Graph-based Conditional Generative Adversarial Networks for Major Depressive Disorder Diagnosis with Synthetic functional Brain Network Generation

Ji-Hye Oh, Deok-Joong Lee, Chang-Hoon Ji, Dong-Hee Shin, Ji-Wung Han, Young-Han Son, and Tae-Eui Kam

Abstract— **Major Depressive Disorder (MDD) is a pervasive disorder affecting millions of individuals, presenting a significant global health concern. Functional connectivity (FC) derived from resting-state functional Magnetic Resonance Imaging (rs-fMRI) serves as a crucial tool in revealing functional connectivity patterns associated with MDD, playing an essential role in precise diagnosis. However, the limited data availability of FC poses challenges for robust MDD diagnosis. To tackle this, some studies have employed Deep Neural Networks (DNN) architectures to construct Generative Adversarial Networks (GAN) for synthetic FC generation, but this tends to overlook the inherent topology characteristics of FC. To overcome this challenge, we propose a novel Graph Convolutional Networks (GCN)-based Conditional GAN with Class-Aware Discriminator (GC-GAN). GC-GAN utilizes GCN in both the generator and discriminator to capture intricate FC patterns among brain regions, and the class-aware discriminator ensures the diversity and quality of the generated synthetic FC. Additionally, we introduce a topology refinement technique to enhance MDD diagnosis performance by optimizing the topology using the augmented FC dataset. Our framework was evaluated on publicly available rs-fMRI datasets, and the results demonstrate that GC-GAN outperforms existing methods. This indicates the superior potential of GCN in capturing intricate topology characteristics and generating high-fidelity synthetic FC, thus contributing to a more robust MDD diagnosis.**

Index Terms— **Conditional generative adversarial networks, Graph convolutional networks, Major depressive disorder, Resting-state functional Magnetic Resonance Imaging (rs-fMRI), Synthetic functional connectivity**

I. INTRODUCTION

J.-H. Oh, D.-J. Lee, C.-H. Ji, D.-H. Shin, J.-W. Han, Y.-H. Son, and T.-E. Kam are with the Department of Artificial Intelligence, Korea University, Seoul, Republic of Korea (e-mail: meeeo_@korea.ac.kr; deokjoong@korea.ac.kr; ckdgns0611@korea.ac.kr; dongheeshin@korea.ac.kr; daniel_han@korea.ac.kr; yhson135@korea.ac.kr; kamte@korea.ac.kr)

MAJOR Depressive Disorder (MDD) is one of the most prevalent mental disorders, characterized by its persistent and severe symptoms such as depressed mood, cognitive impairment, concentration difficulties, and anhedonia [1]. According to the World Health Organization (WHO) mental health report [2], approximately 264 million individuals worldwide suffer from MDD, resulting in significant economic burdens and raising concerns in public health. Therefore, in recognizing the economic and societal burdens caused by MDD, the field of public health emphasizes the urgency of effective treatment, highlighting the need for accurate and timely diagnostic methods [3].

Research on understanding the complex brain mechanisms of MDD is essential in establishing effective strategies for diagnosis, treatment, and prevention [4]. For the purpose, many researchers have considerably utilized resting-state functional Magnetic Resonance Imaging (rs-fMRI) over the past few decades [5]. This non-invasive neuroimaging technique enables an objective analysis of regional brain activity and alterations in networks related to MDD by closely observing subtle changes in blood oxygen level-dependent (BOLD) signals [6]. In particular, for a comprehensive understanding of the complex functional interactions between brain regions involved in MDD, it is essential to utilize functional connectivity (FC) between averaged BOLD signals derived from voxels in each pair of Regions of Interest (ROIs) [7]. The FC provides relevant information about the regional connectivity patterns of a particular brain network, as well as enables the investigation of functional connectivity patterns in the context of the whole brain [8], [9]. Through the analysis of FC, researchers can better discern distinct patterns of brain activity associated with MDD, leading to deepening the insight into MDD and playing a pivotal role in creating more effective strategies for early diagnosis and treatment [4].

With the recent advancements in deep learning (DL), there has been a significant shift towards using DL-based methods in FC analysis for MDD diagnosis [10]–[12]. These cutting-edge methodologies have yielded remarkable results in diagnosing MDD, significantly advancing our understanding of functional connectivity and its relation to MDD. Specifically, at the early stage, DNN [13] was particularly useful for extracting

patterns and features from high-dimensional FC with their ability to model complex non-linear relationships. In addition, BrainNetCNN [14] which was designed with Convolutional Neural Networks (CNN) [15] allowed it to effectively capture local and spatial dependencies in brain connectivity data.

Despite their effectiveness, these methods can make it difficult to capture complex network structures in the FC, which is typically represented as a graph. The utilization of graph theory methods [16], [17] has been traditionally predominant in FC studies due to their ability to capture and examine intricate interactions among different brain regions. It has shown notable effectiveness in understanding the complexities of MDD and facilitating an organized analysis of the FC patterns inherent to the disorder [18]. Motivated by this traditional approach, and with the recent advent of Graph Neural Networks (GNN) [19], specifically designed for graph data, many researchers are actively employing GNN, particularly, Graph Convolutional Networks (GCN) for FC studies [11], [12], [20].

One of the main issues in GCN-based FC analysis is how to define the topology, as it directly affects the GCN's ability to learn fundamental MDD-related FC patterns by guiding information processing and propagation through the network [21]. However, most GCN-based MDD diagnoses heavily rely on a small amount of FC to define the topology. For complex disorders like MDD, which have diverse symptoms and causes, defining the topology based on limited data can create difficulties in capturing the essential information necessary for a precise MDD diagnosis [22]. Additionally, using GCN methods with only limited amounts of FC presents another significant hurdle, as DL methods such as GCN inherently require significant amounts of data to ensure more accurate and generalizable performance.

To overcome the limited data availability, only a few recent studies have employed Generative Adversarial Networks (GAN) [23] for data augmentation [24]–[26]. Even though previous studies have illustrated promising results, there are two notable issues. First, they utilized DNN and BrainNetCNN as a classifier and the backbone architectures of GAN as well [24]–[26]. However, for the reasons mentioned above, we believe that employing a GCN as the backbone architecture for GAN would facilitate the FC generation with its proven capability to capture essential topological characteristics of FC for MDD diagnosis.

Second, the previous research carries the risk of mode collapse [27] since these studies employ the Auxiliary Classifier GAN (ACGAN) [28] or similar frameworks, which can lead to the issue of low intra-class diversity [29]. The low intra-class diversity implies a decrease in the diversity of data generated by the GAN within the same class [29]. Consequently, the GAN tends to repetitively generate data following the same pattern within a restricted range, leading to mode collapse, a situation where it fails to reflect the overall diversity of the training data [27]. This risk poses a significant problem, especially when using GAN to overcome data scarcity and generate diverse and novel data, as it contradicts the primary objective.

To address these issues, in this study, we propose a novel method, GCN-based Conditional GAN with Class-Aware Discriminator (GC-GAN), to improve FC generation quality for robust MDD diagnosis. First of all, GC-GAN employs GCN as the generator and discriminator of the GAN. This allows for a more accurate capture of the complex relationships and patterns between brain regions of FC. Therefore, GC-GAN can precisely reflect important information related to MDD and enable the generation of more realistic FC. Additionally, by employing a class-aware discriminator, GC-GAN addresses the problem of low intra-class diversity [29] to mitigate the potential risk of mode collapse [27]. As a result, our proposed GC-GAN is capable of generating high-fidelity synthetic FC, providing an effective countermeasure to the challenge of data scarcity.

Furthermore, we introduce a graph topology refinement technique by utilizing the augmented FC dataset, including real and synthetic FC, for constructing a robust GCN classifier. Notably, the graph topology directly influences the GCN's capacity to learn essential MDD-related FC patterns, guiding the flow of information processing and propagation throughout the network. The refined topology, constructed on a more comprehensive and diverse dataset, enables a more precise comprehension of the interconnections between brain regions and allows for a generalized diagnosis of MDD.

To sum up, the main contributions of our study include:

1) We present a novel GAN that uses a GCN-based generator and discriminator to effectively capture the complex relationships among brain regions in FC, thereby creating data that more accurately represents the vital information necessary for MDD diagnosis. To the best of our knowledge, GC-GAN is the first to generate FC based on a GCN architecture.

2) Our approach not only uses the synthetic FC in training the classifier but also finely refines the graph topology, making it more robust and specialized in MDD diagnosis. Through this topology refinement technique, we have elevated the performance of MDD diagnosis and have gained a deeper understanding of the brain regions implicated in MDD.

We assess GC-GAN on a publicly available rs-fMRI dataset for MDD diagnosis and demonstrate its superiority by conducting comparison experiments with the previous competing methods. We also perform ablation studies to analyze the effectiveness of GC-GAN on classification performance.

II. MATERIALS AND METHODS

A. Data Acquisition and Pre-processing

In our work, we used the publicly available REST-meta-MDD dataset[1] [30], the most extensive MDD rs-fMRI dataset that is provided by the Depression Imaging REsearch ConsorTium (DIRECT) [31]. Among the 25 diverse sites constituting the dataset, we utilized data specifically from Site 20, which is the largest site consisting of 249 MDD subjects and 228 Normal Control (NC) subjects from the studies approved by local Institutional Review Boards (IRB) [30], [31]. The data
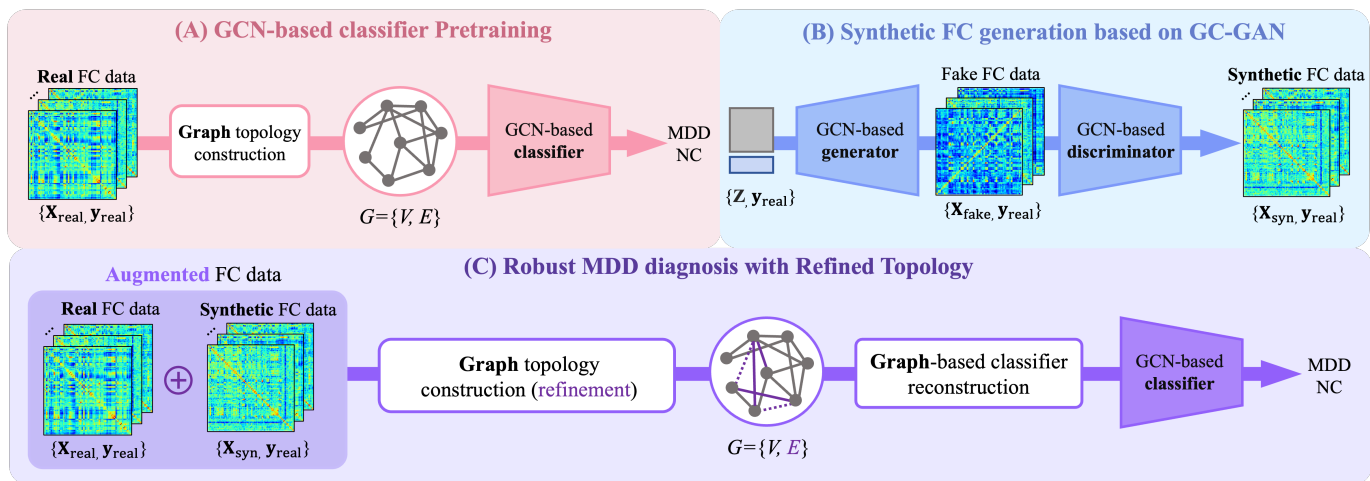
[1] https://rfmri.org/REST-meta-MDD

Fig. 1. An illustration of the proposed method, which refines the topology through the utilization of synthesized FC generated by GC-GAN for robust MDD diagnosis.

was collected using a Simens Tim Trio 3T scanner with the following parameters: TR/TE = 2,000/30 ms, flip angle = 90°, slice thickness = 3.0 mm, interslice gap = 1.0 mm, the number of slices = 32, number of time points = 242, voxel size = 3.44 × 3.44 × 4.00 mm^3, field of view (FOV) = 220 × 220 [30].

The rs-fMRI data were pre-processed using the Data Processing Assistant for Resting-State fMRI (DPARSF) [32]. Initially, the first 10 time points were discarded. The subsequent processing steps included slice-timing correction, head motion correction, band-pass filtering, and removal of confounding factors [20]. Following this, co-registration between T1-weighted images and mean functional images was performed, along with a transformation from individual native space to the Montreal Neurological Institute (MNI) template space [20]. The processed volumes are partitioned into 112 ROIs based on Harvard Oxford atlas [33]. Subsequently, the time series of rs-fMRI BOLD signals were extracted and averaged. Finally, a 112 × 112 FC matrix was constructed for each subject by calculating Pearson's correlation coefficient between ROIs with Fisher's z transformation [34].

B. Proposed Method

Fig. 1 illustrates an overview of our proposed method, which refines the graph topology through the utilization of real and synthesized FC data together for robust MDD diagnosis. As shown in Fig. 1(A), we first pre-train a GCN-based classifier with the topology defined from real FC data. We then construct GC-GAN to generate synthetic FC data as shown in Fig. 1(B). Here, the topology and GCN-based classifier in Fig. 1(A) are adopted to construct GC-GAN. Finally, as shown in Fig. 1(C), we refine the topology and reconstruct the GCN-based classifier based on the augmented FC dataset including both the real and generated synthetic FC data. The detailed overall procedure is outlined in Algorithm 1 in Supplementary Materials.

1) GCN-based Classifier Pretraining: In this work, as shown in Fig. 1(A), each real FC derived from rs-fMRI data is represented as an undirected graph $\mathbf{G} = (\mathbf{V}, \mathbf{E})$. Here, $\mathbf{V}$ is a set of R nodes corresponding to the set of ROIs, and $\mathbf{E}$ denotes the set of edges indicating connections between pairs of nodes. Notably, in our study, the term 'topology' specifically refers to these edges. Each node has its own feature vector $\mathbf{v} \in \mathbb{R}^{1 \times R}$, equivalent to the row vector of the node in the FC matrix. The topology of graph $\mathbf{G}$ is mathematically represented by an adjacency matrix $\mathbf{A} \in \mathbb{R}^{R \times R}$. To define a topology capturing distinct connection patterns between MDD and NC, we employed the minimum-Redundancy-Maximum-Relevance (mRMR) feature selection algorithm [35] on real FC data. Since the algorithm primarily assesses relative importance using mutual information, it enables us to define a topology by focusing on connections highly relevant to MDD in a data-driven manner. By retaining only approximately 10% of the connections [36]–[38], we formed a more effective sparse topology as compared to a fully connected one [39], [40]. The selected feature of FC is visualized in Supplementary Materials Fig. S1.

To extract informative features from the FC graph, we employed spectral graph convolution [19], [41], a widely utilized technique in various studies focused on brain disease diagnosis [42]. In spectral graph analysis, the graph Laplacian plays a pivotal role and is denoted as $\mathbf{L} = \mathbf{D} - \mathbf{A}$. Here, $\mathbf{D} \in \mathbb{R}^{R \times R}$ represents the diagonal degree matrix, and its elements are defined as $\mathbf{D}(i,i) = \sum_j \mathbf{A}(i,j)$. The normalized graph Laplacian is represented as $\mathbf{L} = \mathbf{I_R} - \mathbf{D}^{-\frac{1}{2}} \mathbf{A} \mathbf{D}^{-\frac{1}{2}}$, where $\mathbf{I_R}$ is the identity matrix. The Laplacian can be diagonalized by the matrix of eigenvectors of the normalized graph Laplacian $\mathbf{U}$ such that $\mathbf{L} = \mathbf{U} \boldsymbol{\Lambda} \mathbf{U}^\top$, where $\boldsymbol{\Lambda} = \text{diag}([\lambda_0, ..., \lambda_{R-1}])$ is the diagonal matrix of eigenvalues [41].

The spectral graph convolution can be defined as the multiplication of a feature $\mathbf{v} \in \mathbb{R}^{1 \times R}$ with a filter $\mathbf{g}_\theta = \text{diag}(\theta)$, parameterized by $\theta \in \mathbb{R}^{1 \times R}$:

$$\mathbf{g}_\theta * \mathbf{v} = \mathbf{U} \mathbf{g}_\theta(\boldsymbol{\Lambda}) \mathbf{U}^\top \mathbf{v}, \qquad (1)$$

where $*$ is the convolution operator on graph [37], [41].

However, this non-parametric filter is not localized in space and entails computationally expensive matrix multiplication
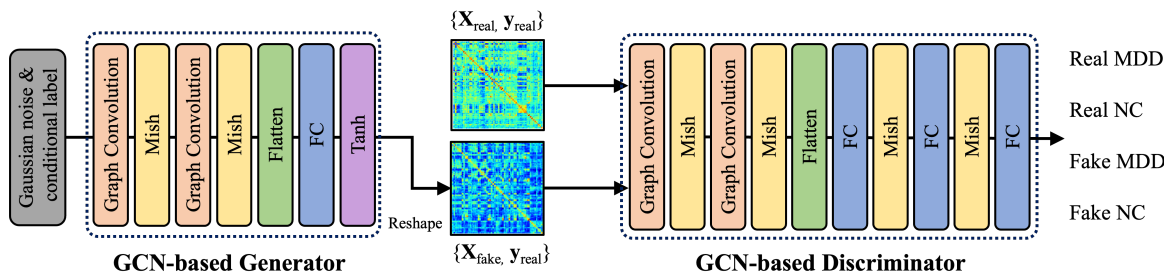
Fig. 2. An illustration of the GC-GAN architecture for synthetic FC generation.

[41]. To mitigate the computational complexity of calculating $\mathbf{U}\mathbf{g}_\theta(\boldsymbol{\Lambda})\mathbf{U}^\top$, the calculation is simplified using k-order Chebyshev polynomials to approximate the filter $\mathbf{g}_\theta(\boldsymbol{\Lambda})$, expressed as:

$$\mathbf{g}_\theta(\boldsymbol{\Lambda}) * \mathbf{v} \approx \sum_{k=0}^{K-1} \theta_k T_k(\tilde{\mathbf{L}})\mathbf{v}, \tag{2}$$

where $\tilde{\mathbf{L}}$ is the scaled graph Laplacian and θ_k are the trainable parameters [43]. Recall that the Chebyshev polynomial $T_k(\tilde{\mathbf{L}})$ of order k may be computed by the stable recurrence relation $T_k(\tilde{\mathbf{L}}) = 2\tilde{\mathbf{L}}T_{k-1}(\tilde{\mathbf{L}}) - T_{k-2}(\tilde{\mathbf{L}})$ with $T_0(\tilde{\mathbf{L}}) = 1$ and $T_1(\tilde{\mathbf{L}}) = \tilde{\mathbf{L}}$ [37], [41].

In this study, given the node features $\mathbf{H} = [\mathbf{v}_0, ..., \mathbf{v}_{R-1}]^\top$ of the FC graph, we can define the output feature of the l-th ChebConv as:

$$\mathbf{H}^{(l+1)} = \sum_{k=0}^{K-1} \theta_k^{(l)} T_k(\tilde{\mathbf{L}})\mathbf{H}^{(l)}. \tag{3}$$

This formulation enables effective feature extraction by leveraging the intrinsic properties of the graph structure [37]. The objective function of the classifier C employs the cross entropy loss [44] as follows:

$$\mathcal{L}_c = -y \cdot \log(\hat{y}) - (1 - y) \cdot \log(1 - \hat{y}), \tag{4}$$

where $\hat{y}$ denotes the predicted label of the GCN-based classifier.

We adopt a pre-training approach in which we first train a GCN-based classifier with the defined topology based on the real FC, and then adopt the GCN layer of the trained classifier as the initial feature extractor of the discriminator of GC-GAN. By using the classification capabilities of the GCN-based classifier trained on the real FC, our discriminator can efficiently discriminate between real/fake FC of each class, i.e., MDD and NC, from an early stage.

2) Synthetic FC generation based on GC-GAN: In this phase in Fig. 1(B), we construct GC-GAN to generate synthetic FC. GC-GAN is composed of a pair of GCN-based generator and discriminator, which are trained adversarially. Fig. 2 illustrates the GC-GAN architecture.

The class-aware discriminator Dis is architected in a manner that facilitates simultaneous classification of whether data is real or synthetic, and the determination of its corresponding class label. Specifically, the discriminator receives real FC or synthetic FC as input and tries to classify real FC as 'real MDD/NC' and synthetic FC as 'fake MDD/NC' (i.e. $y = $

0: real NC; $y = 1$: real MDD; $y = 2$: fake NC; $y = 3$: fake MDD). The structure of the discriminator is designed by borrowing parts of Auxiliary Discriminative Classifier GAN (ADC-GAN) [29] discriminator, which is proposed to mitigate the mode collapse [27] that may be caused by low intra-class diversity [29]. To enhance the discriminative capacity of the discriminator, we used the weights of the GCN layer from a pre-trained GCN-based classifier as the initial weights for the discriminator. The objective function of the discriminator employs cross-entropy loss [44] for both real FC and synthetic FC, thereby enabling the discriminator to acquire the ability to accurately discern between real FC and synthetic FC through the minimization of the cross-entropy loss. Mathematically the objective function of the discriminator can be expressed as follows:

$$\mathcal{L}_d = -y_{\text{real}} \cdot \log(Dis(X_{\text{real}})) - y_{\text{fake}} \cdot \log(Dis(X_{\text{fake}})), \tag{5}$$

where the y_{real} indicates the label of 'real MDD/NC', y_{fake} denotes the label of 'fake MDD/NC' and X_{fake} is the generated synthetic FC.

The generator Gen first takes (56×112) dimensional Gaussian random noise matrix Z and class label y_{real}, as input. Given that the discriminator evaluates both the class information (MDD or NC) and the authenticity of the FC (real or fake) [29], [45], the generator takes labels from the real class as input and tries to fool the discriminator by creating realistic synthetic FC, thus guiding it to categorize the synthetic FC as belonging to the corresponding input class [29], [45]. For instance, it takes the 'real MDD/NC' label as input to generate data that the discriminator identifies as 'real MDD/NC'. Through this process, the generator gains the capability to generate realistic synthetic FC while still preserving the information of the respective class. To enhance the representation of the generator, we used the weights of the GCN layer from a pre-trained GCN-based autoencoder as the initial weights for the generator. The objective function of the generator consists of cross-entropy (CE) loss [44], and mean squared error (MSE) loss [46], which allows the generator to create more realistic FC identified as real class labels by the discriminator and ensures the generated FC aligns with the scale of the real FC on an element-wise basis, respectively. The regularization parameter α of MSE loss is initialized by 1 and will gradually reduce to 0 later in the training process. Mathematically, the objective function of the generator can be

expressed as follows:

$$\mathcal{L}_g = -y_{\text{real}} \cdot \log(Dis(X_{\text{fake}})) - \alpha(X_{\text{real}} - X_{\text{fake}})^2. \quad (6)$$

3) Robust MDD diagnosis with refined topology: After training GC-GAN, we utilize the learned generator and discriminator to generate synthetic FC for data augmentation in MDD classification. Only the synthetic FC classified by the discriminator as belonging to the 'real MDD/NC' is selected and combined with real FC to create an augmented FC dataset. Subsequently, as shown in Fig. 1(C), we utilize the augmented FC dataset to reconstruct the GCN-based classifier for robust MDD diagnosis. Specifically, we apply the mRMR feature selection algorithm in the augmented FC dataset to refine the topology, a technique we refer to as 'topology refinement'. Through this process, some of the initially selected features that were deemed less important may disappear, while new features initially not selected but considered significant may emerge. In other words, by leveraging the diverse and comprehensive information obtained from the augmented FC dataset, the existing topology is finely adjusted to create a structure more suitable for MDD diagnosis. Utilizing this refined topology, we reconstruct a GCN-based classifier for robust MDD diagnosis and train the classifier using the augmented FC dataset. To objectively evaluate the robustness of our approach, the reconstructed GCN-based classifier is trained in the same experimental setting as a baseline classifier.

III. EXPERIMENTAL RESULTS

A. Experimental Setting

In this study, the discriminator and classifier were designed with a two-layer GCN and three fully connected layers, respectively. After each layer, except for the last layer, the Mish activation function [47] was implemented. Additionally, the architecture of the classifier included two batch normalization layers. Similarly, the generator was designed as a two-layer GCN architecture accompanied by a single fully connected layer. Here as well, we applied the Mish activation function after all layers, except the final layer. Furthermore, the Tanh function was used following the last layer to restrict the output range within -1 and 1.

For training, we set the discriminator's learning rate to $9 \cdot e^{-05}$ with a weight decay of $5 \cdot e^{-03}$. The generator used a learning rate of $1 \cdot e^{-04}$ with the same weight decay. We optimized each model using the Adam optimizer [48] with parameters (0.5, 0.9), an exponential scheduler with a decay factor of 0.998, 1,000 training epochs, and a batch size of 100. For the classifier, we employed a learning rate of $5 \cdot e^{-06}$ and a weight decay of $1 \cdot e^{-4}$. The Adam optimizer with parameters (0.5, 0.9), and an exponential scheduler with a decay factor of 0.998 were applied. The training process extended over 2,000 epochs with a batch size of 200.

Notably, all experiments were conducted using the same amount of synthetic FC as the real FC, i.e., 100% augmentation ratio according to the results in Table V. The source codes of the proposed method are available online[2] with the detailed

parameters for all the models presented in the experimental results.

For robustness and unbiased evaluation, we employed a 5-fold cross-validation and repeated this process five times with different random seeds and evaluated the comparison results with four metrics: classification accuracy (ACC), sensitivity (SEN), specificity (SPEC), and F1 score (F1).

B. Competing Methods

To validate the effectiveness of our proposed GC-GAN, we compared the performance of GC-GAN with the several competing methods, i.e., Semi-Supervised GAN (SSGAN) [24], Wasserstein GAN with Gradient Penalty (WGAN-GP) [25], and Auxiliary Classifier GAN (ACGAN) [26]. In SS-GAN, the discriminator classifies data into $c + 1$ classes, where the first c classes correspond to real data from c different categories, and the last $(c + 1)^{th}$ class corresponds to fake data [24]. WGAN-GP incorporates the Wasserstein loss function with a gradient penalty term and utilizes a classifier to guide the data generation process, thus allowing the model to conditionally generate data for specific classes [25]. Lastly, ACGAN enhances the discriminator with an auxiliary classifier to distinguish between real and generated samples and classify them based on additional label information [26].

Notably, to ensure a fair evaluation of data generation performance, we utilized the same GCN-based generator and discriminator for both GC-GAN and the competing methods. Furthermore, to assess the efficacy of our topology refinement (TR) technique in enhancing performance, we conducted ablation experiments comparing the results obtained with (marked as '✓' in the 'TR' column in Table I) and without (marked as '-' in the 'TR' column in Table I) the topology refinement technique.

C. Comparison Results

In the experiments, as shown in Table I, our GC-GAN achieved outstanding performance compared to the competing method regardless of whether TR was applied ('✓') or not ('-'). GC-GAN attained the highest ACC of 66.84% with TR, exhibiting improvements ranging from 1.26 to 2.82% compared to the other methods, by achieving the best SEN of 70.24%, with 1.44 to 3.70% improvements and the best SPEC of 63.14% with 0.70 to 2.30% improvements compared to the competing methods. Lastly, in terms of the F1, GC-GAN achieved the best performance of 68.72%, with improvements from 1.34% to 2.95% over the competing methods as shown in Table I. Notably, in all the proposed GC-GAN and competing methods, adopting topology refinement ('✓') resulted in performance improvements across all evaluation metrics. Especially, GC-GAN achieved the greatest performance improvements with topology refinement in terms of ACC, SEN, SPEC, and F1. The results demonstrate the superiority of GC-GAN compared to other competing methods and emphasize the positive impact of TR within our framework. In conclusion, GC-GAN demonstrated improved performance across all metrics compared to competing methods, underscoring its effectiveness in generating realistic synthetic FC data.

---

[2]https://github.com/lunyy/GC-GAN

TABLE I

MDD CLASSIFICATION RESULTS USING AUGMENTED SYNTHETIC FC FROM THE PROPOSED GC-GAN AND COMPETING GAN METHODS. THE WILCOXON SIGNED-RANK TEST WAS CONDUCTED TO VALIDATE THE STATISTICAL SIGNIFICANCE (*: $p < 0.05$, **: $p < 0.001$) COMPARED TO THE RESULTS OF THE PROPOSED GC-GAN WITH TOPOLOGY REFINEMENT (TR).

Classifier	GAN framework	TR	ACC (%)	SEN (%)	SPEC (%)	F1 (%)
GCN [19]	-	-	64.77 ± 5.70*	68.00 ± 8.39	61.27 ± 9.76	66.72 ± 5.57*
	SSGAN [24]	-	64.62 ± 4.60*	68.07 ± 8.00	60.84 ± 7.20	66.62 ± 4.98*
		✓	65.16 ± 4.41	68.15 ± 7.96	61.90 ± 6.15	66.96 ± 5.03
	WGAN-GP [25]	-	64.02 ± 4.50**	66.54 ± 7.12**	61.28 ± 6.61	65.77 ± 4.77**
		✓	65.12 ± 3.91	67.58 ± 6.82	62.44 ± 5.51	66.79 ± 4.56
	ACGAN [26]	-	64.91 ± 4.42*	67.68 ± 7.56*	61.90 ± 8.68	66.71 ± 4.36*
		✓	65.58 ± 3.37	68.48 ± 7.29	62.42 ± 7.48	67.38 ± 3.66*
	GC-GAN (Ours)	-	65.18 ± 5.08*	68.80 ± 8.09	61.28 ± 7.53	67.23 ± 5.27*
		✓	**66.84 ± 4.25**	**70.24 ± 7.89**	**63.14 ± 8.35**	**68.72 ± 4.57**

TABLE II

MDD CLASSIFICATION RESULTS WERE DERIVED FROM AN ABLATION STUDY WHERE DIFFERENT GAN BACKBONE ARCHITECTURES (GENERATOR (G) — DISCRIMINATOR (D)) WERE EMPLOYED WITHIN OUR PROPOSED GC-GAN FRAMEWORK. THE WILCOXON SIGNED-RANK TEST WAS CONDUCTED TO VALIDATE THE STATISTICAL SIGNIFICANCE (*: $p < 0.05$, **: $p < 0.001$) COMPARED TO THE RESULTS OF THE PROPOSED GCN-GCN WITH TOPOLOGY REFINEMENT (TR).

Classifier	GC-GAN backbone(G−D)	TR	ACC (%)	SEN (%)	SPEC (%)	F1 (%)
GCN [19]	-	-	64.77 ± 5.70*	68.00 ± 8.39	61.27 ± 9.76	66.72 ± 5.57*
	DNN−DNN	-	65.24 ± 4.60*	69.60 ± 7.38	60.48 ± 8.69*	67.54 ± 4.50
		✓	65.88 ± 3.75	69.44 ± 7.45	61.98 ± 6.55	67.85 ± 4.36
	DNN−BrainNetCNN	-	63.05 ± 3.97**	67.27 ± 7.82**	58.48 ± 8.44**	65.39 ± 4.3**
		✓	63.81 ± 3.50**	68.23 ± 8.11	59.02 ± 7.82**	66.14 ± 4.2**
	DNN−GCN	-	65.11 ± 5.40	69.36 ± 7.61	60.47 ± 7.77*	67.40 ± 5.29
		✓	66.16 ± 4.63	69.76 ± 8.02	62.24 ± 6.30	68.12 ± 5.05
	GCN−GCN (Ours)	-	65.18 ± 5.08*	68.80 ± 8.09	61.28 ± 7.53	67.23 ± 5.27*
		✓	**66.84 ± 4.25**	**70.24 ± 7.89**	**63.14 ± 8.35**	**68.72 ± 4.57**

D. Influence of backbone architecture in GC-GAN

To demonstrate the superiority of our proposed GC-GAN framework, which is based on a GCN generator and discriminator, we conducted experiments applying various backbone architectures, which have been utilized in prior studies [24]–[26]. Additionally, same as Table I, to validate the efficacy of our topology refinement (TR) technique in enhancing performance, we also conducted ablation experiments comparing the results obtained with and without the topology refinement technique on the various backbone architectures.

In Table II, utilizing GCN as the backbone architecture of GC-GAN demonstrates superior performance compared to other backbone architectures. This result implies that GCN possesses the capability to capture essential topological characteristics of functional connectivity for MDD diagnosis. Furthermore, same as the results of Table I, the application of topology refinement ('✓') resulted in performance improvements across all metrics for all GAN models. This validates that our topology refinement technique operates effectively in a general context, not limited to a specific GAN model.

E. Application of GC-GAN to other GNN-based models

To demonstrate the reliability of our work, we have conducted an additional experiment to investigate the following two points. First, we sought to investigate whether the classification results from our model achieved meaningful improvement compared to those from other classification models by implementing widely recognized GNN-based models, i.e., Graph Attention Networks (GAT) [49], GraphSAGE [50], and Ensemble [12]. Specifically, GAT applies an attention mechanism to assign varying levels of importance to neighboring nodes for each node in the graph [49]. GraphSAGE employs a learnable function to aggregate information from sampled neighbors, enabling it to capture intricate relationships and patterns within the graph [50]. The Ensemble, recently proposed in the field of MDD diagnosis, combines various judgment techniques such as GCN, GAT, and GraphSAGE to enhance classification performance [12].

Next, we demonstrated that our GC-GAN-based data augmentation and topology refinement techniques can be generally applicable for GNN-based models including the baseline GCN classifier as well as the existing models, i.e., GAT, GraphSAGE, and Ensemble, resulting in further performance improvement with our topology refinement approach. For the purpose, we have further compared their results depending on whether the proposed method was applied ('✓' in the GC-GAN column) or not ('-' in the GC-GAN column) to each classification model as shown in Table III. Notably, for all results with our GC-GAN ('✓') in Table III, we generated augmented FC data based on our GC-GAN without retraining to optimize for each classifier, and then refined the topology by using the augmented FC dataset.

In the comparison results, we first want to highlight that the proposed method with the baseline GCN classifier achieves the best performance in ACC, SEN, and F1, and the second-best performance in SPEC as shown in Table III. Moreover, the results imply that the proposed GC-GAN-based method can be applicable not only to the GCN baseline classifier but also

TABLE III

MDD CLASSIFICATION RESULTS USING VARIOUS CLASSIFIERS WITH ('✓') AND WITHOUT ('-') SYNTHETIC FC GENERATED FROM THE PROPOSED GC-GAN BASED ON THE HARVARD OXFORD ATLAS. THE WILCOXON SIGNED-RANK TEST WAS CONDUCTED TO VALIDATE THE STATISTICAL SIGNIFICANCE (*: $p < 0.05$, **: $p < 0.001$) COMPARED TO THE RESULTS OF THE PROPOSED GC-GAN WITH TOPOLOGY REFINEMENT (TR).

Classifier	GC-GAN	ACC (%)	SEN (%)	SPEC (%)	F1 (%)
GAT [49]	-	63.14 ± 3.67**	65.56 ± 6.66**	60.48 ± 4.08*	64.89 ± 4.07**
	✓	$\mathbf{64.39 \pm 3.93}$*	$\mathbf{67.34 \pm 6.31}$*	$\mathbf{61.22 \pm 6.91}$	$\mathbf{66.32 \pm 4.00}$*
GraphSAGE [50]	-	65.21 ± 3.21*	67.49 ± 6.55*	62.69 ± 6.25	66.82 ± 3.76*
	✓	$\mathbf{66.04 \pm 4.03}$	$\mathbf{69.02 \pm 6.86}$	$\mathbf{62.77 \pm 5.95}$	$\mathbf{67.85 \pm 4.47}$
Ensemble [12]	-	65.17 ± 3.84	$\mathbf{68.79 \pm 7.14}$	61.17 ± 8.43	67.22 ± 3.98
	✓	$\mathbf{66.16 \pm 3.84}$	68.79 ± 9.46	$\mathbf{63.30 \pm 9.80}$	$\mathbf{67.95 \pm 4.04}$
GCN [19]	-	64.77 ± 5.70*	68.00 ± 8.39	61.27 ± 9.76	66.72 ± 5.57*
	✓(ours)	$\mathbf{66.84 \pm 4.25}$	$\mathbf{70.24 \pm 7.89}$	$\mathbf{63.14 \pm 8.35}$	$\mathbf{68.72 \pm 4.57}$

TABLE IV

MDD CLASSIFICATION RESULTS ON DIFFERENT SITE DATASETS WITH ('✓') AND WITHOUT ('-') SYNTHETIC FC GENERATED FROM THE PROPOSED GC-GAN BASED ON THE HARVARD OXFORD ATLAS.

Train	Test	GC-GAN	ACC (%)	SEN (%)	SPEC (%)	F1 (%)
S20	S1	-	57.52 ± 1.74	$\mathbf{52.88 \pm 5.94}$	62.20 ± 5.54	55.32 ± 3.32
		✓	$\mathbf{59.04 \pm 1.93}$	52.32 ± 6.37	$\mathbf{65.74 \pm 3.75}$	$\mathbf{55.90 \pm 4.30}$
	S21	-	56.52 ± 2.79	54.70 ± 5.82	$\mathbf{59.10 \pm 7.42}$	57.76 ± 3.45
		✓	$\mathbf{57.88 \pm 2.13}$	$\mathbf{57.98 \pm 3.83}$	57.82 ± 5.17	$\mathbf{60.16 \pm 2.35}$

TABLE V

MDD CLASSIFICATION RESULTS WITH VARYING SYNTHETIC FC AUGMENTATION RATIOS USING THE PROPOSED GC-GAN ON HARVARD OXFORD ATLAS.

Augmentation ratio	ACC (%)	SEN (%)	SPEC (%)	F1 (%)
0%	64.77 ± 5.70	68.00 ± 8.39	61.27 ± 9.76	66.72 ± 5.57
50%	66.17 ± 5.18	69.28 ± 8.86	62.77 ± 8.53	67.95 ± 5.60
100%	$\mathbf{66.84 \pm 4.25}$	$\mathbf{70.24 \pm 7.89}$	$\mathbf{63.14 \pm 8.35}$	$\mathbf{68.72 \pm 4.57}$
150%	66.38 ± 4.44	70.00 ± 7.88	62.39 ± 7.81	68.37 ± 4.88
200%	65.24 ± 3.95	69.02 ± 11.88	61.05 ± 12.64	66.99 ± 6.12

to other GNN-based models. Specifically, it helps to improve the performance of the other GNN-based baseline models in all the metrics. Specifically, by applying our GC-GAN to the other GNN-based models, ACC improves by 0.83 to 2.07%. Moreover, the highest improvement in SEN is 2.24%, and SPEC shows an increase of up to 2.13%. Finally, with our GC-GAN, F1 exhibits an improvement ranging from 0.73% to 2.00% compared to other GNN-based baseline models. These results highlight the potential of our GC-GAN to enhance the performance of various classifiers.

F. Application of GC-GAN on other MDD sites

To verify the effectiveness of our proposed GC-GAN method, we conduct additional experiments with various sites. Here, we conducted an additional experiment using multiple datasets from the three largest sites, i.e., Site 20, Site 1, and Site 21 within the REST-meta-MDD dataset. Specifically, Site 20 has a total of 477 samples, Site 1 has a total of 148 samples, Site 21 has a total of 145 samples. We have utilized the data from Site 20 as training dataset and the data from Site 1 and Site 21 as test dataset , respectively, to evaluate the GCN-based classification model trained on Site 20 with ('✓' in the GC-GAN column) and without ('-' in the GC-GAN column) our GC-GAN. In this experiment, although the domain shift problem [20] still remains in this experiment, it could be possible to indirectly analyze the efficiency of our GC-GAN on different datasets. Table IV illustrates the performance comparison between the baseline GCN classifier without our GC-GAN ('-') and the GCN classifier enhanced with our GC-GAN ('✓'). From this comparison, the proposed method achieved higher ACC and F1 compared to the baseline.

G. Analysis of augmentation ratio

In our study, we augmented synthetic FC data by the same amount as real FC data (i.e., 100% augmentation ratio) according to the experimental findings from the recent study [51], showing that using synthetic data more than real data can potentially reduce classification performance. In addition, the competing studies [24], [25] also augmented the synthetic data by the same amount as real FC data as well. Thus, based on the previous research, all experiments were conducted with the same amount of synthetic FC data as real FC data to ensure robust results and maintain consistency in comparison. Notably, to address the class imbalance issue with 100% synthetic data, we applied inverted sample proportions. For instance, in the REST-meta MDD dataset on Site 20, with 249/228 real FC data for MDD/NC, we generated 228/249 synthetic FCs for MDD/NC, respectively.

Table V presents the MDD classification performance employing the proposed GC-GAN with varying ratios of synthetic FC augmentation. In the 'Augmentation ratio' column, '0%' indicates experiments conducted without any data augmentation, thus reflecting the performance only based on the real FC data. The augmented results across all augmentation ratios outperform the results obtained using only real data in all the cases except SPEC on 200% augmentation ratio highlighting the robust data generation capability of our method. Moreover, it is worth noting that our method achieved the most effective performance enhancement with 100% augmentation ratio in all the metrics as shown in Table I.

H. Application of GC-GAN on AAL atlas

To verify the effectiveness of our proposed GC-GAN method, we conducted additional experiments using the Automatic Anatomical Labeling (AAL) atlas [52] which is widely

TABLE VI

MDD CLASSIFICATION RESULTS USING AUGMENTED SYNTHETIC FC FROM THE PROPOSED GC-GAN AND COMPETING GAN METHODS BASED ON THE AAL ATLAS. THE WILCOXON SIGNED-RANK TEST WAS CONDUCTED TO VALIDATE THE STATISTICAL SIGNIFICANCE (*: $p < 0.05$, **: $p < 0.01$) COMPARED TO THE RESULTS OF THE PROPOSED GC-GAN WITH TOPOLOGY REFINEMENT (TR).

| Classifier	GAN framework	TR	ACC (%)	SEN (%)	SPEC (%)	F1 (%)
GCN [19]	-	-	63.23 ± 4.80**	65.21 ± 4.62*	61.07 ± 7.00	64.92 ± 4.35*
	SSGAN [24]	-	63.84 ± 5.75	67.02 ± 5.74	60.35 ± 7.38	65.90 ± 5.32
		✓	64.64 ± 4.94	68.72 ± 5.70	60.18 ± 7.02	66.95 ± 4.62
	WGAN-GP [25]	-	62.73 ± 4.87*	64.97 ± 6.23	60.28 ± 6.80	64.44 ± 4.99*
		✓	63.44 ± 3.24	67.51 ± 4.43	59.02 ± 5.30	65.78 ± 3.18
	ACGAN [26]	-	63.49 ± 4.30*	66.13 ± 5.00	60.60 ± 7.42	65.36 ± 3.91*
		✓	64.64 ± 4.66	67.34 ± 4.45	61.69 ± 7.15	66.52 ± 4.13
	GC-GAN (Ours)	-	64.05 ± 5.00	66.36 ± 5.17	61.51 ± 7.20	65.79 ± 4.60
		✓	**65.40 ± 5.61**	**67.92 ± 6.28**	**62.65 ± 6.21**	**67.14 ± 5.53**
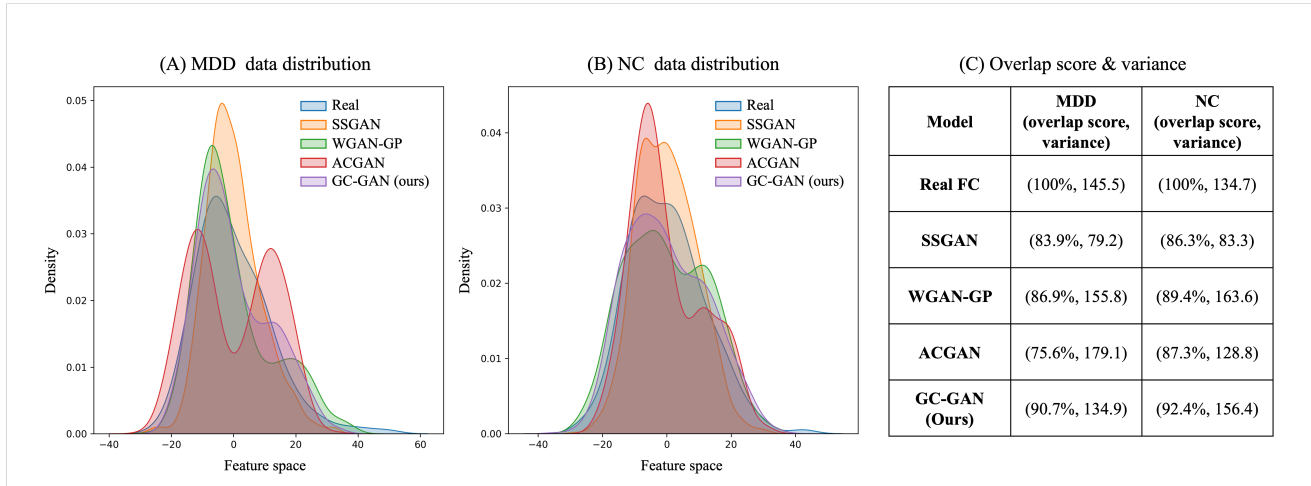
Fig. 3. The class-wise (MDD/NC) data distributions of the real FC (blue) and synthetic FC generated by the proposed (purple) and competing methods according to various GAN frameworks, i.e., SSGAN (orange), WGAN-GP (green), ACGAN (red), with the variance of each data distribution and the overlap score (%) between the distributions of synthetic and real FC, respectively.

used not only in MDD research but also in brain diseases such as Alzheimer's disease [53] and autism spectrum disorder [54]. In the experiment, we employed the same settings as our initial experiment conducted on the Harvard Oxford atlas in Table I.

Table VI shows the comparison results with the proposed and competing methods based on the AAL atlas. In the experiments, the proposed method achieved outstanding performance compared to the competing method regardless of whether TR was applied ('✓') or not ('-'). Our GC-GAN attained the highest ACC of 65.40% with TR, exhibiting improvements ranging from 0.76 to 2.67% compared to the other methods, by achieving the best SPEC of 62.65%, with 0.96 to 3.63% improvements and the second best SEN of 67.92% with -0.8 to 2.95% improvements compared to the competing methods. Lastly, in terms of the F1, our method achieved the best performance of 67.14%, with improvements from 0.19% to 2.7% over the competing methods. Notably, in all the proposed and competing methods, adopting topology refinement ('✓') showed performance improvements across all evaluation metrics, except SPEC for SSGAN and WGAN-GP. The results demonstrate the superiority of our GC-GAN compared to other competing methods and emphasize the positive impact of TR within our framework. Furthermore, experiments conducted with the additional atlas indicate the promising potential of GC-GAN, not limited to a specific atlas.

IV. DISCUSSION

A. Comparison of Synthetic FC Distribution

Fig. 3(A) and (B) show the class-wise (MDD/NC) data distributions of the real FC (blue) and synthetic FC generated by the proposed (purple) and competing methods according to various GAN framework, i.e., SSGAN (orange), WGAN-GP (green), ACGAN (red), with the Kernel Density Estimation (KDE) measure [55].

KDE is a non-parametric technique for estimating density using a kernel function and has been widely used in GAN studies to evaluate model distributions [56], [57]. In addition, as shown in Fig. 3(C), we measured the overlap score (%) between the distributions of synthetic and real FC and the variance of each data distribution. The measurements indicate how well each model produced realistic synthetic FC with the amount of mode collapse, i.e., the overlap score represents how similar the generated data is to real data while the variance shows the diversity of the generated data.

According to the measurements, we can first check that the proposed method has the highest overlap score of 90.7% for MDD data and 92.4% for NC data, respectively. The results indicate that the proposed method generated a more realistic
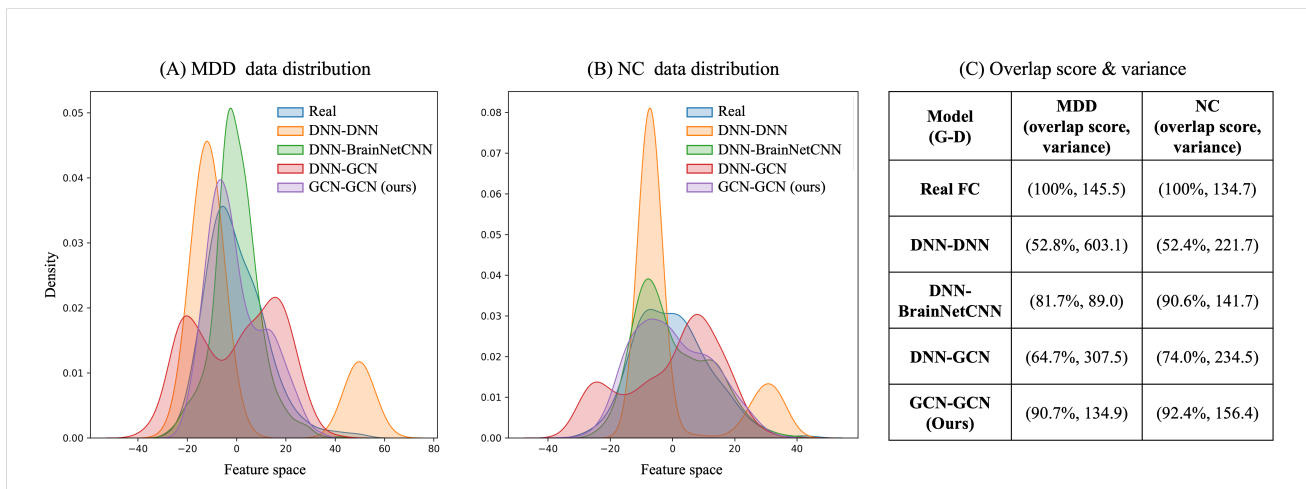
Fig. 4. The class-wise (MDD/NC) data distributions of the real FC (blue) and synthetic FC generated by the proposed (purple) and ablation models according to various GAN architectures, i.e., DNN-DNN (orange), DNN-BrainNetCNN (green), and DNN-GCN (red), with the variance of each data distribution and the overlap score (%) between the distributions of synthetic and real FC, respectively.

synthetic FC compared to the competing methods. In the case of the variance, it can be seen that our method does not have the largest variance but has the values that are closest and second closest to the variance of the actual data distribution for MDD and NC data, respectively. Considering the results on both measurements, the results mean that the synthetic FC generated by the proposed method has a distribution most similar to real FC while showing high variability. In Fig. (A), and (B), we can visually check the similarity between their distributions, i.e., the distribution from synthetic FC derived from the proposed method (purple) has the most overlap to that of the real FC (blue) for each class.

As shown in Fig.4, we also compared the class-wise data distributions of the real FC (blue) and synthetic FC generated by the proposed (GCN-GCN, purple) and ablation models according to various GAN architectures, i.e., DNN-DNN (orange), DNN-BrainNetCNN (green), and DNN-GCN (red) with the overlap score and their variability as well. Here, we can also see that the proposed method generated the most realistic synthetic FC with the highest overlap score for both classes as well as the closest and second closest variance to that of real data distribution for each class, respectively, compared to the others.

B. Qualitative Comparison of Real and Generated FCs

To validate the quality of the synthetic FC, in Fig. 5, we showed the results of the averaged synthetic FC generated by each of the proposed (GC-GAN) and the competing methods, i.e., SSGAN, WGAN-GP, and ACGAN, as well as the averaged real FC for each class (i.e., MDD and NC), respectively. We also present the differences (with their normalized averaged values, NAvg. Diff) between the averaged real and synthetic FC generated by each method, for each class (i.e., MDD-Diff and NC-Diff), respectively. In the results, our method showed the lowest MDD-Diff and NC-Dff (refer to their NAvg. Diff). According to the results, it can be confirmed that the proposed method generated the most realistic synthetic FC compared to the others. Fig. 6 also
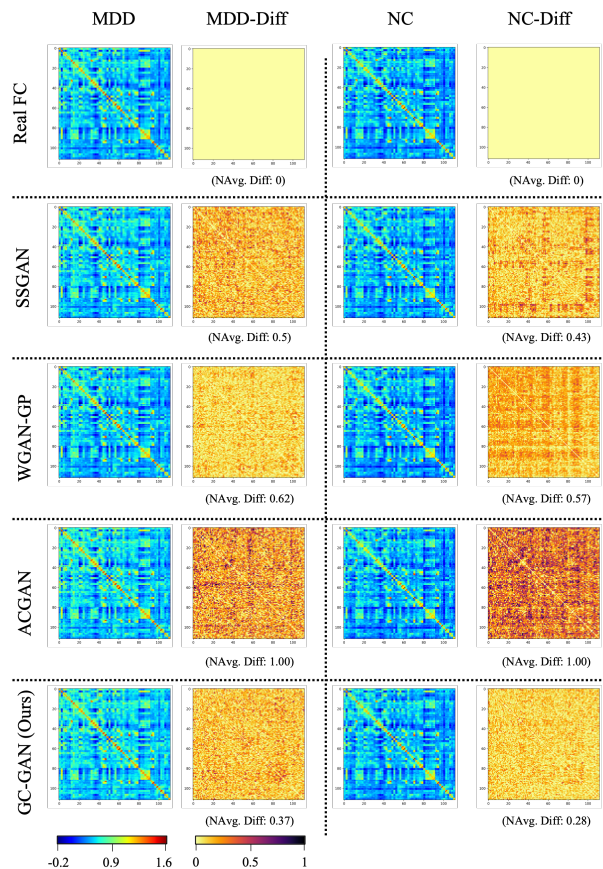

Fig. 5. The averaged synthetic FC generated by each proposed (GC-GAN) and competing methods (i.e., SSGAN, WGAN-GP, and ACGAN), for each class (i.e., MDD and NC), respectively. MDD-Diff and NC-Diff present the difference for each class between the averaged real and synthetic FC generated by each method, respectively, with their normalized averaged values (NAvg. Diff).

showed the averaged results for the proposed (GCN-GCN) and the ablation models, i.e., DNN-DNN, DNN-BrainNetCNN, and DNN-GCN, and their differences (MDD-Diff and NC-
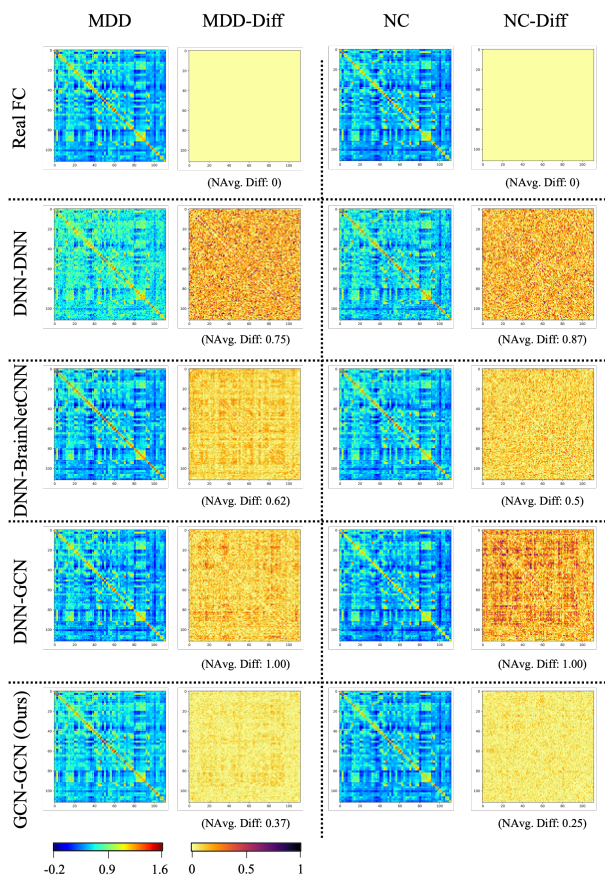
Fig. 6. The averaged synthetic FC generated by each proposed (GCN-GCN) and ablation methods (i.e., i.e., DNN-DNN, DNN-BrainNetCNN, and DNN-GCN), for each class (i.e., MDD and NC), respectively. MDD-Diff and NC-Diff present the difference for each class between the averaged real and synthetic FC generated by each method, respectively, with their normalized averaged values (NAvg. Diff).

Diff, respectively) compared to the real FC (Real). In this comparison, the proposed method also generated the most realistic synthetic, achieving the best similarity, i.e., the lowest NAvg. Diff, to the real ones compared to the others.

C. Topology-based analysis of significant connections

To show the significance of our framework from a neurological perspective, we have conducted a detailed analysis comparing the initial topology, which is defined based on the real FC data, with the refined topology, which is defined based on the augmented FC data (i.e., the combination of real and synthetic FC data) through GC-GAN. As shown in figure 7, for better visualization, we categorized FCs into three groups: (1) common FCs, which have connections in both the initial and refined topologies, (2) newly formed FCs, which have connections in the refined topology but not in the initial topology, (3) disappeared FCs, which have connections in the initial topology but not in the refined topology. To ensure a more focused analysis, we concentrated on the top 10% FCs with the highest scores from the mRMR feature selection algorithm [35] among all FCs.

First, in the common FCs, it is observed that the frontal lobe, which is widely known to be closely associated with

MDD [58], exhibits the highest number of connections. These connections are predominantly linked to the temporal and parietal lobes. Upon examining the connections between the frontal and the parietal lobe, it is observed that left subcallosal cortex (SC.L) exhibits the highest connectivity within the frontal lobe. Previous studies have reported a significant association between left subcallosal cortex and the treatment of MDD [59]. Furthermore, left supramarginal gyrus (SGa.L) and left postcentral gyrus (POG.L) both exhibiting the highest connectivity within the parietal lobe and linked to left subcallosal cortex, have been identified as pivotal regions for discriminating between MDD and NC in previous research [60]. Examining the connections between the frontal and the temporal lobe, it is observed that right frontal orbital cortex (FOC.R) has the highest number of connections with the temporal lobe. This region has previously been reported to exhibit significant volumetric differences between NC and those with MDD [61].

Second, a notable feature of the newly formed FCs is the significant increase in connections within the subcortical gray matter (SCGM), which previously had no connections in common FCs. Specifically, there is a substantial increase in connections between the frontal lobe and the SCGM, as well as between the parietal lobe and the SCGM. In particular, the thalamus (Thal) which showed a significant increase in connections in all lobes, is a region previously reported as the most causal hub for MDD [62]. Furthermore, connections have also emerged from various regions such as left putamen (Put.L) [63], right bed nucleus of the stria terminalis (Bst.R) [64], and amygdala [65] all of which are associated with MDD. Among them, the connections between amygdala and frontal orbital cortex have been reported to be crucial in determining MDD in previous study [65].

Lastly, the most noticeable feature of the disappeared FCs is that there has been a significant reduction in intra-lobe connections. Connections within the frontal lobe, occipital lobe, SCGM have all decreased. Considering the mRMR algorithm that we use to define the topology, we can infer that this outcome is because connections within the same lobe provide less new information compared to connections between different lobes. Although it requires further elaborate verification, the results might lead us to focus more on inter-lobe connections for better MDD diagnosis [66]. Within the brain's interregional connections, SC, which previously exhibited many connections, has experienced a significant decrease in connections. Specifically, in the case of anterior right temporal fusiform (TFa.R) and lateral occipital (OLi), all connections on both sides have disappeared. Both areas are associated with visual processing. Considering that the regions connected to SC were clearly related to the diagnosis and treatment of MDD, it is plausible to consider that the areas of temporal fusiform and lateral occipital may contain relatively less crucial information.

V. CONCLUSION

In this study, we introduced a novel GCN-based Conditional GAN model that generates representative and realistic

This article has been accepted for publication in IEEE Journal of Biomedical and Health Informatics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2023.3340325

OH *et al.*: GRAPH-BASED CONDITIONAL GAN FOR MDD DIAGNOSIS WITH SYNTHETIC FUNCTIONAL BRAIN NETWORK GENERATION 11
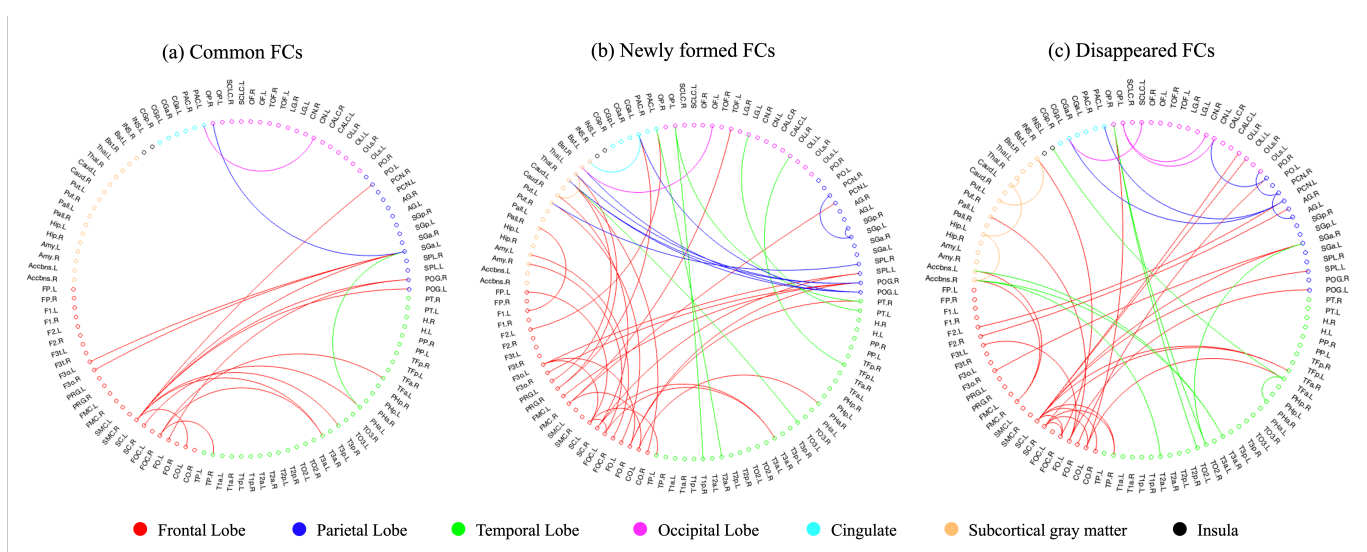

Fig. 7. The connectome ring of the 112 brain regions of Harvard Oxford atlas. The connections in both the initial and refined topologies (a) Common FCs, in the refined topology but not in the initial topology (b) Newly formed FCs, and the connections in the initial topology but not in the refined topology (c) Disappeared FCs.

synthetic FC, aiming to provide a more precise and robust MDD diagnosis. By integrating GCN architecture into our GAN model, our framework demonstrated the capability to accurately capture the complex topology characteristics and intricate patterns present in FC. Additionally, the class-aware discriminator further enhanced the quality and effectiveness of synthetic FC generation in GC-GAN framework. Furthermore, we introduced a topology refinement technique that optimizes the graph structure using the augmented synthetic FC, thereby improving the performance of MDD diagnosis.

REFERENCES

[1] Z. Li, M. Ruan, J. Chen, and Y. Fang, "Major depressive disorder: advances in neuroscience research and translational applications," *Neuroscience bulletin*, vol. 37, pp. 863–880, 2021.

[2] W. H. Organization, "Depression and other common mental disorders: global health estimates," World Health Organization, Tech. Rep., 2017.

[3] K. A. McLaughlin, "The public health impact of major depression: a call for interdisciplinary prevention efforts," *Prevention Science*, vol. 12, pp. 361–371, 2011.

[4] M. Demirtaş, C. Tornador, C. Falcón, M. López-Solà, R. Hernández-Ribas, J. Pujol, and Menchón, "Dynamic functional connectivity reveals altered variability in functional connectivity among patients with major depressive disorder," *Human brain mapping*, vol. 37, no. 8, pp. 2918–2930, 2016.

[5] C. Zhuo, G. Li, X. Lin, D. Jiang, Y. Xu, H. Tian, W. Wang, and X. Song, "The rise and fall of MRI studies in major depressive disorder," *Translational psychiatry*, vol. 9, no. 1, p. 335, 2019.

[6] S. M. Smith, D. Vidaurre, C. F. Beckmann, M. F. Glasser, M. Jenkinson, K. L. Miller, T. E. Nichols, E. C. Robinson, and Salimi-Khorshidi, "Functional connectomics from resting-state fMRI," *Trends in Cognitive Sciences*, vol. 17, no. 12, pp. 666–682, 2013.

[7] K. Helm, K. Viol, T. M. Weiger, P. A. Tass, C. Grefkes, D. Del Monte, and G. Schiepek, "Neuronal connectivity in major depressive disorder: a systematic review," *Neuropsychiatric disease and treatment*, pp. 2715–2737, 2018.

[8] W. Li, L. Zhang, L. Qiao, and D. Shen, "Toward a better estimation of functional brain network for mild cognitive impairment identification: a transfer learning view," *IEEE journal of biomedical and health informatics*, vol. 24, no. 4, pp. 1160–1168, 2019.

[9] J. Ji, Z. Chen, and C. Yang, "Convolutional neural network with sparse strategies to classify dynamic functional connectivity," *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 3, pp. 1219–1228, 2021.

[10] J. Z. Ji and Y. Yao, "A novel CNN framework to extract multi-level modular features for the classification of brain networks," *Applied Intelligence*, vol. 52, no. 6, pp. 6835–6852, 2022.

[11] M. Zhu, Y. Quan, and X. He, "The classification of brain network for major depressive disorder patients based on deep graph convolutional neural network," *Frontiers in Human Neuroscience*, vol. 17, 2023.

[12] S. Venkatapathy, M. Votinov, L. Wagels, S. Kim, M. Lee, U. Habel, I.-H. Ra, and H.-G. Jo, "Ensemble graph neural network model for classification of major depressive disorder using whole-brain functional connectivity," *Frontiers in Psychiatry*, vol. 14, 2023.

[13] G. E. Hinton, S. Osindero, and Y.-W. Teh, "A fast learning algorithm for deep belief nets," *Neural computation*, vol. 18, no. 7, pp. 1527–1554, 2006.

[14] J. Kawahara, C. J. Brown, S. P. Miller, B. G. Booth, V. Chau, R. E. Grunau, J. G. Zwicker, and G. Hamarneh, "Brainnetcnn: Convolutional neural networks for brain networks; towards predicting neurodevelopment," *NeuroImage*, vol. 146, pp. 1038–1049, 2017.

[15] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017.

[16] J. Wang, X. Zuo, and Y. He, "Graph-based network analysis of resting-state functional mri," *Frontiers in systems neuroscience*, p. 16, 2010.

[17] O. Sporns, "Graph theory methods: applications in brain networks," *Dialogues in clinical neuroscience*, 2022.

[18] M. Ye, T. Yang, P. Qing, X. Lei, J. Qiu, and G. Liu, "Changes of functional brain networks in major depressive disorder: a graph theoretical analysis of resting-state fMRI," *PloS one*, vol. 10, no. 9, p. e0133775, 2015.

[19] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," *arXiv preprint arXiv:1609.02907*, 2016.

[20] Y. Fang, M. Wang, G. G. Potter, and M. Liu, "Unsupervised cross-domain functional mri adaptation for automated major depressive disorder identification," *Medical Image Analysis*, vol. 84, p. 102707, 2023.

[21] S.-Y. Louis, A. Nasiri, F. J. Rolland, C. Mitro, and J. Hu, "Node-select: a graph neural network based on a selective propagation technique," *Neurocomputing*, vol. 494, pp. 396–408, 2022.

[22] H. Lu and S. Uddin, "A weighted patient network-based framework for predicting chronic diseases using graph neural networks," *Scientific reports*, vol. 11, no. 1, p. 22607, 2021.

[23] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks," *Communications of the ACM*, vol. 63, no. 11, pp. 139–144, 2020.

[24] J. L. Zhao, J. J. Huang, D. M. Zhi, W. Z. Yan, X. H. Ma, X. Yang, X. B. Li, Q. Ke, T. Z. Jiang, V. D. Calhoun, and J. Sui, "Functional network connectivity (FNC)-based generative adversarial network (GAN)

and its applications in classification of mental disorders," *Journal of Neuroscience Methods*, vol. 341, 2020.

[25] C. Li, Y. Wei, X. Chen, and C.-B. Schönlieb, "Brainnetgan: Data augmentation of brain connectivity using generative adversarial network for dementia classification," in *Deep Generative Models, and Data Augmentation, Labelling, and Imperfections: First Workshop, DGM4MICCAI 2021, and First Workshop, DALI 2021, Held in Conjunction with MICCAI 2021, Strasbourg, France, October 1, 2021, Proceedings 1*. Springer, 2021, pp. 103–111.

[26] D. Yan, S. Wu, M. T. Sami, A. Almudaifer, Z. Jiang, H. Chen, D. Rangaprakash, G. Deshpande, and Y. Ma, "Improving brain dysfunction prediction by gan: A functional-connectivity generator approach," in *2021 IEEE International Conference on Big Data (Big Data)*. IEEE, 2021, pp. 1514–1522.

[27] D. Bau, J.-Y. Zhu, J. Wulff, W. Peebles, H. Strobelt, B. Zhou, and A. Torralba, "Seeing what a gan cannot generate," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2019, pp. 4502–4511.

[28] A. Odena, C. Olah, and J. Shlens, "Conditional image synthesis with auxiliary classifier gans," in *International conference on machine learning*. PMLR, 2017, pp. 2642–2651.

[29] L. Hou, Q. Cao, H. Shen, S. Pan, X. Li, and X. Cheng, "Conditional gans with auxiliary discriminative classifier," in *International Conference on Machine Learning*. PMLR, 2022, pp. 8888–8902.

[30] C. G. Yan, X. Chen, L. Li, F. X. Castellanos, T. J. Bai, Q. J. Bo, J. Cao, G. M. Chen, and Chen, "Reduced default mode network functional connectivity in patients with recurrent major depressive disorder," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 116, no. 18, pp. 9078–9083, 2019.

[31] X. Chen, B. Lu, H.-X. Li, X.-Y. Li, and Wang, "The DIRECT consortium and the REST-meta-MDD project: towards neuroimaging biomarkers of major depressive disorder," *Psychoradiology*, vol. 2, no. 1, pp. 32–42, 2022.

[32] Y. Chao-Gan and Z. Yu-Feng, "DPARSF: A MATLAB toolbox for "pipeline" data analysis of resting-state fMRI," *Front Syst Neurosci*, vol. 4, p. 13, 2010.

[33] D. N. Kennedy, N. Lange, N. Makris, J. Bates, J. Meyer, and J. Caviness, V. S., "Gyri of the human neocortex: an mri-based analysis of volume and variance," *Cereb Cortex*, vol. 8, no. 4, pp. 372–84, 1998.

[34] B. Bhuvaneshwari and A. Kavitha, "Assessment of brain connectivity patterns in progression of alzheimer's disease," *Age (years)*, vol. 72, pp. 1–2, 2016.

[35] H. Peng, F. Long, and C. Ding, "Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy," *IEEE Trans Pattern Anal Mach Intell*, vol. 27, no. 8, pp. 1226–38, 2005.

[36] S. I. Ktena, S. Parisot, E. Ferrante, M. Rajchl, M. Lee, B. Glocker, and D. Rueckert, "Metric learning with spectral graph convolutions on brain connectivity networks," *Neuroimage*, vol. 169, pp. 431–442, 2018.

[37] H. Zhang, R. Song, L. Wang, L. Zhang, D. Wang, C. Wang, and W. Zhang, "Classification of brain disorders in rs-fMRI via local-to-global graph neural networks," *IEEE Transactions on Medical Imaging*, 2022.

[38] T. Yang, M. A. Al-Duailij, S. Bozdag, and F. Saeed, "Classification of autism spectrum disorder using rs-fMRI data and graph convolutional networks," in *2022 IEEE International Conference on Big Data (Big Data)*. IEEE, 2022, pp. 3131–3138.

[39] Y. Zhang, H. Zhang, X. Chen, M. Liu, X. Zhu, S.-W. Lee, and D. Shen, "Strength and similarity guided group-level brain functional network construction for MCI diagnosis," *Pattern Recognition*, vol. 88, pp. 421–430, 2019.

[40] M. P. Van Den Heuvel and H. E. H. Pol, "Exploring the brain network: a review on resting-state fMRI functional connectivity," *European neuropsychopharmacology*, vol. 20, no. 8, pp. 519–534, 2010.

[41] M. Defferrard, X. Bresson, and P. Vandergheynst, "Convolutional neural networks on graphs with fast localized spectral filtering," *Advances in neural information processing systems*, vol. 29, 2016.

[42] A. Kazi, S. Shekarforoush, S. Arvind Krishna, H. Burwinkel, G. Vivar, K. Kortüm, S.-A. Ahmadi, S. Albarqouni, and N. Navab, "Inceptiongcn: receptive field aware graph convolutional network for disease prediction," in *Information Processing in Medical Imaging: 26th International Conference, IPMI 2019, Hong Kong, China, June 2–7, 2019, Proceedings 26*. Springer, 2019, pp. 73–85.

[43] D. K. Hammond, P. Vandergheynst, and R. Gribonval, "Wavelets on graphs via spectral graph theory," *Applied and Computational Harmonic Analysis*, vol. 30, no. 2, pp. 129–150, 2011.

[44] T. M. Cover, *Elements of information theory*. John Wiley & Sons, 1999.

[45] M. Mirza and S. Osindero, "Conditional generative adversarial nets," *arXiv preprint arXiv:1411.1784*, 2014.

[46] L. Zhang, L. Wang, D. Zhu, A. D. N. Initiative *et al.*, "Predicting brain structural network using functional connectivity," *Medical image analysis*, vol. 79, p. 102463, 2022.

[47] D. Misra, "Mish: A self regularized non-monotonic activation function," *arXiv preprint arXiv:1908.08681*, 2019.

[48] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[49] P. Velickovic, G. Cucurull, A. Casanova, A. Romero, P. Lio, Y. Bengio *et al.*, "Graph attention networks," *stat*, vol. 1050, no. 20, pp. 10–48 550, 2017.

[50] W. Hamilton, Z. Ying, and J. Leskovec, "Inductive representation learning on large graphs," *Advances in neural information processing systems*, vol. 30, 2017.

[51] P. Seth, A. Bhandari, and K. Lakara, "Analyzing effects of fake training data on the performance of deep learning systems," *arXiv preprint arXiv:2303.01268*, 2023.

[52] N. Tzourio-Mazoyer, B. Landeau, D. Papathanassiou, F. Crivello, O. Etard, N. Delcroix, B. Mazoyer, and M. Joliot, "Automated anatomical labeling of activations in spm using a macroscopic anatomical parcellation of the mni mri single-subject brain," *Neuroimage*, vol. 15, no. 1, pp. 273–289, 2002.

[53] A. Khazaee, A. Ebrahimzadeh, and A. Babajani-Feremi, "Application of advanced machine learning methods on resting-state fMRI network for identification of mild cognitive impairment and alzheimer's disease," *Brain imaging and behavior*, vol. 10, pp. 799–817, 2016.

[54] T. Eslami, V. Mirjalili, A. Fong, A. R. Laird, and F. Saeed, "Asd-diagnet: a hybrid learning approach for detection of autism spectrum disorder using fMRI data," *Frontiers in neuroinformatics*, vol. 13, p. 70, 2019.

[55] B. W. Silverman, *Density estimation for statistics and data analysis*. CRC press, 1986, vol. 26.

[56] A. Casanova, M. Careil, J. Verbeek, M. Drozdzal, and A. Romero Soriano, "Instance-conditioned gan," *Advances in Neural Information Processing Systems*, vol. 34, pp. 27 517–27 529, 2021.

[57] J. Cho, G. Hwang, and C. Suh, "A fair classifier using kernel density estimation," *Advances in neural information processing systems*, vol. 33, pp. 15 088–15 099, 2020.

[58] E. A. Murray, S. P. Wise, and W. C. Drevets, "Localization of dysfunction in major depressive disorder: prefrontal cortex and amygdala," *Biological psychiatry*, vol. 69, no. 12, pp. e43–e54, 2011.

[59] F. A. Kozel, U. Rao, H. Lu, P. A. Nakonezny, B. Grannemann, T. McGregor, P. E. Croarkin, K. S. Mapes, C. A. Tamminga, and M. H. Trivedi, "Functional connectivity of brain structures correlates with treatment outcome in major depressive disorder," *Frontiers in psychiatry*, vol. 2, p. 7, 2011.

[60] J. S. Suh, M. A. Schneider, L. Minuzzi, G. M. MacQueen, S. C. Strother, S. H. Kennedy, and B. N. Frey, "Cortical thickness in major depressive disorder: A systematic review and meta-analysis," *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, vol. 88, pp. 287–302, 2019.

[61] T.-J. Lai, M. E. Payne, C. E. Byrum, D. C. Steffens, and K. R. R. Krishnan, "Reduction of orbital frontal cortex volume in geriatric depression," *Biological psychiatry*, vol. 48, no. 10, pp. 971–975, 2000.

[62] E. Bora, B. J. Harrison, C. G. Davey, M. Yücel, and C. Pantelis, "Meta-analysis of volumetric abnormalities in cortico-striatal-pallidal-thalamic circuits in major depressive disorder," *Psychological medicine*, vol. 42, no. 4, pp. 671–681, 2012.

[63] Y. Lu, H. Liang, D. Han, Y. Mo, Z. Li, Y. Cheng, X. Xu, Z. Shen, C. Tan, W. Zhao *et al.*, "The volumetric and shape changes of the putamen and thalamus in first episode, untreated major depressive disorder," *NeuroImage: Clinical*, vol. 11, pp. 658–666, 2016.

[64] M. A. Lebow and A. Chen, "Overshadowed by the amygdala: the bed nucleus of the stria terminalis emerges as key to psychiatric disorders," *Molecular psychiatry*, vol. 21, no. 4, pp. 450–463, 2016.

[65] K. D. Young, V. Zotev, R. Phillips, M. Misaki, W. C. Drevets, and J. Bodurka, "Amygdala real-time functional magnetic resonance imaging neurofeedback for major depressive disorder: A review," *Psychiatry and clinical neurosciences*, vol. 72, no. 7, pp. 466–481, 2018.

[66] M. S. Korgaonkar, N. J. Cooper, L. M. Williams, and S. M. Grieve, "Mapping inter-regional connectivity of the entire cortex to characterize major depressive disorder: a whole-brain diffusion tensor imaging tractography study," *Neuroreport*, vol. 23, no. 9, pp. 566–571, 2012.