Multi-Granularity Graph Convolution Network for Major Depressive Disorder Recognition

Xiaofang Sun, *Student Member, IEEE*, Yonghui Xu, Yibowen Zhao, Xiangwei Zheng, Yongqing Zheng, and Lizhen Cui, *Senior Member, IEEE*

***Abstract*— Major depressive disorder (MDD) is the most common psychological disease. To improve the recognition accuracy of MDD, more and more machine learning methods have been proposed to mining EEG features, i.e. typical brain functional patterns and recognition methods that are closely related to depression using resting EEG signals. Most existing methods typically utilize threshold methods to filter weak connections in the brain functional connectivity network (BFCN) and construct quantitative statistical features of brain function to measure the BFCN. However, these thresholds may excessively remove weak connections with functional relevance, which is not conducive to discovering potential hidden patterns in weak connections. In addition, statistical features cannot describe the topological structure features and information network propagation patterns of the brain's different functional regions. To solve these problems, we propose a novel MDD recognition method based on a multi-granularity graph convolution network (MGGCN). On the one hand, this method applies multiple sets of different thresholds to build a multi-granularity functional neural network, which can remove noise while fully retaining valuable weak connections. On the other hand, this method utilizes graph neural network to learn the topological structure features and brain saliency patterns of changes between brain functional regions on the multi-granularity functional neural network. Experimental results on the benchmark datasets validate the superior performance and time complexity of MGGCN. The analysis shows that as the granularity increases, the connectivity defects in the right frontal(RF) and right temporal (RT) regions, left temporal(LT) and left posterior(LP) regions increase. The brain functional connections in these regions can serve as potential biomarkers for MDD recognition.**

***Index Terms*— Major depressive disorder recognition, multi-granularity, graph convolution network, EEG.**

Manuscript received 17 March 2023; revised 14 July 2023 and 9 August 2023; accepted 26 August 2023. Date of publication 4 September 2023; date of current version 1 February 2024. This work was supported in part by the Taishan Scholar Program of Shandong Province of China under Grant tstp20221108, in part by the Shandong Provincial Natural Science Foundation under Grant ZR2022QF018, in part by the Shandong Provincial Outstanding Youth Science Foundation under Grant 2023HWYQ-039, and in part by the Fundamental Research Funds of Shandong University. *(Corresponding authors: Yonghui Xu; Lizhen Cui.)*

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Medical Ethics Committee of Qilu Hospital of Shandong University. The research was conducted in accordance with the principles embodied in the Declaration of Helsinki and in accordance with local statutory requirements.

Xiaofang Sun, Yonghui Xu, Yibowen Zhao, Yongqing Zheng, and Lizhen Cui are with the Joint SDU-NTU Centre for Artificial Intelligence Research (C-FAIR) and the School of Software, Shandong University, Jinan, Shandong 250000, China (e-mail: xiaofangsun2019@163.com; xu.yonghui@hotmail.com; ybw.zhao@mail.sdu.edu.cn; zhengyongqing@dareway.com.cn; clz@sdu.edu.cn).

Xiangwei Zheng is with the School of Information Science and Engineering, Shandong Normal University, Jinan, Shandong 250000, China, and also with the Shandong Provincial Key Laboratory for Distributed Computer Software Novel Technology, Jinan, Shandong 250000, China (e-mail: xwzhengcn@163.com).

Digital Object Identifier 10.1109/TNSRE.2023.3311458

I. Introduction

MAJOR depressive disorder (MDD) is a widespread emotional disorder, which can lead to various psychological, physiological and cognitive disorders [1]. The traditional diagnosis include the hamilton depression rating scale-17 (HAMD-17) [2], [3], [4], montgomery asperger depression rating scale (MADRS) [5], patient health questionnaire-9 (PHQ-9) [6]. The above scales require a lot of domain knowledge, professional inquiry and filling in time. The diagnosis of these traditional MDD is in urgent need of innovation. The recent development of artificial intelligence (AI) has brought new opportunities for the innovation of MDD diagnostic techniques [7], [8], [9], [10].

Early researchers try to utilize traditional machine learning methods, i.e. support vector machine (SVM) [11], [12], [13], [14], naive bayesian (NB) [15], [16], random forest (RF) [17], k-nearest neighbor (KNN) [18], logistic regression (LR) [19], [20] and decision tree (DT) [21] for MDD recognition. Among these methods, feature extraction is crucial for the accuracy improvement of MDD recognition. However, it is worth noting that existing MDD recognition methods may not fully capture the interaction between brain regions and brain saliency patterns when extracting static statistical features from the BFCN. Additionally, these methods may inadvertently remove weak connections that possess functional relevance.

Recently, some studies [22], [23], [24] utilize deep neural networks to explore the relationship between the EEG channel location and depression, e.g. Seal et al. propose a deep convolution neural network framework called deep learning-based convolution neural network (CNN) (DeprNet) for detecting depression using EEG [25]. To understand which channels are related to depression, DeprNet applies different spatial coordinate systems. Although DeprNet explore the relationship

between spatial channels and MDD from the perspective of coordinate system, it did not explore the relationship between spatial location and brain maps of MDD. To learn the relationship between MDD brain map and spatial location area, Liu et al. propose a signal preprocessing method by converting EEG signals into brain maps, and propose a low complexity network architecture (one layer CNN and one layer GRU) to recognize MDD and normal control (NC) [26]. Similarly, Goswami et al. also pay attention to the relationship between the time, frequency and space of EEG signals with depression by converting EEG activity into a series of topologically preserving multi-spectral images, and propose a method to recognize MDD from multi-channel EEG [27]. However, [26] and [27] both ignore some long-distance brain connections in the steps of spatial location. Simply defining neighbors according to distance may lead to the loss of important information, as well as the importance of information transmission between brain regions and brain salient patterns for MDD recognition. Brain functional connection can simulate the dynamic interaction between brain neurons, so it can effectively avoid the above problems.

To learn the information transmission in the brain space and the brain core topological structure, some researchers attempt to utilize the graph neural network (GNN) to learn the brain topology and brain functional connectivity network (BFCN) for recognizing MDD, i.e. Zhu et al. utilize correlation to construct the BFCN by extracting the four linear features of EEG, namely *activity*, *mobility complexity* and *power spectral density*, apply the attention mechanism to calculate the brain connection in BFCN and input it into the GNN to recognize MDD [28]. However, there are a large number of weak connections and pseudo connections in the BFCN, which often blur the topological structure of core connections. Thus, the threshold-based binary method to solve this problem may be one of the best candidates for removing weak connections [29]. Zhang et al. construct BFCN based on phase lag index (PLI), transform BFCN into a binary brain network (BBN) using threshold with a large difference in the proportional threshold, and extract static statistical features from the BBN, such as the *average shortest path length*, *clustering coefficient*, *small world index*, etc. However, this threshold method may excessively remove weak connections with functional relevance, which is not conducive to finding hidden patterns in weak connections. In addition, static statistical features can not describe the topological structure of the brain network, brain salient patterns and information transmission between brain regions.

To solve the above problems, a MDD recognition method based on multi-granularity graph convolution network (MGGCN) is proposed. The main contributions of this paper are as follows:

1) Compared with the existing method that recognizes MDD based on static statistical features of BFCN, there is a defect in ignoring the brain core topology and the significant pattern of information transmission between brain regions. This paper proposes a multi-granularity method to continuously learn the brain network topology, and combines the brain network core topology with the GNN to recognize MDD for the first time, achieving a relatively advanced recognition effect.
2) Compared to existing methods that roughly filter weak and pseudo connections in BFCN, there is a drawback of roughly filtering out weak connections with functional correlation. This paper proposes a MDD recognition method based on a multi-granularity graph convolution network, which achieves the effect of preserving the weak connections with functional correlation while removing the noise.

The organizational structure of this paper is as follows. Section II introduces the MDD recognition framework based on MGGCN; Section III introduces the relevant information of the datasets applied in this paper, as well as the specific steps and results of the experiment; Section IV is a brief summary of research work.

II. Methods

A. Overview of MGGCN Model

To learn the core topological structure of the brain network and the brain salience pattern in depth, remove the noise in the BFCN while preserving the weak connection with functional correlation, a MDD recognition based on MGGCN is proposed.

As shown in Fig.1, we preprocess the original resting state EEG signal; Secondly, the wavelet packet transform (WPT) has a high time-frequency resolution, so as to utilize WPT to decode and reconstruct the EEG signal into Delta, Theta, Alpha, Beta, and Gamma frequency bands, and retain the original undivided All frequency bands; Phase synchronization (PS) is more sensitive to phase change than other brain network construction methods, there by, apply phase locked value (PLV) and PLI to construct the BFCN; However, there are a lot of redundant and pseudo connections in the constructed BFCN. To learn the overall coordination of the brain and the core brain network topology, multi-granularity can both remove redundant connections and prevent the removal of core connections. Thus, transform BFCN into sub BBNs pyramid with different granularity; Train the sub BBNs pyramid of the original All band of MDD and NC constructed by the two methods to the proposed MGGCN. And then test the core topology recognition under different human states by learning the models of the core topology differences of different subjects' brain networks; Finally, softmax activation function is adopted to recognize MDD and NC.

The core part of MGGCN model is shown in Fig.2 and describe in detail below:

B. EEG Frequency Processing

In the preprocessing stage, considering that the length of the original EEG signal of all datasets and the experiment are easily affected by the previous stimulus, the EEG signals of the original MDD and NC are normalized, and the EEG signals of the two are intercepted, respectively. The unified experimental signal is designed for the experimental paradigm with a duration of 80 seconds, and the EEG signal is processed in segments using a time window of 10 seconds. We apply FIR

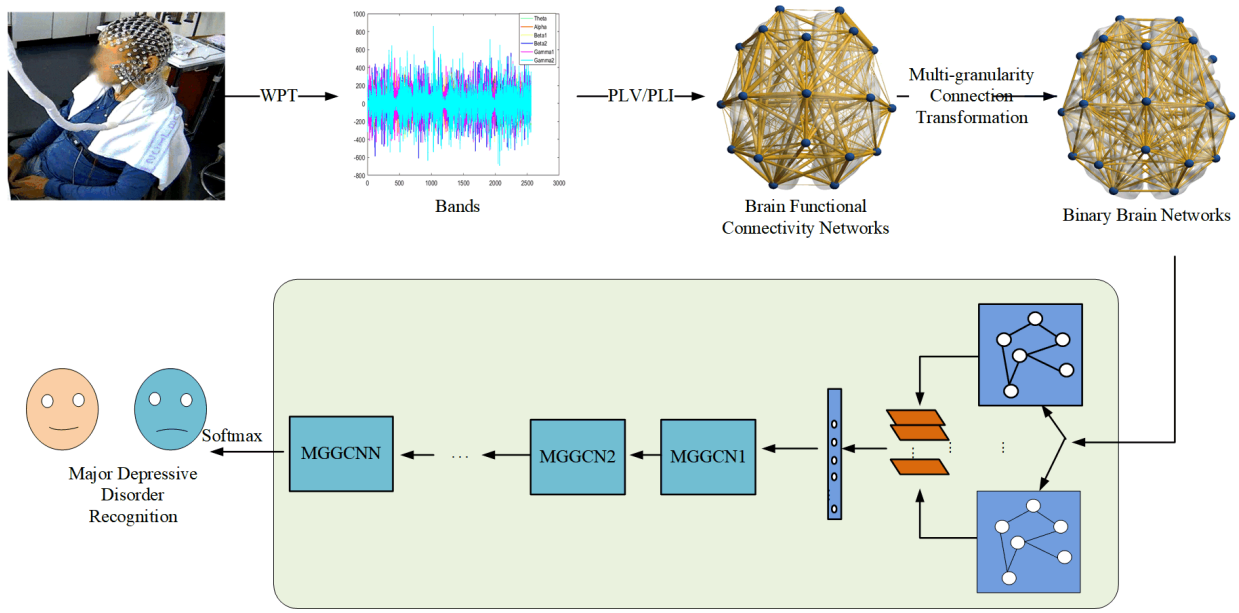

Fig. 1. Framework of multi-granularity graph convolution network.

TABLE I
WAVELET COEFFICIENT INDEX AND FREQUENCY BANDS

Bands	Wavelet coefficient	Frequency range
Delta	W6.1 - W6.3	1 Hz - 3 Hz
Theta	W6.4 - W6.7	4 Hz - 7 Hz
Alpha	W6.8 - W6.12	8 Hz - 12 Hz
Beta	W6.13 - W6.30	13 Hz - 30 Hz
Gamma	W6.31 - W6.60	31 Hz - 60 Hz
All	/	/

band-pass filter 0.5 Hz∼40 Hz to filter EEG signals and utilize REST re-reference, and FastICA to remove ocular artifacts.

WPT is utilized to convert the original EEG signal into Delta, Theta, Alpha, Beta and Gamma frequency bands. We utilize the Wavelet Toolbox in MATLAB for wavelet transform. Daubechies4 (db4) is adopted as mother wavelet to decompose EEG data. The most important point is that db4 is closer to the EEG generated by normal brain activity. Therefore, the db4 is applied to decompose the EEG signal into six layers of wavelet, and then the wavelet coefficients per 1Hz are extracted in the sixth layer.

Input EEG signal S, and the WT can be expressed as vector matrix:

$$\alpha = W^T f \tag{1}$$

among them, it contains N wavelet transform coefficients, and W is an orthogonal matrix composed of basis vectors.

$$\alpha_0 = f = \{f_0, f_1, \ldots f_{N-1}\} \tag{2}$$

$$\alpha^m = G\alpha^{m-1} \tag{3}$$

$$d^m = H\alpha^{m-1} \tag{4}$$

among them, α_0 represents the wavelet coefficient of layer 0, m represents the number of resolution layers, and the sampling interval is related to the number of resolution layers. The sampling interval of layer m is equal to 2^m. G represents the low pass filter, H represents the high pass filter, and d represents the wavelet coefficients obtained after each layer decomposition.

C. Computational Multi-Granularity BBN Pyramid

1) Construction BFCN: Secondly, the BFCNs based on PLV and PLI are constructed respectively.

The research of BFCN is an important research topic in the field of brain computer interface (BCI). The BFCN regards the EEG electrode as a node in the network, and constructs the relationship between electrodes through different methods as the weighted connection of the BFCN [30], [31]. Functional connectivity simulates the dynamic coordination and time interaction between different neurons in the brain [32], [33], [34].

Candra et al. found that the segmentation of 3-12 seconds can better describe emotions in their study on the classification of emotions based on the size of EEG segmentation [35]. Comprehensive analysis shows that a 10 second segmentation is a good choice. Haartsen et al. investigated the performance of and debiased weighted PLI (dwPLI) in measuring the reliability of EEG alpha connections with different time periods and numbers. They found that whole brain dwPLI has higher reliability in many shorter time periods (1 or 2s), while whole brain PLI has higher reliability in fewer longer time periods (5 or 6s) [36].

Similarly, Velde et al. also studied the reliability of PLI and dwPLI in BFCN [37]. They found that compared to global PLI, global dwPLI had higher inter subject variability and lower reliability. Based on the above viewpoints, we chose a longer period of 10 seconds to construct BFCN using PLI with high reliability.

In addition, Ghazani and Phan proposed a graph convolutional neural network for analyzing EEG signals, and evaluated various connection methods such as Corr, Icoh,

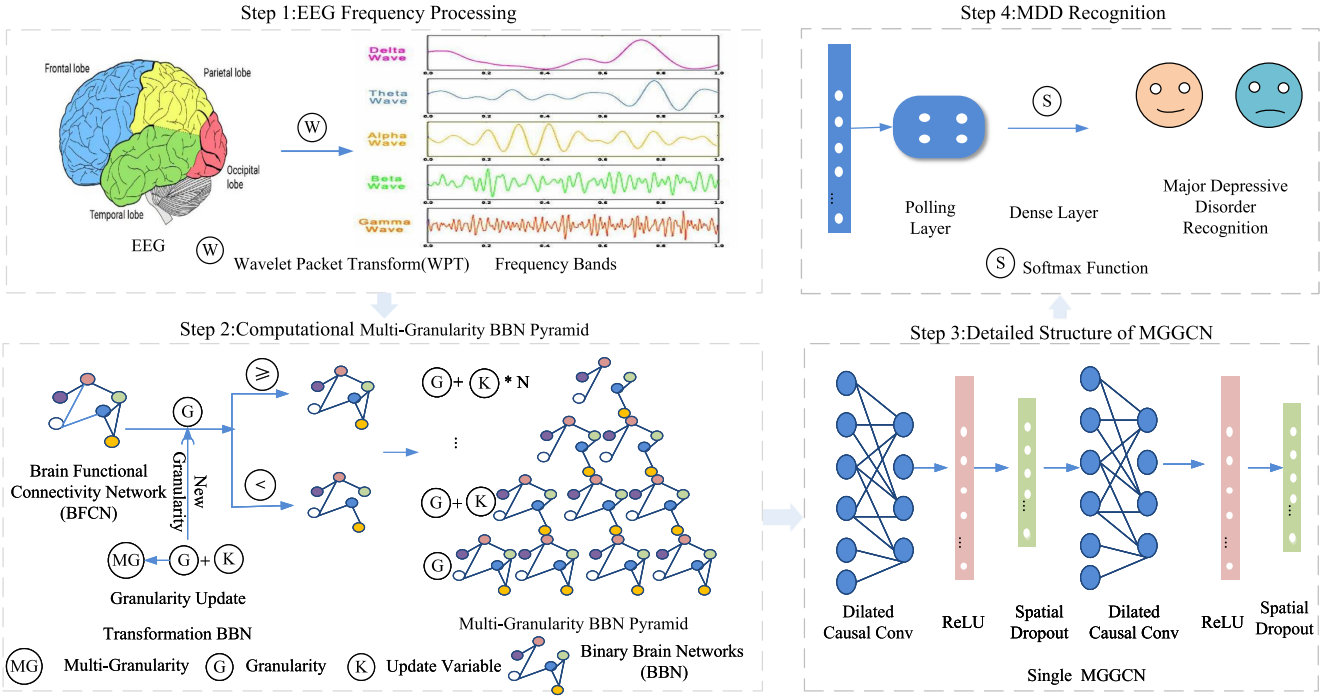

Fig. 2. Structure of multi-granularity graph convolution network based on brain network core topology.

dPLI, IPLV, PLI, wPLI and dwPLI [38]. They found that PLV outperformed other methods in estimating connectivity, while PLI method outperformed wPLI and dwPLI. Therefore, we chose PLV and PLI to construct BFCN to explore the differences in the core topology of the brain network between MDD and NC.

Each EEG signal is regarded as a node in the BFCN, and PLI and PLV are respectively utilized to calculate the edges between nodes in the BFCN.

Assuming that the EEG signals of any two electrodes are $x(t)$ and $y(t)$, the phases of these two signals at instant t are expressed as $\phi_x(t)$ and $\phi_y(t)$, respectively, and the phase change of the instant t is expressed:

$$\Delta\varphi(\mathrm{t}) = \left|\varphi_{\mathrm{x}}(\mathrm{t}) - \varphi_{\mathrm{y}}(\mathrm{t})\right| \tag{5}$$

among $\phi_x(t)$ and $\phi_y(t)$ is the expansion of EEG signals x and y under time series t.

PLV is calculated as:

$$PLV = \left|\frac{1}{N}\right| \sum_{j=0}^{N-1} e^{i\,\Delta\varphi(t)} \tag{6}$$

PLI is a method to measure the asymmetry of phase difference distribution between two signals [39].

$$PLI = |< \mathrm{sign}\left[\Delta\varphi\,(\mathrm{t})\right] >| \tag{7}$$

The asymmetry index of phase difference distribution can be obtained from the phase difference $\Delta(t)$ of t time series.

2) Transformation Multi-Granularity BBN Pyramid: So far, we have obtained the BFCN based on PLV and PLI in all bands. There are a lot of weak connections and pseudo connections caused by noise in the BFCN, which usually blurs the core topology of the BFCN.

To select a reasonable connection selection strategy, reduce the weak connection and pseudo connection of the unknown BFCN caused by noise, the multi-granularity connectivity transformation method is utilized to convert BFCN to sub BBNs pyramid. Through the transformation of different granularity, the pseudo connection can be eliminated and the core connection can be prevented from over-filtering.

Considering that the two construction methods are utilized to build the BFCN respectively, there is PS difference after phase capture, and the range of brain functional connection is different, the optimal granularity is not selected to prevent the two construction methods from having different recognition effects due to different optimal granularity and control variables [40].

Multi-granularity based functional connectivity transformation method is described in detail as follows: First, set the initial granularity and granularity update variable according to experience; Then, compare each functional connection in the BFCN with the granularity. If the functional connection is greater than the granularity, it is deemed that there is a connection between any two electrodes. Otherwise, it is deemed that there is no connection. Continue this operation until the BFCN is converted. From this, a sub BBN with a core topology is generated. Then, the granularity is updated to generate a new sub BBN, and this step is continued until the zero-valued BBN is further generated. So far, we have obtained a number of sub BBN pyramid that are constantly generated by changing multi-granularity.

With the help of GCN, the core topological structure set of brain network generated by multi-granularity is continuously learned to comprehend the brain significant pattern to recognize MDD.

D. Detailed Structure of MGGCN

To overcome above disadvantages, MGGCN is designed to remove weak connections and pseudo connections in BFCN, at the same time, it prevents filtering out the core connections and affecting the extraction of brain significant patterns.

Algorithm 1 MGGCN Model

Input: the weight of BFCN W_i, the filter F_w, the bias b and label La

Output: $\hat{Y}_t$ *and* *Loss*

// BBN Layer;
Initialization granularity M and update variable K
$A(j,k) \leftarrow W_i$
for each $BFCN_i \in BFCN$ **do**
  if $A(j,k) >= M$ **then**
    $B(j,k) \leftarrow 1$
  else
    $B(j,k) \leftarrow 0$
  end if
  $M \leftarrow M + K$
end for
$subBBN_i \leftarrow B(j,k)$
// MGGCN Layer;
for each $\hat{S}_t^{(j,l)} \in subBBN_i$ **do**
  while $(W^{(i)} \in \mathbb{R}^{F_w \times F_w}) and (b \in \mathbb{R}^{F_w})$ **do**
    $\hat{S}_t^{(j,l)} = f\left(W^{(1)} S_{t-s}^{(j,l-1)} + W^{(2)} S_t^{(j,l-1)} + b\right)$;
  end while
end for
while $Z^{(0)} \in \mathbb{R}^{F_w \times T}$ **do**
  $Z_t^{(0)} = ReLU\left(\sum_{j=1}^{B} S_t^{(j,L)}\right)$;
end while
// Pooling Layer;
$W = W^{(i)}{}_{i=1}^{F_l}$
while $(W^{(i)} \in \mathbb{R}^{d \times F_{l-1}}) and (b \in \mathbb{R}^{F_l})$ **do**
  $E^{(l)} = max\_pooling\left(f\left(W * E^{(l-1)} + b\right)\right)$;
end while
// Dense Layer;
while $(U \in \mathbb{R}^{C \times F_w}) and (c \in \mathbb{R}^{C})$ **do**
  $\hat{Y}_t = softmax\left(U Z_t^{(1)} + c\right)$;
end while
$Loss = cross\_entropy(\hat{Y}_t, La) + \alpha \|W\|_2$
return $\hat{Y}_t$ *and* *Loss*

The architecture of temporal convolution network (TCN) [41], [42] is proved to be particularly effective in predicting events with time components, such as the development of critical diseases and sequential operation of single electronic medical records [43], [44]. TCN has some advantages over other neural networks. Firstly, it is faster than the LSTM based recurrent neural network model; Secondly, TCN is better at capturing temporal dependencies, and because of the use of convolution, it can capture local information; Last but not least, the size of receptive field can be flexibly adjusted. Different granularity sequences are utilized to learn the brain network core topology, therefore, the updating sequence of different granularity is applied to design the MGGCN, and the four layers MGGCN are utilized to deeply learn the core topology of the brain network:

TABLE II
DETAILED PARAMETER LIST FOR MGGCN

Model parameter type	Parameter size
Kernel size	20
Dilation	1,2,4,8
Padding	19,38,57,76
Stride	1
Dropout	0.2
Hidden_layer	4
learning rate	0.001
batch_size	512

As shown in Fig. 3, the single MGGCN model is composed of the Dilated Causal Convolution, ReLU, Layer Norm and Spatial Dropout modules, and each block is composed of L convolution layer sequences. The result of dilated convolution at granularity t is:

$$\hat{S}_t^{(j,l)} = f\left(W^{(l)} S_{t-s}^{(j,l-1)} + W^{(2)} S_t^{(j,l-1)} + b\right) \tag{8}$$

where each layer has the same number of filters F_w, which enables us to combine activations from different layers using skip connections later. F_w size is equal to the number of channels in the experimental dataset. Each layer consists a set of dilated convolutions with rate parameter s, a non-linear activation $f(.)$ and a residual connection than combines the layer's input and the convolution signal. The filters are parameterized by $W^{(i)} \in \mathbb{R}^{F_w \times F_w}$ and bias vector $b \in \mathbb{R}^{F_w}$.

The output of each block applies a set of skip connections to sum $Z^{(0)} \in \mathbb{R}^{F_w \times T}$. It is concluded that:

$$Z_t^{(0)} = ReLU\left(\sum_{j=1}^{B} S_t^{(j,L)}\right) \tag{9}$$

where L is the number of convolutional layers.

After the sub BBN pyramid or the MGGCN of the previous state is input into the model, the dilated causal convolution is performed. The number of hidden layers is $[N\ N\ N\ N]$, and the expansion coefficient d is 2^i (i=1,2,4, ..., n, $n \leq N$). The number of input channels N is the number of EEG signal channels, the number of output channels after convolution is $N * N$. Optimizer select the Adam optimizer.

To understand the role of multi granularity in overall model MDD recognition, we conducted ablation experiments in the experimental section and conducted experiments on the benchmark dataset.

E. MDD Recognition

The pooling layer adopts maximum pooling, which reduces the size of spatial information and improves operation efficiency. Reducing spatial information can also decrease the number of parameters and the risk of overflow. Additionally, it helps to achieve invariance of translation-rotation scale.

The collection of filters in each layer are defined as $W = W^{(i)}{}_{i=1}^{Fl}$ for $W^{(i)} \in \mathbb{R}^{d \times F_{l-1}}$ with a corresponding bias vector

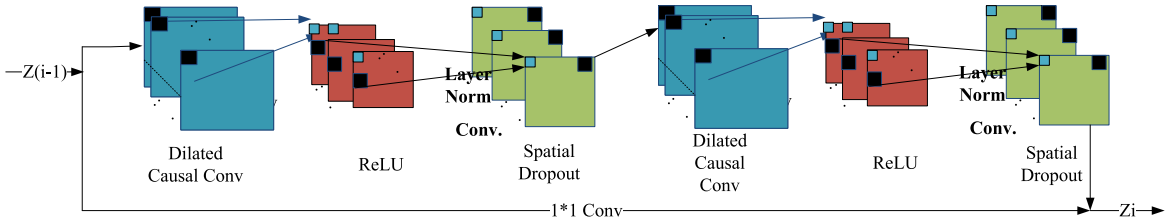

Fig. 3. Schematic diagram of single MGGCN model.

$b \in \mathbb{R}^{Fl}$. Given the signal from the previous layer $E^{(l-1)}$, we compute activations $E(l)$ with

$$E^{(l)} = max\_pooling\left(f\left(W * E^{(l-1)} + b\right)\right) \tag{10}$$

Finally, in the Dense layer, all core topology of the brain networks are integrated, and the Softmax activation function is utilized to recognize MDD and NC.

$$\hat{Y}_t = softmax\left(U Z_t^{(1)} + c\right) \tag{11}$$

where weight matrix $U \in \mathbb{R}^{C \times F_w}$ and bias $c \in \mathbb{R}^C$.

Cross_entropy is selected as the loss function of the model.

$$Loss = cross\_entropy(\hat{Y}_t, La) + \alpha \|W\|_2 \tag{12}$$

where $\hat{Y}_t$ is the predictive value of the model, La is the label, and W is all the parameters of the model, α is the coefficient of the regular term.

The cross_entropy function $cross\_entropy(\hat{Y}_t, La)$ aims to measure the difference between the actual label and the predicted value of the model, while the regularization term $\alpha\|W\|_2$ aims to reduce the over fitting probability of the model's learning parameters.

III. Experiments and Analysis

A. Materials

The public dataset MODMA is utilized to construct BFCNs to recognize MDD. 24 MDD (female/male=11/13, 30.88 ± 10.37 years old) are the experimental subjects, and 29 NC (female/male=9/20, 31.45 ± 9.15 years old) are the control group.There is no significant difference in age (t=0.214, p=0.832) and gender (x2=1.224, p=0.269) between the two groups. The experimental data are recorded continuously by hydrocel geodesic sensor network (HCGSN) with 128 channels of Cz reference at a sampling frequency of 250 Hz [45].

In dataset MPHCE, 64 subjects are recruited, including 34 MDD and 30 NC. The status of subjects is confirmed by talking with professional doctors, including men/women=38/26, 39.4 ± 14.11 years old. EEG equipment adopts 19 electrode caps, sampling frequency is 256Hz, and EEG data are collected in the closed eyes resting state, open eyes resting state and task state. After screening the dataset, the status of all data sets is consistent, the resting EEG data of 25 NC and 23 MDD are utilized [46].

In the dataset EDR, 121 subjects are recruited, including 46 MDD and 77 NC. The diagnosis method is doctor's consultation and Beck depression scale judgment, in which male/female=47/74, 18.9 ± 1.2 years old. The EEG equipment utilizes 66 lead EEG electrode caps with a sampling frequency of 500 Hz. The resting EEG data is also applied to remove some subjects with missing data and mild depression patients. 40 MDD and 67 NC are selected for MDD recognition experiment [47].

All subjects were diagnosed by doctors' dialogue, Beck Depression Inventory and other medical evaluation scales. To investigate the serious impact of depression on the brain, subjects with mild and moderate depression were excluded and subjects with major depression selected as MDD subjects.

B. Comparison Results on Benchmark Datasets

To demonstrate the impact of MGGCN in MDD recognition, we compare its performance with advanced methods using Accuracy (Acc), Precision (Prec), Recall, and F1_Score on benchmark datasets including MODMA, MPHCE, and EDR.

To provide the MGGCN with more comprehensive functional connectivity network information and facilitate visualization of BFCN, the MGGCN incorporates the complete BFCN as the feature.

We divide the dataset into 70% for model training and 30% for testing purposes. Furthermore, we perform 5-fold cross-validation. Table III shows our comparison results with traditional machine learning methods based on BFCN feature extraction and advanced GICN methods. Our MGGCN model achieves better recognition performance and lower time complexity than other methods due to the multi-granularity capture of the core topology of the brain network and the flexible receptive field of the model.

1) Result on MODMA Dataset: On the MODMA dataset, we compare the MDD recognition performance of PLI+SVM [12], PLI+PLV+SVM [13], and correlation+GICN [28] with that of MGGCN proposed by us. The analysis shows that the features fused by two BFCN construction methods in PLI+PLV+SVM [13] are better than those extracted by PLI+SVM [12] from traditional BFCN static statistical features, which supplement some complementary information so as to the MDD recognition effect is better. Correlaton+GICN [28] extracts BFCN features and utilize GCN method to recognize MDD. Compared with traditional machine learning method based PLI+SVM [12] and PLI+PLV+SVM [13], it has a better MDD recognition effect by deeply learning the relationship between brain network features and MDD. However, they all ignore the effect of redundant connections on feature extraction, as well as the role of brain core network structure and brain salient patterns. Therefore, the MDD recognition effect of the above work is lower than our MGGCN model.

TABLE III
COMPARISON OF MDD RECOGNITION EFFECT (%) AND TIME COST(S)

Dataset	Method	Acc	Prec	Recall	F1_Score	Time cost(s)
MODMA	PLI+SVM [12]	84.18±2.18	85.43±3.22	78.18±1.34	83.83±1.81	211.87±0.41
	PLI+PLV+SVM [13]	92.86±3.92	91.31±9.02	91.94±2.50	91.62±1.54	250.82±0.11
	Correlation+GICN [28]	95.78±2.25	93.55±1.97	96.72±1.02	95.11±1.51	243.93±5.45
	PLI/PLV+MGGCN(Ours)	**99.68±0.43**	**96.91±2.34**	**98.44±1.36**	**97.67±2.01**	**185.33±0.69**
MPHCE	PLI+SVM [12]	81.06±1.48	78.74±4.66	79.28±4.21	73.17±2.88	58.99±0.73
	PLI+PLV+SVM [13]	93.38±4.84	92.71±8.13	93.37±5.46	93.03±5.96	72.38±0.76
	Correlation+GICN [28]	88.91±4.50	79.21±7.95	88.91±4.50	83.74±6.46	61.13±1.34
	PLI/PLV+MGGCN(Ours)	**98.44±0.78**	**96.30±2.32**	**98.13±1.18**	**97.20±1.76**	**18.27±0.78**
EDR	PLI+SVM [12]	86.83±1.86	86.20±2.63	86.48±2.11	86.15±1.98	109.62±0.14
	PLI+PLV+SVM [13]	92.58±1.10	94.92±0.46	93.91±0.51	95.77±0.34	136.40±5.27
	Correlation+GICN [28]	95.63±7.96	91.47±3.76	95.62±1.96	93.50±2.90	118.15±0.71
	PLI/PLV+MGGCN(Ours)	**96.72±1.50**	**96.91±1.87**	**98.44±0.96**	**97.66±1.42**	**105.97±0.42**

2) Result on MPHCE Dataset: On the MPHCE dataset, the analysis shows that the performance comparison results of PLI+SVM [12] and PLI+PLV+SVM [13] are consistent with the result of MODMA dataset, but the MDD recognition effect of Correlation+GICN [28] is weaker than that of PLI+PLV+SVM [13]. The reason may be that the number of EEG channels in MPHCE dataset is 19, and SVM has high robustness in small data samples. However, our MGGCN still has some advantages, because our model has flexible receptive fields, which can adjust the input range of causal dilation convolution and can flexibly learn the core topology of the brain network and brain salient patterns with different data sizes.

3) Result on EDR Dataset: On the EDR dataset, the analysis shows that the comparison results of PLI+SVM [12] and MDD recognition of other works are consistent. However, in terms of Acc and Recall, the MDD recognition effect of Correlation+GICN [28] is stronger than that of PLI+PLV+SVM [13]. In terms of Prec and F1_Score, the MDD recognition effect of Correlation+GICN [28] is weaker than that of PLI+PLV+SVM [13]. This may be because the EEG channel number of EDR dataset is 64 and Correlation+GICN [28] is not sensitive to false negative. Our MGGCN still has some advantages, which shows that our MGGCN model is relatively stable, and the flexible local receptive field can ensure the MDD recognition effect of the model.

It can be seen from the synthesis that our MGGCN model is superior to the other three models which based on statistical features in BFCN. Because of the following reasons: we utilize multi-granularity to learn the core topology differences and brain salient patterns between MDD and NC in the brain network; The MGGCN model has flexible local receptive fields. Thus, MGGCN model has advantages in MDD recognition effect and model generalization.

C. Ablation Study

To understand the role of multi-granularity in learning the core topology of the brain network, we adopt the three public datasets of MODMA, MPHCE and EDR to measure the MDD recognition effect of single BFCN and multi-granularity The results of ablation study are shown in Table IV, Table V and Table VI.

1) Result on MODMA Dataset: As shown in Table IV, we compare the MDD recognition effect of a single BFCN based on PLV and PLI with that of a multi-granularity. Firstly, we fix the construction method, and detect that multi-granularity is superior to a single BFCN in most metrics of Delta, Theta, Alpha, Beta, Gamma frequency bands and Aver effect. However, the multi-granularity is weaker than that of a single BFCN on the Prec and F1_Score of Theta band in PLV. The reason may be that MGGCN is more sensitive to false negative rate. In terms of Prec, recall and F1_Score metrics on the Delta band, the multi-granularity is weaker than that of a single BFCN in PLI. The reason may be that redundant connection in BFCN leads to better recognition effect. In general, MGGCN utilizes multi-granularity filtering redundant connections to learn the core topology and brain salient patterns, which is better than a single BFCN.

Secondly, fixing a single BFCN or multi-granularity, we detect that a single BFCN based on PLV has stronger Prec, Recall, and F1_Score than PLI on the Beta band in single BFCN; The four metrics of PLV in the Delta band are stronger than those of PLI in multi-granularity. However, PLI are stronger than PLI in its other band and Aver effect. The reason may be that different construction methods have different PS capture sensitivity in different frequency bands. In general, the MDD recognition effect of PLI is better than that of PLV.

2) Result on MPHCE Dataset: As shown in Table V, firstly, we fix the construction method and detect that the four metrics of a single BFCN in the Delta band and the Acc in the Theta band are stronger than multi-granularity in PLV; Prec, Recall, F1_Score of a single BFCN in the Delta band and Prec, Recall, F1_Score of the Beta band are stronger than multi-granularity in PLI. However, the multiple metrics of other frequency bands and Aver effect of multi-granularity are better. Secondly, fixing a single BFCN or multi-granularity, we can detect that PLV

TABLE IV
COMPARISON OF MDD RECOGNITION EFFECTS (%) OF MODMA DATASET

Methods			Delta	Theta	Alpha	Beta	Gamma	Aver
Single BFCN	PLV-BFCN	Acc	97.66	89.84	94.53	93.75	94.53	94.06±2.80
		Prec	86.43	87.89	83.55	92.34	89.36	87.91±3.28
		Recall	92.97	93.75	91.41	96.09	94.53	93.75±1.74
		F1_Score	89.58	90.73	87.30	94.18	91.87	90.73±2.56
	PLI-BFCN	Acc	98.44	97.66	96.88	96.88	94.53	**96.88±1.46**
		Prec	93.85	90.84	87.89	89.36	90.84	**90.56±2.21**
		Recall	96.88	95.31	93.75	94.53	95.31	**95.16±1.16**
		F1_Score	95.34	93.03	90.73	91.87	93.03	**92.80±1.71**
Multi-granularity	PLV-BBN	Acc	100.00	99.22	100.00	99.22	99.22	99.53±0.43
		Prec	93.85	86.43	90.84	96.90	95.37	92.68±4.15
		Recall	96.88	92.97	95.31	98.44	97.66	96.25±2.17
		F1_Score	95.34	89.58	93.03	97.66	96.50	94.42±3.20
	PLI-BBN	Acc	99.22	99.22	100.00	100.00	100.00	**99.69±0.43**
		Prec	92.34	98.44	96.90	98.44	98.44	**96.91±2.64**
		Recall	96.09	99.22	98.44	99.22	99.22	**98.44±1.36**
		F1_Score	94.18	98.83	97.66	98.83	98.83	**97.67±2.01**

TABLE V
COMPARISON OF MDD RECOGNITION EFFECTS (%) OF MPHCE DATASET

Methods			Delta	Theta	Alpha	Beta	Gamma	Aver
Single BFCN	PLV-BFCN	Acc	95.34	98.44	94.53	96.09	95.31	95.94±4.50
		Prec	93.85	92.34	95.37	84.99	87.89	90.90±4.32
		Recall	96.88	96.09	97.66	92.19	93.75	95.31±2.28
		F1_Score	95.34	94.18	96.5	88.44	90.73	93.04±3.36
	PLI-BFCN	Acc	96.09	96.88	99.22	97.66	96.09	**97.19±1.31**
		Prec	95.37	90.84	93.85	95.37	96.9	**94.47±2.30**
		Recall	97.66	95.31	96.88	97.66	98.44	**97.19±1.19**
		F1_Score	96.5	93.03	95.34	96.5	97.66	**95.81±1.76**
Multi-granularity	PLV-BBN	Acc	94.53	96.09	96.88	96.88	97.66	96.41±1.19
		Prec	89.36	92.34	96.9	89.36	92.34	92.06±3.09
		Recall	94.53	96.09	98.44	94.53	96.09	95.94±1.60
		F1_Score	91.87	94.18	97.66	91.87	94.18	93.95±2.37
	PLI-BBN	Acc	99.22	99.22	97.66	98.44	97.66	**98.44±0.78**
		Prec	93.85	96.9	98.44	93.85	98.44	**96.30±2.32**
		Recall	96.88	98.44	99.22	96.88	99.22	**98.13±1.18**
		F1_Score	95.34	97.66	98.83	95.34	98.83	**97.20±1.76**

based construction methods are better than PLI in Theta and Alpha frequency bands in single BFCN. However, all metrics of all bands based on PLI are better than PLV in multi-granularity. In general, the MDD recognition effect of PLI based construction methods is better than that of PLV.

3) Result on EDR Dataset: As shown in Table VI, firstly, we fix the construction method and find that the Acc of a single BFCN in the Gamma band is stronger than that of multi-granularityn in PLV; The Acc of a single BFCN in the Delta and Theta bands is stronger than that of multi-granularity in PLI. However, multi-granularity are better than that of a single BFCN in multiple metrics of the other bands and Aver effect.

Secondly, fixing a single BFCN or multiple granularity, we can detect that PLV based construction methods are better than PLI in Delta frequency band. However, all metrics of all bands based on PLI are better than PLV in multi-granularity. In general, the MDD recognition effect of PLI based construction methods is better than that of PLV.

Based on the ablation study, we can draw the following conclusions, as shown in Tables IV, V and VI. First of all, through the measurement of multiple metrics of different datasets, we can conclude that the MDD recognition effect of MGGCN based on multi-granularity is better than that of a single BFCN. By comparison, we can see that multi-granularity is eliminating the interference of redundant connections on

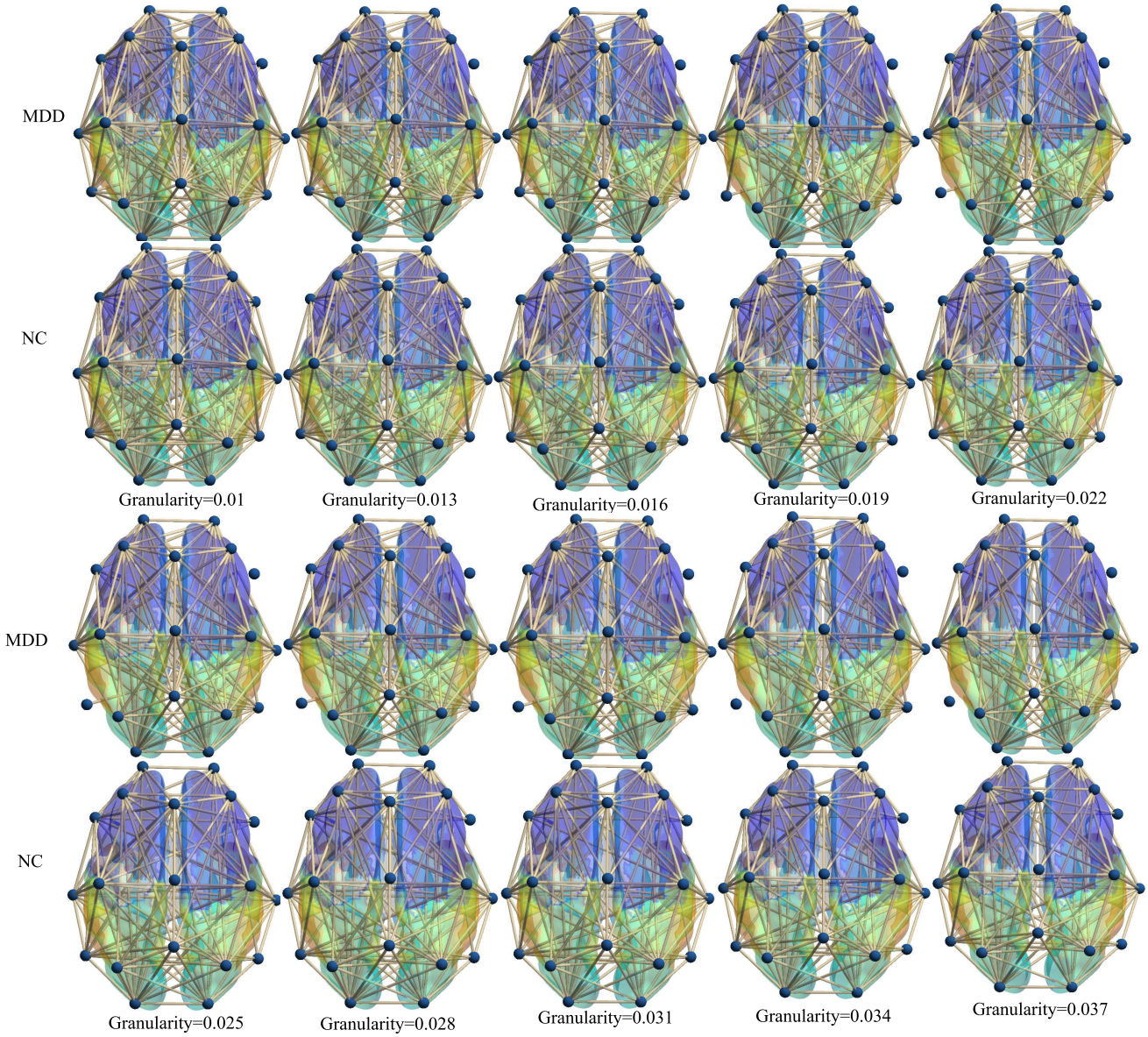

Fig. 4. Comparison of core topological structures of brain networks between MDD and NC based on multi-granularity.

feature extraction. The recognition effect of this model in multiple frequency bands and Aver effect is better than that based on a single BFCN. Secondly, fixing the single model, for example, fixing the single BFCN or multi-granularity, we can draw a conclusion that the MDD recognition effect of PLI is better than PLV, which indicates that PLI is more sensitive than PLV in capturing EEG phase changes.

D. Sensitivity Analysis

1) The Changes of Multi-Granularity: To explore the differences in brain core topology between MDD and NC, we compared and analyzed the core topology of brain networks at different granularity. Due to the different PS intensities captured by different construction methods, the same initial and update granularity is used to ensure the consistency of variables and prevent model MDD based on different construction methods from producing different recognition effects due to different granularity.

MGGCN recognizes MDD by learning the differential core topological structure of brain networks between MDD and NC. To prevent excessive filtering at multiple granularities that might exclude redundant connections with functional relevance or introduce redundant connections due to insufficient granularity, we set the initial granularity to 0.01 based on prior work experience. We then update the granularity value to 0.003 for 10 iterations (as a zero-value network might emerge with further refinement).

As shown in Fig.4, the scale area of the brain is divided into two hemispheres (left and right); The frontal region (dark blue) is divided into left frontal region (LF), right frontal region (RF); The temporal region (orange) is divided into left temporal region (LT), and right temporal region (RT); The central region(light green) is divided into left central region (LC) and right central region (RC); The posterior region (light blue) is divided into left posterior region (LP) and right posterior region(RP) [48]. Fig.4 shows the pyramid of

TABLE VI
COMPARISON OF MDD RECOGNITION EFFECTS (%) OF EDR DATASET

Methods			Delta	Theta	Alpha	Beta	Gamma	Aver
Single BFCN	PLV-BFCN	Acc	96.88	95.31	90.63	93.75	96.09	94.53±2.47
		Prec	84.99	89.36	89.36	90.84	93.85	89.68±3.20
		Recall	92.19	94.53	94.53	95.31	96.88	94.69±1.69
		F1_Score	88.44	91.87	91.87	93.03	95.34	92.11±2.49
	PLI-BFCN	Acc	96.09	97.66	96.09	94.53	96.09	**96.09±1.10**
		Prec	92.34	93.85	93.85	95.37	96.90	**94.46±1.73**
		Recall	96.09	96.88	96.88	97.66	98.44	**97.19±0.89**
		F1_Score	94.18	95.34	95.34	96.50	97.66	**95.80±1.32**
Multi-granularity	PLV-BBN	Acc	97.66	95.31	96.09	96.88	92.19	95.63±2.11
		Prec	96.90	98.44	95.37	96.90	95.37	96.60±1.28
		Recall	98.44	99.22	97.66	98.44	97.66	98.28±0.65
		F1_Score	97.66	98.83	96.50	97.66	96.50	97.43±0.97
	PLI-BBN	Acc	95.31	96.09	96.88	96.09	99.22	**96.72±1.50**
		Prec	93.85	98.44	96.90	98.44	96.90	**96.91±1.87**
		Recall	96.88	99.22	98.44	99.22	98.44	**98.44±0.96**
		F1_Score	95.34	98.83	97.66	98.83	97.66	**97.66±1.42**

sub BBNs observed from PLI in the Beta frequency bands of Subject 1 (MDD) and Subject 1 (NC) from the MPHCE dataset.

By comparing the core topological structure of the BDCN between MDD and NC at different granularity, we found that as the granularity increased, the core topological structure of the brain network became more obvious. The connection between the RF and RT regions continues to decrease, while the brain connection between LT and LP continues to decrease, resulting in increasingly obvious brain network connectivity defects. However, the core extension structure of NC's brain network does not have connectivity defects.

From this, we can conclude that the brain functional connections between RF and RT, LP and LP can serve as potential biomarkers for MDD recognition.[1]

IV. CONCLUSION AND FUTURE WORK

This paper proposes a multi-granularity graph convolution model (MGGCN) for MDD recognition. Particularly, MGGCN constructs the multi-granularity BFCN in the graph for the first time. MGGCN recognizes MDD and NC by building a multi-granularity neural network to learn the core topology of the brain network. With the advance of the multi-granularity neural network, MGGCN can retain valuable weak connections while filtering redundant connections in BFCN. Experimental results on three benchmark datasets show that MGGCN has a flexible local receptive field and strong generalization ability. In addition, we also conduct an ablation study, and the results show that the multi-granularity method has advantages over the single BFCN method. Our proposed model can accurately and quickly learn the differences in core brain network topology between MDD and NC, and thus recognize MDD. Increasing granularity shows a corresponding increase in connectivity defects in the RF and RT regions, as well as the LT and LP regions. These brain functional connections could potentially serve as biomarkers for recognizing MDD.

In future research work, due to the current lack of attention to the differences in subjects with mild, moderate, and severe depression, we tend to pay more attention to the mechanism of depression and its evolution from mild to severe. We will combine AI technology with clinical diagnosis, model the development trend of depression, and study how to control and early intervene in the development of depression.

[1]BFCN map is drawn with the help of BrainNet Viewer [49].

REFERENCES

[1] H. Choi, S. Mun, E.-J. Joo, K. Y. Lee, H.-G. Kang, and J. Lee, "Serum proteomic analysis of major depressive disorder patients and their remission status: Novel biomarker set of zinc-alpha-2-glycoprotein and keratin type II cytoskeletal 1," *Int. J. Biol. Macromol.*, vol. 183, pp. 2001–2008, Jul. 2021.

[2] M. von Glischinski, R. von Brachel, C. Thiele, and G. Hirschfeld, "Not sad enough for a depression trial? A systematic review of depression measures and cut points in clinical trial registrations," *J. Affect. Disorders*, vol. 292, pp. 36–44, Sep. 2021.

[3] C. N. Epperson et al., "Effect of brexanolone on depressive symptoms, anxiety, and insomnia in women with postpartum depression: Pooled analyses from 3 double-blind, randomized, placebo-controlled clinical trials in the HUMMINGBIRD clinical program," *J. Affect. Disorders*, vol. 320, pp. 353–359, Jan. 2023.

[4] E. Ng and P. Giacobbe, "P367. Seasonal depression does not affect depression outcomes from repetitive transcranial magnetic stimulation," *Biol. Psychiatry*, vol. 91, no. 9, pp. S235–S236, May 2022.

[5] M. Vestin, M. Åsberg, M. Wiberg, E. Henje, and I. Dennhag, "Psychometric validity of the Montgomery and Åsberg depression rating scale for youths (MADRS-Y)," *Nordic J. Psychiatry*, vol. 77, no. 5, pp. 421–431, 2022.

[6] K. Kroenke, "PHQ-9: Global uptake of a depression scale," *World Psychiatry*, vol. 20, no. 1, pp. 135–136, Feb. 2021.

[7] S. Yasin, S. A. Hussain, S. Aslan, I. Raza, M. Muzammel, and A. Othmani, "EEG based major depressive disorder and bipolar disorder detection using neural networks: A review," *Comput. Methods Programs Biomed.*, vol. 202, Apr. 2021, Art. no. 106007.

[8] J. Tang et al., "Energy-efficient sensory data collection based on spatiotemporal correlation in IoT networks," *Int. J. Crowd Sci.*, vol. 6, no. 1, pp. 34–43, Apr. 2022.

[9] Y. Liu, S. Yang, Y. Xu, C. Miao, M. Wu, and J. Zhang, "Contextualized graph attention network for recommendation with item knowledge graph," *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 1, pp. 181–195, Jan. 2023.

[10] Y. Cao et al., "KdINet: Knowledge-driven interpretable network for medical imaging diagnosis," in *Proc. IEEE Int. Conf. Bioinf. Biomed. (BIBM)*, Dec. 2022, pp. 1457–1460.

[11] S. Mahato, N. Goyal, D. Ram, and S. Paul, "Detection of depression and scaling of severity using six channel EEG data," *J. Med. Syst.*, vol. 44, no. 7, pp. 1–12, Jul. 2020.

[12] S. Sun, H. Chen, X. Shao, L. Liu, X. Li, and B. Hu, "EEG based depression recognition by combining functional brain network and traditional biomarkers," in *Proc. IEEE Int. Conf. Bioinf. Biomed. (BIBM)*, Dec. 2020, pp. 2074–2081.

[13] X. Sun, X. Zheng, Y. Xu, L. Cui, and B. Hu, "Major depressive disorder recognition and cognitive analysis based on multi-layer brain functional connectivity networks," 2021, *arXiv:2111.01351*.

[14] A. O. Khadidos, K. H. Alyoubi, S. Mahato, A. O. Khadidos, and S. N. Mohanty, "Computer aided detection of major depressive disorder (MDD) using electroencephalogram signals," *IEEE Access*, vol. 11, pp. 41133–41141, 2023.

[15] M. Čukić, M. Stokić, S. Simić, and D. Pokrajac, "The successful discrimination of depression from EEG could be attributed to proper feature extraction and not to a particular classification method," *Cogn. Neurodyn.*, vol. 14, no. 4, pp. 443–455, Aug. 2020.

[16] Y. Zhang, L. Cui, W. He, X. Lu, and S. Wang, "Behavioral data assists decisions: Exploring the mental representation of digital-self," *Int. J. Crowd Sci.*, vol. 5, no. 2, pp. 185–203, Aug. 2021.

[17] Y. Liu, C. Pu, S. Xia, D. Deng, X. Wang, and M. Li, "Machine learning approaches for diagnosing depression using EEG: A review," *Transl. Neurosci.*, vol. 13, no. 1, pp. 224–235, Aug. 2022.

[18] H. Cai, Z. Qu, Z. Li, Y. Zhang, X. Hu, and B. Hu, "Feature-level fusion approaches based on multimodal EEG data for depression recognition," *Inf. Fusion*, vol. 59, pp. 127–138, Jul. 2020.

[19] S. Liu et al., "Alterations in patients with first-episode depression in the eyes-open and eyes-closed conditions: A resting-state EEG study," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 30, pp. 1019–1029, 2022.

[20] M. S. Oduro, H. Yu, and H. Huang, "Predicting the entrepreneurial success of crowdfunding campaigns using model-based machine learning methods," *Int. J. Crowd Sci.*, vol. 6, no. 1, pp. 7–16, Apr. 2022.

[21] J. Rabcan, V. Levashenko, E. Zaitseva, and M. Kvassay, "EEG signal classification based on fuzzy classifiers," *IEEE Trans. Ind. Informat.*, vol. 18, no. 2, pp. 757–766, Feb. 2022.

[22] P. P. Thoduparambil, A. Dominic, and S. M. Varghese, "EEG-based deep learning model for the automatic detection of clinical depression," *Phys. Eng. Sci. Med.*, vol. 43, no. 4, pp. 1349–1360, Dec. 2020.

[23] G. Sharma, A. Parashar, and A. M. Joshi, "DepHNN: A novel hybrid neural network for electroencephalogram (EEG)-based screening of depression," *Biomed. Signal Process. Control*, vol. 66, Apr. 2021, Art. no. 102393.

[24] X. Song, D. Yan, L. Zhao, and L. Yang, "LSDD-EEGNet: An efficient end-to-end framework for EEG-based depression detection," *Biomed. Signal Process. Control*, vol. 75, May 2022, Art. no. 103612.

[25] A. Seal, R. Bajpai, J. Agnihotri, A. Yazidi, E. Herrera-Viedma, and O. Krejcar, "DeprNet: A deep convolution neural network framework for detecting depression using EEG," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–13, 2021.

[26] W. Liu, K. Jia, Z. Wang, and Z. Ma, "A depression prediction algorithm based on spatiotemporal feature of EEG signal," *Brain Sci.*, vol. 12, no. 5, p. 630, May 2022.

[27] A. Goswami et al., "Depression detection using spatial images of multichannel EEG data," in *Applications of Artificial Intelligence and Machine Learning: Select Proceedings of ICAAAIML*. Singapore: Springer, 2022, pp. 569–579.

[28] J. Zhu et al., "EEG based depression recognition using improved graph convolutional neural network," *Comput. Biol. Med.*, vol. 148, Sep. 2022, Art. no. 105815.

[29] B. Zhang, G. Yan, Z. Yang, Y. Su, J. Wang, and T. Lei, "Brain functional networks based on resting-state EEG data for major depressive disorder analysis and classification," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 29, pp. 215–229, 2021.

[30] E. Bullmore and O. Sporns, "Complex brain networks: Graph theoretical analysis of structural and functional systems," *Nature Rev. Neurosci.*, vol. 10, no. 3, pp. 186–198, Mar. 2009.

[31] V. Sakkalis, "Review of advanced techniques for the estimation of brain connectivity measured with EEG/MEG," *Comput. Biol. Med.*, vol. 41, no. 12, pp. 1110–1117, Dec. 2011.

[32] J. Liu, J. Ji, X. Jia, and A. Zhang, "Learning brain effective connectivity network structure using ant colony optimization combining with voxel activation information," *IEEE J. Biomed. Health Informat.*, vol. 24, no. 7, pp. 2028–2040, Jul. 2020.

[33] J.-P. Lachaux, E. Rodriguez, J. Martinerie, and F. J. Varela, "Measuring phase synchrony in brain signals," *Human Brain Mapping*, vol. 8, no. 4, pp. 194–208, 1999.

[34] Y. Dasdemir, E. Yildirim, and S. Yildirim, "Analysis of functional brain connections for positive–negative emotions using phase locking value," *Cogn. Neurodyn.*, vol. 11, no. 6, pp. 487–500, Dec. 2017.

[35] H. Candra et al., "Investigation of window size in classification of EEG-emotion signal with wavelet entropy and support vector machine," in *Proc. 37th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug. 2015, pp. 7250–7253.

[36] R. Haartsen, B. van der Velde, E. J. H. Jones, M. H. Johnson, and C. Kemner, "Using multiple short epochs optimises the stability of infant EEG connectivity parameters," *Sci. Rep.*, vol. 10, no. 1, p. 12703, Jul. 2020.

[37] B. Velde, R. Haartsen, and C. Kemner, "Test-retest reliability of EEG network characteristics in infants," *Brain Behav.*, vol. 9, no. 5, May 2019, Art. no. e01269.

[38] M. M. Ghazani and A. H. Phan, "Graph convolutional neural networks for analysis of EEG signals, BCI application," 2020, *arXiv:2006.14540*.

[39] C. J. Stam, G. Nolte, and A. Daffertshofer, "Phase lag index: Assessment of functional connectivity from multi channel EEG and MEG with diminished bias from common sources," *Hum. Brain Mapping*, vol. 28, no. 11, pp. 1178–1193, 2007.

[40] X. Sun, X. Zheng, T. Li, Y. Li, and L. Cui, "Multimodal emotion classification method and analysis of brain functional connectivity networks," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 30, pp. 2022–2031, 2022.

[41] C. Lea, M. D. Flynn, R. Vidal, A. Reiter, and G. D. Hager, "Temporal convolutional networks for action segmentation and detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 156–165.

[42] F. Zhu et al., "Temporal hypergraph for personalized clinical pathway recommendation," in *Proc. IEEE Int. Conf. Bioinf. Biomed. (BIBM)*, Dec. 2022, pp. 718–725.

[43] S. M. Lauritsen et al., "Explainable artificial intelligence model to predict acute critical illness from electronic health records," *Nature Commun.*, vol. 11, no. 1, pp. 1–11, Jul. 2020.

[44] Z. Qu, L. Cui, and Y. Xu, "Disease risk prediction via heterogeneous graph attention networks," in *Proc. IEEE Int. Conf. Bioinf. Biomed. (BIBM)*, Dec. 2022, pp. 3385–3390.

[45] H. Cai et al., "A multi-modal open dataset for mental-disorder analysis," *Sci. Data*, vol. 9, no. 1, pp. 1–10, Apr. 2022.

[46] W. Mumtaz, "MDD patients and healthy controls EEG data (new)," Figshare, Dataset, 2016. [Online]. Available: https://doi.org/10.6084/m9.figshare.4244171.v2

[47] J. F. Cavanagh, "EEG: Depression rest," OpenNeuro, [Dataset], 2021, doi: 10.18112/openneuro.ds003478.v1.1.0.

[48] Z. Bian, Q. Li, L. Wang, C. Lu, S. Yin, and X. Li, "Relative power and coherence of EEG series are related to amnestic mild cognitive impairment in diabetes," *Frontiers Aging Neurosci.*, vol. 6, p. 11, Feb. 2014.

[49] M. Xia, J. Wang, and Y. He, "BrainNet viewer: A network visualization tool for human brain connectomics," *PLoS ONE*, vol. 8, no. 7, Jul. 2013, Art. no. e68910.