Biomedical Signal Processing and Control 79 (2023) 104155

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

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

Seizure detection algorithm based on fusion of spatio-temporal network constructed with dispersion index

Yuhuan Xiong [a], Jinghan Li [a], Duanpo Wu [a,b,c,*], Fang Dong [d], Junbiao Liu [e], Lurong Jiang [f], Jiuwen Cao [c,e], Yuansheng Xu [g,**]

[a] School of Communication Engineering, Hangzhou Dianzi University, Hangzhou 310018, China
[b] Zhejiang Provincial Key Laboratory of Information Processing, Communication and Networking, Hangzhou 310027, China
[c] Machine Learning and I-health International Cooperation Base of Zhejiang Province, Hangzhou Dianzi University, Hangzhou 310018, China
[d] College of Information and Electric Engineering, Zhejiang University City College, Hangzhou 310015, China
[e] Artificial Intelligence Institute, Hangzhou Dianzi University, Hangzhou 310018, China
[f] School of Information Science and Technology, Zhejiang Sci-Tech University, Hangzhou 310018, China
[g] Affiliated Hangzhou First People's Hospital, Zhejiang University School of Medicine, Hangzhou 310003, China

ARTICLE INFO

Keywords:
Multivariate variational mode decomposition
Dispersion index
Weighted horizontal visibility graph
Seizure detection

ABSTRACT

Epilepsy is one of the common brain disorders. Traditional seizure detection is done by electroencephalography (EEG) technicians, which takes a lot of time. Therefore, this paper proposes a new method for automatic seizure detection. In this paper, we use multivariate variational mode decomposition (MVMD) method for modal decomposition of multi-channel EEG signals and introduce dispersion index (DI) as a new brain network weight calculation method which is based on dispersion entropy (DE) to extract the features of single-channel temporal brain network and multi-channel spatial brain network for seizure detection. In single-channel temporal brain network, the network is constructed by weighted horizontal visibility graph (WHVG) method which considers the sample points as network nodes. In multi-channel spatial brain networks, DI is used to calculate the correlation between channels and select the threshold to construct the ternary-valued network of channels. We perform the validation on two publicly available datasets. The results show that the classification results for F1, AUC, ACC, PRE, SEN and SPE on CHB-MIT dataset are 97.89%, 97.81%, 97.83%, 97.56%, 98.24% and 97.39%, respectively. In addition, the results for F1, AUC, ACC, PRE, SEN and SPE on Siena scalp dataset are 99.21%, 99.19%, 99.19%, 99.28%, 99.14%, and 99.24%, respectively. The method proposed in this paper achieves good results on both datasets. In general, the joint detection of temporal and spatial networks is promising, and DI can serve as a valid indicator of correlation.

1. Introduction

Epilepsy is a chronic, recurrent, and sudden-onset neurological disorder that affects more than 70 million people worldwide, in varying degrees [1]. A canonical view holds that seizures, the characteristic sign of epilepsy, occur at random [2]. The traditional clinical medical methods for judging epileptic seizures relies on the doctor's visual observation of electroencephalography (EEG) data which is highly likely to be biased and inefficient. Therefore, the introduction of intelligent detection of EEG signals is very necessary.

In 2004, Huang et al. proposed the method of empirical mode decomposition (EMD) [3], providing new ideas for the analysis of nonsmooth, nonlinear signals. Basing on the shortcomings of EMD which is sensitive to sampling and noise, Dragomiretskiy et al. proposed an adaptive variational method called variational mode decomposition (VMD) [4]. In comparison with EMD, VMD avoids modeling individual modes as signals with explicit intrinsic mode functions (IMFs) with better robustness, and with less frequency aliasing at the same number of decomposed modes [5]. In 2019, Rehman et al. proposed multivariate variational mode decomposition (MVMD) [6], which was built on the basis of VMD and the joint or common frequency components present to multi-channel data. MVMD exhibits superior performance over the multi-channel expanded multivariate empirical mode decomposition (MEMD) [7] and is gradually applied in fault detection [7,8], schizophrenia detection [9] and etc.

In terms of constructing networks and feature extraction, complex network theory has been increasingly applied to the field of EEG.

* Corresponding author at: School of Communication Engineering, Hangzhou Dianzi University, Hangzhou 310018, China.
** Corresponding author.
E-mail addresses: wuduanpo@hdu.edu.cn (D. Wu), xys0912@163.com (Y. Xu).

https://doi.org/10.1016/j.bspc.2022.104155
Received 16 April 2022; Received in revised form 26 July 2022; Accepted 4 September 2022
Available online 18 September 2022

From 2008 to 2009, Lacasa et al. successively proposed visibility graph (VG) [10] and horizontal visibility graph (HVG) [11] to map time series into complex networks in a new perspective. Later, VG and HVG continued to develop into weighted visibility graph [12], and multiscale limited penetrable HVG [13], which were applied to sleep EEG classification [14,15], autism spectrum disorder analysis [16], and emotion recognition [17]. In addition to temporal network, spatial network which considers channels as nodes has been extensively constructed, and applied in many researches [18,19]. In order to construct spatial network, many methods have been proposed to calculate the weight between different nodes, including mutual information [20,21], imaginary part of phase lag index [22] and etc. Although the research on single network is relatively mature, there is still a lack of research on the fusion of these two networks.

As both of the networks constructed from temporal and spatial domain can provide effective information to detect seizures, this paper proposes a novel seizure detection algorithm based on spatio-temporal network constructed by dispersion index (DI). On the basis of dispersion entropy (DE) [23], DI is introduced as a new weight metric, which consists of three steps. Firstly, the signal points are mapped and classified by the normal cumulative distribution function (NCDF). Secondly, each embedding vector is mapped to the dispersion mode. Finally, the same dispersion mode occurs simultaneously between channels using Shannon entropy formula. A brief description of the algorithm is as follows. The raw EEG signals are first decomposed by MVMD, and then a model based on the decomposed individual modes is constructed as a two-layer complex network in both temporal and spatial domain. Among them, the temporal network is constructed by a fast HVG algorithm with radix function weighting strategy [24]. For the spatial network, we use DI as the inter-channel weight and ternary-value the network by setting threshold. Finally, we classify the features of both networks by random forest (RF) classifier, thus demonstrate the effectiveness of the proposed seizure detection algorithm.

The overall structure of the paper is as follows. We details the dataset used in this paper and the proposed seizure detection method in Section 2, including data preprocessing, the idea and computational process of DI, feature selection and extraction, and the introduction of the classifier. Then, in Section 3, the experimental results are evaluated and analyzed. Finally, further discussion and conclusions are drawn in Section 4.

2. Method

In this section, we present the overall approach and idea of the seizure detection algorithm. Firstly, this paper performs the preprocessing of data and divides the segments with sliding time windows. Secondly, the modal decomposition of multi-channel EEG signals is performed with MVMD to obtain the information of different components. For the spatial channel network, the correlation between channels is firstly calculated by DI, and then the spatial network is further changed into a ternary-valued network by setting threshold. For the temporal network, we use a fast HVG algorithm to construct an undirected weighted network. Finally, after feature extraction and selection, RF classifier is used for classification. The specific processing process is shown in Fig. 1.

2.1. Datasets description

The data for this study are derived from two publicly available independent datasets, including CHB-MIT dataset and Siena scalp EEG dataset.

The CHB-MIT dataset [25] is taken from Boston Children's Hospital and includes EEG recordings collected from pediatric subjects. Twenty-four subjects are taken from 22 patients (5 males, aged 3–22 years; 17 females, aged 1.5–19 years). All signals are sampled with 256 samples per second at 16-bit resolution. Most files contain 23 channel EEG signals. The EEG that is collected from international 10–20 electrode positions are selected as raw data. The dataset can be freely obtained in [26].

The Siena scalp EEG dataset [27] contains EEG recordings of 14 patients obtained from the Department of Neurology and Neurophysiology of the University of Siena. Subjects include 9 males (25–71 years) and 5 females (20–58 years). Subjects are monitored using Video-EEG with a sampling rate of 512 Hz and electrodes are arranged according to the international 10–20 system. The dataset can be freely obtained in [28].

In this experiment, the channel is double-lead, the FP1 and FP2 related channels are removed, and the 'P7-T7', 'T7-FT9', 'FT9-FT10' and 'FT10-T8' are removed according to the 10–20 lead system. Therefore, the number of channels remaining is 14 groups, and for more information, we take another 6 channels into consideration, including 'F7-F3', 'T3-C3', 'T5-P3', 'F4-F8', 'C4-T4' and 'P4-T6' [29].

2.2. Preprocess

In this paper, we first segment the EEG dataset into equal-length segments using sliding time windows. This is because the seizure segments have less data, and by adding windows we are able to extract more data from the seizure segments and ensure that the data are equally informative. Seizures are short in duration and tend to recur. In order to effectively extract the signal features at the moment of seizures, EEG signal is divided into several segments of 4 s and the overlap rate of the seizure segments is 0.5. After that, a band-pass filter of 0.5–48 Hz is used to filter out the electrooculography, out-of-band noise, and power line noise to improve the signal-to-noise ratio [30].

2.3. Multivariate variational mode decomposition

MVMD separates the modal components of different center frequencies from the original signal, which facilitates more specific analysis of the signal for different modes. The main goal of MVMD is to extract predefined K number of multivariate modulated oscillations from input data, which can be expressed as follows.

$$x_c(t) = \sum_{k=1}^{K} u_{k,c}(t). \tag{1}$$

where $x_c(t)$ is the input signal of the cth channel. The modulated oscillation signal of the cth channel of the kth mode is $u_{k,c}(t)$. K denotes the number of multivariate modulated oscillations.

In MVMD, the bandwidth of each mode $u_{k,c}(t)$ per channel is estimated by: (i) computing the analytic signal representation $u_{+}^{k,c}(t)$ for $u_{k,c}(t)$; (ii) shifting the spectrum to baseband via harmonic mixing with a complex exponential of frequency ω_k; (iii) taking the squared L_2 norm of the gradient of the harmonically mixed signal obtained in the previous step. Finally, MVMD is formulated as a constrained optimization problem as follows:

$$\begin{aligned} &\underset{\{u_{k,c}(t)\},\{\omega_k\}}{\text{minimize}} \quad \left\{ \sum_k \sum_c \left\| \partial_t \left[u_{+}^{k,c}(t) e^{-j\omega_k t} \right] \right\|_2^2 \right\} \\ &\text{subject to} \quad \sum_k u_{k,c}(t) = x_c(t), c = 1, 2, \ldots, C. \end{aligned} \tag{2}$$

where the symbol ∂_t represents partial derivative operation with respect to time, and C is the total number of channels. ω_k is defined as the kth common frequency component in all channels. $\{u_{k,c}(t)\}$ and $\{\omega_k\}$ respectively denote the set of all kth modes of the cth channel and the set of kth mode center frequencies.

The above optimization problem can be turned into corresponding augmented Lagrangian functions, and then transformed into multiple sub-optimization problems by alternating directional multiplier method (ADMM) to simplify the solution and perform mode update, as described in [6]. In order to apply MVMD method, it is necessary to properly set K and the parameter α which defines the bandwidth of

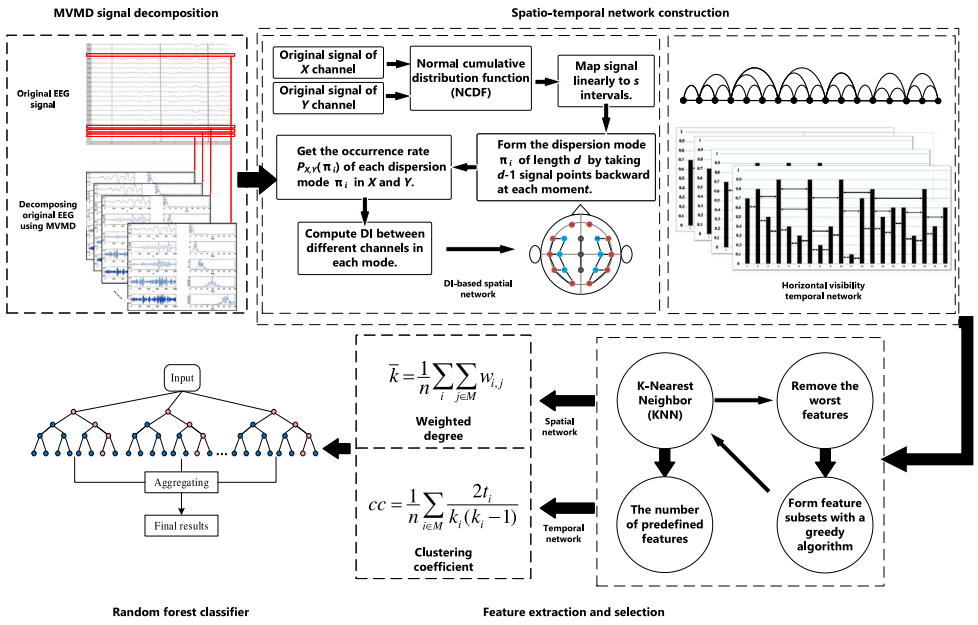

Fig. 1. Block diagram of DI-based spatio-temporal network for seizure detection.

extracted modes in ADMM. If K and α are very low, mode mixing problem will occur. On the other hand, when K and α are high, the algorithm will be difficult to converge and the noise in signal may be perceived as a mode. The results are tested on datasets when K is 4, 5, 6 and 7, where $K = 4$ performs the best. According to the simulation of the EEG dataset in [6], $\alpha = 2000$ is chosen.

2.4. DI-based spatial network construction

The idea of DI is derived from DE proposed by Rostaghi et al. [23]. DE originates from permutation entropy (PE) and sample entropy (SE) and maps signal points to dispersion modes by NCDF to perform computation, which includes the following three steps:

(1) First, NCDF maps the signal $u_{k,c}(t)(t = 1,2,\dots,N)$ to $y_{k,c}(t)(t = 1,2,\dots,N)$, which lies between 0 and 1.

$$y_{k,c}(t) = \int_{-\infty}^{u_{k,c}(t)} \frac{1}{\sqrt{2\pi}\sigma} \exp(-\frac{(x-\mu)^2}{2\sigma^2})\, dx. \tag{3}$$

where σ and μ represent the standard deviation and mean of the signal $u_{k,c}(t)$, respectively. Then, [0,1] can be linearly mapped to s intervals. If the value of $y_{k,c}(t)$ is in the jth interval, it will be quantized to j.

$$z_{k,c}(t) = j(\frac{j-1}{s} \le y_{k,c}(t) < \frac{j}{s}, 1 \le j < s). \tag{4}$$

where $z_{k,c}(t)$ represents the linearly mapped signal of $y_{k,c}(t)$.

(2) $z_{k,c}(t)$ can be mapped into a matrix, consisting of $N-(m-1)d$ rows and m columns, with m being the embedding dimension and d being the time lag. $\mathbf{Z}_{k,c}(t,m,d)$ is expressed as follows:

$$\mathbf{Z}_{k,c}(t,m,d) = \begin{bmatrix} z_{k,c}(t) & z_{k,c}(t+1) & \cdots & z_{k,c}(t+m-1) \\ z_{k,c}(t+d) & z_{k,c}(t+d+1) & \cdots & z_{k,c}(t+d+m-1) \\ \vdots & \vdots & \ddots & \vdots \\ z_{k,c}(t+Ld) & z_{k,c}(t+Ld+1) & \cdots & z_{k,c}(t+Ld+m-1) \end{bmatrix} \tag{5}$$

where L denotes the coefficient of d, $L = N-(m-1)d-1$.

(3) Each row vector of $\mathbf{Z}_{k,c}(t,m,d)$ can be represented as one of the dispersion modes, which totally include s^m possible dispersion modes. Let the dispersion mode of the ith row vector be θ_i, $\theta_i \in \{\pi_1, \dots \pi_{s^m}\}$. Then $\mathbf{Z}_{k,c}(t,m,d)$ is further written as:

$$\mathbf{Z}_{k,c}(t,m,d) = [\theta_1, \theta_2, \dots, \theta_i, \dots, \theta_{L+1}]^T. \tag{6}$$

Let the total number of ith dispersion mode in the signal be $n(\pi_i)$, then its probability can be calculated as follows:

$$P(\pi_i) = \frac{n\{\pi_i\}}{N-(m-1)d} (1 \le i \le s^m). \tag{7}$$

Basing on Shannon's definition of entropy, the DE with embedding dimension m, time lag d, and the number of linear map intervals s is calculated as follows:

$$DE(m,s,d) = -\sum_{i=1}^{s^m} P(\pi_i) \ln(P(\pi_i)). \tag{8}$$

However, DE is a measure of single-channel complexity. In order to quantify the correlation between different channels, DI is proposed as a new indicator. We define the number of simultaneous occurrence of the same dispersion mode in two channels X and Y as $n_{X,Y}\{\pi_i\}$. Then the probability of occurrence of the ith dispersion mode in X and Y channels simultaneously can be denoted as:

$$P_{X,Y}(\pi_i) = \frac{n_{X,Y}\{\pi_i\}}{\sum_i n_{X,Y}\{\pi_i\}} (1 \le i \le s^m). \tag{9}$$

The formula for calculating DI is as follows:

$$DI(X,Y) = -\sum_{i=1}^{s^m} P_{X,Y}(\pi_i) \ln(P_{X,Y}(\pi_i)). \tag{10}$$

The Fig. 2 shows intuitive calculation process. First, the signals of X and Y channels are processed by NCDF, and it can be seen that X channel shows fluctuations, and Y channel is with almost no change. After that, the NCDF-processed signals of X and Y channels are linearly mapped to s intervals, and the result is that X and Y channels have completely consistent dispersion modes, which cannot directly judged from the original signal. This shows that DI is able to present the internal relationship between channels.

In this paper, the weight matrix obtained by DI is ternary-valued, as shown in Fig. 3. We calculate the distribution histogram of DI and divide the frequency into three equal parts. The first tertile is in the range of [1.3,1.4] and the second tertile is in the range of [1.8,1.9]. The

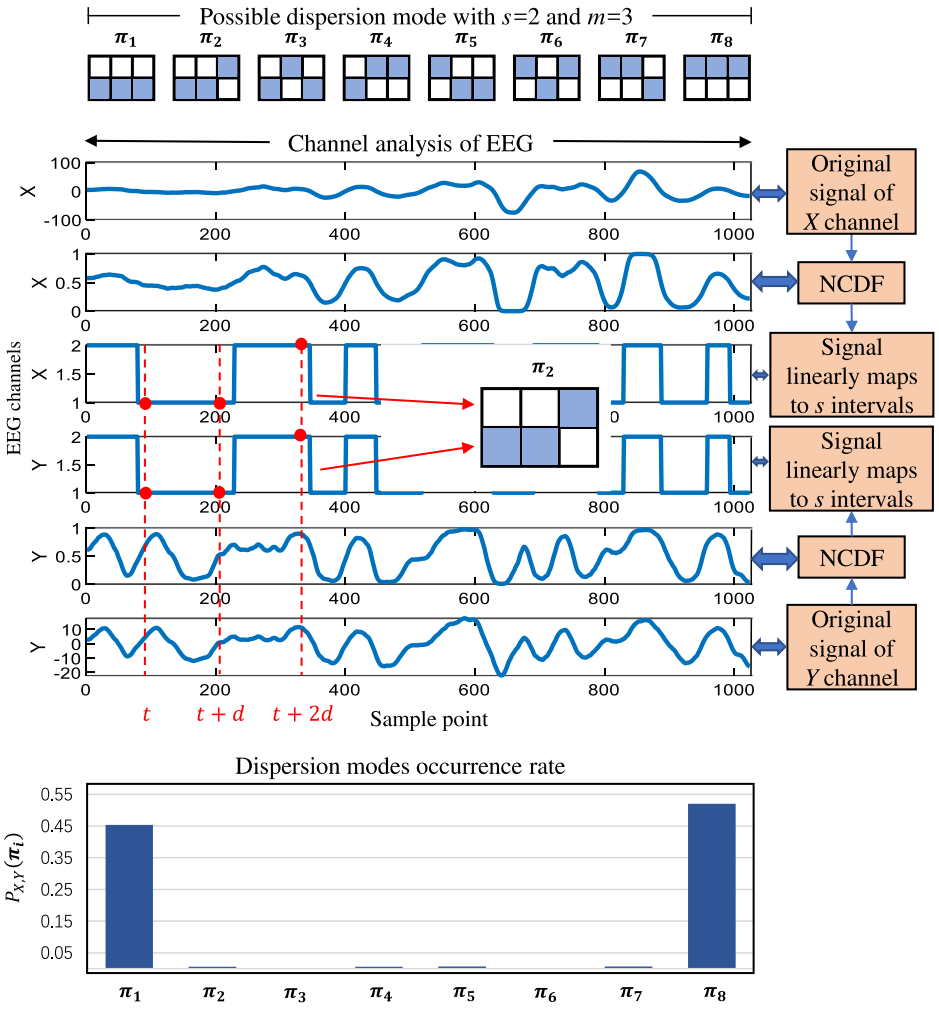

Fig. 2. Given two channels of EEG signals X and Y, the parameters of m, s, t and d. We can project the signals of X and Y channels in the kth mode into vectors $y_{k,X}(t)$ and $y_{k,Y}(t)$, respectively, by NCDF and then linearly map to s intervals to obtain dispersion modes. When $m = 3$ and $s = 2$, there are 8 possible dispersion modes. At time point t in this figure, the same dispersion mode π_2 appears in X and Y channels at the same time. This process is repeated for different t and d, so we can get the occurrence rate $P_{X,Y}(\pi_i)$ of each dispersion mode π_i in X and Y.

threshold of different part is the end point or median of the interval in which the tertile of statistics is located. Finally, in the spatial network, the weight of edge between channel i and j can be ternary-valued as follows:

$$e_{i,j} = \begin{cases} 1, & DI(i,j) > T_{h,2} \\ 0.5, & DI(i,j) > T_{h,1} \\ 0, & else. \end{cases} \tag{11}$$

where $T_{h,1}$ and $T_{h,2}$ represent the thresholds determined according to the interval of the first and second tertile, respectively. In addition, this method of determining thresholds can be used for multi-valuation of weight in network.

2.5. Fast HVG algorithm to build temporal network

The spatial channel network characterizes the spatial information of the network, but ignores the information in temporal, so we construct the network by the fast HVG algorithm for the time domain information of each channel. Each sample point of the sequence $u_{k,c}(t)(t = 1, 2, \ldots, N)$ is considered as a node of the HVG, and ith and jth nodes are connected by a horizontal line when two nodes in the sequence satisfy the following conditions [11].

$$u_{k,c}(i), u_{k,c}(j) > u_{k,c}(n)(n = i+1, i+2, \ldots, j-1). \tag{12}$$

where $u_{k,c}(n)$ is intermediate points between ith node and jth node.

In this paper, the weights are calculated using the radix function proposed by Supriya et al. [12] to construct a temporal undirected weighted network, and the weight between ith and jth sample points can be calculated as:

$$w_{i,j} = \left| \arctan\left(\frac{u_{k,c}(i) - u_{k,c}(j)}{i - j} \right) \right|. \tag{13}$$

where arctan helps to detect the sudden change in EEG signals.

2.6. Feature extraction and selection

There is similar information in features, so selecting the most representative features is the key to improve the classification performance. In spatial network, we extract four features from different segments, including weighted degree, harmonic centrality, closeness centrality and eigenvector centrality. The weighted degree and clustering coefficient are extracted as features in temporal network. Weighted degree and closeness centrality measure the importance of node in network. Harmonic centrality reflects the ease of reaching between nodes. Eigenvector centrality reflects the influence of neighboring nodes on the importance of node. Clustering coefficient represents the degree of aggregation of nodes in a graph. Specific calculation formula is shown in Table 1, where M is the set of all nodes in the network, and n is the

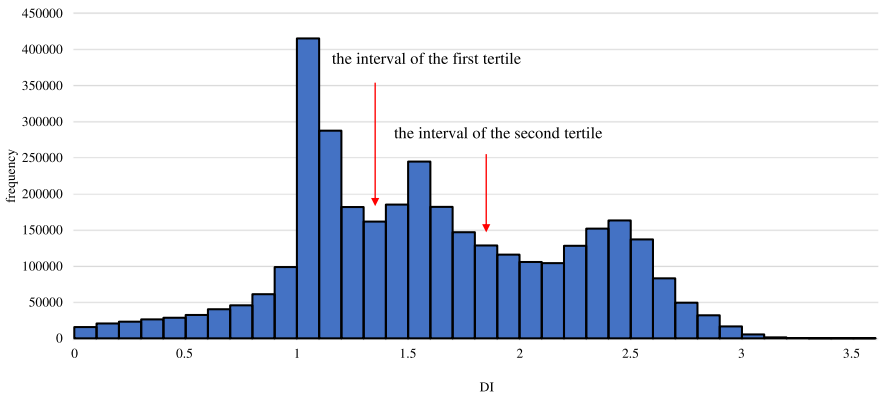

Fig. 3. The distribution histogram of DI.

Table 1
The extracted features.

Feature	Descriptions
Weighted degree[a,b]	$k_i = \sum_{j \in M} w_{i,j}$
Harmonic centrality[a]	$hc_i = \frac{1}{n-1} \sum_{j=1}^{n-1} \frac{1}{d_{i,j}}$
Closeness centrality[a]	$c_i = \frac{n-1}{\sum_{j \neq i} d_{i,j}}$
Eigenvector centrality[a]	$e_i = \frac{1}{\lambda} \sum_{j=1}^{n} w_{i,j} e_j$
Clustering coefficient[b]	$cc = \frac{1}{n} \sum_{i \in M} \frac{\sum_{j,h} (w_{i,j} w_{i,h} w_{j,h})^{\frac{1}{3}}}{k_i(k_i-1)}$

[a]Denotes that the feature is extracted from spatial network.
[b]Denotes that the feature is extracted from temporal network.

number of nodes. $w_{i,j}$ and $d_{i,j}$ are the shortest path length and weight between nodes i and j.

In feature selection, basing on K-Nearest Neighbor (KNN), we removes (backward selection) features to form a feature subset in a greedy algorithm. At each stage, KNN chooses the worst feature to remove which is based on the cross-validation accuracy score. The specific feature extraction and selection steps are shown in Fig. 4. Since we want to get information from different networks, the clustering coefficient and the weighted degree are obtained as features after feature selection for the temporal and spatial networks separately. They are calculated as follows.

Since there are only six features extracted in this paper, and all of them are network features, the computational power of features is small. However, the computational power of the overall model mainly depends on MVMD and feature extraction, so the computational power required by the model is within an acceptable range.

3. Result

In terms of the presentation of results, we have mainly carried out the following works.

1. Discussing the effect of the embedding dimension m and the number of linear map intervals s of DI on the detection results after feature selection.
2. Demonstrating the effectiveness of DI as a spatial network weight and the effectiveness of spatio-temporal network by comparing the experiments.

3.1. Random forest classifier

We use RF proposed by Breiman et al. [31] in our work. It generates a new training sample set by integrating multiple weak classifiers (trees) and repeats randomly selected samples using bootstrap resampling technique, and the final classification results are obtained by voting method. RF has faster training, higher model accuracy and generalization ability when there are more samples and features [32] and is more suitable for classification of large datasets in this paper. In the following, this paper will show a comparison of the results of RF, support vector machine (SVM) and logistic regression (LR) classifiers.

In our results, 75% of all data are randomly selected as the training set and 25% as the test set. Of course, the data in both the training and test sets are equalized. We hope that the generalization ability of the model can be enhanced to be more applicable in real situations by considering all data.

3.2. Evaluation index

In this paper, we use 5-fold cross-validation to test the performance of our model on training set to select the locally optimal classifier parameters by grid search. Finally, we calculate the six indicators of accuracy (ACC), precision (PRE), sensitivity (SEN), specificity (SPE), F1 score (F1) and area under the ROC curve (AUC) on test set as classification results, which calculate through confusion matrices as follows.

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

$$PRE = \frac{TP}{TP + FP}. \tag{15}$$

$$SEN = \frac{TP}{TP + FN}. \tag{16}$$

$$SPE = \frac{TN}{TN + FP}. \tag{17}$$

$$F1 = \frac{2PRE \cdot SEN}{PRE + SEN}. \tag{18}$$

3.3. CHB-MIT dataset result

The lag d is usually taken as 1, while the embedding dimension and the number of linear map intervals have a large impact on the results, depending on different situations [23]. The ternary value of DI may cause different thresholds under different parameter networks, resulting in different accuracy of classification. Therefore, we use the unternary-valued network for the analysis of parameters. The experiments results for the embedding dimension are shown in Table 2.

In Table 2, this paper compares the differences in feature selection between two classifiers, RF and KNN, for different m values. The results show that when $s = 3$, the features selected by KNN and RF in the cases of $m = 4$ and $m = 5$ are different, and the features selected by RF for the spatial network are closeness centrality and eigenvector centrality, respectively, while KNN selects both weighted degree. Comparison of average shows that $m = 4$ performs the best and KNN is better than RF.

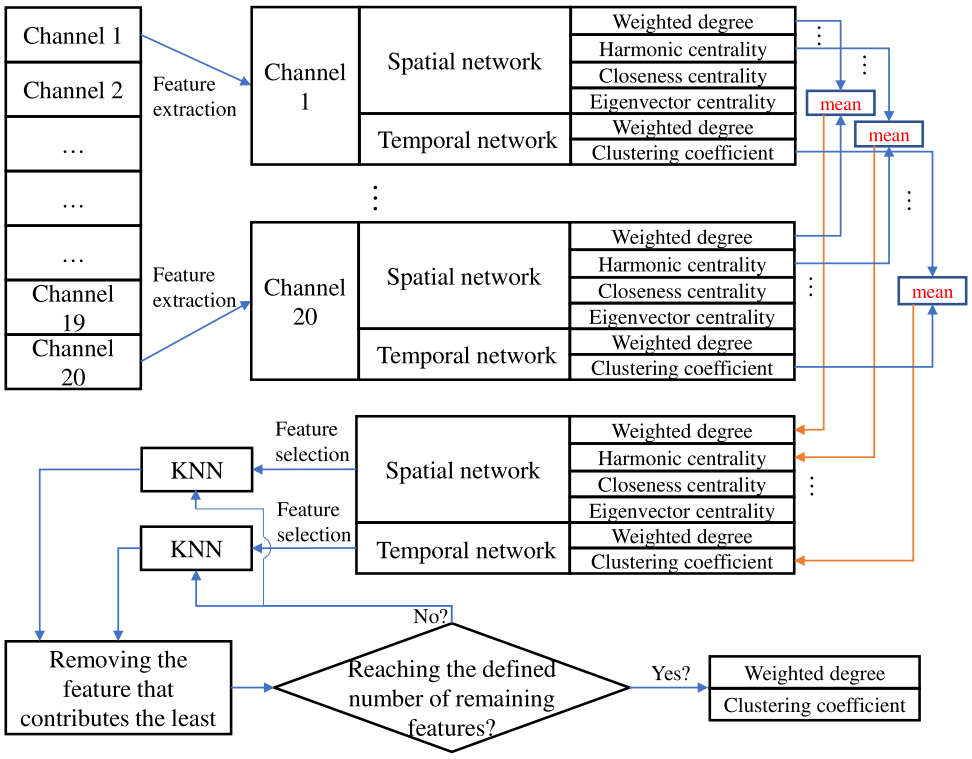

Fig. 4. Feature extraction and selection. Firstly, we extract four features and two features from spatial and temporal networks separately for every channel. Secondly, we characterize the segments by calculating the average of all channels. Finally, we remove the worst feature based on KNN until the defined number of features reaches.

Table 2
Changes in feature selection due to different m when $s = 3$ on CHB-MIT dataset.

Indicator	Classifier	m = 2	m = 3	m = 4	m = 5	m = 6	m = 7	m = 8
F1	RF	86.60%	87.67%	85.73%	73.46%	87.65%	86.73%	85.73%
	KNN	86.69%	87.67%	88.10%	87.39%	87.71%	86.75%	85.75%
AUC	RF	84.33%	85.63%	83.18%	69.16%	85.52%	84.58%	83.68%
	KNN	84.41%	85.62%	86.06%	85.05%	85.53%	84.59%	83.69%
ACC	RF	84.74%	86.01%	83.66%	69.81%	85.92%	84.98%	83.98%
	KNN	84.83%	86.01%	86.44%	85.55%	85.95%	85.00%	83.99%
PRE	RF	85.60%	86.49%	84.56%	72.60%	86.71%	85.46%	85.20%
	KNN	85.60%	86.47%	87.26%	85.78%	86.61%	85.45%	85.18%
SEN	RF	87.61%	88.87%	86.94%	74.34%	88.60%	88.04%	86.26%
	KNN	87.81%	88.90%	88.97%	89.06%	88.83%	88.09%	86.32%
SPE	RF	81.05%	82.39%	79.42%	63.99%	82.44%	81.11%	81.10%
	KNN	81.01%	82.35%	83.15%	81.05%	82.23%	81.09%	81.06%
Average	RF	84.99%	86.18%	83.92%	70.56%	86.14%	85.15%	84.32%
	KNN	85.06%	86.17%	86.66%	85.65%	86.14%	85.16%	84.33%

Then the effect of the number of linear map intervals on the results is analyzed and shown in Table 3. When $s = 7$, the feature chosen by RF in spatial networks is harmonic centrality while KNN chooses eigenvector centrality. Combining two parameters and classifier analysis, KNN performs the best results when $s = 2$ and $m = 3$. This set of parameters is used for the next analysis. The results in Tables 2 and 3 run 100 times, so the results of RF tend to be stable after multiple runs, and will fluctuate within 0.23%. All classification results shown in this paper are then run 100 times.

To demonstrate the effectiveness of DI as network weights, we respectively compare the results of seizure detection with mutual information (MI) and DI as network weights. The results in Fig. 5 show that the results of DI on CHB-MIT dataset are lower than MI in all indicators, with an average of 0.11% lower. However, this difference is very small, so this paper considers DI to be as effective as MI on CHB-MIT dataset.

We also compare the results of spatial network, temporal network and spatio-temporal network for seizure detection in this paper (Table 4). Among them, spatial network extracts four features such as weighted degree, harmonic centrality, closeness centrality and eigenvector centrality, while temporal network extracts two features, including weighted degree and clustering coefficient. The results on CHB-MIT dataset show that spatio-temporal network significantly outperforms the single temporal network or spatial network, with an average increase of 2.13% and 4.26%, respectively.

When using DI to classify all subjects separately, the results are shown in Table 5. According to Table 5, chb06, chb12 and chb13 have poor indicators. The performance of chb6 is the worst whose F1, AUC, ACC, PRE, SEN, and SPE are 92.58%, 92.48%, 92.48%, 91.36%, 93.86% and 91.11%, respectively. For the other 21 subjects, almost all of the indicators are above 97%. In chb05 and chb19, all indicators even exceed 99.5%. In addition, among all subjects, F1, AUC, ACC, and SEN are less volatile, with variances of 0.000335, 0.000398, 0.000384,

Table 3
Changes in feature selection due to different s when $m = 3$ on CHB-MIT dataset.

Indicator	Classifier	s = 2	s = 3	s = 4	s = 5	s = 6	s = 7	s = 8
F1	RF	91.03%	87.67%	90.09%	89.71%	89.53%	85.80%	87.47%
	KNN	91.00%	87.67%	90.09%	89.69%	89.53%	75.10%	87.45%
AUC	RF	90.04%	85.63%	88.72%	88.35%	88.24%	83.40%	85.27%
	KNN	90.04%	85.62%	88.73%	88.34%	88.24%	71.85%	85.26%
ACC	RF	90.13%	86.01%	88.91%	88.54%	88.40%	83.85%	85.72%
	KNN	90.12%	86.01%	88.91%	88.52%	88.40%	72.21%	85.70%
PRE	RF	91.06%	86.49%	89.84%	89.37%	89.31%	84.52%	85.97%
	KNN	91.16%	86.47%	89.85%	89.38%	89.30%	75.37%	85.97%
SEN	RF	90.99%	88.87%	90.34%	90.06%	89.75%	87.11%	89.02%
	KNN	90.84%	88.90%	90.34%	90.01%	89.77%	74.84%	88.97%
SPE	RF	89.09%	82.39%	87.10%	86.64%	86.73%	79.69%	81.52%
	KNN	89.24%	82.35%	87.11%	86.67%	86.71%	68.86%	81.54%
Average	RF	90.39%	86.18%	89.17%	88.78%	88.66%	84.06%	85.83%
	KNN	90.40%	86.17%	89.17%	88.77%	88.66%	73.04%	85.82%

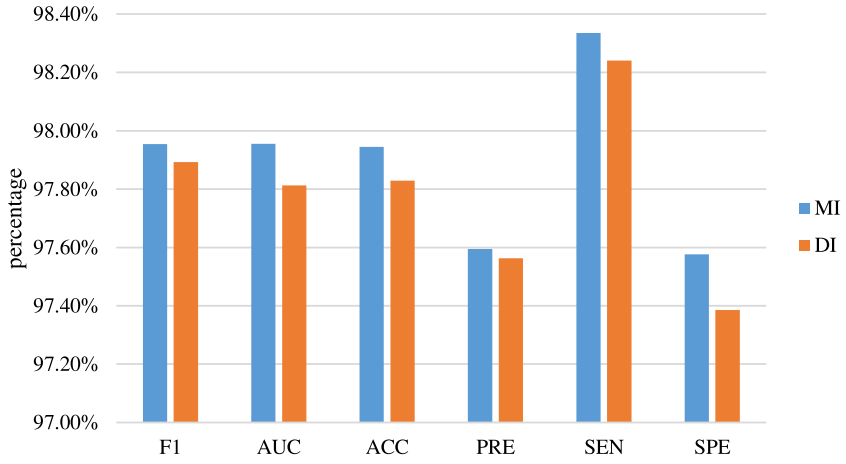

Fig. 5. Comparison of MI with the results of DI on CHB-MIT dataset.

Table 4
Comparison of spatial, temporal and spatio-temporal networks on CHB-MIT dataset.

Network	F1	AUC	ACC	PRE	SEN	SPE
Spatio	95.90%	92.30%	94.30%	94.06%	97.82%	86.78%
Temporal	95.82%	95.50%	95.50%	96.11%	95.53%	95.47%
Spatio-temporal	**97.89%**	**97.81%**	**97.83%**	**97.56%**	**98.24%**	**97.39%**

and 0.000227, respectively, while PRE and SPE are more volatile, with variances of 0.000567 and 0.00076.

In comparison with the results of other papers which are listed in Table 6, our method exhibits good results. Relative to the results of other papers, our method considers more number of subjects, more number of channels, and less number of extracted features. In the results of [34,36], they use fewer channels and achieve better results on certain indicators than in this paper. The ACC and SPE of [34] are 99.62% and 99.72%, respectively, while the ACC of [36] is 99.84%. In contrast, our results are more comprehensive in terms of the performance of each indicator, with F1, AUC, ACC, PRE, SEN and SPE of 97.89%, 97.81%, 97.83%, 97.56%, 98.24% and 97.39%, respectively. In addition, the sensitivity reaches 98.24% in proposed algorithm, which is better than the results of other papers. Our proposed method shows better results in most indicators and is more balanced among the indicators than other papers. This paper achieves efficient automatic seizure detection on short segment length. Compared with the features automatically extracted by deep learning (DL), the manually extracted features in this paper have specific formulas for calculation. In addition, the learning process of DL is a black-box operation and does not have a good interpretation, while the method proposed in this paper has.

Table 5
Results for all subjects on CHB-MIT dataset.

Subjects	F1	AUC	ACC	PRE	SEN	SPE
chb01	98.74%	98.79%	98.81%	99.01%	98.47%	99.11%
chb02	99.05%	99.07%	99.06%	98.56%	99.55%	98.59%
chb03	99.39%	99.46%	99.43%	98.80%	99.98%	98.95%
chb04	98.34%	98.38%	98.38%	98.46%	98.23%	98.52%
chb05	99.65%	99.66%	99.66%	99.62%	99.68%	99.63%
chb06	92.58%	92.48%	92.48%	91.36%	93.86%	91.11%
chb07	99.03%	99.04%	99.04%	99.45%	98.62%	99.45%
chb08	97.35%	97.45%	97.43%	96.68%	98.03%	96.87%
chb09	98.85%	98.84%	98.82%	99.54%	98.17%	99.52%
chb10	99.20%	99.16%	99.17%	98.87%	99.53%	98.79%
chb11	99.57%	99.58%	99.59%	99.93%	99.22%	99.93%
chb12	95.91%	95.54%	95.56%	95.99%	95.83%	95.25%
chb13	94.86%	93.82%	94.07%	94.14%	95.59%	92.04%
chb14	97.24%	97.18%	97.17%	97.64%	96.85%	97.52%
chb15	97.92%	97.94%	97.94%	97.70%	98.14%	97.74%
chb16	94.68%	94.09%	94.25%	91.53%	98.05%	90.12%
chb17	99.43%	99.42%	99.42%	99.21%	99.66%	99.19%
chb18	98.03%	97.97%	97.98%	96.84%	99.24%	96.70%
chb19	99.80%	99.82%	99.81%	99.61%	99.98%	99.67%
chb20	96.31%	96.27%	96.27%	95.18%	97.46%	95.09%
chb21	98.13%	98.11%	98.11%	97.79%	98.48%	97.73%
chb22	98.98%	99.01%	99.00%	98.79%	99.18%	98.84%
chb23	98.56%	98.61%	98.61%	98.55%	98.58%	98.64%
chb24	97.83%	97.82%	97.81%	98.27%	97.40%	98.23%
Average	97.89%	97.81%	97.83%	97.56%	98.24%	97.39%

Table 6
Comparison of performance on CHB-MIT dataset.

Author	Sub. \ Cha. \ Fea. \ Len.	F1	AUC	ACC	PRE	SEN	SPE
Sirpa et al.(2019) [33]	– \ – \ – \ –	–	94.00%	98.20%	–	95.90%	92.10%
Cao et al.(2019) [34]	22 \ 5 \ – \ 2	93.90%	–	99.62%	–	97.17%	**99.72%**
Gao et al.(2020) [35]	11 \ – \ 1 \ 4	–	–	92.60%	–	92.60%	97.10%
Anuragi et al.(2021) [36]	21 \ 1 \ 24 \ 10	–	–	**99.84%**	–	–	–
Peng et al.(2021) [37]	24 \ – \ – \ –	–	–	95.06%	–	95.38%	94.33%
Liu et al.(2022) [38]	24 \ 23 \ 2 \ 4	–	–	97.18%	–	97.10%	97.77%
Our works	24 \ 20 \ 2 \ 4	**97.89%**	**97.81%**	97.83%	**97.56%**	**98.24%**	97.39%

Sub.: Considered number of subjects.
Cha.: Considered number of channels.
Fea.: Extracted number of features.
Len.: Extracted length of each segment (Unit: s).

Table 7
Changes in feature selection due to different m when $s = 3$ on Siena dataset.

Indicator	Classifier	m = 2	m = 3	m = 4	m = 5	m = 6	m = 7
F1	RF	76.42%	73.32%	71.75%	69.12%	92.97%	92.28%
	KNN	76.43%	73.32%	71.80%	69.12%	92.87%	92.15%
AUC	RF	76.92%	73.85%	72.55%	70.35%	92.89%	92.34%
	KNN	76.85%	73.87%	72.55%	70.31%	92.81%	92.21%
ACC	RF	76.89%	73.83%	72.47%	70.34%	92.89%	92.33%
	KNN	76.83%	73.84%	72.47%	70.29%	92.81%	92.21%
PRE	RF	78.73%	75.58%	75.33%	72.33%	92.59%	91.64%
	KNN	78.49%	75.61%	75.25%	72.22%	92.68%	91.58%
SEN	RF	74.25%	71.20%	68.51%	66.19%	93.34%	92.92%
	KNN	74.48%	71.17%	68.65%	66.27%	93.06%	92.72%
SPE	RF	79.58%	76.50%	76.60%	74.52%	92.43%	91.75%
	KNN	79.22%	76.56%	76.45%	74.34%	92.56%	91.70%
Average	RF	77.13%	74.05%	72.87%	70.48%	92.85%	92.21%
	KNN	77.05%	74.06%	72.86%	70.42%	92.80%	92.10%

Table 8
Changes in feature selection due to different s when $m = 3$ on Siena dataset.

Indicator	Classifier	s = 2	s = 3	s = 4	s = 5	s = 6	s = 7
F1	RF	95.04%	73.32%	92.85%	92.90%	88.85%	91.86%
	KNN	95.09%	73.32%	92.79%	92.92%	88.93%	92.18%
AUC	RF	94.93%	73.85%	92.83%	92.99%	89.11%	91.78%
	KNN	94.98%	73.87%	92.77%	93.01%	89.18%	92.05%
ACC	RF	94.92%	73.83%	92.82%	92.98%	89.13%	91.78%
	KNN	94.96%	73.84%	92.77%	93.00%	89.20%	92.05%
PRE	RF	95.73%	75.58%	92.17%	92.17%	89.95%	91.42%
	KNN	95.79%	75.61%	92.13%	92.18%	89.96%	91.11%
SEN	RF	94.37%	71.20%	93.53%	93.63%	87.78%	92.31%
	KNN	94.39%	71.17%	93.46%	93.66%	87.92%	93.28%
SPE	RF	95.50%	76.50%	92.12%	92.34%	90.44%	91.25%
	KNN	95.57%	76.56%	92.08%	92.35%	90.44%	90.81%
Average	RF	95.08%	74.05%	92.72%	92.84%	89.21%	91.73%
	KNN	95.13%	74.06%	92.67%	92.85%	89.27%	91.91%

3.4. Siena scalp EEG dataset result

Similarly, we discuss the effect of parameters and classifiers on the results. It can be seen from Table 7 that with the increase of embedding dimension, all indicators show a general upward trend. The KNN and RF classifiers select the same features. In terms of the overall mean, the results show best at $m = 6$. We then analyze the number of linear map intervals. According to the results from Table 8, when $s = 7$, feature chosen by RF in spatial network is harmonic centrality, while KNN chooses weighted degree. At $s = 2$, mean value of the classification results reaches maximum. Combining the analysis of these two parameters, it has best results when $s = 2$ and $m = 3$. So we select this set of parameters to calculate the results in the subsequent experiments. In addition, the results of RF and KNN classifiers on Siena dataset are essentially the same. Combining with results of classifiers on CHB-MIT dataset, KNN is chosen for feature selection over RF.

We compare the results of MI and DI after ternary value. As shown in Fig. 6, DI as a weight in Siena scalp EEG dataset yields significantly higher results than MI for all indicators, with an average of 1.05% higher. This demonstrates the validity of DI as a channel-to-channel metric. In Siena scalp EEG dataset, all indicators exceed 99% and the SEN reaches 99.14%. The MI and DI classification results are compared after running multiple times, so that differences of more than 1% are considered significant. This proves that our proposed algorithm is effective for seizure detection.

We likewise compare the effects of spatial, temporal and spatio-temporal networks on this dataset. The results in Table 9 show that spatio-temporal network obtains better seizure detection than temporal network, with an average of 1.85% higher. Compared with the results of spatial network, the results of spatio-temporal network is 1.69% higher on average. Overall, the results of spatio-temporal network show a more balanced and comprehensive performance and are better than those obtained from temporal or spatial network.

Table 9
Comparison of spatial, temporal and spatio-temporal network on Siena scalp EEG dataset.

Network	F1	AUC	ACC	PRE	SEN	SPE
Spatio	98.74%	96.52%	97.89%	98.91%	98.56%	94.48%
Temporal	97.36%	97.31%	97.30%	97.53%	97.19%	97.42%
Spatio-temporal	**99.21%**	**99.19%**	**99.19%**	**99.28%**	**99.14%**	**99.24%**

Table 10
Comparison of classification results of different classifiers on two datasets.

Dataset	Classifier	F1	AUC	ACC	PRE	SEN	SPE
CHB-MIT	RF	97.89%	97.81%	97.83%	97.56%	98.24%	97.39%
	SVM	87.10%	87.09%	87.08%	89.56%	85.29%	88.89%
	LR	74.54%	74.99%	74.98%	78.95%	71.34%	78.64%
Siena	RF	99.21%	99.19%	99.19%	99.28%	99.14%	99.24%
	SVM	87.52%	87.62%	87.55%	90.20%	85.00%	90.24%
	LR	74.45%	76.21%	75.99%	82.13%	68.08%	84.35%

In addition, this paper also compares the classification effects of RF, SVM and LR in two datasets (Table 10). Among them, the kernel function of SVM is radial basis function. According to Table 10, results show that RF is the most effective one.

4. Conclusion

In this paper, a new metric DI between channels is proposed, and the effectiveness of DI as a channel association metric is explored by comparing the experimental results of DI and MI. In addition, this study also explores the improvement of spatio-temporal networks for classification accuracy. In the previous section we test the method proposed in this paper on CHB-MIT dataset and Siena scalp EEG dataset, respectively. For the CHB-MIT dataset, the classification results for F1, AUC, ACC, PRE, SEN and SPE are 97.89%, 97.81%, 97.83%, 97.56%, 98.24% and

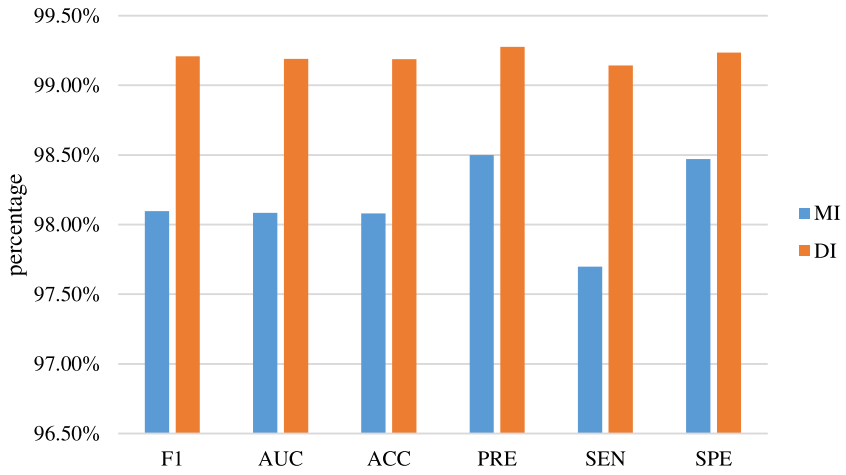

Fig. 6. Comparison of mutual information(MI) with the results of DI on Siena scalp EEG dataset.

97.39%, respectively. For the Siena scalp EEG dataset, the classification results for F1, AUC, ACC, PRE, SEN and SPE are 99.21%, 99.19%, 99.19%, 99.28%, 99.14% and 99.24%, respectively. The experimental results show that DI has better detection performance than MI, and the spatio-temporal network is more effective than a single temporal or spatial network. The main contributions of this paper are as follows:

(1) A new metric DI of channel correlation is proposed.
(2) Integrating temporal network and spatial network for seizure detection
(3) The effect of DI on different parameters is tested, and a new idea for automatic seizure detection is proposed.

Despite the excellent accuracy of seizure classification obtained in our work, there is still room for improvement. Next, we can further consider different length of segments for research to find the optimal length of EEG signal segments. In addition, the research in this paper focuses on classifying EEG signals into seizure and non-seizure. Further multiple classifications of pre-seizure and post-seizure cases can be considered.

CRediT authorship contribution statement

Yuhuan Xiong: Conceptualization, Methodology, Investigation, Writing – original draft. **Jinghan Li:** Validation, Funding acquisition, Methodology, Writing – review & editing. **Duanpo Wu:** Conceptualization, Methodology, Writing – review & editing, Supervision, Funding acquisition. **Fang Dong:** Conceptualization, Formal analysis, Writing, Funding acquisition. **Junbiao Liu:** Methodology, Formal analysis. **Lurong Jiang:** Validation, Formal analysis. **Jiuwen Cao:** Methodology, Investigation, Funding acquisition. **Yuansheng Xu:** Validation, Investigation.

Declaration of competing interest

One or more of the authors of this paper have disclosed potential or pertinent conflicts of interest, which may include receipt of payment, either direct or indirect, institutional support, or association with an entity in the biomedical field which may be perceived to have potential conflict of interest with this work. For full disclosure statements refer to https://doi.org/10.1016/j.bspc.2022.104155. Duanpo Wu reports financial support was provided by Zhejiang Province Natural Science Foundation. Jiuwen Cao reports financial support was provided by National Natural Science Foundation of China. Jiuwen Cao reports financial support was provided by Ministry of Science and Technology of the People's Republic of China. Duanpo Wu reports financial support was provided by Zhejiang Province Science and Technology Department. Jiuwen Cao reports financial support was provided by Zhejiang University-Zhijiang Campus. Duanpo Wu reports financial support was provided by Zhejiang University.

Data availability

The data used in the paper is open access.

Acknowledgments

This project has received funding from the National Key R & D Program, China (2021YFE0100100), Joint Fund of Zhejiang Provincial Natural Science Foundation of China (LBY21H090001), Key R & D Programs in Zhejiang Province, China (2020C03038), NSFC-Zhejiang Integration Joint Fund, China (U1909209), the Open Research Projects of Zhejiang Lab, China (2021MC0AB04), the open project of Zhejiang Provincial Key Laboratory of Information Processing, Communication and Networking, Zhejiang, China, and Hangzhou Agricultural and Social Development Scientific Research Project, China (20201203B145).

References

[1] R.D. Thijs, R. Surges, T.J. O'Brien, J.W. Sander, Epilepsy in adults, Lancet 393 (10172) (2019) 689–701.
[2] P.J. Karoly, V.R. Rao, N.M. Gregg, G.A. Worrell, C. Bernard, M.J. Cook, M.O. Baud, Cycles in epilepsy, Nature Rev. Neurol. 17 (5) (2021) 267–284.
[3] N.E. Huang, Z. Shen, S.R. Long, M.C. Wu, H.H. Shih, Q. Zheng, N.-C. Yen, C.C. Tung, H.H. Liu, The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. 454 (1971) (1998) 903–995.
[4] K. Dragomiretskiy, D. Zosso, Variational mode decomposition, IEEE Trans. Signal Process. 62 (3) (2013) 531–544.
[5] F. Xiao, D. Yang, Z. Lv, X. Guo, Z. Liu, Y. Wang, Classification of hand movements using variational mode decomposition and composite permutation entropy index with surface electromyogram signals, Future Gener. Comput. Syst. 110 (2020) 1023–1036.
[6] N. ur Rehman, H. Aftab, Multivariate variational mode decomposition, IEEE Trans. Signal Process. 67 (23) (2019) 6039–6052.
[7] C. Gu, X. Qiao, Y. Jin, Y. Liu, A novel fault diagnosis method for diesel engine based on MVMD and band energy, Shock Vib. 2020 (2020).
[8] C. Zhao, J. Sun, S. Lin, Y. Peng, Fault diagnosis method for rolling mill multi row bearings based on AMVMD-MC1DCNN under unbalanced dataset, Sensors 21 (16) (2021) 5494.
[9] K. Das, R.B. Pachori, Schizophrenia detection technique using multivariate iterative filtering and multichannel EEG signals, Biomed. Signal Process. Control 67 (2021) 102525.
[10] L. Lacasa, B. Luque, F. Ballesteros, J. Luque, J.C. Nuno, From time series to complex networks: The visibility graph, Proc. Natl. Acad. Sci. 105 (13) (2008) 4972–4975.
[11] B. Luque, L. Lacasa, F. Ballesteros, J. Luque, Horizontal visibility graphs: Exact results for random time series, Phys. Rev. E 80 (4) (2009) 046103.

[12] S. Supriya, S. Siuly, H. Wang, J. Cao, Y. Zhang, Weighted visibility graph with complex network features in the detection of epilepsy, IEEE Access 4 (2016) 6554–6566.
[13] Z.-K. Gao, Q. Cai, Y.-X. Yang, W.-D. Dang, S.-S. Zhang, Multiscale limited penetrable horizontal visibility graph for analyzing nonlinear time series, Sci. Rep. 6 (1) (2016) 1–7.
[14] H. Xiong, P. Shang, F. Hou, Y. Ma, Visibility graph analysis of temporal irreversibility in sleep electroencephalograms, Nonlinear Dynam. 96 (1) (2019) 1–11.
[15] S. Supriya, S. Siuly, H. Wang, Y. Zhang, EEG sleep stages analysis and classification based on weighed complex network features, IEEE Trans. Emerg. Top. Comput. Intell. 5 (2) (2018) 236–246.
[16] T. Wadhera, Brain network topology unraveling epilepsy and ASD association: Automated EEG-based diagnostic model, Expert Syst. Appl. 186 (2021) 115762.
[17] L. Yao, M. Wang, Y. Lu, H. Li, X. Zhang, EEG-based emotion recognition by exploiting fused network entropy measures of complex networks across subjects, Entropy 23 (8) (2021) 984.
[18] Z. Gao, S. Li, Q. Cai, W. Dang, Y. Yang, C. Mu, P. Hui, Relative wavelet entropy complex network for improving EEG-based fatigue driving classification, IEEE Trans. Instrum. Meas. 68 (7) (2018) 2491–2497.
[19] S. Lotfan, S. Shahyad, R. Khosrowabadi, A. Mohammadi, B. Hatef, Support vector machine classification of brain states exposed to social stress test using EEG-based brain network measures, Biocybern. Biomed. Eng. 39 (1) (2019) 199–213.
[20] P.E. Davis, K. Kapur, R. Filip-Dhima, S.K. Trowbridge, E. Little, A. Wilson, A. Leuchter, E.M. Bebin, D. Krueger, H. Northrup, et al., Increased electroencephalography connectivity precedes epileptic spasm onset in infants with tuberous sclerosis complex, Epilepsia 60 (8) (2019) 1721–1732.
[21] M. Ma, X. Wei, Y. Cheng, Z. Chen, Y. Zhou, Spatiotemporal evolution of epileptic seizure based on mutual information and dynamic brain network, BMC Med. Inform. Decis. Mak. 21 (2) (2021) 1–11.
[22] G. Sdoukopoulou, M. Antonakakis, G. Modde, C.H. Wolters, M. Zervakis, Interictal spike classification in pharmacoresistant epilepsy using combined EEG and MEG, in: 2021 IEEE 21st International Conference on Bioinformatics and Bioengineering (BIBE), IEEE, 2021, pp. 1–6.
[23] M. Rostaghi, H. Azami, Dispersion entropy: A measure for time-series analysis, IEEE Signal Process. Lett. 23 (5) (2016) 610–614.
[24] G. Zhu, Y. Li, P.P. Wen, Epileptic seizure detection in EEGs signals using a fast weighted horizontal visibility algorithm, Comput. Methods Programs Biomed. 115 (2) (2014) 64–75.
[25] A.H. Shoeb, Application of Machine Learning to Epileptic Seizure Onset Detection and Treatment (Ph.D. thesis), Massachusetts Institute of Technology, 2009.
[26] CHB-MIT scalp EEG database (version 1.0.0), 2022, PhysioNet. [Online]. Available: https://www.physionet.org/content/chbmit/1.0.0/.
[27] P. Detti, G. Vatti, G. Zabalo Manrique de Lara, Eeg synchronization analysis for seizure prediction: A study on data of noninvasive recordings, Processes 8 (7) (2020) 846.
[28] P. Detti, Siena scalp EEG database (version 1.0.0), 2022, PhysioNet. [Online]. Available: https://www.physionet.org/content/siena--scalp--eeg/1.0.0/.
[29] D. Wu, Z. Wang, L. Jiang, F. Dong, X. Wu, S. Wang, Y. Ding, Automatic epileptic seizures joint detection algorithm based on improved multi-domain feature of cEEG and spike feature of aEEG, IEEE Access 7 (2019) 41551–41564.
[30] F.A. Alturki, K. AlSharabi, A.M. Abdurraqeeb, M. Aljalal, EEG signal analysis for diagnosing neurological disorders using discrete wavelet transform and intelligent techniques, Sensors 20 (9) (2020) 2505.
[31] L. Breiman, Random forests, Mach. Learn. 45 (1) (2001) 5–32.
[32] L. Zhang, Z. Liu, T. Ren, D. Liu, Z. Ma, L. Tong, C. Zhang, T. Zhou, X. Zhang, S. Li, Identification of seed maize fields with high spatial resolution and multiple spectral remote sensing using random forest classifier, Remote Sens. 12 (3) (2020) 362.
[33] P. Sirpal, A. Kassab, P. Pouliot, D.K. Nguyen, F. Lesage, fNIRs improves seizure detection in multimodal EEG-fNIRS recordings, J. Biomed. Opt. 24 (5) (2019) 051408.
[34] J. Cao, J. Zhu, W. Hu, A. Kummert, Epileptic signal classification with deep EEG features by stacked CNNs, IEEE Trans. Cognit. Dev. Syst. 12 (4) (2019) 709–722.
[35] Y. Gao, B. Gao, Q. Chen, J. Liu, Y. Zhang, Deep convolutional neural network-based epileptic electroencephalogram (EEG) signal classification, Front. Neurol. 11 (2020) 375.
[36] A. Anuragi, D.S. Sisodia, R.B. Pachori, Automated FBSE-EWT based learning framework for detection of epileptic seizures using time-segmented EEG signals, Comput. Biol. Med. 136 (2021) 104708.
[37] H. Peng, C. Li, J. Chao, T. Wang, C. Zhao, X. Huo, B. Hu, A novel automatic classification detection for epileptic seizure based on dictionary learning and sparse representation, Neurocomputing 424 (2021) 179–192.
[38] H. Liu, Y. Gao, J. Zhang, J. Zhang, Epilepsy EEG classification method based on supervised locality preserving canonical correlation analysis, Math. Biosci. Eng. 19 (1) (2022) 624–642.