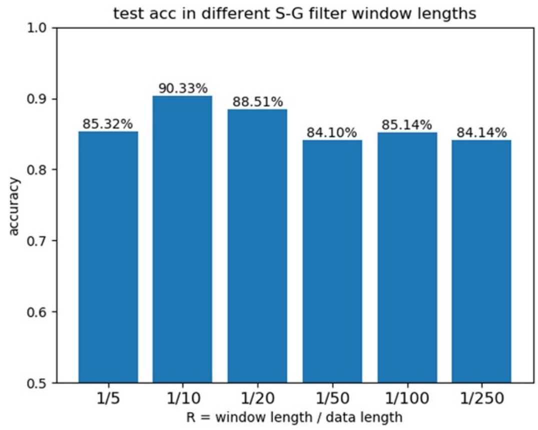

Figure 2. A patient's test ACC in different S-G filter window lengths.

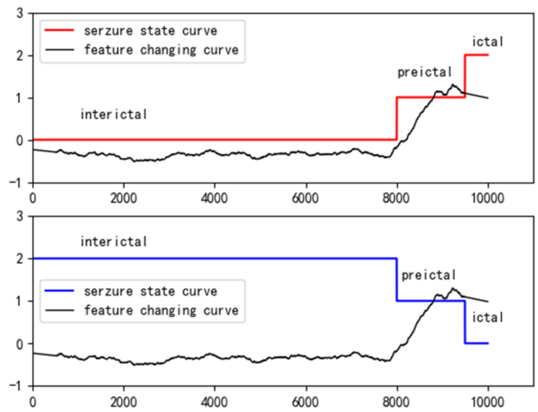

Figure 3. Comparison of the consistency between feature changing and uptrend/downtrend state curves.

the patients' states over time, the Pearson correlation coefficient is further utilized to calculate the similarity between the features, thereby removing the features with high redundancy. For 134 dimension features, a certain number of features are subtracted from the feature space each time, and finally the feature subset with the best result is selected. The upper limit of the feature subset number is set to 80. This paper trains the learning algorithm when the numbers of selected features are 80, 70, 60, 50, 40, 30, 20, and 10, and calculates evaluation indicators such as accuracy.

3.2.4. Feature subset search

After obtaining the feature sets with different numbers of features, we adopt the idea of sequence backward search and use the AlexNet [45] as a learning model to select the best feature subset. The AlexNet was proposed by Alex Krizhevsky in ImageNet LSVRC- 2010. The input of the net has two 2 dimensions, including feature number and channel number, which are regarded as the two dimensions of an image. In addition, some other methods are used in the paper for comparison, such as ANOVA [46], mRmR [47], and XGBoost [48].

3.3. Seizure detection and prediction algorithm

3.3.1. Res-TCN module

To effectively and flexibly learn the temporal information in EEG, this paper adopts the Res-TCN, as shown in figure 4. The block I in figure 4(a) contains two rounds of dilated causal convolution, weight normalization, activation function and dropout. The dilated convolution has a larger receptive field than the normal convolution, as shown in figure 4(d). The parameter dilation d is 2^{i-1}, where $i \in \{1,2,3\}$ is the block number. Weight normalization can increase the convergence speed. Dropout module is utilized which has a probability p to zero the output in order to prevent over-fitting. p is set to 0.6 in the experiments. After two rounds, the original input data is added to the output if the input and output match dimensions. Otherwise, a 1×1 conv in block II is used to match the dimensions, as shown in figure 4(b). The input of the Res-TCN is $x \in R^{T\times 1\times F}$, where T and F represent the time and feature dimensions, respectively. In this paper, $T = 19$ and $F = 40$. The convolution kernel parameters have a dimension of $R^{k\times 1\times c_\mathrm{i}\times c_\mathrm{o}}$, where k, c_i and c_o are kernel size, input channel and output channel of the convolution layer, respectively. k is set to 4 in the experiments. The c_o of all blocks is set to 10 in LOOCV-based experiments and 22 in CPV-based experiments, respectively. The c_i depends on the c_o of the previous convolution or the input feature dimension F.

3.3.2. Channel self-selection module

Since different patients may have the seizure occurring in different EEG channels, we adopt the channel attention method to calculate the importance of each channel before sending the features to the Res-TCN module, and finally the weighted sum of each channel is sent to the TCN network. It converts the number of EEG channels to one, for which the detection and prediction algorithm can process dataset with different numbers of channels. The overall calculation process is shown in figure 5. A perception layer is adopted to nonlinearly transform the input channel vector, which can be expressed as

$$M = \tanh\left(W^\mathrm{T} H\right) \tag{1}$$

where $H = [h_1, h_2, \ldots, h_n]$ is the input channel vector, and W is the weight matrix. The result M is then sent to the Softmax to generate weight coefficients,

$$\alpha = \mathrm{soft\,max}\left(M\right). \tag{2}$$

Finally, the weighted summation α is multiplied with the input feature matrix, which is then sent to the temporal convolution module.

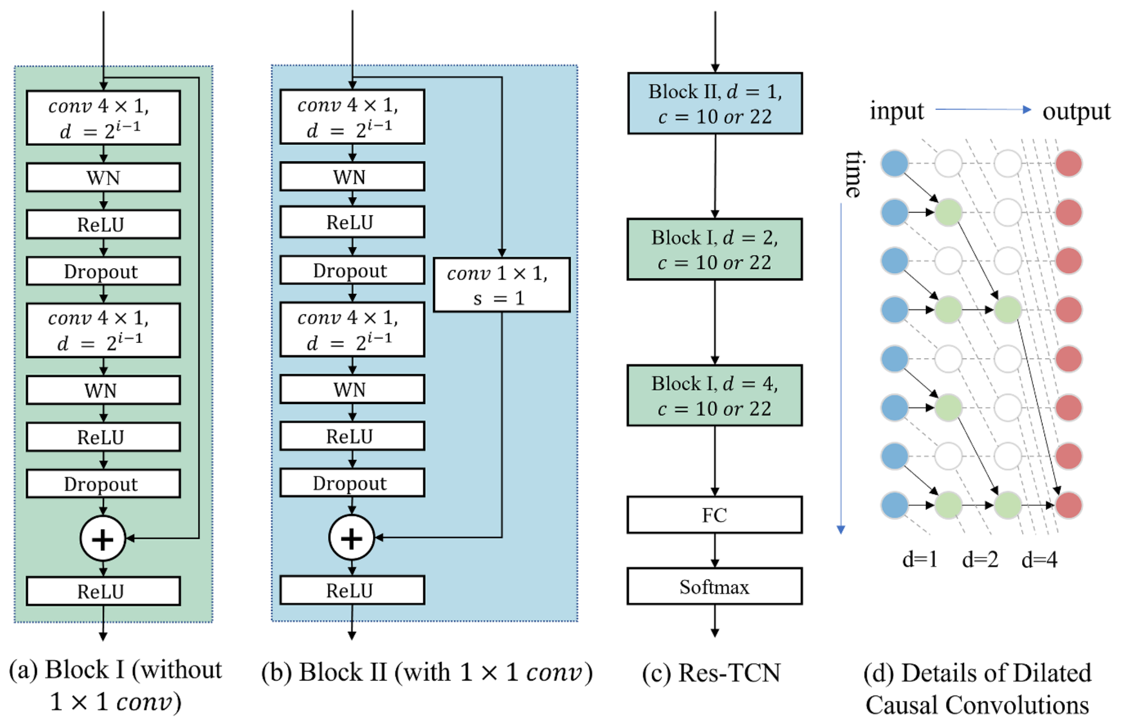

Figure 4. Architecture of Res-TCN. (a) the block without 1 × 1 conv (b) the block with 1 × 1 conv (c) the Res-TCN (d) how the dilated convolution works.

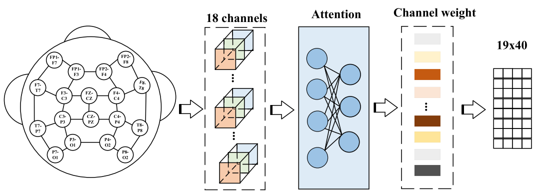

Figure 5. The schematic diagram of channel self-selection module.

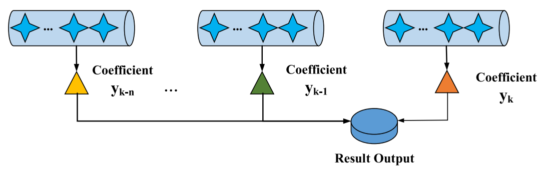

Figure 6. Schematic diagram of stream computing module.

3.3.3. Streaming module

To further enhance the correlation before and after the continuous data and the stickiness of the prediction results, the streaming computing mode is developed, as shown in figure 6. The prediction result of the current moment depends on the previous n moments, which can be expressed as

$$y_k = w_0\hat{y}_k + w_1 y_{k-1} + w_2 y_{k-2} + \cdots + w_n y_{k-n}. \quad (3)$$

We consider that the last four outputs have effect on the current result. Therefore, n is set to four and the weight values are trained as a linear regression network.

3.3.4. Adaptive focal loss function

For binary classification problems, the most commonly used cross-entropy loss function can be expressed as

$$FL(p_t) = -\alpha_t(1-p_t)^{\gamma}\log(p_t). \quad (4)$$

In this paper, we regard α_t as an automatically adjusted parameter to adapt to different patients. Its specific formula is

$$\alpha_t = \frac{N_n}{N_n + N_p} - \beta. \quad (5)$$

Among them, N_p is the number of positive samples and N_n is the number of negative samples. In this paper, β is set to 0.2 and 0 in the seizure detection and prediction tasks, respectively. The hyper-parameter γ decreases the loss contribution of the samples which are easy to distinguish. We have conducted a pre-experiment to figure out the best γ value. We set γ to 0.5, 1, 2, 3, 4 and 5. The result in figure 7 shows it reaches the highest accuracy when γ is set to 2. In follow-up experiments, γ is set to 2.

3.3.5. K-of-N

It is a very common phenomenon to produce isolated false positives during the interictal period. To reduce the frequency of its occurrence and thus reduce the false alarm rate, we adopt the K-of-N voting strategy

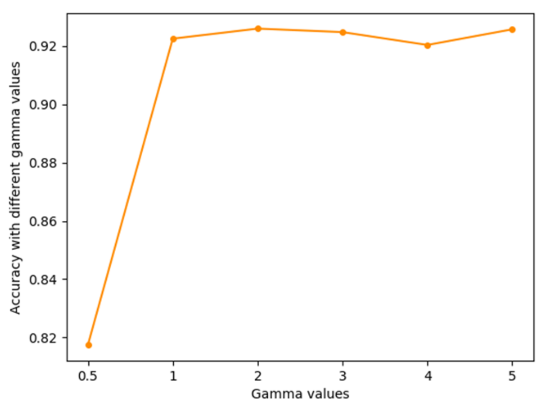

Figure 7. The result of the pre-experiment to find the best γ.

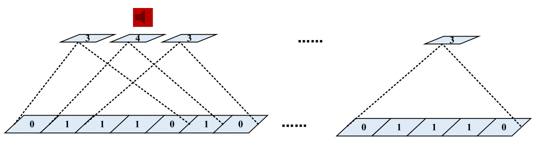

Figure 8. The schematic diagram of the K-of-N module.

from [9]. Specifically, the alarm is only issued when there are at least K positives in N predictions. The specific implementation is shown in figure 8. In this paper, N is 5 and K is 4.

4. Datasets

The Bonn dataset is widely used in EEG analysis, which was collected by the University Hospital of Bonn. The dataset includes five subsets A, B, C, D, and E. Among them, the subsets A and B contain the normal EEG data collected by five healthy volunteers with their eyes open and closed, using the standard 10–20 system scalp EEG acquisition method; the subsets C and D are the interictal data of epilepsy patients, which were collected by intracranial electrodes; E is the EEG data of epilepsy patients during seizures, which was also collected by intracranial electrodes. In addition, each subset of the Bonn dataset contains 100 EEG clips, each of which is manually processed data with a time length of 23.6 s, a sampling frequency of 173.61 Hz, and 4096 data points. This paper adopts all five sub-categories of Bonn and uses a sliding window to cut the 23.6 s data into 2 s segments with a window overlap of 50%.

CHB-MIT is a children's EEG dataset publicly available on PhysioNet, with intractable epilepsy collected by Boston Children's Hospital using the international 10–20 standard. It should be noted that chb21 is the data obtained 1.5 years after chb01, and chb24 is the latest data added in December 2010 with no detailed information, so there are a total of 22 available subjects. Among them, five are male from 3 to 22 years old; 17 are females between 1.5 and 19. The data set contains a total of 198 seizures, with a total recording time of about 960 h. Following the dataset splitting method in [40, 43], the CHB-MIT dataset is divided into three categories including interictal, preictal, and ictal. Specifically, the interictal period is the recording at least 4 h before the seizures or at least 4 h after the ictal, and the preictal period is the recording at least t_{sop} minutes before the ictal, where t_{sop} is determined by the length of set the time of onset period (SOP). In addition, if there are multiple epileptic seizures in a short period of time, they will be regarded as the same seizure, and only the leading seizure will be used as the ictal period when the data is trimmed.

The Kaggle competition dataset, sponsored by the National Institutes of Health, the Epilepsy Foundation, and the American Epilepsy Society, includes intracranial EEG data from five dogs and two patients. In the process of collecting and arranging the intracranial EEG of dogs, the data channels are 16, the sampling frequency is 400 Hz, the interictal period is defined as the data 1 week before or after the ictal, and preictal is the data between 65 min and 5 min before onset. In the patients' EEG data, patient 1 has 15 electrode channels, patient 2 has 24 electrode channels, and the sampling frequency of both patients is 5000 Hz. The data collection for patients is more difficult, so the interictal period is defined 4 h before and after ictal. Dog-specific or patient-specific data are sliced into 10 min long EEG segments. There are labeled and unlabeled data in the dataset, and only the publicly labeled data is used in this paper. The Kaggle dataset is pre-processed in the same way as CHB-MIT.

Another dataset containing scalp EEG signals of 14 epileptic patients was acquired by the Unit of Neurology and Neurophysiology at the University of Siena, Italy. The EEG signals of were collected according to the international 10–20 system with a sampling rate of 512 Hz. One of the patients has 21 EEG data channels, while the rest of them have 29 data channels. Every patient in the Siena dataset has a folder which consists of 1–5 EEG recordings of the patient in EDF format, and a text file summarizing the detailed information of the EDF files, such as the seizure onset/end time and sampling channels. The dataset is divided into ictal, preictal and interictal. However, unlike the CHB-MIT dataset, the interictal in Siena dataset is 1 hour away from the ictal, because of the less length of the original signals. The other preprocessing steps of Siena dataset are conducted in the same way as CHB-MIT.

5. Experiments

5.1. Model validation methods

In the field of seizure detection and prediction, there are three common validation methods, namely whole

Table 2. main tasks conducted in the experiment.

Part	Validation method	Task		Dataset
Part I	WCV	Feature selection	Two-class detection Two-class prediction Three-class	CHB-MIT, Bonn (for validation only)
Part II	LOOCV	Detection	Segment-based Event-based	CHB-MIT, Siena
		Prediction	Event-based	CHB-MIT, Siena, Kaggle
	CPV	Detection	Segment-based Event-based	CHB-MIT, Siena
		Prediction	Event-based	CHB-MIT, Siena, Kaggle

cross-validation (WCV), LOOCV, and CPV. WCV refers to sorting and shuffling all the cut data of the patients, and then randomly dividing the training set and the test set. The data segmentation method for LOOCV refers to taking a certain seizure data of a patient as the test set and the rest of the data of this patient as the training set. LOOCV can be adopted for the seizure detection task [22, 23, 30, 36, 37, 49] and the seizure prediction task [24, 26, 40, 41, 43]. CPV refers to using the data of a certain patient as the test set and the data of the rest patients as the training set. Both seizure detection and prediction experiments can be carried out with CPV [44, 50–53].

5.2. Experiment environment

In this paper, both feature selection experiment and seizure detection and prediction experiment are conducted.

The feature selection experiment is conducted on CHB-MIT dataset. The screened feature subsets are applied to Bonn dataset for cross-validation. In this task, the WCV method is used, which reflects the separability of the data and the performance of the model. It is worth noting that due to the insufficient data before and during the onset, the sliding window method is used for data cutting. After the original EEG data is cut into 2 s segments and the multi-domain feature extraction process is completed, the whole data set is broken up and randomly divided into a 5:5 training set and test set, and then the training set is further divided into 9:1, where 1 is the validation set.

To make the algorithm more suitable for medical application scenarios, we use the CHB-MIT, Siena and Kaggle competition datasets to conduct seizure daily detection and prediction experiments. In detection and prediction tasks, we use the same architecture of the model, and train the model with different categories of samples. For the detection task, ictal samples are regarded as the positives. Both interictal and preictal samples are regarded as the negatives. For the prediction task, preictal and interictal samples are the positives and negatives, respectively. Ictal samples are not used in the prediction task. In each task, we train the model to identify the positive samples from the negative samples. In those tasks, both LOOCV and CPV methods are used. To capture temporal information, the length of the data segment in the seizure detection and prediction experiment is 20 s [32], and the specific data dimension is $x \in R^{C\times T\times F}$, where $C = 18$, $T = 19$, $F = 40$. Because the amount of interictal data is large, only part of the data is randomly selected for training. To make it clear, table 2 shows the detailed description of the main tasks that are conducted in the experiment.

Table 3. The feature evaluation indicators for feature selection task. *TP*, *TN*, *FP*, and *FN* are the sample numbers of true positive, true negative, false positive, and false negative, respectively.

Indicators	Equations
Accuracy (ACC)	$\frac{TP+TN}{TP+TN+FP+FN}$
Sensitivity (SEN)	$\frac{TP}{TP+FN}$
Specificity (SPE)	$\frac{TN}{TN+FP}$

5.3. System evaluation

In the feature selection task, the feature evaluation indicators used are accuracy (ACC), sensitivity (SEN), specificity (SPE), and the area under the receiver operating characteristic (ROC) curve. The specific calculation formulas of ACC, SEN and SPE are provided in table 3.

In the seizure detection and prediction tasks, there are two evaluation methods, segment-based and event-based. The segment-based evaluation metrics are ACC, SEN, and SPE, whose definitions are the same as given in the table 3. The event-based evaluation metrics are SEN and SPE, but the definitions of TP, TN, FP, and FN are different. TP refers to the number of seizure events in which there is at least one successful positive result; FP refers to a positive detection during the non-seizure period, i.e. false alarm; TN refers to the correct categorization of the non-seizure period data as a non-seizure period; FN refers to the number of seizure events in which the model failed to successfully detect the current seizure. The seizure detection tasks are conducted in both segment-based and event-based methods. The seizure prediction tasks are conducted in event-based

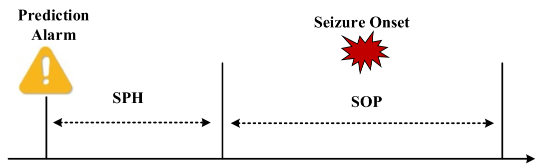

Figure 9. Schematic diagram of SOP and SPH.

Table 4. Experimental results of different feature selection algorithms on three-classification task of seizure.

	proposed		ANOVA		mRmR		XGBoost	
	Acc	Var	Acc	Var	Acc	Var	Acc	Var
134	94.1	14.33	94.1	14.33	94.1	14.33	94.1	14.33
80	94.63	13.46	94.30	13.9	94.29	10.38	94.02	14.13
70	94.75	13.52	94.40	12.5	95.54	10.35	94.19	14.08
60	98.57	2.11	98.48	2.11	95.90	4.87	97.91	2.99
50	99.13	1.65	98.99	1.71	97.11	3.69	98.72	2.15
40	99.20	1.31	99.07	1.55	96.27	4.34	98.63	2.33
30	99.05	1.55	98.95	1.79	94.23	4.83	98.31	2.66
20	98.36	2.79	98.63	2.37	93.33	5.4	97.93	2.86
10	98.27	3.09	98.76	2.39	91.60	6.02	97.26	2.89

method. To make it more clinically meaningful, we use false prediction rate (FPR) in seizure prediction tasks. The FPR in this paper is defined as the number of false predictions per hour in the interictal period [54].

The definition of successful early warning of seizure is to give early warning within a certain range before the onset. Due to the uncertainty of the seizure onset, SOP and seizure prediction horizon (SPH) are adopted [40, 43]. A schematic diagram of SOP and SPH is shown in figure 9. The SOP and SPH are set to 30 min and 5 min, respectively. Therefore, in LOOCV and CPV experiments, TP is defined as having at least one successful warning during the preictal time period defined by SOP and SPH.

5.4. Results

5.4.1. Feature selection results

Table 4 shows the comparison results of the CHB-MIT three-classification tasks (interictal, preictal and ictal) between our proposed feature selection method and three comparative algorithms (ANOVA [46], mRmR [47] and XGBoost [48]). From the longitudinal perspective of the table, our proposed feature selection method achieves the best performance when the number of feature subsets is 40. The average accuracy of the three-classification is 99.20% and the standard deviation is 1.31%. When the number of features is 10, the accuracy still can reach 98%. The ANOVA algorithm also achieves the best accuracy when the feature subset size is 40, with an accuracy rate and standard deviation of 99.07% and 1.55%, respectively. From the horizontal view of the table, our proposed method achieves the highest accuracy when the feature subset sizes are 80, 70, 60, 50, 40, and 30.

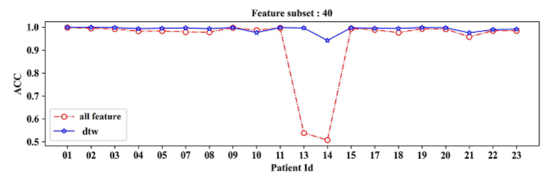

Figure 10. The comparison results before and after feature screening.

Figure 10 shows patient-specific performance on the CHB-MIT seizure three-classification task using the all-feature dataset and the optimal feature subset with the size being 40. It can be seen that, using all the features, the average accuracy of the three-classifications of seizure in all patients can reach 94%. However, the accuracy of patients 13 and 14 is only about 50%. In contrast, when using the optimal feature subset for this three-classification task, the model's performance is more stable on various patients.

Table 5 shows the comparison results with state-of-the-art works on CHB-MIT for the seizure detection dichotomous, prediction dichotomous, and three-classification tasks. The experimental results show that our proposed method outperforms most of the related works in terms of ACC, SEN and SPE, even though [55] has 634 $\times$ 1 feature vector dimension. The SEN and SPE of [32] have 0.34% and 0.03% higher than our proposed method in two-class detection, respectively. Parvez and Paul [28] has 0.14% higher SEN in two-class prediction and 0.7% higher ACC in three-class detection as well. But that is at the cost of adopting complicated hierarchical GCN with a feature vector dimension of 274 $\times$ 1.

Table 5. Comparison of seizure dichotomous and three-classification results with existing studies on the CHB-MIT dataset.

Method	Task	ACC	SEN	SPE	AUC
RNN [55]	Pred Two-class	99.37	99.37	99.6	99.37
CNN [13]	Detect Two-class	99.33	/	/	/
	Three-class	98.62	/	/	
GCN [32]	Detect Two-class	/	100.00	100.00	/
	Pred Two-class	/	99.8	99.9	/
	Three-class	99.9			
RNN [18]	Pred Two-class	96.55	96.52	97.54	/
GCN [33]	Detect Two-class	99.3	98.82	99.43	98.57
CNN [56]	Detect Two-class	99.29	99.29	99.86	/
Proposed	Detect Two-class	99.95	99.66	99.97	99.81
	Pred Two-class	99.95	99.66	99.97	99.81
	Three-class	99.2	/	/	/

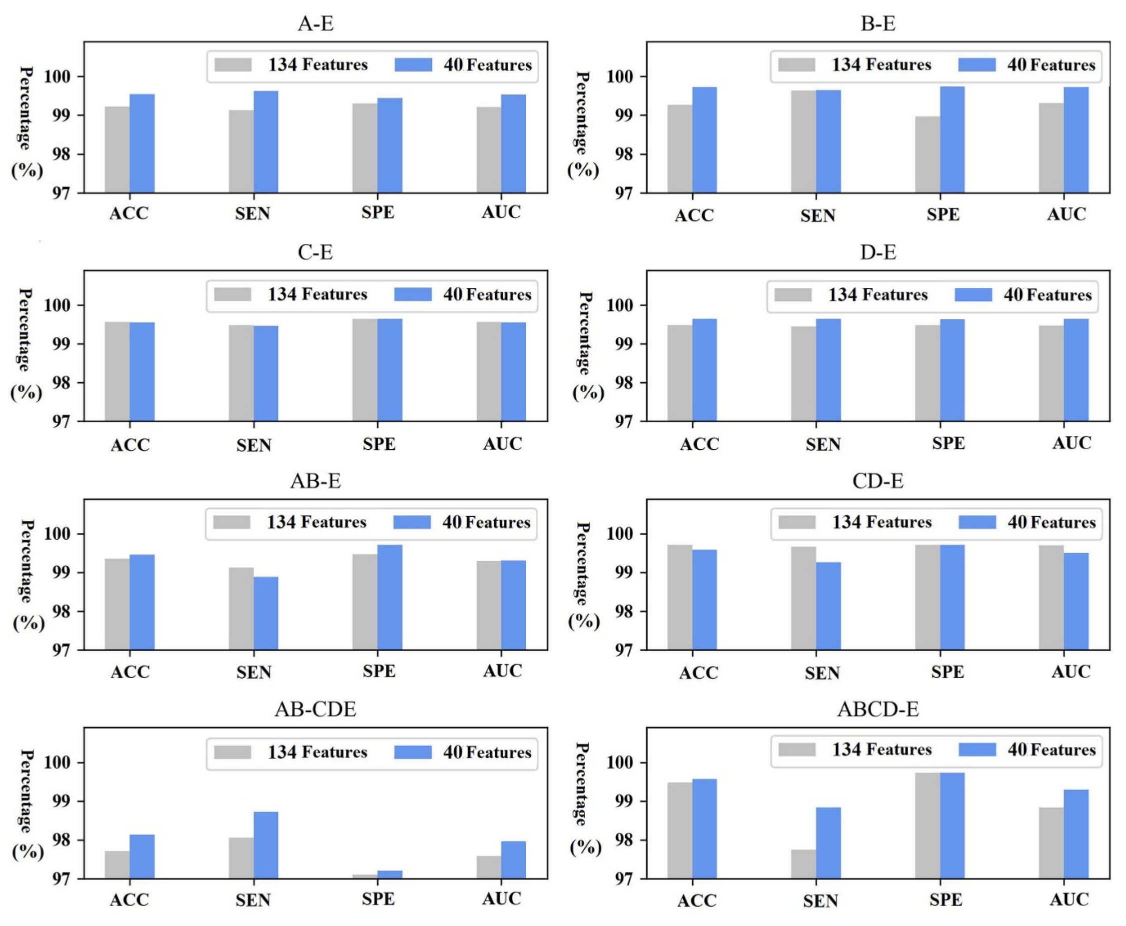

Figure 11. Comparison of the results on the Bonn dataset before and after feature screening.

Table 6. Experimental results of seizure detection and prediction based on LOOCV in CHB-MIT.

	Seizure detection											
	Segment based				Event based			Seizure prediction				
Patient Id	N	ACC (%)	SEN (%)	SPE (%)	N+	SEN (%)	SPE (%)	N	N+	SEN (%)	FPR (h^{-1})	Tp(min)
01	7	100.00	100.00	100.00	7	100.00	100.00	5	5	100.00	0.00	34.17
02	2	96.41	75.81	98.19	2	100.00	100.00	3	3	100.00	0.00	33.94
03	7	99.74	97.71	100.00	7	100.00	100.00	7	6	85.71	0.75	33.88
04	3	98.15	62.97	99.89	2	66.67	100.00	3	3	100.00	0.00	33.89
05	5	100.00	100.00	100.00	5	100.00	100.00	5	5	100.00	0.00	34.17
06	/	/	/	/	/	/	/	7	7	100.00	0.05	26.10
07	3	99.52	96.77	100.00	3	100.00	100.00	3	3	100.00	0.00	34.17
09	4	100.00	100.00	100.00	4	100.00	100.00	4	4	100.00	0.83	34.17
10	7	100.00	100.00	100.00	7	100.00	100.00	7	7	100.00	0.00	34.14
11	2	98.79	98.37	99.32	2	100.00	100.00	2	2	100.00	0.00	34.17
13	5	99.93	99.23	100.00	5	100.00	100.00	7	7	100.00	0.64	34.02
14	/	/	/	/	/	/	/	6	6	100.00	0.00	32.81
15	14	100.00	100.00	100.00	14	100.00	100.00	13	13	100.00	0.10	33.28
16	/	/	/	/	/	/	/	3	3	100.00	0.00	27.67
17	3	99.51	95.83	100.00	3	100.00	99.91	3	3	100.00	0.00	33.44
18	5	99.90	98.40	100.00	5	100.00	100.00	4	3	75.00	0.86	29.17
19	3	100.00	100.00	100.00	3	100.00	100.00	2	2	100.00	0.00	34.00
20	5	100.00	100.00	100.00	5	100.00	100.00	5	5	100.00	0.00	34.17
21	3	100.00	100.00	100.00	3	100.00	100.00	4	4	100.00	0.00	33.88
22	3	100.00	100.00	100.00	3	100.00	100.00	3	3	100.00	0.00	33.83
23	4	100.00	100.00	100.00	4	100.00	100.00	4	4	100.00	0.00	34.13
Total	85	99.55	95.84	99.86	84	98.15	100.00	100	98	98.13	0.15	33.01

To further verify the effectiveness of the proposed feature selection method, the feature subset selected from the CHB-MIT data is used in the Bonn dataset for validation. The specific results are shown in figure 11. The effect after feature selection is better than that using all features in the following cases: A-E and B-E (normal EEG and ictal EEG), D-E and CD-E (interictal period and ictal period), AB-CDE (normal and epileptic patients), and ABCD-E (non-ictal and ictal) experiment. However, in the C-E experiment, the effect before and after feature selection is almost the same. And in the CD-E experiment, the effect before feature selection is about 0.1% higher in ACC than that after feature selection. After analysis, the reason may be that in the Bonn dataset, the electrode types used in the data of the three datasets C, D and E are all intracranial electrodes, which are different from the scalp EEG of CHB-MIT. The features selected on the dataset are not perfect for the experiment on the C, D and E datasets.

5.4.2. LOOCV experimental results

The LOOCV experiments are conducted on CHB-MIT, Siena and Kaggle datasets. Table 6 shows both the seizure detection and prediction results on CHB-MIT dataset. In event-based seizure detection experiments, the algorithm is successful in most cases, except for one episode in patient 04. Moreover, our proposed STCNN algorithm achieves an average SPE of approximately 100%. In the segment-based seizure detection experimental results, the average ACC, SEN and SPE are 98.72%, 95.72% and 99.22%, respectively. In seizure prediction task, the patients with patient numbers 01, 02, 04, 05, 07, 10, 11, 14, 16, 17, 19, 20, 21, 22, and 23 achieve 100% in SEN and 0.00 h^{-1} in FPR. The average SEN, FPR, and early prediction time are 98.13%, 0.15 h^{-1}, and 33.01 minutes, respectively.

Table 7 shows the detection and prediction results on Siena dataset. The numbers of recorded seizures of some patients are less than two, so the experiments based on LOOCV could not be conducted on those patients. According to table 7, the average ACC in segment-based detection is 96.54%. In event-based detection, SEN and SPE are both 100%. SEN and FPR in event-based prediction are 100% and 0.14 h^{-1}, respectively. The averaged prediction time is 33.36 min.

Table 8 shows the seizure prediction experimental results on Kaggle dataset which has 50 seizures on 2 patients and 5 dogs. In the animal epilepsy data, 44 seizures in five dogs are successfully predicted; in the human epilepsy data, only patient 01 has one seizure with an unsuccessful alert. Overall, the averaged early prediction SEN is 95.24% and FPR is 0.07 h^{-1}.

The LOOCV-based seizure detection and prediction comparison results with the conventional methods are tabulated in tables 9 and 10. The experiments in table 9 are all conducted on the CHB-MIT dataset. In the seizure detection task, [29] uses multi-resolution dynamic pattern decomposition to extract features and uses RUSBoost for detection. Li

Table 7. Experimental results of seizure detection and prediction based on LOOCV in Siena.

	Seizure detection							Seizure prediction				
	Segment based				Event based							
Patient Id	N	ACC (%)	SEN (%)	SPE (%)	N+	SEN (%)	SPE (%)	N	N+	SEN (%)	FPR (h^{-1})	Tp(min)
PN03	2	99.24	95.62	100	2	100	100	2	2	100	0.98	35
PN05	2	87.46	66.67	96.74	2	100	100	3	3	100	0.00	33.98
PN06	2	97.05	86.15	100	3	100	100	2	2	100	0.00	31.8
PN12	2	98.87	97.83	100	2	100	100	3	3	100	0.00	34.05
PN13	2	98.97	95.83	100	3	100	100	3	3	100	0.00	31.49
PN16	2	98.71	93.48	100	2	100	100	2	2	100	0.00	34.01
PN17	2	97.61	89.49	99.09	2	100	100	2	2	100	0.00	33.19
Total	16	96.84	89.3	99.4	16	100	100	17	17	100	0.14	33.36

Table 8. Experimental results of seizure prediction based on LOOCV in Kaggle.

ID	N	N+	SEN (%)	FPR (h^{-1})	Tp (min)
Dog_1	4	4	100.00	0.06	26.96
Dog_2	7	7	100.00	0.07	31.14
Dog_3	12	12	100.00	0.26	29.69
Dog_4	16	16	100.00	0.11	32.94
Dog_5	5	5	100.00	0.00	29.30
Patient_1	3	2	66.67	0.00	24.22
Patient_2	3	3	100.00	0.00	29.61
Total	50	49	95.24	0.07	29.12

Table 9. The comparison results of seizure detection and prediction based on LOOCV in CHB-MIT.

		Segment based				Event based		
Task	Methods	N	SEN (%)	SPE (%)	ACC (%)	N+	SEN (%)	SPE(%)
Seizure detection	RUSBoost [29]	198	/	/	/	/	93.7	99.2
	SENet [36]	/	92.41	96.05	95.96	/	98.93	/
	SVM [22]	/	/	/	/	/	97.2	/
	1D-CNN [30]	145	88.14	99.62	99.54	144	99.31	/
	SVM [23]	131	97.34	97.5	97.49	129	98.47	97.5
	Bi-GRU [37]	/	93.89	98.49	98.49	/	95.49	/
	STCNN (proposed)	85	95.84	99.86	99.55	84	98.15	100.00
		Event based						
Task	Methods	N	N+	SOP (min)	SPH (min)	SEN (%)	FPR (h^{-1})	Tp(min)
Seizure prediction	CSP and LDA [26]	170	/	/	/	89	0.39	68.71
	STFT+CNN [40]	64	54	30	5	81.2	0.16	/
	3D-CNN [41]	77		30	1	87.01	0.19	/
	Wavelet transform [57]	83	74	30	1	90.92	/	/
	CNN [42]	89		/	/	95.51	0.118	27.63
	SVM, CNN, LSTM [24]	/	/	/	/	96.28	/	33
	SVM [28]	77	73	30	3	95.40	0.36	/
	STCNN (proposed)	100.00	98	30	5	98.13	0.15	33.01

et al [36] adopts SENet model. Tang *et al* [22] uses tree model for feature selection and SVM for detection. This work only carried out event-based experiments. Wang *et al* [30] uses two superimposed 1D CNNs to jointly extract high-level temporal features. This work conducts segment based experiments. From table 9, the ACC, SEN, and SPE with those methods are all lower than those with STCNN by about 3%–8% in segment based experiments. Li *et al* [23] adopts EMD and co-spatial mode to extract features, and [37] adopts WT and Bi-LSTM. The SEN, SPE and ACC of [37] in segment based experiments and SEN in event based experiments are 1.95%, 1.37%, 1.06% and 2.66% lower than STCNN, respectively. Li *et al* [23] has 1.48% higher SEN but 2.36% lower SPE and 2.06% lower ACC than STCNN in segment based experiments. As in event based experiments, it has 0.32% higher SEN but 2.5% lower SPE than STCNN. Therefore, the comprehensive performance of the models is lower than STCNN. In short, based on the evaluation metrics such as SEN and SPE in the segment based and event based experiments,

Table 10. The comparison results of seizure prediction based on LOOCV in Siena and Kaggle.

		Event based						
Task	Methods	N	N+	SOP (min)	SPH (min)	SEN (%)	FPR (h^{-1})	Tp(min)
Prediction on Siena	Ensembling model [58]	/	/	/	/	93.18	/	/
	Geometric DL [59]	/	/	60	/	96.05	/	/
	Synchronization analysis [60]	/	/	15	/	100	2	7.78
	STCNN (proposed)	17	17	30	5	100.00	0.14	33.36
Prediction on Kaggle	STFT+CNN [40]	48	36	30	5	75	0.21	/
	Ensemble learning [24]	/	/	/	/	94.2	/	/
	Wavelet transforms [57]	111	103	30	1	91.5	/	/
	STCNN (proposed)	50	49	30	5	95.24	0.07	29.12

STCNN performs well in the seizure detection experiment, especially the SPE is higher than 99% in both experiments.

In the seizure prediction task on CHB-MIT, as shown in table 9, [26] uses a common spatial pattern to extract features in the CHB-MIT dataset, and a linear discriminator as a prediction model. Truong *et al* [40] extracts only the frequency domain features and uses CNN for prediction. Ozcan and Erturk [41] uses spectral band energy, statistical moments and Hjorth parameters as the feature extraction method, and a 3D-CNN modeling to realize early prediction. Gotman [57] uses a synchronous extraction of linearly modulated WTs. The SOP and SPH are set to 30 min and 1 min, respectively. Tang *et al* [42] uses a combination of CNN and RNN. An ensemble learning classifier composed of SVM, CNN and LSTM is used for learning in [24]. Parvez and Paul [28] extracts undulated global/local features and uses LS-SVM as classifier. Among the work mentioned above, our proposed method has the highest SEN of 98.13%. The FPR of our proposed method are 0.24 h^{-1}, 0.01 h^{-1}, 0.04 h^{-1} and 0.21 h^{-1} lower than those of [26, 28, 40, 41], respectively. In summary, the STCNN proposed in this article has good early prediction capabilities.

Table 10 is the comparison results of the LOOCV-based seizure prediction methods on Siena and Kaggle datasets. In Siena dataset, [58] uses statistical, wavelet, and entropy features and an ensemble classifier that consists of the Adaboost, random forest and DT. It achieves the SEN of 93.18%, which is 6.82% lower than our work. Dissanayake *et al* [59] uses geometric deep learning method, obtaining the SEN of 96.05%, which is 3.95% lower than our proposed model. Detti *et al* [60] has proposed a seizure prediction method based on the detection of synchronization patterns in the EEG. Even though it achieves the same SEN as our proposed method, the FPR of [60] is approximately 14 times higher than our method.

In the Kaggle dataset, [40] uses STFT and CNN as feature extractor and prediction model, respectively. It uses a total of 48 seizures, 36 successful warnings. In [57], 103 of the 111 seizures are successfully predicted. In [24], the averaged SEN on the Kaggle dataset is the best among those works, but it is still 1.04% lower than STCNN. Truong *et al* [40] achieves the FPR of 0.21 h^{-1}, which is 0.14 h^{-1} higher than our method. In conclusion, our proposed STCNN prediction algorithm learns both temporal and spatial information, resulting an SEN of 95.24% and an FPR of 0.07 h^{-1} on the Kaggle dataset. The model has good robustness and seizure predicting performance.

To prove the effectiveness of each module in STCNN, the ablation experiment is carried out on CHB-MIT dataset after removing the channel self-selection module, adaptive loss function, stream computing module, and K-of-N module. Both seizure detection and prediction abilities of different models are evaluated. The specific experimental results in table 11 prove the effectiveness of each module.

The experimental results of the weight distribution of different channels are shown in figure 12, where patients 01 to 05 are selected. It can be clearly seen from the figure that the proportion of each channel during epileptic seizures varies from person to person, which proves that the channel self-selection module can effectively deal with the channel importance in different patients. When it comes to seizure prediction task, the data demonstrate the superiority of STCNN and the effectiveness of each module.

5.4.3. CPV experimental results

Since the LOOCV-based experiment is still aimed at specific patients, it does not take into account the variability of different patients. CPV is carried out on both seizure detection and prediction. During the experiment, the data of a single patient is used as the test set, and the data of the rest of the patients is the training set. The final experimental result is the average of multiple patients. The experimental results on CHB-MIT and Siena dataset are shown in tables 12 and 13, respectively.

For the Kaggle dataset, we also carry out CPV-based experiments. But considering the difference between animal epilepsy data and human epilepsy data, CPV-based experiments are carried out in the same species. Specifically, for the epileptic seizure data of five dogs, the data of one dog is used as the test set in turn, and the data of the other dogs is used as

Table 11. Results of ablation experiments based on LOOCV seizure detection and prediction in CHB-MIT.

	Seizure detection						Seizure prediction				
	Segment based			Event based							
Methods	ACC (%)	SEN (%)	SPE (%)	N+	SEN (%)	SPE (%)	N	N+	SEN (%)	FPR(h^{-1})	Tp(min)
No channel selection	97.51	94.74	98.38	84	98.15	99.37	100	95	95.19	0.20	32.35
No adaptive focal loss	97.23	90.02	99.30	82	95.03	99.98	100	95	95.46	0.18	31.30
No stream	97.53	91.21	99.15	82	95.09	99.97	100	93	94.52	0.22	31.18
No K-of-N	99.55	95.84	99.86	84	98.15	99.86	100	98	98.13	0.25	33.88
STCNN (proposed)	99.55	95.84	99.86	84	98.15	100.00	100	98	98.13	0.15	33.01

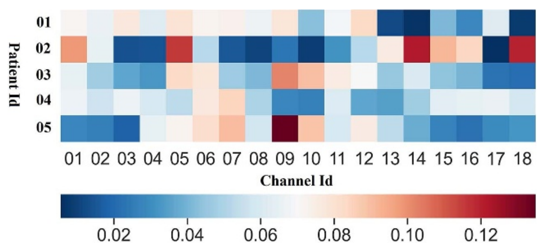

Figure 12. Example of seizure detection channel weights for some patients in CHB-MIT.

Table 12. Experimental results of seizure detection and prediction based on CPV in CHB-MIT.

	Seizure detection							Seizure prediction				
	Segment based				Event based							
Patient Id	N	ACC (%)	SEN (%)	SPE (%)	N+	SEN (%)	SPE (%)	N	N+	SEN (%)	FPR(h^{-1})	Tp(min)
1	7	99.61	95.63	100	7	100	100	5	5	100	0.00	32.83
2	2	99.62	95.24	100	2	100	100	3	3	100	0.00	32
3	7	99.79	97.79	100	7	100	100	7	5	71.43	0.00	32.98
4	3	98.42	82.52	100	2	66.67	100	3	3	100	0.83	34.17
5	5	99.6	97.42	99.93	5	100	100	5	5	100	0.00	33.07
6	/	/	/	/	/	/	/	7	7	100	0.00	26.1
7	3	99.59	96.3	100	3	100	100	3	3	100	1.28	34.17
8	4	95.87	79.73	100	4	100	100	4	4	100	0.00	33.83
9	4	99.87	99.01	99.93	4	100	100	4	4	100	2.45	22.75
10	7	99.57	96.23	99.81	7	100	100	7	6	85.71	0.00	21.02
11	2	99.12	97.59	100	2	100	100	2	2	100	0.93	33.92
13	5	98.66	82.22	99.61	5	100	100	7	5	71.43	0.81	28.95
14	/	/	/	/	/	/	/	6	6	100	0.00	32.81
15	14	97.25	90.36	99.83	13	92.86	100	13	10	76.92	0.14	29.15
16	/	/	/	/	/	/	/	3	3	100	0.00	27.67
17	3	99.5	96.64	99.81	3	100	100	3	3	100	1.45	33.72
18	5	99.56	94.68	99.92	5	100	100	4	2	50	0.00	24.96
19	3	99.7	96.67	100	3	100	100	2	2	100	0.94	34
20	5	99.29	94.87	99.39	5	100	100	5	4	80	0.00	27.57
21	3	99.71	95.45	100	3	100	100	4	4	100	0.00	31.17
22	3	100	100	100	3	100	100	3	3	100	0.00	28.17
23	4	99.55	95.41	99.86	4	100	100	4	4	100	0.00	29.04
Total	89	99.17	93.88	99.9	87	97.87	100	104	93	92.52	0.40	29.69

Table 13. Experimental results of seizure detection and prediction based on CPV in Siena.

	Seizure detection							Seizure prediction				
	Segment based				Event based							
Patient Id	N	ACC (%)	SEN (%)	SPE (%)	N+	SEN (%)	SPE (%)	N	N+	SEN (%)	FPR (h^{-1})	Tp(min)
PN01	1	100	100	100	1	100	100	1	1	100	0.00	33.81
PN03	2	99.03	95.65	100	1	50	100	2	2	100	0.00	35
PN05	2	93.68	66.67	100	1	50	100	3	3	100	0.00	35
PN06	3	96.55	100	100	1	33.33	100	2	2	100	4.50	34.25
PN07	1	97.98	100	100	1	100	100	1	1	100	0.00	35
PN09	1	98.96	100	100	1	100	100	3	3	100	0.00	34.69
PN11	1	100	100	100	1	100	100	1	1	100	0.00	34.9
PN12	2	93.55	58.54	100	1	50	100	3	3	100	0.00	32.91
PN13	3	98.64	100	98.72	1	33.33	100	3	3	100	0.00	33.94
PN16	2	96.9	81.4	99.71	1	50	100	2	2	100	0.00	34.16
PN17	2	99.21	96.43	100	1	50	100	2	2	100	0.00	34.99
Total	24	97.68	90.79	99.86	11	65.15	100	23	23	100	0.41	34.42

Table 14. Experimental results of seizure prediction based on CPV in Kaggle.

ID	N	N+	SEN (%)	FPR (h^{-1})	Tp (min)
Dog_1	4	4	100.00	0.51	\
Dog_2	7	5	71.43	0.64	19.64
Dog_3	12	12	100.00	0.27	24.38
Dog_4	16	14	87.5	0.77	21.22
Dog_5	5	5	100.00	0.01	25.3
Patient_1	3	3	100.00	0.42	30.78
Patient_2	3	3	100.00	0.00	26.67
Total	50	46	94.13	0.37	24.77

Table 15. Literature comparison results of seizure detection and prediction based on CPV in CHB-MIT.

		Segment based			Event based		
Task	Method	SEN (%)	SPE (%)	ACC (%)	N+	SEN (%)	SPE (%)
Seizure detection	Deep CNN [50]	90	91.65	98.05	/	99.46	/
	GAN+CNN [52]	72.11	95.89	84	/	90.75	/
	LSTM [44]	87.3	88.3	83.89	/	/	/
	STCNN (proposed)	93.88	99.9	99.17	87	97.87	100.00
		Segment based			Event based		
Task	Methods	ACC (%)	SEN (%)	SPE (%)	N	SEN (%)	FPR (h^{-1})
Seizure prediction	Transform learning [44]	58.55	/	/	/	/	/
	STCNN (proposed)	69.76	64.83	74.28	104	92.52	0.40

the training set. It should be noted that the number of channels of Dog-5 is 15. To keep all animal channels consistent, the zero channel filling operation is used. For the data of the two patients, one is used as the training set and the other is used as the test set. To align the number of channels, the zero channel filling operation is performed on the data of the 01 patient. The rest of the experimental settings are the same as the CHB-MIT dataset. The experimental results are shown in table 14. The averaged SEN and FPR are 94.13% and 0.37 h^{-1}, respectively, and the warning time is 24.77 min. The experimental results prove that STCNN has a good cross-patient warning effect on the Kaggle dataset.

The comparative results of the CPV-based experiment carried out on CHB-MIT are shown in table 15. In [50], CNN is used to extract high-level features from the original signal and frequency domain features. In the segment-based results, the SEN, SPE and ACC of STCNN are 3.88%, 8.25% and 1.12% higher than [50], respectively. In [52], adversarial generative network is used for data enhancement, and CNN is used to extract high-dimension features. However, its SEN, SPE and ACC are 21.77%, 4.01% and 15.17% lower than STCNN, respectively. In [44], attention mechanism and Bi-LSTM network are used to extract spatial features and temporal features, respectively. This work uses fragment experiments to evaluate the

Table 16. Results of ablation experiments based on CPV seizure detection and prediction in CHB-MIT.

	Seizure detection						Seizure prediction			
	Segment based			Event based						
Methods	ACC (%)	SEN (%)	SPE (%)	N+	SEN (%)	SPE (%)	N+	SEN (%)	FPR(h^{-1})	Tp(min)
No channel Selection	98.21	91.52	99.24	85	95.5	99.83	86	87.46	0.45	29.39
No adaptive focal loss	98.8	90.46	99.55	86	96.44	99.94	91	91.35	0.42	30.43
No stream	98.07	90.23	99.18	85	95.5	99.35	92	91.86	0.43	28.86
No K-of-N	99.17	93.88	99.9	87	97.87	99.9	94	93.82	0.65	33.88
STCNN (proposed)	99.17	93.88	99.9	87	97.87	100	93	92.52	0.40	29.69

experimental results, and its SEN, SPE and ACC on the CPV-based seizure detection task are all lower than 90%.

The CPV-based seizure prediction experiment is carried out based on the Siamese network in [44]. The accuracy of STCNN is above 10% higher than [44]. In summary, our proposed STCNN algorithm can learn and capture the differences of different patients, and has good detection and early warning performance.

Table 16 shows the CPV-based ablation experimental result on CHB-MIT dataset. Similar conclusion can be drawn as table 11.

6. Conclusion

Long-term detection and prediction of EEG in epilepsy patients has great research significance. In this paper, an seizure detection and prediction algorithm based on feature selection and spatio-temporal information has been proposed, which has realized the screening of feature space and reduced the difficulty of model fitting. In addition, to fulfill various medical application scenarios, seizure detection and prediction experiments based on leave-one-out method and cross-patient have been carried out. Besides CHB-MIT, the Siena dataset and Kaggle dataset are used in the experiment as well. The experimental results have proved the effectiveness and robustness of our proposed STCNN algorithm.

Data availability statement

The data that support the findings of this study are available upon reasonable request from the authors.

ORCID iD

Yuanming Zhang https://orcid.org/0000-0002-3224-7708

References

[1] Hussein A F, Hashim S J, Aziz A F A, Rokhani F Z B and Adnan W A W 2017 A real time ECG data compression scheme for enhanced Bluetooth low energy ECG system power consumption *J. Ambient Intell. Human. Comput.* **8** 1–14

[2] Huang H-Y and Lo P-C 2009 EEG dynamics of experienced zen meditation practitioners probed by complexity index and spectral measure *J. Med. Eng. Technol.* **33** 314–21

[3] Diykh M, Li Y and Wen P 2016 EEG sleep stages classification based on time domain features and structural graph similarity *IEEE Trans. Neural Syst. Rehabil. Eng.* **24** 1159–68

[4] Devarajan K, Bagyaraj S, Balasampath V, Jyostna E and Jayasri K 2014 EEG-based epilepsy detection and prediction *Int. J. Eng. Technol.* **6** 212–6

[5] Esteller R, Vachtsevanos G J, Echauz J R and Litt B 2001 A comparison of waveform fractal dimension algorithms *IEEE Trans. Circuits Syst.* I **48** 177–83

[6] Kumar Y, Dewal M L and Anand R S 2014 Epileptic seizures detection in EEG using DWT-based APEN and artificial neural network *Signal, Image Video Process.* **8** 1323–34

[7] Song Y, Crowcroft J and Zhang J 2012 Automatic epileptic seizure detection in EEGs based on optimized sample entropy and extreme learning machine *J. Neurosci. Methods* **210** 132–46

[8] Molla M K I, Hassan K M, Islam M R and Tanaka T 2020 Graph eigen decomposition-based feature-selection method for epileptic seizure detection using electroencephalography *Sensors* **20** 2020

[9] Hussain W *et al* 2019 Epileptic seizure detection with permutation fuzzy entropy using robust machine learning techniques *IEEE Access* **7** 182238–58

[10] Atyabi A, Shic F and Naples A 2016 Mixture of autoregressive modeling orders and its implication on single trial EEG classification *Expert Syst. Appl.* **65** 164–80

[11] Polat K and Güneş S 2008 Artificial immune recognition system with fuzzy resource allocation mechanism classifier, principal component analysis and FFT method based new hybrid automated identification system for classification of EEG signals *Expert Syst. Appl.* **34** 2039–48

[12] Cao J, Hu D, Wang Y, Wang J and Lei B 2022 Epileptic classification with deep-transfer-learning-based feature fusion algorithm *IEEE Trans. Cogn. Dev. Syst.* **14** 684–95

[13] Cao J, Zhu J, Hu W and Kummert A 2020 Epileptic signal classification with deep EEG features by stacked CNNS *IEEE Trans. Cogn. Dev. Syst.* **12** 709–22

[14] Li M, Chen W and Zhang T 2017 Application of MODWT and log-normal distribution model for automatic epilepsy identification *Biocybern. Biomed. Eng.* **37** 679–89

[15] Yuan Y, Xun G, Ma F, Suo Q, Xue H, Jia K and Zhang A 2018 A novel channel-aware attention framework for multi-channel EEG seizure detection via multi-view deep learning *2018 IEEE EMBS Int. Conf. on Biomedical & Health Informatics (BHI)* pp 206–9

[16] Zhang Z and Parhi K K 2016 Low-complexity seizure prediction from iEEG/sEEG using spectral power and ratios of spectral power *IEEE Trans. Biomed. Circuits Syst.* **10** 693–706

[17] Hassan K M, Islam M R, Nguyen T T and Molla M K I 2022 Epileptic seizure detection in EEG using mutual information-based best individual feature selection *Expert Syst. Appl.* **193** 116414

[18] Borhade R R and Nagmode M S 2020 Modified atom search optimization-based deep recurrent neural network for epileptic seizure prediction using electroencephalogram signals *Biocybern. Biomed. Eng.* **40** 1638–53

[19] Javidan M, Yazdchi M, Baharlouei Z and Mahnam A 2021 Feature and channel selection for designing a regression-based continuous-variable emotion recognition system with two EEG channels *Biomed. Signal Process. Control* **70** 102979

[20] Manjusha M and Harikumar R 2016 Performance analysis of knn classifier and k-means clustering for robust classification of epilepsy from EEG signals *2016 Int. Conf. on Wireless Communications, Signal Processing and Networking (WiSPNET)* pp 2412–6

[21] Siuly S, Alcin O F, Bajaj V, Sengur A and Zhang Y 2019 Exploring hermite transformation in brain signal analysis for the detection of epileptic seizure *IET Sci. Meas. Technol.* **13** 35–41

[22] Tang F-G, Liu Y, Li Y and Peng Z-W 2020 A unified multi-level spectral–temporal feature learning framework for patient-specific seizure onset detection in EEG signals *Knowl.-Based Syst.* **205** 106152

[23] Li C, Zhou W, Liu G, Zhang Y, Geng M, Liu Z, Wang S and Shang W 2021 Seizure onset detection using empirical mode decomposition and common spatial pattern *IEEE Trans. Neural Syst. Rehabil. Eng.* **29** 458–67

[24] Usman S M, Khalid S and Bashir S 2021 A deep learning based ensemble learning method for epileptic seizure prediction *Comput. Biol. Med.* **136** 104710

[25] Radman M, Moradi M, Chaibakhsh A, Kordestani M and Saif M 2021 Multi-feature fusion approach for epileptic seizure detection from EEG signals *IEEE Sens. J.* **21** 3533–43

[26] Alotaiby T N, Alshebeili S A, Alotaibi F M and Alrshoud S R 2017 Epileptic seizure prediction using CSP and LDA for scalp EEG signals *Comput. Intell. Neurosci.* **2017** 1240323

[27] Ein Shoka A A, Alkinani M H, El-Sherbeny A S, El-Sayed A and Dessouky M M 2021 Automated seizure diagnosis system based on feature extraction and channel selection using EEG signals *Brain Inform.* **8** 1

[28] Parvez M Z and Paul M 2016 Seizure prediction using undulated global and local features *IEEE Trans. Biomed. Eng.* **64** 208–17

[29] Bilal M, Rizwan M, Saleem S, Khan M M, Alkatheir M S and Alqarni M 2019 Automatic seizure detection using multi-resolution dynamic mode decomposition *IEEE Access* **7** 61180–94

[30] Wang X, Wang X, Liu W, Chang Z, Kärkkäinen T and Cong F 2021 One dimensional convolutional neural networks for seizure onset detection using long-term scalp and intracranial EEG *Neurocomputing* **459** 212–22

[31] Tanveer M A, Khan M J, Sajid H and Naseer N 2021 Convolutional neural networks ensemble model for neonatal seizure detection *J. Neurosci. Methods* **358** 109197

[32] Zeng D, Huang K, Xu C, Shen H and Chen Z 2021 Hierarchy graph convolution network and tree classification for epileptic detection on electroencephalography signals *IEEE Trans. Cogn. Dev. Syst.* **13** 955–68

[33] Zhao Y, Dong C, Zhang G, Wang Y, Chen X, Jia W, Yuan Q, Xu F and Zheng Y 2021 EEG-based seizure detection using linear graph convolution network with focal loss *Comput. Methods Programs Biomed.* **208** 106277

[34] Tuncer E and Doğru Bolat E 2022 Classification of epileptic seizures from electroencephalogram (EEG) data using bidirectional short-term memory (Bi-LSTM) network architecture *Biomed. Signal Process. Control* **73** 103462

[35] Saichand N V and Gopiya N S 2021 Epileptic seizure detection using novel multilayer LSTM discriminant network and dynamic mode Koopman decomposition *Biomed. Signal Process. Control* **68** 102723

[36] Li Y, Liu Y, Cui W-G, Guo Y-Z, Huang H and Hu Z 2020 Epileptic seizure detection in EEG signals using a unified temporal-spectral squeeze-and-excitation network *IEEE Trans. Neural Syst. Rehabil. Eng.* **28** 782–94

[37] Zhang Y, Yao S, Yang R, Liu X, Qiu W, Han L, Zhou W and Shang W 2022 Epileptic seizure detection based on bidirectional gated recurrent unit network *IEEE Trans. Neural Syst. Rehabil. Eng.* **30** 135–45

[38] Golmohammadi M, Ziyabari S, Shah V, Obeid I and Picone J 2018 Deep architectures for spatio-temporal modeling: automated seizure detection in scalp EEGs *2018 17th IEEE Int. Conf. on Machine Learning and Applications (ICMLA)* pp 745–50

[39] Shah V, Golmohammadi M, Ziyabari S, Von Weltin E, Obeid I and Picone J 2017 Optimizing channel selection for seizure detection *2017 IEEE Signal Processing in Medicine and Biology Symp. (SPMB)* pp 1–5

[40] Truong N D, Nguyen A D, Kuhlmann L, Bonyadi M R, Yang J, Ippolito S and Kavehei O 2018 Convolutional neural networks for seizure prediction using intracranial and scalp electroencephalogram *Neural Netw.* **105** 104–11

[41] Ozcan A R and Erturk S 2019 Seizure prediction in scalp EEG using 3D convolutional neural networks with an image-based approach *IEEE Trans. Neural Syst. Rehabil. Eng.* **27** 2284–93

[42] Tang L, Xie N, Zhao M and Wu X 2020 Seizure prediction using multi-view features and improved convolutional gated recurrent network *IEEE Access* **8** 172352–61

[43] Sun B, Lv J-J, Rui L-G, Yang Y-X, Chen Y-G, Ma C and Gao Z-K 2021 Seizure prediction in scalp EEG based channel attention dual-input convolutional neural network *Physica* A **584** 126376

[44] Yao X, Li X, Ye Q, Huang Y, Cheng Q and Zhang G-Q 2021 A robust deep learning approach for automatic classification of seizures against non-seizures *Biomed. Signal Process. Control* **64** 102215

[45] Krizhevsky A, Sutskever I and Hinton G E 2017 Imagenet classification with deep convolutional neural networks *Commun. ACM* **60** 84–90

[46] Fisher R A 1992 *Statistical Methods for Research Workers* (Springer)

[47] Peng H, Long F and Ding C 2005 Feature selection based on mutual information criteria of max-dependency, max-relevance and min-redundancy *IEEE Trans. Pattern Anal. Mach. Intell.* **27** 1226–38

[48] Chen T and Guestrin C 2016 XGBoost: a scalable tree boosting system *Proc. of the 22nd acm Sigkdd International Conference on Knowledge Discovery and Data Mining* pp 785–94

[49] Avcu M T, Zhang Z and Chan D W S 2019 Seizure detection using least EEG channels by deep convolutional neural network *ICASSP 2019 - 2019 IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)* pp 1120–4

[50] Hossain M S, Amin S U, Alsulaiman M and Muhammad G 2019 Applying deep learning for epilepsy seizure detection and brain mapping visualization *ACM Trans. Multimedia Comput. Commun. Appl.* **15** 1–17

[51] Currey D, Hsu D, Ahmed R, Venkataraman A and Craley J 2021 Cross-site epileptic seizure detection using convolutional neural networks *2021 55th Annual Conf. on Information Sciences and Systems (CISS)* pp 1–6

[52] Wei Z, Zou J, Zhang J and Xu J 2019 Automatic epileptic EEG detection using convolutional neural network with improvements in time-domain *Biomed. Signal Process. Control* **53** 101551

[53] Dissanayake T, Fernando T, Denman S, Sridharan S and Fookes C 2021 Deep learning for patient-independent epileptic seizure prediction using scalp EEG signals *IEEE Sens.* J. **21** 9377–88

[54] Gadhoumi K, Lina J-M, Mormann F and Gotman J 2016 Seizure prediction for therapeutic devices: a review *J. Neurosci. Methods* **260** 270–82

[55] Tsiouris K M, Pezoulas V C, Zervakis M, Konitsiotis S, Koutsouris D D and Fotiadis D I 2018 A long short-term memory deep learning network for the prediction of epileptic seizures using EEG signals *Comput. Biol. Med.* **99** 24–37

[56] Rout S K, Sahani M, Dora C, Kumar Biswal P and Biswal B 2022 An efficient epileptic seizure classification system using empirical wavelet transform and multi-fuse reduced deep convolutional neural network with digital implementation *Biomed. Signal Process. Control* **72** 103281

[57] Gotman J 1982 Automatic recognition of epileptic seizures in the EEG *Electroencephalogr. Clin. Neurophysiol.* **54** 530–40

[58] Kapoor B, Nagpal B, Jain P K, Abraham A and Gabralla L A 2022 Epileptic seizure prediction based on hybrid seek optimization tuned ensemble classifier using EEG signals *Sensors* **23** 423

[59] Dissanayake T, Fernando T, Denman S, Sridharan S and Fookes C 2021 Geometric deep learning for subject independent epileptic seizure prediction using scalp EEG signals *IEEE J. Biomed. Health Inform.* **26** 527–38

[60] Detti P, Vatti G and Zabalo Manrique de Lara G 2020 EEG synchronization analysis for seizure prediction: a study on data of noninvasive recordings *Processes* **8** 846