A Local-Global Graph Convolutional Network for Depression Recognition using EEG Signals

Yu Chen[1], Xiuxiu Hu[2], Lihua Xia[3]*
College of Computer and Control Engineering, Northeast Forestry University, Harbin, China[1, 2]
TaiZhou University, TaiZhou, China[3]

***Abstract*—Graph Convolutional Networks (GCNs) have shown remarkable capabilities in learning the topological relationships among electroencephalogram (EEG) channels for recognizing depression. However, existing GCN methods often focus on a single spatial pattern, disregarding the relevant connectivity of local functional regions and neglecting the data dependency of the original EEG data. To address these limitations, we introduce the Local-Global GCN (LG-GCN), a novel GCN inspired by brain science research, which learns the local-global graph representation of EEG. Our approach leverages discriminative features extracted from EEG signals as auxiliary information to capture dynamic multi-level spatial information between EEG channels. Specifically, the representation learning of the topological space in brain regions comprises two graphs: one for exploring augmentation information in local functional regions and another for extracting global dynamic information. The aggregation of multiple graphs enables the GCN to acquire more robust features. Additionally, we develop an Information Enhancement Module (IEM) to capture multi-dimensional fused features. Extensive experiments conducted on public datasets demonstrate that our proposed method surpasses state-of-the-art (SOTA) models, achieving an impressive accuracy of 99.30% in depression recognition.**

***Keywords*—Electroencephalogram; depression recognition; Local-Global Graph Convolutional Network (LG-GCN); multilevel spatial information; brain regions; multiple graphs**

I. Introduction

Depression is a prevalent mental disorder affecting a large population worldwide. According to the World Health Organization (WHO), more than 350 million individuals globally suffer from depression [1]. It is characterized by significant mental impairment and negative emotions, including feelings of sadness, fatigue, and hopelessness. Currently, the primary method for diagnosing depression relies on doctor-patient communication. However, factors such as patient subjectivity, low sensitivity, and denial pose significant challenges to the diagnostic process [2]. Therefore, there is a pressing need for an objective and accurate method for detecting depression.

In recent years, electroencephalography (EEG) has emerged as a widely adopted technique for classifying depression due to its advantages, such as high temporal resolution, low acquisition cost, and ease of operation and recording [3]. It has become a commonly used and effective tool for assessing brain function. Previous approaches have involved models based on recurrent neural networks (RNNs) and convolutional neural networks (CNNs), which analyze EEG signals in the time-frequency domain, extracting features independently from individual channels. However, research has demonstrated that the brainwave patterns of individuals with depression arise from interactions between multiple channels, and EEG electrodes are positioned in a spherical space. Traditional CNNs face limitations in handling irregular and non-Euclidean data. In contrast, graphs, which are effective in handling irregular data, are better suited for modeling signals in a three-dimensional spherical space. In this context, each electrode can be seen as a node in the graph, and the spatial relationships or correlations between electrodes can be represented as edges. Graph Neural Networks (GNNs) leverage the adjacency relationships among nodes to jointly learn the spatial patterns of EEG signals [4].

Integrating prior knowledge derived from neuro-psychological research into the design of Graph Neural Networks (GNNs) presents significant potential for decoding psychological states from EEG signals. In the case of depression, it has been observed that the manifestation of this condition in individuals may involve interactions within specific brain regions [5]. Activation of a particular brain area can also trigger simultaneous activation in other regions. Research has indicated the presence of high-level connections between electrodes on the left and right hemispheres, offering additional insights for biomedical analysis [6]. Ding et al. [7] have demonstrated that many cognitive functions rely on the cooperation between different brain regions rather than being confined to a specific area. While previous research [8] utilized a globally connected adjacency matrix with learnable connections, it failed to consider the local activities within each functional region. On the other hand, the Regularized Graph Neural Network (RGNN) [9] established local connections based on spatial distances between electrodes but struggled to effectively capture the complex relationships between functional regions. Thus, it is crucial to strengthen the connectivity within local functional regions while establishing connection patterns between channels based on global dynamics. Additionally, relying solely on spatial patterns would overlook important discriminative features present in the original EEG data, which are vital for identifying depression in EEG signals. Consequently, our research focuses on appropriately constructing the brain topology based on EEG data and addressing challenges related to information loss.

To tackle the aforementioned issues, we propose an algorithm for depression recognition based on a local-global graph convolutional network (LG-GCN). LG-GCN constructs

*Corresponding Author.

both a global dynamic graph and a local functional graph to explore the multi-level spatial information across EEG channels. To capture adaptive multi-dimensional features of EEG, we introduce an Information Enhancement Module (IEM) that incorporates data dependencies and cleverly integrates them with spatial information. Inspired by brain science research [10], LG-GCN incorporates topological information from both local and global perspectives in the graph convolutional layers. Furthermore, recognizing that the graph structure may not extract all discriminative features from the original signals, we adopt Gated Convolutional Networks to capture the dependency relationships between raw temporal data and advanced features, thereby enhancing the model's performance. Finally, to assess the effectiveness of incorporating prior knowledge into the LG-GCN model, we analyze the differences between various brain region partitioning methods and conduct extensive visualization experiments.

The main contributions of this study are as follows:

1) Introducing a local-global multi-graph fusion framework that explores multi-level features of graph topology spaces. This framework overcomes the limitation of insufficient extraction of depression-related information in a single encoding path by utilizing an information enhancement module to adaptively supplement multi-dimensional fused features.

2) Proposing an adaptive global dynamic graph and local functional connectivity graph that are integrated into the global graph convolutional network. This integration allows for the incorporation of local information, capturing the multi-level spatial dependencies in EEG data. It effectively associates the spatial distribution of EEG signal channels with depth-encoded depression features, resulting in improved classification performance.

3) Validating the effectiveness of the LG-GCN framework on a public dataset, where it achieves state-of-the-art (SOTA) classification performance. Additionally, multiple ablation experiments are conducted, and LG-GCN is compared with other methods, further confirming its efficacy.

The remaining structure of the article is outlined as follows: Section II briefly reviews related works. The proposed framework will be described in detail in Section III. Section IV provides extensive experiments and analyses. Section V provides a discussion of the results obtained, and finally, the study is concluded in Section VI.

II. Related Work

A significant number of previous studies have attempted to use machine learning methods to detect depression. Saeedi et al. [11] used a genetic algorithm to select significant features from linear and nonlinear features and employed an enhanced K-nearest neighbor algorithm to classify the EEG signals of depressed patients. Mumtaz et al. [12] performed time-frequency decomposition of EEG signals based on wavelet transform to construct feature matrices, which were then inputted into an LR classifier. Although machine learning has achieved significant success in classification tasks, it faces challenges in feature selection complexity and high accuracy performance requirements. Therefore, deep learning methods are gradually becoming popular among researchers. Kang et al. [13] proposed transforming the asymmetry of EEG into a matrix image as input to a CNN, but the calculation of electrode spatial positions in the two-dimensional image transformation is challenging. Acharya et al. [14] introduced a 13-layer CNN model based on deep learning, but it has complex training and requires a significant amount of time. Chen et al. [15] used a lightweight DCTNet model, which consists of a six-layer neural network with CNN-LSTM and achieved high classification performance. Compared to machine learning methods that rely on manual feature extraction, deep learning algorithms have better performance in improving the accuracy of depression diagnosis. However, most current deep learning methods focus on mining information in the time-frequency domain, while the spatial distribution information of channels is an important but often overlooked feature.

In recent years, GNN models have shown promising performance in handling graph-structured data, providing new research directions for the development of EEG signal analysis. Zhang et al. [16] proposed a depression recognition framework based on the fusion of time-space ubiquitous EEG features at the feature level. The model selected 19 optimal time EEG features and four optimal spatial metrics features and analyzed the intrinsic connections among EEG signal channels. Zhu et al. [17] modeled correlation using graphs and proposed an attention based GCN that achieved a recognition rate of 96.50% for depression. Wu et al. [18] combined space-time graph convolutional networks (ST-GCN) with a depression-related functional connectivity graph in a deep learning approach, improving the ability to represent spatiotemporal features. Sun et al. [19] used a complex network graph convolutional neural network (CN-GCN) with a multi-branch structure to explore deep information between channels, achieving a recognition rate of 99.29% for depression. However, these models only used one type of functional connectivity in the brain. In this article, the proposed model considered both structural and functional connectivity for the first time. Jang et al. [20] suggested that individuals with depression tend to exhibit greater relative right frontal lobe activity. Mumtaz et al. [21] found that individuals with depression exhibit greater activity in the front EEG. Therefore, in depression research, considering both global brain connectivity and local functional connectivity is critical while constructing multi-channel EEG signal dependency relationships. The proposed LG-GCN model in this paper analyzes the spatial relationships within and between different functional areas, extracting more comprehensive spatial features for more effective depression identification.

III. Methodology

To fully exploit the potential of EEG data, LG-GCN is proposed for EEG-based depression recognition, which consists of a local-global graph convolution module (LG-GCM) and IEM. The overall framework is shown in Fig. 1, which can be divided into the following steps:

1) Considering the frequent local connections in the brain network and the need for better differentiation of brain regions in relation to depression, we construct local regions based on prior research, each region's information are aggregated as nodes in the local functional graph, and node similarity are computed to represent functional connectivity.

2) The brain network adopts an adaptive adjacency matrix to dynamically calculate the connection strength between channel nodes, dynamically reflecting the information under depression from a global perspective and introducing a GCN to aggregate local-global features.

3) The IEM is used to further extract information from the original signal, where the weighted averaging fusion mechanism is used to extract high-dimensional features outputted by the GCN, and then inputted into an adaptive attention fusion network to obtain complementary information. Next, this article will present detailed information related to this.

A. Local-Global Graph Convolution Module

The results of pattern analysis of neuroimaging data indicate that the depression group exhibits distinct patterns related to emotional categories in the activity of the distributed neural system spanning across cortical and subcortical regions [22]. Therefore, to simulate the interconnections between different brain regions, we propose a multi-graph representation, namely LG-GCM. By considering spatial information from different perspectives, the adaptive global dynamic graph can dynamically integrate EEG channel information, while the local functional connectivity graph can characterize static relationships between different brain regions. Finally, multi-layer GCN is introduced to capture multi-graph and multi-level representation features by adaptively fusing local and global spaces.

1) Local functional graph: Different regions of the human brain cortex are highly connected and concentrated. Based on prior knowledge from the field of neuroscience [23], we construct two types of local brain regions by aggregating local EEG features to capture more robust features. Fig. 2(a) shows the first partition based on the reference brain cortex anatomy. The brain cortex is typically divided into four lobes: the frontal, parietal, temporal, and occipital lobes, each of which is responsible for different tasks. Due to the importance of electrodes located in the frontal lobe (FP1, FP2, F7, F3, FZ, F4, F8) in depression detection tasks [20], we divide the regions in more detail to meet the needs of structural and functional connectivity. Fig. 2(b) shows another effective method based on the division of regions between the brain's two hemispheres.

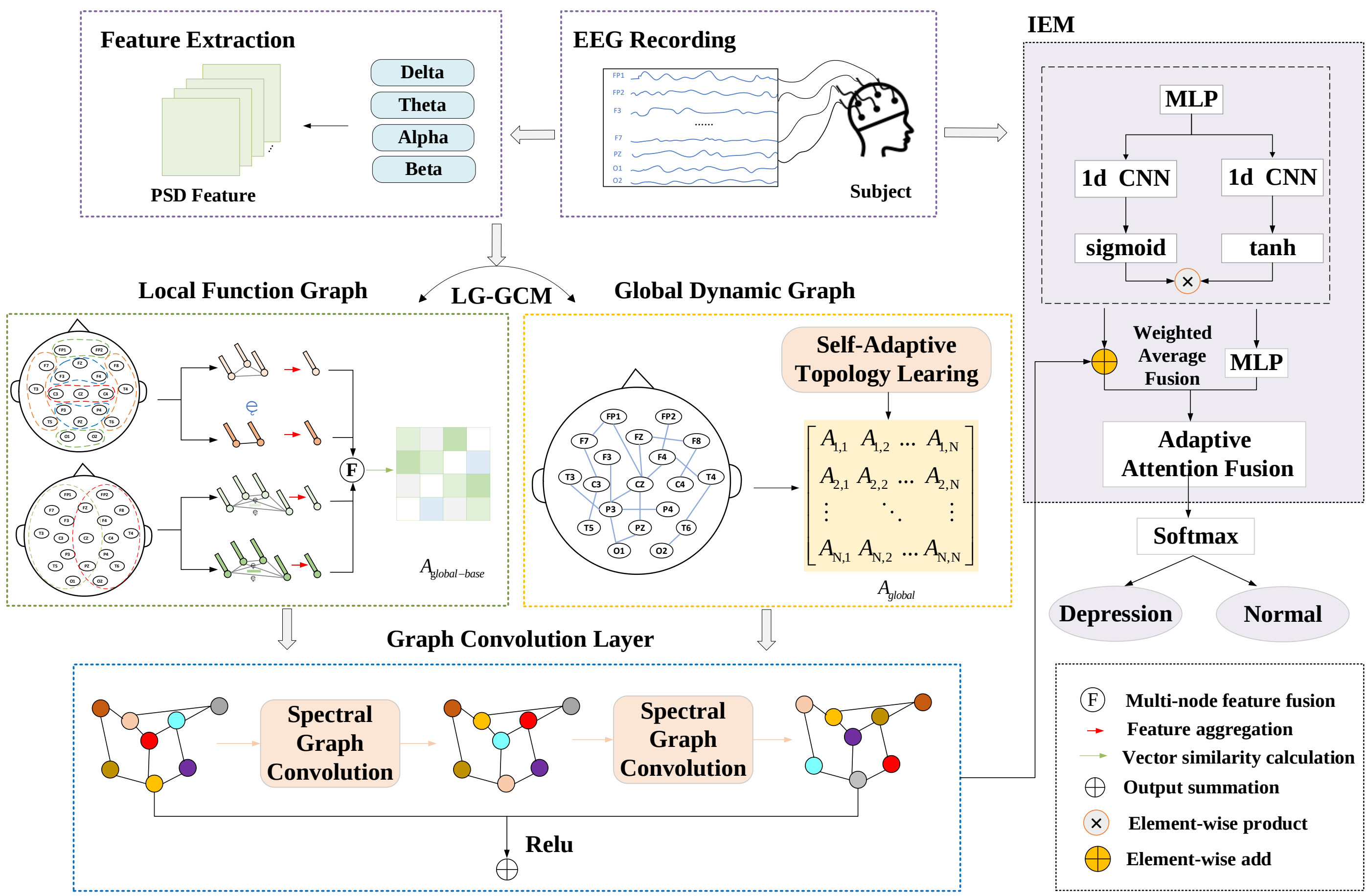


Fig. 1. Illustration of the proposed model LG-GCN.

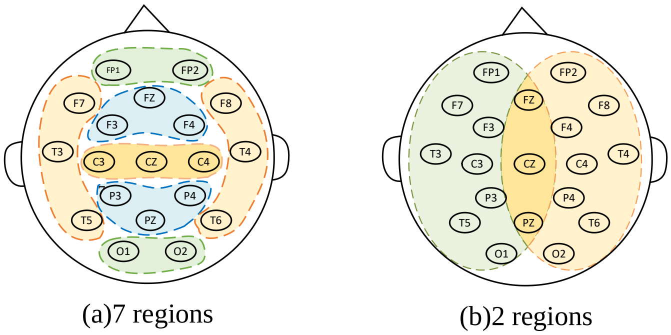


Fig. 2. Two different region partition methods for graph coarsening.

We represent multiple graphs as weighted undirected graph G= (V, ε, A), where V is the set of nodes, ε represents the connections between nodes, the adjacency matrix $A \in R^{N\times N}$represents the connectivity between nodes, and N is the number of EEG channels. The region functional graph containing P subgraphs is represented as $G_R^{(1)}, G_R^{(2)} = [G_{R_1}, G_{R_2}, \cdots, G_{R_P}]$, For kth region functional subgraph G_{R_k}, the input data is denoted as $G_{R_k} \in R^{S\times n_k\times D}$, where S represents the number of samples, n_k is the number of channels in the kth subgraph, and D represents the frequency band, including δ band(0.5~4 Hz), θ band(4~8 Hz), α band(8~13 Hz), and β band(13~30 Hz).These frequency bands have been used in previous studies to investigate differences between patients with depression and healthy controls [24]. In this work, PSD features in the frequency domain of the EEG signals are extracted using the Welch method [19] for each channel node, which can be obtained from the MATLAB software with default parameters. Owing to EEG original data containing noise and redundant information, singular value decomposition (SVD) is commonly used to perform dimensionality reduction and extract relevant information. It can be represented as

$$G_{R_k} = W_{R_k} \cdot G_{R_k}^{b^T} \tag{1}$$

Where W_{R_k} represents the left singular matrix, $G_{R_k}^b \in R^{N_k\times\mu}$ is the channel-level matrix containing advanced channel information, with dimensions of $n_k \times \mu$, Where n_k is the number of channels in the k-th subgraph $(\sum n_k = N)$, and μ represents the number of features for each EEG channel. Furthermore, the local subgraph can be represented as $G_{R_k} = \left[g_{R_k}^1, \cdots g_{R_k}^n, \cdots g_{R_k}^{n_k}\right]$, where $g_{R_k}^n$ is the node feature vector of the local subgraph. The aggregation function aggregates the node feature vectors in each local subgraph, which can be the maximum, minimum, average, etc. This process is known as graph coarsening. In LG-GCN, the average aggregation is selected as the aggregation function. Hence, the output of the local graph, $G_{local}^{(1)}, G_{local}^{(2)} \in R^{P\times\mu}$, can be calculated by

$$\begin{aligned} G_{local}^{(1)}, G_{local}^{(2)} &= \mathcal{F}_{aggregate}\left(\left[G_{R_1}, G_{R_2}, \cdots, G_{R_P}\right]\right) \\ &= \left[\frac{1}{N_1}\sum_{n=1}^{N_1} g_{R_k}^1, \cdots, \frac{1}{N_k}\sum_{n=1}^{N_k} g_{R_k}^{n_k}\right] \\ &= [h_{local}^1, \cdots, h_{local}^P] \end{aligned} \tag{2}$$

Where p is the index of each node in the local graph, h_{local} represents the latent representation of the local subgraph. Then, the local subgraphs are concatenated to obtain $G_{local} = concat(G_{local}^{(1)}, G_{local}^{(2)}) \in R^{V\times\mu}$, which results in the feature fusion matrix of the two partitioned local subgraphs.

The definition of the local subgraph is incorporated into the basic adjacency matrix of the brain to model the correlation between different brain regions, which is used in the subsequent graph convolutional layers to enhance the feature propagation between more important local edges. The relationships between local graphs are utilized as the edges of the basic global graph. Neuroscience research suggested that activating one specific brain region also tends to activate other regions in the group for advanced cognitive processes [25]. The dot products between local graph representations for each EEG instance are calculated to reflect the relations among local graphs. Thus, we have $A_{global-base} \in R^{V\times V}$, which is calculated as follows.

$$\begin{aligned} A_{global-base} &= \begin{bmatrix} A_{global-base}^{1,1} & \cdots & A_{global-base}^{1,V} \\ & \ddots & \\ A_{global-base}^{V,1} & \cdots & A_{global-base}^{V,V} \end{bmatrix} \\ &= \begin{bmatrix} h_{local}^1 \cdot h_{local}^1 & \cdots & h_{local}^1 \cdot h_{local}^V \\ & \ddots & \\ h_{local}^1 \cdot h_{local}^V & \cdots & h_{local}^V \cdot h_{local}^V \end{bmatrix} \end{aligned} \tag{3}$$

where $\cdot$ operation is dot product. $h_{local}^i \cdot h_{local}^j$ represents the similarity between local subgraphs as the edge weight in the adjacency matrix. $A_{global-base}^{i,j}$ represents the ith row and jth column of the basic adjacency matrix corresponding to the brain. Each local subgraph is treated as a node, forming the basic functional connectivity matrix of the brain. The specific process is as follows.

$$A_{global-base}^{i,j} = exp\left(-\frac{\left\|h_{local}^i - h_{local}^j\right\|^2}{2\sigma^2}\right) \tag{4}$$

Finally, $A_{global-base}$ forms the basic adjacency matrix of the local functional graph. The graph coarsening method aggregates the characteristics of local EEG signals, reducing redundant information and lowering computational complexity, thereby obtaining more robust features.

2) Global dynamic graph: Due to the complex relationships between functional regions of the brain, they exhibit temporal variability and high dynamics [26]. However, existing methods are often restricted by static prior knowledge or weak modification, making it difficult to fully capture the complex dynamic functional connections between brain regions. We aim to learn an adaptive brain network adjacency matrix [27], which is self-learned during the model training process to understand the global connectivity relationships between EEG channels and better understand the interactions between channels. Therefore, we define a non-negative function $F(X_m, X_n)$ to quantify the strength of functional connections between channels, where X_m and X_n respectively represent the features between any two channels.

Function F calculates the connection weight between nodes based on signal differences between channel nodes and can dynamically reflect signal patterns under depression. The formula for calculating function F is shown below.

$$A_{m,n} = F(X_m, X_n) = \frac{exp\left(Relu(W_t|X_m - X_n|)\right)}{\sum_{m=1}^{N}\sum_{n=1}^{N} exp\left((W_t|X_m - X_n|)\right)} \quad (5)$$

where W_t is a learnable parameter, and $|X_m - X_n|$ represents the distance between the features of the two channels. The activation function Relu is applied to constrain weak channel coupling and is applied to the output to ensure that $A_{m,n}$ is a non-negative element. The output of the final global dynamic layer of the brain is a non-negative adjacency matrix A.

3) Multilayer GCNs: Once the local and global graph representations are obtained, we aggregate the information from these two types of graphs and enable global information interaction among the channels of EEG signals using GCN. The global dynamic graph extracts common features of depression patterns under coarse granularity, while the dense local functional graph can better capture functional connectivity features between different regions, containing more detailed information., LG-GCN captures the spatial information of EEG channels at multiple levels and dimensions, thereby extracting topological spatial features that are most favorable for the recognition of depression in EEG.

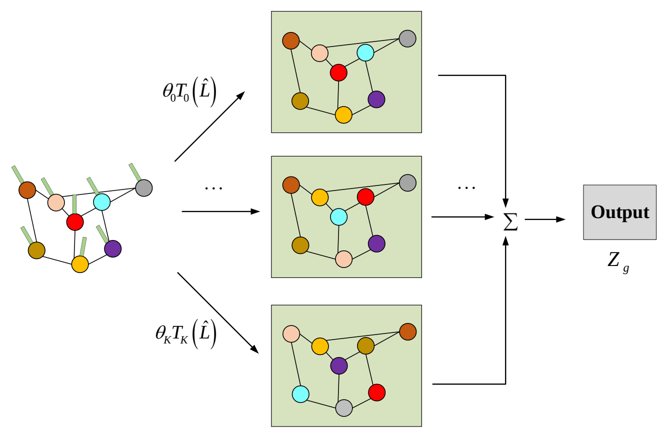


Fig. 3. The structure of graph convolution network.

In terms of node initialization, we use the value of $X^0 = X(G_{local})^T$ as the initialized graph node features of the graph, where X refers to the high-level feature set of the input signal, $X \in R^{N\times\mu}$. The notation $(G_{local})^T$ represents the aggregation relationship of the node features of the local subgraph. Then, we utilize the adjacency matrix A and $A_{global-base}$ to implement multilayer graph convolution and achieve feature propagation of local and global information of EEG signal channels. we first normalize the adjacency matrix A as shown in equation (6), where D is a diagonal matrix with $D_{ii} = \sum_{j=1}^{n} A_{ij}$, Then, we calculate the graph Laplacian matrix $L = D - A$ and perform spectral decomposition to obtain the eigenvalues and eigenvectors corresponding to L, as shown in equation (7), where $U = (\overrightarrow{u_1}, \overrightarrow{u_2}, \cdots \overrightarrow{u_n})$ and Λ is the eigenvalues and eigenvectors of L, respectively, and λ_{max} is set to 2. Assuming that the signal $x \in R^N$ undergoes Fourier transformation as $x = U\hat{x}$, we can calculate the Kth order Chebyshev polynomials $T_{k-1}(\hat{L})$ as shown in equation (8). Given X^{l-1} as the input to the l-th graph convolutional layer, the output X^l can be calculated using equation (9), where θ_k^l is the trainable parameter corresponding to the graph convolutional layer. A key step in GCN is the aggregation of node features, as illustrated in Fig. 3. As the convolutional layers iterate, node features propagate through the graph structure, allowing LG-GCN to obtain the aggregation of global information.

$$A = D^{-\frac{1}{2}}(I_N + A)D^{\frac{1}{2}} \quad (6)$$

$$\hat{L} = U\Lambda U^T = \frac{2L}{\lambda_{max}} - I_N \quad (7)$$

$$T_i(x) = 2xT_{i-1}(x) - T_{i-2}(x), T_0(x) = 1, T_1(x) = x \quad (8)$$

$$X^l = \sigma\left(\sum_{k=0}^{K-1} \theta_k^l\, T_k(\hat{L}) X^{l-1} A_{global-base}\right) \quad (9)$$

Since the multi-channel EEG graph only contains 19 nodes, to avoid the loss of important dynamic information, we do not perform graph pooling and instead directly sum up the outputs of each graph convolutional layer, integrating all node features into one graph representation $Z_g = \sum_{i=0}^{L} X^l \in R^{N\times\mu'}$, μ' is the feature dimension of the hidden layer.

B. Information Enhancement Module

LG-GCN can extract effective multi-graph and multi-level information, but it may ignore some important discriminative features in the raw EEG data. Hence, IEM is proposed to extract the dependency relationships between data and higher-level features, which are adaptively aggregated these features into the graph topology representation to capture a more comprehensive range of information within the EEG signals.

1) Feature Extractor (IEM-FE): To address the noise and high-dimensional characteristics of the raw EEG data and capture its discriminative features, we introduce a gated convolutional network as an FE for capturing enhancing information. The structure of the Gated Convolutional Networks, as shown in Fig. 1, utilizes two types of activation functions. Throughout the entire process, the computation and gate weight adjustment of the gate units are utilized to filter and update the output of the convolutional layers. This allows for the preservation of the ability to extract nonlinear features and further explores the discriminative features dependent on EEG data. Simultaneously, it enhances the network's generalization capability and robustness. Given the reshaped input data $X \in R^{1\times(N*d)}$, the operation of the gated convolution can be expressed as follows:

$$Z_i = S(\Theta_1 \cdot X + b_1) \odot T(\Theta_2 \cdot X + b_2) \quad (10)$$

where Θ_1 and Θ_2 are two 1D convolution operators, b_1 and b_2 are model parameters, $S(\cdot)$ and $T(\cdot)$ denote sigmoid and tanh functions, and $\odot$ is the element-wise multiplication. The output $\odot$ of the gate mechanism is obtained by multiplying the outputs of these two functions, which is used to represent the importance and information gain of each channel at a specific time point.

2) Weighted Average Fusion mechanism (IEM-AF): To obtain more comprehensive information from EEG data, AF mechanism is designed to fuse data-dependent auxiliary information into spatial information. Specifically, we flatten and map the output features Z_i and Z_g to a consistent dimension $\hat{z}_i, \hat{z}_g \in R^{1\times\beta}$, These features are then fused together using the following representation:

$$\hat{z}_f = \omega(\hat{z}_i, \hat{z}_g) \tag{11}$$

where $\omega(\cdot)$ represents the weighted average fusion function. Subsequently, a fully connected layer is utilized to reduce the dimensionality of the fused spatial features and auxiliary features, denoted as z_g and z_f respectively. These features are then passed through an adaptive attention fusion network to achieve effective feature fusion and extract complementary information. The final representation for EEG depression recognition is as follows:

$$z = \varphi_f \cdot z_f + \varphi_g \cdot z_g \tag{12}$$

where $z \in R^{1\times B}$, B is the binary class number for depression and health. The attention values φ_i and φ_g with embeddings z_i and z_g are self-learned by the proposed model. Specifically, we adopt a non-linear transformation and a shared matrix q to calculate the attention scores using equation (13):

$$AS_f = Q^T \cdot \tanh(W \cdot z_f^T + b) \tag{13}$$

where W and b are the transformation matrix and bias vector, respectively. Similarly, AS_g is obtained in the same way. We then normalize the attention scores AS_f and AS_g using softmax $(\cdot)$ function, yielding the attention values $\varphi_f, \varphi_g \epsilon R^{N\times 1}$, which represent the importance and relevance of the nodes. This can be expressed as:

$$\varphi_f = \frac{exp(AS_f)}{exp(AS_f)+exp(AS_g)} \tag{14}$$

Ultimately, the embeddings z are normalized by the softmax$(\cdot)$ function to calculate the predicted probabilities $\hat{y}$.

C. Optimizing LG-GCN

LG-GCN can extract effective information from EEG data and utilize this information to optimize the model. The construction of the loss function is crucial in this process. To obtain the optimal network parameters, the backpropagation (BP) algorithm is applied to iteratively update the network parameters until the best or suboptimal solution is obtained. The loss function during the training phase consists of a classification optimization term and a spatial regularization term. The classification optimization term uses the cross-entropy loss function, which is defined as:

$$\mathcal{L}_{cla} = -\frac{1}{S}\sum_{i=1}^{S}\sum_{k=0}^{K-1} y_{i,k} log_{\hat{y}_{i,k}} \tag{15}$$

Where S represents the samples and N represents the number of nodes. On the other hand, a spatial regularization term is introduced into the loss function to incorporate the smoothness and sparsity of the learned adjacency matrix, which better considers the spatial relationships among the nodes. This term can be defined as:

$$\mathcal{L}_{reg} = \frac{\lambda}{C}\sum_{i=1}^{C}\sum_{j=0}^{N}\sum_{k=0}^{N}|A_{jk}| \tag{16}$$

The total loss function consists of the above two terms, that is, $Loss = \mathcal{L}_{cla} + \mathcal{L}_{reg}$. λ represents the regularization parameter that controls the trade-off between the classification and regularization terms.

IV. Experimental Results And Analysis

In this section, we perform EEG-based depression recognition to demonstrate the effectiveness of the proposed LG-GCN model on public datasets. Then, experimental details and results are described, and finally, the ablation experiments and interpretability of EEG-based depression recognition are presented. The training procedure of the proposed model is shown in Algorithm 1.

Algorithm 1 LG-GCN Training Procedure

Require: EEG training samples S, ground-truth labels Y, training epoch e, number of partitions P, parameter K in LG-GCM, and learning rate r.
Ensure: Prediction of the model $\hat{y}$
1: Initialization;
2: Aggregate node features for each local graph by 1-3;
3: Calculate the adjacency matrix $A_{global-base}$ by 4;
4: Compute the global dynamic representation;
5: **for** j = 1 → e **do**
6: Aggregate Z_g by 6-9;
7: Compute the EEG discriminative features Z_i by 10;
8: Conduct fusion to obtain the high-dimensional features by 11;
9: Compute the final representation by 12;
10: Update graph convolution parameters;
11: **end for**
12: **return** Prediction $\hat{y}$.

A. Dataset

We evaluate the proposed EEG-based depression diagnosis method using a publicly available dataset provided by Mumtaz et al. [11]. The dataset consists of two groups of participants: 1) 34 patients with major depressive disorder (MDD) (mean age 40.33, SD = ±12.861), and 2) 30 age-matched healthy control (HC) subjects (mean age 38.227, SD = ±15.64). The study is conducted on outpatient participants at the Hospital University Sains Malaysia (HUSM) and record their EEG signals in the resting state with eyes closed (5 min) and eyes open (5 min). The experimental procedures of this study are approved by the HUSM Ethics Committee, and all participants involved in the study have a thorough understanding of the entire process and provide informed consent.

B. Preprocessing

In this study, data preprocessing uses the PyCharm software and the MNE tool. Original EEG signals are always accompanied by artifacts and noise, including electrooculogram (EOG), electromyogram (EMG), and noise related to data acquisition. A bandpass filter is applied to extract the frequency band of 0.5-30Hz before applying feature extraction techniques. The independent component analysis (ICA) algorithm is then employed to further remove redundant artifacts and noise in the EEG signals, such as blinking and eye movement. The FastICA algorithm is used in this study. After these steps, clean EEG data is obtained, providing a basis for further analysis.

Considering the inaccuracy of signal boundary values, only the middle portion of the 120s EEG data is retained for analysis. To increase the number of samples and decoding accuracy, a cropping strategy is employed to process the data [28]. First, a sample set is established using all available EC and EO files in the dataset. For each data file, the 120-second data is segmented into 120 samples using a non-overlapping sliding window of 1 second in length, resulting in 240 samples for each participant. Ultimately, there are 8,160 and 7,200 samples for MDD and HC, respectively.

C. Implementation Details

We conduct all the experiments on the platform of NVIDIA GeForce RTX 3090 GPU. The proposed model is implemented using the PyTorch framework, a popular deep learning toolkit. The number of graph convolutional layers is set to 2, and the number of gate convolutional layers is set to 1. We use Chebyshev polynomials of order K = 3 for graph convolutions. The electrodes are divided into 7 and 2 regions, respectively. During the training process, we train the model for 200 epochs with a batch size of 256 and a learning rate of 0.01. The LG-GCN model is trained using Adam optimizer [29], which implements the stochastic gradient descent algorithm to update network parameters, weights, and model biases. Adam applies biases to each node in the graph. The training set and test set are set in a ratio of 9:1. Similarly, the validation set is established using 10% of the training set.

D. Evaluating Metrics

Due to the existence of false positive and false negative samples, using only accuracy to measure the performance of classifiers is far from sufficient. In previous studies [30], [11], [13], three typical performance metrics, accuracy, sensitivity, and specificity, are used to measure the performance of classifiers. Therefore, this study still chooses these three metrics to facilitate better comparison with other studies. These metrics can be calculated according to the following formulas 17-19:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{17}$$

$$Sensitivity = \frac{TP}{TP+FN} \tag{18}$$

$$Specificity = \frac{TN}{TN+FP} \tag{19}$$

where TP is the number of MDD patients detected correctly, FN is the number of MDD patients detected as healthy individuals, TN is the number of healthy individuals detected correctly, and FP is the number of healthy individuals detected as MDD patients.

E. Classification Performance

To demonstrate the superiority of the LG-GCN algorithm in depression recognition, we carefully select the most representative baseline models by comparing different methods. One category includes traditional machine learning algorithms, namely Support Vector Machine (SVM) [11] and Multi-Layer Perceptron Neural Network (MLPNN) [31]. Another category includes the most popular deep learning algorithms in the field of depression, namely CNN [13] and MDCNN [32]. The last category includes graph-based models. Since GCN has been applied less frequently to EEG-based depression recognition tasks, this paper only lists one of them, namely CN-GCN [19]. Table I presents the performance of LG-GCN compared to these three categories of baseline models on four frequency bands and all frequency bands (using δ、θ、α and β bands together).

TABLE I. ACCURACY ON DIFFERENT FREQUENCY BANDS

Model	$\boldsymbol{\delta}$ band	$\boldsymbol{\theta}$ band	$\boldsymbol{\alpha}$ band	$\boldsymbol{\beta}$ band	all（$\boldsymbol{\delta,\theta,\alpha,\beta}$）
MLPNN	83.33	86.67	91.67	81.67	93.33
SVM	65.52	74.14	81.03	77.55	89.96
CNN	95.50	95.90	98.85	96.07	/
MDCNN	82.31	86.00	87.30	94.10	97.27
CN-GCN	78.09	80.39	78.15	96.68	99.29
Ours	84.36	88.67	91.44	96.90	**99.30**

Our proposed model demonstrates strong advantages in each frequency band. the best accuracy is achieved in the α and β frequency bands. Furthermore, the performance when using all frequency bands together is superior to using the four frequency bands individually. This finding reveal that the overall structure allows for complementation and integration of information from each frequency band, resulting in more comprehensive information for depression recognition. This validates previous works [31].

To further validate the superiority of the proposed model, Table II shows the classification performance of our proposed model compared to existing methods on the same dataset. Mumtaz et al. [12][33][21] manually extracted features, including wavelet features, synchronous likelihood (SL) features, hemispherical asymmetry features, and frequency band energy features, and fed them into traditional machine learning models (LR and SVM) for classification. Mahato et al. [31][34] used frequency band energy and asymmetry features as input features for their classifiers, but their accuracy did not exceed 95%. Similarly, Saeedi et al. [11] considered different band powers and entropies. Dang et al. [32] combined multiple frequency band brain networks with deep learning algorithms, achieving an accuracy of 97.27%. In addition, the results of CN-GCN are comparable to the results of our proposed method, as it learns node features based on the topological connections of the brain network, rather than selecting local features between nodes. Existing studies on building brain networks based on GCN methods only consider the relationship between adjacent nodes, but ignore the activation relationship between brain regions. Soni et al. [35] fused information from three channels in the frontal lobe and used the KNN algorithm to achieve an accuracy of 92.80% in detecting depression. This finding is consistent with the results of this paper and previous studies, demonstrating that the features of depression patients are closely related to the activity in the frontal lobe of the brain.

TABLE II. CLASSIFICATION PERFORMANCE OF OUR PROPOSED METHOD AND EXISTING STUDIES ON THE SAME DATASET

Existing study	Year	Methods or features+Model	Accuracy	Sensitivity	Specificity
Mumtaz et al.	2017	Wavelet features+LR	87.50	95.00	80.00
Mumtaz et al.	2017	SL, coherence,MI+SVM	94.70	98.30	91.40
Mumtaz et al.	2017	Asymmetry+SVM	98.40	96.66	100
Mahato et al.	2019	Alpha power,RWE+MLPNN	93.33	94.44	87.78
Mahato et al.	2020	Power, Asymmetry+SVM	86.96	86.00	89.92
Saeedi et al.	2020	Bandpower,ApEn+ E-KNN	98.44	100	97.10
Kang et al.	2020	Asymmetry Image+CNN	98.85	99.15	98.51
Dang et al.	2020	FDMB+MDCNN	97.27	97.22	97.35
Sun et al.	2022	Multilayer networks+ CN-GCN	99.29	99.37	99.32
Chen et al.	2022	Frequency matrix+DCTNet	99.15	99.30	99.01
Soni et al.	2023	Sparse graph network+KNN	92.80	/	/
Ours	2023	Local-Global graph+LG-GCN	**99.30**	**99.41**	**99.17**

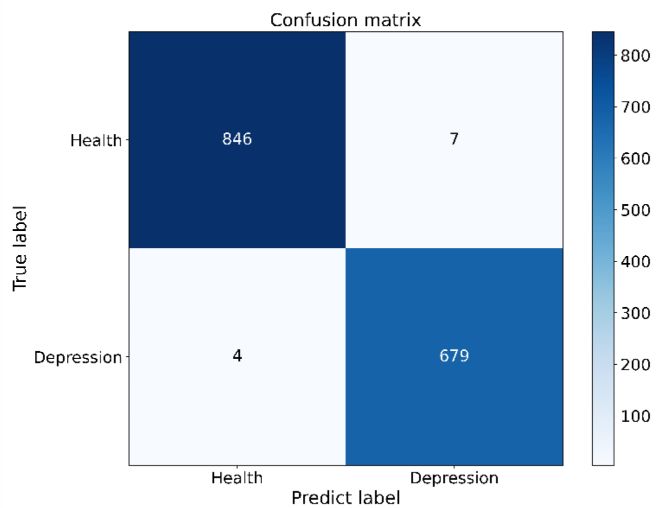


Fig. 4. Confusion matrix of the proposed LG-GCN on the public dataset.

In summary, the LG-GCN model outperforms other methods in depression recognition, with accuracy, sensitivity, and specificity of 99.30%, 99.38%, and 99.16%, respectively. Fig. 4 shows the confusion matrix of the proposed LG-GCN model for EEG-based depression recognition results.

F. Ablation Experiment

To better understand the robustness and individual contributions of the local functional graph, global dynamic graph, and IEM in LG-GCN, ablation studies are conducted by removing these modules from the LG-GCN to demonstrate the effectiveness of the model. The results are shown in Table III.

Each module is removed one by one to demonstrate the effectiveness of the combination of multiple graphs. From the results in the first and second rows of the table, Specifically, removing L from LG-GCN and not embedding the global graph led to a performance decrease in terms of Acc, Spe, and Sen. This fully verifies previous research that functional connections between local areas are important for discriminating different human emotional patterns. Hence, we adopt a graph coarsening method to aggregate local EEG features to make the model more robust. Completely removing G and directly performing graph convolution on L with full integration with IEM, results in a decrease in Acc from 99.30% to 97.90%, a decrease of 1.40% in Acc, 1.21% in Sen, and 1.59% in Spe, which highlights the importance of the global dynamic graph.

When the LG-GCN model lacked the IEM and the multi-graph representations were input directly into the GCN, flattened to the fully connected layer and softmax layer for depression recognition, the ACC dropped from 99.30% to 96.62%. This comparison demonstrates the effectiveness of IEM in achieving better performance. Additionally, it indicates that the discriminative features of the original EEG signal data can effectively complement the information captured by the LG-GCN, providing a more comprehensive representation for depression recognition. In conclusion, the information captured by multiple graph representations is crucial, and the ablation experiments strongly validate the necessity of constructing multiple graphs while demonstrating the benefits of extracting multidimensional information for model learning.

TABLE III. THE RESULTS OF ABLATION EXPERIMENTS ON PUBLIC DATASETS USING LG-GCN

L	G	IEM	Acc	Changes	Sen	Changes	Spe	Changes
×	√	√	97.52	-1.78	97.98	-1.43	97.35	-1.82
√	×	√	97.90	-1.40	98.20	-1.21	97.58	-1.59
√	√	×	96.62	-2.68	97.10	-2.31	96.38	-2.79
√	√	√	**99.30**	-	**99.41**	-	**99.17**	-

√: Keep the component
L: local function graph
G: Global dynamic graph
IEM: Information Enhancement Module
Changes: Compared with the original DGM-GCN.

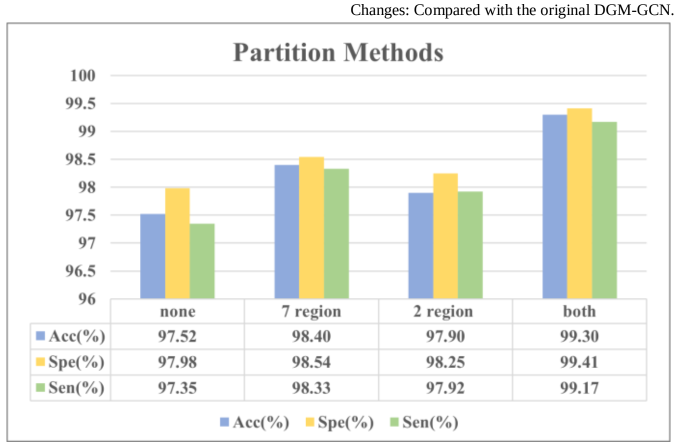


Fig. 5. Performance comparison of different region partitioning methods.

The local functional graph includes two distinct partition types, one comprising seven brain regions, and the other comprising two hemispheres, to estimate whether there is a significant difference in depression recognition tasks under coarsening at different region scales. We remove the submodules of this local functional graph and present the results in Fig. 5. It can be observed that the performance of classification is lower when the number of regions is fewer. By introducing the 7 region and 2 region subgraphs, the network in the local functional graph enhances some key information features, improving the fitting ability of LG-GCN. This suggests that the functional connectivity between local regions is closely related to depression.

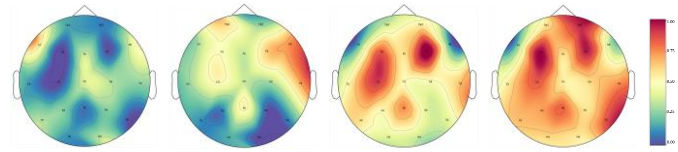
G. Interpretability and Visualization

In this section, to explain that significant connections mainly exist in certain brain regions, which have been found to be related to the pathology of depression. We visualize the distribution of degree centrality in the adjacency matrix of the global dynamics of the brain on four frequency bands in Fig. 6, providing an intuitive mapping of brain functional connections to the brain's topological structure. The degree centrality evaluates the connection strength of a node with the other nodes, which has been widely used to measure the importance of the nodes in a graph [36], calculated as follows:

$$\mathbb{C}_i = \sum_{n=1}^{19} A_{i,n} + \sum_{m=1}^{19} A_{m,i} - 2A_{i,i}, (i = 1, \dots, 19) \quad (20)$$

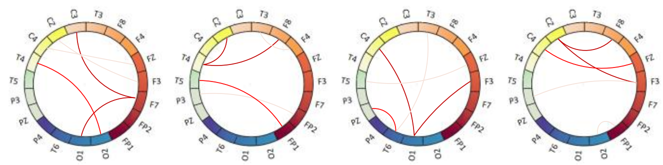
Based on the distribution of significant electrodes, we can identify brain regions that are favorable for EEG-based depression identification. This finding explains the results in Table II, where the model performs better on the α and β bands, and the connectivity is stronger than in other bands. Especially in the frontal and temporal lobes, as the frontal and temporal lobes are highly related to the onset of depression, the centrality degree is very high. Our study is consistent with [20][21], which found that the coherence of activity in the frontal and temporal regions in depressed patients is significantly higher than in healthy controls. This suggests that considering the connections between local EEG channels is crucial for accurate depression identification.

To display the connections between nodes, Fig. 7 depicts the top five connections learned by the proposed model in the adjacency matrix on the four frequency bands (δ、θ、α and β bands). Unlike the heatmap plot of degree centrality distribution, we remove the diagonal elements and just plot the connections between nodes. The positions of the 19 electrodes are displayed around the circles in the subplots. The lines represent the connections between channels, with darker colors indicating stronger connectivity. It can be observed that the critical connections are primarily located in the frontal and temporal lobes, where nodes corresponding to electrodes in these regions have a higher representation in the brain. This indicates the crucial role of the frontal and temporal lobes in the recognition of depression. In addition, two electrodes of multiple electrode pairs belong to different brain functional regions, indicating that the correlation between these regions not only has local structural connections but also has functional connections from global electrode channels.



(a) (b) (c) (d)

Fig. 6. Degree centrality distribution of the learned global adjacency matrices by LG-GCN over four frequency bands. (a) δ band：0.5-4 Hz (b) θ band：4-8 Hz (c) α band：8-13 Hz (d) β band：13-30 Hz.



(a) (b) (c) (d)

Fig. 7. Visualization of the top five functional connections between EEG channels in the learned adjacency matrices over four frequency bands. (a) δ band：0.5-4 Hz (b) θ band：4-8 Hz (c) α band：8-13 Hz (d) β band：13-30 Hz.

V. Discussion

In this research, we are thoroughly validating the effectiveness of LG-GCN. By partitioning the brain regions into local and global divisions, we discover high-level connections between different brain areas. Through the fusion of locally functional maps from prior research and adaptive adjacency matrices, our graph convolutional model accurately discerns the correlation between brain regions and depression. The fusion of local-global features is playing a crucial role in enhancing the representational power of the features. LG-GCN demonstrates a palpable edge over existing methodologies in data processing and feature extraction. Furthermore, our model evinces exceptional interpretability, elucidating the saliency accorded to specific brain regions and connections. Nodes situated within the frontal and temporal lobes exhibit heightened self-loop weights, indicative of their preeminent roles in the network, forging robust connections with other nodes. This, in turn, culminates in superlative depression identification capabilities.

VI. Conclusion

In this work, we propose LG-GCN, a local-global graph representation that simultaneously explores local brain functional connectivity and global dynamics, constructs the connections between different brain regions with prior knowledge in neuroscience research, and aggregates them into multi-layer GCNs to capture hierarchical dynamic graph topology spatial relations. In addition, IEM enables the proposed model to adaptively include discriminative depression features while retaining high-dimensional information. Extensive experimental results and visualizations demonstrate that the accuracy, sensitivity, and specificity of LG-GCN on public datasets are 99.30%, 99.41%, and 99.17%, respectively. All of these are superior to existing SOTA research, which indicates enormous potential for decoding depression based on EEG signals. However, a challenging issue in EEG-based depression recognition is the inter-individual variability of brain signals. In future work, it is crucial to take this factor into account and construct a common graph structure that is independent of individuals, thus enhancing the generalizability of the model.

ACKNOWLEDGMENT

This work was supported by the National Natural Science Foundation of China (61300098), the Natural Science Foundation of Heilongjiang Province (F201347), and the Fundamental Research Funds for the Central Universities (2572015DY07).

REFERENCES

[1] F. Chen, D. Zheng, J. Liu, Y. Gong, Z. Guan, and D. Lou, “Depression and anxiety among adolescents during COVID-19: A cross-sectional study,” Brain, behavior, and immunity, vol. 88, pp. 36, 2020.

[2] H. AlSagri, and M. Ykhlef, “Quantifying feature importance for detecting depression using random forest,” International Journal of Advanced Computer Science and Applications, vol. 11, no. 5, 2020.

[3] Y.-J. Huang, C.-Y. Wu, A. M.-K. Wong, and B.-S. Lin, “Novel active comb-shaped dry electrode for EEG measurement in hairy site,” IEEE Transactions on Biomedical Engineering, vol. 62, no. 1, pp. 256-263, 2014.

[4] T. N. Kipf, and M. Welling, “Semi-supervised classification with graph convolutional networks,” arXiv preprint arXiv:1609.02907, 2016.

[5] A. D. Friederici, N. Chomsky, R. C. Berwick, A. Moro, and J. J. Bolhuis, “Language, mind and brain,” Nature human behaviour, vol. 1, no. 10, pp. 713-722, 2017.

[6] Y. Yang, Y. Chen, H. Tang, N. Zong, and X. Jiang, “Microfluidics for biomedical analysis,” Small methods, vol. 4, no. 4, pp. 1900451, 2020.

[7] Y. Ding, N. Robinson, C. Tong, Q. Zeng, and C. Guan, “LGGNet: Learning from local-global-graph representations for brain–computer interface,” IEEE Transactions on Neural Networks and Learning Systems, 2023.

[8] H. Wang, L. Xu, A. Bezerianos, C. Chen, and Z. Zhang, “Linking attention-based multiscale CNN with dynamical GCN for driving fatigue detection,” IEEE Transactions on Instrumentation and Measurement, vol. 70, pp. 1-11, 2020.

[9] P. Zhong, D. Wang, and C. Miao, “EEG-based emotion recognition using regularized graph neural networks,” IEEE Transactions on Affective Computing, vol. 13, no. 3, pp. 1290-1301, 2020.

[10] J. D. Power, A. L. Cohen, S. M. Nelson, G. S. Wig, K. A. Barnes, J. A. Church, . . . B. L. Schlaggar, “Functional network organization of the human brain,” Neuron, vol. 72, no. 4, pp. 665-678, 2011.

[11] M. Saeedi, A. Saeedi, and A. Maghsoudi, “Major depressive disorder assessment via enhanced k-nearest neighbor method and EEG signals,” Physical and Engineering Sciences in Medicine, vol. 43, pp. 1007-1018, 2020.

[12] W. Mumtaz, L. Xia, M. A. Mohd Yasin, S. S. Azhar Ali, and A. S. Malik, “A wavelet-based technique to predict treatment outcome for major depressive disorder,” PloS one, vol. 12, no. 2, pp. e0171409, 2017.

[13] M. Kang, H. Kwon, J.-H. Park, S. Kang, and Y. Lee, “Deep-asymmetry: Asymmetry matrix image for deep learning method in pre-screening depression,” Sensors, vol. 20, no. 22, pp. 6526, 2020.

[14] U. R. Acharya, S. L. Oh, Y. Hagiwara, J. H. Tan, H. Adeli, and D. P. Subha, “Automated EEG-based screening of depression using deep convolutional neural network,” Computer methods and programs in biomedicine, vol. 161, pp. 103-113, 2018.

[15] Y. Chen, S. Wang, and J. Guo, “DCTNet: hybrid deep neural network-based EEG signal for detecting depression,” Multimedia Tools and Applications, pp. 1-15, 2023.

[16] B. Zhang, D. Wei, G. Yan, T. Lei, H. Cai, and Z. Yang, “Feature-level fusion based on spatial-temporal of pervasive EEG for depression recognition,” Computer Methods and Programs in Biomedicine, vol. 226, pp. 107113, 2022.

[17] J. Zhu, C. Jiang, J. Chen, X. Lin, R. Yu, X. Li, and B. Hu, “EEG based depression recognition using improved graph convolutional neural network,” Computers in Biology and Medicine, vol. 148, pp. 105815, 2022.

[18] H. Wu, and J. Liu, “A Multi-stream Deep Learning Model for EEG-based Depression Identification.” 2022 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). IEEE, pp. 2029-2034, 2022.

[19] X. Sun, C. Ma, P. Chen, M. Li, H. Wang, W. Dang, . . . Z. Gao, “A Novel Complex Network-Based Graph Convolutional Network in Major Depressive Disorder Detection,” IEEE Transactions on Instrumentation and Measurement, vol. 71, pp. 1-8, 2022.

[20] K.-I. Jang, C. Lee, S. Lee, S. Huh, and J.-H. Chae, “Comparison of frontal alpha asymmetry among schizophrenia patients, major depressive disorder patients, and healthy controls,” BMC psychiatry, vol. 20, no. 1, pp. 1-9, 2020.

[21] W. Mumtaz, L. Xia, S. S. A. Ali, M. A. M. Yasin, M. Hussain, and A. S. Malik, “Electroencephalogram (EEG)-based computer-aided technique to diagnose major depressive disorder (MDD),” Biomedical Signal Processing and Control, vol. 31, pp. 108-115, 2017.

[22] L. Pessoa, “Understanding emotion with brain networks,” Current opinion in behavioral sciences, vol. 19, pp. 19-25, 2018.

[23] G. E. Bruder, J. W. Stewart, and P. J. McGrath, “Right brain, left brain in depressive disorders: clinical and theoretical implications of behavioral, electrophysiological and neuroimaging findings,” Neuroscience & Biobehavioral Reviews, vol. 78, pp. 178-191, 2017.

[24] V. A. Grin-Yatsenko, I. Baas, V. A. Ponomarev, and J. D. Kropotov, “Independent component approach to the analysis of EEG recordings at early stages of depressive disorders,” Clinical Neurophysiology, vol. 121, no. 3, pp. 281-289, 2010.

[25] J. J. Allen, P. M. Keune, M. Schönenberg, and R. Nusslock, “Frontal EEG alpha asymmetry and emotion: From neural underpinnings and methodological considerations to psychopathology and social cognition,” vol. 55, pp. e13028, 2018.

[26] L. Zhang, J. Q. Gan, Y. Zhu, J. Wang, and H. Wang, “EEG source‐space synchrostate transitions and Markov modeling in the math‐gifted brain during a long‐chain reasoning task,” Human brain mapping, vol. 41, no. 13, pp. 3620-3636, 2020.

[27] R. Li, S. Wang, F. Zhu, and J. Huang, "Adaptive graph convolutional neural networks." In Proceedings of the AAAI conference on artificial intelligence. vol. 32, 2018.

[28] X. Li, X. Zhang, J. Zhu, W. Mao, S. Sun, Z. Wang, . . . B. Hu, “Depression recognition using machine learning methods with different feature generation strategies,” Artificial intelligence in medicine, vol. 99, pp. 101696, 2019.

[29] D. P. Kingma, and J. Ba, “Adam: A method for stochastic optimization,” arXiv preprint arXiv:1412.6980, 2014.

[30] S.-L. Developers, “Metrics and Scoring: Quantifying the Quality of Predictions,” User Guide,[entre 2007 e 2019]. Disponível em: https://scikit-learn. org/stable/modules/model_evaluation. html. Acesso em, vol. 26, 2021.

[31] S. Mahato, and S. Paul, “Detection of major depressive disorder using linear and non-linear features from EEG signals,” Microsystem Technologies, vol. 25, pp. 1065-1076, 2019.

[32] W. Dang, Z. Gao, X. Sun, R. Li, Q. Cai, and C. Grebogi, “Multilayer brain network combined with deep convolutional neural network for detecting major depressive disorder,” Nonlinear Dynamics, vol. 102, no. 2, pp. 667-677, 2020.

[33] W. Mumtaz, S. S. A. Ali, M. A. M. Yasin, and A. S. Malik, “A machine learning framework involving EEG-based functional connectivity to diagnose major depressive disorder (MDD),” Medical & biological engineering & computing, vol. 56, pp. 233-246, 2018.

[34] S. Mahato, and S. Paul, “Classification of depression patients and normal subjects based on electroencephalogram (EEG) signal using alpha power and theta asymmetry,” Journal of medical systems, vol. 44, pp. 1-8, 2020.

[35] S. Soni, A. Seal, S. K. Mohanty, and K. Sakurai, “Electroencephalography signals-based sparse networks integration using a fuzzy ensemble technique for depression detection,” Biomedical Signal Processing and Control, vol. 85, pp. 104873, 2023.

[36] X. Zhang, G. Cheng, and Y. Qu, " S. Saedi, A. A. F. Fini, M. Khanzadi, J. Wong, M. Sheikhkhoshkar, and M. Banaei, “Applications of electroencephalography in construction,” Automation in Construction, vol. 133, pp. 103985, 2022