ITC 2/55
Information Technology and Control
Vol. 55 / No. 2/ 2026
pp. 901-916
DOI 10.5755/j01.itc.55.2.44579

WG-Mamba: A Wavelet-Graph Mamba Network with Multi-Scale Spatiotemporal-Spectral Fusion for Major Depressive Disorder Diagnosis

Received 2026/03/18 | Accepted after revision 2026/06/19

HOW TO CITE: Yan, J., Gui, R., Liang, H. (2026) WG-Mamba: A Wavelet-Graph Mamba Network with Multi-Scale Spatiotemporal-Spectral Fusion for Major Depressive Disorder Diagnosis. *Information Technology and Control*, 55(2), 901-916. https://doi.org/10.5755/j01.itc.55.2.44579

WG-Mamba: A Wavelet-Graph Mamba Network with Multi-Scale Spatiotemporal-Spectral Fusion for Major Depressive Disorder Diagnosis

Jian Yan, Renzhou Gui*, Hao Liang
College of Electronic and Information Engineering, Tongji University, Shanghai 201804, China

Corresponding author: 2180133@tongji.edu.cn

Major Depressive Disorder (MDD) is a mental disorder negatively impacting the lives of numerous patients worldwide. Currently, functional magnetic resonance imaging (fMRI) data are widely used for computer-aided MDD diagnosis. However, fMRI samples exhibit high-dimensional complexity and dynamic dependency, leading to suboptimal classification accuracy in traditional deep learning networks. To address these challenges, we propose WG-Mamba, a novel Wavelet-enhanced Graph Mamba network with multi-scale spatiotemporal-spectral fusion architecture, for feature extraction and classification of fMRI data between MDD patients and healthy controls (HCs). First, the frequency block is designed to decompose fMRI signals into multi-scale frequency features and filter noise through wavelet transform. Second, the spatial block utilizes a graph neural network with a brain region adjacency matrix to model cross-regional dependencies, capturing spatial features of the brain network. Third, the temporal block is designed based on Mamba to extract temporal dynamics. Finally, a dynamic weight mechanism is designed to achieve deep coupling of multi-scale features. The five-fold cross-validation on the REST-meta-MDD dataset shows WG-Mamba achieves 83.72% classification accuracy. The model also exhibits strong interpretability by identifying core brain regions associated with MDD, including the prefrontal cortex, anterior cingulate cortex, and limbic system, providing novel insights into MDD neurobiological mechanisms.

KEYWORDS: Functional magnetic resonance imaging, Deep learning, Functional connectivity, Mamba, Major depressive disorder.

1. Introduction

Major depressive disorder (MDD) represents a highly prevalent affective disorder, with core clinical features encompassing persistent low mood, impaired cognitive function, and elevated risk of suicidal behavior [23]. Statistics recently released by the World Health Organization (WHO) indicate that the global number of individuals affected by depressive disorders has reached approximately 280 million [25].

This psychiatric condition is characterized by chronic low mood and anhedonia, which can contribute to a wide spectrum of psychological disturbances, physical comorbidities, and potentially life-threatening outcomes [22]. At present, the clinical diagnostic workflow for major depressive disorder mainly depends on two core dimensions: self-reported symptom descriptions from patients, and subjective assessment scales administered by psychiatrists. Both of these approaches lack quantitative biological markers, making it difficult to precisely capture the aberrant neural activity underlying the pathogenesis of the disease [4]. A key challenge in MDD research and clinical practice is its pronounced heterogeneity. Patients exhibit diverse symptom profiles, disease trajectories, and treatment responses, rooted in distinct neurobiological mechanisms. There is thus a pressing clinical demand to establish reliable, quantitative approaches for detecting depressive disorders and exploring their associated neurobiological markers to serve as supplementary tools for clinical diagnosis [17].

Currently, a variety of in-vivo neuroimaging modalities have been widely applied as core tools in neurological disorder research, covering structural and functional magnetic resonance imaging (MRI), scalp and intracranial electroencephalography (EEG), magnetoencephalography (MEG) for neural oscillation mapping, positron emission tomography (PET) for molecular imaging, as well as diffusion tensor imaging (DTI) for white matter tract analysis; these techniques play an indispensable role in routine clinical evaluation, pathological mechanism exploration, and computer-aided diagnostic decision support [11].

Among in-vivo neuroimaging modalities, blood oxygen level-dependent functional magnetic resonance imaging (BOLD-fMRI) has been widely adopted to investigate the pathophysiological mechanisms underlying major depressive disorder, primarily owing to its three core advantages: no ionizing radiation exposure, superior spatial localization precision, and the ability to capture whole-brain neural activity patterns across multiple dimensions [20]. Fang et al. distinguished two discrete MDD neural subtypes with entirely divergent functional connectivity profiles using resting-state functional MRI (rs-fMRI) recordings, providing key insight into the heterogeneous nature of depression [8]. These findings confirm that fMRI has deepened the understanding of the pathogenesis of MDD and established a new diagnostic foundation.

Driven by the rapid development of artificial intelligence techniques, especially the widespread application of deep learning (DL) architectures, the interdisciplinary fusion of data-driven algorithmic models and resting-state fMRI neuroimaging datasets has introduced an innovative research framework to address the shortcomings of conventional neuroimaging analysis pipelines, and to support the identification of individualized neurobiological biomarkers for major depressive disorder [21]. Many studies have employed conventional ML methods such as support vector machines (SVM) and logistic regression (LR) for classification [32]. Zhang et al. utilized SVM to distinguish between subclinical depression, MDD, and healthy controls (HCs), achieving a classification accuracy of approximately 77.6% [35]. Yamashita et al. developed a logistic regression algorithm for classifying MDD from HCs, with a classification accuracy of about 70% [28].

Nevertheless, conventional machine learning frameworks rely heavily on the quality of handcrafted feature extraction performed during the preprocessing step, and present inherent constraints when it comes to modeling the topological organizational patterns and temporal dynamic characteristics of brain networks. Deep learning methods, capable of capturing higher-order nonlinear features, are increasingly being applied in psychiatric classification. Wei et al. adopted a 3D deep CNN to co-train heterogeneous features obtained from rs-fMRI and sMRI scans, attaining a classification accuracy of 69.38% in MDD patient identification tasks [24]. Dai et al. applied a Transformer encoder framework to extract functional coupling features from large multi-center rs-fMRI cohorts, obtaining an average classification accuracy

of 78.11% for MDD versus healthy control identification [5]. In practice, given the complexity of brain network topology, graph convolutional neural network (GCN) models may be more suitable than deep learning architectures such as CNNs for processing fMRI data reflecting human brain functional information [26]. Dai et al. proposed a novel feature extraction method incorporating dynamic temporal information, capturing functional connectivity (FC) at specific time points by calculating Pearson correlation coefficients between brain region sequences across time points. These features were then imported into a GCN to carry out MDD classification, attaining 75.8% classification accuracy and effectively pinpointing depression-related brain regions [6].

The remarkable performance of deep learning stems from its ability to extract effective representations from data, thereby supporting predictive modeling tasks [13]. Blood oxygen level-dependent (BOLD) signals captured by fMRI are inherently time-series containing multi-frequency band neural rhythm information, encompassing multi-scale dynamic information ranging from fast transient neural activities to slow fluctuating brain network states. However, the aforementioned deep learning methods primarily extract features from the spatial domain of the brain. This focus may limit their capacity to capture temporal dynamics of interregional interactions. Recurrent neural networks (RNNs) are suitable for time-series feature extraction, but their inherent gradient vanishing/explosion problem makes it difficult to model long-term temporal dependencies, often leading to loss of critical early-stage brain activity features [31]. Although Transformer-based models achieve efficient modeling of global temporal relationships through self-attention mechanisms, their self-attention operations exhibit quadratic-time complexity with sequence length, resulting in high computational costs [33]. In recent years, modern structured state-space models (SSMs) [9], such as Mamba [10], have emerged as efficient long-sequence modeling methods with linear complexity, providing novel approaches for modeling both local and global dependencies [19].

In this paper,we propose a novel multi-scale spatio-temporal-spectral fusion architecture, the WG-Mamba network, for feature extraction and classification of fMRI data from patients with MDD and HCs. Specifically, considering the non-stationary characteristics of fMRI signals, the framework is designed as follows: First, a frequency block is implemented to decompose fMRI signals into multi-scale frequency features and filter noise through wavelet transform, while simultaneously extracting dynamic fluctuation features across different frequency bands of fMRI signals without compromising temporal localization accuracy. Second, a spatial block is designed to model cross-regional dependencies of each frequency component using a graph neural network with a brain region adjacency matrix, capturing spatial features of the brain network. Third, an inverse DWT is applied to reconstruct time series from the fused frequency components, forming a temporal block that modifies Mamba to extract temporal dynamic features. Finally, the architecture deeply integrates multi-scale spatio-temporal and frequency features to generate a comprehensive representation of brain activity.

2. Dataset and Preprocessing

2.1. Dataset Description

In this study, the fMRI data were sourced from the publicly available multicenter dataset REST-meta-MDD (http://rfmri.org/REST-meta-MDD). This dataset was jointly constructed by the research team led by Yan Chaogan from the Institute of Psychology, Chinese Academy of Sciences, in collaboration with 25 depression research teams from 17 institutions across China. All data collection and preprocessing followed a unified standardized protocol, and the multicenter data integration was completed after rigorous quality control [29]. The REST-meta-MDD dataset includes 1,300 patients with MDD and 1,128 demographically matched healthy controls, making it the largest publicly available resting-state functional imaging dataset for depression worldwide. It contains complete demographic information, clinical assessment data, and imaging data, and is open for global researchers to access and share. Detailed information on enrolment criteria and data formats can be found in the supplementary documentation on the official website. To minimize the interference of multicenter scanning parameter heterogeneity on model generalization performance, this study analyzed data from a single-center site designated as S20, which included 533 valid samples: 282 clinically diagnosed MDD

patients and 251 healthy controls. There were no statistically significant differences in age or gender distribution between the two groups.

2.2. Preprocessing

In this study, all resting-state fMRI data preprocessing was performed using the DPARSF toolkit, with the official access address being http://www.rfmri.org/DPARSF [30]. The preprocessing steps included: (1) Exclusion of the first 10 time points in each participant's rs-fMRI sequence acquisition phase to eliminate signal drift interference caused by incomplete stabilization of the magnetic field gradient during the scan initiation phase. (2) Application of time slice correction and six-parameter (rigid) linear transformation for image realignment. (3) Rigid co-registration of individual T1-weighted structural images with the corresponding average functional images of the participants using a 6-degree-of-freedom linear transformation, without resampling during registration to preserve spatial resolution information of the original images to the greatest extent. (4) Division of brain tissue into gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF) based on the registered T1-weighted structural images using an automatic segmentation algorithm. (5) Standardization of all images to the Montreal Neurological Institute (MNI) space. (6) Regression of head motion artifacts using the Friston 24-parameter head motion model, while simultaneously regressing white matter, CSF signals, and linear drift terms to further reduce interference from non-neural activity signals. (7) Filtering of all time series data using a time bandpass filter (0.01- 0.1 Hz). The frequency range was selected because it aligns with the typical slow oscillatory patterns of spontaneous brain activity that have been widely validated in fMRI studies .

2.3. Brain Functional Connectivity Construction

The construction of functional connectivity matrices is implemented in three sequential steps, including region of interest (ROI) parcellation, BOLD time series extraction, and functional connectivity calculation.

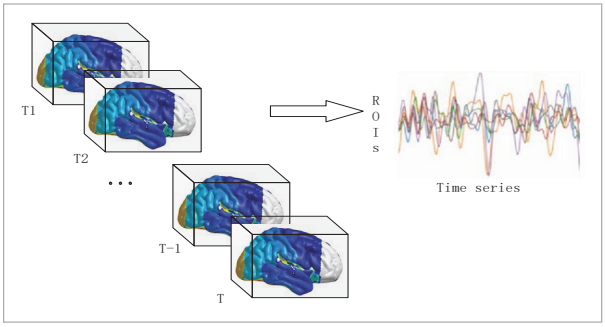
1 **ROI Parcellation**

Subsequent to the completion of rs-fMRI data preprocessing, the 3D whole-brain volume is parcellated into discrete regions of interest via standardized anatomical/functional atlas templates, a widely adopted approach for reducing data dimensionality in neuroimaging analysis. The atlas-based parcellation framework maps each individual voxel in the brain volume to a predefined ROI based on established anatomical boundaries or functional coherence patterns, following which the average BOLD signal of all voxels within each ROI is computed to generate region-level feature representations. This parcellation procedure converts high-dimensional voxel-wise features into macroscopic ROI-level features, which not only reduces computational complexity in subsequent analysis but also significantly improves the neurobiological interpretability of derived results. Widely used parcellation templates in resting-state fMRI studies include the Automated Anatomical Labeling (AAL) atlas, Craddock functional atlas, Harvard Oxford (HO) anatomical atlas, and POWER functional connectivity atlas [3]. In this work, the POWER functional atlas is employed for brain parcellation, which partitions the entire brain into 264 functionally homogeneous ROIs covering cortical, subcortical, and cerebellar regions.

2 **Time Series Extraction**

The fMRI data is a four-dimensional data, which is composed of three-dimensional brain image information and one-dimensional time information. The extraction of time series is to convert four-dimensional data into two-dimensional N×T matrix M. Where, N is the number of brain regions, T is the time, and Mij is the characteristic value of brain region i at time j. Thwerefore, fMRI data can be converted into time series data of each brain region. The schematic diagram of extracting time series is shown in Figure 1.

Figure 1
The extraction of time series.



3 **functional connectivity**

Quantification of inter-regional functional coupling in resting-state fMRI analysis relies on three mainstream correlation metrics: Pearson's correlation coefficient, partial correlation coefficient, and wavelet correlation coefficient, each with distinct applicability scenarios. For this study, Pearson's correlation coefficient is adopted for functional connectivity estimation, as it demonstrates superior stability for low-sampling-rate rs-fMRI data and is the most extensively validated metric for capturing linear temporal dependencies between blood oxygen level-dependent signal sequences of paired brain regions. The mathematical expression for calculating the pairwise Pearson's correlation coefficient is given as:

$$\rho_{ij}=corr(\mathrm{x_i},\mathrm{x_j})=\frac{\sum_{t=1}^{N}(x_{it}-\overline{x_i})(x_{jt}-\overline{x_j})}{\sqrt{\sum_{t=1}^{N}((x_{it}-\overline{x_i})^2)\sum_{t=1}^{N}(x_{jt}-\overline{x_j})^2}}. \quad (1)$$

In the above formula, x_i, x_j refer to the time series of the i-th and j-th ROI respectively, $\overline{x_i}$ and $\overline{x_j}$ represent the time series mean of the i-th and j-th ROI respectively. ρ_{ij} is the correlation coefficient between two ROIs, with a range of $-1\le\rho_{ij}\le 1$. If ρ_{ij} is positive, it indicates a positive correlation between the signal intensities of the two ROIs over time, suggesting functional synergy. If ρ_{ij} is negative, it indicates functional antagonism between the two ROIs, with a stronger correlation as the absolute value increases. When $\rho_{ij}=0$, it can be concluded that there is no correlation between the signal intensities of the two ROIs over time. After ρ_{ij} is computed, each value is converted to its absolute value to focus on connection strength rather than directional functional relationships. This yields a correlation coefficient matrix, where each row or column represents the sequence of Pearson correlation coefficients between a specific ROI and all other ROIs.

3. Methods

3.1. Preliminaries

As a representative state space model (SSM), Mamba excels at modeling long-range dependencies in sequential data, outperforming traditional recurrent neural networks and Transformers in both efficiency and accuracy [27]. Mamba2 recently proposed the concept of State Space Duality (SSD), simplifying matrix A into a scalar form [7]. Mamba2 is an advanced Structured State Space Model (SSM) specifically designed for efficient long-sequence modeling. Mamba2 operates using a structured state space representation. This special case of selective state space models (SSMs) is applicable to both linear and quadratic scenarios. Without loss of generality, the matrix transformation form of the selective state space model can be expressed as.

$$y(t)=\sum_{i=1}^{t}C_t^T A_{t:i+1}B_i x(i) \quad (2)$$

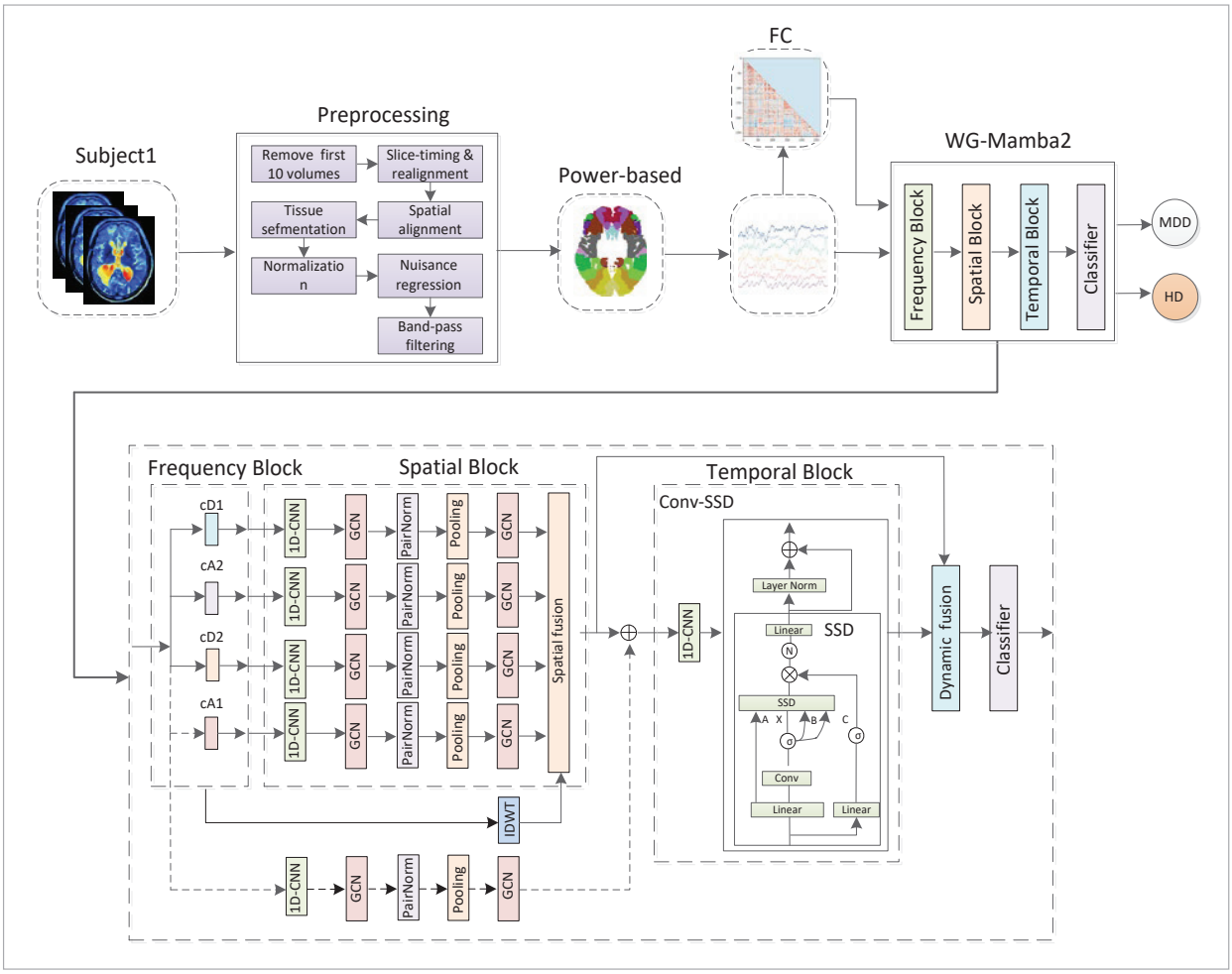
$$A_{t:i}=\prod_{i=2}^{t}A_i, \quad (3)$$

where $t\in\{1,\ldots T\}$ represents the current time step, and A_i, B_i, C_t^T are dynamically updated parameterized matrices.

3.2. Architecture Overview

The overall architecture of WG-Mamba is illustrated in Figure 2. WG-Mamba primarily consists of a frequency block layer (wavelet transform), a spatial layer (graph neural network), a temporal layer (Mamba), a dynamic fusion layer and a classifier layer. The fMRI data used in this study were sourced from the REST-meta-MDD dataset. For each subject, we utilized preprocessed fMRI data to extract blood oxygen level-dependent signal sequences corresponding to regions of interest based on a brain atlas template. Functional connectivity matrices were constructed through pairwise Pearson correlation analysis of time series, serving as adjacency matrices for brain region maps to characterize functional associations between brain regions. The fMRI time series were first processed through the frequency block layer, decomposed into different frequency components to achieve hierarchical capture of multi-scale brain activity features. Local frequency patterns were extracted from each scale feature via 1D-CNN, which was then input into the spatial layer along with the functional connectivity matrix to effectively model spatial dependencies between

Figure 2
The overall architecture.



brain regions and output refined spatial features of the brain network. Multi-scale frequency features incorporating spatial information were transformed and reconstructed into time-series data that retains spatial context. These sequences were input into the temporal layer to capture long-range temporal dependencies and extract temporal features, followed by fusion of spatial and temporal features. Ultimately, these aggregated multi-scale features were fed into the classifier to complete the prediction task.

3.3. Frequency Block

The Frequency Block serves as the core multi-scale feature extraction module in the WG-Mamba architecture. By decomposing raw BOLD signals into distinct frequency components, this module overcomes the limitation of traditional single-band analysis in distinguishing temporal scales of brain activity. A two-level discrete wavelet transform (DWT) is employed to decompose fMRI time series into frequency bands, aiming to capture hierarchical brain activity features ranging from short-term signal bursts and dynamic fluctuations to rapid neural oscillations. We utilize the Daubechies-2 (db2) wavelet, which is a tight-branch orthogonal wavelet with two vanishing moments. Its corresponding low-pass filter achieves smooth fitting of local quadratic trends, while the high-pass filter precisely extracts first-and second-order abrupt change features of the signal. These characteristics are highly compatible with the coexisting slow-wave activity, transient responses, and nonlinear fluctuations in BOLD time

series [34]. In MDD fMRI signal analysis, the db2 wavelet has been demonstrated to exhibit higher matching with the time-domain morphology of the hemodynamic response function (HRF) [18]. Given the original fMRI time series $X=[x_1,x_2,...,x_T]\in R^T$, the DWT decomposes the BOLD signal into i-level high-frequency detail coefficients cD_i and i-level low-frequency approximation coefficients cA_i at level L using orthogonal high-pass and low-pass filters f_{flow} and f_{high}, where X is cA_0. To enhance the discriminative power of frequency features, a one-dimensional convolution layer with kernel size 1 is applied to encode the high-frequency and low-frequency detail coefficients into hidden states X_{hfi} and X_{lfi}, respectively. Adaptive average pooling is then employed to unify features of different scales into the same dimension. The calculation formula for the Frequency Block is as follows:

$$cA_i[k]=\sum_{n=0}^{N} cA_{i-1}[n]:f_{flow}[2k-n] \tag{4}$$

$$cD_i[k]=\sum_{n=0}^{N} cA_{i-1}[n]:f_{high}[2k-n] \tag{5}$$

where N denotes the filter length, f_{low} and f_{high} are the low-pass and high-pass filter coefficients of the db2 wavelet, k represents the filter scale parameter for signal transformation, and n ranges from 1 to N/2.

We decompose the depth limit to L=2, where the first layer (Level-1) captures short-term signal bursts and dynamic fluctuations, while the second layer (Level-2) isolates low-frequency oscillatory components that organize predictable functional networks. This is because each additional layer in wavelet transformation reduces temporal resolution by half while doubling frequency resolution. Excessive layers would lead to signal over-smoothing, causing transient physiological activity to be lost and failing to capture dynamic connectivity changes in the brain. Typically, only a few key frequency bands are of interest, making a two-layer decomposition sufficient.

A two-level discrete wavelet transformation using the db2 wavelet is applied to decompose the original fMRI time series into three frequency subbands, including the Level-1 detail component (cD_1) that captures high-frequency signal bursts and rapid neural fluctuations, the Level-2 detail component (cD_2) that captures medium-frequency oscillatory patterns, and the Level-2 approximation component (cA_2) that isolates low-frequency baseline brain activity.

Inverse discrete wavelet transform (IDWT) is then applied to suppress high-frequency noise in fMRI time series.

$$cD_i[k]=\begin{cases} cD_i[k], \ if \ |cD_i[k]|\geq\lambda \\ 0, \ otherwise \end{cases} \tag{6}$$

$$X_{iF}=IDWT(cA_2,cD_1,cD_2), \tag{7}$$

where λ denotes the threshold value for adaptive thresholding of high-frequency.

$X_{iF}\in R^{1\times T}$ denotes the frequency-domain feature vector of the i-th ROI, reconstructed via inverse discrete wavelet transform from the raw frequency coefficients. Concatenating across all 264 ROIs yields, $X_F=[X_{1F},X_{2F},...,X_{264F}]\in R^{264\times T}$, and $X_F\in R^{264\times 1\times T}$ is reshaped.

3.4. Spatial Block

The brain network carries topological information, and graph neural networks exhibit significant advantages in processing such node-edge connection data. This design employs graph neural networks to model spatial dependencies between brain regions at the network level. The spatial block primarily consists of ROI graph convolutional layers, normalization layers (PairNorm), pooling layers (SAGPooling), and spatial fusion. First, each subject is transformed into an undirected graph $G_F=(V,E)$, where V represents the fMRI frequency coefficient feature data of ROI extracted by the Frequency Block layer, and E represents the edge weights reflecting the connection relationships between nodes, which are calculated using the Pearson correlation coefficient ρ_{ij} mentioned earlier. The resulting adjacency matrix A is as follows:

$$A_{ij}=\begin{cases} |\rho_{ij}| \ , \ if \ v_i \ and \ v_j \ are \ adjacent, \\ 1 \ , \ if \ i=j, \\ 0 \ , \ otherwise \end{cases} \tag{8}$$

The edge weights between nodes represent their Pearson correlation coefficient values. The adjacency matrix $A_{ij\in R^{n\times n}}$, where n=264.

For each subband, a 1D convolutional layer with kernel size 1 expands the feature dimension from 1 to C (C=16) channels. After ReLU activation, dimension permutation is performed to transpose the channel and temporal axes, enabling AdaptiveAvgPool1D to operate along the channel dimension. The pooled features are then permuted back to yield a single-channel temporal representation. The proposed method ensures that the temporal dimension remains unchanged for each subband, while simultaneously achieving feature compression and extraction of critical information.

$$X_{hfi}=AdaptiveAvgPool1D(ReLU(Conv1d(c\ \mathrm{D}_i))) \quad (9)$$

$$X_{lfi}=AdaptiveAvgPool1D(ReLU(Conv1d(cA_i))) \quad (10)$$

$$X_G=[X_{lfi},X_{hfi}],, \quad (11)$$

where X_G is formed by concatenating the coefficients of each subband along the temporal dimension.The calculation formulas for the graph convolutional layer and the graph normalization layer are as follows:

$$X_G^{l+1}=\sigma(GCN(X_G,A_G))=\sigma(\dot{D}_G^{-1/2}\cdot\dot{A}_G\cdot\dot{D}_G^{1/2}\cdot X_G^l\cdot W_G^l) \quad (12)$$

$$X_G^{l+1}{}'=Norm(X_G^{l+1}), \quad (13)$$

where G is the input graph, W_G^l denotes the learnable weight matrix, $\dot{A}_G=A_G+I$ is the adjacency matrix with added self-loops, D_G is the degree matrix, $\dot{D}_G^{-1/2}\cdot\dot{A}_G\cdot\dot{D}_G^{1/2}$ represents the normalized Laplacian matrix, σ is the activation function, and X_G^l and X_G^{l+1} respectively denote the feature representations learned at the l-th and (l+1)-th layers.

Research indicates that in certain scenarios, the classification performance of Graph Neural Networks may not improve with increasing layers, but could instead experience a significant decline. This phenomenon is attributed to the over-smoothing problem inherent in GNNs. The node normalization layer in our model prevents all node embeddings from converging toward similar directions by standardizing node data. This approach effectively mitigates overfitting while addressing the over-smoothing issue, thereby enhancing the robustness of GNNs.

First, perform centralization processing on the features of each node to obtain:

$$\bar{h}_i^{l+1}=h_i^{l+1}-\frac{1}{n}\sum_{i=1}^{n}h_i^{l+1}. \quad (14)$$

Assuming the scaling factor is s, then scale $\bar{h}_i^{l+1}$ to obtain:

$$\tilde{h}_i^{l+1}=s\cdot\left(\frac{\bar{h}_i^{l+1}}{\sqrt{\frac{1}{n}\sum\|\bar{h}_i^{l+1}\|_2}}\right). \quad (15)$$

Replace the term $\Sigma\|h_i^{-l+1}\|_2$ in the above formula with $n\Sigma\|h_i^{-l+1}\|_2$ to obtain PairNorm in the scaling individually (SI) mode. Therefore, the above formula can be rewritten as:

$$\tilde{h}_i^{l+1}=s\cdot\left(\frac{\bar{h}_i^{l+1}}{\|\bar{h}_i^{l+1}\|_2}\right). \quad (16)$$

A self-attention guided graph pooling module is proposed to extract discriminative node features and filter functionally relevant key nodes. Core nodes associated with abnormal connectivity patterns are retained via hierarchical node pooling layers, while the graph structure size is reduced and network parameters are minimized simultaneously to mitigate overfitting risks. High-level semantic information of the brain functional network is effectively encoded in the node representations generated from the coarsened graph structure.

For the node feature matrix X_F^l of the current layer, the self-attention score matrix S_F^l is computed via graph convolution operation, based on which a top-k node selection strategy is implemented to preserve a subset of informative nodes from the brain functional graph, with k denoting the predefined number of nodes for constructing the coarsened graph.

During the node selection process, nodes are ranked according to the indices idx corresponding to the top k values in S_G^l. The tanh activation function is then applied to these top k elements in S_G^l to derive the normalized self-attention scores s of the retained nodes. The pooled node feature matrix X_G^{l+1} is obtained by performing element-wise multiplication between the indexed features of the retained nodes and s. Correspondingly, rows and columns matching the indices idx are extracted from the original adjacency matrix A_G^{l+1} to generate the adjacency matrix $A_{idxl,idxl}^l$ for the coarsened graph.

$$S_G^l = GCN(X_G^l, A_G) \tag{17}$$

$$s = tanh(S_G^l) \tag{18}$$

$$idx_G^l = TOP_k(s) \tag{19}$$

$$X_G^{l+1} = X_{idxl}^l \cdot s \tag{20}$$

$$A_G^{l+1} = A_{idxl,idxl}^l, \tag{21}$$

where the $X_G^{l+1} \in R^{k\times 1\times T}$. The time series of the corresponding K ROIs X_F^k are extracted from the X_F time series based on the K ROI brain region indices.

The Spatial Fusion Layer is designed to effectively integrate temporal dynamics and spatial domain information.

$$X_{rec} = \text{cat}\,(X_F^k, X_G^{l+1}). \tag{22}$$

The final output signal of the spatial block is X_{rec}.

3.4. Temporal Block

The temporal block serves as the core temporal feature extraction module in the WG-Mamba architecture. It employs a Selective State Space Model (S4) integrated with Gate-Locked Linear Unit (GLU) to capture long-range dependencies from fMRI time series. This module takes the spatially enhanced temporal data output by the spatial block as input, efficiently modeling the dynamic evolution of brain activity to extract temporal features across the full sampling duration, capturing dynamic functional connectivity patterns from short-term signal fluctuations to sustained network states within the fMRI scanning session.

This provides critical temporal dimension information for final classification tasks. The underlying layer of the temporal block is the Conv-SSD layer, which captures temporal dependencies. The core module Mamba2 in this layer incorporates 1D convolutional layers (1D-CNN), LayerNorm, and residual connections to enhance stability and performance during training. First, a 1D convolutional layer with kernel size 1 encodes the initial fMRI time series X_{rec} into a hidden state X_M. Next, within the Mamba architecture, this module scans the entire X_M sequence to capture style features. Finally, the Conv-SSD layer outputs result through residual connections with the spatial block output. The computational formulas for each layer of the temporal block are as follows:

$$X_M = Conv1d(X_{rec}) \tag{23}$$

$$X_R = SSD(X_M, A, B, C, \Delta) + X_M. \tag{24}$$

3.5. Dynamic Fusion Layer

The dynamic weight mechanism operates on the outputs of the spatial block and temporal block. Rather than treating these features equally, we compute adaptive fusion weights through a lightweight attention mechanism:

$$W_s = \sigma(W \cdot GAP(X_{rec}) + b_s) \tag{25}$$

$$W_t = \sigma(W \cdot GAP(X_R) + b_t), \tag{26}$$

where GAP(.) denotes global average pooling. σ(.) is the sigmoid function. The normalized weights are obtained via softmax:

$$\alpha_s = \frac{exp(W_s)}{exp(W_s) + exp(W_t)} \tag{27}$$

$$\alpha_t = \frac{exp(W_t)}{exp(W_s) + exp(W_t)}. \tag{28}$$

The fused feature is then computed as:

$$Y=\alpha_s \cdot X_{rec}+\alpha_t \cdot X_R. \quad (29)$$

3.6. Classification Layer

The feature vector z serves as the input to the fully connected layer, which is the residual output concatenated from the temporal block and spatial block. The fully connected layer translates features to a format usable for the cross entropy loss function to evaluate classification performance. The loss of function of WG-Mamba model is defined as in equation:

$$Loss=Min(\sum_j CrossEntropy(y_j,z_j)+H(z) \quad (30)$$

$$z_j=SoftMax(W \cdot f(x_j)+b) \quad (31)$$

$$H(z)=\sum_i z_i log\ z_i). \quad (32)$$

Among them, (x_j,y_j) is a labeled sample in the fmri data, and the function $f(x_j)$ is the feature vector extracted by the WG-Mamba model. The parameter z_j is the prediction, and the model parameter W and b can be trained on the small samples data set. H(z) is a regularization.

The WG-Mamba model and its classification accuracy and error are generated by multiple rounds of computer iterative calculation.

4. Experiment and Result Discussion

4.1. Training Setup

In this study, a stratified 5-fold cross-validation strategy was adopted to rigorously assess the generalization performance of the proposed algorithm.

The experiment used the following hardware: Intel i9-14900K CPU, NVIDIA RTX 4090 GPU with 24GB VRAM, running on a 64-bit Linux operating system.

The training phase parameters were set as follows: an initial learning rate of 0.0001, dropout regularization rate to 0.5, weight decay coefficient 0.000001, graph pooling ratio 0.3. The optimizer was Adam, batch size is set to 4. The training process is scheduled for a maximum of 100 epochs, with early stopping enabled to prevent overfitting if the validation loss plateaus for 10 consecutive epochs.

4.2. Statistical Metrics

To quantitatively evaluate the classification performance of the proposed model, accuracy (ACC), precision, recall, and F1-score were selected as core evaluation metrics, which are commonly used in medical imaging diagnostic tasks to comprehensively assess model effectiveness from different dimensions.

For the binary classification scenario of major depressive disorder diagnosis, where MDD subjects are defined as the positive category and healthy participants as the negative group, four core entries in the confusion matrix are specified as follows: True Positive (TP) corresponds to the count of MDD participant samples accurately assigned to the MDD group, False Positive (FP) describes healthy control samples incorrectly labeled as having MDD, False Negative (FN) refers to MDD patient samples wrongly categorized into the healthy control cohort, and True Negative (TN) represents the count of healthy participant samples correctly identified as non-MDD cases.

The calculation formulas of each evaluation metric are defined as follows:

$$Accuracy=\frac{TP+TN}{TP+TN+FP+FN} \quad (33)$$

$$Precision=\frac{TP}{TP+FP} \quad (34)$$

$$Recall=\frac{TP}{TP+FN} \quad (35)$$

$$F1=2\times\frac{Precision\times Recall}{Precision+Recall}. \quad (36)$$

Accuracy represents the ratio of correctly predicted instances within the full evaluation sample set. Precision metric measures the ratio of accurately identified positive cases to all samples predicted to belong to the positive category. Recall is the proportion of positive samples in the dataset that are correctly predicted.

The F1-score is computed as the harmonic average of precision and recall metrics. In general, higher precision and recall values are preferable, yet the two metrics often show a trade-off in practical scenarios. Therefore, the F1 score was introduced as a comprehensive metric to balance precision and recall.

4.3. Classification of fMRI Data

The training process of the WG-Mamba architecture was monitored across all 5 cross-validation folds, with early stopping enabled to prevent overfitting. The model demonstrated stable convergence, with classification accuracy and cross-entropy loss stabilizing within a 0.015 interval after the 62nd epoch, and no significant overfitting was observed. The temporal dynamics of average training accuracy and average loss function values across the five cross-validation folds are visualized in Figures 3-4, respectively.

On the held-out test partition from the REST-meta-MDD S20 scanning site, our presented architecture reached a mean peak classification accuracy of 83.72% across all cross-validation folds, demonstrating its robust discriminative performance in identifying aberrant brain functional patterns correlated with major depressive disorder.

To rigorously verify the performance advantages and generalizability of our presented WG-Mamba architecture, we constructed a multi-group benchmark evaluation workflow, with all tests uniformly executed on the public REST-meta-MDD dataset to rule out performance deviations caused by inconsistent data sources.A unified preprocessing pipeline, evaluation protocol, and experimental environment are adopted across all baseline methods and the proposed framework:

The conventional machine learning baseline SVM was implemented using the Scikit-learn (version 1.3.0) toolbox, with hyperparameters tuned through stratified 5-fold cross-validation conducted on the training partition, covering kernel function selection (linear/RBF) and regularization parameter adjustment ($C \in \{0.01, 0.1, 1, 10\}$).

Graph neural network baselines, including GCN and GAT, were constructed using the PyTorch Geometric (version 2.3.1) library, with their network structures and hyperparameters strictly consistent with the settings in their original publications to ensure baseline performance reaches the published level (Table 1).

4.4. Ablation Experiment

We conduct ablation studies on the held-out test partition to validate the performance gains brought by each component of the proposed framework and associated technical designs. The WG-Mamba is mainly composed of the frequency block(wavelet-based) and the temporal block(conv-ssd-based), spatial block (GCN-based). For the ablation trials, model performance was evaluated using five key metrics: accuracy (ACC), sensitivity (SEN), specificity (SPE), area under the receiver operating characteristic curve (AUROC), and area under the precision-recall curve (AUPRC).

Figure 3
The model training accuracy.

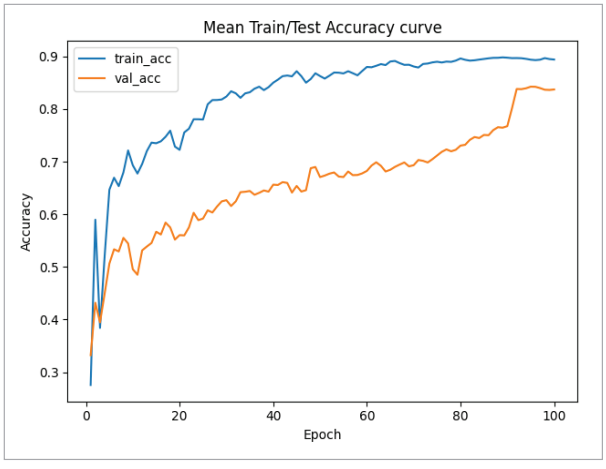


Figure 4
The Loss function value.

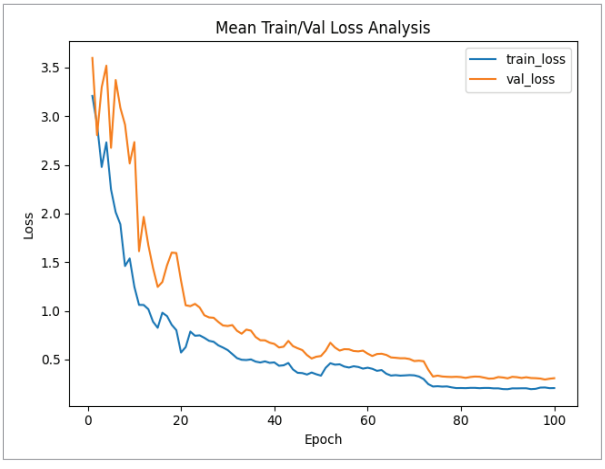


Table 1
Comparison results with baselines.

Method	Samples	Accuracy	Precision	Recall	F1
SVM	533	0.581±0.042	0.589±0.038	0.695±0.051	0.638±0.045
GCN	533	0.6921±0.035	0.5981±0.041	0.6995±0.048	0.645±0.039
GAT	533	0.705±0.032	0.602±0.036	0.779±0.041	0.678±0.036
MSC-Mamba [15]	1642	0.6991±0.028	0.7062±0.031	0.6796±0.034	-
BrainGNN [16]	615	0.709±0.030	0.750±0.027	0.667±0.039	0.703±0.033
MAF-GNN [2]	533	0.7580±0.026	0.7129±0.029	0.8889±0.022	0.7912±0.025
MFGCN [17]	533	0.776±0.023	0.819±0.021	0.797±0.024	0.793±0.022
MSSTAN [14]	533	0.8001±0.019	0.7325±0.025	0.7843±0.023	0.7578±0.021
This Paper	533	0.8372±0.012	0.8345±0.011	0.8768±0.010	0.8492±0.012

Table 2
Ablation experiments.

Ablation settings	ACC	SEN	SPE	AUROC	AUPRC
This Paper	0.8372±0.012	0.8650±0.01	0.802±0.013	0.885±0.009	0.852±0.011
w/o Frequency block	0.765±0.021	0.7928±0.018	0.7298±0.022	0.8128±0.016	0.7798±0.019
w/o Temporal block	0.782±0.019	0.8098±0.017	0.7468±0.02	0.8298±0.015	0.7968±0.018
w/o Spatial block	0.798±0.017	0.8258±0.015	0.7628±0.019	0.8458±0.013	0.8128±0.016
w/o Dynamic Layer	0.8111±0.014	0.8422±0.012	0.7731±0.016	0.8631±0.011	0.8315±0.013

Predictive performance of all candidate baseline approaches is evaluated via stratified 5-fold cross-validation across the entire dataset partition, and quantitative comparison outcomes are summarized in Table 2.

As presented in Table 2, the ablation experimental results demonstrate that the integration of the WG-Mamba module yields the most substantial improvement in classification performance, while the incorporation of the wavelet-based frequency block not only accelerates the training convergence speed by reducing redundant high-frequency noise components in raw BOLD signals, but also further enhances the model's discriminative capability for subtle abnormal brain activity patterns. The final optimized configuration of the proposed WG-Mamba framework consists of the full combination of the GCN-based spatial block, conv-SSD temporal block, and wavelet frequency block, with dynamic weight fusion applied across multi-scale features. The proposed framework achieved optimal classification performance on the held-out test set, with an accuracy of 83.72%, a sensitivity of 86.50%, a specificity of 80.2%, an AUROC of 88.5%, and an AUPRC of 85.2%, outperforming most existing end-to-end deep learning methods for resting-state fMRI-based major depressive disorder identification.

All ablation experiment results consistently demonstrate that each core module of the proposed WG-Mamba framework contributes significantly to the performance improvement, and the multi-scale spatiotemporal-spectral fusion design is reasonable and effective.

4.5. Analysis of Relevant Brain Regions

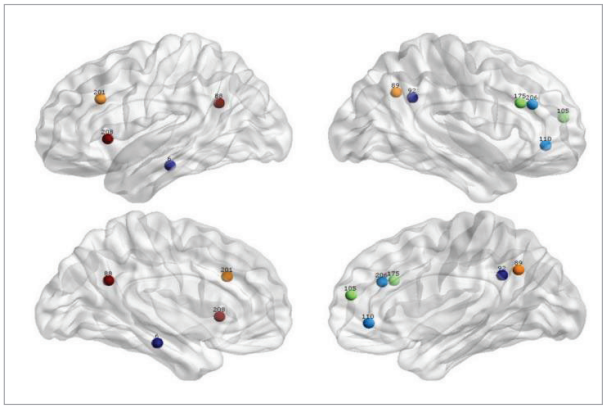
This section primarily analyzes brain regions associated with depression. The parameter weights of each

Region of Interest in the brain regions output by the self-attention pooling layer in the WG-Mamba module are crucial for generating discriminative features.

First, fMRI data covering all regions of interest were fed into the WG-Mamba model for forward propagation. Second, spatial features of the fMRI data were extracted at the spatial level using a graph neural network. Hierarchical graph dimensionality reduction was achieved through self-attention TopK pooling operations, which compress the graph structure while selecting K key nodes and retaining information from these critical nodes. These K nodes also play a significant role in depression classification. In this study, the value of K was set to 10. Finally, after multiple iterations of model computation, the top 10 most important brain regions after TopK pooling were statistically analyzed, and the 10 regions with the highest frequency of occurrence were selected as biomarkers. The top 10 ROIs with the highest average node pooling scores were numbered as 20 8,201,175,206,92,88,89,105,110, and 6. These brain regions correspond to the left ventral lateral prefrontal cortex (vlPFC), left dorsal lateral prefrontal cortex (DLPFC), right dorsal lateral prefrontal cortex (DLPFC), ventral anterior cingulate cortex/medial geniculate cortex (sgACC), medial prefrontal cortex/frontal pole (mPFC/FP), posterior cingulate cortex (PCC), precuneus, Superior Frontal Gyrus, Medial Prefrontal Cortex, and left parahippocampal gyrus. The distribution of these ROIs in the brain is shown in Figure 5.

The study results indicate that the highest-ranked prefrontal cortex, anterior cingulate cortex, and limbic system regions exhibit significant correlations with MDD. The prefrontal cortex, primarily composed of the dorsolateral prefrontal cortex (DLPFC) and ventral lateral prefrontal cortex (vlPFC), is responsible for higher-order cognitive control, emotional regulation, and reward processing. Functional abnormalities in this region directly mediate anhedonia, decision-making impairments, and cognitive reappraisal deficits in MDD [1]. The anterior cingulate cortex serves as a critical hub for emotional and cognitive integration, with hyperactivation of the superior geniculate nucleus (sgACC) being a core pathological feature of negative affective disturbance in MDD and a key target for current antidepressant therapies. The limbic system, including structures such as the hippocampus and amygdala, is primarily involved in emotional memory processing and stress response regulation [12]. Extensive research has confirmed that hippocampal volume atrophy and amygdaloid hyperactivation serve as stable biomarkers for MDD.

Figure 5
Illustration of public relevant brain regions.



5. Conclusions and Future Work

This study proposes WG-Mamba, a novel multiscale spatiotemporal-spectral fusion framework for distinguishing MDD patients from HCs using resting-state fMRI. The network integrates wavelet-based frequency decomposition, brain network spatial modeling, and temporal dynamics capture through a dynamic weight fusion mechanism, enabling end-to-end identification of neurobiological biomarkers for depression classification. Specifically, the Frequency module employs db2 wavelet transform to decompose fMRI signals into multiscale frequency components, mitigating information loss inherent in single-band analysis while preserving temporal localization. The Spatial module utilizes graph convolution with brain region adjacency matrices to model cross-regional functional dependencies, where the self-attention guided TopK pooling mechanism automatically identifies core ROI biomarkers with demonstrated neuropathological validity. The Temporal module leverages selective state space modeling to capture long-range temporal dependencies with linear computational complexity. Experimental evaluation on the REST-meta-MDD dataset (S20 site) demonstrates that

WG-Mamba achieves 83.72% classification accuracy under stratified 5-fold cross-validation, outperforming existing end-to-end deep learning approaches. By analyzing high-weight nodes output from the graph pooling module, we pinpointed core brain regions that serve as reliable neuroimaging biomarkers for distinguishing major depressive disorder cases from demographically matched healthy individuals. The clinical interpretability was validated by neuropathological studies. Furthermore, the WG-Mamba model can be extended to other fields of neuroimaging research.

The model enables early MDD screening via fMRI in routine checkups and pinpoints dysfunctional connectivity in core regions, offering interpretable biomarkers to guide personalized treatment.

To homogenize cross-site data quality, the current work exclusively utilizes neuroimaging data collected from the S20 scanning site within the REST-meta-MDD multi-center cohort. However, practical clinical deployment scenarios require proposed frameworks to exhibit better cross-population generalization and reliable predictive stability. In the future, we will address this limitation by developing technologies specifically designed for multi-center fMRI data, including domain-adaptive algorithms and federated learning techniques. These approaches aim to mitigate disparities in data distribution across centers while preserving clinically meaningful biological signals. Concurrently, we plan to integrate multimodal neuroimaging modalities such as sMRI, DTI, and PET into the WG-Mamba architecture. This cross-modal fusion technology will enable a comprehensive elucidation of the pathophysiological mechanisms underlying MDD through the integration of multimodal data. This integrated approach is expected to reveal the multidimensional neurobiological characteristics of MDD, facilitating more precise subtype classification and personalized treatment strategies.

Acknowledgement

This work was supported by the National Natural Science Foundation of China (41827807 and 61271351), the Science and Technology Innovation Plan of Shanghai Science and Technology Commission (22DZ1209500).

References

1. Bertocci, M.A., Afriyie-Agyemang, Y., Rozovsky, R., Iyengar, S., Stiffler, R., Aslam, H.A., Bebko, G., Phillips, M. L. Altered Patterns of Central Executive, Default Mode and Salience Network Activity and Connectivity Are Associated with Current and Future Depression Risk in Two Independent Young Adult Samples. Molecular Psychiatry, 2023, 28(3), 1046-1056. https://doi.org/10.1038/s41380-022-01899-8
2. Chen, Z., Wu, J., Pu, L., Li, S., Zhang, Y., Yao, D., Yan, H. MAF-GNN: Graph Neural Network-Based Multi-Atlas Brain Functional Information Fusion for Major Depressive Disorder Diagnosis With rs-fMRI. Information Processing & Management, 2026, 63(3), 104524. https://doi.org/10.1016/j.ipm.2025.104524
3. Craddock, R.C., James, G.A., Holtzheimer, P.E., Hu, X.P., Mayberg, H.S. A Whole Brain fMRI Atlas Generated via Spatially Constrained Spectral Clustering. Frontiers in Systems Neuroscience, 2012, 33(8), 1994-1928. https://doi.org/10.1002/hbm.21333
4. Dai, P., Huang, K., Shi, Y., Xiong, T., Zhou, X., Liao, S., Huang, Z., Yi, X., Grecucci, A., Chen, B. T. Effective Connectivity Between the Cerebellum and Fronto-Temporal Regions Correctly Classifies Major Depressive Disorder: fMRI Study Using a Multi-Site Dataset. Journal of Affective Disorders, 2025, 390, 119783. https://doi.org/10.1016/j.jad.2025.119783
5. Dai, P., Zhou, Y., Shi, Y., Lu, D., Chen, Z., Zou, B., Liu, K., Liao, S. Classification of MDD Using a Transformer Classifier with Large-Scale Multisite Resting-State fMRI Data. Human Brain Mapping, 2024, 45(1), e26542. https://doi.org/10.1002/hbm.26542
6. Dai, P., Lu, D., Shi, Y., Zhou, Y., Xiong, T., Zhou, X., Chen, Z., Zou, B., Tang, H., Huang, Z., Liao, S. Classification of Recurrent Major Depressive Disorder Using a New Time Series Feature Extraction Method Through Multisite rs-fMRI Data. Journal of Affective Disorders, 2023, 339, 511-519. https://doi.org/10.1016/j.jad.2023.07.077
7. Dao, T., Gu, A. Transformers Are SSMs: Generalized Models and Efficient Algorithms Through Structured State Space Duality. arXiv Preprint, 2024. arXiv:2405.21060. https://doi.org/10.48550/arXiv.2405.21060

8. Fang, K., Niu, L., Wen, B., Liu, L., Tian, Y., Yang, H., Hou, Y., Han, S., Sun, X., Zhang, W. Individualized Resting-State Functional Connectivity Abnormalities Unveil Two Major Depressive Disorder Subtypes with Contrasting Abnormal Patterns of Abnormality. Translational Psychiatry, 2025, 15(1), 45. https://doi.org/10.1038/s41398-025-03268-9

9. Gu, A., Dao, T. Mamba: Linear-Time Sequence Modeling with Selective State Spaces. arXiv Preprint, 2023. arXiv:2312.00752. https://doi.org/10.48550/arXiv.2312.00752

10. Gu, A., Goel, K., Ré, C. Efficiently Modeling Long Sequences with Structured State Spaces. arXiv Preprint, 2021. arXiv:2111.00396. https://doi.org/10.48550/arXiv.2111.00396

11. Hasan, M., Yasmin, F., Yu, X., Karim, A. Advances in Brain Imaging Technologies: A Comprehensive Overview. Brain Networks in Neuroscience: Personalization Unveiled via Artificial Intelligence. River Publishers, 2025, 11-40. https://www.taylorfrancis.com/chapters/edit/10.1201/9788770047371-2

12. Javelle, F., Suhrkamp, J., Wählen, L., Donath, L., Bloch, W. The Role of Physical Exercise on Hippocampal Volume in Depressive Symptoms: A Systematic Review and Multi-Level Meta-Analysis. NeuroImage, 2025, 320, 121386. https://doi.org/10.1016/j.neuroimage.2025.121386

13. Ju, W., Fang, Z., Gu, Y., Liu, Z., Long, Q., Qiao, Z., Qin, Y., Shen, J., Sun, F., Xiao, Z., Yang, J. A Comprehensive Survey on Deep Graph Representation Learning. Neural Networks, 2024, 173, 106207. https://doi.org/10.1016/j.neunet.2024.106207

14. Kong, Y., Zhang, X., Wang, W., Zhou, Y., Li, Y., Yuan, Y. Multi-Scale Spatial-Temporal Attention Networks for Functional Connectome Classification. IEEE Transactions on Medical Imaging, 2024, 44(1), 475-488. https://doi.org/10.1109/TMI.2024.3448214

15. Li, S., Bo, Y., Chen, Y., Cao, J., Bi, B., Ma, T., Ye, C. Multiscale Contextual Mamba: Advancing Psychiatric Disorder Detection Across Multisite Functional Magnetic Resonance Imaging Datasets via State Space Modeling. Health Data Science, 2025, 5, 0224. https://doi.org/10.34133/hds.0224

16. 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, 2021, 74, 102233. https://doi.org/10.1016/j.media.2021.102233

17. Liu, S., Gui, R. Fusing Multi-Scale fMRI Features Using a Brain-Inspired Multi-Channel Graph Neural Network for Major Depressive Disorder Diagnosis. Biomedical Signal Processing and Control, 2024, 90, 105837. https://doi.org/10.1016/j.bspc.2023.105837

18. Luo, F. F., Wang, J. B., Yuan, L.X., Zhou, Z. W., Xu, H., Ma, S. H., Zang, Y. F., Zhang, M. Higher Sensitivity and Reproducibility of Wavelet-Based Amplitude of Resting-State fMRI. Frontiers in Neuroscience, 2020, 14, 224. https://doi.org/10.3389/fnins.2020.00224

19. Ma, J., Li, F., Wang, B. U-Mamba: Enhancing Long-Range Dependency for Biomedical Image Segmentation. arXiv Preprint, 2024. arXiv:2401.04722. https://doi.org/10.48550/arXiv.2401.04722

20. Mousavian, M., Chen, J., Traylor, Z., Greening, S. Depression Detection From sMRI and rs-fMRI Images Using Machine Learning. Journal of Intelligent Information Systems, 2021, 57(2), 395-418. https://doi.org/10.1007/s10844-021-00653-w

21. Pilmeyer, J., Huijbers, W., Lamerichs, R., Jansen, J. F., Breeuwer, M., Zinger, S. Functional MRI in Major Depressive Disorder: A Review of Findings, Limitations, and Future Prospects. Journal of Neuroimaging, 2022, 32(4), 582-595. https://doi.org/10.1111/jon.13011

22. Pitsik, E. N., Maximenko, V. A., Kurkin, S.A., Sergeev, A. P., Stoyanov, D., Paunova, R., Kandilarova, S., Simeonova, D., Hramov, A. E. The Topology of fMRI-Based Networks Defines the Performance of a Graph Neural Network for the Classification of Patients with Major Depressive Disorder. Chaos, Solitons & Fractals, 2023, 167, 113041. https://doi.org/10.1016/j.chaos.2022.113041

23. Su, S., Ning, Y., Guo, Z., Yang, W., Zhu, M., Zhou, Q., He, X. Classification of Major Depressive Disorder Using Graph Attention Mechanism with Multi-Site rs-fMRI Data. Neuroinformatics, 2025, 23(2), 34. https://doi.org/10.1007/s12021-025-09731-8

24. Wei, Y., Gao, J. Classification of Major Depressive Disorder Based on Functional and Structural MRI. In Fourth International Conference on Computer Vision and Pattern Analysis (ICCPA 2024). SPIE, 2024, 13256, 129-135. https://doi.org/10.1117/12.3037818

25. World Health Organization. Depression (Depressive Disorder). Fact Sheet, 2023. https://www.who.int/news-room/fact-sheets/detail/depression

26. Xiaoai, G., Yujing, X., Lin, L., Lin, T. An Overview of Disease Prediction Based on Graph Convolutional Neural

Network. In Proceedings of the 6th International Conference on Intelligent Information Processing, 2021, 27-32. https://doi.org/10.1145/3480571.3480575

27. Xing, Z., Ye, Y., Yang, Y., Liu, G., Zhu, L. SegMamba: Long-Range Sequential Modeling Mamba for 3D Medical Image Segmentation. arXiv Preprint, 2024. arXiv:2401.13560. https://doi.org/10.1007/978-3-031-72111-3_54
28. Yamashita, A., Sakai, Y., Yamada, T., Yahata, N., Kunimatsu, A., Okada, N., Itahashi, T., Hashimoto, R., Mizuta, H., Ichikawa, N., and Takamura, M. Generalizable Brain Network Markers of Major Depressive Disorder Across Multiple Imaging Sites. PLoS Biology, 2020, 18(12), e3000966. https://doi.org/10.1371/journal.pbio.3000966
29. Yan, C. G., Chen, X., Li, L., Castellanos, F. X., Bai, T. J., Bo, Q. J., Cao, J., Chen, G. M., Chen, N. X., Chen, W., Cheng, C. Reduced Default Mode Network Functional Connectivity in Patients with Recurrent Major Depressive Disorder. Proceedings of the National Academy of Sciences, 2019, 116(18), 9078-9083. https://doi.org/10.1073/pnas.1900390116
30. Yan, C., Zang, Y. DPARSF: A MATLAB Toolbox for Pipeline Data Analysis of Resting-State fMRI. Frontiers in Systems Neuroscience, 2010, 4, 1377. https://doi.org/10.3389/fnsys.2010.00013
31. Yan, W., Qu, G., Hu, W., Abrol, A., Cai, B., Qiao, C., Plis, S. M., Wang, Y. P., Sui, J., Calhoun, V. D. Deep Learning in Neuroimaging: Promises and Challenges. IEEE Signal Processing Magazine, 2022, 39(2), 87-98. https://doi.org/10.1109/MSP.2021.3128348
32. Yin, S. Q., Li, Y. H. Advancing the Diagnosis of Major Depressive Disorder: Integrating Neuroimaging and Machine Learning. World Journal of Psychiatry, 2025, 15(3), 103321. https://doi.org/10.5498/wjp.v15.i3.103321
33. Yue, Y., Li, Z. MEDMAMBA: Vision Mamba for Medical Image Classification. arXiv Preprint, 2024. arXiv:2403.03849. https://doi.org/10.48550/arXiv.2403.03849
34. Zaynidinov, H., Juraev, U., Tishlikov, S., Modullayev, J. Application of Daubechies Wavelets in Digital Processing of Biomedical Signals and Images. International Conference on Intelligent Human Computer Interaction, 2023, 194-206. https://doi.org/10.1007/978-3-031-53827-8_19
35. Zhang, B., Liu, S., Liu, X., Chen, S., Ke, Y., Qi, S., Wei, X., Ming, D. Discriminating Subclinical Depression from Major Depression Using Multi-Scale Brain Functional Features: A Radiomics Analysis. Journal of Affective Disorders, 2022, 297, 542-552. https://doi.org/10.1016/j.jad.2021.10.122