

# Self-Supervised Learning for Anomalous Channel Detection in EEG Graphs: Application to Seizure Analysis

Thi Kieu Khanh Ho, Narges Armanfard

Department of Electrical and Computer Engineering, McGill University  
Mila - Quebec AI Institute, Montreal, QC, Canada  
{thi.k.ho, narges.armanfard}@mcgill.ca

## Abstract

Electroencephalogram (EEG) signals are effective tools towards seizure analysis where one of the most important challenges is accurate detection of seizure events and brain regions in which seizure happens or initiates. However, all existing machine learning-based algorithms for seizure analysis require access to the labeled seizure data while acquiring labeled data is very labor intensive, expensive, as well as clinicians dependent given the subjective nature of the visual qualitative interpretation of EEG signals. In this paper, we propose to detect seizure channels and clips in a self-supervised manner where no access to the seizure data is needed. The proposed method considers local structural and contextual information embedded in EEG graphs by employing positive and negative sub-graphs. We train our method through minimizing contrastive and generative losses. The employ of local EEG sub-graphs makes the algorithm an appropriate choice when accessing to all EEG channels is impossible due to complications such as skull fractures. We conduct an extensive set of experiments on the largest seizure dataset and demonstrate that our proposed framework outperforms the state-of-the-art methods in the EEG-based seizure study. The proposed method is the only study that requires no access to the seizure data in its training phase, yet establishes a new state-of-the-art to the field, and outperforms all related supervised methods.

## Introduction

Epilepsy is one of the most prevalent neurological disorders affecting 65 million people (Johnson 2019). Patients with epilepsy suffer from sudden and unforeseen seizures, during which they are vulnerable to injury, suffocation and death (Surges et al. 2021). Epileptic seizure detection makes it possible to identify more accurately the epileptogenic zone (EZ), which is the brain area responsible for initiating seizures (Sharma, Rai, and Tewari 2018), and the surgical resection of EZ renders epilepsy patients seizure-free. As epilepsy can start at any age, early detection is crucial to prevent further damage during physiological development and to increase life expectancy (Beghi and Giussani 2018).

Scalp electroencephalogram (EEG) has been considered as the most common tool to detect seizures (Kuhlmann et al.

2018). EEG measures the voltage changes between electrodes by the ionic currents flowing within the brain neurons and provides spatial-temporal information of the brain. However, detecting seizures in EEGs requires a direct examination by experienced EEG readers that requires a huge amount of time and effort. Moreover, different opinions of experts can cause discrepant diagnostic results (Yan et al. 2018). Therefore, the development of automated and objective methods for epileptic seizure detection is needed.

Although many studies have proposed deep learning (DL) based models for automated seizure detection (Saab et al. 2020; Shoeibi et al. 2021; Abdelhameed and Bayoumi 2021; Thuwajit et al. 2021; Khalkhali et al. 2021; Rashed-Al-Mahfuz et al. 2021; Mahajan, Somaraj, and Sameer 2021; Saichand et al. 2021; Shen et al. 2022; Gao et al. 2022), several challenges still remain unsolved. First, these studies train their proposed model in a supervised approach – i.e. availability of labeled seizure data during training is required. This does not address the challenge from a clinical perspective as seizure labels obtaining is difficult and labour expensive during the process of manually seizure marking, hence very scarce in real-world applications.

Second, these studies apply deep convolutional neural networks (CNNs) directly to the time-series signals or spectrograms hence ignore the critically important information physical distance-based connectivity and functional-based connectivity between different brain regions. Several studies have recently proposed graph learning techniques to capture the relationships between EEG electrodes (aka EEG nodes) (Wang et al. 2020; Wagh and Varatharajah 2020; Saboksayr, Mateos, and Cetin 2021; Tang et al. 2021; Saboksayr, Mateos, and Cetin 2021; Mathur and Chakka 2022). However, they all fail to take into consideration the *local* patterns (e.g. local sub-graphs and sub-structures) when learning EEG graphs. Such local information could be effective when detecting anomalies in EEG graphs. The effectiveness of such local information has been recently demonstrated in applications such as analyzing social, traffic, citation, and financial transaction networks (Liu et al. 2021a,b; Zheng et al. 2021). However, its effectiveness in analyzing EEG signals, that are inherently non-stationary and dynamic, has remained unexplored to date.

Third, as is discussed before, in real-world applications, there is not enough training data from seizure class; while

the existence of adequate data from both normal and seizure classes is essential when training supervised algorithms. To handle the imbalance-data issue, some graph-based methods employ graph augmentation (Wagh and Varatharajah 2020; Tang et al. 2021). However, in EEG graphs, not every augmentation technique is effective as some may corrupt the underlying brain region connectivities. Hence, pinpointing appropriate augmentation strategies in EEG graphs that preserve the underlying semantic information is necessary towards proving accurate seizure detection and localization.

Based on the above observations, we propose a novel approach for the detection of abnormal brain regions and EEG channels, called Contrastive and Generative Self-supervised Learning for EEG Graphs (EEG-CGS). Although the effectiveness of contrastive learning methods, as a self-supervised learning technique, has been demonstrated for anomaly detection (Li et al. 2021; Hojjati, Ho, and Armanfard 2022) and graph analysis in general (Zeng and Xie 2021; Xie et al. 2022), its ability on analysing the EEG graphs has remained unexplored.

The main contributions of this paper are as follows:

- We propose a self-supervised method for identifying abnormal brain regions and EEG channels without having access to the abnormal class data during the training phase. To the best of our knowledge, this is the first study for unsupervised identification of abnormal EEG channels and regions.
- We propose to model brain regions and their connectivities using attributed graphs where each EEG channel is considered as a graph node. Each node is associated with a feature vector constructed from the corresponding EEG signal. Nodes are connected based on four different metrics including nodes Euclidean distance, randomly connection of nodes, node features correlations, and directed transfer function; the first two are meant to capture the geometry of EEG channels and the last two are for capturing connectivity of brain regions.
- We propose an effective augmentation approach to create the positive and negative pairs. Augmentations are based on sub-graphs hence allowing to capture the local structural and contextual information.
- The proposed self-supervised method is realized through employing contrastive and generative learning techniques. We define a channel-based anomaly score function which is a linear combination of the contrastive and reconstruction losses.
- Performance of the proposed abnormal EEG node and region detection is demonstrated on the largest and most comprehensive EEG seizure dataset TUSZ. We show that the proposed EEG-CGS establishes a new state-of-the-art on this dataset. EEG-CGS significantly outperforms all existing (supervised) seizure detection techniques.
- The proposed technique can be considered as a general approach for the detection and localization in other brain disorders. The sub-graph based nature of EEG-CGS makes it a suitable choice for applications where not all EEG channels are available or reliable during

training. E.g. in coma outcome prediction application, the scalp of comatose patients is usually uneven with fractions so that recording of all EEG channels is rarely possible (Armanfard et al. 2018).

## Proposed Method

The block diagram of the proposed method is shown in Figure 1. The proposed EEG-CGS for node anomaly detection consists of four components namely EEG graph construction, positive and negative pair sampling, contrastive and generative learning based on graph neural networks (GNN), and anomaly score computation. The following sections discuss these components in detail.

### EEG Graph Construction

An EEG graph is denoted as  $\mathcal{G} = (\mathcal{A}, \mathcal{X})$  where  $\mathcal{A} = (\mathcal{V}, \mathcal{E})$ .  $\mathcal{A} \in \mathbb{R}^{n \times n}$  is the adjacency matrix which its element on the  $i^{th}$  row and  $j^{th}$  column, i.e.  $a_{ij}$ , is greater than zero if a connectivity exists between EEG electrodes  $v_i$  and  $v_j$ .  $\mathcal{V} = \{v_1, \dots, v_n\}$  is the set of  $n$  electrodes (aka nodes),  $\mathcal{E}$  is the set of  $m$  edges,  $\mathcal{X} \in \mathbb{R}^{n \times d}$  is the feature matrix,  $d$  denotes the number of attributes (aka features) representing an EEG signal. The  $i^{th}$  row of  $\mathcal{X}$ , denoted by  $x_i$ , is the feature vector representing the  $i^{th}$  EEG channel. For a given EEG clip, we build four types of EEG graphs:

- *Dist-EEG-Graph* endeavors to embed the structure of electrode locations in the graph’s adjacency matrix, using Euclidean distance between electrodes (Tang et al. 2021). As the electrode locations are fixed in an EEG recording cap, the same adjacency matrix is obtained for all EEG clips. More specifically,  $a_{ij}$  of Dist-EEG-Graph is obtained as follows:

$$a_{ij} = \begin{cases} \exp(-\frac{\|v_i - v_j\|^2}{\zeta^2}) & \text{if } \|v_i - v_j\| \leq k, \\ 0 & \text{if } O.W. \end{cases}, \quad (1)$$

where  $\|\cdot\|$  denotes  $\ell_2$ -norm and  $\zeta$  is a scaling constant. The closer the two electrodes  $v_i$  and  $v_j$  are, the closer  $a_{ij}$  is to 1. In all our experiments  $k$  is set to 0.9 for all EEG clips. Setting  $a_{ij} = 0$  for far away electrodes imposes sparsity to the graph.

- *Rand-EEG-Graph* is built based on the assumption that all electrodes are connected and equally contribute in brain activities. This graph is realized through forming an adjacency matrix as follows:

$$a_{ij} = \begin{cases} 0.5 & \text{if } i \neq j \\ 1 & \text{if } O.W. \end{cases}. \quad (2)$$

In this way, every edge has the chance of being present in the graph.

- *Corr-EEG-Graph* is meant to capture the functional connectivity between electrodes. It is encoded in the adjacency matrix elements defined as below:

$$a_{ij} = \begin{cases} \frac{\text{xcorr}(x_i, x_j)}{\|x_i\| \|x_j\|} & \text{if } v_j \in \mathcal{N}(v_i), \\ 0 & \text{if } O.W. \end{cases}, \quad (3)$$

where  $\text{xcorr}(\cdot, \cdot)$  denotes cross correlation function,  $\mathcal{N}(v_i)$  presents the top-3 neighborhood nodes of  $v_i$  that

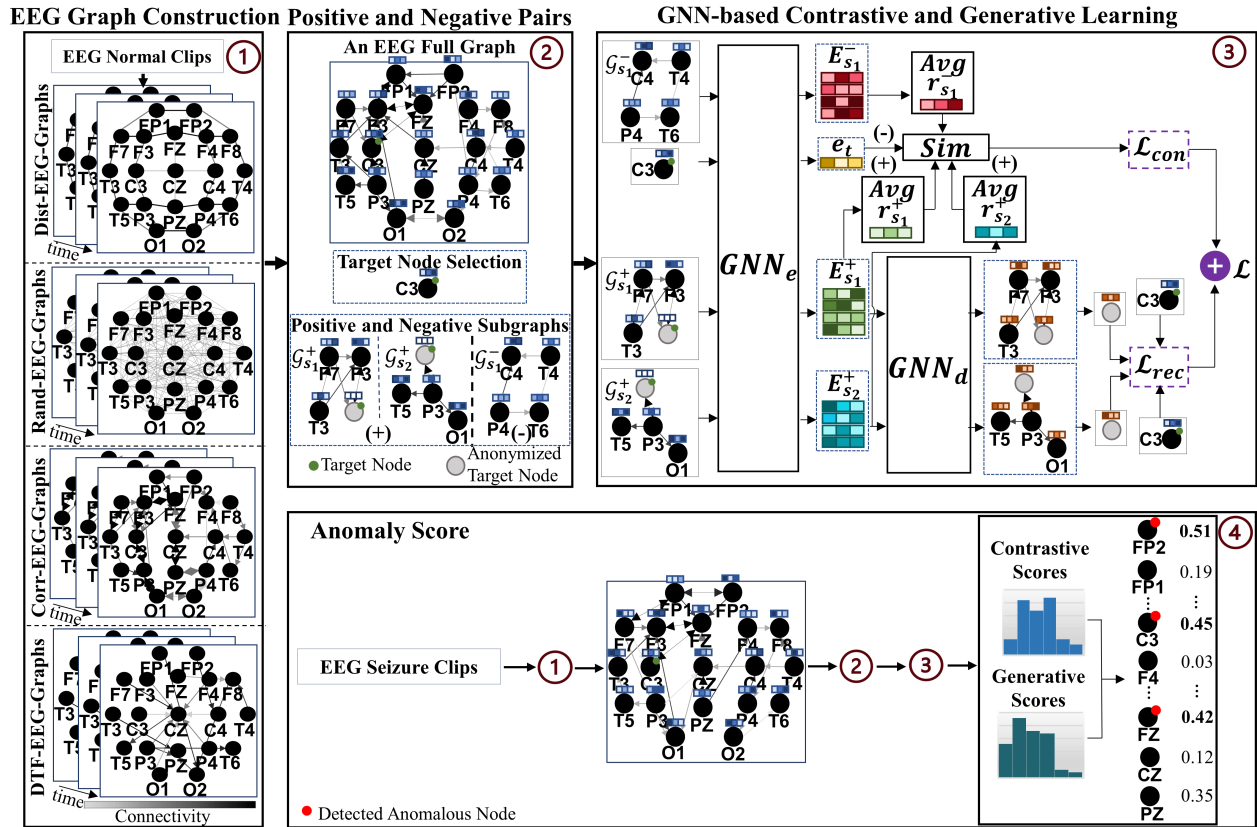


Figure 1: The overall framework of EEG-CGS where  $d' = 3$ ,  $\alpha = 4$  and  $\tau_1 = 0.4$ . The positive and negative pairs are shown for a typical target node C3.

have the highest normalized correlation.  $\mathcal{N}(v_i)$  is set to the top-3 neighborhood nodes to avoid overly connected graphs. We only keep the top three edges (i.e.,  $m = 3$ ) for each node to avoid overly connected graphs.

- *DTF-EEG-Graph* is Directed Transfer Function Graph (Blinowska 2011) that is meant to capture the mutual influence between EEG channels, hence it models functional connectivity of the brain regions. The adjacency matrix for this graph is as follows:

$$a_{ij} = \begin{cases} \frac{\text{xcorr}(x_i, x_j)}{\sqrt{\sum_{m=1, m \neq i, j}^n \|\text{xcorr}(x_i, x_m)\|^2}} & \text{if } v_j \in \mathcal{N}(v_i) \\ 0 & \text{if } O.W. \end{cases}, \quad (4)$$

$a_{ij}$  takes a value from 0 to 1, and denotes a ratio between the correlation of the attributes of  $v_i$  and  $v_j$  to the correlations of the features of all the remaining nodes and  $x_i$ . As similar to Corr-EEG-Graph, we set  $\mathcal{N}(v_i)$  to the top-3 neighbors of  $v_i$ .

### Positive and Negative Pair Sampling

For every EEG clip in an input mini-batch  $\mathcal{B}$ , we first construct the four graphs discussed in the EEG Graph Construction section (see this construction block in Figure 1). Then, we create positive and negative sub-graphs for each clip to be used in its following step where the algorithm's networks are trained using reconstruction and contrastive losses.

We create two positive and one negative sub-graphs for every node in every constructed EEG graph. To this end, we first select an electrode as target node  $v_t$  and then select the two positive sub-graphs  $\mathcal{G}_{s_1}^+$  and  $\mathcal{G}_{s_2}^+$ , and one negative sub-graph  $\mathcal{G}_{s_1}^-$ . Each target node is assigned to its own sub-graphs. To form  $\mathcal{G}_{s_1}^+$  and  $\mathcal{G}_{s_2}^+$ , we leverage a random walk with restart (RWR) technique (Tong, Faloutsos, and Pan 2006) by which sub-graphs are centered at the target node with the fixed size  $\alpha$  which controls the radius of the surrounding contexts. In the other words,  $\alpha$  determines the number of nodes within the sub-graphs. To impose generalization to the system, we anonymize the target node in the positive sub-graphs by replacing its feature vector with an all-zero vector. To form the negative sub-graph  $\mathcal{G}_{s_1}^-$ , we first find the farthest electrode from the target node and then apply the RWR to the farthest node with a same  $\alpha$  value as of the one used for the positive sub-graphs. In Figure 1,  $\alpha$  is set to 4 indicating 4 nodes are considered in the positive and negative sub-graphs where the farthest node to the target node (i.e. C3) is T4.

### Contrastive and Generative Learning

As is shown in Figure 1, we employ GNN autoencoder to realize the unsupervised contrastive and generative learning objectives. The autoencoder consists of two networks: en-

coder and decoder, respectively denoted as  $\text{GNN}_e(\cdot, \cdot)$  and  $\text{GNN}_d(\cdot, \cdot)$ . Encoder takes a graph as input and embed it to a lower dimensional space while decoder takes the encoder's output as input and endeavors to reconstruct the encoder's input.

Assume a sub-graph  $\mathcal{G}_s = (\mathcal{A}_s, \mathcal{X}_s)$ , where  $\mathcal{A}_s \in \mathbb{R}^{\alpha \times \alpha}$  and  $\mathcal{X}_s \in \mathbb{R}^{\alpha \times d}$  respectively denote the sub-graph adjacency matrix and the node attributes. The encoder takes  $\mathcal{G}_s$  as input and embed it to a lower dimensional space as below:

$$E_s = \text{GNN}_e(\mathcal{X}_s, \mathcal{A}_s) \triangleq \text{ReLu}(\hat{\mathcal{A}}_s \mathcal{X}_s W_e), \quad (5)$$

where  $E_s \in \mathbb{R}^{\alpha \times d'}$ ,  $d' \ll d$ ,  $W_e$  denotes the encoder trainable parameters, and ReLu is the Rectified Linear Unit.  $\hat{\mathcal{A}}_s = \tilde{D}_s^{-\frac{1}{2}} \tilde{\mathcal{A}}_s \tilde{D}_s^{-\frac{1}{2}}$ ,  $\tilde{\mathcal{A}}_s = \mathcal{A}_s + I$  is the adjacency matrix with self-loop,  $\tilde{D}_s$  is the diagonal node degree matrix of the sub-graph where its  $i^{\text{th}}$  diagonal element is equal to  $\sum_j \tilde{a}_{s_{ij}}$ .  $I$  is the identity matrix. For more details, see (Zhang et al. 2021).

We also embed the target node to the lower dimensional space defined by the encoder. Since a single node  $v_i$  has no adjacency matrix, we define its embedding as below:

$$e_i = \text{ReLu}(x_i W_e). \quad (6)$$

In summary, for each constructed graph, we feed three sub-graphs and one target node to the  $\text{GNN}_e$  to create their embeddings in the lower dimensional space, as is illustrated in the output of  $\text{GNN}_e$  in Figure 1 where  $d'$  is set to 3.

**Contrastive Learning Module:** To learn the local topological structure of the graphs, we propose to employ contrastive learning module. In order to create positive and negative *pairs* to be used for the purpose of contrastive learning, we first take an average over all the  $\alpha$  rows of  $E_s$  to map it to a  $d'$  dimensional vector, called  $r_s \in \mathbb{R}^{1 \times d'}$ .

Since we have formed two positive and one negative sub-graphs for each target node, we define two positive and one negative pairs. The positive pairs are  $\mathcal{P}_1^+ = (e_t, r_{s_1}^+)$  and  $\mathcal{P}_2^+ = (e_t, r_{s_2}^+)$ , and the negative pair is  $\mathcal{P}_1^- = (e_t, r_{s_1}^-)$ . Embedding of the target node  $v_t$  is denoted by  $e_t$ ,  $r_{s_{\{1,2\}}}^+$  indicates the  $d'$  dimensional vector obtained by taking an average of the  $E_{s_{\{1,2\}}}^+ = \text{GNN}_e(\mathcal{G}_{s_{\{1,2\}}}^+)$ , and  $r_{s_1}^-$  indicates the  $d'$  dimensional vector obtained by taking average of the  $E_{s_1}^- = \text{GNN}_e(\mathcal{G}_{s_1}^-)$ .

We introduce a trainable scoring matrix  $W_s \in \mathbb{R}^{d' \times d'}$  where the similarity of  $e_t$  to  $r_{s_1}^+$  is obtained as below:

$$\text{Sim}_{t,1}^+ = \sigma(e_t W_s r_{s_1}^{+\top}), \quad (7)$$

where  $\sigma(\cdot)$  is the logistic function that outputs a value between 0 and 1 and  $(\cdot)^\top$  denotes the transpose operator. Similarly,  $\text{Sim}_{t,2}^+$  and  $\text{Sim}_{t,1}^-$  can be obtained by replacing  $r_{s_1}^+$  in (7) with  $r_{s_2}^+$  and  $r_{s_1}^-$ , respectively.

Finally, inspired by (Zheng et al. 2021), we define the contrastive loss function,  $\mathcal{L}_{con}$  as below:

$$\mathcal{L}_{con} = -\frac{1}{2n|\mathcal{B}|} \sum_{q=1}^2 \sum_{t=1}^n \left( \log(\text{Sim}_{t,q}^+) + \log(1 - \text{Sim}_{t,1}^-) \right), \quad (8)$$

where  $|\mathcal{B}|$  indicates the number of EEG clips in the mini-batch  $\mathcal{B}$ , and  $n$  denotes the number of all existing EEG electrodes (e.g.,  $n$  is 19 in a 10-20 EEG cap recording system).

**Generative Learning Module:** This module is to learn the contextual information hidden in the graph through reconstructing the target node anonymized in the positive sub-graphs  $\mathcal{G}_1^+$  and  $\mathcal{G}_2^+$ , using the other node features and edges of the sub-graph. More specifically, we reconstruct the input sub-graph  $\mathcal{G}_{s_{\{1,2\}}}^+$  using its embedded version  $E_{s_{\{1,2\}}}^+$ , i.e.  $\hat{\mathcal{G}}_{s_{\{1,2\}}}^+ = \text{GNN}_d(E_{s_{\{1,2\}}}^+)$  while minimizing the below reconstruction loss:

$$\mathcal{L}_{rec} = \frac{1}{2n|\mathcal{B}|} \sum_{q=1}^2 \sum_{t=1}^n \|\hat{x}_{t,q} - x_t\|^2, \quad (9)$$

where  $\hat{x}_{t,q}$  is the reconstructed vector of the target node  $x_t$ , taken from the reconstructed graph  $\hat{\mathcal{G}}_{s_q}^+ = \text{GNN}_d(E_{s_q}^+)$ . We denote the parameters of the decoder as  $W_d$ .

Finally, we define the total loss  $\mathcal{L}$  through weighted sum of the reconstruction and contrastive losses, as below:

$$\mathcal{L} = \lambda \mathcal{L}_{con} + (1 - \lambda) \mathcal{L}_{rec}, \quad (10)$$

where  $\lambda$  is the balancing hyperparameter.

## Anomaly Score

From the contrastive and generative modules, for a node  $v_i \in \mathcal{V}$ , we define the anomaly scoring function as below:

$$f(v_i) = \lambda f_{con}(v_i) + (1 - \lambda) f_{rec}(v_i), \quad (11)$$

where  $f_{con}(v_i) = \frac{1}{2} \sum_{q=1}^2 \delta(\text{Sim}_{i,1}^- - \text{Sim}_{i,q}^+)$  is the contrastive anomaly score, and  $f_{rec}(v_i) = \frac{1}{2} \sum_{q=1}^2 \delta(\|\hat{x}_{i,q} - x_i\|^2)$  is the reconstruction anomaly score, and  $\delta$  is the Min-MaxScaler function (Bisong 2019) that scales the scores to the range  $[0, 1]$ .

The anomaly score for all nodes is calculated. If  $v_i$  is an anomaly node from structural point of view, both  $\text{Sim}_{i,q}^+$ ,  $q \in \{1, 2\}$  and  $\text{Sim}_{i,1}^-$  would be close to 0.5 as there exists a mismatch between  $v_i$  and the corresponding node in sub-graphs  $\mathcal{G}_{s_q}^+, \mathcal{G}_{s_1}^-$  used for training the algorithm networks, because we only use normal clips and nodes in the training phase. Here, the scaling function  $\delta$  maps 0 to 1, hence  $f_{con}(v_i)$  would be equal to 1. If  $v_i$  is an anomaly node from contextual point of view,  $f_{rec}(v_i)$  would be close to 1 as the decoder cannot reconstruct an anomaly since the networks are trained on only normal nodes and clips. If  $v_i$  is a normal node, both of  $f_{con}(v_i)$  and  $f_{rec}(v_i)$  would be close to 0. In other words, an anomaly (normal) node would result into a high (low) anomaly score.

In the inference phase, a node is considered as an anomaly if its corresponding anomaly score is higher than a threshold  $\tau_1$ . An EEG clip is considered as seizure-free (i.e. normal) if there are less than  $\tau_2$  abnormal nodes within its corresponding EEG graph, otherwise it is detected as a seizure clip. The inference phase for a seizure clip is illustrated in Figure 1 at block (4).

Data	Patients (% SZ)	EEG Files (% SZ)	EEG Clips (% SZ)
Train <sup>Supervised</sup>	591 (34.0%)	4,599 (18.9%)	38,613 (9.3%)
Train <sup>EEG-CGS</sup>	493 (0%)	4,028 (0%)	35,019 (0%)
Test	45 (77.8%)	900 (25.6%)	8,848 (14.7%)

Table 1: Train and test sets of TUSZ used in the supervised method and EEG-CGS. The percentages of the seizure data (SZ) is indicated in parenthesis.

## Experiments

We use the public benchmark dataset, Temple University Hospital EEG Seizure Corpus (TUSZ) v1.5.2 (Shah et al. 2018), which contains the largest number of seizure categories and patient samples. TUSZ is recorded over years, and recorded by different equipment generations from subjects of all ages. Thus, the variance is much more than other available EEG datasets in the seizure study, hence it is the most challenging dataset for seizure detection. The number of existing channels is 19 recorded in the standard EEG 10-20 system, as is shown in Figure 1. Table 1 summarizes the TUSZ dataset used in our experiments. As is shown in the table, in the training phase, we use the same number of normal clips as is used in other supervised methods, and do not use any seizure clips. In the test phase, the same number of test clips, including seizure and normal clips, is used in our comparison supervised methods and our method. To evaluate the model’s ability in seizure location detection, we use the available annotations on focal and generalized seizure types from 23 distinct patients. Note that compared with other types of seizures, focal and generalized types are most present in epilepsy patients.

The performance of our proposed method is quantified using five criteria including Receiver Operating Characteristics - Area Under the Curve (ROC-AUC), Precision (Pre.), F1 score (F1), Sensitivity (Sen.) and Specificity (Spec.).

We compare our proposed EEG-CGS with two streams of DL-based methods<sup>1</sup>: (1) DL models in the EEG time-series and/or spectrograms domain, including EEGNet (Lawhern et al. 2018), EEG-TL (Khalkhali et al. 2021), Dense-CNN, LSTM and CNN-LSTM (Tang et al. 2021); and (2) DL models in the EEG graph domain (Tang et al. 2021). Unlike our method, these DL models make use of the seizure data and their corresponding labels in the training phase. For a fair comparison, we test the methods on the same test data, whilst our method is trained on a much smaller number of training sample because we do not include any seizure data in the training phase.

We introduce four variations of our EEG-CGS method based on the type of constructed graph used as the input. In the following, we refer to these variations as EEG<sub>d</sub>-CGS, EEG<sub>r</sub>-CGS, EEG<sub>c</sub>-CGS, EEG<sub>f</sub>-CGS and EEG<sub>t</sub>-CGS, respectively corresponding to the cases where the employed input graph is Dist-EEG-Graph, Rand-EEG-Graph, Corr-EEG-Graph, DTF-EEG-Graph and the case where all the

<sup>1</sup>To provide a fair comparison, we do not include algorithms that require training their model on an extra EEG dataset (using transfer learning), in addition to the given TUSZ data.

Graph Type	Pre	F1	Sen	Spec
EEG <sub>d</sub> -CGS	0.846	0.771	0.688	0.945
EEG <sub>r</sub> -CGS	0.877	0.881	0.887	0.979
EEG <sub>c</sub> -CGS	0.861	0.843	0.825	0.953
EEG <sub>f</sub> -CGS	0.844	0.832	0.822	0.967
EEG <sub>t</sub> -CGS	<b>0.901</b>	<b>0.883</b>	<b>0.920</b>	<b>0.989</b>

Table 2: Synthetic anomalous node detection performance. The best performance for each metric is shown in bold.

four graph types are concatenated and fed to the system as the input representing the given input EEG clip. In our experiments, following (Tang et al. 2021), we simply use the Fast Fourier Transform coefficients as attributes of EEG channels, although adding other attributes such as Wavelet coefficients might also be useful, which is out of the scope of this paper.

The hyperparameters  $d, k, m, d', \alpha, \lambda, \tau_1, \tau_2$  of the proposed method are respectively set to 6000, 0.9, 3, 256, 4, 0.6, 0.4, 1. To have a fair comparison, all hyperparameters of the proposed method are fixed over all experiments. Regarding the EEG<sub>t</sub>-CGS case, we also set the hyperparameters to be the same as the values mentioned above; where at the inference phase, we take an average of anomaly scores computed for all nodes of the four graph types – a node is predicted as an anomaly if the average score is greater than  $\tau_1$ . Based on the average anomaly scores for all nodes, an EEG clip is predicted as seizure if the number of predicted anomalous nodes (over all clips) is greater than  $\tau_2$ . Hyperparameters of our comparison methods are set to the values suggested by their authors.

The performance of our proposed method is demonstrated for anomalous channel detection, seizure clip detection and seizure channel detection.

### Detection of Synthetic Anomalous Channels

This section is to verify the performance of the proposed method for reliable anomalous channel detection. To this end, we create a synthetic test set using the normal clips used in the training phase. More specifically, we first average every 35 normal clips (with no overlap). We then, with probability of 3%, inject an anomaly node to the average clip. A node of the selected average clip will be corrupted with a probability of 0.03%, and no more than one node is corrupted per average clip. A node is corrupted both contextually and structurally. We perform two types of corruptions to make  $v_i$  abnormal: (1) we connect  $v_i$  to all other nodes in the average clip – i.e. we set  $a_{ij} = 1$  for  $j = 1, \dots, n$ ; (2) we corrupt the attribute vector of  $v_i$  by replacing its feature vector with the feature vector of the node in the average clip that has the largest Euclidean distance to  $v_i$ . We then feed the average clips (some have injected anomalies) to the EEG-CGS networks trained on the pure normal clips and let the trained system compute the anomaly score for all channels. If the anomaly score is greater than the pre-defined threshold  $\tau_1$ , the corresponding channels are marked as anomalies.

(Liu et al. 2021a) demonstrated that, for node anomaly detection in the test phase, it would be more effective to create

Approach	Method	F1	Sen	Spec
Supervised	EEGNet	0.474	0.299	0.902
	EEG-TL	0.420	NA	NA
	Dense-CNN	0.404	0.451	0.869
	LSTM	0.365	0.463	0.814
	CNN-LSTM	0.330	0.363	0.857
	<i>Dist</i> -DCRNN	0.341	0.326	0.932
	<i>Corr</i> -DCRNN	0.448	0.457	0.900
Self-Supervised (ours)	EEG <sub>d</sub> -CGS	0.487	0.481	0.932
	EEG <sub>r</sub> -CGS	0.496	0.465	0.942
	EEG <sub>c</sub> -CGS	0.521 <sup>†</sup>	0.497 <sup>†</sup>	0.942
	EEG <sub>f</sub> -CGS	0.516	0.474	0.952 <sup>†</sup>
	EEG <sub>t</sub> -CGS	<b>0.534</b>	<b>0.501</b>	<b>0.974</b>

Table 3: Seizure clips detection. The best and the second-best performances for each metric are denoted in bold and <sup>†</sup>.

Graph Type	Pre	F1	Sen	Spec
EEG <sub>d</sub> -CGS	0.668	0.474	0.367	0.656
EEG <sub>r</sub> -CGS	0.631	0.334	0.327	0.748
EEG <sub>c</sub> -CGS	0.667	0.433	0.424	0.745
EEG <sub>f</sub> -CGS	0.659	0.394	0.378	0.735
EEG <sub>t</sub> -CGS	<b>0.704</b>	<b>0.551</b>	<b>0.432</b>	<b>0.775</b>

Table 4: Seizure channels detection performance. The best performance for each metric is shown in bold.

multiple contrastive pairs to capture more neighbors of the target node. (Note that in the training phase, we created only two contrastive pairs.) As such, following (Liu et al. 2021a), we compute the anomaly score for 80 randomly selected sub-graphs that include both close and far nodes around the target node. The final anomaly score of a target node is then the average of the 80 calculated ones. This strategy makes the final anomaly score more reliable. Our experiments show high detection performance can be obtained for a wide range of contrastive pair numbers.

Node anomaly detection results for the four variations of our EEG-CGS method are reported in Table 2. The results demonstrate the effectiveness of our unsupervised method for anomalous EEG channel detection. Note that this is the only existing study for anomalous EEG channel detection with access to no abnormal data. The reported results, confirm the effectiveness of the reconstructed graphs and that including all graphs for detection results in a more consistent performance across all metrics.

### Detection of Seizure Clips and Channels

The performance of the proposed method along with our comparison methods, for seizure clip detection, are shown in Table 3. Note that the proposed method is a self-supervised 1-class classification technique while all the comparison methods are supervised 2-class classification methods. The compared methods are the current state-of-the-art in the seizure detection research field. Our method significantly outperforms all the comparison methods in all metrics and set a new SOTA in the field though it is trained with only normal data compared to the other methods that are trained with much more data including both normal and abnormal.

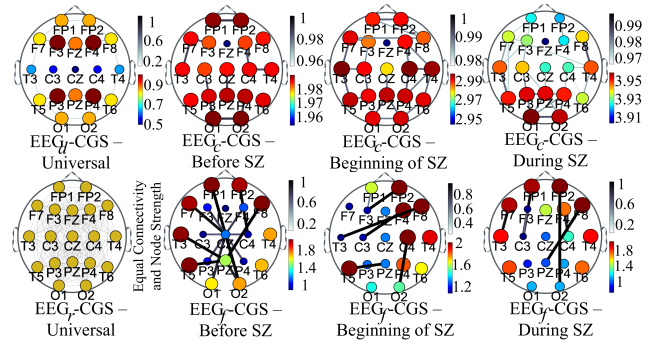


Figure 2: Visualization of EEG graphs before, beginning and during a typical seizure (SZ) incident. The top and bottom bars in each graph indicate the level of connectivity and node strength, respectively. Darker colors indicate higher values.

To show the effectiveness of our method for anomalous channel detection, in a seizure test clip, we mark the channel with an anomaly score greater than  $\tau_1$  as an anomaly and compare them with the ground-truth of seizure channels provided in the TUSZ dataset. The performance is reported in Table 4. As it can be seen, our method can effectively identify the anomalous node in a real-world dataset, as well.

Figure 3 visualizes the anomaly score of the EEG channels for a typical focal seizure clip and a typical generalized seizure clip. Subfigure (a) shows that, as is expected, the channels with high anomaly scores (e.g., FP2 to C4 and F8 to T6) are focused on specific brain regions that also match the ground-truth. For the generalized seizure case shown in subfigure (b), as is expected, the EEG-CGS method correctly tells us that anomalous channels are spread across the brain, not focused on a specific region as opposed to the focal seizure case shown in (a).

### Visualization of Graph's Nodes Connectivity

This section is to visualize the EEG graph level of nodes connectivity (i.e. adjacency matrix) changes when a seizure happens. At each graph, the intensity of an edge between nodes  $v_i$  and  $v_j$  depends on the value of  $a_{ij}$  in the graph's adjacency matrix. This figure also visualizes the level of node strength, where the level of strength at node  $v_i$  is defined as  $\sum_{j=1, j \neq i}^n a_{ij}$ , i.e. the summation of all connectivities per node – if a node is connected to channels with high adjacency values, a high level of strength is assigned to the node.

As is discussed in the EEG Graph Construction section, the adjacency matrix of Dist-EEG and Rand-EEG graphs do not change when a seizure happens since they are defined based on the Euclidean distance of node locations and independent from the node attributes. However, the connectivities (i.e. adjacency matrices) vary in the Corr-EEG and DTF-EEG graphs when a seizure happens. These changes are noticeable when one compares the node connectivities and level of strengths before, at the onset, and during a seizure incident. As is shown in Figure 2, before a seizure incident (i.e. in normal brain activity), in the Corr-EEG-Graph, all nodes have almost similar level of strengths and every

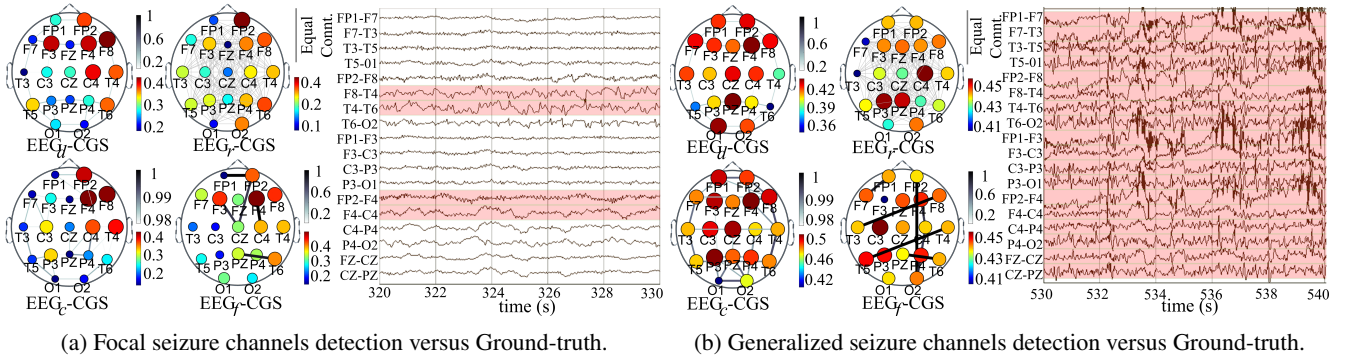


Figure 3: Visualization of seizure hannels detection for (a) Focal- and (b) Generalized seizures. Ground-truth channels are shaded in pink. Each EEG graph shows the level of the node connectivity (top bar) and node anomaly score (bottom bar). A darker color indicates higher node connectivity and anomaly score.

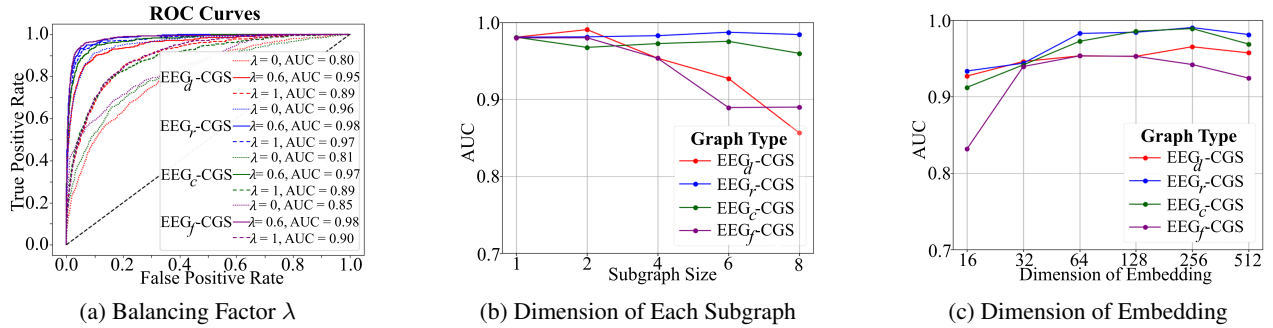


Figure 4: Hyper-parameter sensitivity for synthetic anomalous node detection.

node is connected to almost similar number of other nodes. This uniform structure changes when the seizure starts and deepen during a seizure episode. The changes happen to the DTF-EEG graph are also visualized in the figure. As it can be seen, this graph also carries important information and its variations help in distinguishing a seizure activity from normal. This figure demonstrates the effectiveness of the proposed constructed graphs in identifying seizure channels and clips in the self-supervised way.

### Hyperparameters Sensitivity

In this section, we investigate the effect of hyperparameters  $\lambda$ ,  $\alpha$ , and  $d'$  on the performance of the proposed method.

In Figure 4a, we explore the importance of the contrastive and generative components through changing the hyperparameter  $\lambda$  in the loss function  $\mathcal{L}$  defined in (10), where  $\lambda \in \{0, 0.6, 1\}$ .  $\lambda = 0$  (1) is equivalent to removing the contrastive (generative) component of the proposed method, while setting  $\lambda$  to a moderate value such as 0.6 allows the method to include both components when training its networks. We explored the performance of  $EEG_{\{d,r,c,f\}}-CGS$  for these three values. The results show that for all method variations,  $\lambda = 0.6$  provides a better AUC compared to the extreme cases  $\lambda = 0$  and  $\lambda = 1$ .

In Figure 4b, we explore the impact of subgraph size, i.e.  $\alpha$  by selecting  $\alpha$  from the set  $\{1, 2, 4, 6, 8\}$ . We observe that more reliable performance (with high AUC value) can be

obtained for  $\alpha = 4$ , i.e. 4 nodes are included in a sub-graph.

To investigate the effect of the embedding space dimension  $d'$  on the proposed method, we run our method variations using different  $d'$  values from set  $\{16, 32, 64, 128, 256, 512\}$ . The results shown in Figure 4c demonstrate that the method is not too sensitive to this hyperparameter and a reliable and accurate performance can be obtained for a wide range of  $d'$ .

### Conclusion

This paper is the first attempt on providing a self-supervised method for anomalous EEG channel detection. The proposed EEG-CGS method, consisting of contrastive and generative components, is trained only on normal EEG data with no-access to the abnormal samples. We propose to construct EEG attributed sub-graphs to capture the local structural and contextual information. The performance of the proposed method is demonstrated on both synthetic and real-world anomalies. The proposed method sets a new state-of-the-art for seizure detection on the largest and most difficult publicly available seizure dataset. EEG-CGS can be considered as an anomalous channel detection technique that can be employed for the detection of other brain disorders. The employ of sub-graphs rather than the whole EEG graph also makes the algorithm appropriate in applications, such as coma outcome prognosis, where access to all EEG channels is not possible, e.g. due to skull fractures.

## Ethics Statement

The Temple University Hospital EEG Seizure Corpus is publicly available with full Institutional Review Board (IRB) approval (Shah et al. 2018) at [https://isip.piconepress.com/projects/tuh\\_eeg/html/downloads.shtml](https://isip.piconepress.com/projects/tuh_eeg/html/downloads.shtml). The authors have no conflicts of interest to declare. No harmful insights are observed by our proposed method described in this paper. Source code is publicly available at <https://github.com/Armanfard-Lab/EEG-CGS>.

## Acknowledgments

The authors would like to acknowledge the financial support of the Natural Sciences, Engineering Research Council of Canada (NSERC), Fonds de recherche du Quebec, and the Department of Electrical and Computer Engineering at McGill University. The authors also wish to thank Calcul Quebec and Compute Canada for providing the necessary computational resources to conduct our experiments.

## References

- Abdelhameed, A.; and Bayoumi, M. 2021. A deep learning approach for automatic seizure detection in children with epilepsy. *Frontiers in Computational Neuroscience*, 15: 650050.
- Armanfard, N.; Komeili, M.; Reilly, J. P.; and Connolly, J. F. 2018. A machine learning framework for automatic and continuous MMN detection with preliminary results for coma outcome prediction. *IEEE journal of biomedical and health informatics*, 23(4): 1794–1804.
- Beghi, E.; and Giussani, G. 2018. Aging and the epidemiology of epilepsy. *Neuroepidemiology*, 51(3-4): 216–223.
- Bisong, E. 2019. Introduction to Scikit-learn. In *Building machine learning and deep learning models on Google cloud platform*, 215–229. Springer.
- Blinowska, K. J. 2011. Review of the methods of determination of directed connectivity from multichannel data. *Medical & biological engineering & computing*, 49(5): 521–529.
- Gao, B.; Zhou, J.; Yang, Y.; Chi, J.; and Yuan, Q. 2022. Generative adversarial network and convolutional neural network-based EEG imbalanced classification model for seizure detection. *Biocybernetics and Biomedical Engineering*, 42(1): 1–15.
- Hojjati, H.; Ho, T. K. K.; and Armanfard, N. 2022. Self-Supervised Anomaly Detection: A Survey and Outlook. *arXiv preprint arXiv:2205.05173*.
- Johnson, E. L. 2019. Seizures and epilepsy. *Medical Clinics*, 103(2): 309–324.
- Khalkhali, V.; Shawki, N.; Shah, V.; Golmohammadi, M.; Obeid, I.; and Picone, J. 2021. Low Latency Real-Time Seizure Detection Using Transfer Deep Learning. In *2021 IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*, 1–7. IEEE.
- Kuhlmann, L.; Lehnertz, K.; Richardson, M. P.; Schelter, B.; and Zaveri, H. P. 2018. Seizure prediction—ready for a new era. *Nature Reviews Neurology*, 14(10): 618–630.
- Lawhern, V. J.; Solon, A. J.; Waytowich, N. R.; Gordon, S. M.; Hung, C. P.; and Lance, B. J. 2018. EEGNet: a compact convolutional neural network for EEG-based brain-computer interfaces. *Journal of neural engineering*, 15(5): 056013.
- Li, C.-L.; Sohn, K.; Yoon, J.; and Pfister, T. 2021. Cutpaste: Self-supervised learning for anomaly detection and localization. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 9664–9674.
- Liu, Y.; Li, Z.; Pan, S.; Gong, C.; Zhou, C.; and Karypis, G. 2021a. Anomaly detection on attributed networks via contrastive self-supervised learning. *IEEE transactions on neural networks and learning systems*, 33(6): 2378–2392.
- Liu, Y.; Pan, S.; Wang, Y. G.; Xiong, F.; Wang, L.; Chen, Q.; and Lee, V. C. 2021b. Anomaly detection in dynamic graphs via transformer. *IEEE Transactions on Knowledge and Data Engineering*.
- Mahajan, A.; Somaraj, K.; and Sameer, M. 2021. Adopting artificial intelligence powered ConvNet to detect epileptic seizures. In *2020 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES)*, 427–432. IEEE.
- Mathur, P.; and Chakka, V. K. 2022. Graph Signal Processing Based Cross-Subject Mental Task Classification Using Multi-Channel EEG Signals. *IEEE Sensors Journal*, 22(8): 7971–7978.
- Rashed-Al-Mahfuz, M.; Moni, M. A.; Uddin, S.; Alyami, S. A.; Summers, M. A.; and Eapen, V. 2021. A deep convolutional neural network method to detect seizures and characteristic frequencies using epileptic electroencephalogram (EEG) data. *IEEE Journal of Translational Engineering in Health and Medicine*, 9: 1–12.
- Saab, K.; Dunnmon, J.; Ré, C.; Rubin, D.; and Lee-Messer, C. 2020. Weak supervision as an efficient approach for automated seizure detection in electroencephalography. *NPJ digital medicine*, 3(1): 1–12.
- Saboksayr, S. S.; Mateos, G.; and Cetin, M. 2021. EEG-based emotion classification using graph signal processing. In *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 1065–1069. IEEE.
- Saichand, N. V.; et al. 2021. Epileptic seizure detection using novel Multilayer LSTM Discriminant Network and dynamic mode Koopman decomposition. *Biomedical Signal Processing and Control*, 68: 102723.
- Shah, V.; Von Weltin, E.; Lopez, S.; McHugh, J. R.; Veloso, L.; Golmohammadi, M.; Obeid, I.; and Picone, J. 2018. The temple university hospital seizure detection corpus. *Frontiers in neuroinformatics*, 12: 83.
- Sharma, A.; Rai, J.; and Tewari, R. 2018. Epileptic seizure anticipation and localisation of epileptogenic region using EEG signals. *Journal of Medical Engineering & Technology*, 42(3): 203–216.
- Shen, M.; Wen, P.; Song, B.; and Li, Y. 2022. An EEG based real-time epilepsy seizure detection approach using discrete wavelet transform and machine learning methods. *Biomedical Signal Processing and Control*, 77: 103820.



Shoeibi, A.; Khodatars, M.; Ghassemi, N.; Jafari, M.; Moridian, P.; Alizadehsani, R.; Panahiazar, M.; Khozeimeh, F.; Zare, A.; Hosseini-Nejad, H.; et al. 2021. Epileptic seizures detection using deep learning techniques: a review. *International Journal of Environmental Research and Public Health*, 18(11): 5780.

Surges, R.; Shmuelly, S.; Dietze, C.; Ryvlin, P.; and Thijs, R. D. 2021. Identifying patients with epilepsy at high risk of cardiac death: signs, risk factors and initial management of high risk of cardiac death. *Epileptic Disorders*, 23(1): 17–39.

Tang, S.; Dunnmon, J.; Saab, K. K.; Zhang, X.; Huang, Q.; Dubost, F.; Rubin, D.; and Lee-Messer, C. 2021. Self-Supervised Graph Neural Networks for Improved Electroencephalographic Seizure Analysis. In *International Conference on Learning Representations*.

Thuwajit, P.; Rangpong, P.; Sawangjai, P.; Autthasan, P.; Chaisaen, R.; Banluesombatkul, N.; Boonchit, P.; Tatsaringkansakul, N.; Sudhawiyangkul, T.; and Wilaiprasitporn, T. 2021. EEGWaveNet: Multiscale CNN-Based Spatiotemporal Feature Extraction for EEG Seizure Detection. *IEEE Transactions on Industrial Informatics*, 18(8): 5547–5557.

Tong, H.; Faloutsos, C.; and Pan, J.-Y. 2006. Fast random walk with restart and its applications. In *Sixth international conference on data mining (ICDM'06)*, 613–622. IEEE.

Wagh, N.; and Varatharajah, Y. 2020. Eeg-gcnn: Augmenting electroencephalogram-based neurological disease diagnosis using a domain-guided graph convolutional neural network. In *Machine Learning for Health*, 367–378. PMLR.

Wang, J.; Liang, S.; Wang, Y.; Zhang, Y.; He, D.; Ma, J.; Ruan, C.; Wu, Y.; Hong, X.; and Shen, J. 2020. A weighted overlook graph representation of eeg data for absence epilepsy detection. In *2020 IEEE International Conference on Data Mining (ICDM)*, 581–590. IEEE.

Xie, Y.; Xu, Z.; Zhang, J.; Wang, Z.; and Ji, S. 2022. Self-supervised learning of graph neural networks: A unified review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.

Yan, T.; Dong, X.; Mu, N.; Liu, T.; Chen, D.; Deng, L.; Wang, C.; and Zhao, L. 2018. Positive classification advantage: tracing the time course based on brain oscillation. *Frontiers in human neuroscience*, 11: 659.

Zeng, J.; and Xie, P. 2021. Contrastive self-supervised learning for graph classification. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 10824–10832.

Zhang, H.; Lu, G.; Zhan, M.; and Zhang, B. 2021. Semi-supervised classification of graph convolutional networks with Laplacian rank constraints. *Neural Processing Letters*, 1–12.

Zheng, Y.; Jin, M.; Liu, Y.; Chi, L.; Phan, K. T.; and Chen, Y.-P. P. 2021. Generative and contrastive self-supervised learning for graph anomaly detection. *IEEE Transactions on Knowledge and Data Engineering*.