This MICCAI paper is the Open Access version, provided by the MICCAI Society. It is identical to the accepted version, except for the format and this watermark; the final published version is available on SpringerLink.

# Multi-view Graph Contrastive Learning with Dynamic Self-aware and Cross-sample Topology Augmentation for Brain Disorder Diagnosis

Hao Zhang[1], Xiaoyun Liu[2], Shuo Huang[1], Yonggui Yuan[2(✉)], Daoqiang Zhang[3(✉)], and Li Zhang[1(✉)]

[1] College of Information Science and Technology & Artificial Intelligence, Nanjing Forestry University, Nanjing 210037, China
lizhang@njfu.edu.cn

[2] Department of Psychosomatics and Psychiatry, Zhongda Hospital, School of Medicine, Jiangsu Provincial Key Laboratory of Brain Science and Medicine, Southeast University, Nanjing, 210009, China
yygylh2000@sina.com

[3] College of Artificial Intelligence, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China
dqzhang@nuaa.edu.cn

**Abstract.** Resting-state functional MRI (rs-fMRI) has been increasingly employed to aid in brain disorder diagnosis and reveal the pathological mechanisms underlying neurological diseases. However, clinical applications of current automated diagnosing techniques remain constrained by the complexity of brain topology structures and the high costs associated with expert-derived biomarkers. Recent advancements in research have shown that Graph Contrastive Learning (GCL) holds substantial potential for overcoming these challenges and improving diagnosis accuracy. Nevertheless, existing GCL-based methods predominantly generate a static augmented brain network during graph augmentation and primarily focus on the semantic differences between the original and augmented views. To address above issues, we introduce MGCL-DA (**M**ulti-view **G**raph **C**ontrastive **L**earning with **D**ynamic Self-aware and Cross-sample Topology **A**ugmentation), a novel framework aimed at generating two complementary augmentations of brain networks that account for both individual-specific and inter-subject functional heterogeneity, as well as dynamically regulating the update of augmented views to optimize the transmission of discriminative features. Furthermore, we incorporate multi-view graph contrastive learning with min-max constraints, applying distinct contrastive constraints based on specific augmentation semantics to enable pairwise comparisons between the original network and its two augmented views. Extensive experiments on the MDD dataset demonstrate the superior classification performance of MGCL-DA over several state-of-the-arts. Code is available at https://github.com/goodcodingagoodboy/MGCL-DA.

**Keywords:** Multi-view graph contrastive learning · Dynamic graph topology augmentation · Resting-state functional MRI · Brain disorder.

## 1 Introduction

Resting-state functional MRI (rs-fMRI) [4, 15] has been increasingly employed to aid in disorder diagnosis and reveal the pathological mechanisms. Conditions such as major depressive disorder (MDD) [6, 25] pose a range of significant challenges related to brain development and function. However, the clinical applications of current techniques are still limited by the complexity of brain topology structures and the high costs associated with expert-derived biomarkers [7].

Recently, Graph Contrastive Learning (GCL) has demonstrated significant potential in enhancing the exploration of brain topology structure and clinical model generalization, particularly in scenarios with limited labeled data [7, 17]. Nevertheless, its application in brain disorder diagnosis remains relatively unexplored. Existing contrastive methods predominantly generate a static augmented brain network during graph augmentation [13, 23], which overlooks their dynamic adjustment capabilities and hinders the model's ability to capture the complex and evolving nature of brain networks over time or under different conditions [16]. Furthermore, conventional multi-view contrastive learning is typically restricted to capturing differences between the original view and the augmented views [9, 17, 23], neglecting the interrelationships among augmented views and essential semantic information relevant to specific application contexts. This limitation leads to the omission of certain critical feature representations.

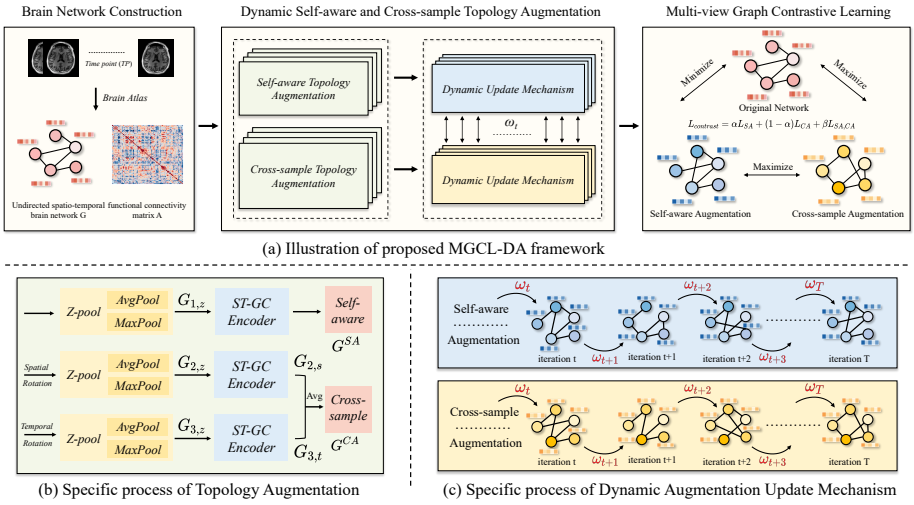



**Fig. 1.** Illustration of proposed MGCL-DA and significant components.

To address these challenges, we introduce MGCL-DA (Multi-view Graph Contrastive Learning with Dynamic Self-aware and Cross-sample Topology Augmentation), a novel framework designed to advance the automated diagnosis of

brain disorders. As illustrated in Fig. 1, the main contributions of proposed MGCL-DA framework can be summarized as follows: (1) We develop a dynamic self-aware and cross-sample topology augmentation to iteratively regulate the dynamic update of augmented views and generate two complementary brain network views that account for both individual-specific and inter-subject functional heterogeneity. (2) Multi-view graph contrastive learning with min-max constraints is further proposed, applying distinct contrastive constraints based on specific augmentation semantics to enable pairwise comparisons between the original network and its two augmented views. (3) Extensive experiments on the MDD dataset collected from two hospitals demonstrate that our MGCL-DA has superior classification performance than existing state-of-the-arts methods. Additionally, ablation studies and parameter analysis confirm the effectiveness of proposed dynamic update mechanism and multi-view contrastive constraints.

## 2 Methodology

As shown in Fig. 1(a), the proposed MGCL-DA framework consists of three parts: (1) brain network construction, (2) dynamic self-aware and cross-sample topology augmentation, and (3) multi-view graph contrastive learning, with details introduced below.

### 2.1 Brain Network Construction

In this study, we employ rs-fMRI to construct an undirected spatio-temporal brain network $G$ and its corresponding functional connectivity matrix $A$.

For each subject, blood-oxygen-level-dependent (BOLD) signals from rs-fMRI represent an undirected spatio-temporal network $G = (V, E) \in R^{T\times N}$. The vertex set $V = \{v_{ti} \mid t = 1, \dots T; i = 1, \dots, N\}$ includes each ROI (region of interest) at every time point $t$, where $T$ denotes the number of time points during scanning, and $N$ is the number of brain regions. The edge set $E$ contains two subsets [3]: (1): $E_R = \{(v_{ti}, v_{tj})\}$, representing the set of connections between the $i$-th and $j$-th ROIs at the same time point $t$. (2): $E_T = \{(v_{ti}, v_{t+1,i})\}$, representing the connections between consecutive time points for the same $i$-th ROI. This approach allows the undirected spatio-temporal network $G$ of each subject to express topology expressions from both spatial and temporal dimensions [10].

Subsequently, the functional connectivity matrix $A \in R^{N\times N}$ is generated by computing the Pearson correlation coefficient for all ROIs at each time point, aiming to identify ROIs with dysregulation in patients with brain disorders.

### 2.2 Dynamic Self-aware and Cross-sample Topology Augmentation

To fully leverage the topology structure of the brain network and regulate the transmission of discriminative information in augmented views, we propose a dynamic self-aware and cross-sample topology augmentation to generate two complementary augmentations of brain network for both individual-specific and inter-subject functional heterogeneity [24].

**Self-aware and Cross-sample Topology Augmentation.** Considering the effect of sample sizes in graph topology augmentation, we extend the brain network $G$ and $A$ into three dimensions, represented as $R^{B\times T\times N}$ and $R^{B\times N\times N}$.

As for self-aware augmentation, first, the input $G$ is fed into Z-pool layer[12], producing the $G_z$. The transformation equation for the Z-pool layer is given by:

$$\begin{aligned} G_z &= \text{Z-pool}(G) = \text{MaxPool}(G) \oplus \text{AvgPool}(G) \\ &= \left[ \max_{b\in R} (G_{b,T,N}) \oplus \frac{1}{|R|} \sum_{b\in R} G_{b,T,N} \right] \end{aligned} \tag{1}$$

where $G_{b,T,N}$ denotes the $b$-th subject brain network and $R$ represents the local neighborhood in the spatio-temporal domain, with size $|R|$. MaxPool selects the most prominent edges in the brain network, amplifying core functional connections while reducing noise and redundancy. In contrast, AvgPool smooths the network by averaging functional connections, weakening less important ones while preserving the network's overall topology. The operation $\oplus$ merges prominent features with the overall network structure. Next, the Spatio-temporal GC encoder, consisting of spatial and temporal two components, encodes $G_z$ to capture the spatial structure and temporal dynamics of brain connectivity, which operations are consistent with the module proposed by Fang et al [2]. After the above operations, we obtain the self-aware augmentation $G^{SA}$, which represents an augmented brain network that captures the self-specific topological structure.

For cross-sample augmentation, the procedure follows the same steps as described previously. The only difference is that, before inputting, $G$ is rotated separately along the spatial $N$ and temporal $T$ axes and finally obtain two outputs, $G_{2,s}$ and $G_{3,t}$. Then, we employ an equally weighted averaging approach to fuse two augmented views, yielding $G^{CA}$ by leveraging inter-sample variability.

**Dynamic Augmentation Update Mechanism.** To dynamically update augmented views over multiple iterations, we introduce a dynamic augmentation update mechanism. At iteration $t$-th, the augmented brain networks $G_t^{SA}, G_t^{CA}$ are updated from the previously $(t-1)$-th representation, as expressed below:

$$G_t^{SA} = (1-\omega_t)\cdot G_{t-1}^{SA} + \omega_t \cdot G_t^{SA} \qquad G_t^{CA} = (1-\omega_t)\cdot G_{t-1}^{CA} + \omega_t \cdot G_t^{CA} \tag{2}$$

where, $\omega_t$ is a dynamically adjusted weight factor, which plays a significant role in balancing the fusion of discriminative information contained in two consecutive augmented views. Meanwhile, it is uniformly shared among all subjects within the same iteration for two topology augmentation, but dynamically varies across iterations. Specifically, the implementation of $\omega_t$ is provided as follows:

$$\omega_t = \omega_{min} + (\omega_{max} - \omega_{min}) \cdot \sigma[\frac{1}{2}(\frac{L_t - L_{t-1}}{L_{t-1}} + \frac{ACC_t - ACC_{t-1}}{ACC_{t-1}})] \tag{3}$$

where $L$ is the objective loss function of MGCL-DA and $ACC$ is the training accuracy for MDD classification. When the model exhibits suboptimal performance in the early stages of training, $\omega_t$ is increased to maximally eliminate the

interfering information from the previous brain network representation. Conversely, when the loss remains low and accuracy is high, $\omega_t$ is decreased to retain the important information from the previous iteration while also facilitating the generation of the current augmented network. $\omega_{min}$ and $\omega_{max}$ are set to 0.25 and 0.95, to constrain the boundaries of dynamic updates. The lower bound $\omega_{min}$ that each epoch makes at least a minimal contribution to the augmented view, preserving essential discriminative information throughout training. The upper bound $\omega_{max}$, on the other hand, allows important information to propagate across epochs, preventing premature convergence and enabling gradual refinement of representations. Moreover, the sigmoid function $\sigma$ is used to constrain fluctuations within the range of $[0, 1]$.

### 2.3 Multi-view Graph Contrastive Learning

We introduce a multi-view graph contrastive learning approach that integrates the min-max contrastive loss $L_{contrast}$ with the classification loss $L_{class}$.

**Min-Max Multi-view Contrastive Loss.** In conventional graph contrastive learning [9, 23], the contrastive loss is typically computed between the augmented and original views. In our semantic scenario, we enforce a minimization constraint between $G^{SA}$ and $G$ to capture disorder-related functional connectivity, while promoting topological diversity between $G^{CA}$ and $G$, as specified below:

$$L_{SA} = \frac{1}{B}\sum_{i=1}^{B} \log \frac{\exp(\text{sim}(G_i^{SA}, G_i)/\tau)}{\sum_{j=1,j\neq i}^{B} \exp(\text{sim}(G_i^{SA}, G_j)/\tau)} \tag{4}$$

$$L_{CA} = -\frac{1}{B}\sum_{i=1}^{B} \log \frac{\exp(\text{sim}(G_i^{CA}, G_i)/\tau)}{\sum_{j=1,j\neq i}^{B} \exp(\text{sim}(G_i^{CA}, G_j)/\tau)} \tag{5}$$

where, $\tau$ is the temperature parameter that controls the smoothness of the similarity which is set to 0.5, consistent with Xu et al [20]. And $\text{sim}(G_i^{SA}, G_i) = \frac{G_i^{SA\top} G_i}{\|G_i^{SA}\|\|G_i\|}$ is the cosine similarity function [19].

To further explore discriminative features between two augmented views $G^{SA}$ and $G^{CA}$, we introduce maximized contrastive constraints $L_{SA,CA}$. This objective is designed to amplify the semantic differences between the two views, as $G^{SA}$ emphasizes individual-specific semantics while $G^{CA}$ captures inter-subject variability. By maximizing their discrepancy, MGCL-DA is encouraged to distinguish and preserve both intra-individual uniqueness and cross-individual differences in expression patterns.

$$L_{SA,CA} = -\frac{1}{B}\sum_{i=1}^{B} \log \frac{\exp(\text{sim}(G_i^{SA}, G_i^{CA})/\tau)}{\sum_{j=1,j\neq i}^{B} \exp(\text{sim}(G_i^{SA}, G_j^{CA})/\tau)} \tag{6}$$

Combining the above multi-view losses, we can obtain the total graph contrastive loss $L_{contrast}$ of our proposed MGCL-DA framework as follows:

$$L_{contrast} = \alpha \cdot L_{SA} + (1-\alpha) \cdot L_{CA} + \beta \cdot L_{SA,CA} \tag{7}$$

where $\alpha$ and $\beta$ are non-negative tuning parameters.

**Objective Loss Function of MGCL-DA.** For downstream brain disorder diagnosis tasks, we employ a composite objective loss incorporating binary cross-entropy loss $L_{class}$. The classification loss is computed based on classification probabilities, which are obtained from a convolutional operation (kernel size: $1 \times 1$) on $G$ followed by a sigmoid activation function. The final objective loss $L$ can be defined as:

$$L = L_{class} + \lambda L_{contrast} \tag{8}$$

where $\lambda$ is a hyper-parameter that balances the contributions of these two loss terms. The subsequent experimental results are all obtained when $\lambda$ is set to 1.

## 3 Experiment

**Dataset and Data Pre-processing.** We evaluate our proposed MGCL-DA framework using one Major Depression Disorder (MDD) rs-fMRI dataset[11, 22]. This dataset includes subjects from two hospitals: Zhongda Hospital of Southeast University and Xinxiang Medical University's Second Affiliated Hospital, which comprises 64 healthy controls (HCs) (31 males, 33 females, age 41.48 $\pm$ 13.44, education 11.84 $\pm$ 4.56 years) and 107 MDD patients (52 males, 55 females, age 39.55 $\pm$ 14.86, education 10.22 $\pm$ 4.36 years). HCs are recruited via advertisements, while MDD patients are from inpatient and outpatient psychiatry departments. MDD patients met DSM-IV criteria, have a first depressive episode after 18, a HAM-D24 score $\geq$ 20, and no major psychiatric or physical illnesses. It is noteworthy that, to ensure consistency in scanning parameters and minimize differences between the two study locations, a designated individual is invited to adjust the magnetic resonance scanning parameters at both sites prior to data collection [11].

The rs-fMRI image data in MDD dataset are preprocessed using DPARSF [21] and the Automated Anatomical Labeling (AAL) template [14], dividing the brain into 116 regions. After discarding the first 10 time points, pre-processing includes: (1) head motion correction (exclusion if motion $>$ 2.0 mm/degrees), (2) co-registration of T1 and functional images, (3) spatial normalization (3 mm voxels), (4) spatial smoothing (6-mm Gaussian kernel), (5) trend removal, (6) nuisance signal regression, (7) band-pass filtering (0.01–0.08 Hz), and (8) time series normalization (zero mean, unit variance).

**Implementation Details and Evaluation Metric.** Moreover, we conduct experiments on PyTorch with a 3090 GPU. In our experiments, the training/test data is randomly split by five-fold cross-validation, and repeated 5 times using random seeds. Classification performance is tested using five metrics: accuracy (ACC), the area under the subject operating characteristic curve (AUC), precision (PRE), recall (REC) and F1 score (F1). Final results are the aggregation of mean values and standard deviation of all methods.

**Table 1.** Experimental results for different methods on MDD datasets.

| Model | ACC (%) | AUC (%) | PRE (%) | REC (%) | F1 (%) |
|---|---|---|---|---|---|
| CC+SVM | $64.23_{\pm 1.11}$ | $50.20_{\pm 0.50}$ | $63.55_{\pm 1.10}$ | $69.83_{\pm 0.78}$ | $66.68_{\pm 0.70}$ |
| SPL+RF | $62.59_{\pm 1.06}$ | $50.25_{\pm 0.74}$ | $62.57_{\pm 1.04}$ | $70.26_{\pm 0.78}$ | $66.16_{\pm 0.68}$ |
| BrainGNN | $73.44_{\pm 1.56}$ | $74.51_{\pm 2.25}$ | $68.74_{\pm 1.99}$ | $74.31_{\pm 3.53}$ | $71.37_{\pm 1.95}$ |
| UFA-NET | $75.67_{\pm 1.37}$ | $76.56_{\pm 2.24}$ | $66.29_{\pm 1.90}$ | $75.60_{\pm 2.30}$ | $70.63_{\pm 1.48}$ |
| UCGL | $74.33_{\pm 1.60}$ | $76.34_{\pm 2.08}$ | $64.92_{\pm 1.88}$ | $75.44_{\pm 2.25}$ | $69.75_{\pm 1.45}$ |
| GCDA | $80.95_{\pm 2.61}$ | $84.26_{\pm 3.68}$ | $72.62_{\pm 7.27}$ | $75.92_{\pm 6.73}$ | $72.25_{\pm 3.17}$ |
| MGCL-DA | $\mathbf{84.76}_{\pm 3.98}$ | $\mathbf{87.99}_{\pm 3.30}$ | $\mathbf{73.07}_{\pm 8.49}$ | $\mathbf{76.22}_{\pm 4.86}$ | $\mathbf{73.46}_{\pm 4.34}$ |

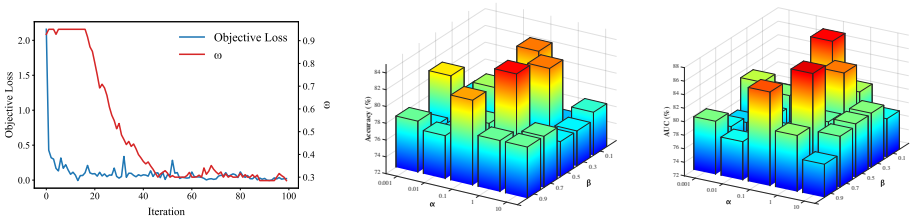
**Table 2.** Ablation study results for MGCL-DA framework on MDD dataset.

| Model | ACC (%) | AUC (%) | PRE (%) | REC (%) | F1 (%) |
|---|---|---|---|---|---|
| MGCL-DA-I | $77.14_{\pm 9.16}$ | $79.74_{\pm 9.36}$ | $64.83_{\pm 4.36}$ | $73.58_{\pm 5.50}$ | $65.36_{\pm 6.74}$ |
| MGCL-DA-II | $78.10_{\pm 5.43}$ | $77.89_{\pm 9.36}$ | $61.43_{\pm 7.25}$ | $65.56_{\pm 9.20}$ | $64.37_{\pm 7.53}$ |
| MGCL-DA-III | $79.05_{\pm 2.61}$ | $80.10_{\pm 5.37}$ | $64.84_{\pm 9.80}$ | $58.33_{\pm 7.59}$ | $59.62_{\pm 8.68}$ |
| MGCL-DA | $\mathbf{84.76}_{\pm 3.98}$ | $\mathbf{87.99}_{\pm 3.30}$ | $\mathbf{73.07}_{\pm 8.49}$ | $\mathbf{76.22}_{\pm 4.86}$ | $\mathbf{73.46}_{\pm 4.34}$ |

**Competing Methods.** We compare the proposed MGCL-DA with six state-of-the-art competing methods on the MDD dataset. These include two traditional machine learning methods, **CC+SVM** [5] and **SPL+RFs** [1], two graph convolutional networks, **Brain-GNN** [8] and **UFA-NET** [2], as well as two recent graph contrastive learning-based models, **UCGL** [17] and **GCDA** [18].

**Classification Performance Results.** The experimental results shown in Table 1 demonstrate that the MGCL-DA achieves state-of-the-art performance in MDD classification, attaining an accuracy of 84.76% and an AUC of 87.99%. It significantly outperforms conventional machine learning methods, with a 20.53% accuracy gain over CC+SVM, and surpasses recent GCN baselines, showing an 11.43% improvement of AUC compared to UFA-NET. Additionally, the 3.81% improvement in accuracy over MGCL-DA highlights the advantage of integrating dynamic network topology augmentation with multi-view contrastive learning compared to the current state-of-the-art (SOTA) GCL methods. Moreover, the balanced precision-recall characteristics indicate effective mitigation of class imbalance issues commonly encountered in neuroimaging-based diagnosis.

**Ablation Study.** To further validate the contribution of the dynamic augmentation update mechanism, we compare MGCL-DA with two of its variants. Specifically, MGCL-DA-I sets $\omega_t = 1$ in Eq. 2, which means the augmented brain network depends solely on the current iteration. MGCL-DA-II sets $\omega_t = 0.5$ to evaluate the influence of a fixed weight on topology augmentation. Meanwhile, to validate the effectiveness of maximizing the discrepancy between $G^{SA}$ and $G^{CA}$, MGCL-DA-III is optimized with $\beta$ initialized to 0 in Eq. 7.




(a) Analysis of Dynamic Weight Factor (b) Analysis of Multi-view Contrastive Loss Hyper-parameters

**Fig. 2.** Analysis of (a) dynamic factor $\omega_t$, (b) multi-view contrast params $\alpha$, $\beta$.

As shown in Table 2, all variants of the MGCL-DA framework exhibit a decline in metrics. Specifically, MGCL-DA-I shows a 7.62% drop in accuracy. The lack of adaptive brain network adjustments during training leads to the loss of critical feature representations, ultimately degrading performance. Similarly, MGCL-DA-II, which fixes the adjusted weight factor, causes an accuracy decrease of 6.66%. This circumstance hinders information flow during training and overlooks differences in functional connectivity expression between consecutive augmented brain networks. MGCL-DA-III exhibits a decrease of 7.89% in AUC, indicating the presence of valuable discriminative information to be explored within the two complementary augmented brain network views.

**Analysis of Dynamic Adjusted Weight Factor.** In Fig. 2(a), we present the variation trend of $\omega_t$ over the first 100 iterations as a representative case. It can be observed that during training, $\omega_t$ consistently decreases as the objective loss progressively reduces. During the early training stages, the framework's limited brain network representation ability and lower classification accuracy cause the augmented views to rely heavily on immediate results. At this stage, the significant decrease in $\omega_t$ reflects the dynamic propagation and updating of discriminative information within the brain network. As the model's fitting capability improves and it learns better to extract discriminative representations of brain networks, $\omega_t$ decreases, assigning greater weight to the augmented views generated in the previous iteration, which aligns with our expectations. Eventually, $\omega_t$ stabilizes around 0.3 at the end of training, indicating a convergence point where the model effectively balances the previous discriminative information and preservation of inter-subject variability for robust representation learning.

From above update mechanism analysis, it can be further hypothesized that, during the training process, the discriminative representation behind brain network augmentation for each subject is selectively retained or removed under the control of the dynamically adjusted $\omega_t$. Throughout the updating process, the framework learns to identify unusual functional connections conducive to MDD classification, while filtering out irrelevant or redundant information.

**Analysis of Multi-view Contrastive Loss Hyper-parameters.** To examine the effects of hyper-parameters in multi-view contrastive loss of Eq. 7, we adjust the values of $\alpha$ and $\beta$. The parameter $\alpha$, which regulates the contribution of self-aware and cross-sample augmentation, is varied from 0.0 to 1.0. Meanwhile, $\beta$, reflecting the impact of differences between augmented brain networks, is tested at 0.001, 0.01, 0.1, 1, and 10. Fig. 2(b) presents a subset of the results for the parameter combinations. Experimental results indicate that the MGCL-DA framework achieves optimal performance when $\alpha = 0.7$ and $\beta = 1$.

## 4 Conclusion

In this work, we introduce MGCL-DA for automated brain disorder diagnosis using rs-fMRI data. Dynamic self-aware and cross-sample topology augmentations effectively capture the complex and evolving nature of brain networks over time. Meanwhile, multi-view graph contrastive learning incorporating min-max contrastive constraints extracts critical discriminative connectivity between the augmented and original brain networks. Extensive experiments and ablation studies on the MDD dataset demonstrate superior performance over SOTA methods and the feasibility of the two proposed components. Moving forward, we will focus on the dynamic brain network representation mechanisms and the contrastive constraints of specific semantic information under multiple views, in order to effectively handle the challenges of understanding complex brain activity patterns and improving the interpretability of multi-view semantic analysis.

**Acknowledgments.** This work was supported in part by the National Natural Science Foundation of China (Nos. 62136004, 62276130), in part by the Zhongda Hospital Affiliated to Southeast University, Jiangsu Province High-Level Hospital Construction Funds (GSP-ZXY13).

**Disclosure of Interests.** The authors have no competing interests to declare that are relevant to the content of this article.

## References

1. Cohen, A.D., Chen, Z., Parker Jones, O., Niu, C., Wang, Y.: Regression-based machine-learning approaches to predict task activation using resting-state fmri. Human brain mapping **41**(3), 815–826 (2020)
2. Fang, Y., Wang, M., Potter, G.G., Liu, M.: Unsupervised cross-domain functional mri adaptation for automated major depressive disorder identification. Medical image analysis **84**, 102707 (2023)
3. Gao, B., Yu, A., Qiao, C., Calhoun, V.D., Stephen, J.M., Wilson, T.W., Wang, Y.P.: An explainable unified framework of spatio-temporal coupling learning with application to dynamic brain functional connectivity analysis. IEEE Transactions on Medical Imaging (2024)
4. Han, X., Xue, R., Du, S., Gao, Y.: Inter-intra high-order brain network for asd diagnosis via functional mris. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 216–226. Springer (2024)

5. Jin, J., Huang, L.: A region-based feature extraction method for rs-fmri of depressive disorder classification. In: 2020 International Conference on Computer Vision, Image and Deep Learning (CVIDL). pp. 707–710. IEEE (2020)
6. Li, L., Zhang, L., Cao, P., Yang, J., Wang, F., Zaiane, O.R.: Exploring spatio-temporal interpretable dynamic brain function with transformer for brain disorder diagnosis. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 195–205. Springer (2024)
7. Li, T., Guo, Y., Zhao, Z., Chen, M., Lin, Q., Hu, X., Yao, Z., Hu, B.: Automated diagnosis of major depressive disorder with multi-modal mris based on contrastive learning: a few-shot study. IEEE Transactions on Neural Systems and Rehabilitation Engineering (2024)
8. Li, X., Zhou, Y., Dvornek, N., Zhang, M., Gao, S., Zhuang, J., Scheinost, D., Staib, L.H., Ventola, P., Duncan, J.S.: Braingnn: Interpretable brain graph neural network for fmri analysis. Medical Image Analysis **74**, 102233 (2021)
9. Lin, Y., Gou, Y., Liu, X., Bai, J., Lv, J., Peng, X.: Dual contrastive prediction for incomplete multi-view representation learning. IEEE Transactions on Pattern Analysis and Machine Intelligence **45**(4), 4447–4461 (2022)
10. Liu, R., Huang, Z.A., Hu, Y., Zhu, Z., Wong, K.C., Tan, K.C.: Spatial–temporal co-attention learning for diagnosis of mental disorders from resting-state fmri data. IEEE transactions on neural networks and learning systems (2023)
11. Liu, S., Fan, D., He, C., Liu, X., Zhang, H., Zhang, H., Zhang, Z., Xie, C.: Resting-state cerebral blood flow and functional connectivity abnormalities in depressed patients with childhood maltreatment: Potential biomarkers of vulnerability? Psychiatry and Clinical Neurosciences **78**(1), 41–50 (2024)
12. Misra, D., Nalamada, T., Arasanipalai, A.U., Hou, Q.: Rotate to attend: Convolutional triplet attention module. In: Proceedings of the IEEE/CVF winter conference on applications of computer vision. pp. 3139–3148 (2021)
13. Peng, L., Wang, N., Xu, J., Zhu, X., Li, X.: Gate: Graph cca for temporal self-supervised learning for label-efficient fmri analysis. IEEE Transactions on Medical Imaging **42**(2), 391–402 (2022)
14. Rolls, E.T., Huang, C.C., Lin, C.P., Feng, J., Joliot, M.: Automated anatomical labelling atlas 3. Neuroimage **206**, 116189 (2020)
15. Shoeibi, A., Khodatars, M., Jafari, M., Ghassemi, N., Moridian, P., Alizadehsani, R., Ling, S.H., Khosravi, A., Alinejad-Rokny, H., Lam, H.K., et al.: Diagnosis of brain diseases in fusion of neuroimaging modalities using deep learning: A review. Information Fusion **93**, 85–117 (2023)
16. Wang, M., Huang, J., Liu, M., Zhang, D.: Modeling dynamic characteristics of brain functional connectivity networks using resting-state functional mri. Medical image analysis **71**, 102063 (2021)
17. Wang, X., Chu, Y., Wang, Q., Cao, L., Qiao, L., Zhang, L., Liu, M.: Unsupervised contrastive graph learning for resting-state functional mri analysis and brain disorder detection. Human Brain Mapping **44**(17), 5672–5692 (2023)
18. Wang, X., Fang, Y., Wang, Q., Yap, P.T., Zhu, H., Liu, M.: Self-supervised graph contrastive learning with diffusion augmentation for functional mri analysis and brain disorder detection. Medical Image Analysis **101**, 103403 (2025)
19. Wang, Y., Han, Y., Wang, H., Zhang, X.: Contrast everything: A hierarchical contrastive framework for medical time-series. Advances in Neural Information Processing Systems **36**, 55694–55717 (2023)
20. Xu, Y., Wang, J., Guang, M., Yan, C., Jiang, C.: Graph contrastive learning with min-max mutual information. Information Sciences **665**, 120378 (2024)

21. Yan, C., Zang, Y.: Dparsf: a matlab toolbox for" pipeline" data analysis of resting-state fmri. Frontiers in systems neuroscience **4**, 1377 (2010)
22. Zhang, L., Pang, M., Liu, X., Hao, X., Wang, M., Xie, C., Zhang, Z., Yuan, Y., Zhang, D.: Incorporating multi-stage diagnosis status to mine associations between genetic risk variants and the multi-modality phenotype network in major depressive disorder. Frontiers in Psychiatry **14**, 1139451 (2023)
23. Zhang, S., Chen, X., Shen, X., Ren, B., Yu, Z., Yang, H., Jiang, X., Shen, D., Zhou, Y., Zhang, X.Y.: A-gcl: Adversarial graph contrastive learning for fmri analysis to diagnose neurodevelopmental disorders. Medical Image Analysis **90**, 102932 (2023)
24. Zhang, Z., Ran, R., Tian, C., Zhou, H., Li, X., Yang, F., Jiao, Z.: Self-aware and cross-sample prototypical learning for semi-supervised medical image segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 192–201. Springer (2023)
25. Zhao, T., Zhang, G.: Detecting major depressive disorder by graph neural network exploiting resting-state functional mri. In: International Conference on Neural Information Processing. pp. 255–266. Springer (2022)