

Article

DepressionGraph: A Two-Channel Graph Neural Network for the Diagnosis of Major Depressive Disorders Using rs-fMRI

Zhiqiang Xia ^{1,2}, Yusi Fan ^{1,2}, Kewei Li ^{1,2}, Yueying Wang ^{1,2,*}, Lan Huang ^{1,2} and Fengfeng Zhou ^{1,2,*} 

¹ College of Computer Science and Technology, Jilin University, Changchun 130012, China; xiazq20@mails.jlu.edu.cn (Z.X.); ysfan23@mails.jlu.edu.cn (Y.F.); likw21@mails.jlu.edu.cn (K.L.); huanglan@jlu.edu.cn (L.H.)

² Key Laboratory of Symbolic Computation and Knowledge Engineering of Ministry of Education, Jilin University, Changchun 130012, China

* Correspondence: wyy180320@163.com (Y.W.); fengfengzhou@gmail.com or ffzhou@jlu.edu.cn (F.Z.)

Abstract: Major depressive disorder (MDD) is a prevalent psychiatric condition with a complex and unknown pathological mechanism. Resting-state functional magnetic resonance imaging (rs-fMRI) has emerged as a valuable non-invasive technology for MDD diagnosis. By utilizing rs-fMRI data, a dynamic brain functional connection network (FCN) can be constructed to represent the complex interacting relationships of multiple brain sub-regions. Graph neural network (GNN) models have been widely employed to extract disease-associated information. The simple averaging or summation graph readout functions of GNNs may lead to a loss of critical information. This study introduces a two-channel graph neural network (DepressionGraph) that effectively aggregates more comprehensive graph information from the two channels based on the node feature number and node number. Our proposed DepressionGraph model leverages the transformer–encoder architecture to extract the relevant information from the time-series FCN. The rs-fMRI data were obtained from a cohort of 533 subjects, and the experimental data show that DepressionGraph outperforms both traditional GNNs and simple graph readout functions for the MDD diagnosis task. The introduced DepressionGraph framework demonstrates efficacy in extracting complex patterns from rs-fMRI data and exhibits promising capabilities for the precise diagnosis of complex neurological disorders. The current study acknowledges a potential gender bias due to an imbalanced gender distribution in the dataset. Future research should prioritize the development and utilization of gender-balanced datasets to mitigate this limitation and enhance the generalizability of the findings.

Keywords: major depressive disorder (MDD); resting-state functional magnetic resonance imaging (rs-fMRI); graph neural network (GNN); brain functional connection network (FCN); transformer–encoder; MDD identification



Citation: Xia, Z.; Fan, Y.; Li, K.; Wang, Y.; Huang, L.; Zhou, F. DepressionGraph: A Two-Channel Graph Neural Network for the Diagnosis of Major Depressive Disorders Using rs-fMRI. *Electronics* **2023**, *12*, 5040. <https://doi.org/10.3390/electronics12245040>

Academic Editor: Duc Thanh Nguyen

Received: 10 November 2023

Revised: 13 December 2023

Accepted: 14 December 2023

Published: 18 December 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Major depressive disorder (MDD) represents one of the most prevalent mental diseases worldwide, and it stands as a leading cause of disabilities like depressed mood, marked loss of interest in activities, and difficulty in concentrating [1,2]. The heterogeneity among the MDD patients leads to significant variations in symptoms and poses a considerable challenge in the accurate diagnosis [3]. Technology innovations may facilitate the early diagnosis and timely treatment of MDD patients.

Research on automated, computer-aided medical decision-making systems has made significant progress with the advancement of medical imaging technologies. Bayareh-Mancilla et al. integrated the morphological shape and skin thickness in the mammogram-based asymmetry screening and demonstrated that such explainable features could significantly improve the mammogram screening performance [4]. Explainable medical image features like segmentation-based skin thickness were further supported by the efficacy of early cancer detection [5].

Resting-state functional magnetic resonance imaging (rs-fMRI) has emerged as a widely-used tool for analyzing MDD as it provides insights into the brain's functional connectivities through the dynamic changes in cerebral blood flows [6,7]. The human brain is recognized as an integrated network comprising multiple brain regions that are structurally separated yet functionally interconnected. The functional connectivity (FC) between two brain regions is determined by analyzing the correlations between the two sequences of blood-oxygen-level-dependent signals [8].

Many studies have demonstrated the feasibility of diagnosing MDD based on the changes in brain FCs. The FCs of MDD patients at rest underwent abnormal changes over time, primarily affecting the brain regions associated with the default mode, subcortical, and sensorimotor functions [9,10]. Furthermore, the degree of the FC abnormality correlated with the severity of depression in the MDD patients, and the brain network markers derived from FCs proved effective in distinguishing MDD patients from healthy controls (HCs) [11]. Preti et al. highlighted the advantages of utilizing the dynamic FCs to deliver richer temporal dynamic information against the static versions [12]. Zheng et al. proposed a high-order FC network (FCN) to capture dynamic and high-level brain FCs and achieved a significantly improved accuracy in the MDD diagnosis [13]. These findings underscored the potential of utilizing the FC information for the MDD diagnosis.

The classification tasks of rs-fMRI-based FCNs can be broadly categorized into traditional machine learning and modern deep learning approaches. Several machine learning algorithms were designed to handle the brain FCN data in the form of an FC matrix. Vergun et al. employed Fisher's z-transformation to normalize the brain FC matrix and used these features for age classification based on the classifier support vector machine (SVM) [14]. Lei et al. constructed FCNs using multiple brain region segmentation templates and employed SVM with a sigmoid kernel for the multi-class classification based on the fused features from these FC matrices [15]. The other studies proposed differential feature extraction algorithms from FCNs and applied various classifiers for MDD detections [7,16–18].

Deep learning techniques, especially graph neural networks (GNNs), have recently gained attention for their capability to model graph-structured data effectively. GNNs have also been investigated for their capability in diagnosing brain disorders such as MDD and MCI [19]. Kim et al. utilized a GCN-based architecture for early Alzheimer's disease (AD) diagnosis in cognitively unimpaired individuals [20]. Pitsik et al. combined GNN and long short-term memory (LSTM) networks to capture the dynamic changes in the functional brain networks for MDD prediction [18]. Liu et al. utilized rs-fMRI to generate multi-index representations for each subject and proposed a tensor-based multi-task learning model with a tensor-based regularizer for effective MDD predictions [21]. In addition, Wang et al. collected information from three modalities (DTI, fMRI, and sMRI) for each sample, utilized a transformer and CNN as feature extractors and employed an ensemble learning voting mechanism to obtain predictions. This is a multi-modal deep-learning framework designed for the automatic diagnosis of MDD [22]. Venkatapathy et al. integrated multiple graph models for the challenging MDD classification task across 16 sites [23]. Zhao et al. employed the graph-encoder-based architecture for multi-site MDD classification [24]. Both studies revealed that cross-site batch effects substantially compromised the accuracy of MDD classification, resulting in significantly reduced performance metrics [23,24].

This study proposes a two-channel graph neural network, DepressionGraph, to extract the global graph information in the two channels based on node feature number and node number. This setting alleviates the limitation of the simple readout function in the graph representations. We also integrate the transformer-encoder part of the vision transformer (ViT) [25] to capture the dynamic information at different time points of the dynamic FCNs. The proposed framework, DepressionGraph, demonstrates superior classification performance compared to existing approaches. The ablation experiment further validates the effectiveness and necessity of the main modules in DepressionGraph. However, our

study has a limitation regarding potential gender bias, as the dataset used for analysis contains a larger number of female samples than male samples.

2. Materials and Methods

2.1. Dataset

The public rs-fMRI dataset utilized in this study was generously provided by the REST-meta-MDD consortium and comprises data from 17 Chinese hospitals/sites [26]. This study specifically selects the site 20 with the largest number of subjects, similar to [27]. The data were downloaded from <http://rfmri.org/REST-meta-MDD> accessed on 27 March 2021. The dataset consists of 533 subjects, among whom 282 individuals were diagnosed with MDD—the other 251 subjects served as the matched healthy controls (HCs).

The diversity in demographic characteristics of the participants is important in influencing the manifestation of MDD symptoms and will impact the generalizability of the MDD prediction models. The demographic characteristics of the dataset are summarized in Table 1. The proportion of female participants in the MDD and HC classes is similar. The mean and standard deviation values of the participants' ages are also similar in the two classes. The HC participants have a slightly longer education length (12.97 years) than the MDD patients (10.78 years). Due to that, this whole cohort was recruited from Chinese hospitals/sites. The majority of the participants are Chinese.

Table 1. Demographics of the subjects in the MDD dataset. MDD: major depressive disorder; HC: healthy control; M/F: male/female; Edu: education years; std: standard deviation.

	MDD ($n = 282$)	HC ($n = 251$)
Gender (M/F)	99/183	87/164
Age (mean \pm std)	38.74 \pm 13.65	39.64 \pm 15.87
Edu (mean \pm std)	10.78 \pm 3.61	12.97 \pm 3.94

The rs-fMRI scans were acquired using the Siemens Tim Trio 3T MRI scanner at Southwestern University. The scanning parameters used for data acquisition were as follows: a repetition time (TR) of 2000 ms, an echo time (TE) of 30 ms, and a flip angle of 90°. The image slice thickness/gap was set to 3.0 mm/1.0 mm, resulting in a total of 32 slices. Each rs-fMRI scan consisted of 242-time points, and the voxel size was 3.44 \times 3.44 \times 4.00 mm³, with a field of view (FOV) measuring 220 \times 220 mm².

We applied the preprocessing step using the Data Processing Assistant for rs-fMRI (DPARSF) [28,29]. During this step, the rs-fMRI BOLD signals were extracted for each subject. The preprocessing operations included (1) discarding the first ten volumes to achieve magnetization equilibrium, (2) slice timing correction and head motion correction, (3) normalization to the Montreal Neurological Institute (MNI) template and resampling to isotropic 3 mm voxel size, (4) smoothing, (5) detrending and band-pass filtering, and (6) regressing covariates. Ultimately, each subject's data was transformed as a matrix of size $T \times N$, where T denotes the number of time points and N is the number of brain regions of interest (ROIs).

Compared to other fMRI data preprocessing software, such as SPM version 12 (<https://www.fil.ion.ucl.ac.uk/spm/software/download/spmreg.php>), dparxf is more user-friendly and efficient, significantly improving the efficiency of image analysis. Its streamlined workflow minimizes the risk of operational and subjective errors in the rs-fMRI image analysis.

2.2. Overall Framework

The proposed DepressionGraph framework aims to identify MDD by effectively capturing the time-varying information of brain FCNs based on rs-fMRI time-series data, as depicted in Figure 1. Each subject is denoted as $X \in \mathbb{R}^{N \times T}$, where there are N brain regions of interest (ROIs) and T length of the sequence of the BOLD signals over time. The ROIs

are segmented according to the brain region segmentation template. We adopt a sliding window approach by dividing the time series data of each ROI into T_{dyn} segments of length L . As a result, each subject is represented as $X_{dyn} = \{X_1, \dots, X_{T_{dyn}}\} \in \mathbb{R}^{T_{dyn} \times N \times L}$.

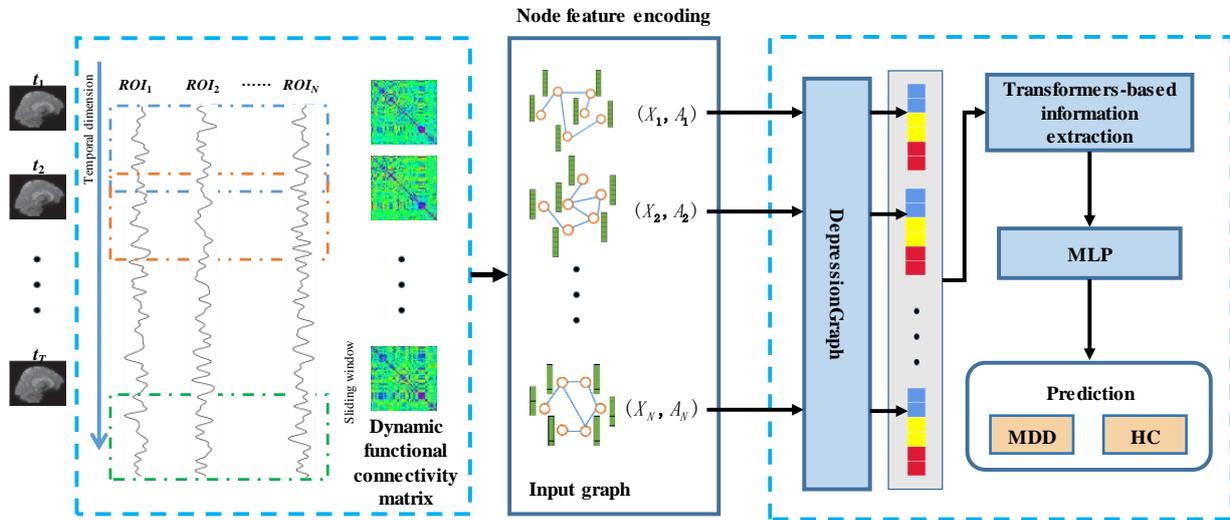


Figure 1. Overview of the workflow for detecting MDD based on rs-fMRI data.

For $X_t \in \mathbb{R}^{N \times L}$ at the t th time point, we construct an FCN with the ROIs as the nodes and the BOLD signal value of the t th time point as the node feature of the t th FCN. An edge is defined between two nodes if the correlation coefficient between their node features satisfies a predefined threshold. Let $G_t = (X_t, A_t)$ be the FCN of the t th time point. The feature matrix is denoted by $X_t = \{x_t^1, \dots, x_t^N\}$, and the adjacency matrix is denoted by $A_t \in \{0, 1\}^{N \times N}$.

Our proposed DepressionGraph framework uses a gate recurrent unit (GRU) network to encode the node features and the two-channel graph neural network to obtain the feature representation of the entire graph at the t th time point G_fea_t . Subsequently, we introduce a transformer-based network to capture temporal information from the FCN graphs, denoted as G_fea_{dyn} . Finally, we employ a multi-layer perceptron (MLP) to predict the labels, expressed as $Label_{pred} = MLP(G_fea_{dyn})$.

The next two sections give detailed descriptions of the proposed DepressionGraph model and the transformer-based timing information extraction module.

2.3. DepressionGraph: Two-Channel Graph Neural Network

The DepressionGraph framework aims to generate a one-dimensional feature vector that comprehensively captures the global information of an input graph (Figure 2). It leverages two distinct channels for graph representation learning: one for the fine-grained node features and the other for the coarse-grained features of the entire graph. Given the graph, $G = (X, A)$ with N nodes and M features for each node. The global features of the input graph obtained through the DepressionGraph can be represented as:

$$G_fea = concatenate(G_fea_{channel1}, G_fea_{channel2}), \tag{1}$$

Here, $G_fea_{channel1} \in \mathbb{R}^N$ represents the graph features extracted through Channel 1, where the dimension of node features is transformed into a one-dimensional vector of length N . Similarly, $G_fea_{channel2} \in \mathbb{R}^M$ represents the graph features extracted through Channel 2, where the dimension of node count is transformed into a one-dimensional vector of length M . The final feature representation of the entire graph is obtained by concatenating the two vectors acquired from the two channels.

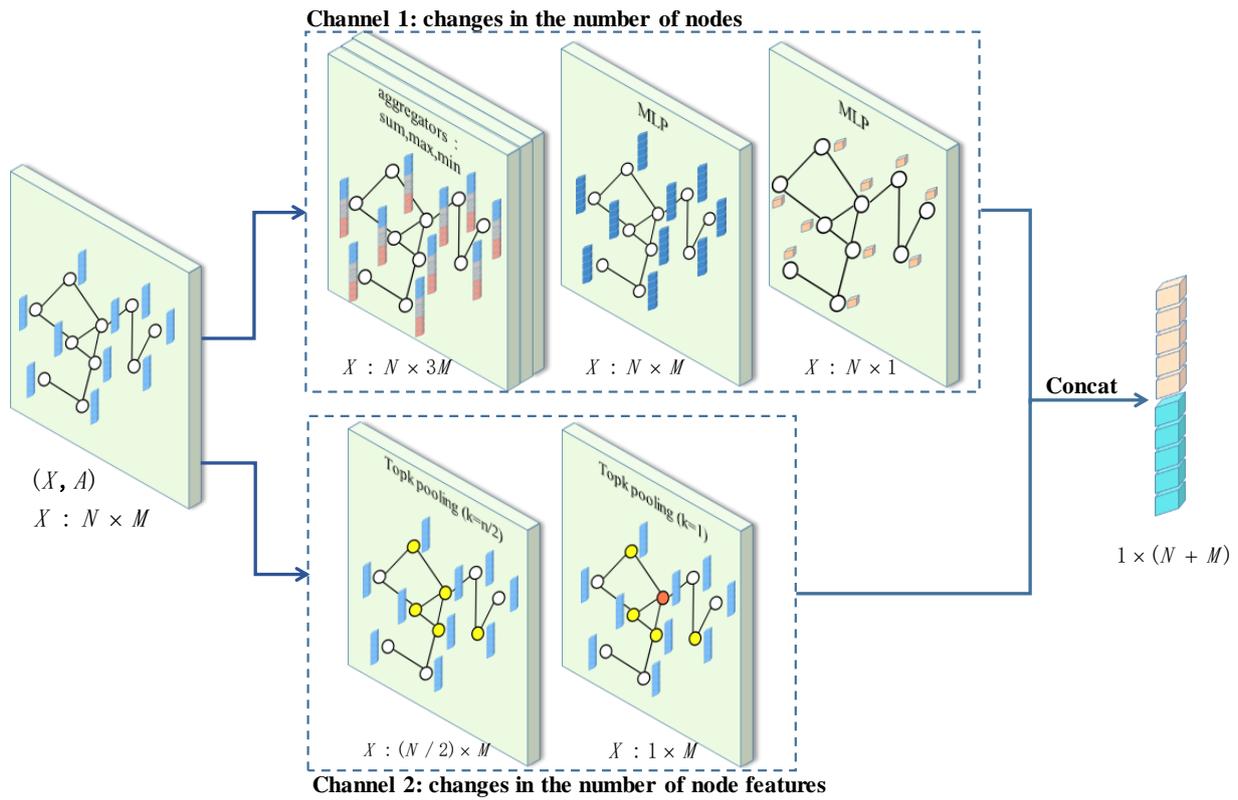


Figure 2. Illustration of the proposed DepressionGraph framework. The upper half focuses on extracting features based on changes in the number of nodes (Channel 1), while the lower half focuses on extracting features based on changes in the number of node features (Channel 2). Specifically, during the channel2 process, the graph of n input nodes undergoes the first graph pooling, aggregating into $n/2$ nodes, represented by the yellow nodes in Figure 2. After the second graph pooling, the nodes further aggregate into a single node, represented by the orange node in Figure 2.

2.4. Channel 1: Graph Feature Extraction by Changing the Number of Node Features

This channel modifies the number of node features to enable the extraction of essential graph information. Firstly, we employ three aggregation methods to combine information from neighboring nodes: (1) summing the features of neighboring nodes to obtain X_{Sum_agg} , (2) taking the maximum feature value of neighboring nodes to obtain X_{Max_agg} , and taking the minimum feature value of neighboring nodes to obtain X_{Min_agg} . After aggregation, the node feature matrix becomes $X_{agg} \in \mathbb{R}^{N \times 3M}$, represented as:

$$X_{agg} = \text{concatenate}(X_{Sum_agg}, X_{Max_agg}, X_{Min_agg}). \tag{2}$$

Next, we employ a multi-layer perceptron (MLP) to reduce the dimensionality of node features. Through one layer of MLP, the node features of length $3 \times M$ are transformed into features of length M . Another layer of MLP further reduces the length of each node’s features to one. These one-dimensional node features are concatenated to form a one-dimensional vector of length N , which serves as the representation of the graph for Channel 1. This representation captures the fine-grained features of each node.

$$G_fea_{channel1} = \text{MLP}_2(\text{MLP}_1(X_{agg})). \tag{3}$$

2.5. Channel 2: Graph Feature Extraction by Changing the Number of Nodes

Channel 2 aims to capture critical graph information by altering the number of nodes. The top-k pooling algorithm [30] is employed to successively reduce the N nodes to $N/2$

nodes and then further reduce the $N/2$ nodes to a single node. Therefore, all node features in the graph are aggregated into this single node while preserving the original number of features (M). The resulting one-dimensional vector of length M (denoted as $G_fea_{channel2}$) serves as the representation of the graph for Channel 2, capturing the global information of the graph.

$$G_fea_{channel2} = \text{TopK}_2(\text{TopK}_1(X, A)). \tag{4}$$

2.6. ViT-Based Timing Information Extraction

Figure 3 shows that the brain FCN at each time period can be viewed as a graph structure. The proposed DepressionGraph framework represents a graph as a one-dimensional vector $G_fea_i \in \mathbb{R}^C$, where $C = M + N$. Therefore, each subject is represented as a time-varying graph representation $H \in \mathbb{R}^{T \times C}$. A simple weighted averaging of the graph representation at different time points would not adequately capture the temporal dynamics. To address this issue, we draw inspiration from the vision transformer (ViT) network [31] and introduce location encoding via the encoder part of the transformer network to effectively capture information across time.

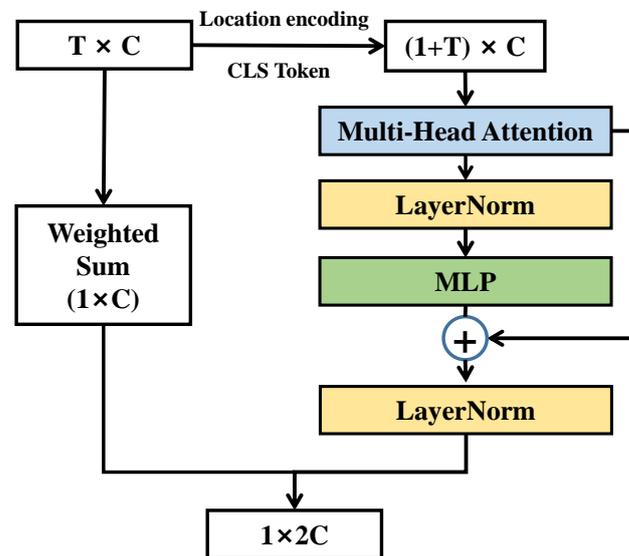


Figure 3. The architecture of transformers-based timing information extraction.

The dynamic graph features G_fea_{dyn} for each subject can be represented as:

$$G_fea_{dyn} = \text{concatenate}(f_{Sum}(H^T \times W), f_{encoder}(H_{loc})). \tag{5}$$

Here, the variable W represents the weight assigned to different time points, $f_{Sum} \in \mathbb{R}^C$ is the vector obtained by weighted summing of the different time point vectors, H_{loc} is the matrix obtained by adding location encoding to H , and $f_{encoder} \in \mathbb{R}^C$ is the vector obtained after passing through the encoder. We concatenate the vectors obtained from these two methods to obtain the dynamic graph representation vector $G_fea_{dyn} \in \mathbb{R}^{2C}$.

2.7. Fusion of Prediction Results of Multiple Brain Region Segmentation Templates

The utilization of various brain region segmentation templates has been demonstrated to enhance the accuracy of brain disease diagnosis [32,33]. This study selects three widely adopted brain segmentation templates to construct brain FCNs, i.e., Harvard–Oxford atlas [34], Automated Anatomical Labeling (AAL) atlas [35], and Craddock’s clustering 200 ROIs (CK) [36]. We construct a brain FCN using each template and employ the majority voting strategy from the predicted labels of the three trained classification models.

This prediction strategy allows us to leverage the diverse information captured by different brain region segmentation templates, and we aim to achieve an improved and more reliable diagnosis of MDD using rs-fMRI.

2.8. Implementation Details

The code implementation was carried out in the Python programming language, utilizing PyTorch version 1.7.1 and Scikit-learn version 0.24.2. All the experiments were conducted on a system equipped with an NVIDIA P100 GPU featuring 16 GB VRAM.

The graph pooling in channel two was implemented by the Top-K pooling adjusted to the problem setting in this study [30]. The approach allowed us to effectively reduce the number of nodes while preserving essential graph information.

The original BOLD signal was partitioned into multiple segments using a sliding window strategy. We set the window size to 50 and the window stride to 3. For example, if the BOLD signal has a length of 232 time points, it will be partitioned into 61 segments.

A graph was constructed by determining an edge between two nodes if their correlation coefficient was within the top 30% of all the correlation coefficients. We employed the cross-entropy loss function and utilized the OneCycleLR technique to adjust the learning rate dynamically. The maximum learning rate was set to 0.001. If a fixed learning rate is used, satisfying the learning rate requirements of both the initial and final stages of the model becomes challenging. If the learning rate is too large, the model may struggle to find a better solution. On the other hand, if the learning rate is too small, the optimization process at the initial stage becomes slow, resulting in slow convergence. We used dropout with a 0.5 probability instead of regularization methods. This was done because there were no apparent signs of overfitting during the model training process. Our model was trained for 50 epochs, and by the 50th epoch, the model had already converged. The training batch size, either 1 or 4, was determined by the number of nodes in the graph. Specifically, for experiments using the AAL and HO templates, we set the batch size to 4, while for experiments using the CK template, the batch size was set to 1. We used a grid search strategy to obtain the optimal values of the hyper-parameters. Regarding training speed, for the AAL and HO templates, each epoch took approximately 4 min 30 s. Completing the five-fold cross-validation required approximately 19 h. Due to the larger number of brain regions in the CK template resulting in a higher computational load, each epoch took approximately 6 min. Completing the five-fold cross-validation for the CK template required approximately 25 h.

Slow training may pose challenges in the practical implementation of Depression-Graph for clinical situations with new cohorts of samples or new ethnic populations. Taking into account the factor of model training time, we have designed the model structure with a moderately sized number of MLP layers and graph pooling layers. This helps control the parameter count of the model and ensures that it does not require excessively long training time. Neural network distillation technology may be utilized to reduce the model training time further.

2.9. Performance Measurements

This study formulates the determination of whether a sample suffers from MDD as a binary classification task. The prediction model is evaluated by the following five widely used measurements, i.e., accuracy (Acc), precision (Pre), sensitivity (Sen), specificity (Spe), F1 score (F1), and the area under the receiver operating characteristic curve (AUC). The numbers of correctly predicted positive (MDD) and negative (HC) samples are denoted as TP and TN, respectively, while those of the incorrectly predicted positive and negative samples are denoted as FN and FP, respectively.

$$\text{Acc} = (\text{TP} + \text{TN}) / (\text{TP} + \text{FN} + \text{TN} + \text{FP}), \quad (6)$$

$$\text{Pre} = \text{TP} / (\text{TP} + \text{FP}), \text{Sen} = \text{TP} / (\text{TP} + \text{FN}), \text{Spe} = \text{TN} / (\text{TN} + \text{FP}) \quad (7)$$

and

$$F1 = 2 \times \text{Pre} \times \text{Sen} / (\text{Pre} + \text{Sen}). \quad (8)$$

All the classifiers are trained using the training subset and evaluated on the test subset. To ensure the experimental robustness and reduce bias, we employ a five-fold cross-validation strategy for the evaluation of all the models. This approach involves randomly dividing a dataset into five equally sized subsets and using four subsets for training and the remaining one subset for testing iteratively. The final performance is averaged across the five iterations.

3. Results and Discussion

3.1. Comparison of Different Graph Neural Networks

In this section, we assess the performance of various graph neural networks on the MDD prediction task (Figure 4). Traditional graph neural networks, such as graph convolutional networks (GCNs) and graph isomorphism networks (GINs), often employ simple averaging or summation methods for graph readout in graph-based classification tasks. However, this simplistic approach may overlook critical graph information. The proposed DepressionGraph utilizes two channels of graph information to capture more comprehensive global information from the graph.

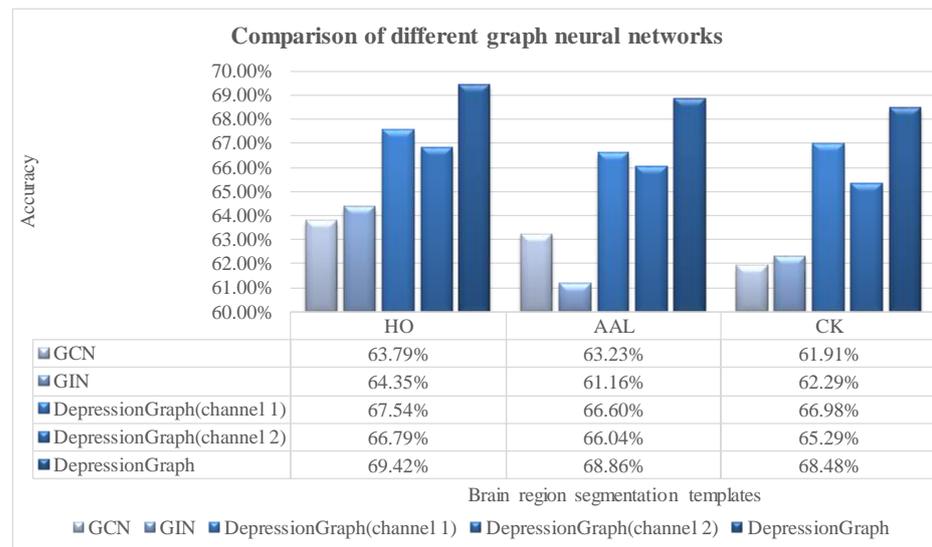


Figure 4. Classification accuracies of different graph neural networks for the MDD prediction task. The horizontal axis gives the three brain region segmentation templates, and the vertical axis is the classification accuracy.

Figure 4 presents the MDD prediction accuracies of five models, including GCN, GIN, and two variants of DepressionGraph using only channel one and channel two independently, together with the proposed DepressionGraph framework. DepressionGraph outperforms all four other models using each of the three templates HO/AAL/CK. At least a 2.24% improvement in accuracy has been achieved by DepressionGraph. The DepressionGraph variant with channel one only achieves the second-best accuracy using each of the three templates. It is evident that utilizing features from two channels significantly enhances the model's performance compared to using features from just one channel across three templates. The experimental results in Figure 4 validate the effectiveness of our proposed model and demonstrate that the features extracted from Channel 1, which focuses on capturing local attention, and Channel 2, which emphasizes global attention, are complementary to each other.

3.2. Comparison with Other Classification Algorithms

We constructed a graph structure and used the local clustering coefficients as the features, similarly to [37]. These constructed features were then classified using the classifiers, including support vector machine (SVM), random forest (RF), extreme gradient boosting (XGBoost), multilayer perceptron (MLP), and convolutional neural network (CNN). Figure 5 presents the performance of these machine learning or deep learning methods as well as our proposed DepressionGraph. The experimental data demonstrated that DepressionGraph outperformed other algorithms in terms of metrics, such as Acc = 68.86% and AUC = 67.90%. Particularly, the Sen = 84.40% of DepressionGraph was significantly higher than that of other algorithms. Sen measures the detection rate of the model for diseased samples, and a higher Sen value can reduce the rate of false negatives, which is crucial for the MDD diagnosis task. DepressionGraph achieves Spe = 51.39%, which is similar to the classifiers SVM, XGBoost, and MLP. The other two classifiers, RF and CNN, outperform DepressionGraph in Spe, but their Sen values are the worst two in Figure 5. The main reason for the poor performance of neural network algorithms like MLP and CNN may be that the complexity of the feature extraction models is insufficient. It may require tens of thousands of training samples to train a better model. However, our proposed DepressionGraph offers a more targeted approach to extract features from changes in brain networks caused by diseases such as MDD or MCI. Even with fewer training samples, it can still achieve excellent performance.

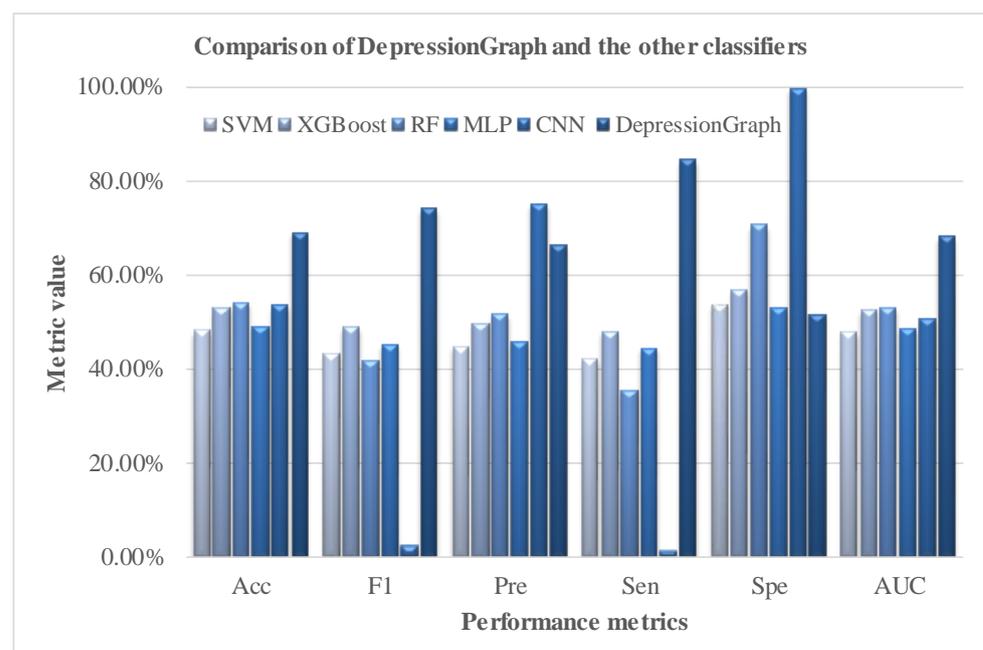


Figure 5. Performance comparison between DepressionGraph and the other machine learning or deep learning algorithms. The horizontal axis gives the performance metrics Acc, F1, Pre, Sen, Spec, and AUC. The vertical axis lists the metric values. The experiment evaluates six classification models, including SVM, XGBoost, RF, MLP, CNN, and DepressionGraph.

3.3. Necessity of Dynamic Brain Functional Connection Network

To demonstrate the necessity of adopting the dynamic functional connectivity network (DFCN) approach, we compared the performance of DepressionGraph using the static functional connectivity network (SFCN) versus DFCN. The AAL template was used. As shown in Table 2, it is evident that DFCN outperforms SFCN in all metrics, highlighting the importance of incorporating time-varying features in our method. This further validates the effectiveness of DepressionGraph based on DFCN in capturing patterns of functional connectivity that evolve over time.

Table 2. Comparison of performance between DepressionGraph using DFCN and SFCN. The bolded values indicate the best-performing method for a specific indicator.

Algorithm	Acc	F1	Pre	Sen	Spe	AUC
DepressionGraph (SFCN)	56.47%	61.07%	57.96%	64.54%	47.41%	55.97%
DepressionGraph (DFCN)	68.86%	74.14%	66.11%	84.40%	51.39%	67.90%

3.4. Contribution of Transformer–Encoder Module

We further investigate the impact of incorporating the transformer–encoder for the timing information extraction in the MDD prediction task (Figure 6). We use the weighted summation module to replace the transformer encoder module in the DepressionGraph framework and denote this simplified variant as DepressionGraph-NoTE. The DepressionGraph framework outperforms DepressionGraph-NoTE on the MDD prediction task on all three brain region segmentation templates HO/AAL/CK. The experimental data show the valuable addition of the transformer encoder module to the DepressionGraph framework.

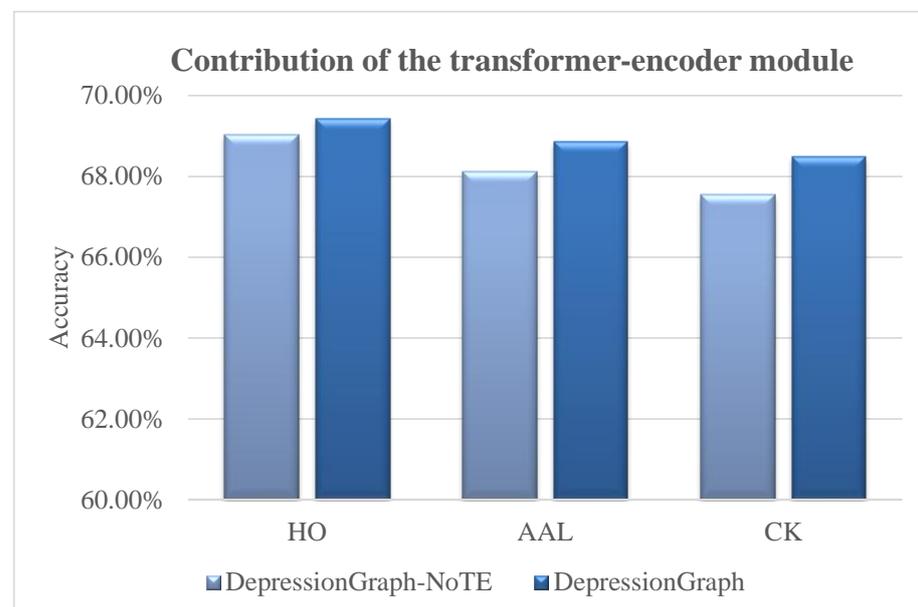


Figure 6. Evaluation of the transformer encoder module in the DepressionGraph framework. The model “DepressionGraph-NoTE” denotes the DepressionGraph without the transformer encoder module. Three brain region segmentation templates, HO/AAL/CK, are evaluated on the horizontal axis. The vertical axis gives the prediction accuracies.

3.5. Contributions of Brain Segmentation Templates

We evaluate the contributions of the three brain segmentation templates HO/AAL/CK and their fusion to the DepressionGraph framework (Table 3). The complete version of DepressionGraph fuses all three templates and achieves the best performances in three overall measurements, including Acc = 71.48%, F1 = 77.91%, and AUC = 70.03%. The variant DepressionGraph(AAL) achieves the highest Pre = 66.11%, slightly better than that (66.01%) of DepressionGraph. The other variant, DepressionGraph(CK) (Sen = 98.58%), outperforms the complete version DepressionGraph (Sen = 95.04%) in the measurement Sen.

We analyzed the features encoded by the second-last layer of the HO-based model. We chose the HO brain parcellation template because it outperformed the models based on the AAL and CK templates (Table 3). The features in the layer before the final layer of the model are one-dimensional vectors with a length of 480. We separately calculated the Pearson correlation coefficients (PCCs) between these 480 features and the class labels separately. PCC ranges from -1 to 1 . After analyzing the distribution of the latent features

encoded by the HO-based model, we found that the highest absolute value of these PCCs was less than 0.1 (Figure 7). We speculate that this might be because the features before the final layer of the model still need to pass through a linear layer for classification. The weight parameters of the linear layer and the cross-entropy loss function are also crucial for the classification results, as they can potentially lead to a lack of strong correlation between the features before the final layer and the labels. We believe that improving the correlation between the latent features and the labels would further contribute to enhancing the model's performance. In our future work, we plan to design additional loss functions specifically for enhancing the correlations of the latent features with the labels.

Table 3. Contribution evaluation of DepressionGraph using different brain segmentation templates. The column “Model” includes the complete version of DepressionGraph and the three variants using only one of the three templates HO/AAL/CK. The next columns give the five performance measurements, i.e., Acc, F1, Pre, Sen, Spe, and AUC. The bolded values indicate the best-performing method for a specific indicator.

Model	Acc	F1	Pre	Sen	Spe	AUC
DepressionGraph(HO)	69.42%	75.11%	65.95%	87.23%	49.40%	68.32%
DepressionGraph(AAL)	68.86%	74.14%	66.11%	84.40%	51.39%	67.90%
DepressionGraph(CK)	68.48%	76.80%	62.90%	98.58%	34.66%	66.62%
DepressionGraph	71.48%	77.91%	66.01%	95.04%	45.02%	70.03%

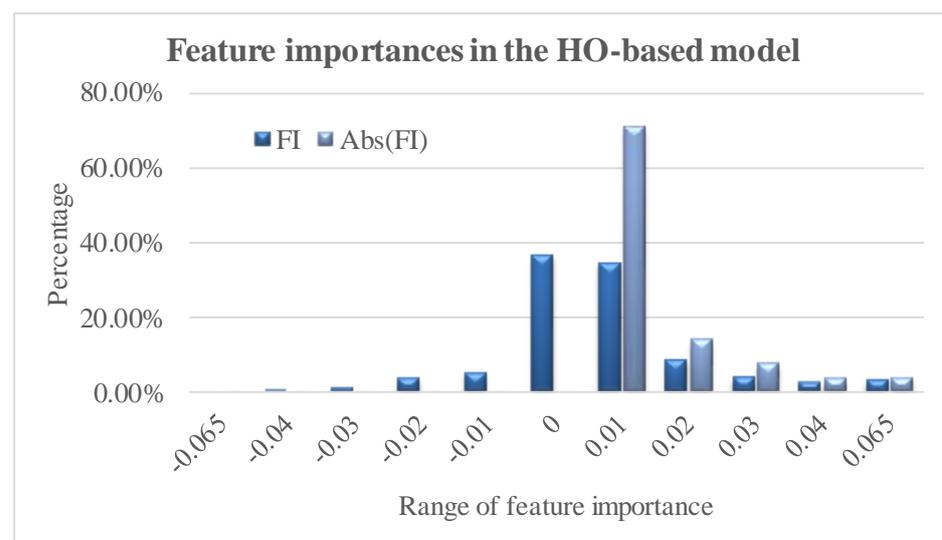


Figure 7. Distribution of the feature importance in the HO-based model. The horizontal axis gives the value ranges of the feature importance metrics in the HO-based model. The vertical axis lists the percentages of the features with the importance metrics in the HO-based model. The data series “FI” and “Abs(FI)” refer to the feature importance and the absolute values, respectively, of the feature importance of the latent features encoded by the HO-based model.

In summary, the DepressionGraph framework achieves similarly well, with values of Acc = 69.42%, 68.86%, and 68.48% for the three templates HO, AAL, and CK, respectively. The fusion of all three templates further improves the MDD prediction model to Acc = 71.48%. However, the fusion of three brain segmentation templates increased the algorithmic complexity and model training time. Future work may consider the construction of a better single template-based prediction framework or the designing of refined brain segmentation templates.

3.6. Comparison with the Existing Studies

The overall performance of DepressionGraph is compared with the existing studies (Figure 8). The comparison experiment focuses on the models using the same dataset as in this study, including Gu et al. [38], Li et al. [39], Jie et al. [40], Guo et al. [41], Yao et al. [27], Zhang et al. [42], and Zhu et al. [43]. Figure 8 shows that DepressionGraph outperforms all the other studies in both the two performance measurements, Acc and F1. Previous studies did not release the performance measurements Pre, Sen, Spe, and AUC. The proposed DepressionGraph even improves the MDD prediction accuracy of Yao et al. [27] by 11.34%. Zhang et al. proposed a multi-view graph neural network to detect MDD across ten sites and achieved Acc = 65.61% and F1 = 64.55%, which were worse than DepressionGraph [42]. Zhu et al. evaluated their model across 16 sites, and DepressionGraph outperformed their model by 0.78% in Acc and 6.78% in F1 on the same site [43]. The multi-site investigations referenced in [42,43] reported F1 scores around 65%, whereas the DepressionGraph model attained a more elevated F1 score of 71.13%. This disparity implies the presence of an inter-site batch effect, underscoring the necessity for future research focused on the development and implementation of sophisticated batch effect mitigation algorithms.

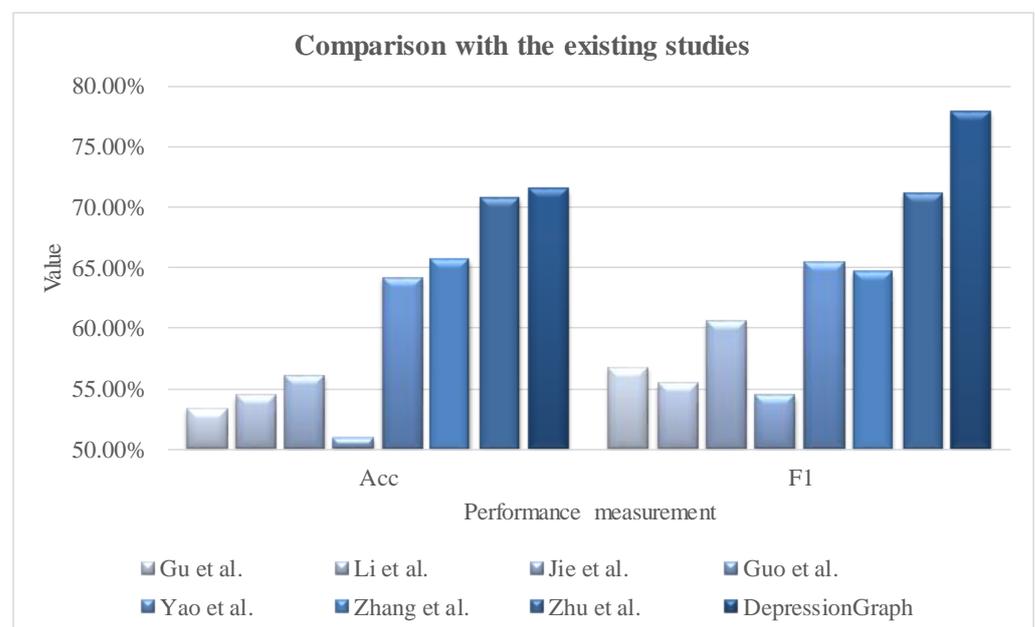


Figure 8. Performance comparison with the existing studies. The horizontal axis gives the two measurements, Acc and F1, and the vertical axis gives the values of these performance measurements. The proposed DepressionGraph framework is compared with the previous studies, including Gu et al. [38], Li et al. [39], Jie et al. [40], Guo et al. [41], Yao et al. [27], Zhang et al. [42], and Zhu et al. [43].

3.7. Validation of DepressionGraph on Mild Cognitive Impairment

To further validate the effectiveness of DepressionGraph, we downloaded 149 samples of early mild cognitive impairment (EMCI), 105 samples of mild cognitive impairment (MCI), and 133 samples of normal controls (NC) from the ADNI database [44]. The data were preprocessed using dparfs following the approach described in Yang et al. [45]. We employed the commonly used automated anatomical labeling (AAL) template for brain region segmentation.

We set up two binary classification tasks: EMCI vs. NC and LMCI vs. NC. We compared the performance of DepressionGraph with [45] to further demonstrate the effectiveness of the proposed algorithm. Table 4 presents the performance of our proposed method compared to the method proposed in [45]. The experimental data showed that our method outperformed Yang et al. in most metrics and provided evidence that De-

pressionGraph not only performed well in diagnosing MDD but also exhibited excellent performance in predicting other disorders such as MCI. We anticipated that DepressionGraph can also achieve good results on the MCI dataset because both disorders primarily arise from abnormalities within the brain and exhibit symptomatic similarities. Compared to the MDD dataset, the MCI dataset's shorter temporal signals per brain region result in quicker training and testing speeds. The performance also suggests that EMCI and LMCI are easier to detect than MDD against the normal control samples.

Table 4. Performance of DepressionGraph and Yang et al.'s method on the MCI dataset. The bolded values indicate the best-performing method for a specific indicator.

Algorithm	Task	Acc	Pre	Rec	AUC
Yang et al. [45]	EMCI/NC	82.76%	79.58%	69.73%	88.23%
DepressionGraph	EMCI/NC	86.52%	93.58%	76.69%	86.00%
Yang et al. [45]	LMCI/NC	87.23%	69.13%	75.58%	92.34%
DepressionGraph	LMCI/NC	88.66%	88.97%	90.98%	88.35%

3.8. Model Discussion

The above results indicated that our proposed DepressionGraph framework could extract more comprehensive information from graphs compared to the existing graph neural network methods, such as GCN and GIN. By utilizing the dynamic functional connectivity network (DFCN), we can capture the temporal features of functional connectivity networks (FCN). This approach significantly outperforms the SFCN method, with an improvement of 12.39% in accuracy (ACC). Recent studies have also adopted DFCN for their experiments. While most papers use the AAL template for brain parcellation, it may not be suitable for all brain disorders. To address this limitation, we employed three different brain parcellation templates and fused the results to compensate for the weakness of a single template. The experimental results further supported the effectiveness of fusing multiple brain parcellation templates.

DepressionGraph extracts high-level abstracted information from the complex brain rs-fMRI data, which lacks explainable features to guide the future clinical practice of MDD diagnosis. Explainable artificial intelligence (XAI) technologies may be considered in future studies. Combining expert medical knowledge with XAI (explainable artificial intelligence) technology is also more likely to achieve excellent performance.

3.9. Clinical Implications

The diagnosis of major depressive disorder (MDD) currently relies heavily on clinical observations. The spatiotemporal changes in brain blood flow can be detected by fMRI technology and have been related to neural activities of various neural disorders, including MDD [46]. The rapidly innovated fMRI technology offers high temporal and spatial resolution of imaging human brains, and the fMRI-based machine learning and deep learning algorithms have been extensively explored for their contributions to diagnosing MDD [47]. Precise classification of MDD patients from other neural disorders and normal controls may facilitate the computational diagnosis of MDD in future clinical practice and regular health examinations.

4. Conclusions

This study presents a two-channel graph neural network framework, DepressionGraph, for MDD prediction using rs-fMRI data. DepressionGraph leverages the capabilities of graph neural network and transformer-encoder-based timing information extraction and effectively captures rich temporal dynamics in the brain FCNs. We also address the limitation of a single brain segmentation template by fusing the prediction results from models trained with three different templates.

The experimental results clearly demonstrate the superiority of DepressionGraph over traditional GNNs as well as the existing studies. At least a 6.78% improvement in F1 is

achieved by our proposed DepressionGraph over the other models on the same dataset. Similarly to [42,43], the cross-site generalizability of DepressionGraph needs to be explored in future work. The impact of the biased ethnicity in the cohort will be evaluated with additional diversified datasets.

Long-term trends associated with MDD in the time series data may greatly facilitate a better understanding of MDD's biological mechanism and the early diagnosis of MDD in clinical practice. However, most of the existing datasets overlooked this issue, probably due to the difficulties in identifying and tracking MDD patients before and after the disease onset in the long term. The lack of long-term time-series data means most of the computational studies focus on detecting the short-term fluctuations that only reflect acute changes in the mental status of neurological disorders like MDD. The accumulation of long-term time series brain imaging data for various neurological disorders can only be resolved by the establishment of a large international consortium.

DepressionGraph extracted complex high-level information across the time series activities of the whole brain, which captured the subtle patterns of MDD. The main limitation of the DepressionGraph-extracted features is the lack of explainability, and it cannot intuitively decompose the precise MDD classification model into the causal relations of activated brain regions and the in-depth characterization of functional neurodynamic patterns. In our future work, we aim to extract clinically meaningful types of node features from the current model. We plan to visualize the changes in brain networks over time using visualization tools to explore clinically actionable conclusions. Additionally, we will attempt to optimize the network structure by adding additional constraints to the node features and analyzing the changes in brain networks over time based on the node features and the connections between nodes. This analysis will help us identify specific brain regions activated in MDD patients, which facilitates a better understanding of the causal relations of activated brain regions and enhances the interpretability and clinical applicability of our model.

The other limitations of DepressionGraph include the confounding effect of unbalanced gender ratios and the high requirement for computational resources due to the FCN construction and the evaluation of the time series data. In future work, we will explore network modules that are more computationally efficient for the MDD prediction task. Additionally, we will conduct experiments on more datasets to ensure a balanced male-female ratio in both positive and negative samples, thereby avoiding the confounding effect of gender on MDD classification. The imbalanced ratio between the two genders will be evaluated by building two separate gender-specific models. We will attempt to mitigate the limitations of the model in characterizing functional neurodynamics and explainability by utilizing more advanced algorithms to extract temporal dynamics information of individual brain regions. Additionally, we will refer to the image feature extraction methods evaluated by El-Gayar et al. [48] to extract intrinsic features from the magnetic resonance imaging itself, aiming to further enhance the performance of the model. Explainable artificial intelligence (XAI) algorithms will also be evaluated for their MDD prediction performance and the possible roles in characterizing the causal relations of brain activations and functional neurodynamics. Hierarchical clustering and k-means clustering strategies of the XAI-extracted explainable features will be employed to group the MDD patients into clinically meaningful subpopulations of MDD biotypes in future work.

Author Contributions: Conceptualization, F.Z.; Methodology, Z.X. and F.Z.; Software, K.L.; Validation, Z.X. and Y.W.; Formal analysis, Z.X. and Y.W.; Investigation, Z.X., Y.F. and Y.W.; Resources, Y.F.; Data curation, K.L.; Writing—original draft, Z.X.; Writing—review & editing, Y.F., K.L., L.H. and F.Z.; Supervision, F.Z.; Project administration, L.H. and F.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Senior and Junior Technological Innovation Team (20210509055RQ), Guizhou Provincial Science and Technology Projects (ZK2023-297), the Science and Technology Foundation of Health Commission of Guizhou Province (gzwkj2023-565), Science and Technology Project of Education Department of Jilin Province (JJKH20220245KJ and JJKH20220226SK),

the National Natural Science Foundation of China (U19A2061), the Jilin Provincial Key Laboratory of Big Data Intelligent Computing (20180622002JC), and the Fundamental Research Funds for the Central Universities, JLU.

Data Availability Statement: The datasets used in our study are publicly available. The MDD dataset used can be obtained by applying at <https://rfmri.org/REST-meta-MDD>, while the MCI dataset can be downloaded from <https://adni.loni.usc.edu/>. Polished translation: The datasets used in our study are publicly available. The MDD dataset can be obtained by applying at <https://rfmri.org/REST-meta-MDD>, and the MCI dataset can be downloaded from <https://adni.loni.usc.edu/>.

Acknowledgments: The insightful suggestions from the three anonymous reviewers have substantially improved the experimental convincingness and the readability of this work and are greatly appreciated.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Yazdkhasti, F. A study on cognitive and depressive disorders in the elderly and their correlation with each other. *Adv. Cogn. Sci.* **2010**, *11*, 12–17.
2. Kalibatseva, Z.; Leong, F.T. Depression among Asian Americans: Review and recommendations. *Depress. Res. Treat.* **2011**, *2011*, 320902. [[CrossRef](#)] [[PubMed](#)]
3. Bordini, B.J.; Stephany, A.; Kliegman, R. Overcoming diagnostic errors in medical practice. *J. Pediatr.* **2017**, *185*, 19–25.e11. [[CrossRef](#)] [[PubMed](#)]
4. Bayareh-Mancilla, R.; Medina-Ramos, L.A.; Toriz-Vázquez, A.; Hernández-Rodríguez, Y.M.; Cigarroa-Mayorga, O.E. Automated Computer-Assisted Medical Decision-Making System Based on Morphological Shape and Skin Thickness Analysis for Asymmetry Detection in Mammographic Images. *Diagnostics* **2023**, *13*, 3440. [[CrossRef](#)]
5. Medina-Ramos, L.A.; Toriz-Vázquez, A.; Hernández-Rodríguez, Y.M.; Bayareh-Mancilla, R.; Cigarroa-Mayorga, O.E. Automated Segmentation of Breast Skin for Early Cancer Diagnosis: A Multi-otsu Region Growing Approach for Detecting Skin Thickness Variations. In *Proceedings of Congreso Nacional de Ingeniería Biomédica*; Springer Nature: Cham, Switzerland, 2023; pp. 211–221.
6. Sen, B.; Mueller, B.; Klimes-Dougan, B.; Cullen, K.; Parhi, K.K. Classification of major depressive disorder from resting-state fMRI. In *Proceedings of the 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Berlin, Germany, 23–27 July 2019; pp. 3511–3514.
7. Sen, B.; Cullen, K.R.; Parhi, K.K. Classification of adolescent major depressive disorder via static and dynamic connectivity. *IEEE J. Biomed. Health Inform.* **2020**, *25*, 2604–2614. [[CrossRef](#)]
8. Dai, P.; Xiong, T.; Zhou, X.; Ou, Y.; Li, Y.; Kui, X.; Chen, Z.; Zou, B.; Li, W.; Huang, Z. The alterations of brain functional connectivity networks in major depressive disorder detected by machine learning through multisite rs-fMRI data. *Behav. Brain Res.* **2022**, *435*, 114058. [[CrossRef](#)]
9. Kaiser, R.H.; Whitfield-Gabrieli, S.; Dillon, D.G.; Goer, F.; Beltzer, M.; Minkel, J.; Smoski, M.; Dichter, G.; Pizzagalli, D.A. Dynamic resting-state functional connectivity in major depression. *Neuropsychopharmacology* **2016**, *41*, 1822–1830. [[CrossRef](#)]
10. Long, Y.; Cao, H.; Yan, C.; Chen, X.; Li, L.; Castellanos, F.X.; Bai, T.; Bo, Q.; Chen, G.; Chen, N. Altered resting-state dynamic functional brain networks in major depressive disorder: Findings from the REST-meta-MDD consortium. *NeuroImage Clin.* **2020**, *26*, 102163. [[CrossRef](#)]
11. Okada, G.; Yoshioka, T.; Yamashita, A.; Itai, E.; Yokoyama, S.; Kamishikiryo, T.; Shinzato, H.; Masuda, Y.; Mitsuyama, Y.; Kan, S. Verification of the brain network marker of major depressive disorder: Test-retest reliability and anterograde generalization performance for newly acquired data. *J. Affect. Disord.* **2023**, *326*, 262–266. [[CrossRef](#)]
12. Preti, M.G.; Bolton, T.A.; Van De Ville, D. The dynamic functional connectome: State-of-the-art and perspectives. *Neuroimage* **2017**, *160*, 41–54. [[CrossRef](#)]
13. Zheng, Y.; Chen, X.; Li, D.; Liu, Y.; Tan, X.; Liang, Y.; Zhang, H.; Qiu, S.; Shen, D. Treatment-naive first episode depression classification based on high-order brain functional network. *J. Affect. Disord.* **2019**, *256*, 33–41. [[CrossRef](#)] [[PubMed](#)]
14. Vergun, S.; Deshpande, A.S.; Meier, T.B.; Song, J.; Tudorascu, D.L.; Nair, V.A.; Singh, V.; Biswal, B.B.; Meyerand, M.E.; Birn, R.M. Characterizing functional connectivity differences in aging adults using machine learning on resting state fMRI data. *Front. Comput. Neurosci.* **2013**, *7*, 38. [[CrossRef](#)] [[PubMed](#)]
15. Lei, B.; Zhao, Y.; Huang, Z.; Hao, X.; Zhou, F.; Elazab, A.; Qin, J.; Lei, H. Adaptive sparse learning using multi-template for neurodegenerative disease diagnosis. *Med. Image Anal.* **2020**, *61*, 101632. [[CrossRef](#)] [[PubMed](#)]
16. Shi, Y.; Zhang, L.; Wang, Z.; Lu, X.; Wang, T.; Zhou, D.; Zhang, Z. Multivariate machine learning analyses in identification of major depressive disorder using resting-state functional connectivity: A multicenter study. *ACS Chem. Neurosci.* **2021**, *12*, 2878–2886. [[CrossRef](#)] [[PubMed](#)]
17. Wang, X.; Ren, Y.; Zhang, W. Depression disorder classification of fMRI data using sparse low-rank functional brain network and graph-based features. *Comput. Math. Methods Med.* **2017**, *2017*, 3609821. [[CrossRef](#)] [[PubMed](#)]

18. Pitsik, E.N.; Maximenko, V.A.; Kurkin, S.A.; Sergeev, A.P.; Stoyanov, D.; Paunova, R.; Kandilarova, S.; Simeonova, D.; Hramov, A.E. The topology of fMRI-based networks defines the performance of a graph neural network for the classification of patients with major depressive disorder. *Chaos Solitons Fractals* **2023**, *167*, 113041. [[CrossRef](#)]
19. Zhang, S.; Yang, J.; Zhang, Y.; Zhong, J.; Hu, W.; Li, C.; Jiang, J. The Combination of a Graph Neural Network Technique and Brain Imaging to Diagnose Neurological Disorders: A Review and Outlook. *Brain Sci.* **2023**, *13*, 1462. [[CrossRef](#)] [[PubMed](#)]
20. Kim, S.Y. Personalized Explanations for Early Diagnosis of Alzheimer’s Disease Using Explainable Graph Neural Networks with Population Graphs. *Bioengineering* **2023**, *10*, 701. [[CrossRef](#)]
21. Liu, L.; Wen, G.; Cao, P.; Hong, T.; Yang, J.; Zhang, X.; Zaiane, O.R. BrainTGL: A dynamic graph representation learning model for brain network analysis. *Comput. Biol. Med.* **2023**, *153*, 106521. [[CrossRef](#)]
22. Wang, J.; Li, T.; Sun, Q.; Guo, Y.; Yu, J.; Yao, Z.; Hou, N.; Hu, B. Automatic Diagnosis of Major Depressive Disorder Using a High- and Low-Frequency Feature Fusion Framework. *Brain Sci.* **2023**, *13*, 1590. [[CrossRef](#)]
23. Venkatapathy, S.; Votinov, M.; Wagels, L.; Kim, S.; Lee, M.; Habel, U.; Ra, I.-H.; Jo, H.-G. Ensemble graph neural network model for classification of major depressive disorder using whole-brain functional connectivity. *Front. Psychiatry* **2023**, *14*, 1125339. [[CrossRef](#)] [[PubMed](#)]
24. Zhao, T.; Zhang, G. *Detecting Major Depressive Disorder by Graph Neural Network Exploiting Resting-State Functional MRI*; Springer: Singapore, 2022; pp. 255–266.
25. Dosovitskiy, A.; Beyer, L.; Kolesnikov, A.; Weissenborn, D.; Zhai, X.; Unterthiner, T.; Dehghani, M.; Minderer, M.; Heigold, G.; Gelly, S. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv* **2020**, arXiv:2010.11929.
26. Yan, C.-G.; Chen, X.; Li, L.; Castellanos, F.X.; Bai, T.-J.; Bo, Q.-J.; Cao, J.; Chen, G.-M.; Chen, N.-X.; Chen, W. Reduced default mode network functional connectivity in patients with recurrent major depressive disorder. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 9078–9083. [[CrossRef](#)] [[PubMed](#)]
27. Yao, D.; Yang, E.; Guan, H.; Sui, J.; Zhang, Z.; Liu, M. Tensor-Based Multi-index Representation Learning for Major Depression Disorder Detection with Resting-State fMRI. In Proceedings of the Medical Image Computing and Computer Assisted Intervention—MICCAI 2021: 24th International Conference, Strasbourg, France, 27 September–1 October 2021; Proceedings, Part V 24. pp. 174–184.
28. Yan, C.-G.; Wang, X.-D.; Zuo, X.-N.; Zang, Y.-F. DPABI: Data processing & analysis for (resting-state) brain imaging. *Neuroinformatics* **2016**, *14*, 339–351. [[CrossRef](#)] [[PubMed](#)]
29. Yan, C.-G.; Zang, Y.-F. DPARSF: A MATLAB Toolbox for “Pipeline” Data Analysis of Resting-State fMRI. *Front. Syst. Neuroscience* **2010**, *4*, 13. [[CrossRef](#)] [[PubMed](#)]
30. Gao, H.; Ji, S. Graph u-nets. In Proceedings of the International Conference on Machine Learning, Long Beach, CA, USA, 9–15 June 2019; pp. 2083–2092.
31. Han, K.; Wang, Y.; Chen, H.; Chen, X.; Guo, J.; Liu, Z.; Tang, Y.; Xiao, A.; Xu, C.; Xu, Y. A survey on vision transformer. *IEEE Trans. Pattern Anal. Mach. Intell.* **2022**, *45*, 87–110. [[CrossRef](#)] [[PubMed](#)]
32. Langs, G.; Golland, P.; Ghosh, S.S. Predicting activation across individuals with resting-state functional connectivity based multi-atlas label fusion. In Proceedings of the Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015: 18th International Conference, Munich, Germany, 5–9 October 2015; Proceedings, Part II 18. pp. 313–320.
33. Long, Z.; Li, J.; Liao, H.; Deng, L.; Du, Y.; Fan, J.; Li, X.; Miao, J.; Qiu, S.; Long, C. A Multi-Modal and Multi-Atlas Integrated Framework for Identification of Mild Cognitive Impairment. *Brain Sci.* **2022**, *12*, 751. [[CrossRef](#)] [[PubMed](#)]
34. Kennedy, D.N.; Lange, N.; Makris, N.; Bates, J.; Meyer, J.; Caviness, V.S., Jr. Gyri of the human neocortex: An MRI-based analysis of volume and variance. *Cerebral Cortex* **1998**, *8*, 372–384. [[CrossRef](#)]
35. Tzourio-Mazoyer, N.; Landeau, B.; Papathanassiou, D.; Crivello, F.; Etard, O.; Delcroix, N.; Mazoyer, B.; Joliot, M. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* **2002**, *15*, 273–289. [[CrossRef](#)]
36. Craddock, R.C.; James, G.A.; Holtzheimer III, P.E.; Hu, X.P.; Mayberg, H.S. A whole brain fMRI atlas generated via spatially constrained spectral clustering. *Hum. Brain Mapp.* **2012**, *33*, 1914–1928. [[CrossRef](#)]
37. Jie, B.; Liu, M.; Lian, C.; Shi, F.; Shen, D. Designing weighted correlation kernels in convolutional neural networks for functional connectivity based brain disease diagnosis. *Med. Image Anal.* **2020**, *63*, 101709. [[CrossRef](#)] [[PubMed](#)]
38. Gu, L.; Huang, L.; Yin, F.; Cheng, Y. Classification of depressive disorder based on rs-fMRI using multivariate pattern analysis with multiple features. In Proceedings of the 2017 4th IAPR Asian Conference on Pattern Recognition (ACPR), Nanjing, China, 26–29 November 2017; pp. 61–66.
39. Li, M.; Das, T.; Deng, W.; Wang, Q.; Li, Y.; Zhao, L.; Ma, X.; Wang, Y.; Yu, H.; Li, X. Clinical utility of a short resting-state MRI scan in differentiating bipolar from unipolar depression. *Acta Psychiatr. Scand.* **2017**, *136*, 288–299. [[CrossRef](#)] [[PubMed](#)]
40. Jie, N.-F.; Zhu, M.-H.; Ma, X.-Y.; Osuch, E.A.; Wammes, M.; Théberge, J.; Li, H.-D.; Zhang, Y.; Jiang, T.-Z.; Sui, J. Discriminating bipolar disorder from major depression based on SVM-FoBa: Efficient feature selection with multimodal brain imaging data. *IEEE Trans. Auton. Ment. Dev.* **2015**, *7*, 320–331. [[CrossRef](#)] [[PubMed](#)]
41. Guo, W.; Liu, F.; Dai, Y.; Jiang, M.; Zhang, J.; Yu, L.; Long, L.; Chen, H.; Gao, Q.; Xiao, C. Decreased interhemispheric resting-state functional connectivity in first-episode, drug-naive major depressive disorder. *Prog. Neuro-Psychopharmacol. Biol. Psychiatry* **2013**, *41*, 24–29. [[CrossRef](#)] [[PubMed](#)]

42. Zhang, M.; Long, D.; Chen, Z.; Fang, C.; Li, Y.; Huang, P.; Chen, F.; Sun, H. Multi-view graph network learning framework for identification of major depressive disorder. *Comput. Biol. Med.* **2023**, *166*, 107478. [[CrossRef](#)]
43. Zhu, M.; Quan, Y.; He, X. The classification of brain network for major depressive disorder patients based on deep graph convolutional neural network. *Front. Hum. Neurosci.* **2023**, *17*, 1094592. [[CrossRef](#)]
44. Fan, L.; Zhu, X.; Borenstein, A.; Huang, X.; Shrubsole, M.; Dugan, L.; Dai, Q. Association of Circulating Caprylic Acid with Risk of Mild Cognitive Impairment and Alzheimer's Disease in the Alzheimer's Disease Neuroimaging Initiative (ADNI) Cohort. *J. Prev. Alzheimer's Dis.* **2023**, *10*, 513–522. [[CrossRef](#)]
45. Yang, P.; Zhou, F.; Ni, D.; Xu, Y.; Chen, S.; Wang, T.; Lei, B. Fused Sparse Network Learning for Longitudinal Analysis of Mild Cognitive Impairment. *IEEE Trans. Cybern.* **2021**, *51*, 233–246. [[CrossRef](#)]
46. Bondi, E.; Maggioni, E.; Brambilla, P.; Delvecchio, G. A systematic review on the potential use of machine learning to classify major depressive disorder from healthy controls using resting state fMRI measures. *Neurosci. Biobehav. Rev.* **2023**, *144*, 104972. [[CrossRef](#)]
47. Sijtsma, M.; Marjoram, D.; Gallagher, H.L.; Grealy, M.A.; Brennan, D.; Mathias, C.; Cavanagh, J.; Pollick, F.E. Major Depression and the Perception of Affective Instrumental and Expressive Gestures: An fMRI Investigation. *Psychiatry Res. Neuroimaging* **2023**, *336*, 111728. [[CrossRef](#)]
48. El-gayar, M.M.; Soliman, H.; meky, N. A comparative study of image low level feature extraction algorithms. *Egypt. Inform. J.* **2013**, *14*, 175–181. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.