

# Empowering Cross-Patient Epilepsy Diagnosis from Diverse-Sampling Low-Quality EEG Signals

Yanshen Sun

Department of Computer Science  
Virginia Tech  
Falls Church, VA, USA  
yansh93@vt.edu

Jianger Yu

Department of Computer Science  
Virginia Tech  
Falls Church, VA, USA  
jianger@vt.edu

Chang-Tien Lu

Department of Computer Science  
Virginia Tech  
Falls Church, VA, USA  
clu@vt.edu

**Abstract**—Electroencephalograms (EEGs) are essential tools for detecting epilepsy using artificial intelligence, despite their inherent variability in format and fidelity. Most studies focus on patient-specific, high-quality EEGs with fixed sampling rates. However, the real-world challenge lies in diagnosing epilepsy across diverse patients with varying sampling rates, low-quality signals, and incomplete data. To tackle these challenges, we introduce the CPEDNet model, designed for Cross-Patient Epilepsy Diagnosis using EEGs that may have diverse sampling, missing data, high noise, insufficient labels, and volatility. Specifically, we employ a latent Neural Ordinary Differential Equation (NODE) method to enhance EEG signals, mitigating the challenges posed by irregular sampling, missing data, and outliers. Then, we represent the enhanced EEG signals as brain network flows, capturing neural representations across spatial and temporal dimensions. Furthermore, we integrate a score-based, two-stage self-supervised strategy to strengthen multi-channel EEG feature learning and ensure stability in the temporal dynamics of the brain network flow. Extensive testing on two real-world datasets shows that our CPEDNet surpasses current methods, effectively learning spatiotemporal patterns from diverse EEG formats. Abundant ablation studies validate the effectiveness of each module within our model.

**Index Terms**—cross-patient epilepsy diagnosis, low-quality EEG signal, neural ordinary differential equation, Transformer, Anomaly detection

## I. INTRODUCTION

Epilepsy is a common neurological disorder characterized by brief and sudden abnormal disturbances in brain neurons. These disturbances lead to severe, widespread, and prolonged convulsions during an episode, posing significant dangers and, in severe cases, potentially life-threatening risks. Over 50 million people worldwide suffer from epilepsy, affecting individuals of all age groups [1]. Thus, it is vital to have an early and accurate diagnosis that facilitates seizure detection in all patients, ensuring effective treatment and management.

Traditionally, neuroscience experts have studied and highlighted Electroencephalograms (EEGs) that do not follow usual patterns to report epilepsy. However, human recognition is time-consuming and labor-intensive, heavily relying on human expertise. Additionally, the massive amount of EEG data can easily lead to fatigue and pose significant challenges to manual processing. In contrast, machine learning methods have demonstrated exceptional performance in large-scale EEG signal analysis tasks. Classical machine

learning techniques [2] often extract features from EEGs from various perspectives, and approach epilepsy detection as a classification problem. Deep learning methods [3] automatically learn underlying nonlinear features from EEGs, thereby more effectively detecting the occurrence of epilepsy. Despite the advancements, most current automated epilepsy detection research focuses on patient-specific scenarios [4]. These methods require learning the EEG patterns unique to each patient, making it difficult to diagnose newly incoming epilepsy patients. Cross-patient EEG-based epilepsy detection is more complex than patient-specific diagnosis but is closer to real-world conditions. Despite the common use of machine learning methods, the potential for cross-patient detection remains largely unexplored.

Furthermore, these models predominantly focus on EEGs with fixed formats, while real-world EEG data often exhibit variable sampling rates, missing values, low signal-to-noise ratio, and instability. EEGs are prone to loss due to electrode detachment and data corruption, and are affected by various factors, such as movement, muscle activity, and electrical interference. EEG-based epilepsy detection methods generally handle missing data through either omission or imputation. Omission uses only available data, but high missing rates can lead to inaccurate results and misleading models. Imputation fills in missing values, using techniques from simple methods to advanced deep learning approaches like Long Short-Term Memory (LSTM), Generative Adversarial Network (GAN), and variational autoencoder (VAE) [5]. However, these methods often assume a fixed missing rate, while real-world EEGs usually have variable missing observations, adding complexity to multivariate time series analysis.

Similarly, we can employ traditional or deep learning-based methods for noise reduction in EEG-based epilepsy detection. Conventional methods include regression, blind source separation, wavelet decomposition, empirical mode decomposition (EMD), and hybrid approaches like EMD-CCA, wavelet-ICA, and EMD-ICA [6]. Deep learning methods use VAEs, GANs, residual convolutional neural networks (CNNs), RNNs, graph neural networks (GCNs), and Transformer-based denoising networks [7]. Despite advancements, effectively leveraging multi-channel signal features and dynamic inter-/intra-channel relations to improve cross-patient diagnostic capability re-

mains challenging. Moreover, the lack of labeled data and the diversity of anomalies make it difficult for supervised learning models, which are effective in other data mining areas [8]. Therefore, this paper tackles a more challenging yet practical scenario: cross-patient epilepsy diagnosis from diverse-sampling low-quality EEGs.

To address these challenges, we present **CPEDNet**, an innovative **Cross-Patient Epilepsy Detection** system for diverse-sampling low-quality EEG signals. CPEDNet aims to explore spatiotemporal EEG anomaly patterns across patients using powerful dynamic neural networks. It leverages latent Neural Ordinary Differential Equations (NODE) to enhance EEG signals, addressing sampling variations and signal degradations. CPEDNet effectively extracts spatiotemporal relations, identifying pattern variations through a combination of a spatial graph convolution network and an anomaly-score temporal Transformer. The former precisely identifies fine-grained spatial correlations in EEGs by efficiently pruning unnecessary edges with a graph convolution mechanism. Meanwhile, using an anomaly-score-based training approach, the latter effectively detects abnormal EEG feature changes, further enhancing model generalization performance. The contributions of this paper are as follows:

- **Proposing CPEDNet, a novel Cross-Patient Epilepsy Detection Network.** CPEDNet learns cross-patient EEG variation patterns through an encoder-decoder framework that combines a spatial graph convolution network and an anomaly-score temporal Transformer. It also uses NODE networks to handle the unfixed format of EEG signals.
- **Employing latent NODE networks for EEG signal enhancement.** NODE can reconstruct the original signals by learning the differential dynamics of EEG signals. This approach recovers missing data, removes noise, and aligns data with varying sampling rates, enriching the original EEG data.
- **Designing an EEG brain network flow framework and a score-based two-stage training model.** In the spatial domain, graph convolution networks extract fine-grained high-order correlations from multi-channel EEG signals. In the temporal domain, a score-based two-stage training Transformer captures temporal dependencies.
- **Evaluating CPEDNet's performance with extensive experiments.** Experimental results on real-world datasets show that CPEDNet outperforms state-of-the-art methods, especially in detecting cross-patient EEG anomalies with diverse sampling and low-quality signals. Each module in CPEDNet is proved effective.

## II. RELATED WORK

EEG-based epilepsy detection is an active and evolving research field, categorized into statistical methods, feature extraction methods, and deep learning methods.

### A. Statistical method

Statistical methods were the earliest approaches, focusing on converting non-stationary EEG signals into stationary ones and

using statistical measures like mean and variance to detect epilepsy. Techniques like Keogh's "Hot SAX" [9] and Ren's interval probability distribution [10] have been employed for patient-specific epilepsy detection. Barz et al. used Kullback-Leibler divergence for unsupervised patient-specific epilepsy detection in multi-channel EEG [11].

### B. Feature extraction method

Feature extraction methods are widely used in EEG-based epilepsy detection, focusing on extracting dynamic rhythm changes in EEG signals through time-domain, frequency-domain, time-frequency, and nonlinear features. De Aguiar Neto FS et al. use phase lag index with Support Vector Machine (SVM) [12], Mumtaz et al. using synchronization likelihood with SVM, logistic regression, and naive Bayesian methods [13], and Faust improving time-frequency analysis for better resolution [14]. Methods such as discrete wavelet transforms, energy spectral density, and functional connectivity networks have also been used for epilepsy detection. Ashokumar employed bicoherence amplitude, normalized bispectral entropy, and normalized quadratic entropy for feature extraction [15]. Adeli analysis to isolate relevant information from specific EEG sub-bands and changes in the maximum Lyapunov exponent for seizure detection [16].

### C. Deep learning method

Deep learning methods have become the most popular approach for epilepsy detection due to their ability to automatically learn complex features from raw EEG data. Gramacki et al. [17] proposed a CNN based on spatial-temporal attention mechanisms, while Djamel et al. [18] combined CNN and RNN for sleep state EEG anomaly detection. Researchers have also explored graph neural networks and hierarchical attention mechanisms to enhance detection accuracy. Transformer-based models show promise, with Panchavati et al. [19] employing a pre-trained Transformer model for seizure detection. Ho et al. [20] applied self-supervised contrastive learning, and Chen et al. [21] used data augmentation to achieve superior performance with fewer labeled data. Despite these advancements, most research has focused on patient-specific scenarios, often overlooking the cross-patient applicability of these methods.

## III. METHODS

As depicted in Fig. 1, the proposed CPEDNet consists of four main parts. NODE-ENHANCE utilizes latent NODE to enhance raw EEG signals. EEG-CORRELATION models these enhanced signals as brain network flows, leveraging a graph convolutional neural network to learn high-order spatial correlations. EEG-SLIDING converts these embeddings into sliding window and context sequences, capturing temporal dependencies. IST-TRANSFORMER aggregates temporal information, assigns higher scores to slices with abnormal feature changes, and projects the learned dynamic graph representation to the expected output.

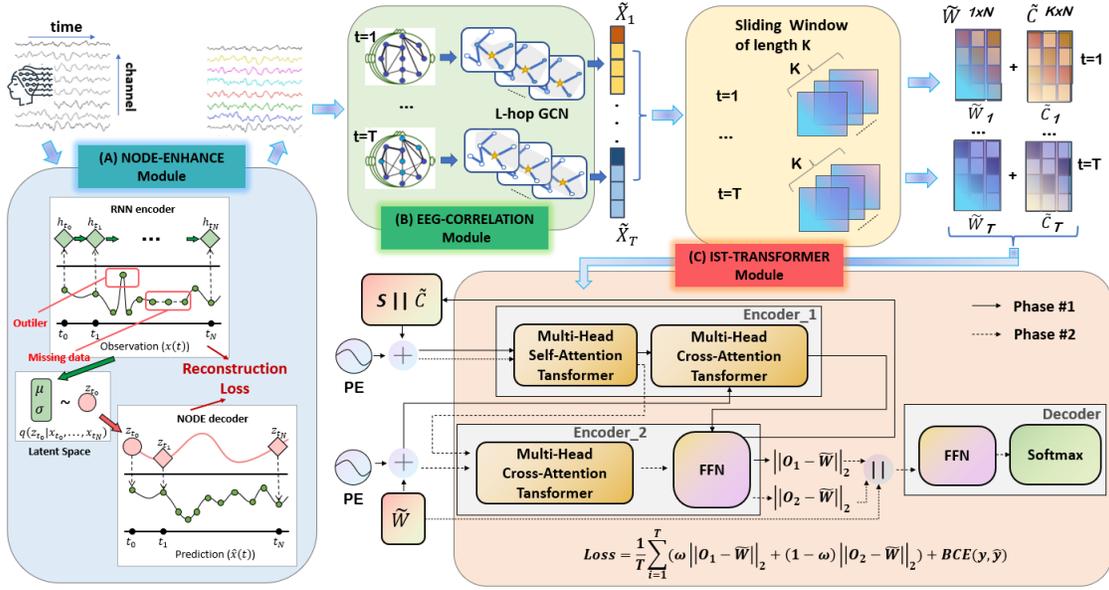


Fig. 1: The CPEDNet framework. (A) **NODE-ENHANCE** receives raw EEGs and outputs enhanced EEGs. (B) **EEG-CORRELATION** constructs a brain network and learns high-order spatial correlations at each timestamp. (C) **IST-TRANSFORMER** learns "anomaly scores"  $S$  for each time slice and projects these scores to the expected output.

### A. NODE-ENHANCE

This module uses neural ODE generation technology to enhance raw EEGs, aiming to repair missing and damaged data, remove noise, and align data with varying sampling rates.

Let the initial state at time step  $t_0$  be  $h_0$ . The dynamic representation of the hidden layer state is formulated as a feedforward neural network, as shown in (1).

$$\frac{dh(t)}{dt} = f(\mathbf{h}(t), t, \theta) \quad (1)$$

Where  $f(\cdot)$  is the function to be solved,  $h(t)$  represents the hidden state at time  $t$ , and  $\theta$  is the parameters of the function  $f(\cdot)$ . The output state can be expressed as the integral of the dynamic function over time, as shown in (2).

$$\mathbf{h}(T) = \text{ODESolver}(\mathbf{h}(0), f, [0, T], \theta) \quad (2)$$

We adopt the latent VAE framework for EEG signal enhancement [22]. Given EEG signal of each channel  $X(t) = \{x_{t_0}, x_{t_1}, \dots, x_{t_T}\}$ , observation time  $\{t_0, t_1, \dots, t_T\}$  and its latent initial state  $z_{t_0}$ , an ODE solver can produce  $\{z_{t_1}, \dots, z_{t_T}\}$ , which describe the latent state at each observation. We define this generative model formally as (3)-(4).

$$z_{t_0}, z_{t_1}, \dots, z_{t_T} = \text{ODESolver}(z_{t_0}, f, \theta_f, t_0, t_1, \dots, t_T) \quad (3)$$

$$z_{t_i} \sim p(x|z_{t_i}, \theta_x) \quad (4)$$

Where function  $f(\cdot)$  is a time-invariant function that takes each  $z_{t_i}$  as input and outputs the gradient in (1). We parametrize this function using a neural network. The latent VAE is trained with sequence-valued observations as follows.

Step 1: Run an RNN encoder through EEG series to infer the parameters for the posterior distribution over  $z_{t_0}$ , as (5).

$$q(z_{t_0} | \{x_{t_i}, t_i\}_i, \phi) = \mathcal{N}(z_{t_0} | \mu_{z_{t_0}}, \sigma_{z_0}) \quad (5)$$

Step 2: Sample  $z_{t_0}$  from the posterior distribution of  $z_{t_0}$ .

Step 3: Given a sample of  $z_{t_0}$ , we can find the latent state at any time by solving an ODE initial-value problem as (3).

Step 4: Train both the encoder and decoder jointly by maximizing the evidence lower bound (ELBO) in (6).

$$\begin{aligned} \mathcal{L} &= \left[ \sum_{i=1}^T \log p(x_{t_i} | z_{t_i}, \theta_x) \right] + \log p(z_{t_0}) - \log (q(z_{t_0} | \{x_{t_i}, t_i\}_i, \phi)) \\ &= 0.5 \left( \sum_{i=1}^T \frac{(x_{t_i} - \hat{x}_{t_i})^2}{\sigma_x^2} - (\log \sigma_{z_{t_0}}^2 + \mu_{z_{t_0}}^2 + \sigma_{z_{t_0}}^2) \right) \end{aligned} \quad (6)$$

Since  $\sigma_x^2$  relies only on the input, the loss function can be simplified as (7).

$$\mathcal{L} = 0.5 \left( \sum_{i=1}^T (x_{t_i} - \hat{x}_{t_i})^2 - (\log \sigma_{z_{t_0}}^2 + \mu_{z_{t_0}}^2 + \sigma_{z_{t_0}}^2) \right) \quad (7)$$

Through the process above, the RNN-ODE network enhances EEGs for each channel. This enhancement involves removing noise, repairing damaged or missing data, and aligning EEG signals with varying sampling rates.

### B. EEG-CORRELATION

1) **Brain network flow construction:** To capture the fine-grained spatial correlations of multi-channel EEGs, we construct correlation networks between EEG channels for each time slice, modeling the enhanced EEG signals as brain network flows. The process is as follows.

The EEG sequence of each person is divided into several time slices through a sliding window, each with a length of  $T$ . The brain network representation of a time slice is  $G = (V, E)$ , where  $V$  represents the set of nodes (EEG channels) and  $E$  is the set of edges. Each edge is assigned a weight to indicate the degree of node correlation. The nodes  $V$  can be represented by a feature matrix  $X \in \mathbb{R}^{N \times T}$  and an adjacency matrix  $A \in \mathbb{R}^{N \times N}$ .  $X = \{x_1, x_2, \dots, x_N\}$  represents  $N$  channels each with  $T$  sampling points along time.

For any two channels  $i$  and  $j$  ( $i, j = 1, 2, \dots, N$ ), we calculate the Pearson Correlation Coefficient Distance (PCCD) between them, as shown in (8):

$$p_{ij} = \frac{\sum_{t=1}^T (x_i(t) - \bar{x}_i)(x_j(t) - \bar{x}_j)}{\sqrt{\sum_{t=1}^T (x_i(t) - \bar{x}_i)^2} \sqrt{\sum_{t=1}^T (x_j(t) - \bar{x}_j)^2}} \quad (8)$$

Where  $x_i(t)$  and  $x_j(t)$  represent the EEG observations of channels  $i$  and  $j$ , respectively. The mean values of the EEG signals for channels  $i$  and  $j$  are denoted by  $\bar{x}_i$  and  $\bar{x}_j$ , respectively. Thus, we can derive the PCCD matrix for the EEG signals of a time slice as  $P \in \mathbb{R}^{N \times N}$ . The adjacency matrix  $A = P - \lambda I$ , where  $I$  is the identity matrix, and  $\lambda$  is an adjustment factor.

2) **Brain network embedding**: To explore the correlations between EEG signals, we apply graph convolution operations on the brain networks  $G = (V, E)$ . This approach helps us examine both local and global spatial dependencies among channels within each time slice. To reduce the computational complexity of the convolution process, we employ Chebyshev polynomials extended to  $L$  orders. This enables the EEG-CORRELATION module to capture  $L$ -order correlation between channels, as depicted in (9).

$$\tilde{\mathbf{X}} = \sum_{l=0}^L \Theta_l \cdot T_l(\tilde{\mathbf{L}}) \cdot \mathbf{X} \quad (9)$$

where  $\tilde{\mathbf{X}}$  is the output tensor,  $\mathbf{X}$  is the input feature tensor,  $\Theta_l$  are the learnable parameters,  $T_l(\tilde{\mathbf{L}})$  are the Chebyshev polynomials, and  $\tilde{\mathbf{L}}$  is the normalized Laplacian matrix.

### C. IST-TRANSFORMER

Since brain signals in epilepsy patients typically contain a small number of epileptic seizure states and a large number of normal states, identifying and utilizing seizure states is crucial in epilepsy classification. Therefore, inspired by [23], this module assists classification by learning ‘‘anomaly scores’’ of EEG signals using a score-based two-stage training strategy.

1) **Preprocessing**: At each timestamp of an EEG signal slice, we compute its anomaly score by evaluating the difference between its values and the values from several preceding timestamps. Specifically, we extract small EEG slices with a sliding window of length  $K$ , where  $K \ll T$ . Denote the embedding of a timestamp  $t$  after the graph convolution as  $\tilde{\mathbf{X}}_t$ . The sliding window extracts a set of embeddings  $\{\tilde{\mathbf{X}}_{t-K+1}, \tilde{\mathbf{X}}_{t-K+2}, \dots, \tilde{\mathbf{X}}_t\}$ , where  $\tilde{\mathbf{W}} = \tilde{\mathbf{X}}_t$  is the timestamp to be assigned anomaly scores and the entire set of

embeddings,  $\tilde{\mathbf{C}} = \{\tilde{\mathbf{X}}_{t-K+1}, \tilde{\mathbf{X}}_{t-K+2}, \dots, \tilde{\mathbf{X}}_t\}$  is considered the contexts of  $\tilde{\mathbf{X}}_t$ . For  $t < K$ , we pad a vector with a length of  $K - t$  to maintain the window length of  $K$  for each time slice. In total, there are  $T$  pairs of  $\tilde{\mathbf{C}}$  and  $\tilde{\mathbf{W}}$  in total.

2) **Module architecture**: IST-Transformer depicted in Fig. 1(C) is a composite Transformer that integrates two encoders and one decoder. Encoder\_1 processes the complete sequence up to the current timestamp  $\tilde{\mathbf{C}}$ , to assign anomaly scores and generate the initial representation of  $\tilde{\mathbf{W}}$ , denoted as  $\mathbf{O}_1$ . Encoder\_2 utilizes the anomaly scores to produce a second representation of  $\tilde{\mathbf{W}}$ ,  $\mathbf{O}_2$ . The representations  $\mathbf{O}_1$  and  $\mathbf{O}_2$  are then concatenated with  $\tilde{\mathbf{W}}$  and passed to the Decoder to obtain the classification result. During training, the differences between  $\mathbf{O}_1$  and  $\tilde{\mathbf{W}}$  as well as between  $\mathbf{O}_2$  and  $\tilde{\mathbf{W}}$  are minimized alongside the classification loss, enabling the model to classify EEG signals while effectively identifying abnormal patterns. The core components of the two encoders are multi-head attention layers, as described in (10).

$$MultiHeadAtt(Q, K, V) = Concat(H_1, \dots, H_h)W^O \quad (10)$$

Where each  $H_h$ , computed as (11), presents the latent state of the  $h$ -th head computed with attention.

$$H_h = softmax\left(\frac{(QW_h^Q)(KW_h^K)^T}{\sqrt{N_h}}\right)(VW_h^V) \quad (11)$$

Where  $Q$ ,  $K$ , and  $V$  are query, key, and value matrices and  $W_h^Q$ ,  $W_h^K$ ,  $W_h^V$  are their corresponding linear projection parameter matrices.  $N_h$  is the number of channels per head. We adopted positional encoding to indicate temporal characteristics while obtaining  $Q$ ,  $K$ , and  $V$ . The position encodings are acquired from in (12) and (13).

$$PE_{(pos, 2i)} = \sin\left(\frac{pos}{10000^{2i/d}}\right) \quad (12)$$

$$PE_{(pos, 2i+1)} = \cos\left(\frac{pos}{10000^{2i/d}}\right) \quad (13)$$

Where  $pos$  is the index of the timestamp of the input sequence.  $i$  is the index of the embedding dimension, and  $d$  is set to  $N$ . After the attention layers, a feed-forward network module is employed to project the attention outputs to the shape of  $\tilde{\mathbf{W}}$ .

3) **Workflow**: IST-Transformer’s workflow is as follows.

Step 1: Encode the context embedding  $\tilde{\mathbf{C}}$  with self-attention. We first add the positional encoding to  $\tilde{\mathbf{C}}$ . The anomaly score tensor  $\mathbf{S}^{K \times N}$  is initialized as an all-zero tensor and concatenated with the context embedding. We denote the result of the operations above as  $I_1$ . The self-attention module of Encoder\_1 then performs the operations in (14) and (15).

$$I_1' = LayerNorm(I_1 + MultiHeadAtt(I_1, I_1, I_1)) \quad (14)$$

$$I_1'' = LayerNorm(I_1' + FeedForward(I_1')) \quad (15)$$

Step 2: Encode the target sequence  $\tilde{\mathbf{W}}$  with cross-attention with  $I_1''$ . Similar as Step 1, we combine the positional encoding with  $\tilde{\mathbf{W}}$  and feed them into Encoder\_2, denoted as  $I_2$ . Encoder\_2 will then perform the operations in (16) – (18).

$$I'_2 = \text{Mask}(\text{MultiHeadAtt}(I_2, I''_1, I''_1)) \quad (16)$$

$$I''_2 = \text{LayerNorm}(I'_2 + \text{FeedForward}(I'_2)) \quad (17)$$

$$\mathbf{O}_1 = \text{Sigmoid}(WI''_2 + b) \quad (18)$$

The anomaly score  $S$  can then be updated with  $(\tilde{\mathbf{C}} - \mathbf{O}_1)^2$ , where  $\mathbf{O}_1 \in \mathbb{R}^{1 \times N}$  repeats  $K$  times to match the shape of  $\tilde{\mathbf{C}} \in \mathbb{R}^{K \times N}$ .

Step 3: Encode the target sequence  $\tilde{\mathbf{W}}$  with the updated score. We repeat Step 1 and Step 2 with the updated anomaly scores. Note that Step 3 reuses the parameters of the modules represented by (14), (15), and (18). The output of this step is denoted as  $\mathbf{O}_2 \in \mathbb{R}^{1 \times N}$ .

Step 4: Decode the target timestamp embedding into classification results. For each timestamp, we concatenate  $\|\tilde{\mathbf{W}} - \mathbf{O}_1\|_2$ ,  $\|\tilde{\mathbf{W}} - \mathbf{O}_2\|_2$ , and  $\tilde{\mathbf{W}}$  along the channel dimension, where  $\|(\cdot)\|_2$  is the second order difference. Consider the stacked sequence of the embeddings as  $\mathbf{O} \in \mathbb{R}^{T \times 3N}$ . We flatten the embedding and apply a feed-forward network to acquire the estimated classification result  $\hat{\mathbf{y}}$  as in (19).

$$\hat{\mathbf{y}} = \text{FeedForward}(\text{flatten}(\mathbf{O})) \quad (19)$$

**4) Two-phase self-supervised training:** Given that EEG signals during epileptic seizures typically represent a small portion of the overall EEG trajectory, it's crucial to dynamically pinpoint the EEG time window indicative of epilepsy. Under the assumption that *segments affected by epilepsy exhibit more pronounced changes than other sections*, we employ the two-phase adversarial training technique proposed in [23] to achieve anomaly score assignment.

Phase 1: Generate an approximate reconstruction of  $\tilde{\mathbf{W}}$  without anomaly score. Initially, all the anomaly scores in  $\mathbf{S}$  are zero. After Step 1 and Step 2 above, a new anomaly score for the entire window is computed as  $\mathbf{S} = (\tilde{\mathbf{C}} - \mathbf{O}_1)^2$ , where  $\mathbf{O}_1$  is an approximate reconstruction of  $\tilde{\mathbf{W}}$  and broadcasted  $K$  times along the time dimension to match the shape of  $\tilde{\mathbf{C}}$ .

Phase 2: Generate an approximate reconstruction of  $\tilde{\mathbf{W}}$  with anomaly score. With the new anomaly score  $\mathbf{S}$ , we go through Step 3 to get the second reconstruction of  $\tilde{\mathbf{W}}$ , denoted as  $\mathbf{O}_2$ .

The reconstruction loss function integrates the reconstruction errors from both phases, as illustrated in (20).

$$\mathcal{L}_{rec} = \omega \|\mathbf{O}_1 - \mathbf{W}\|_2 + (1 - \omega) \|\mathbf{O}_2 - \mathbf{W}\|_2 \quad (20)$$

Where  $\omega$  is a hyper-parameter decreases during training. In this work, we set  $\omega = \frac{1}{n}$ , where  $n$  is the index of training iterations. The eventual version of the loss is (21).

$$\mathcal{L} = \omega \|\mathbf{O}_1 - \mathbf{W}\|_2 + (1 - \omega) \|\mathbf{O}_2 - \mathbf{W}\|_2 + \text{BCE}(\mathbf{y}, \hat{\mathbf{y}}) \quad (21)$$

Where  $\text{BCE}(\cdot)$  denotes binary cross entropy loss evaluating the differences between two distributions.

## A. Dataset

Our experiments utilize the CHB-MIT database from Boston Children's Hospital [24] and the TUSZ database from Temple University [25]. EEGs' collection in the CHB-MIT dataset follows the international standard 10-20 system at a 256 Hz sampling frequency. The TUSZ dataset, part of Temple University Hospital's EEG database, is a prominent resource for epileptic EEG identification comprising 3,050 epilepsy annotations. Records in the TUSZ dataset have diverse sampling rates in the standard 10/20 placement. Dataset statistics are summarized in Table I.

TABLE I: Dataset Statistics.

Database	Patients	Records	Rate	Channels
CHB-MIT	24	686	256	23
TUH-TUSZ	637	5612	250-1000	20

## B. Evaluation metrics and experimental configuration

In this study, we employ five key metrics to gauge model performance: Accuracy (ACC), sensitivity (SEN), specificity (SPE), F1 score, and the area under the ROC curve (AUC) which is the integral value of the ROC curve.

Due to computational resource constraints, we randomly sampled subsets of records for the experiments in this study. For each dataset, after assigning patients to training, validation, or test groups, we divided the recordings into ten-second samples. Specifically, for the CHB-MIT dataset, we sampled 3780 samples (2400 of which were normal) from 24 subjects. For the TUH-TUSZ dataset, we used 3773 samples (1892 of which were normal) from 53 subjects.

The EEG signal enhancement task is designed as a self-supervised reconstruction task, assuming all EEG signals are available for training the RNN-ODE model, as EEG records are typically accessible before diagnosis in real-world scenarios. If a new patient's record outside the existing dataset is introduced, the enhancement model can be fine-tuned to adapt to the distribution of these new records with minimal additional effort. During training, we randomly sample  $T/8$ ,  $T/4$ ,  $T/2$ , and  $T$  timestamps from the original data to ensure that the trained model remains robust to various time intervals and temporal dependencies. For enhancement, we reconstruct sequences of length  $T$  to be inputs for the classification task.

For classification, the data were then divided into training, validation, and test sets in the ratio of 8:1:1. To enhance the reliability of our model, we performed ten-fold cross-validation. The experiments were conducted in a 64-bit Linux system with an Intel(R) Core(TM) i7-12700k processor, 32GB of RAM, RTX 3060 GPU, and 12GB VRAM. Hyper-parameters of our CPEDNet model is as follows: learning rate is  $1^{-4}$ , number of epochs to train is 500, number of units in the hidden layer is 64, the dropout rate is 0.2, weight\_decay is  $1^{-4}$ , L-hop neighbors for GCN is 4, the batch size is 128, and optimizer in training process is RMSprop.

### C. Baseline

To validate the effectiveness of our model in detecting epilepsy, we compared it with advanced methods using the same datasets. The baseline methods considered in this paper include: SVM [12], LR [13], Reservoir-LIF [26], CNN [17], LSTM [27], 1D-CNN-LSTM [28], DeepConvNet [29], ResNet18 [30], ST-GCN [31], MGCL-ACO [32], Bi-GRU [33], GAT+Bi-LSTM [34], and EEGNet [35].

### D. Comparison result

1) **Comparison with baselines:** We present comparative analyses of our CPEDNet model with other baseline models on the two datasets in Table II and III. The results demonstrate that our model surpasses the baseline methods in the majority of indicators for CHI-MIT and TUSZ.

TABLE II: Comparison results on the CHB-MIT Dataset (%). **Bold** means the best, underline means the second best.

Method	ACC	SPE	SEN	F1	AUC
SVM	76.65	66.26	66.19	64.60	74.21
LR	87.40	88.43	87.11	85.16	87.34
Reservoir-LIF	96.40	96.56	86.35	-	-
CNN	96.70	95.40	92.35	95.72	94.38
LSTM	95.00	93.21	91.32	94.51	96.27
1D-CNN-LSTM	95.75	95.77	95.93	-	96.62
DeepConvNet	96.80	97.21	96.53	-	-
ResNet18	98.44	98.57	98.13	95.36	96.51
EEGNet	90.21	81.08	93.02	86.64	88.58
ST-GCN	96.32	95.29	91.02	93.25	96.15
MGCL-ACO	96.62	96.39	97.25	95.47	96.02
Bi-GRU	98.49	93.89	<b>98.49</b>	-	-
GAT+Bi-LSTM	98.52	97.75	94.34	<b>95.90</b>	96.81
CPEDNet	<b>99.17</b>	<b>98.85</b>	97.68	95.73	<b>98.26</b>

TABLE III: Comparison results on the TUSZ Dataset (%). **Bold** means the best, underline means the second best.

Method	ACC	SPE	SEN	F1	AUC
SVM	59.35	60.21	58.75	62.14	58.53
LR	47.27	43.98	73.10	54.92	50.08
CNN	75.42	74.62	75.05	75.73	73.28
LSTM	78.31	77.33	77.74	85.37	76.25
1D-CNN-LSTM	81.35	77.54	83.75	80.53	78.97
DeepConvNet	80.30	79.67	<b>92.31</b>	85.52	80.85
ResNet18	84.20	83.52	84.61	84.69	83.51
EEGNet	70.91	72.64	85.03	76.12	77.25
ST-GCN	86.95	85.39	84.27	87.29	<b>86.48</b>
MGCL-ACO	84.62	83.28	84.15	88.74	80.35
Bi-GRU	86.36	84.75	85.24	-	-
GAT+Bi-LSTM	87.55	<b>87.36</b>	86.41	86.07	86.24
CPEDNet	<b>88.42</b>	86.57	87.04	<b>89.25</b>	85.93

(1) Traditional machine learning such as LR and SVM have shown limited effectiveness in seizure detection due to their inability to capture the complexity of EEG signals.

(2) Deep learning methods such as CNN and LSTM have improved seizure detection by recognizing complex EEG signal features. However, they often fail to fully utilize the complementary nature of spatial and temporal information.

(3) Spatial-temporal integration method ST-GCN attempts to integrate spatial and temporal EEG features, leading to

better anomaly detection. Yet, it falls short in capturing long-term dependencies, which is critical for seizure detection.

(4) Bidirectional networks like Bi-GRU and GAT+Bi-LSTM enhance seizure detection by modeling spatial-temporal dependencies more effectively. They leverage the bidirectional context in EEG signals, improving detection accuracy.

(5) FC+GA-LASSO+SVM approach mines time-frequency information and assigns discriminative weights to different features. It also connects features through graph topology, enhancing seizure detection performance.

(6) CPEDNet excels in epileptic seizure detection. Its integration of NODE for signal enhancement, coupled with a two-stage self-supervised learning strategy and a Transformer architecture, ensures robust feature extraction and accurate capture of long-range dependencies, positioning CPEDNet as a leading model for seizure detection.

2) **Leave-One-Out Experiment result:** This section verifies the cross-patient performance of the model. For the CHB-MIT dataset, we employed the Leave-One-Out Cross-Validation method, using each patient's data as the test set and combining data from the other 21 patients as the training and validation set. The outcomes are summarized in Table IV.

TABLE IV: Detection Results for each CHB-MIT subject (%).

Subject	ACC	SPE	SEN	F1	AUC
1	99.87	99.56	99.68	99.93	99.99
2	99.92	99.37	99.51	98.76	97.64
3	98.89	98.61	99.53	99.87	99.14
4	99.21	98.75	98.58	96.75	98.91
5	100.00	99.35	98.64	99.95	99.01
6	98.53	99.03	98.05	98.62	97.25
7	99.36	98.63	99.05	98.76	98.48
8	98.65	98.53	98.56	99.04	98.95
9	99.54	99.07	99.72	98.52	98.81
10	98.45	97.56	98.57	98.39	97.83
11	98.96	99.42	99.08	98.67	99.31
12	100.00	99.24	98.62	99.53	99.01
13	99.26	98.40	99.17	99.58	99.43
14	99.74	99.05	99.93	99.47	98.79
15	98.62	99.28	97.49	98.05	99.27
16	98.37	97.53	98.69	99.02	99.51
17	98.06	98.29	97.58	99.49	98.06
18	98.35	97.94	98.53	98.59	99.18
19	99.45	99.42	99.35	99.27	99.60
20	99.65	99.67	99.73	98.37	99.03
21	98.56	97.31	97.63	96.58	97.07
22	99.97	99.48	99.30	99.58	99.99
23	99.65	98.51	98.68	99.39	99.20
<b>Average</b>	<b>99.17</b>	<b>98.68</b>	<b>98.85</b>	<b>98.87</b>	<b>98.84</b>

From the table, it is clear that the detection accuracy for all subjects exceeds 98%. The average accuracy of the CPEDNet model reaches 99.17%, and the average values of other metrics also remain above 98%. Our proposed CPEDNet model demonstrates superior performance in accurately identifying epilepsy and consistently shows stable and reliable detection performance across multiple experiments.

### E. Ablation experiment

To assess the effectiveness of each CPEDNet module, we conducted a series of ablation studies on the CHB-MIT dataset, as shown in Fig. 2.

- (1) Importance of EEG signal enhancement.
  - Variant-q: EEG signals without enhancement.
  - Variant-r: EEG signals enhanced by NODE, with the enhanced signals as features for epilepsy diagnosis.
- (2) Importance of spatiotemporal fusion.
  - Variant-l: using GCN for spatial convolution, and feed the feature embedding into a classifier for epilepsy diagnosis.
  - Variant-m: using Transformer to reconstruct the original EEG signals and detect epilepsy based on the differences between the reconstructed and original signals.
  - Variant-n: using GCN and Transformer to explore temporal and spatial correlations for anomaly detection.
- (3) Effectiveness of two-stage training strategy.
  - Variant-o: Without using the two-stage training strategy, the reconstruction loss from the first stage is used as the anomaly score for epilepsy detection.
  - Variant-p: Using the two-stage training strategy, the reconstruction loss from the first stage is used as input for the second stage in epilepsy detection.

Figure 2(a) shows that variant-q did not achieve optimal detection results due to varying sampling frequencies, high noise, and signal loss in the original EEGs. In contrast, variant-r, using NODE for EEG signal enhancement, captures continuous dynamic changes, reduces noise, and achieves better anomaly detection performance.

Figure 2(b) shows that variant-l identifies spatial relationships between EEGs to diagnose seizures based on feature differences. Variant-m detects anomalies by leveraging temporal dependencies in EEGs. Variant-n combines the strengths of the both, exploring deeper spatiotemporal features for superior performance. This demonstrates that spatiotemporal fusion improves EEG anomaly detection.

Figure 2(c) shows that variant-o, which uses the difference between the Transformer's output and input as the anomaly score, is prone to misjudgments during normal signal fluctuations. In contrast, variant-p employs a two-stage self-supervised method, using the first stage's reconstruction error as the anomaly score input for the second stage. This captures short-term trends in anomalies, enhancing detection performance, and indicates the effectiveness of the two-stage self-supervised strategy in detecting EEG anomalies.

## F. Discussion

1) **Impact of window size:** This section analyzes the impact of window size on CPEDNet's performance. As shown in Fig. 3, window size affects various performance metrics of epilepsy detection. With increasing window length, accuracy and specificity improve, while sensitivity decreases. Experiments demonstrate that a window length of 10 strikes a good balance, achieving high accuracy and specificity while maintaining high sensitivity.

2) **Visualization of anomalous EEGs detection:** We visualized the EEG signals and diagnostic results for two epileptic patients in the CHB-MIT dataset, as shown in Fig. 4. EEG signals of Subject 1 show significant fluctuations with three

seizure-induced anomalies. Although Subject 2's EEG signals are relatively stable, three distinct anomalies are still observed. For both subjects, the CPEDNet model assigns high anomaly scores to durations of epileptic seizures, demonstrating the CPEDNet's effectiveness in detecting epileptic seizures.

## V. CONCLUSION

This paper presents cross-patient epilepsy diagnosis from diverse-sampling low-quality EEG signals. The model enhances EEG signals using latent NODE, models them as brain network sequences, and employs graph convolution to extract spatial features. A Transformer captures temporal dependencies in the brain network flow, and a two-stage training process improves robustness. Experiments show our method's superior performance over baseline models. Future work will integrate multimodal data for richer and more comprehensive EEG anomaly detection and analysis, improving the identification of abnormal patterns.

## REFERENCES

- [1] W. H. Organization, *Intersectoral global action plan on Epilepsy and other neurological disorders 2022–2031*. World Health Organization, 2023.
- [2] Y. Yuan, G. Xun, K. Jia, and A. Zhang, "A multi-view deep learning framework for eeg seizure detection," *IEEE journal of biomedical and health informatics*, vol. 23, no. 1, pp. 83–94, 2018.
- [3] A. O'Shea, R. Ahmed, G. Lightbody, E. Pavlidis, R. Lloyd, F. Pisani, W. Marnane, S. Mathieson, G. Boylan, and A. Temko, "Deep learning for eeg seizure detection in preterm infants," *International journal of neural systems*, vol. 31, no. 08, p. 2150008, 2021.
- [4] Z. Wang, S. Hou, T. Xiao, Y. Zhang, H. Lv, J. Li, S. Zhao, and Y. Zhao, "Lightweight seizure detection based on multi-scale channel attention," *International Journal of Neural Systems*, pp. 2350061–2350061, 2023.
- [5] J. Yoon, J. Jordon, and M. Schaar, "Gain: Missing data imputation using generative adversarial nets," in *International conference on machine learning*. PMLR, 2018, pp. 5689–5698.
- [6] V. Bono, S. Das, W. Jamal, and K. Maharatna, "Hybrid wavelet and emd/lca approach for artifact suppression in pervasive eeg," *Journal of neuroscience methods*, vol. 267, pp. 89–107, 2016.
- [7] X. Pu, P. Yi, K. Chen, Z. Ma, D. Zhao, and Y. Ren, "Eegdnnet: Fusing non-local and local self-similarity for eeg signal denoising with transformer," *Computers in Biology and Medicine*, vol. 151, p. 106248, 2022.
- [8] A. Hajisafi, H. Lin, Y.-Y. Chiang, and C. Shahabi, "Dynamic gnns for precise seizure detection and classification from eeg data," in *Pacific-Asia Conference on Knowledge Discovery and Data Mining*. Springer, 2024, pp. 207–220.
- [9] E. Keogh, J. Lin, S.-H. Lee, and H. V. Herle, "Finding the most unusual time series subsequence: algorithms and applications," *Knowledge and Information Systems*, vol. 11, pp. 1–27, 2007.
- [10] H. Ren, M. Liu, X. Liao, L. Liang, Z. Ye, and Z. Li, "Anomaly detection in time series based on interval sets," *IEEJ Transactions on Electrical and Electronic Engineering*, vol. 13, no. 5, pp. 757–762, 2018.
- [11] B. Barz, E. Rodner, Y. G. Garcia, and J. Denzler, "Detecting regions of maximal divergence for spatio-temporal anomaly detection," *IEEE transactions on pattern analysis and machine intelligence*, vol. 41, no. 5, pp. 1088–1101, 2018.
- [12] F. S. de Aguiar Neto and J. L. G. Rosa, "Depression biomarkers using non-invasive eeg: A review," *Neuroscience & Biobehavioral Reviews*, vol. 105, pp. 83–93, 2019.
- [13] R. Li, S. Li, J. Roh, C. Wang, and Y. Zhang, "Multimodal neuroimaging using concurrent eeg/fnirs for poststroke recovery assessment: an exploratory study," *Neurorehabilitation and Neural Repair*, vol. 34, no. 12, pp. 1099–1110, 2020.
- [14] O. Faust, U. R. Acharya, H. Adeli, and A. Adeli, "Wavelet-based eeg processing for computer-aided seizure detection and epilepsy diagnosis," *Seizure*, vol. 26, pp. 56–64, 2015.

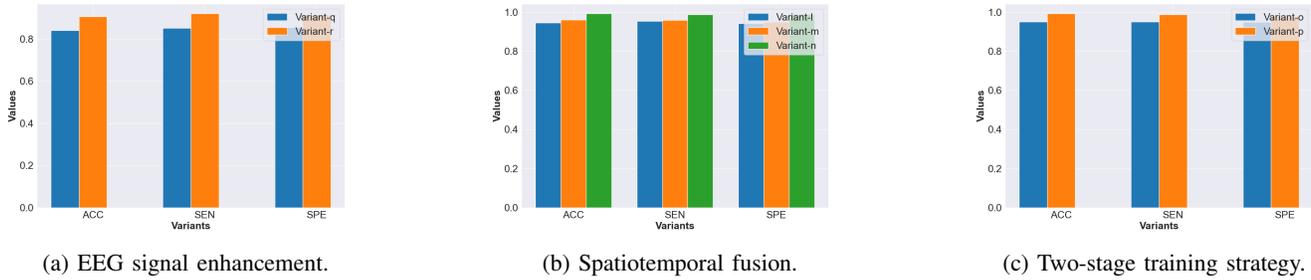


Fig. 2: Ablation experimental results

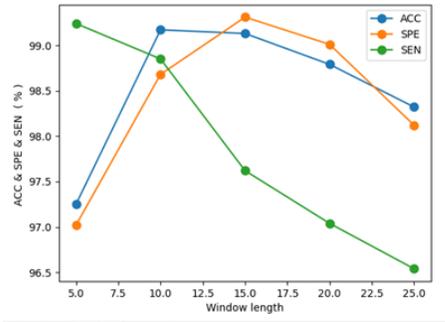


Fig. 3: Impact of window size on CPEDNet's performance.

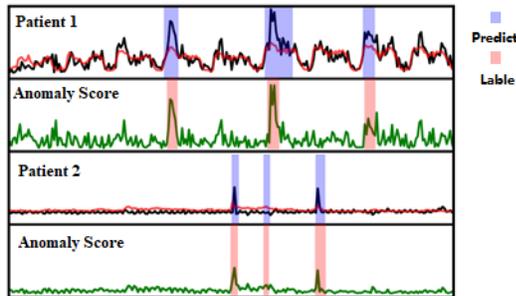


Fig. 4: Two cases of anomalous EEGs detection.

- [15] N. Mahmoodian, J. Haddadnia, A. Illanes, A. Boese, and M. Friebe, "Seizure prediction with cross-higher-order spectral analysis of eeg signals," *Signal, Image and Video Processing*, vol. 14, pp. 821–828, 2020.
- [16] H. Adeli, S. Ghosh-Dastidar, and N. Dadmehr, "A wavelet-chaos methodology for analysis of eegs and eeg subbands to detect seizure and epilepsy," *IEEE Transactions on Biomedical Engineering*, vol. 54, no. 2, pp. 205–211, 2007.
- [17] A. Gramacki and J. Gramacki, "A deep learning framework for epileptic seizure detection based on neonatal eeg signals," *Scientific reports*, vol. 12, no. 1, p. 13010, 2022.
- [18] E. C. Djamal, I. S. Mufadhol, and F. Kasyidi, "Sleep stage identification of split wave eeg signals using cnn-rnn," in *2023 International Conference on Electrical Engineering and Informatics (ICEEI)*. IEEE, 2023, pp. 1–6.
- [19] S. Panchavati, S. Vander Dussen, H. Semwal, A. Ali, J. Chen, H. Li, C. Arnold, and W. Speier, "Pretrained transformers for seizure detection," in *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2023, pp. 1–2.
- [20] T. K. K. Ho and N. Armanfard, "Self-supervised learning for anomalous channel detection in eeg graphs: Application to seizure analysis," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 37, no. 7, 2023, pp. 7866–7874.
- [21] T. Chen, S. Kornblith, M. Norouzi, and G. Hinton, "A simple framework for contrastive learning of visual representations," in *International conference on machine learning*. PMLR, 2020, pp. 1597–1607.
- [22] R. T. Chen, Y. Rubanova, J. Bettencourt, and D. K. Duvenaud, "Neural ordinary differential equations," *Advances in neural information processing systems*, vol. 31, pp. 1–18, 2018.
- [23] S. Tuli, G. Casale, and N. R. Jennings, "Tranad: Deep transformer networks for anomaly detection in multivariate time series data," *arXiv preprint arXiv:2201.07284*, 2022.
- [24] J. Prasanna, M. Subathra, M. A. Mohammed, R. Damaševičius, N. J. Sairamya, and S. T. George, "Automated epileptic seizure detection in pediatric subjects of chb-mit eeg database—a survey," *Journal of Personalized Medicine*, vol. 11, no. 10, p. 1028, 2021.
- [25] V. Shah, E. Von Weltin, S. Lopez, J. R. McHugh, L. Veloso, M. Golmohammadi, I. Obeid, and J. Picone, "The temple university hospital seizure detection corpus," *Frontiers in neuroinformatics*, vol. 12, pp. 1–6, 2018.
- [26] Z. Yang, K. Liu, R. Yuan, X. Wu, L. Cai, T. Zhang, Y. Tao, Y. Jin, and Y. Yang, "Seizure detection using dynamic memristor-based reservoir computing and leaky integrate-and-fire neuron for post-processing," *APL Machine Learning*, vol. 1, no. 4, pp. 1–10, 2023.
- [27] K. M. Tsiouris, V. C. Pezoulas, M. Zervakis, S. Konitsiotis, D. D. Koutsouris, and D. I. Fotiadis, "A long short-term memory deep learning network for the prediction of epileptic seizures using eeg signals," *Computers in biology and medicine*, vol. 99, pp. 24–37, 2018.
- [28] W. Hussain, M. T. Sadiq, S. Siuly, and A. U. Rehman, "Epileptic seizure detection using 1 d-convolutional long short-term memory neural networks," *Applied Acoustics*, vol. 177, p. 107941, 2021.
- [29] R. T. Schirrmeyer, J. T. Springenberg, L. D. J. Fiederer, M. Glasstetter, K. Eggenesperger, M. Tangermann, F. Hutter, W. Burgard, and T. Ball, "Deep learning with convolutional neural networks for eeg decoding and visualization," *Human brain mapping*, vol. 38, no. 11, pp. 5391–5420, 2017.
- [30] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.
- [31] H. Wu, J. Liu, and Y. Zhao, "Eeg-based depression identification using a deep learning model," in *2022 IEEE 6th Conference on Information and Communication Technology (CICT)*. IEEE, 2022, pp. 1–5.
- [32] S. Zhang, H. Wang, Z. Zheng, T. Liu, W. Li, Z. Zhang, and Y. Sun, "Multi-view graph contrastive learning via adaptive channel optimization for depression detection in eeg signals," *International Journal of Neural Systems*, vol. 33, no. 11, p. 2350055, 2023.
- [33] Y. Zhang, S. Yao, R. Yang, X. Liu, W. Qiu, L. Han, W. Zhou, and W. Shang, "Epileptic seizure detection based on bidirectional gated recurrent unit network," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 30, pp. 135–145, 2022.
- [34] J. He, J. Cui, G. Zhang, M. Xue, D. Chu, and Y. Zhao, "Spatial-temporal seizure detection with graph attention network and bi-directional lstm architecture," *Biomedical Signal Processing and Control*, vol. 78, p. 103908, 2022.
- [35] V. J. Lawhern, A. J. Solon, N. R. Waytowich, S. M. Gordon, C. P. Hung, and B. J. Lance, "Eegnet: a compact convolutional neural network for eeg-based brain-computer interfaces," *Journal of neural engineering*, vol. 15, no. 5, p. 056013, 2018.