G-026

Multimodal Hypergraph Attention for Identification of Depression using rs-fMRI Data

シャ　ヨウギ‡　　　鎌田　清一郎‡
Yangyi SHE　　　Sei-ichiro Kamata

1. Abstract

Major Depressive Disorder (MDD) is a prevalent mental illness that poses significant diagnostic challenges due to complex brain network disruptions. Resting-state functional MRI (rs-fMRI) has emerged as a promising tool for identifying MDD-related neural alterations. However, conventional models often overlook the temporal dynamics and higher-order interactions of brain connectivity. In this study, I propose a Multimodal Hypergraph Attention Network that integrates multi-scale dynamic functional connectivity networks (dFCNs), curvelet-based frequency analysis, and phenotypic data for robust MDD classification. The framework constructs personalized hypergraphs that capture fine-grained temporal and spatial variations in brain activity. Experimental results on a large-scale rs-fMRI dataset demonstrate superior performance compared to state-of-the-art methods, validating the effectiveness of the proposed multimodal approach.

2. Introduction

Major Depressive Disorder (MDD) is a serious and prevalent mental health condition affecting over 280 million people globally [1]. The disorder presents a wide array of symptoms, including persistent sadness, anhedonia, sleep disturbances, fatigue, and cognitive dysfunction, leading to a significant decline in daily functioning and quality of life [1][7]. Despite its high prevalence and profound impact, the diagnosis of MDD remains largely subjective, relying on self-reported symptoms and clinical judgment, which may lead to inconsistent and delayed interventions.

Recent advances in neuroimaging, particularly resting-state functional magnetic resonance imaging (rs-fMRI), offer promising tools for objective diagnosis and characterization of MDD. Rs-fMRI leverages Blood Oxygen Level-Dependent (BOLD) signals to map brain activity and has been used to infer functional connectivity (FC) between different regions of interest (ROIs) [7]. Several studies have revealed that MDD is associated with altered FC in key networks such as the default mode network (DMN), frontoparietal network (FPN), and salience network (SN) [7][11]. However, many existing approaches to FC analysis rely on static FCNs (sFCNs), which are derived from the entire scanning session and assume consistent interactions over time. This simplification ignores the intrinsic dynamics of brain networks, especially in psychiatric disorders where functional variability may be a critical biomarker [9].

Dynamic FCNs (dFCNs), constructed using sliding window techniques, allow the capture of transient states in brain connectivity. These networks can reveal fluctuating patterns that better reflect the temporal nature of neural interactions in MDD [9][10]. Yet, current models often treat each connection equally or combine multiple scales in a naive way, without accounting for differences in relevance or frequency across temporal windows.

To overcome limitations of pairwise-only FC modeling, recent research has proposed the use of hypergraphs. Unlike conventional graphs, hypergraphs enable modeling of high-order relationships by connecting more than two ROIs within a single hyperedge. This structure captures complex inter-regional interactions that are more representative of true brain dynamics [5][6]. For example, a hyperedge may represent synchronized activity across a functional subnetwork rather than simple dyadic interactions.

Additionally, existing diagnostic models rarely integrate phenotypic information such as age, sex, or symptom severity. These features may significantly influence neural connectivity patterns and can enhance the personalization and generalizability of machine learning models for MDD classification [11].

To address these challenges, I propose a multimodal hypergraph attention model that integrates (1) multi-scale dFCNs, capturing brain dynamics at various temporal frequencies; (2) curvelet-transformed frequency-aware features to weight different graph scales intelligently; and (3) phenotypic data for subject-specific modeling. My approach aims to construct a robust, interpretable, and clinically relevant framework for supporting MDD diagnosis based on rs-fMRI.

3. Related Work

3.1 Functional Connectivity Networks in MDD

Functional connectivity (FC) analysis using rs-fMRI has become a cornerstone in the study of Major Depressive Disorder (MDD). Traditional static FCNs assume time-invariant connectivity by computing Pearson correlations between pairs of ROIs across the entire scanning session. This approach is limited because it neglects transient connectivity changes that are essential in psychiatric conditions like MDD [7]. To address this, dynamic FCNs (dFCNs) have been introduced, where connectivity is computed over sliding windows, allowing the model to detect rapid transitions in brain network states [9][10]. These dynamic patterns offer improved sensitivity in distinguishing patients from controls and provide a richer temporal characterization of brain activity.

3.2 Hypergraph Learning for Brain Network Modeling

Conventional graph-based models capture only pairwise relationships between ROIs, which limits their ability to represent higher-order dependencies in brain function. Hypergraphs, where a hyperedge can connect more than two nodes, offer a natural extension to model multi-region interactions. Recent studies have applied hypergraph neural networks (HGNNs) to neuroimaging data, demonstrating their effectiveness in disorders like MDD and migraine [5][6]. Hypergraphs are particularly suited to brain data due to their ability to encode modular, overlapping, and hierarchical connectivity patterns—properties that are common in functional brain networks.

3.3 Multimodal and Personalized MDD Modeling

Another important trend in MDD research is the incorporation of multimodal data, such as combining rs-fMRI with clinical or demographic information. Models that integrate phenotypic features like age, sex, or symptom severity have shown improved classification accuracy and better generalization across individuals and cohorts [11]. Moreover, recent works have emphasized the need for personalized representations that adapt to subject-specific brain dynamics. Such strategies are especially important in mental health applications, where symptom profiles and neurobiological correlates vary widely across patients.

4. Methodology

The proposed framework aims to improve the classification of Major Depressive Disorder (MDD) by modeling brain dynamics across multiple temporal and spectral scales, and incorporating both functional and non-functional (phenotypic) data for personalized predictions. As shown in Figure 1, the architecture is composed of four major components:

1. Multi-scale Dynamic Functional Connectivity Construction

Resting-state fMRI time series are segmented into multiple overlapping time windows of varying lengths to extract dynamic connectivity patterns at different temporal resolutions.

2. Frequency-Aware Hypergraph Fusion via Curvelet Attention

Curvelet transform is applied to each ROI time series to extract frequency-domain features. These features are then used to compute attention weights across different temporal scales, guiding the fusion of hypergraphs generated from each dFCN.

3. Phenotypic Feature Integration

Subject-level demographic and clinical features (e.g., age, sex, symptom scores) are embedded and fused with hypergraph representations to enhance personalization and generalization.

4. Hypergraph Convolutional Classification

A Hypergraph Convolutional Network (HGCN) processes the fused graph to classify subjects as MDD or normal control, capturing higher-order interactions across ROIs.

Each component is designed to work synergistically, providing robustness and flexibility in modeling complex neural dynamics relevant to depression.

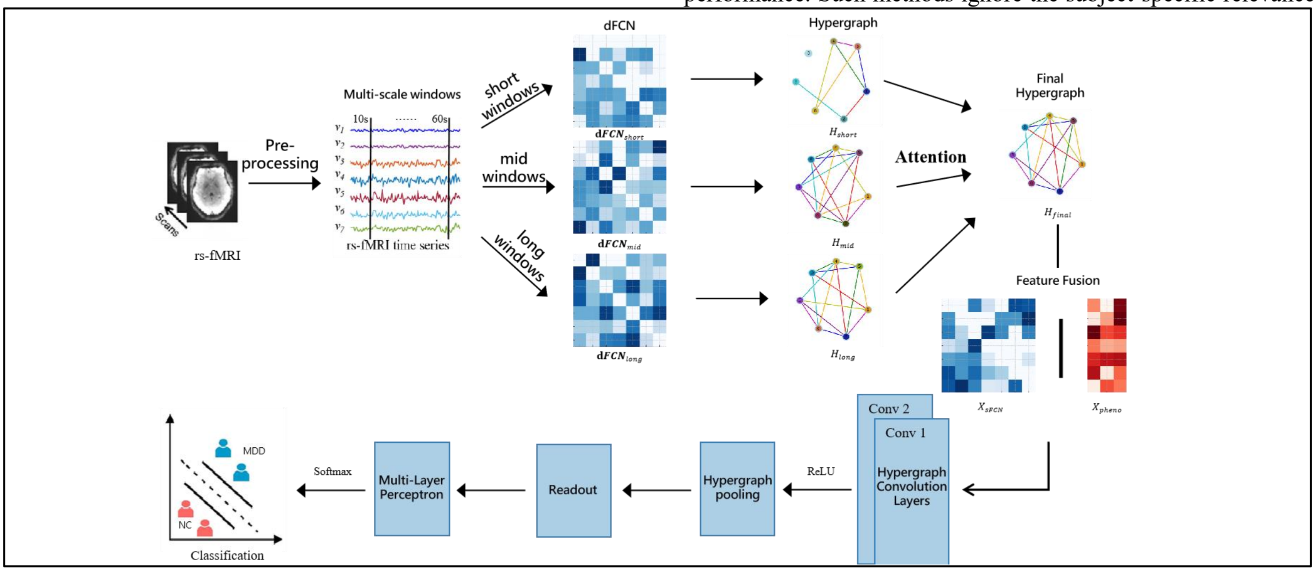

Fig1. The overall pipeline of the proposed model, which consists of multi-scale dFCN construction, Curvelet-based feature extraction, α-weighted hypergraph fusion, and final classification via HGCN.

4.1 Multi-Scale Dynamic Functional Connectivity Construction

Resting-state brain dynamics manifest across a range of temporal resolutions. To capture these nuances, I segment each subject's rs-fMRI time series into three predefined temporal scales:

- Short (10s, 14s, 18s)
- Mid (24s, 28s, 32s)
- Long (40s, 50s, 60s)

For each scale, I construct dynamic functional connectivity networks (dFCNs) by sliding the window across the full time series. Within each window, I compute Pearson correlation coefficients between pairs of regions of interest (ROIs), resulting in a sequence of functional connectivity matrices that reflect time-varying inter-regional interactions.

To avoid manual hyperparameter tuning and ensure optimal performance, I apply Bayesian Optimization to select the most informative window length within each temporal scale. This optimization process is guided by preliminary classification performance, enabling a data-driven selection of the most discriminative temporal configuration either globally or per subject.

Each selected window then yields one dFCN, capturing transient but meaningful connectivity patterns. These dFCNs form the basis for further hypergraph construction, enabling my model to represent short- and long-term brain dynamics simultaneously.

To ensure sparsity and reduce noise, I later apply a fixed threshold (0.7) on the correlation values during hypergraph generation, retaining only strong connections. This helps prevent overfitting and ensures that only salient patterns are preserved for downstream analysis.

4.2 Frequency-Aware Hypergraph Attention using Curvelet Transform

While multi-scale dFCNs capture dynamic interactions at different temporal resolutions, directly fusing the resulting graphs using uniform or heuristic weights can lead to suboptimal performance. Such methods ignore the subject-specific relevance of different time scales and neglect the frequency-domain characteristics of neural signals.

To address this, I introduce a frequency-aware attention mechanism that leverages the Curvelet Transform, a multi-scale, multi-directional analysis tool particularly effective in capturing localized and directional variations in time-series data.

My approach proceeds as follows:

1. Curvelet Feature Extraction

For each ROI's time series in each selected dFCN, I apply the Fast Discrete Curvelet Transform (FDCT). The resulting coefficients capture energy distributions across multiple scales and angles.

2. Mean Frequency Descriptor Computation

I compute the mean Curvelet coefficients across scales and orientations for each ROI. These form a robust, compressed descriptor of temporal frequency patterns unique to each region and scale.

3. Attention Weight Estimation

Using the frequency descriptors from all ROIs within a dFCN, we compute scalar attention weights that reflect the relative importance of each scale for each subject. These weights are normalized to ensure interpretability and comparability.

4. α-Weighted Hypergraph Fusion

Each hypergraph, constructed from a dFCN, is scaled by its corresponding attention weight. The fused hypergraph is obtained by summing all weighted hypergraphs, resulting in a single, personalized graph that emphasizes the most informative time-frequency patterns for that subject.

This fusion strategy preserves critical temporal information while suppressing redundant or noisy representations, thereby enabling the model to focus on discriminative neural dynamics. Compared to equal-weight fusion, my frequency-aware method significantly improves robustness and interpretability across subjects.

4.3 Integration of Phenotypic Data

In my framework, phenotypic data, such as age, gender, and clinical assessment scores, are not treated as peripheral metadata, but are explicitly embedded into the graph learning pipeline to enhance subject-specific modeling.

Rather than injecting phenotypic features directly into the classifier, I integrate them with the subject's static functional connectivity network (sFCN), which is computed by correlating the full-length rs-fMRI time series across all ROI pairs. This yields a stable, subject-specific connectivity matrix reflecting long-term co-activation patterns.

To fuse phenotypic data with the sFCN, I proceed as follows:

1. sFCN Construction: A 160 × 160 correlation matrix is computed using the entire time series.
2. Phenotypic Feature Embedding: A low-dimensional feature vector is extracted from metadata and normalized across the population.
3. Matrix Fusion: The phenotypic vector is projected and either concatenated with the sFCN, producing a phenotype-informed matrix.

Importantly, this phenotype-enhanced sFCN is used as an auxiliary input, complementing the primary α-weighted hypergraph (H_{final}) obtained from dynamic multi-scale dFCNs and frequency-aware attention. Both the fused matrix and the final hypergraph are jointly fed into the Hypergraph Convolutional Network (HGCN), which learns to integrate both dynamic and static sources of connectivity − modulated by individual traits−into a unified predictive representation.

This dual-input design improves the model's ability to capture both long-term and short-term connectivity alterations in MDD, while leveraging subject-specific priors for enhanced personalization and robustness.

4.4 Hypergraph Convolutional Network for Classification

To perform final subject-level classification, I utilize a Hypergraph Convolutional Network (HGCN) that takes as input two complementary representations:

The fused hypergraph (H_{final}), which encodes multi-scale, frequency-aware dynamic brain connectivity patterns; and

An auxiliary phenotype-informed connectivity matrix, derived from static functional connectivity (sFCN) modulated by subject metadata.

The HGCN is designed to model high-order interactions among ROIs, which are naturally represented in a hypergraph structure. Unlike traditional GCNs that operate on pairwise edges, HGCNs allow each node (ROI) to aggregate information from all other nodes within a shared hyperedge−making them well-suited for capturing groupwise co-activation and complex neural dependencies.

The network architecture consists of the following key modules:

- Dual Input Fusion: Both H_final and the phenotype-aware matrix are embedded into a unified feature space. The fused matrix can be treated as an additional node feature channel. This joint encoding ensures both temporal and demographic signals contribute to the learned representation.
- Hypergraph Convolution Layers: We apply two successive convolutional layers that propagate signals through hyperedges, allowing each ROI to update its embedding based on the high-order context. Nonlinear activation functions are used between layers to increase model capacity.
- Hypergraph Pooling and Readout: A pooling operation reduces the node-level features to a compact graph-level embedding. This stage aggregates information across the entire brain network and supports subject-level inference.
- Multi-Layer Perceptron and Softmax: The final graph-level embedding is passed to a fully connected MLP and softmax classifier to predict whether the subject is an MDD patient or a normal control.

This architecture ensures that the classification decision is informed by both individualized temporal brain dynamics and subject-level clinical and demographic context.

5. Experiments and Results

To evaluate the effectiveness of the proposed framework, I conducted experiments on the REST-meta-MDD dataset, the largest publicly available rs-fMRI dataset for MDD.

5.1 Dataset

The dataset includes 1,560 subjects (844 MDD patients, 716 controls) from 25 sites in China. Each subject's data consists of ~240 volumes and was parcellated into 160 ROIs using the Dosenbach atlas.

5.2 Preprocessing

The preprocessing pipeline includes the following steps: (1) slice timing correction; (2) head motion correction; (3) spatial normalization to MNI space; (4) band-pass filtering in the range of 0.01–0.1 Hz; and (5) extraction of time series from 160 ROIs based on the Dosenbach 160 atlas. Subjects with excessive head motion (>2.5 mm or >2.5°) were excluded.

5.3 Experiment Design

Data were split into training (80%) and testing (20%) sets, stratified by diagnosis label to maintain class balance. A 0.7 correlation threshold was used for graph construction. Performance was evaluated using accuracy, precision, recall, F1, and AUC.

5.4 Results

The proposed method achieved over 72% accuracy and AUC > 0.7. It outperformed static FCN, single-scale dFCN, and HGCN variants without phenotypic data or Curvelet attention.

Table 1. Classification accuracy and AUC comparison across baseline and state-of-the-art models on rs-fMRI data for MDD identification.

Method	Subjects (MDD/NC)	Accuracy (%)	AUC
SVM	285	61.0	0.67
SVM+ MLP	300	<60.0	0.66
GCN	450	68.5	0.73
DGCNN	830/771	72.1	-
Ensemble GNN (GCN + GAT +GraphSAGE)	-	71.2	-
My	**400**	**72.4**	**0.76**

5.5 Ablation Study

Removing multi-scale modeling, frequency-aware attention, phenotypic integration, or using GCN instead of HGCN all led to performance drops, confirming the contribution of each component.

Table 2. Ablation Study of Model Components on MDD Classification Performance

Variant	**Removed Component**	**Accuracy (%)**	**AUC**
Full model	-	72.4	0.76
w/o multi-scale dFCN	Only 1-scale	71.3	0.74
w/o HAT (no attention)	Equal weight for hyperedges	70.4	0.71
w/o phenotypic fusion	No age/gender features	72.0	0.75

6. Conclusion and Future Work

6.1 Conclusion

I proposed a novel multimodal framework for MDD diagnosis using rs-fMRI. By combining multi-scale dynamic FCNs, Curvelet-based hypergraph attention, and phenotypic data integration, My model achieved state-of-the-art results on the REST-meta-MDD dataset.

6.2 Future Work

In future work, I plan to (1) validate my framework on additional large-scale, multi-center datasets; (2) incorporate other modalities such as structural MRI or genetic profiles; and (3) explore the applicability of my model to early-stage prediction and longitudinal monitoring of depression progression.

References

[1] Organization, W.H., et al.: Depression and Other Common Mental Disorders: Global Health Estimates. World Health Organization, Technical report (2017)
[2] Frédérique LiégeoisRachael Elward, Neurocognitive Development: Disorders and Disabilities. Handbook of Clinical Neurology (2020)
[3] Dai, Q., Gao, Y. . Mathematical Foundations of Hypergraph. In: Hypergraph Computation. Artificial Intelligence: Foundations, Theory, and Algorithms. Springer, Singapore. https://doi.org/10.1007/978-981-99-0185-2_2 (2023)
[4] Guo H, Huang X, Wang C, Wang H, Bai X, Li Y. High-Order line graphs of fMRI data in major depressive disorder. Med Phys, 51: 5535‐5549. https://doi.org/10.1002/mp.17119 (2024)
[5] Jingyu Liu, Wenxin Yang, Yulan Ma, Qunxi Dong, Yang Li, Bin Hu, Effective hyper-connectivity network construction and learning: Application to major depressive disorder identification, Computers in Biology and Medicine, Volume 171, (2024), 108069, ISSN 0010-4825, https://doi.org/10.1016/j.compbiomed.2024.108069.
[6] Shen, G., Zeng, W. & Yang, J. Research on migraine classification model based on hypergraph neural network. J Supercomput 80, 25403‐25423 (2024). https://doi.org/10.1007/s11227-024-06387-0
[7] Yan, C.G., et al.: Reduced default mode network functional connectivity in patients with recurrent major depressive disorder. Proc. Nat. Acad. Sci. 116(18), 9078‐9083 (2019)
[8] Venkatapathy S, Votinov M, Wagels L, Kim S, Lee M, Habel U, Ra I-H and Jo H-GEnsemble graph neural network model for classification of major depressive disorder using whole-brain functional connectivity. Front. Psychiatry 14:1125339. doi: 10.3389/fpsyt.2023.1125339 (2023)
[9] Zhu M, Quan Y and He X, The classification of brain network for major depressive disorder patients based on deep graph convolutional neural network. Front. Hum. Neurosci. 17:1094592. doi: 10.3389/fnhum.2023.1094592 (2023)
[10] Long D, Zhang M, Yu J, Zhu Q, Chen F and Li F Intelligent diagnosis of major depression disease based on multi-layer brain network. Front. Neurosci. 17:1126865. doi: 10.3389/fnins.2023.1126865 (2023)
[11] Gallo, S., El-Gazzar, A., Zhutovsky, P. et al. Functional connectivity signatures of major depressive disorder: machine learning analysis of two multicenter neuroimaging studies. Mol Psychiatry 28, 3013 - 3022 (2023). https://doi.org/10.1038/s41380-023-01977-5
[12] Wang et al. Diagnosis of major depressive disorder using a novel interpretable adaptive-propagation GCN. Neuroscience. PMID: 39730018. (2024)
[13] Xia, Z. et al. DepressionGraph: A Two-Channel GNN for the Diagnosis of MDD Using rs-fMRI. Electronics, 12(24), 5040. (2023)
[14] El Gazzar, A., Thomas, R. M., & Van Wingen, G. Improving the diagnosis of psychiatric disorders with self-supervised graph-state space models. arXiv preprint [Online] 2206.03331. (2022)
[15] Ye et al. Changes of functional brain networks in MDD: A graph-theoretical analysis of resting-state fMRI. PLOS ONE, 10(9): e0133775. (2015)
[16] Steinsträter et al. Altered resting-state functional connectome in major depressive disorder. Translational Psychiatry, 11, 511. (2021)
[17] Pan et al. Multi-atlas ensemble graph neural network model for major depressive disorder detection using functional MRI. Frontiers in Computational Neuroscience. (2025)
[18] El Gazzar, A., Thomas, R. M., & Van Wingen, G. Benchmarking Graph Neural Networks for fMRI analysis. arXiv preprint 2211.08927. (2022)