

EEG Signal Epilepsy Detection With a Weighted Neighbor Graph Representation and Two-Stream Graph-Based Framework

Jialin Wang^{1b}, Shen Liang, Jiawei Zhang, Yingpei Wu, Lanying Zhang, Rui Gao^{1b}, *Member, IEEE*,
Dake He, and C.-J. Richard Shi^{2b}, *Fellow, IEEE*

Abstract—Epilepsy is one of the most common neurological diseases. Clinically, epileptic seizure detection is usually performed by analyzing electroencephalography (EEG) signals. At present, deep learning models have been widely used for single-channel EEG signal epilepsy detection, but this method is difficult to explain the classification results. Researchers have attempted to solve interpretive problems by combining graph representation of EEG signals with graph neural network models. Recently, the combination of *graph representations* and *graph neural network* (GNN) models has been increasingly applied to single-channel epilepsy detection. By this methodology, the raw EEG signal is transformed to its graph representation, and a GNN model is used to learn latent features and classify whether the data indicates an epileptic seizure episode. However, existing methods are faced with two major challenges. First, existing graph representations tend to have high time complexity as they generally require each vertex to traverse all other vertices to construct a graph structure. Some of them also have high space complexity for being dense. Second, while separate graph representations can be derived from a single-channel EEG signal in both time and frequency domains, existing GNN models for epilepsy detection can learn from a single graph

representation, which makes it hard to let the information from the two domains complement each other. For addressing these challenges, we propose a *Weighted Neighbour Graph* (WNG) representation for EEG signals. Reducing the redundant edges of the existing graph, WNG can be both time and space-efficient, and as informative as its less efficient counterparts. We then propose a two-stream graph-based framework to simultaneously learn features from WNG in both time and frequency domain. Extensive experiments demonstrate the effectiveness and efficiency of the proposed methods.

Index Terms—EEG signal, graph representation, graph neural network, weighted neighbour graph, seizure detection.

I. INTRODUCTION

EPILEPSY is a common neurological disease. Clinically, neurologists detect epilepsy by analyzing electroencephalography (EEG) signals [1]. However, manual epilepsy detection can often be time-consuming. Therefore, machine learning-enabled automatic epilepsy detection methods have attracted growing attention from both researchers and neurologists [2], [3], [4], [5].

Epilepsy detection is traditionally conducted by first extracting handcrafted features from raw EEG signals [6], [7], and then feeding these features into a classifier to determine whether the given signal contains epileptic segments [8], [9]. Such methods are suitable for cases which the different classes can be relatively distinguished between. However, EEG signals often come with complex waveforms with reach underlying semantics that cannot be easily characterized by handcrafted features [10]. With the development of deep learning techniques, researchers resort to deep learning models that can effectively learn complex features, especially when large amounts of data are available [11], [12].

Recently, researchers have used deep learning models directly to learn latent features from raw EEG signals [13], [14]. However, these raw EEG signals are highly random and heterogeneous. In particular, relationships of two or more data points can be of great value for epilepsy classification [15]. That limits the performance of deep learning models without explicit transformation of data representation [10], [16]. Using such information can be hard for deep learning models to interpret classification results from raw EEG signals.

To address this issue, researchers have turned to graph representations [17], [18], such as Visibility Graph (VG) [19].

Manuscript received 4 February 2023; revised 3 July 2023; accepted 19 July 2023. Date of publication 28 July 2023; date of current version 7 August 2023. This work was supported in part by the Science and Technology Commission of Shanghai Municipality under Grant 2018SHZDZX01 and Grant 2021SHZDZX; in part by the Natural Science Foundation of Shanghai Municipality under Grant 22ZR1406900; in part by the Young Scientist Project of Ministry of Education (MOE) Innovation Platform, Pujiang, Shanghai, under Grant 22PJ1405400; in part by the Guangdong Basic and Applied Basic Research Foundation under Grant 2022A1515110574; in part by the Data Intelligence Institute of Paris (diiP); and in part by IdEx Université Paris Cité under Grant ANR-18-IDEX-0001. (*Corresponding authors: Rui Gao; Dake He.*)

Jialin Wang and Lanying Zhang are with the State Key Laboratory of ASIC and Systems, Institute of Brain-Inspired Circuits and Systems (iBICAS), Fudan University, Shanghai 200433, China.

Shen Liang is with the Data Intelligence Institute of Paris (diiP), Université Paris Cité, 75013 Paris, France.

Jiawei Zhang is with the Department of New Networks, Peng Cheng Laboratory, Shenzhen, Guangdong 518055, China.

Yingpei Wu is with the School of Information Technology, Jiangxi University of Finance and Economics, Nanchang 330013, China.

Rui Gao is with the Department of Naval Architecture and Ocean Engineering, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: rgao@sjtu.edu.cn).

Dake He is with the Xinhua Hospital Affiliated to Shanghai Jiao Tong University School of Medicine, Shanghai 200092, China (e-mail: hedake@xinhua.com.cn).

C.-J. Richard Shi is with the Department of Electrical and Computer Engineering, University of Washington, Seattle, WA 98195 USA.

Digital Object Identifier 10.1109/TNSRE.2023.3299839

For a given single-channel signal, each vertex in its corresponding graph represents a data point in the signal, and each edge represents the pairwise relationship between two data points [20], and multiple edges represent the relationship of the multiple vertices that are connected by them. This method provides a new feature for EEG signals, namely the relationship feature between data points. Compared with the classic VG, limited penetrable visibility graph (LPVG) [21], horizontal visibility graph (HVG) [22], and limited penetrable horizontal visibility graph (LPHVG) [23], the Overlook Graph (OG) and Weighted Overlook Graph (WOG) have the weighted directed edges and the distinguishing ability of different categories have been significantly improved [24]. However, the time complexity and memory requirements of these methods are very large. It is only suitable for offline EEG classification tasks.

Existing Graph-based methods suffer from the following two problems.

1. High time and space complexity. For some existing graph representations [19], [21], it is required that each vertex traverse all other vertices to construct the graph structure, which leads to $O(n^2)$ (n is the number of data points in the raw signal) or high time complexity. Meanwhile, to maximally retain semantic information, some existing graphs [24] resort to dense graph representations that lead to high memory costs.

2. Inability to simultaneously learn from time and frequency domain signals. Raw EEG signals reside in time domain. However, a recent work [25] has highlighted the need for learning from graphs built upon frequency domain signals. While both time and frequency domain streams can be derived from one to the other, their different forms indicate that different semantic information can be better encoded in them. Therefore, it can be assumed that simultaneously learning from time and frequency domain graphs can lead to superior results. However, there exist very few deep learning models that can undertake this task.

In response to these problems, we make the following contributions in this paper.

1. Weighted Neighbour Graph (WNG). We propose a novel graph representation called Weighted Neighbour Graph (WNG), which is essentially a lossy compression of a state-of-the-art dense graph representation called Weighted Overlook Graph (WOG) [24]. Reducing the redundant edges of the WOG, WNG can strike a balance between the amount of information retained. Moreover, unlike previous graphs with high time and (or) space cost, WNG has both linear time and space complexity.

2. A two-stream graph-based framework for time and frequency domain graphs. Inspired by the two-stream convolutional neural network [26] for video analytics, we present a novel two-stream graph-based framework, which utilizes two parallel branches within a single GNN to simultaneously learn from time and frequency domain graphs. Note that this framework is generic. It is compatible with *any* graph representation and *any* GNN model that can learn from this representation.

3. Experimental evaluation. We conduct several real-world experiments to demonstrate the effectiveness and efficiency of the proposed methods.

For the rest of the paper, we introduce our WNG representation and the two-stream graph-based framework in Section II. Afterward, we report our experiments and results in Section III and IV. We discuss our framework in Section V. Finally, we conclude the paper in Section VI.

II. METHOD

A. Preliminaries

We first introduce the preliminaries required in the section. We denote an EEG signal with n data points as $T = t_1, t_2, \dots, t_n$, where t_i is the value of i th point in the signal. A graph representation of $T \in \mathbb{R}^n$ can be written as a graph $G = (V, E)$, where V is the vertex set with n vertices, each corresponding to a data point in T . E is the edge set, in which $e_{i,j} \in E$ denotes the edge connecting vertices v_i and v_j . $e_{i,j}$ is the pairwise relationship between data points t_i and t_j .

Definition (Weighted Overlook Graph (WOG)): Given a single-channel EEG signal $T = t_1, t_2, \dots, t_n$, its WOG representation can be written as a graph $G = (V, E)$. Given two arbitrary data points t_i and t_j in an EEG signal T , an *weighted, directed* edge $e_{i,j}$ is created if it suffices that:

$$t_i > t_j. \quad (1)$$

Let $A_T = \{a_{i,j} | i = 1, 2, \dots, n, j = 1, 2, \dots, n\}$ be the adjacency matrix of the time domain G , where

$$a_{i,j} = \begin{cases} \frac{t_i - t_j}{|i - j|} & \text{if } e_{i,j} \text{ exists;} \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

In this paper, we restrict our discussion to directed graphs. Hence, if $a_{i,j} = \frac{t_i - t_j}{|i - j|}$, then $a_{j,i} = \frac{t_j - t_i}{|j - i|}$.

B. Graph Representation

The graph representation method for EEG signals has a very large time complexity and memory requirements. The existing graph representation method needs to traverse all other vertices when building the graph structure, so its time complexity is large. The existing graph representation methods also require a large number of edges to represent the topology information of signals. In fact, the information of some edges in these graph structures can be inferred from the relationship between other edges, which increases the memory requirements of the graph (see II-B.2 for details). For example, the graph structure generated by the WOG method with the highest classification accuracy currently has a large number of redundant edges, which also greatly increases the time complexity of graph generation and memory requirements. Therefore, our goal is to design a weighted directed graph representation that can maintain the classification accuracy of EEG signals and minimize time complexity and memory requirements.

1) Weighted Rand Graph (WRG): We design a random probability $p \in [0, 1]$, which means that among n vertices in the graph structure, $p \times n$ points will be reconnected randomly [27].

WRG Connection Rule: Given one arbitrary data point t_i in an EEG signal $T \in \mathbb{R}^{n-1}$, a *weighted, directed* edge $e_{i,i+1}$ is created if it suffices that: $t_i > t_{i+1}$. Where $t_r = t_{r_1}, t_{r_2}, \dots, t_{r_{p \times n}}$ points will be disconnected and reconnected.

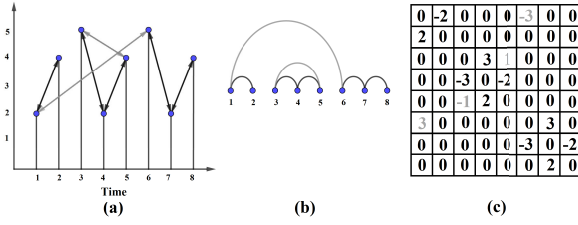


Fig. 1. We convert an EEG signal containing eight sampling points as a Weighted Random Graph (WRG). The vertex originally connected to the neighbouring vertex will disconnect the original connection according to the random probability and establish a connection relationship with another vertex.

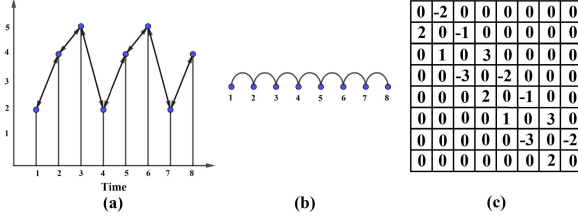


Fig. 2. An example of an EEG signal containing eight sampling points, construct a Weighted Neighbour Graph (WNG) representation structure that only connects the neighbour vertices on the nearest two sides and the corresponding adjacency matrix.

$\{r_i | a_{i,i+1} = 0, a_{i+1,i} = 0, a_{i,r_i} = t_i - t_{r_i}, a_{r_i,i} = t_{r_i} - t_i\}$ is a random reconnected number of i th point.

For example, the vertex $t_i(i, t_i)$ disconnects the original connection ($a_{i,i+1} = 0$ and $a_{i+1,i} = 0$), randomly generates a random connection ($r_i \in [1, n] \cap r_i \neq i$), and rebuilds the connection ($a_{i,r_i} = t_i - t_{r_i}$ and $a_{r_i,i} = t_{r_i} - t_i$) for the vertex $t_i(i, t_i)$. This representation do not increase edges, but effectively reduces the characteristic path length between vertices.

2) *Weighted Neighbour Graph (WNG)*: In the extreme case of the graph ($p = 0$), that is, each vertex only constructs edges with vertices adjacent in the signal. We call it the **Weighted Neighbour Graph (WNG)**, as shown in Fig. 2.

WNG Connection Rule: Given one arbitrary data point t_i in an EEG signal $T \in \mathbb{R}^{n-1}$, a *weighted, directed edge* $e_{i,i+1}$ is created if it suffices that:

$$t_i > t_{i+1}. \quad (3)$$

where the weight is $a_{i,i+1} = t_i - t_{i+1}$ for vertex $t_i(i, t_i)$. It is worth noting that here $a_{i+1,i} = t_{i+1} - t_i$, taking the opposite number of the values of two vertices at a diagonally symmetrical position along the adjacency matrix is a redundant operation. This is for the vertex $t_{i+1}(i+1, t_{i+1})$ to also aggregate the information of the vertex $t_i(i, t_i)$. The concept of positive or negative value means the direction of the edge between vertices. We reduce the time complexity of the graph representation method from $O(n^2)$ to $O(n)$. This reduces the time complexity of the graph generation. The memory requirement of the generated graph is greatly reduced due to the reduction of the edges.

We found the edge of WOG is $a_{i,j} = \frac{t_i - t_j}{i - j}$, where $i > j$, $a_{i,j} = \frac{a_{i-1,i} + a_{i-2,i-1} + \dots + a_{j,j+1}}{i - j}$ can be represented by edges in a WNG by a finite number of addition operations. $a_{i-1,i}, a_{i-2,i-1}, \dots, a_{j,j+1}$ are the adjacent edges of WNG.

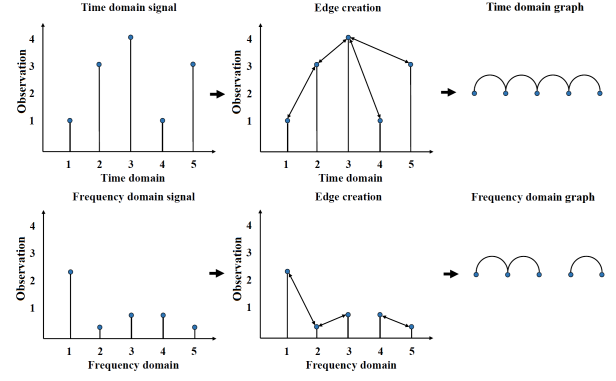


Fig. 3. This process converts a single-channel time domain EEG signal into the frequency domain graph representation.

When we get $a_{i,i+1}$, we can represent all the edges of WOG. The graph representation designed in this paper can effectively represent the information of signal with fewer edges.

3) *Frequency Domain Graph Representation*: These graph representation methods usually represent time domain signals and rarely represent the frequency domain information of EEG signals. There are many limitations in the time domain EEG signals. For example, it is difficult to align the phase of the signals, which results in a large variance of different EEG signals of the same category. The frequency domain feature can alleviate this problem as an essential signal feature [25], [28]. In addition, frequency domain graphs can provide interpretability for classification results (see V-B for details). The signal can be converted to the frequency domain and aligned strictly according to the frequency arrangement. The steps of the frequency domain graph are: 1) convert the signal to the frequency domain through Fast Fourier Transform (FFT); 2) represent the frequency domain signal into a graph structure.

We convert a time domain EEG signal into the frequency domain graph¹ is shown in Fig. 3. First, we transform the time domain EEG signal $T = t_1, t_2, \dots, t_n$ into the frequency domain using FFT. The signal is $|F| = |f_1|, |f_2|, \dots, |f_n|$, where

$$f_k = \sum_{i=1}^n t_i e^{-j \frac{2\pi}{n} ik} \quad (k = 1, 2, \dots, n). \quad (4)$$

Each f_k corresponds to the frequency k for all input time domain signals as long as they have the same sampling rate and length.

Then, we build the frequency domain graphs for the EEG signals. Specifically, we create a vertex for each data point in the signal and create edges based on some connection rule. We use the connection rule of the graph representation methods. However, we note that our frequency domain method is compatible with other graphs. The connection rule of the graph is as follows: For one data point $|f_i|$, if it suffices that: $|f_i| > |f_{i+1}|$. We create an edge connecting the vertices for $|f_i|$ and $|f_{i+1}|$. Each edge indicates the relationship between two frequencies in the graph, which can be written as a $n \times n$ adjacency matrix A_F .

¹In this paper, we only use the real part of the frequency domain signals, $|F| = |f_1|, |f_2|, \dots, |f_n|$.

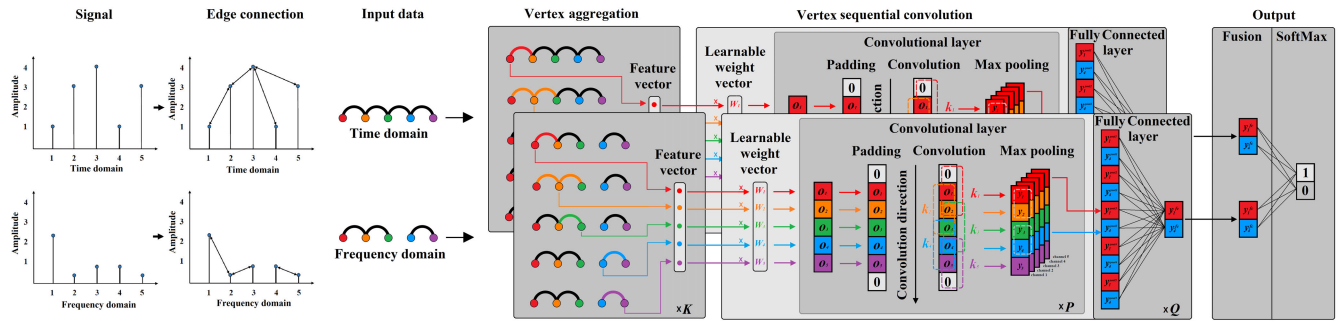


Fig. 4. The two-stream architecture consists of two parts: time and frequency domain. We use different colors to represent the order of vertices.

Then for vertex $i(i, |f_i|)$ there is $a_{i,j} = |f_i| - |f_j|$ (here $a_{j,i} = |f_j| - |f_i|$, which is consistent with the time domain.)

C. Two-Stream Graph-Based Framework

We design a two-stream Graph-based framework to extract features of the time and frequency domain graph representations.

The READOUT function of the graph neural network uses vertex aggregation (such as sum, average, and max) for graph classification [29], [30]. This is because there is only a connection relationship between vertices in the graph classification task, and there are no sequential features. However, the graph of signals contains sequential features between vertices. In addition, these graph neural network models used for graph classification extract the topological structure features of the graph by aggregating the features of neighbour vertices. However, for all graphs in the WNG, the topological structure of the chain structure is challenging to distinguish the differences between different categories. To this end, we designed a two-stream architecture based on deep learning models that simultaneously extracts the sequential features between the time and frequency domain graphs of EEG signals.

For the existing graph classification methods of graph neural networks, 1) multi-hop neighbourhood aggregation, 2) readout aggregation at the graph part [30], [31]. Through multiple iterations of the first step, the vertex obtains the characteristics of neighbour vertices. Through the second part, to finally get the characterization of the whole graph. However, the current graph neural network cannot extract the sequential information between the vertices in the graph. This feature will cause the performance of the existing graph neural network to deteriorate in classifying our graphs. Because there is a sequence relationship between vertices in a graph, these sequence features are essential and cannot be easily discarded.

For this reason, we propose a new graph-based framework. The purpose is to preserve the sequential information between the vertices in the graph while extracting the structure information of the graph.

We first involved two identical aggregation operations for the time and frequency domain graph structures, which have been used to aggregate each vertex in the time and frequency domain graph structures simultaneously. Specifically, the information between vertices is aggregated K ($1 \leq K$) times.

After completing the aggregation of the vertices in the time and frequency domain graph structures, respectively, the

learnable weight is used to extract the features of the vertices. Because the graph structure in time and frequency domain has sequential features, we still introduce the operation of sequential graph convolution to extract the sequential features of the vertices in time and frequency domain graph structures.

The time and frequency domain graph $G = (V, E)$ containing n vertices are input into the model respectively in Fig. 4 (Input data), where n is the number of the time and frequency domain sampling points. We perform vertex aggregation operations on graphs in the time and frequency domain, respectively, shown in Fig. 4 (vertex aggregation). In this part, K -hop aggregation is performed on each vertex of the graph, and K iterations are performed. Specifically, for vertex v_i^t and v_i^f , where N_i^t and N_i^f are its neighbour vertex, $h_{v_i^t}^k$ and $h_{v_i^f}^k$ are the feature values of the K iteration.

we initialize $h_{v_i^t}^0 = 1$ ($h_{v_i^f}^0 = 1$) for all $v_i^t \in V^t$ and $v_i^f \in V^f$. Then, in each iteration k ($1 \leq k \leq K$), we perform the following summation [30] for each v_i^t and v_i^f :

$$\begin{cases} h_{v_i^t}^k = h_{v_i^t}^{k-1} + \sum_{u_i^t} h_{u_i^t}^{k-1} \\ h_{v_i^f}^k = h_{v_i^f}^{k-1} + \sum_{u_i^f} h_{u_i^f}^{k-1} \end{cases} \quad (5)$$

where $u_i^t \in N_i^t \cup \{v_i^t\}$ and $u_i^f \in N_i^f \cup \{v_i^f\}$. After repeating k times, we get the feature values $h_{v_i^t}^k$ and $h_{v_i^f}^k$ of vertices v_i^t and v_i^f . The feature vector of all vertices in the graph are $H_{v^t}^K = \{h_{v_1^t}^K, h_{v_2^t}^K, \dots, h_{v_n^t}^K\}$ and $H_{v^f}^K = \{h_{v_1^f}^K, h_{v_2^f}^K, \dots, h_{v_n^f}^K\}$, where the sequence of the vertices is the same as the sequence of the vertices in the input graph.

After the time and frequency domain graphs are aggregated, the feature vectors corresponding to the time and frequency domain graphs are input into the sequential convolution module of the vertex. This operation aims to extract the sequential features of the vertices of the feature vector through sequential convolution. As is shown in Fig. 4 (vertex sequential convolution), We first separately design time and frequency domain initialization learnable parameters $W = \{w_1, w_2, \dots, w_n\}$ to weight each vertex of the feature vector.

$$\begin{cases} o_i^t = w_i^t \cdot h_{v_i^t}^k \\ o_i^f = w_i^f \cdot h_{v_i^f}^k \end{cases} \quad (6)$$

After completing the weighting of the time and frequency domain feature vectors, respectively, we get the weighted

Algorithm 1 Two-Stream Graph-Based Framework

Input: Data: $T = t_1, t_2, \dots, t_n$; TS-Model: {TS-MLP, TS-GNN, TS-1DCNN, TS-SGCN};
Output: Labels: $Y \in [0, 1]$
 Graph Representation Step:
 $A_T = \text{Connection Rule}(\text{Data})$;
 $\text{Data}_F = \text{FFT}(\text{Data})$
 $A_F = \text{Connection Rule}(\text{Data}_F)$;
 Two-stream Graph-based Framework Training Step:
for $i=1$ to epochs **do**
 for batch to shuffle(A_T, A_F) **do**
 Labels = Softmax($\text{Adam}_{lr}(\text{batch}, \text{TS-Model})$);
 end for
end for

feature vector $O = \{o_1, o_2, \dots, o_n\}$. We can feed O into different deep learning models. The module contains forward operations, which are mapped from O to $y^c \in \mathbb{R}^n$.

$$\begin{cases} y_i^t = \sigma \left(\sum_{i=1}^n \omega_i^t \cdot o_i^t + b^t \right) \\ y_i^f = \sigma \left(\sum_{i=1}^n \omega_i^f \cdot o_i^f + b^f \right) \end{cases} \quad (7)$$

We set ω_i^t and $\omega_i^f \in \mathbb{R}^m$ as the weight parameter in the time and frequency domain. The size of m is the kernel size of the deep learning model, b^t and b^f are the bias parameter in the time and frequency domain, and σ is the ReLU activation function [32]. The stride size of the kernel is 1. The kernel size in the time domain and the frequency domain can be set independently. The size of the kernel means the different vertex information of the feature vector of the graph, and the vertex sequential convolution operation is performed P times.

Then we get the feature vector y^t and $y^f \in \mathbb{R}^{[n]}$. We flatten the time and frequency domain data separately, and we fuse the output of the time and frequency domain:

$$\hat{y} = \{y^t, y^f\}. \quad (8)$$

We fuse the output y^t and y^f as \hat{y} , and then flatten the \hat{y} as the input y^{fc} ($1, \text{width} \times \text{height} \times 2 \times \text{channel}$) to the fully-connected layer. The corresponding fully-connected layer output can be expressed as follows:

$$\hat{y}^{fc} = f \left(\sum_{i=1}^{2n} \omega_i^{fc} \cdot y_i^{fc} + b^{fc} \right), \quad (9)$$

where ω_i^{fc} and b^{fc} are the trainable parameters (weight and bias). The superscript fc indicates the fully-connected layer. In this part of the model, the fully-connected layer can extract the associated features of the time and the frequency domain simultaneously. The fully connection layer is performed Q times. Finally, we use Softmax [33] to obtain the classification result $Y \in [0, 1]$, which means seizure or non-seizure.

The algorithm pseudocode is shown in Algorithm 1, and our source code is available at <https://github.com/anonymous2020-source-code/WNG-TS-Model/>.

III. EXPERIMENTS

In this section, we conduct experiments to demonstrate the effectiveness and efficiency of the proposed methods.

A. Dataset

We use three epileptic EEG signal datasets in our experiments.

The Bonn dataset is a classic epilepsy detection dataset collected at the Bonn University [34]. It contains five subsets, called Set A to Set E. Each subset consists of 100 EEG signals. Each signal has 4097 data points and is sampled at 173.61Hz. Among the five subsets, Sets A to D are epileptic-free EEG signals collected under different conditions, of which only subset E is the epileptic EEG signal. We set up the Bonn dataset as four experiments (A vs. E, B vs. E, C vs. E, D vs. E.) for comparison with existing methods. In our experiments, each EEG signal is divided into 256 sampling points with non-overlapping.

The SSW dataset is a special type of epilepsy, ‘‘absence epilepsy’’, which is a dataset without any convulsive symptoms in clinical practice. It needs to be detected by identifying spike and slow waves in the EEG signal. It is a dataset annotated by neurologists from Xinhua Hospital affiliated to Shanghai Jiaotong University. The sampling frequency is 200Hz. The original signal is divided into segments. Each segment contains 200 sampling points, corresponding to 1 second of recording. We choose this duration according to domain knowledge, a full SSW usually lasts less than 1 second [35]. Neurologists manually screened and marked each segment, ensuring that all segments marked positive contained at least one intact SSW. The dataset includes 10473 positive samples and 10473 negative samples.

The CHB-MIT dataset is the EEG signals of epilepsy patients collected at the Children’s Hospital Boston [36], from 23 subjects, with a total of 24 epilepsy records. The sampling frequency of EEG signals is 256Hz, and each data usually contains 23 channels of signals. The method studied in this paper only uses single-channel EEG signals. We only use data from the first channel. In the CHB-MIT dataset, we collected the epileptic EEG signal of the first channel of each subject, and selected a considerable amount of non-epileptic data from it, and segmented the EEG signal into 256 sampling points in length. Each segment corresponds to a 1-second recording. This duration was chosen to use the same data preprocessing scheme as the comparison method. In this paper, the dataset includes 9255 positive samples and 12200 negative samples. More details about the dataset can be found in [37].

For all datasets, we use the average of five-fold cross-validation as the classification result. In addition, for WRG experiments, we fix r_i to ensure that the WRG in each experiment adopts the same random reconnection.

B. Setting

In the experiment, we use the same indicators as other methods for evaluation, namely Accuracy (Acc), Specificity (Spe), and Sensitivity (Sen).

As parameter settings, in this paper, we use cross-entropy loss function and Adam optimizer, epoch set to 100, and

TABLE I
THE STRUCTURE OF DIFFERENT TWO-STREAM DEEP LEARNING MODELS

| | TS-MLP | | TS-GNN | | | TS-1DCNN | | TS-SGCN | |
|-------------|-----------------------|-----------------------|-----------------------|-----------------------|-------------|------------------------|------------------------|------------------------|------------------------|
| | Time | Frequency | Time | Frequency | | Time | Frequency | Time | Frequency |
| Input | $1 \times n \times n$ | $1 \times n \times n$ | $1 \times n \times n$ | $1 \times n \times n$ | Input | $1 \times n \times n$ | $1 \times n \times n$ | $1 \times n \times n$ | $1 \times n \times n$ |
| Aggregation | $1 \times 1 \times n$ | $1 \times 1 \times n$ | $1 \times 1 \times n$ | $1 \times 1 \times n$ | Aggregation | $1 \times 1 \times n$ | $1 \times 1 \times n$ | $1 \times 1 \times n$ | $1 \times 1 \times n$ |
| layer 1 | $1 \times n$ | $1 \times n$ | $1 \times n$ | $1 \times n$ | layer 1 | $8 \times 1 \times 3$ | $8 \times 1 \times 3$ | $8 \times 1 \times 3$ | $8 \times 1 \times 3$ |
| layer 2 | 1×256 | 1×256 | 1×256 | 1×256 | layer 2 | $8 \times 1 \times 3$ | $8 \times 1 \times 3$ | $8 \times 1 \times 3$ | $8 \times 1 \times 3$ |
| layer 3 | 1×128 | 1×128 | 1×128 | 1×128 | layer 3 | $16 \times 1 \times 9$ | $16 \times 1 \times 9$ | $16 \times 1 \times 9$ | $16 \times 1 \times 9$ |
| layer 4 | 1×64 | 1×64 | 1×64 | 1×64 | layer 4 | $32 \times 1 \times 9$ | $32 \times 1 \times 9$ | $32 \times 1 \times 9$ | $32 \times 1 \times 9$ |
| layer 5 | 1×64 | 1×64 | 1×64 | 1×64 | layer 5 | 1×512 | 1×512 | 1×512 | 1×512 |
| layer 6 | 1×64 | 1×64 | 1×64 | 1×64 | layer 6 | 1×128 | 1×128 | 1×128 | 1×128 |

learning trip to 0.005. The code is implemented using the PyTorch platform. All experiments are performed on Intel i7-6850K3 3.6GHz CPU and 4 NVIDIA Titan XP GPUs.

We verify the scheme proposed in this paper in two parts:

The first part is the performance evaluation of the graph of the data. We will evaluate the performance of the time complexity and memory requirement of the graph method of EEG signals. In the performance evaluation section, we will compare the performance of VG [19], HVG [22], LPVG [21], LPHVG [23], OG, and WOG [24] on single-channel EEG signals.

The second part is the classification performance evaluation. We will verify different two-stream models (TS-MLP, TS-GNN, TS-1DCNN, and TS-SGCN). We will also compare the classification performance with the current best baseline method.

TS-MLP: The TS-MLP model has no sequential convolution structure and only performs vertex aggregation once. It is verified whether the sequential convolution operation and the vertex aggregation produce gains for epilepsy detection tasks.

TS-GNN: The TS-GNN model has no sequential convolution structure. It is verified whether the sequential convolution operation produces gains for epilepsy detection tasks.

TS-SGCN: The TS-SGCN includes vertex aggregation and sequential convolution operation. A two-stream sequential graph convolutional network model based on EEG signal conversion in the time and frequency domain.

TS-1DCNN: The TS-1DCNN only performs vertex aggregation once. To replace the vertex sequential convolution operation in TS-SGCN, we directly perform one-dimensional convolution on the adjacent matrix of the time and frequency domain graph structure.

In order to make the results of the classifiers comparable, the parameters of the same module of each model are consistent with the parameter settings of the corresponding module of TS-SGCN. For the deep learning model, the number of layers is the same. For example, each model consists of 1 aggregation layer and 6 layers in Table I. The number of aggregations K is set to 2 in TS-GNN and TS-SGCN (set to 1 in TS-MLP and TS-1DCNN). Note that all models have performed at least graph vertex aggregation once. Therefore, all the above models are graph neural networks.

IV. RESULTS

A. Analysis of the p Setting

Before starting the experiment, we need to determine the value of the random probability parameter p in the WRG. We select the range of p from 0-50% for experiments.

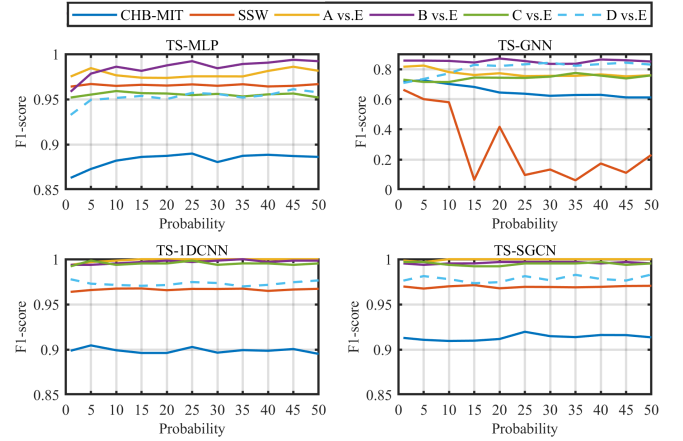


Fig. 5. Accuracy does not correlate with increasing p .

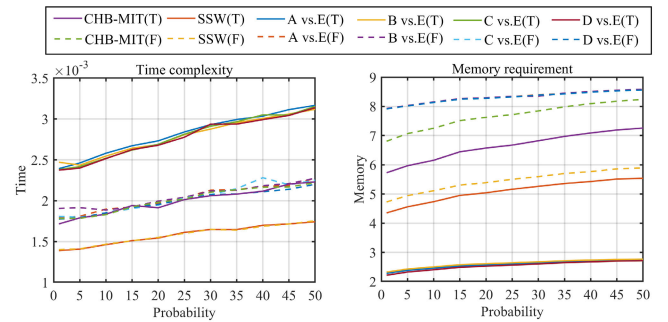


Fig. 6. Time complexity and memory requirement increases with the increase of p .

It should be pointed out that when p is 0, it is a special case of WNG.

We have carried out classification experiments under different p values. From the experimental results, we can find that the classification accuracy of WRG is not affected as the p value increases, as shown in Fig. 5. Even some of the models have a decline in their classification results. For example, in the TS-GNN experiment, with the increase of p , the classification performance shows a decreasing trend.

We further conduct experiments on the efficiency of WRG when p goes to different values. As shown in Fig. 6, we can clearly observe that the curve of time complexity and memory requirement increases continuously with the increase of p .

Increasing p cannot improve classification performance, but it will cause a drop in efficiency. Therefore, in the classification task, we choose a smaller p value to reduce

TABLE II
TIME COMPLEXITY OF DIFFERENT GRAPH REPRESENTATION METHODS (S OF SINGLE GRAPH)

| Methods | CHB-MIT | | SSW | | Bonn | | | | | | | |
|---------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| | | | | | A vs.E | | B vs.E | | C vs.E | | D vs.E | |
| | T | F | T | F | T | F | T | F | T | F | T | F |
| VG | 0.19092 | 0.18706 | 0.11566 | 0.11791 | 0.33859 | 0.19137 | 0.30297 | 0.18688 | 0.31200 | 0.18569 | 0.30216 | 0.19269 |
| LPVG | 0.19287 | 0.19314 | 0.11849 | 0.11821 | 0.33876 | 0.19253 | 0.30834 | 0.19125 | 0.31082 | 0.19230 | 0.30108 | 0.19352 |
| HVG | 0.04285 | 0.03737 | 0.02515 | 0.02194 | 0.18331 | 0.03806 | 0.14820 | 0.03658 | 0.15997 | 0.03596 | 0.15410 | 0.03721 |
| LPHVG | 0.04375 | 0.03863 | 0.02661 | 0.02345 | 0.18498 | 0.03856 | 0.14939 | 0.03739 | 0.16126 | 0.03720 | 0.15535 | 0.03745 |
| OG | 0.13096 | 0.13123 | 0.07787 | 0.08023 | 0.13944 | 0.13340 | 0.12849 | 0.12864 | 0.12740 | 0.12927 | 0.12739 | 0.12844 |
| WOG | 0.25520 | 0.25173 | 0.15544 | 0.15288 | 0.56197 | 0.27258 | 0.48359 | 0.25146 | 0.50171 | 0.24772 | 0.49007 | 0.25018 |
| WRG | 0.00495 | 0.00438 | 0.00412 | 0.00409 | 0.00865 | 0.00641 | 0.00541 | 0.00416 | 0.00534 | 0.00414 | 0.00530 | 0.00413 |
| WNG | 0.00202 | 0.00200 | 0.00150 | 0.00152 | 0.00290 | 0.00200 | 0.00301 | 0.00205 | 0.00296 | 0.00203 | 0.00297 | 0.00208 |

TABLE III
MEMORY REQUIREMENT OF DIFFERENT GRAPH REPRESENTATION METHODS (KB OF SINGLE GRAPH)

| Methods | CHB-MIT | | SSW | | Bonn | | | | | | | |
|---------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| | | | | | A vs.E | | B vs.E | | C vs.E | | D vs.E | |
| | T | F | T | F | T | F | T | F | T | F | T | F |
| VG | 6.347 | 8.003 | 4.287 | 5.856 | 6.008 | 8.467 | 5.980 | 8.482 | 6.226 | 8.822 | 6.273 | 8.858 |
| LPVG | 7.557 | 12.105 | 5.302 | 8.743 | 6.813 | 12.678 | 6.698 | 12.488 | 6.958 | 12.960 | 6.949 | 12.979 |
| HVG | 4.557 | 4.685 | 3.043 | 3.169 | 4.616 | 4.679 | 4.614 | 4.674 | 4.553 | 4.669 | 4.539 | 4.664 |
| LPHVG | 5.721 | 6.458 | 3.900 | 4.434 | 5.717 | 6.392 | 5.708 | 6.361 | 5.657 | 6.396 | 6.949 | 12.979 |
| OG | 10.963 | 13.902 | 6.432 | 8.615 | 12.947 | 12.650 | 13.332 | 12.384 | 11.516 | 12.868 | 11.474 | 12.732 |
| WOG | 347.933 | 454.760 | 229.011 | 253.235 | 270.576 | 445.650 | 276.180 | 445.950 | 270.098 | 445.280 | 270.972 | 445.323 |
| WRG | 6.155 | 7.253 | 4.736 | 5.112 | 4.807 | 7.243 | 4.842 | 7.245 | 4.774 | 7.239 | 4.775 | 7.240 |
| WNG | 5.662 | 6.736 | 4.306 | 4.679 | 4.474 | 6.745 | 4.506 | 6.750 | 4.431 | 6.739 | 4.430 | 6.740 |

the consumption of computing resources. In the subsequent experiments, we choose two smaller values of p , 0 (WNG) and 10% (WRG).

B. The Performance of the Graph

The performance of graph representation consists of time complexity, memory requirements, and classification performance.

The time complexity of the EEG signal graph representation is the time to transform the EEG signal into a graph. The memory requirement of the EEG signal graph representation is the size of the storage space occupied by the EEG signal graph. In addition, we verify the classification performance. Our purpose for testing classification performance is to evaluate the optimal combination of graph representation and classification model. The experiments are carried out on the Bonn, SSW, and CHB-MIT datasets.

1. The time complexity of the graph representations.

We compare the graph representation method proposed in this paper with other existing methods. As shown in Table II, we only use one indicator (Seconds, S). The experiments are all carried out in the same experimental environment. Each graph representation method performs graph representation on all samples in the dataset, then averages the total time to calculate the generation time of each method for a single sample. The method proposed in this paper has the shortest generation time among all graph representations. WOG is a graph with the highest classification accuracy for existing epileptic EEG signals. Compared with WOG, WRG is 56 times faster than WOG on the CHB-MIT dataset (WNG is 125 times), WRG is 37 times faster than WOG on the SSW dataset (WNG is 100 times), and WRG is at least 42 times faster than WOG on the Bonn dataset (WNG is at least 122 times). The time complexity of both WRG and WNG methods is significantly lower than other graph representation methods. This shows that our method only needs less time to convert the EEG signal into the graph representation.

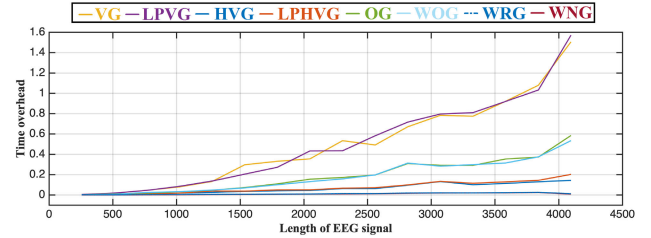


Fig. 7. The generation time (S) of graph representation in the different EEG signal lengths.

To further verify the graph representation time efficiency, we randomly selected from the epileptic EEG signal dataset containing 256 sampling points as the unit of EEG signal segments, and gradually increased the length of EEG signals, and observed the graph representation methods of different EEG signals in these EEG signals. A total of 16 EEG signal lengths were set in the experiment, and EEG signal segments ranged from 256 sampling points to 4096 sampling points. Each graph representation method calculates 10 generation times on each length of EEG signal segment, and takes the average of these 10 generation times as the final result.

It can be seen from Fig. 7 that the generation time of VG and LPVG is the largest, which is $O(n^3)$. The time complexity of OG and WOG methods is $O(n^2)$. The time complexity of HVG and LPHVG methods is $O(n \log n)$. The least time-consuming methods are WRG and WNG methods, and the time complexity is $O(n)$. This also intuitively verifies the fact that the graph representation method proposed in this paper has the least time complexity.

2. The memory requirements of the graph representations.

We compare the graph representation method proposed in this paper with other existing methods. As shown in Table III, we only use one unit (Kilobyte, KB). Each graph representation method is the total memory requirement obtained by performing graph representation on all dataset samples. The memory requirements of OG and WOG are large, which is

TABLE V

COMPARE WITH ONLY TIME OR FREQUENCY DOMAIN METHODS

| Dataset | Method | MR | Acc | Spe | Sen | F1 |
|---------|--------------|--------|--------------|--------------|--------------|--------------|
| CHB-MIT | WRG-T-1DCNN | 6.155 | 0.925 | 0.963 | 0.876 | 0.910 |
| | WRG-F-1DCNN | 7.253 | 0.927 | 0.961 | 0.878 | 0.911 |
| | WRG-TS-1DCNN | 13.408 | 0.931 | 0.942 | 0.918 | 0.895 |
| | WNG-T-1DCNN | 5.662 | 0.903 | 0.963 | 0.826 | 0.867 |
| | WNG-F-1DCNN | 6.736 | 0.922 | 0.968 | 0.862 | 0.892 |
| | WNG-TS-1DCNN | 12.398 | 0.936 | 0.966 | 0.897 | 0.909 |
| SSW | WRG-T-1DCNN | 4.736 | 0.943 | 0.903 | 0.984 | 0.946 |
| | WRG-F-1DCNN | 4.679 | 0.967 | 0.946 | 0.987 | 0.967 |
| | WRG-TS-1DCNN | 9.415 | 0.953 | 0.961 | 0.945 | 0.935 |
| | WNG-T-1DCNN | 4.306 | 0.961 | 0.930 | 0.990 | 0.930 |
| | WNG-F-1DCNN | 4.679 | 0.953 | 0.925 | 0.983 | 0.917 |
| | WNG-TS-1DCNN | 8.985 | 0.958 | 0.971 | 0.944 | 0.944 |
| A vs.E | WRG-T-1DCNN | 4.807 | 1.000 | 1.000 | 1.000 | 1.000 |
| | WRG-F-1DCNN | 7.243 | 1.000 | 1.000 | 1.000 | 1.000 |
| | WRG-TS-1DCNN | 12.500 | 0.999 | 0.999 | 1.000 | 0.999 |
| | WNG-T-1DCNN | 4.474 | 0.997 | 1.000 | 0.993 | 0.997 |
| | WNG-F-1DCNN | 6.745 | 0.994 | 1.000 | 0.988 | 0.994 |
| | WNG-TS-1DCNN | 11.219 | 0.999 | 0.998 | 1.000 | 0.997 |
| B vs.E | WRG-T-1DCNN | 4.842 | 0.995 | 0.994 | 0.997 | 0.995 |
| | WRG-F-1DCNN | 7.245 | 0.997 | 0.997 | 0.997 | 0.997 |
| | WRG-TS-1DCNN | 12.087 | 0.998 | 0.999 | 0.996 | 0.997 |
| | WNG-T-1DCNN | 4.506 | 0.995 | 1.000 | 0.990 | 0.995 |
| | WNG-F-1DCNN | 6.750 | 0.990 | 0.997 | 0.984 | 0.989 |
| | WNG-TS-1DCNN | 11.256 | 0.999 | 0.999 | 0.999 | 0.999 |
| C vs.E | WRG-T-1DCNN | 4.774 | 0.991 | 1.000 | 0.981 | 0.991 |
| | WRG-F-1DCNN | 7.239 | 0.991 | 0.997 | 0.984 | 0.991 |
| | WRG-TS-1DCNN | 12.013 | 0.993 | 0.999 | 0.987 | 0.992 |
| | WNG-T-1DCNN | 4.431 | 0.997 | 1.000 | 0.994 | 0.997 |
| | WNG-F-1DCNN | 6.739 | 0.981 | 1.000 | 0.963 | 0.981 |
| | WNG-TS-1DCNN | 11.170 | 0.998 | 1.000 | 0.996 | 0.998 |
| D vs.E | WRG-T-1DCNN | 4.775 | 0.970 | 0.972 | 0.969 | 0.970 |
| | WRG-F-1DCNN | 7.240 | 0.969 | 0.981 | 0.956 | 0.968 |
| | WRG-TS-1DCNN | 12.015 | 0.986 | 0.989 | 0.982 | 0.981 |
| | WNG-T-1DCNN | 4.430 | 0.992 | 0.997 | 0.988 | 0.991 |
| | WNG-F-1DCNN | 6.740 | 0.967 | 0.981 | 0.953 | 0.957 |
| | WNG-TS-1DCNN | 11.170 | 0.995 | 0.997 | 0.994 | 0.994 |

In the Bonn dataset, the classification accuracy of WNG-TS-1DCNN is better than that of LSTM [2] and AMWCNN [11], which further shows that the signal graph representation method can promote the classification of deep learning models. The experimental results show that our method performs better in classification than methods using only deep learning model. Compared with SeizNet [13] in the SSW dataset, and compared with TF-CNN [39] in CHB-MIT, the classification accuracy of WNG-TS-1DCNN is also better than that of using only the deep learning model.

Compared with VG-F-GIN [25], WNG-TS-1DCNN achieves better classification results, which indicates that the proposed method is more effective than graph neural networks using only graph aggregation. It should be pointed out that GIN is only suitable for undirected graph structures, we only use VG for the graph representation of EEG signal.

Compared with VG-SGCN [25], the classification accuracy of WNG-TS-1DCNN is significantly improved, especially in the SSW and CHB-MIT datasets experiments. The results further verified that WNG-TS-1DCNN brought gains for epilepsy classification. The memory requirement of WNG is bigger than that of VG, but WNG is the sum of the memory requirements represented in both time and frequency domains, while VG is only the memory requirement of the frequency domain.

Compared with WOG-2DCNN [24], WNG-TS-1DCNN has comparable or even better classification accuracy. However, it is worth pointing out that WNG is more lightweight than WOG, and the number of parameters of TS-1DCNN is much smaller than that of 2DCNN. With similar classification

TABLE VI

COMPARE WITH EXISTING METHODS

| Dataset | Method | TC | MR | Acc | Spe | Sen |
|-------------|------------------|----------|---------|--------------|--------------|--------------|
| CHB-MIT | NCOV [38] | - | - | 0.918 | 1.000 | 0.836 |
| | TF-CNN [39] | - | - | 0.890 | - | - |
| | VG-SGCN [25] | $O(n^3)$ | 8.003 | 0.850 | 0.834 | 0.871 |
| | WOG-2DCNN [24] | $O(n^2)$ | 347.933 | 0.911 | 0.964 | 0.843 |
| | WNG-TS-1DCNN | $O(n)$ | 12.398 | 0.936 | 0.966 | 0.897 |
| | AdaBoost [40] | - | - | 0.587 | 0.942 | 0.548 |
| SSW | SeizNet [13] | - | - | 0.886 | 1.000 | 0.772 |
| | WNFG-SSGCN [41] | $O(n)$ | - | 0.951 | 0.984 | 0.918 |
| | VG-F-GIN [25] | $O(n^3)$ | 5.856 | 0.824 | 0.864 | 0.783 |
| | VG-SGCN [25] | $O(n^3)$ | 5.856 | 0.893 | 0.911 | 0.874 |
| | WOG-2DCNN [24] | $O(n^2)$ | 229.011 | 0.965 | 0.974 | 0.960 |
| | WNG-TS-1DCNN | $O(n)$ | 8.985 | 0.958 | 0.971 | 0.944 |
| A vs.E | CT-LS-SVM [42] | - | - | 1.000 | 1.000 | 1.000 |
| | LSTM [2] | - | - | 0.970 | 0.980 | 0.960 |
| | AMWCNN [11] | - | - | 1.000 | 1.000 | 1.000 |
| | DCT-FHE-SVM [43] | - | - | 0.949 | 0.943 | 0.954 |
| | WNFG-SSGCN [41] | $O(n)$ | - | 0.998 | 0.997 | 1.000 |
| | VG-SVMKNN [9] | $O(n^3)$ | 6.008 | 1.000 | 1.000 | 1.000 |
| B vs.E | VG-F-GIN [25] | $O(n^3)$ | 8.467 | 0.968 | 0.983 | 0.953 |
| | VG-SGCN [25] | $O(n^3)$ | 8.467 | 0.998 | 0.995 | 1.000 |
| | WOG-2DCNN [24] | $O(n^2)$ | 270.576 | 0.999 | 0.988 | 1.000 |
| | WNG-TS-1DCNN | $O(n)$ | 11.219 | 0.999 | 0.998 | 1.000 |
| | CT-LS-SVM [42] | - | - | 0.995 | 0.992 | 0.998 |
| | LSTM [2] | - | - | 0.925 | 0.940 | 0.910 |
| C vs.E | AMWCNN [11] | - | - | 1.000 | 1.000 | 1.000 |
| | DCT-FHE-SVM [43] | - | - | 0.990 | 0.987 | 0.993 |
| | WNFG-SSGCN [41] | $O(n)$ | - | 1.000 | 1.000 | 1.000 |
| | VG-SVMKNN [9] | $O(n^3)$ | 5.980 | 0.973 | 0.952 | 0.995 |
| | VG-F-GIN [25] | $O(n^3)$ | 8.482 | 0.925 | 0.923 | 0.928 |
| | VG-SGCN [25] | $O(n^3)$ | 8.482 | 0.995 | 0.995 | 0.995 |
| D vs.E | WOG-2DCNN [24] | $O(n^2)$ | 276.180 | 0.991 | 0.992 | 0.990 |
| | WNG-TS-1DCNN | $O(n)$ | 11.256 | 0.999 | 0.999 | 0.999 |
| | CT-LS-SVM [42] | - | - | 0.964 | 0.948 | 0.980 |
| | LSTM [2] | - | - | 0.920 | 0.950 | 0.890 |
| | AMWCNN [11] | - | - | 0.994 | 0.984 | 0.992 |
| | DCT-FHE-SVM [43] | - | - | 0.975 | 0.970 | 0.980 |
| C vs.E | WNFG-SSGCN [41] | $O(n)$ | - | 0.988 | 0.984 | 0.991 |
| | VG-SVMKNN [9] | $O(n^3)$ | 6.226 | 0.983 | 0.980 | 0.985 |
| | VG-F-GIN [25] | $O(n^3)$ | 8.822 | 0.946 | 0.913 | 0.980 |
| | VG-SGCN [25] | $O(n^3)$ | 8.822 | 0.985 | 0.983 | 0.988 |
| | WOG-2DCNN [24] | $O(n^2)$ | 270.098 | 0.988 | 0.990 | 0.985 |
| | WNG-TS-1DCNN | $O(n)$ | 11.170 | 0.998 | 1.000 | 0.996 |
| D vs.E | CT-LS-SVM [42] | - | - | 0.940 | 0.820 | 0.940 |
| | LSTM [2] | - | - | 0.910 | 0.870 | 0.950 |
| | AMWCNN [11] | - | - | 0.991 | 0.994 | 0.989 |
| | DCT-FHE-SVM [43] | - | - | 0.964 | 0.962 | 0.965 |
| | WNFG-SSGCN [41] | $O(n)$ | - | 0.977 | 0.972 | 0.981 |
| | VG-SVMKNN [9] | $O(n^3)$ | 6.273 | 0.933 | 0.963 | 0.906 |
| B vs.D vs.E | VG-F-GIN [25] | $O(n^3)$ | 8.858 | 0.934 | 0.930 | 0.934 |
| | VG-SGCN [25] | $O(n^3)$ | 8.858 | 0.964 | 0.960 | 0.968 |
| | WOG-2DCNN [24] | $O(n^2)$ | 270.972 | 0.974 | 0.976 | 0.971 |
| | WNG-TS-1DCNN | $O(n)$ | 11.170 | 0.995 | 0.997 | 0.994 |
| | WOG-SVM [24] | $O(n^2)$ | 268.998 | 0.965 | 0.953 | 0.980 |
| | CNN [10] | - | - | 0.887 | 0.900 | 0.950 |
| B vs.D vs.E | WOG-2DCNN [24] | $O(n^2)$ | 268.998 | 0.974 | 0.958 | 0.975 |
| | WNG-TS-1DCNN | $O(n)$ | 4.432 | 0.968 | 0.981 | 0.972 |

accuracy, our framework time complexity and memory requirements are significantly lower than WOG-2DCNN.

Compared with TF-CNN [39], WNG-TS-1DCNN has higher classification accuracy. It is worth noting here that the method proposed in this paper only uses the information of a single-channel of CHB-MIT, while the TF-CNN method uses 23 channels of data. This further shows that the framework in this paper can still maintain a high classification accuracy in the case of limited information.

In summary, WNG-TS-1DCNN is a classification graph-based framework for epileptic EEG signals that combines accuracy, low time complexity, and low memory requirements.

V. DISCUSSION

A. Critical Difference

In order to further illustrate the performance of our method proposed, we analyze the performance of the EEG signal graph

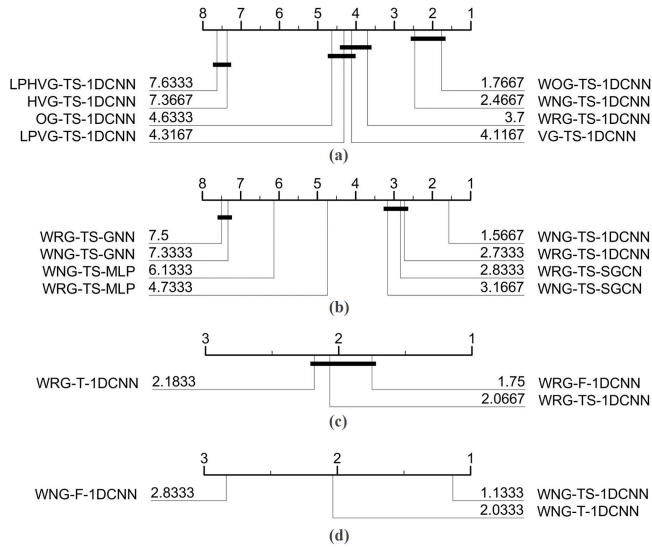


Fig. 8. Critical difference diagram (CDD) of all methods on all datasets.

representation and the two-stream graph-based framework. We use the critical difference diagram (CDD) [44] ($\alpha = 0.05$) to compare the accuracies across all datasets. We use their average accuracies over five runs.

We first evaluated the performance of different graph representations. As shown in Fig. 8 (a), we find that WOG and WNG are superior to all baseline methods with TS-1DCNN. However, WOG contains a large number of redundant edges, so WNG is comprehensively considered to be the graph representation method of EEG signals taking into account classification accuracy, time complexity, and memory consumption.

We evaluate the performance of different two-stream graph-based frameworks. As shown in Fig. 8 (b), we find that WRG and WNG have the best classification performance with TS-1DCNN. This also verifies that the two-stream graph-based framework with only one aggregation and sequential convolution is superior to other ablation baseline methods in the graph representation of EEG signals.

We also evaluated the classification performance of WRG and WNG in time domain, frequency domain, and simultaneous extraction time and frequency domain experiments. As shown in Fig. 8 (c), we find that there is little difference between the time domain, frequency domain, and the method of extracting time and frequency domain simultaneously in WRG. The best performance is the frequency domain WRG, which shows that the random edges do not improve the performance of the graph representation of EEG signals. As shown in Fig. 8 (d), in WNG, simultaneous extraction of time and frequency domain features is significantly better than the other two methods, which also verifies the effectiveness of simultaneous extraction of time and frequency domain features.

B. Interpretability

We look into the interpretability of our method. As was previously explained, our interpretability comes from the vertex aggregation and learnable weight vector. Concretely,

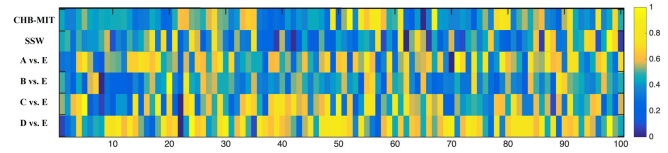


Fig. 9. Learnable weight vectors of the two-stream graph-based framework on all datasets, which can be interpreted as the importance of each frequency to epilepsy detection.

each weight corresponds to a specific frequency and can be seen as the importance of that frequency to epilepsy detection.

On the CHB-MIT dataset and the SSW dataset, we found that the weight distribution is on the frequency bands 30Hz and 60Hz (gamma rhythm), which is consistent with the conclusions of the paper [45]. On the Bonn dataset, we found that the weight distribution is between 3-4Hz (A vs. E, C vs. E), 8-12Hz (D vs. E), and 30-100Hz (A vs. E, B vs. E, C vs. E, D vs. E), which is consistent with the conclusions of the paper [15], [45]. These conclusions further provide explanatory evidence for the classification of epilepsy using EEG signal graph representations and graph-based framework.

C. Robustness of Graph Representation

In order to illustrate the robustness of our graph representation, We analyze the training curve of each EEG signal graph representation with different two-stream graph-based frameworks. The experiment contains a total of five categories. The tasks are CHB-MIT, SSW, A vs. E, B vs. E, C vs. E, and D vs. E.

It can be seen from Fig. 10:

1. In the experiments of TS-MLP, TS-MLP has the weighted feature extraction ability. Therefore, WOG and WRG have the highest classification accuracy among all graphs. This shows that WOG and WRG have better information representation abilities.

2. In the experiments of TS-GNN, the stability of the TS-GNN curve is poor, and the classification accuracy is the lowest among the four models, especially in the SSW and CHB-MIT datasets. This shows that the TS-GNN model without sequential feature extraction capability has very limited classification capability for graph representations of EEG signals.

3. In the experiment of TS-SGCN, the curve is stable, and the classification accuracy converges higher, which further verifies the importance of sequential feature extraction.

4. In the experiment of TS-1DCNN, the curve is stable, and the classification accuracy converges the highest. After the aggregation operation, compared with TS-SGCN, we find that the graph has a further improvement, and TS-1DCNN only uses vertex aggregation once.

Based on the above research, it can be known that the graph representation method of EEG signals proposed in this paper is the method with the least time complexity and memory requirements at present, which greatly accelerates the generation speed of graphs. It can be found from the training curve experiment of the deep learning model that this graph

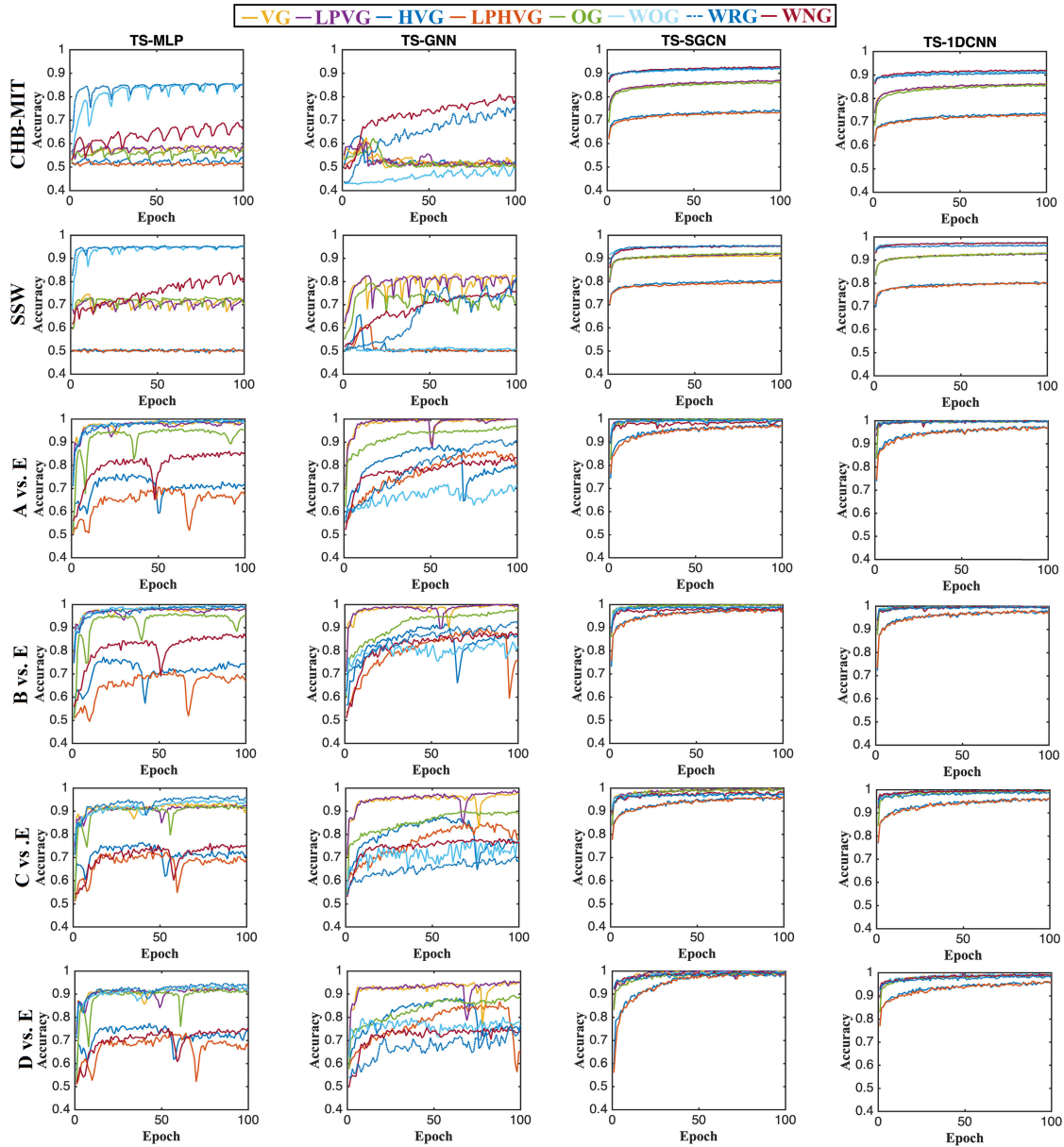


Fig. 10. Training curves of the two-stream graph-based framework in three datasets. The graph representations used here are VG, LPVG, HVG, LPHVG, OG, WOG, WRG, and WNG.

representation method can improve the classification accuracy of the model. In addition, the TS-1DCNN has shown stable performance in the classification task of the epileptic EEG signals graph representation. In summary, it can be known that the graph representation method proposed in this paper and the designed two-stream graph-based framework have important significance for the classification of the graph representation of epileptic EEG signals.

It should be noted that the method proposed in this paper is mainly used to optimize the epilepsy classification task of single-channel EEG signals. In future work, we will design a graph representation method based on multi-channel EEG signals to complete the epilepsy classification task.

VI. CONCLUSION

In this paper, we propose a new graph representation method to represent EEG signals and develop two-stream graph-based

frameworks to extract the time and frequency domain features of graphs. The graph representation method we proposed dramatically reduces the time complexity and memory requirements while retaining the information of the EEG signal. Our two-stream graph-based framework can simultaneously extract features in the time and frequency domain and extract the sequence features of the vertices in the graph. We apply our method to the classification task of epileptic EEG signal, and the results show that the performance of the method proposed in this paper is better than existing methods. We compare our method with the state-of-the-art method. It turns out that our graph representation method has the lowest time complexity and the minor memory requirement. The classification results of the two-stream graph-based framework surpass other methods in many experiments. These results prove that our method is competitive for the epileptic EEG signal classification task.

REFERENCES

- [1] T. A. Glauser et al., "Ethosuximide, valproic acid, and lamotrigine in childhood absence epilepsy," *New England J. Med.*, vol. 362, no. 9, pp. 790–799, Mar. 2010.
- [2] D. Ahmedtaristzabal, C. Fookes, K. Nguyen, and S. Sridharan, "Deep classification of epileptic signals," in *Proc. 40th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2018, pp. 332–335.
- [3] H. Zhao, Q. Zheng, K. Ma, H. Li, and Y. Zheng, "Deep representation-based domain adaptation for nonstationary EEG classification," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 2, pp. 535–545, Feb. 2021.
- [4] Z. Song, B. Deng, J. Wang, G. Yi, and W. Yue, "Epileptic seizure detection using brain-rhythmic recurrence biomarkers and ONASNet-based transfer learning," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 30, pp. 979–989, 2022.
- [5] Y. Zhao, C. Li, X. Liu, R. Qian, R. Song, and X. Chen, "Patient-specific seizure prediction via adder network and supervised contrastive learning," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 30, pp. 1536–1547, 2022.
- [6] L. S. Vidyaratne and K. M. Iftikharuddin, "Real-time epileptic seizure detection using EEG," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 11, pp. 2146–2156, Nov. 2017.
- [7] U. R. Acharya, H. Fujita, V. K. Sudarshan, S. Bhat, and J. E. W. Koh, "Application of entropies for automated diagnosis of epilepsy using EEG signals: A review," *Knowl.-Based Syst.*, vol. 88, pp. 85–96, Nov. 2015.
- [8] P. Boonyakitanont, A. Lek-Uthai, K. Chomtho, and J. Songsiri, "A review of feature extraction and performance evaluation in epileptic seizure detection using EEG," *Biomed. Signal Process. Control*, vol. 57, Mar. 2020, Art. no. 101702.
- [9] S. Supriya, S. Siuly, H. Wang, J. Cao, and Y. Zhang, "Weighted visibility graph with complex network features in the detection of epilepsy," *IEEE Access*, vol. 4, pp. 6554–6566, 2016.
- [10] U. R. Acharya, S. L. Oh, Y. Hagiwara, J. H. Tan, and H. Adeli, "Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals," *Comput. Biol. Med.*, vol. 100, pp. 270–278, Sep. 2018.
- [11] Q. Xin, S. Hu, S. Liu, L. Zhao, and Y.-D. Zhang, "An attention-based wavelet convolution neural network for epilepsy EEG classification," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 30, pp. 957–966, 2022.
- [12] Y. Zhang et al., "Epileptic seizure detection based on bidirectional gated recurrent unit network," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 30, pp. 135–145, 2022.
- [13] M. T. Avcu, Z. Zhang, and D. W. S. Chan, "Seizure detection using least EEG channels by deep convolutional neural network," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2019, pp. 1120–1124.
- [14] S. Mao and E. Sejdic, "A review of recurrent neural network-based methods in computational physiology," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Feb. 7, 2022, doi: [10.1109/TNNLS.2022.3145365](https://doi.org/10.1109/TNNLS.2022.3145365).
- [15] K. Benedek, A. Berényi, P. Gombkötő, H. Piilgaard, and M. Lauritzen, "Neocortical gamma oscillations in idiopathic generalized epilepsy," *Epilepsia*, vol. 57, no. 5, pp. 796–804, May 2016.
- [16] Z. Gao et al., "EEG-based spatio-temporal convolutional neural network for driver fatigue evaluation," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 9, pp. 2755–2763, Sep. 2019.
- [17] Y. Zou, R. V. Donner, N. Marwan, J. F. Donges, and J. Kurths, "Complex network approaches to nonlinear time series analysis," *Phys. Rep.*, vol. 787, pp. 1–97, Jan. 2018.
- [18] J. Yu, D. Tao, and M. Wang, "Adaptive hypergraph learning and its application in image classification," *IEEE Trans. Image Process.*, vol. 21, no. 7, pp. 3262–3272, Jul. 2012.
- [19] L. Lacasa, B. Luque, F. Ballesteros, J. Luque, and J. C. Nuño, "From time series to complex networks: The visibility graph," *Proc. Nat. Acad. Sci. USA*, vol. 105, no. 13, pp. 4972–4975, Apr. 2008.
- [20] S. Supriya, S. Siuly, H. Wang, and Y. Zhang, "Epilepsy detection from EEG using complex network techniques: A review," *IEEE Rev. Biomed. Eng.*, vol. 16, pp. 292–306, 2023.
- [21] Z. Ting-Ting, J. Ning-De, G. Zhong-Ke, and L. Yue-Bin, "Limited penetrable visibility graph for establishing complex network from time series," *Acta Phys. Sinica*, vol. 61, no. 3, 2012, Art. no. 030506.
- [22] B. Luque, L. Lacasa, F. Ballesteros, and J. Luque, "Horizontal visibility graphs: Exact results for random time series," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 80, no. 4, Oct. 2009, Art. no. 046103.
- [23] Z.-K. Gao, Q. Cai, Y.-X. Yang, W.-D. Dang, and S.-S. Zhang, "Multiscale limited penetrable horizontal visibility graph for analyzing nonlinear time series," *Sci. Rep.*, vol. 6, no. 1, p. 35622, Oct. 2016.
- [24] J. Wang et al., "A weighted overlook graph representation of EEG data for absence epilepsy detection," in *Proc. IEEE Int. Conf. Data Mining (ICDM)*, Nov. 2020, pp. 581–590.
- [25] J. Wang, S. Liang, D. He, Y. Wang, Y. Wu, and Y. Zhang, "A sequential graph convolutional network with frequency-domain complex network of EEG signals for epilepsy detection," in *Proc. IEEE Int. Conf. Bioinf. Biomed. (BIBM)*, Dec. 2020, pp. 785–792.
- [26] K. Simonyan and A. Zisserman, "Two-stream convolutional networks for action recognition in videos," 2014, *arXiv:1406.2199*.
- [27] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, vol. 393, nos. 66–84, pp. 440–442, 1998.
- [28] X. Yan, D. Yang, Z. Lin, and B. Vucetic, "Significant low-dimensional spectral-temporal features for seizure detection," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 30, pp. 668–677, 2022.
- [29] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," 2016, *arXiv:1609.02907*.
- [30] K. Xu, W. Hu, J. Leskovec, and S. Jegelka, "How powerful are graph neural networks?" 2018, *arXiv:1810.00826*.
- [31] F. Errica, M. Podda, D. Bacciu, and A. Micheli, "A fair comparison of graph neural networks for graph classification," 2019, *arXiv:1912.09893*.
- [32] J. Yu, M. Tan, H. Zhang, Y. Rui, and D. Tao, "Hierarchical deep click feature prediction for fine-grained image recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 2, pp. 563–578, Feb. 2022.
- [33] E. Jang, S. Gu, and B. Poole, "Categorical reparameterization with Gumbel-Softmax," 2016, *arXiv:1611.01144*.
- [34] R. G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, and C. E. Elger, "Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 64, no. 6, Nov. 2001, Art. no. 061907.
- [35] C.-C. Chiang et al., "Slow moving neural source in the epileptic hippocampus can mimic progression of human seizures," *Sci. Rep.*, vol. 8, no. 1, p. 1564, Jan. 2018.
- [36] A. L. Goldberger et al., "PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. 215–220, Jun. 2000.
- [37] A. H. Shoeb, "Application of machine learning to epileptic seizure onset detection and treatment," Ph.D. dissertation, Massachusetts Inst. Technol., Cambridge, MA, USA, 2009.
- [38] Y. U. Khan, N. Rafiuddin, and O. Farooq, "Automated seizure detection in scalp EEG using multiple wavelet scales," in *Proc. IEEE Int. Conf. Signal Process., Comput. Control*, Mar. 2012, pp. 1–5.
- [39] M. Taherisadr, M. Joneidi, and N. Rahnavard, "EEG signal dimensionality reduction and classification using tensor decomposition and deep convolutional neural networks," in *Proc. IEEE 29th Int. Workshop Mach. Learn. Signal Process. (MLSP)*, Oct. 2019, pp. 1–6.
- [40] Y.-C. Liu, C.-C. Lin, J.-J. Tsai, and Y.-N. Sun, "Model-based spike detection of epileptic EEG data," *Sensors*, vol. 13, no. 9, pp. 12536–12547, Sep. 2013.
- [41] J. Wang, R. Gao, H. Zheng, H. Zhu, and C.-J. R. Shi, "SSGCNet: A sparse spectra graph convolutional network for epileptic EEG signal classification," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Mar. 16, 2023, doi: [10.1109/TNNLS.2023.3252569](https://doi.org/10.1109/TNNLS.2023.3252569).
- [42] Y. Li and P. P. Wen, "Clustering technique-based least square support vector machine for EEG signal classification," *Comput. Methods Programs Biomed.*, vol. 104, no. 3, pp. 358–372, Dec. 2011.
- [43] A. Gupta, P. Singh, and M. Karlekar, "A novel signal modeling approach for classification of seizure and seizure-free EEG signals," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 5, pp. 925–935, May 2018.
- [44] J. Demšar, "Statistical comparisons of classifiers over multiple data sets," *J. Mach. Learn. Res.*, vol. 7, pp. 1–30, Dec. 2006.
- [45] J. O. Willoughby, "Persistent abnormality detected in the non-ictal electroencephalogram in primary generalised epilepsy," *J. Neurol., Neurosurgery Psychiatry*, vol. 74, no. 1, pp. 51–55, Jan. 2003.