

# Deep multi-view feature learning for EEG-based epileptic seizure detection

Xiaobin Tian, Zhaohong Deng, *Senior Member, IEEE*, Kup-Sze Choi, *Member, IEEE*, Dongrui Wu, *Senior Member, IEEE*, Bin Qin, Jun Wan, Hongbin Shen, Shitong Wang

**Abstract**—Epilepsy is a neurological illness caused by abnormal discharge of brain neurons, where epileptic seizure can lead to life-threatening emergencies. By analyzing the encephalogram (EEG) signals of patients with epilepsy, their conditions can be monitored and seizure can be detected and intervened in time. As the identification of effective features in EEG signals is important for accurate seizure detection, this paper proposes a multi-view deep feature extraction method in attempt to achieve this goal. The method first uses fast Fourier transform (FFT) and wavelet packet decomposition (WPD) to construct the initial multi-view features. Convolutional neural network (CNN) is then used to automatically learn deep features from the initial multi-view features, which reduces the dimensionality and obtain the features with better seizure identification ability. Furthermore, the multi-view Takagi-Sugeno-Kang fuzzy system (MV-TSK-FS), an interpretable rule-based classifier, is used to construct a classification model with strong generalizability based on the deep multi-view features obtained. Experimental studies show that the classification accuracy of the proposed multi-view deep feature extraction method is at least 1% higher than that of common feature extraction methods such as principal component analysis (PCA), FFT and WPD. The classification accuracy is also at least 4% higher than the average accuracy achieved with single-view deep features.

**Index Terms**—EEG; Seizure detection; Multi-view; Feature extracting; Deep learning.

## I. INTRODUCTION

About 1-2% of people worldwide suffer from epilepsy. The unpredictability of epileptic seizures is the main cause of disability and even death. Although most people with epilepsy appear the same as able-bodied people during non-seizure periods, the spontaneity of seizures affects the quality of life seriously and can be fatal. Encephalogram (EEG) is an im-

portant means to record the activities of brain neurons. The electrophysiological signals generated by the neurons contain information that reflects the overall brain activities at the cerebral cortex. Since the electrical brain waves resulting from abnormal discharge of neurons during seizures are different from those generated during normal discharges, EEG can be used to detect seizures by identifying the characteristic brain signals, which is instrumental for predicting the onset of seizures and applying in-time interventions. On the other hand, accurate seizure detection is a key to automatic closed-loop treatment, where electrical stimulation, drug infusion, cooling or biofeedback, for example, can be applied to patients when seizure is detected. The types and the extent of closed-loop treatments may also be determined automatically based on the features extracted from the epileptic EEG signals that inform patient's condition. Besides, accurate seizure detection can be used to assess patient's condition during neurological surgery.

With the advance of machine learning, intelligent algorithms have been increasingly applied to improve the accuracy of EEG-based seizure detection. These algorithms include classification methods such as support vector machines (SVM) [6, 7], naive Bayes (NB) [9], neural networks [11] and fuzzy logic systems [13, 14], as well as feature extraction methods like principal component analysis (PCA) [15], wavelet packet decomposition (WPD) [6, 16] and high order crossings (HOC) [17]. For seizure detection, features are extracted from raw EEG signals to train classification models that can identify different states of epilepsy. Despite the proliferation of feature extraction and classification methods, extracting effective features with essential information for accurate seizure detection is still a critical challenge.

In recent years, deep learning emerges as an effective machine learning paradigm that has received extensive attention in feature learning. In deep learning, the idea of multi-level combinations is adopted to achieve complex feature representations through a large number of simple expressions. It learns and adjusts the weights at each layer of the neural network to obtain the features that are more likely to attain the desired output. That is, the input features are optimized at each layer to learn increasingly discriminative features. Deep learning techniques have been applied effectively for EEG signal processing. Different feature extraction methods are adopted to extract features from the EEG signals [18-20], whereby seizure detection is performed using convolutional neural network (CNN).

On the other hand, multi-view learning technology also finds applications in epileptic seizure detection. Various feature extraction methods are applied to obtain multi-view datasets of EEG signals [14, 21]. The datasets are then used for

---

This work was supported in part by the National Key Research Program of China under Grant No. 2016YFB0800803, the NSFC under Grant No. 61772239, the Jiangsu Province Outstanding Youth Fund under Grant No. BK20140001, by the National First-Class Discipline Program of Light Industry Technology and Engineering under Grant LITE2018-02, and the Hong Kong Research Grants Council under Grant No. PolyU 152040/16E. (*Corresponding author: Zhaohong Deng*)

X. B. Tian, Z. H. Deng, B. Qin, J. Wang, S. Wang are with the School of Digital Media, Jiangnan University and Jiangsu Key Laboratory of Digital Design and Software Technology, Wuxi 214122, China (e-mail: 6171611023@stu.jiangnan.edu.cn; dengzhaohong@jiangnan.edu.cn; 569398535@qq.com; qinbin\_sd@126.com; wangjunsy@126.com; wxwangst@aliyun.com)

K.S. Choi is with The Centre for Smart Health, the Hong Kong Polytechnic University (e-mail: thomasks.choi@polyu.edu.hk)

D. Wu is with the Key Laboratory of the Ministry of Education for Image Processing and Intelligent Control, School of Automation, Huazhong University of Science and Technology, Wuhan, China (e-mail: drwu@hust.edu.cn)

H. B. Shen is with the Institute of Image Processing and Pattern Recognition, Shanghai Jiao Tong University, and Key Laboratory of System Control and Information Processing, Ministry of Education of China, Shanghai, 200240, China (e-mail: hbshen@sjtu.edu.cn)

detecting epileptic seizure by using multi-view learning techniques. Multi-view learning technology is a learning paradigm with multi-view data that leverages the similarities and differences among the views. Multi-view learning algorithms can be divided into three types, namely, co-training, multiple kernel learning and subspace learning. Co-training algorithms perform alternate training of different views to maximize data consistency among the views. Multiple kernel learning algorithms train models using different kernels that are associated with different views, such that the kernels are combined linearly or nonlinearly to improve learning performance. Subspace learning algorithms consider data from multiple views from a common subspace and use different techniques to obtain the subspaces of the multi-view data. Although the multi-view learning approaches are significantly different, they are primarily based on the principles of consensus or complementarity to ensure successful multi-view learning.

In order to construct effective EEG features for seizure detection, this paper proposes a deep multi-view feature extraction method for EEG signals, which is based on multi-view and deep learning technology to construct classifier for seizure detection. The deep multi-view feature learning method proposed in this paper is summarized below.

- 1) *Construction of initial multi-view EEG features* Many methods have been proposed to construct features from EEG signals, each with specific advantages. In this paper, multi-view features are used to combine the advantages of these methods for seizure detection. The multi-view features constructed here include frequency domain features acquired by fast Fourier transform (FFT), time-frequency features acquired from WPD, and the original time domain features.
- 2) *Construction of deep multi-view features* To improve the effectiveness of the initial multi-view features, CNN is used to construct deep multi-view features based on the initial multi-view data. Compared with the initial multi-view features, the deep multi-view features extracted has lower dimensionality and higher discriminability.
- 3) *Construction of multi-view learning classifier* Finally, multi-view classifier learning technology is adopted to train the classification model based on the deep multi-view features learned by using CCN, which yields a more generalized multi-view classifier for seizure detection with EEG signals.

The proposed method is advantageous in that it utilizes not only deep learning but also multi-view learning, where shallow features of the EEG signals are first generated from different views to construct the multi-view deep features with deep learning. This can optimize the feature representation effectively for seizure detection. When multi-view deep features are integrated with multi-view learning to generate the multi-view classifier for seizure detection, the generalizability of the proposed method is further enhanced.

This paper is organized as follows. Section II describes the relevant technical background of the proposed algorithm. Section III proposes the seizure detection algorithm based on deep

multi-view feature learning. Experimental analyses are given in Section IV. Finally, conclusions and future work are described in Section V.

## II. RELATED WORK

This section provides a brief introduction of the work related to the proposed method, including the significance of using EEG signals for seizure detection, the application of machine learning and deep learning in epileptic seizure detection, and a review of multi-view learning techniques.

### A. Seizure detection using EEG signals

EEG signals reflect the activities of brain neurons and have been widely used in the fields of epileptic seizure detection. Automatic algorithms are developed to analyze EEG signals so that the information contained inside the signals is converted into distinctive outputs for determining different states of epilepsy, e.g. whether seizure is about to occur or is occurring. A major goal of EEG-based epilepsy detection is to make such conversion as fast and accurate as possible. A variety of epilepsy detection algorithms have been proposed in recent years [8, 10, 12, 14, 22]. In [22], a method adopting transfer learning and semi-supervised learning is used to classify the status of epilepsy with EEG signals. In [14], multi-view learning technology is used for automatic recognition of epileptic EEG signals based on shallow features. In [8, 10, 12], deep learning techniques are used to automatically classify epilepsy by using deep features. These algorithms apply multi-view learning and deep learning techniques separately and demonstrate promising performance. This suggests that the integration of multi-view learning and deep learning is a promising approach to further increase the accuracy of EEG-based epilepsy detection.

### B. EEG epilepsy detection based on machine learning

Machine learning techniques have received considerable attention for automatic seizure detection with EEG signals. In [23], four methods – random forest (RF), decision tree (DT) algorithm C4.5, SVM+RF, SVM+C4.5 – are used to detect seizure, where RF yields the best classification results. In [24], approximate entropy and sample entropy extracted by WPD are used as features, whereby SVM and extreme learning machine are used as classifiers for epileptic seizure detection. Besides, WPD and kernel PCA (KPCA) are adopted in [25] for dimensionality reduction, followed by using Takagi-Sugeno-Kang (TSK) fuzzy logic system as the classifier. Advancement in automatic seizure detection is achieved with these algorithms.

With the development of deep learning, classical algorithms such as stacked auto-encoder [26-28], deep belief networks [29-31], CNN [32-34] and recurrent neural networks [35-38] have been applied to biomedicine effectively. Attempts have been made to use CNN to process EEG signals. For example, CNN is used to perform one-dimensional convolution on the original raw EEG signals to predict epileptic seizure [18]. In [19], EEG signals are transformed to the frequency domain by Fourier transform and classified using CNN. Moreover, EEG

signals are encoded into pixel colors through feature processing to form a two-dimensional pattern [20] for epileptic seizure prediction. In general, deep learning techniques have demonstrated promising performance in EEG signal processing.

### C. Multi-view learning technology

Multi-view learning is a machine learning paradigm developed for datasets with features from different views. Since multi-view cooperation can effectively utilize both the independence of each view and the correlation between different views in the learning process, a better modeling effect is achieved when compared with models obtained based on single-view data. For dataset without natural segmentation of features from different views, it is possible to manually construct different views for multi-view learning to achieve better learning effect than methods that use the original features from a single view only.

Canonical correlation analysis (CCA) [39, 40], co-training [41, 42], sparse multi-view SVM [43] are common algorithms that have been applied to different multi-view data application scenarios. In fact, multi-view learning techniques have been applied to detect epileptic seizure with EEG signals. Tensor decomposition is performed on multi-view features to extract new features which can enhance classification performance for seizure detection [21]. Multi-view fuzzy system classifier is proposed in [14] for epileptic seizure detection with EEG signals, where multi-view data are obtained by adopting different feature extraction methods and utilizing TSK fuzzy system as classifier to train the detection model.

## III. EPILEPSY DETECTION BASED ON EEG SIGNALS USING DEEP MULTI-VIEW FEATURE LEARNING

In the proposed method, the initial multi-view features are constructed using FFT and WPD with the original raw EEG signals. The features are then used to generate deep multi-view features using deep neural networks and CNN. Finally, a multi-view classification model is trained using the deep multi-view features. In this section, the overall framework of the proposed method is described in subsection A. The details of the initial feature representation from three different views of the EEG signals are given in subsection B. The structure of the CNN used to extract deep features from the different views is introduced in subsection C. The procedure of constructing the classifier using the multi-view Takagi-Sugeno-Kang fuzzy system (MV-TSK-FS) and deep multi-view features is described in subsection D.

### A. Framework of Deep Multi-View Learning for Epilepsy Detection

The framework of the proposed method is shown in Fig. 1. It consists of three core components, i.e., the construction of the initial multi-view feature, the automatic learning of deep multi-view features and the training of the multi-view classifier.

### B. Initial Multi-View Features Construction

The original EEG signals are time domain signals. Although

the signals contain some useful time features, features in other domains can be used to extract more discriminative information. Transformation of the EEG signals from the time domain can be conducted using traditional feature extraction techniques. To obtain frequency domain features, Fourier transform can be used to transform the signals from the time domain to the frequency domain. Any continuous signals can be transformed to the frequency domain provided that the signals are periodic. Furthermore, wavelet transform can be used to transform the signals into the time-frequency domain and obtain the instantaneous frequency at each time point while retaining the time features of the signals. When calculating the instantaneous frequency, since the signal length is very short, the frequency features obtained by wavelet transform is more accurate.

In the proposed method, features are extracted from the original EEG signals from three views, i.e., time domain, frequency domain and time-frequency domain, to construct the initial multi-view EEG data.

#### 1) Time Domain

The original EEG signals are time-series signals that change with time. A discrete point in the signals represents the energy intensity at a certain time, or the measured voltage value at that moment. This paper uses the original EEG signals as the features of the time domain view. Fig. 2 plots a channel of the EEG signals from the time domain, where the horizontal axis is time and the vertical axis is the amplitude of the signal.

#### 2) Frequency Domain

EEG signals can be considered as the superposition of signals of different frequencies. The frequency range of interest spans from 0 Hz to 60 Hz. It is divided into six frequency bands: Delta-1 (0-2 Hz), Delta-2 (2-4 Hz), Theta (4-8 Hz), Alpha (8-15 Hz), Beta (15-30 Hz) and Gamma (30-60 Hz).

In this study, in order to reduce the number of features in the frequency domain and to preserve the original features, the sampling interval after discrete Fourier transform is set to 1 Hz. The studies in [44, 45] show that the features of epileptic seizure mainly appear between 4Hz and 30Hz. Hence, we have adopted the frequency band between 4 Hz and 30 Hz in our experiments to construct the initial features in the frequency view. Fig. 3 shows the frequency domain features obtained by transforming the time domain signals shown in Fig. 2, where the horizontal and the vertical axis represent frequency and amplitude respectively.

#### 3) Time-frequency Domain

Time-frequency features describe the instantaneous frequency of the signals at various time points. Wavelet decomposition is a commonly used method to transform time domain signals to the time-frequency domain, where the trigonometric function base of Fourier transform is changed to the wavelet function base. There are two variables in the wavelet functions, i.e.,  $a$  and  $\tau$ , where  $a$  controls the expansion and contraction of the wavelet transform, i.e., frequency; and  $\tau$  controls the translation of the wavelet transform, i.e., time. By controlling these two variables, wavelet transform can realize adap-

tive time-frequency signal analysis on multiple scales.

WPD is a common wavelet transform method [24, 46]. It is used in this paper to obtain time-frequency features of the EEG signals. The wavelet basis function adopted is Daubechies (dbN), which has good regularity. Since the order of wavelet basis functions increases with the smoothness of the functions and the localization ability in the frequency domain, the higher the order of the functions, the better the results of band division. However, high order vanishing moments would lead to increase in computation time which is undesirable and deteriorates real-time performance. In this paper, the order of the wavelet functions is set to 4. Since the signals to be transformed into the time-frequency domain are three-dimensional data (channel \* frequency \* time), the number of features after transformation is usually very high. To reduce the number of features and the computation time, a large sampling interval of 2 Hz is used in the study. As for Fourier transform, the frequency range between 4 Hz and 30 Hz is only considered [44, 45] as discussed above. Fig. 4 shows the time-frequency domain features extracted from the time domain signals shown in Fig. 2, where the number of decomposition layers of the wavelet transform is set to 6. In the figure, the horizontal axis is frequency and the vertical axis is time. The amplitude at different time and frequency is represented using the colors shown in the legend on the right.

### C. Deep Multi-View Feature Learning

In this paper, CNN is used to extract features from EEG signals from different views. The initial features in the time, frequency and time-frequency domains are first constructed following the approach discussed previously. Three different CNNs are then constructed to extract deep features from the initial features. The CNN uses the results of the output layer to calculate the approximation error and performs back propagation to update the network parameters during training.

Since the feature vector calculated by the penultimate layer of the network only passes through one fully connected layer before reaching the output layer, the output of the penultimate layer is also considered to be optimized when optimizing the network structure according to the results of the output layer. Besides the penultimate layer, the CNN also learns a better feature expression of the middle layers through the training process. We have chosen the output of the penultimate layer as the deep features learned by the CNN. The deep features thus obtained not only have lower dimensionality than the original features, but also possess better discrimination ability to enhance the generalizability of the subsequent classification model.

Figs. 5-7 show the CNN architecture used for extracting deep features from the three views. In the figures, the notation  $k@m*n$  indicates the feature map at each layer of the network, where  $k$  is the number of feature maps of the layer, and  $m*n$  is the size of the feature map. The two-dimensional convolution kernels of the network are represented by the notation  $k*m*n$ , where  $k$  is the number of convolution kernels,  $m*n$  is the size of the convolution kernel. Moreover,

the three-dimensional convolution kernels are represented by  $k*m*n*l$ , where  $k$  is the number of convolution kernels, and  $m*n*l$  is the size of convolution kernel. The default step size of the convolution kernel is set to 1. The input of the CNN is the original features from each view, and the output is a vector with length equal to 2, corresponding to a seizure or non-seizure sample. If the sample is a seizure sample, the values for the first and the second dimension are 0 and 1 respectively. The values are 1 and 0 respectively for a non-seizure sample.

#### 1) Time Domain Deep Feature Extraction Network

The CNN architecture shown in Fig. 5 is used for deep feature extraction from the view of the time domain, which includes a total of 4 convolution layers and 3 fully connected layers. In the time domain, the multi-channel EEG signals can be represented as a two-dimensional matrix of channel number and time. Here, the input of the CNN is a two-dimensional matrix of size  $23 * 256$ , i.e., the number of channels is 23.

The first convolution layer of the CNN is shown in Fig. 5. It adopts a  $1*128$  convolution kernel and the step size is 1. A feature map with size of  $23*129$  is then obtained. The second convolution layer of the CNN adopts thirty  $1*65$  convolution kernels and the step size is 1. Thirty feature maps of size  $23*65$  are then obtained. The subsequent convolutional layers are constructed by the same token. The fifth layer of the CNN is a fully connected layer, where the 10 feature maps of size  $13*16$  are first converted into a  $1*2080$  vector, and further into a  $1*1024$  vector. The fully connected layers that follow are constructed in the same way.

#### 2) Frequency Domain Deep Feature Extraction Network

Fig. 6 shows the CNN architecture used for deep feature extraction from the view of the frequency domain. It includes 2 convolution layers and 3 fully connected layers. The initial multi-channel EEG features in the frequency domain can be represented as a two-dimensional matrix of the number of channels and the number of frequencies. Here, the input of the network is a two-dimensional matrix of size  $23 * 27$ , i.e., the number of channels is 23 and the number of sampling points in frequency are 27. The operations of each layer in the CNN is similar to that of the time domain deep feature extraction network described above.

#### 3) Time-frequency Domain Deep Feature Extraction Network

Fig. 7 shows the CNN architecture used for deep feature extraction from the view of the time-frequency domain, which includes a total of 4 three-dimensional convolutional layers and 3 fully connected layers. The initial features of the multi-channel EEG signals in the time-frequency domain can be represented as a three-dimensional matrix of time, number of channels and number of frequencies. In Fig. 7, the input of the network is a three-dimensional matrix with the size of  $256 * 23 * 14$ , i.e., the number of sampling points in time is 256, the number of channels is 23 and the number of frequencies is 14. The operations of each layer in the CNN is also similar to that

of the time domain deep feature extraction network described above. Since the input is a three-dimensional matrix, the convolution kernel used is also three dimensional. The three-dimensional convolution operation is an extension of the two-dimensional operation in three-dimensional space. The operation is the same as that of the two-dimensional convolution operation.

In Figs. 5-7, the  $\tanh$  function is used in the three CNNs as the activation function to implement nonlinear transformation. The  $\tanh$  function is given by

$$\tanh(\mathbf{x}) = \frac{e^{\mathbf{x}} - e^{-\mathbf{x}}}{e^{\mathbf{x}} + e^{-\mathbf{x}}}.$$

Since  $\tanh(x) \in [-1, 1]$  and the mean value of the  $\tanh$  function is 0, it is more suitable for practical applications than the sigmoid function.

The CNNs adopt the *softmax* cross entropy as the loss function, which is defined as follows.

$$\text{softmax: } a_{ji} = \frac{e^{z_{ji}}}{\sum_{k=1}^K e^{z_{jk}}}$$

$$\text{loss: } L = \frac{1}{N} \sum_{j=1}^N \sum_{i=1}^K -y_{ji} \log a_{ji}$$

The *softmax* function calculates the probability  $a_{ji}$  that the  $j$ th sample belongs to the  $i$ th class,  $z_{ji}$  is the output of the  $j$ th sample at the  $i$ th output node.  $\mathbf{y}_j = \{y_{ji} | i = 1, \dots, K\}$  is the true label of the  $j$ th sample,  $K$  is the total number of classes,  $N$  is the total number of samples, and  $L$  is the total average cross entropy loss of the  $N$  samples.

#### D. Classifier Training based on Multi-View Learning

In the previous sections, we have focused on the construction of deep multi-view features using deep learning techniques. In order to exploit the potential of integrating deep feature learning and multi-view learning, we need to train a multi-view classifier based on the multi-view deep features. Multi-view-learning based classification methods have been effectively used for EEG signal recognition [14, 21]. One example is the Multi-view TSK Fuzzy system (MV-TSK-FS) [14], which is proposed to generate multi-view classifier for EEG-based epilepsy detection by using the traditional shallow multi-view features to train TSK fuzzy systems.

In practice, any multi-view-learning based classifiers can be adopted to implement EEG-based epilepsy detection based on the multi-view deep features constructed by the abovementioned strategy. In this work, MV-TSK-FS [14, 47] is adopted due to its distinctive characteristics of good interpretability and uncertainty knowledge representation ability, which are inherited from fuzzy rules and fuzzy logic inference. Note that as a component of the proposed multi-view deep features-based epilepsy detection method, MV-TSK-FS is only a feasible choice. It can be replaced with other multi-view classification methods in the future to enhance the effectiveness. Details of MV-TSK-FS can be found in Part I of the *Supplementary Materials* section. Meanwhile, the algorithm of the

proposed seizure detection is detailed in Table S1 of Part II in that section.

## IV. EXPERIMENTAL STUDIES

This section presents the experiments conducted in this study and is arranged as follows:

1) Subsection IV-A gives the specific details of the dataset, discusses the oversampling method used to cope with data imbalance, and describes the performance indices.

2) Subsection IV-B reports the performance of the proposed algorithm and makes comparison with the related methods. The sliding window and cross-validation strategy adopted are also analyzed.

3) Subsection IV-C compares the effect of the deep feature extraction method with that of the traditional shallow feature extraction methods in order to demonstrate the advantages of the former.

4) Subsection IV-D presents the experiments conducted to verify that the method using the multi-view deep features are advantageous over the methods that only use single-view deep features.

5) Subsection IV-E presents the experiments conducted to demonstrate the effectiveness of the proposed method from the perspective of detection delay.

#### A. Dataset, Data Preprocessing and Performance Indices

The CHB\_MIT dataset provided by the Boston Children's Hospital is adopted for the experiments. The dataset contains EEG signals collected from 23 patients. The data are organized into 24 groups, each group comprising of EEG signals acquired from a patient for more than 12 consecutive hours. Note that the 21st group of the data is the records of the first patient collected a few years later. The EEG data used in our experiments were collected with an 18-channel EEG device by the Boston Children's Hospital. Based on the data, the hospital generated multi-channel data with 23 channels, which are open and have been used extensively for research [1,4,5]. Fig. 8 shows the raw data of the CHB-MIT dataset over a certain period of time in a data group. It contains continuous signals of the 23 channels, where each of these channels is derived from the difference between two channels of the original signals.

Since the data are highly imbalanced, i.e., the ratio of the number of seizure samples to that of non-seizure samples is 1:100, the evaluation would suffer from serious over-fitting problem if all the data are used directly. To reduce the imbalance between the non-seizure and seizure data, part of the non-seizure data is abandoned, as performed in [3-5]. Furthermore, over-sampling is applied to the seizure data, where sliding window is used to capture the data segment. The interval between two samples is less than the width of one sliding window. Since there is a repetition segment between two adjacent samples, the number of epileptic data segments is also increased. In our experiments, the EEG signals are split into multiple signal segments of 1s in length, each containing 256 sample points. The data are shown in Table S2 of Part III in the *Supplementary Materials* section.

Three common performance indices are used here for experimental analysis, i.e., accuracy, sensitivity and specificity [48, 49]. They are defined as follows,

$$\text{Accuracy} = (TN + TP) / (TP + TN + FP + FN),$$

$$\text{Sensitivity} = TP / (TP + FN),$$

$$\text{Specificity} = TN / (TN + FP),$$

where  $TP$ , denoting true positive, is the number of seizure segments detected with seizure segments;  $FN$ , denoting false negative, is the number of non-seizure segments detected with seizure segments;  $FP$ , denoting false positive, is the number of seizure segments detected with non-seizure segments, and  $TN$ , denoting true negative, is the number of non-seizure segments detected with non-seizure segments. Accuracy is the proportion of correctly classified of seizure and non-seizure segments. Sensitivity is the proportion of correctly classified seizure segments. A classifier with high sensitivity has outstanding performance in identifying seizure segments. Conversely, specificity is the proportion of correctly classified non-seizure segments. A classifier with high specificity is good at identifying non-seizure segments.

### B. Epilepsy Detection Performance

In our experiments, the proposed method is evaluated by conducting the experiments on each group of data using  $k$ -fold cross-validation strategy, where the data are divided into  $k$  subsets of the same size to ensure consistent data distribution. Validation is repeated  $k$  times such that at each time, one of the subsets is taken as the testing set while the rest are used as the training set. The final validation result is given by the mean of the results in each individual validation. This strategy is effective in avoiding sampling bias and thus obtaining more convincing experimental results. In this paper, five-fold cross-validation strategy is adopted. The data of each patient are evenly divided into five parts, each containing the same amount of seizure and non-seizure data. Four of the five parts are used as the training set, and the part left behind is used as the validation set. The above procedure is repeated five times to obtain the mean and standard deviation of the results.

Table I shows the average results of the experiments obtained by performing five-fold cross-validation on the 24 groups of data, in terms of the three indices described above. The average accuracy, sensitivity and specificity over all the data groups are  $98.33 \pm 0.18\%$ ,  $96.66 \pm 0.14\%$  and  $99.14 \pm 0.14\%$  respectively, indicating that the performance indices of the proposed method are all above 90%, except the 12th group. Furthermore, among the 24 data groups, the accuracy, sensitivity and specificity of the proposed method exceed 99% in 12, 7 and 19 groups respectively, showing that the performance is particularly outstanding from the perspectives of accuracy and specificity.

We further analyze the proposed method by making comparison with the feature extraction methods proposed in recent studies [1-5,8,10,12] that exploit deep learning and use the same CHB-MIT dataset for seizure detection. Table II gives the feature extraction methods, experimental settings and the

classification performance (in the table, NA meaning “not applicable”, indicating that the data is not used in the study). Note that the studies in [1-3] do not adopt cross-validation strategy, which precludes a clearer analysis of the performance of the seizure detection methods. A possible reason for this is due to the lack of seizure samples in the dataset that cross-validation would result in even fewer seizure samples for the validation set, thus creating large discrepancy between the test results and the real case. To avoid the problem, 25% of the samples are only used for training in [3-5]. In [8,10,12], different over-sampling methods have been adopted to increase the number of seizure samples. As a result of data imbalance, the accuracy and sensitivity of most of the algorithms under comparison are low. On the contrary, the proposed algorithm exhibits better accuracy and sensitivity while maintaining the same level of specificity, as indicated in Table II.

The dataset of Patient 12 (the 12th group) corresponds to a special case that the dataset has been excluded from the experiments in many previous studies [1-5]. In [8,10,12] and our study, among the 24 groups of data used in the experiments, the performance of this 12th group is found to be particularly inferior. This suggests that the EEG signals of the 12th group is highly unstable and irregular which introduces a lot of interference to the learning process of the algorithms, thus leading to poor performance in most cases.

In order to verify the effectiveness of using the strategy of sliding window to capture the seizure segments, we compare the sliding window method with the popular over-sampling method SMOTE [53]. Table III shows the average performance of sliding window versus SMOTE for all the data groups. It can be seen that these two methods have similar performance, thus justifying that the sliding window strategy can be used effectively for over-sampling.

Besides the five-fold cross validation strategy, we also use the strategy of percentage split to construct the training and testing sets. Table IV shows that the performance of using different percentage to split the testing set and the training set. It can be seen that the proposed algorithm is more stable when 90% and 75% of the data are used respectively as the training set. When it is further reduced to 50%, the performance of the algorithm decreases slightly.

### C. Effectiveness of Deep Feature Extraction

To evaluate the effectiveness of the proposed method in extracting deep features, comparison is made by using different feature extraction methods to extract the features and then applying the classifiers SVM, K-nearest neighbors (KNN), NB, DT and TSK-FS respectively to assess the performance of the feature extraction methods. The results are shown in Tables V-VII. For TSK-FS, the evaluation for the feature extraction method WPD is not conducted since the features obtained have high dimensionality. The results indicate that the deep features extracted from the frequency and time-frequency domain yield the best accuracy, specificity and sensitivity for all the classifiers. While the time domain deep features produce the best average sensitivity for all the classifiers, their effect on the enhancement of accuracy and specificity is moderate. It

can be concluded that the deep features obtained by using the deep feature extraction method are a better feature representation than those obtained by using the conventional methods from the three views, and therefore resulting in better classification performance.

#### D. Effectiveness of Multi-View Model Learning

Table VIII compares the effect of different single-view classifiers and the proposed MV-TSK-FS multi-view learning classifier with the deep features extracted from the three views. The results show that adopted MV-TSK-FS classifier achieves the best results in terms of accuracy, sensitivity and specificity, indicating that the multi-view classifier MV-TSK-FS also plays a role in improving the classification performance of the proposed method.

#### E. Performance on Detection Delay

Detection delay reflects the real-time performance of EEG seizure detection algorithms [49]. Here, two indices are used to evaluate detection delay – the proportion of successful detection  $S$  and the average detection delay  $EO_{\text{Latency}}$ .  $S$  is defined by

$$S = \frac{1}{N_s} \sum_{m=1}^{N_s} S_m,$$

where  $N_s$  is the total number of seizure records that are used for evaluation,  $S_m \in \{0, 1\}$  is used to indicate whether a record  $m$  is detected as a seizure record, with 1 and 0 specifying respectively that seizure is detected or not.  $EO_{\text{Latency}}$  is defined by

$$EO_{\text{Latency}} = \frac{1}{K} \sum_{m=1}^{N_s} S_m * EO_{\text{Latency},m},$$

where  $K$  is the total number of seizure records that are successfully detected and  $EO_{\text{Latency},m}$  is the detection delay of record  $m$ .  $EO_{\text{Latency},m}$  is calculated with the procedure below.

First, the seizure segment of the first second is used for testing. If seizure is not detected, then the seizure segment in the next second is tested. The above process is repeated until seizure is detected, or when the segments in the first ten seconds are tested. If seizure is detected with the segment in the  $i$ th second,  $EO_{\text{Latency},m} = i$ . Otherwise, a large constant value is assigned to  $EO_{\text{Latency},m}$  to indicate that seizure is not successfully detected with all the segments in the first ten seconds.

The experimental results are shown in Table IX. It can be seen that the average detection delay is only 1.0431s, showing that the proposed method has a low detection delay. Besides, the proportion of successful detection is 99.95%, indicating that almost all epileptic events can be successfully detected within 10 seconds.

## V. CONCLUSIONS

This study proposes a deep multi-view feature learning approach to develop epileptic seizure detection method with EEG signals. Multi-view classifier is used to integrate deep

features from different views to further enhance the detection performance. The study demonstrates that the deep feature extraction method and the introduction of multi-view learning are instrumental for epilepsy detection with EEG signals. Experimental studies show that the deep features concerned in this study can increase the detection performance when compared with traditional feature extraction methods.

Despite the promising results, there exist issues that deserve further study. For example, the present study only concerns feature from the time, frequency and time-frequency domains, while EEG signals also contain other useful features such as statistical features and nonlinear features, which can be exploited to enhance seizure detection performance. How to make effective use of these features is an interesting work. Although the multi-view classifier MV-TSK-FS adopted in the proposed epilepsy detection method exhibits better performance than single-view classifiers, there is still room for improvement, e.g. further improving the multi-view classifier by using more effective multi-view learning mechanism. In addition, while the focus of the study is to detect epileptic seizure, it can indeed be extended to the prediction of epileptic seizure. In-depth investigations will be conducted along these research directions. The codes are released and downloadable from <https://github.com/Txiaobin/deep-multi-view-feature-learning>.

## REFERENCES

- [1] N. Rafiuddin, Y. Uzzaman Khan, and O. Farooq, "Feature extraction and classification of EEG for automatic seizure detection," in *International Conference on Multimedia, Signal Processing and Communication Technologies*, 2011, pp. 184-187.
- [2] Y. U. Khan, N. Rafiuddin, and O. Farooq, "Automated seizure detection in scalp EEG using multiple wavelet scales," in *IEEE International Conference on Signal Processing, Computing and Control*, 2012, pp. 1-5.
- [3] S. Kiranyaz, T. Ince, M. Zabihi, and D. Ince, "Automated patient-specific classification of long-term Electroencephalography," *Journal of Biomedical Informatics*, vol. 49, pp. 16-31, 2014.
- [4] M. Zabihi, S. Kiranyaz, A. B. Rad, A. K. Katsaggelos, M. Gabbouj, and T. Ince, "Analysis of High-Dimensional Phase Space via Poincaré Section for Patient-Specific Seizure Detection," *IEEE Transactions on Neural Systems & Rehabilitation Engineering A Publication of the IEEE Engineering in Medicine & Biology Society*, vol. 24, pp. 386-398, 2015.
- [5] K. Samiee, S. Kiranyaz, and M. Gabbouj, "Long-term epileptic EEG classification via 2D mapping and textural features," *Expert Systems with Applications An International Journal*, vol. 42, pp. 7175-7185, 2015.
- [6] Y. Zhang, Y. Zhang, J. Wang, and X. Zheng, "Comparison of classification methods on EEG signals based on wavelet packet decomposition," *Neural Computing & Applications*, vol. 26, pp. 1-9, 2015.
- [7] K. Fu, J. Qu, Y. Chai, and Y. Dong, "Classification of seizure based on the time-frequency image of EEG signals using HHT and SVM," *Biomedical Signal Processing & Control*, vol. 13, pp. 15-22, 2014.
- [8] X. Yao, X. Li, Q. Ye, Y. Huang, Q. Cheng, and G. Zhang, "A Robust Deep Learning Approach for Automatic Seizure Detection," *arXiv preprint arXiv:1812.06562*, 2018.
- [9] A. Bhattacharyya and R. B. Pachori, "A multivariate approach for patient specific EEG seizure detection using empirical wavelet transform," *IEEE Trans Biomed Eng.*, vol. PP, pp. 1-1, 2017.
- [10] X. Yao, Q. Cheng, and G.-Q. Zhang, "A Novel Independent RNN Approach to Classification of Seizures against Non-seizures," *arXiv preprint arXiv:1903.09326*, 2019.
- [11] Q. Yuan, W. Zhou, S. Li, and D. Cai, "Epileptic EEG classification based on extreme learning machine and nonlinear features," *Epilepsy Research*, vol. 96, pp. 29-38, 2011.
- [12] H. Ke, D. Chen, X. Li, Y. Tang, T. Shah, and R. Ranjan, "Towards brain big data classification: epileptic EEG identification with a lightweight VGGNet on global MIC," *IEEE Access*, vol. 6, pp. 14722-14733, 2018.

- [13] Z. Deng, P. Xu, L. Xie, K. S. Choi, and S. Wang, "Transductive Joint-Knowledge-Transfer TSK FS for Recognition of Epileptic EEG Signals," *IEEE Transactions on Neural Systems & Rehabilitation Engineering*, vol. PP, pp. 1-1, 2018.
- [14] Y. Jiang, Z. Deng, F. L. Chung, G. Wang, P. Qian, K. S. Choi, *et al.*, "Recognition of Epileptic EEG Signals Using a Novel Multiview TSK Fuzzy System," *IEEE Transactions on Fuzzy Systems*, vol. 25, pp. 3-20, 2017.
- [15] J. R. Villar, M. Menéndez, E. D. L. Cal, J. Sedano, and V. M. González, "Identification of abnormal movements with 3D accelerometer sensors for seizure recognition ☆," *Journal of Applied Logic*, 2016.
- [16] Y. Sugianela, Q. L. Sutino, and D. Herumurti, "Eeg Classification For Epilepsy Based On Wavelet Packet Decomposition And Random Forest," *Jurnal Ilmu Komputer dan Informatika*, vol. 11, pp. 27-33, 2018.
- [17] S. Nasehi, H. Pourghassem, and A. Etesami, "Online Epilepsy Diagnosis Based on Analysis of EEG Signals by Hybrid Adaptive Filtering and Higher-order Crossings," in *International Conference on Intelligent Computation and Bio-Medical Instrumentation*, 2011, pp. 192-195.
- [18] A. Antoniadis, L. Spyrou, C. C. Took, and S. Sanei, "Deep learning for epileptic intracranial EEG data," in *IEEE International Workshop on Machine Learning for Signal Processing*, 2016, pp. 1-6.
- [19] H. Cecotti and A. Graeser, "Convolutional Neural Network with embedded Fourier Transform for EEG classification," in *International Conference on Pattern Recognition*, 2008, pp. 1-4.
- [20] P. Mirowski, D. Madhavan, Y. Lecun, and R. Kuzniecky, "Classification of patterns of EEG synchronization for seizure prediction," *Clinical Neurophysiology*, vol. 120, pp. 1927-1940, 2009.
- [21] L. Spyrou, S. Kouchaki, and S. Sanei, "Multiview Classification and Dimensionality Reduction of Scalp and Intracranial EEG Data through Tensor Factorisation," *Journal of Signal Processing Systems*, pp. 1-12, 2016.
- [22] Y. Jiang, D. Wu, Z. Deng, P. Qian, J. Wang, G. Wang, *et al.*, "Seizure Classification from EEG Signals using Transfer Learning, Semi-Supervised Learning and TSK Fuzzy System," *IEEE Transactions on Neural Systems & Rehabilitation Engineering*, vol. PP, pp. 1-1, 2017.
- [23] G. Wang, Z. Deng, and K. S. Choi, *Detection of Epileptic Seizures in EEG Signals with Rule-Based Interpretation by Random Forest Approach*: Springer International Publishing, 2015.
- [24] Y. Zhang, B. Liu, X. Ji, and D. Huang, "Classification of EEG Signals Based on Autoregressive Model and Wavelet Packet Decomposition," *Neural Processing Letters*, vol. 45, pp. 1-14, 2017.
- [25] C. Yang, Z. Deng, K. S. Choi, and S. Wang, "Takagi-Sugeno-Kang Transfer Learning Fuzzy Logic System for the Adaptive Recognition of Epileptic Electroencephalogram Signals," *IEEE Transactions on Fuzzy Systems*, vol. 24, pp. 1079-1094, 2016.
- [26] R. Heffernan, K. Paliwal, J. Lyons, A. Dehngi, A. Sharma, J. Wang, *et al.*, "Improving prediction of secondary structure, local backbone angles, and solvent accessible surface area of proteins by iterative deep learning," *Scientific reports*, vol. 5, p. 11476, 2015.
- [27] J. Xu, L. Xiang, Q. Liu, H. Gilmore, J. Wu, J. Tang, *et al.*, "Stacked sparse autoencoder (SSAE) for nuclei detection on breast cancer histopathology images," *IEEE transactions on medical imaging*, vol. 35, pp. 119-130, 2016.
- [28] S. Jirayucharoensak, S. Pan-Ngum, and P. Iprasena, "EEG-based emotion recognition using deep learning network with principal component based covariate shift adaptation," *The Scientific World Journal*, vol. 2014, 2014.
- [29] M. Spencer, J. Eickholt, and J. Cheng, "A deep learning network approach to ab initio protein secondary structure prediction," *IEEE/ACM transactions on computational biology and bioinformatics (TCBB)*, vol. 12, pp. 103-112, 2015.
- [30] S. M. Plis, D. R. Hjelm, R. Salakhutdinov, E. A. Allen, H. J. Bockholt, J. D. Long, *et al.*, "Deep learning for neuroimaging: a validation study," *Frontiers in neuroscience*, vol. 8, p. 229, 2014.
- [31] X. An, D. Kuang, X. Guo, Y. Zhao, and L. He, "A deep learning method for classification of EEG data based on motor imagery," in *International Conference on Intelligent Computing*, 2014, pp. 203-210.
- [32] B. Alipanahi, A. Delong, M. T. Weirauch, and B. J. Frey, "Predicting the sequence specificities of DNA-and RNA-binding proteins by deep learning," *Nature biotechnology*, vol. 33, p. 831, 2015.
- [33] F. Ning, D. Delhomme, Y. Lecun, F. Piano, L. Bottou, and P. E. Barbanò, "Toward automatic phenotyping of developing embryos from videos," *IEEE Transactions on Image Processing*, vol. 14, pp. 1360-1371, 2005.
- [34] H. Cecotti and A. Graeser, *Convolutional Neural Networks for P300 Detection with Application to Brain-Computer Interfaces*: IEEE Computer Society, 2011.
- [35] S. K. Sønderby, C. K. Sønderby, H. Nielsen, and O. Winther, "Convolutional LSTM networks for subcellular localization of proteins," in *International Conference on Algorithms for Computational Biology*, 2015, pp. 68-80.
- [36] A. Graves, "Offline handwriting recognition with multidimensional recurrent neural networks," in *International Conference on Neural Information Processing Systems*, 2008, pp. 545-552.
- [37] P. R. Davidson, R. D. Jones, and M. T. Peiris, "EEG-based lapse detection with high temporal resolution," *IEEE Trans Biomed Eng*, vol. 54, pp. 832-839, 2007.
- [38] S. Lee, M. Choi, H. S. Choi, M. S. Park, and S. Yoon, "FingerNet: Deep learning-based robust finger joint detection from radiographs," in *Biomedical Circuits and Systems Conference*, 2015, pp. 1-4.
- [39] Q. Yin, S. Wu, and L. Wang, "Incomplete Multi-view Clustering via Subspace Learning," 2015, pp. 383-392.
- [40] H. Hotelling, "Relations Between Two Sets of Variates," *Biometrika*, vol. 28, pp. 321-377, 1936.
- [41] A. Blum, "Combining labeled and unlabeled data with co-training," in *Conference on Computational Learning Theory*, 1998, pp. 92-100.
- [42] D. R. Hardoon and J. Shawe-Taylor, *Sparse canonical correlation analysis*: Kluwer Academic Publishers, 2011.
- [43] S. Sun and J. Shawe-Taylor, *Sparse Semi-supervised Learning Using Conjugate Functions*: JMLR.org, 2010.
- [44] P. Bashivan, G. M. Bidelman, and M. Yeasin, "Spectrotemporal dynamics of the EEG during working memory encoding and maintenance predicts individual behavioral capacity," *European Journal of Neuroscience*, vol. 40, pp. 3774-3784, 2014.
- [45] O. Jensen and C. D. Tesche, "Frontal theta activity in humans increases with memory load in a working memory task," *European journal of Neuroscience*, vol. 15, pp. 1395-1399, 2002.
- [46] D. Hu, W. Li, and X. Chen, "Feature extraction of motor imagery EEG signals based on wavelet packet decomposition," in *Ieee/icme International Conference on Complex Medical Engineering*, 2011, pp. 694-697.
- [47] T. Zhang, Z. Deng, D. Wu, and S. Wang, "Multi-View Fuzzy Logic System with the Cooperation between Visible and Hidden Views," 2018.
- [48] W. Zhou, Y. Liu, Q. Yuan, and X. Li, "Epileptic Seizure Detection Using Lacunarity and Bayesian Linear Discriminant Analysis in Intracranial EEG," *IEEE Transactions on Biomedical Engineering*, vol. 60, pp. 3375-3381, 2013.
- [49] A. H. Shoeb, "Application of machine learning to epileptic seizure onset detection and treatment," in *International Conference on Machine Learning*, 2009, pp. 975-982.



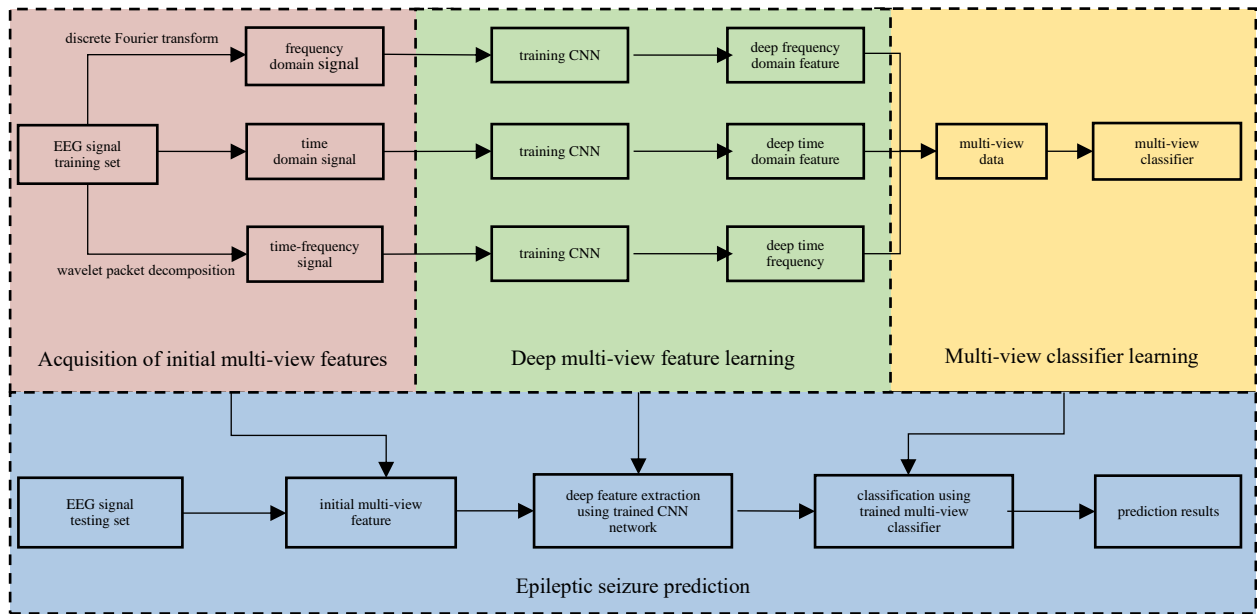


Fig 1 Framework of the proposed method

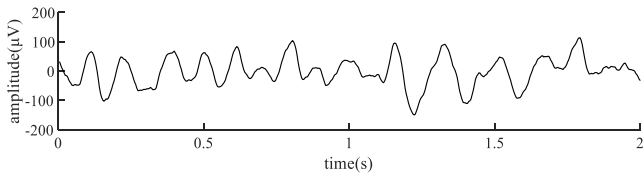


Fig. 2 Time domain features of EEG signals

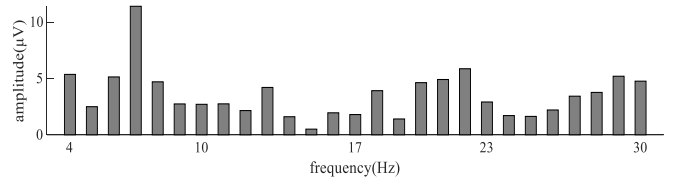


Fig 3 Frequency domain features of EEG signals

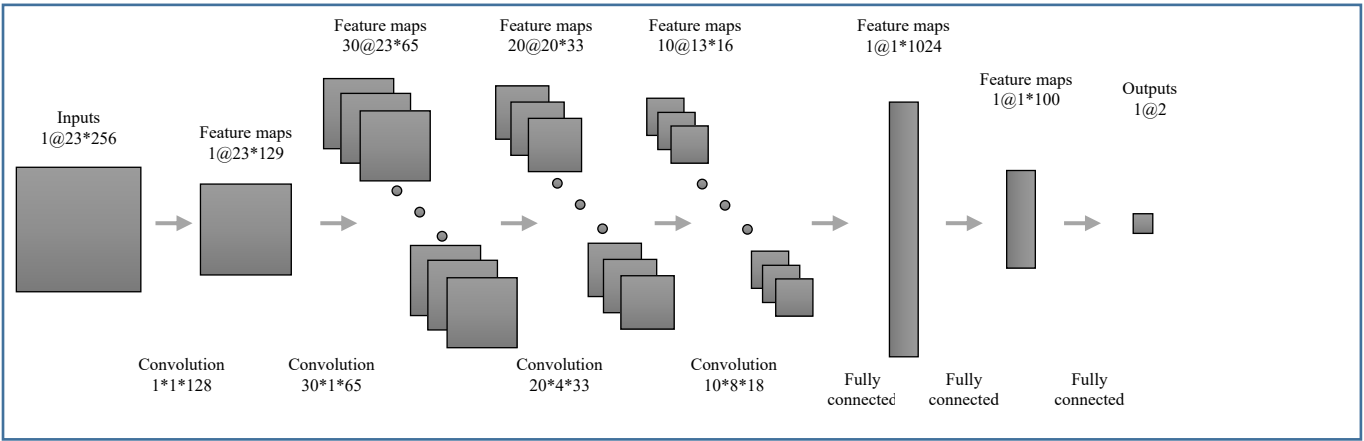


Fig 5 Deep feature extraction network with time domain features

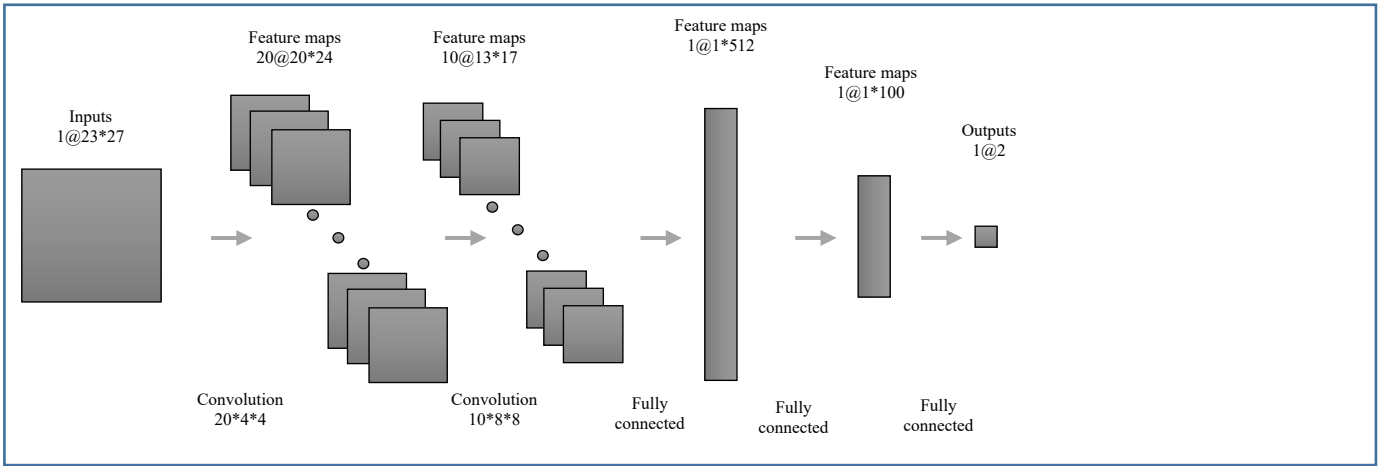


Fig 6 Deep feature extraction network with frequency domain features

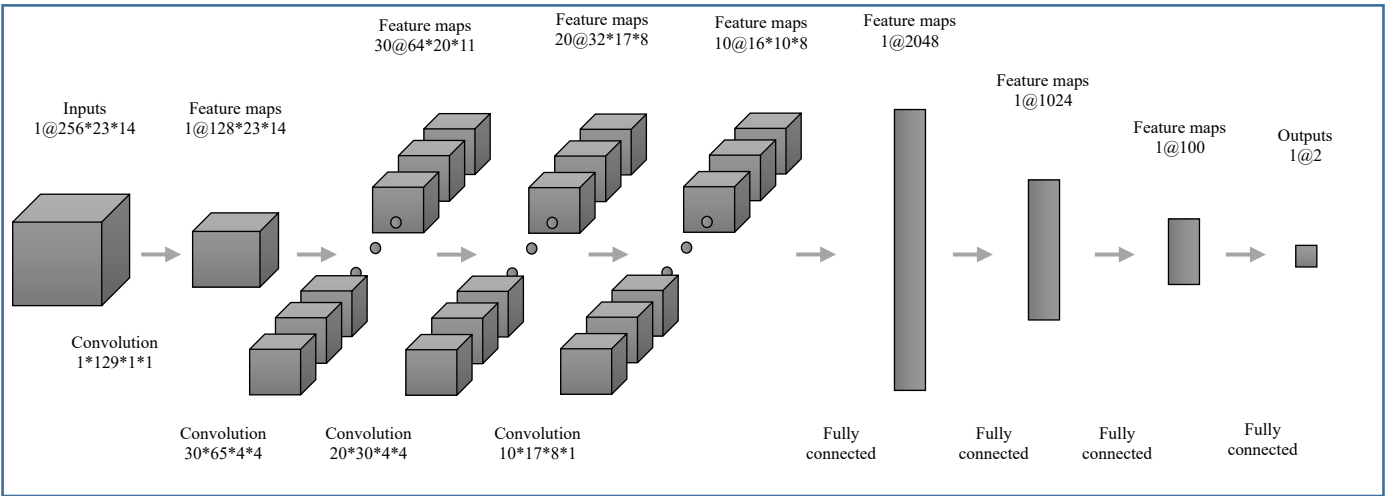


Fig 7 Deep feature extraction network with time-frequency features

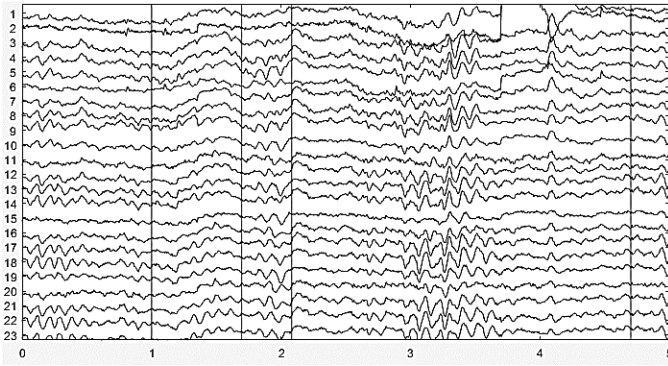


Fig. 8 The original multi-channel EEG signals

TABLE III The effect of over-sampling method on the performance

Over-sampling method	accuracy	sensitivity	specificity
Sliding window	98.33±0.18	96.66±0.14	99.14±0.14
SMOTE	98.91±0.24	97.97±0.12	99.38±0.18

TABLE IV The performance of different division percentage

Percentage of train data	accuracy	sensitivity	specificity
90%	98.43±0.21	96.51±0.15	98.95±0.14
75%	98.33±0.18	96.66±0.14	99.14±0.14
50%	97.16±0.31	94.68±0.21	98.40±0.18

TABLE I Accuracy, sensitivity and specificity of the proposed method on the CHB-MIT dataset

group	accuracy	sensitivity	specificity
1	99.57±0.15	99.37±0.16	99.67±0.13
2	99.71±0.21	99.13±0.19	100±0.15
3	99.09±0.19	98.02±0.2	99.57±0.16
4	99.11±0.2	98.59±0.17	99.36±0.11
5	99.35±0.17	98.28±0.1	99.87±0.12
6	98.85±0.26	97.06±0.13	99.7±0.15
7	99.28±0.13	99.31±0.12	99.28±0.14
8	98.16±0.19	96.54±0.17	98.94±0.14
9	99.64±0.13	99.26±0.15	99.8±0.16
10	98.79±0.17	96.76±0.16	99.81±0.13
11	99.35±0.2	98.35±0.11	99.85±0.11
12	88.07±0.37	81.01±0.2	91.58±0.23
13	98.69±0.24	96.19±0.14	99.92±0.13
14	97.15±0.23	94.17±0.14	98.51±0.12
15	98.39±0.18	96.27±0.16	99.44±0.13
16	99.31±0.12	98.48±0.1	99.68±0.14
17	97.19±0.18	93.8±0.11	98.77±0.18
18	99.79±0.09	99.92±0.02	99.62±0.16
19	99.63±0.13	99.26±0.09	99.8±0.11
20	98.48±0.15	96.9±0.13	99.2±0.1
21	96.61±0.19	92.71±0.21	98.61±0.15
22	99.73±0.06	99.38±0.11	99.89±0.18
23	98.72±0.14	97.63±0.15	99.27±0.13
24	97.24±0.19	93.51±0.12	99.12±0.17
average	98.33±0.18	96.66±0.14	99.14±0.14

TABLE II Performance of existing seizure detection methods on the CHB-MIT dataset

Authors	Feature Extraction Methods	Groups	Channels	acc-sen-spe
Rafiuddin et.al. [1]	Energy and coefficient of wavelet coefficients and statistical feature	23	23	80.16-NA-NA
Khan et.al. [2]	Relative energy and normalized coefficients of wavelet coefficients	5	NA	91.8-83.6-100
Kiranyaz et.al. [3]	Time domain, frequency domain, time-frequency domain and non-linear features.	21	18	NA-89.0-94.7
Zabihi et.al. [4]	Seven features extracted from intersection sequence	23	23	94.6-89.1-94.8
Samiee et.al. [5]	Multivariate textural features extracted by GLCM	23	23	NA-70.1-97.7
Xinghua Yao[8]	Using independent RNN to extract feature	24	17	87-87.3-86.7
Xinghua Yao[10]	An attention mechanism and a bidirectional long short-term memory model exploiting both spatially and temporally discriminating features	24	17	83.9-83.7-84.1
Hengjin Ke[12]	Using a Lightweight VGGNet to extract feature	24	19	98.1-98.9-97.4
Proposed	Time domain, frequency domain, time-frequency domain feature extracted using CNN	24	23	98.3- 96.7 -99.1

TABLE V Accuracy of different feature extraction methods

Classifier	Time Domain				Frequency Domain			Time-Frequency Domain		
	Time	PCA	LDA	Deep features	FFT	FFT-PCA	Deep features	WPD	WPD-PCA	Deep features
SVM	70.93±1.37	<b>86.47±1.28</b>	83.35±1.33	75.82±1.1	<b>97.87±0.08</b>	97.83±0.05	97.47±0.37	91.88±0.08	91.87±0.05	<b>92.95±0.81</b>
KNN	67.81±2.72	<b>83.50±2.72</b>	81.92±1.49	75.79±1.23	97.38±0.26	97.35±0.48	<b>98.16±0.21</b>	91.55±0.06	91.60±0.48	<b>93.31±0.63</b>
NB	<b>84.49±0.19</b>	84.04±0.86	79.71±1.21	76.23±1.09	87.84±0.59	89.43±0.49	<b>97.08±0.48</b>	77.80±0.25	73.91±0.49	<b>92.56±0.62</b>
DT	84.78±0.41	<b>89.93±1.11</b>	83.08±1.39	72.13±1.64	95.49±0.82	95.47±0.05	<b>95.58±0.32</b>	77.87±0.23	87.63±0.05	<b>91.66±0.94</b>
TSK-FS	<b>81.15±0.23</b>	71.47±0.68	71.33±1.49	78.70±0.99	94.61±0.22	<b>98.38±0.08</b>	97.82±0.23	N. A.	91.89±0.35	<b>93.41±0.59</b>
average	77.00±0.98	<b>85.99±1.33</b>	81.71±1.38	74.99±1.21	94.67±0.39	95.02±0.23	<b>97.07±0.32</b>	84.78±0.15	86.25±0.28	<b>92.62±0.72</b>

TABLE VI Sensitivity of different feature extraction methods

Classifier	Time Domain				Frequency Domain			Time-Frequency Domain		
	time	PCA	LDA	Deep features	FFT	FFT-PCA	Deep features	WPD	WPD-PCA	Deep features
SVM	44.84±3.10	29.53±3.21	34.97±2.43	<b>56.92±1.75</b>	88.37±0.15	87.43±0.16	<b>96.11±0.68</b>	62.74±0.23	62.77±0.16	<b>88.91±1.35</b>
KNN	21.54±6.04	17.52±6.04	37.31±3.56	<b>54.35±1.25</b>	82.07±0.74	82.37±1.32	<b>96.12±0.37</b>	47.45±0.68	47.92±1.32	<b>89.87±1.34</b>
NB	<b>84.75±0.56</b>	82.22±1.35	29.87±5.73	64.27±0.86	89.60±0.35	82.41±0.61	<b>96.18±0.49</b>	69.52±0.31	75.68±0.61	<b>92.40±0.9</b>
DT	<b>76.28±1.38</b>	56.13±2.53	36.88±4.76	55.62±2.88	79.98±0.77	78.83±0.96	<b>93.04±0.78</b>	69.82±0.34	50.38±0.96	<b>87.07±1.42</b>
TSK-FS	<b>58.72±2.36</b>	48.98±2.28	46.57±3.39	56.70±1.89	85.43±0.29	97.25±0.12	<b>96.56±0.43</b>	N. A.	84.99±0.53	<b>90.92±1.25</b>
Average	56.85±2.68	46.35±3.08	34.76±3.97	<b>57.79±1.73</b>	85.01±0.46	82.76±0.63	<b>95.36±0.55</b>	62.38±0.39	59.19±0.71	<b>89.56±1.25</b>

TABLE VII Specificity of different feature extraction methods

Classifier	Time Domain				Frequency Domain			Time-Frequency Domain		
	time	PCA	LDA	Deep features	FFT	FFT-PCA	Deep feature	WPD	WPD-PCA	Deep features
SVM	83.46±1.75	<b>93.26±1.70</b>	89.20±1.45	84.93±0.84	<b>99.02±0.06</b>	99.10±0.06	98.13±0.31	<b>95.41±0.08</b>	95.39±0.06	84.9±0.85
KNN	90.11±0.06	<b>91.48±8.72</b>	87.28±2.35	86.17±1.33	<b>99.23±0.06</b>	99.17±0.06	98.18±0.25	96.92±0.08	<b>96.93±0.06</b>	94.96±0.82
NB	84.36±0.36	84.26±1.18	<b>85.74±3.31</b>	81.97±1.28	87.61±0.79	90.28±0.49	<b>97.22±0.59</b>	78.78±0.56	73.68±0.49	<b>92.63±0.93</b>
DT	88.90±0.58	<b>93.93±1.14</b>	87.30±2.21	80.04±1.23	97.48±0.80	<b>97.50±0.41</b>	96.82±0.32	78.81±0.53	92.13±0.41	<b>93.88±1.19</b>
TSK-FS	<b>91.83±0.11</b>	82.24±0.53	83.26±2.53	89.29±0.81	<b>99.10±0.32</b>	98.93±0.08	98.44±0.22	N. A.	<b>95.22±0.04</b>	<b>94.61±0.92</b>
Average	86.71±0.57	90.73±2.65	87.38±2.37	83.28±1.10	95.84±0.41	96.51±0.22	<b>97.59±0.34</b>	87.48±0.31	89.53±0.21	<b>91.59±0.94</b>

TABLE VIII The effect of deep features on the classifiers.

		Accuracy	Sensitivity	Specificity
Time domain deep features	SVM	75.82±1.10	56.92±1.75	84.93±0.84
	KNN	75.79±1.23	54.35±1.25	86.17±1.33
	NB	76.23±1.09	64.27±0.86	81.97±1.28
	DT	72.13±1.64	55.62±2.88	80.04±1.23
	TSK-FLS	78.70±0.99	56.70±1.89	89.29±0.81
Frequency domain deep features	SVM	97.47±0.37	96.11±0.68	98.13±0.31
	KNN	98.16±0.21	96.12±0.37	98.18±0.25
	NB	97.08±0.48	95.18±0.49	97.22±0.59
	DT	95.58±0.32	93.04±0.78	96.82±0.32
	TSK-FLS	97.82±0.23	96.56±0.43	98.44±0.22
Time-frequency deep features	SVM	92.95±0.81	88.91±1.35	84.90±0.85
	KNN	93.31±0.63	89.87±1.34	94.96±0.82
	NB	92.56±0.62	92.40±0.90	92.63±0.93
	DT	91.66±0.94	87.07±1.42	93.88±1.19
	TSK-FLS	92.41±0.59	90.92±1.25	94.61±0.92
The proposed method		<b>98.33±0.18</b>	<b>96.66±0.14</b>	<b>99.14±0.14</b>

TABLE IX Detection delay of the proposed method

Group	EO <sub>Latency</sub> (s)	S (%)
1	1.0045	100
2	1.0029	100
3	1.0338	100
4	1.0213	100
5	1.0541	99.82
6	1.0167	100
7	1.0016	100
8	1.0480	100
9	1.0055	100
10	1.0146	100
11	1.0236	100
12	1.3573	99.38
13	1.0594	100
14	1.0119	100
15	1.1197	99.95
16	1.0148	100
17	1.0552	100
18	1.0032	100
19	1.0064	100
20	1.0308	100
21	1.0684	100
22	1.0074	100
23	1.0497	99.76
24	1.0225	100
Average	1.0431	99.95