Received: 29 December 2023 | Revised: 4 February 2024 | Accepted: 19 February 2024

DOI: 10.1002/ima.23053

RESEARCH ARTICLE

WILEY

Constructing brain functional networks with adaptive manifold regularization for early mild cognitive impairment

Xidong Fu[1] | Shengchang Shan[2] | Chun Liu[1] | Yu Lu[2] | Zhuqing Jiao[1,2]

[1]School of Computer Science and Artificial Intelligence, Changzhou University, Changzhou, China

[2]School of Microelectronics and Control Engineering, Changzhou University, Changzhou, China

Correspondence
Zhuqing Jiao, School of Computer Science and Artificial Intelligence, Changzhou University, Changzhou 213164, China.
Email: jzq@cczu.edu.cn

Funding information
Qing Lan Project of Jiangsu Province; Jiangsu Provincial Key Research and Development Program, Grant/Award Number: BE2021636; National Natural Science Foundation of China, Grant/Award Number: 51877013

Abstract

Brain functional network (BFN) has emerged as a practical path to explore biomarkers for early mild cognitive impairment (eMCI). Currently, most of BFNs only considered the topology structure between two brain regions and ignored the high-order information among multiple brain regions. We proposed an adaptive manifold regularization method to construct a new BFN. Firstly, a traditional hypergraph was constructed through a low-order BFN. Then, an adaptive hypergraph was obtained by updating the traditional hypergraph weight and structure through adaptive hypergraph learning. An adaptive hypergraph manifold regularization term was constructed by the Laplacian matrix of the adaptive hypergraph. Finally, the low-order BFN was optimized through the adaptive hypergraph manifold regularization and L_1 sparse regularization. The experimental results confirmed that the proposed method outperformed other state-of-the-art methods in classification performance and stability. This study revealed the causes of changes in topological properties and provided a reference for the clinical diagnosis of eMCI.

KEYWORDS

adaptive hypergraph learning, brain functional network, early mild cognitive impairment, functional magnetic resonance imaging, manifold regularization term

1 | INTRODUCTION

Alzheimer's disease (AD) is a neurodegenerative disorder characterized by insidious onset and progressive development, representing the most common cause of dementia.[1] Studies have shown that the disease affects patients' memory and communication abilities, potentially leading to fatality.[2] As of now, effective treatments for AD have not been discovered, and the current study focuses on identifying early intervention measures to slow down the progression of the disease.[3] Mild cognitive impairment (MCI) serves as a transitional stage between AD and normal aging, with a conversion rate of 10%–15% developing into AD every year.[4] The brains of early mild cognitive impairment (eMCI) patients have very subtle changes compared with normal people.[5] Timely detection and treatment of eMCI are crucial to prevent its progression to MCI and AD. However, the pathological mechanisms of eMCI remain incompletely understood, making neuroimaging studies on patients essential for meaningful therapeutic insights.

Currently, image acquisition and analysis methods are in the developing stage, allowing researchers to easily and noninvasively study brain function and structural activity.[6] Among them, abnormal changes are detected by diffusion kurtosis imaging (DKI) in white matter, fibers, and neurons.[7] Information is acquired via arterial spin labeling (ASL) in cerebral blood flow.[8] Changes in

blood oxygen concentration are measured through functional magnetic resonance imaging (fMRI).[9] In particular, fMRI has higher temporal resolution and has become a mainstream analytical method for detecting brain activity. For instance, Zhang et al.[9] classified MCI using improved feature extraction methods for fMRI data. Sheng et al.[10] predicted MCI patients' cognitive scores by combining fMRI features. Overall, fMRI provides great convenience for the early diagnosis of neurological diseases.[11]

Numerous studies have indicated that brain functional networks (BFNs) reflected the connectivity and temporal correlations among brain regions, serving as a powerful tool for detecting neurodegenerative diseases.[12,13] Researchers have proposed various methods for constructing BFNs. For instance, Biswal et al.[14] constructed a BFN by calculating the connectivity strength between brain regions. Lee et al.[15] adopted the sparse regularization method to construct a BFN with sparse representation. Notably, BFNs typically involved complex topological structures, and their construction process required better preservation of prior information about the brain.[16] Prior information such as modularity was retained by introducing different regularization terms.[17] For example, Li et al.[18] constructed a BFN with graph manifold regularization, leveraging the premise that similar brain regions often exhibit similar structures. Ji et al.[19] adopted the hypergraph as a substitute for the original graph, introducing hypergraph manifold regularization term to construct a BFN. Xi et al.[20] constructed a BFN with dynamic hypergraph manifold regularization to improve classification accuracy.

Most of the above methods only considered the relationship information between two brain regions, ignored the high-order relationships among multiple brain regions in BFNs, and could not retain more topological structure information. Moreover, the hypergraphs they constructed only focused on attributes from the data itself and were unable to adapt to data changes and mine more potential relationship information. The graph regularization learning and the adaptive hypergraph learning were adopted to solve these problems. On the one hand, the graph regularization learning can preserve topological structure information between brain regions and improve generalization ability. The information contained the connections between nodes, which better represent the abnormal connections and changes of BFNs, improving performance and diagnostic accuracy. The generalization ability was improved to prevent overfitting, allowing the model to better generalize to previously unseen data. On the other hand, the adaptive hypergraph can adaptively adjust the connection structure of the hypergraph according to the characteristics of the data to better capture the high-order relationships between data, compared with traditional hypergraphs. It updates the weight of the hypergraph and enhances the hypergraph structure through hypergraph ranking optimization and adaptive weight learning.

We captured high-order relationship information among multiple brain regions by integrating graph regularization learning and adaptive hypergraph learning. As mentioned above, an adaptive manifold regularization (AMR) method was proposed to construct new BFNs. Firstly, a low-order BFN was constructed by calculating the Pearson correlation coefficients between time series. Then, a traditional hypergraph was constructed via k-nearest neighbors (KNN) algorithm. Subsequently, the traditional hypergraph structure was jointly updated through hypergraph ranking optimization and adaptive weight learning to obtain the adaptive hypergraph. The Laplacian matrix of adaptive hypergraph was taken as the manifold regularization term. Finally, the low-order BFN was optimized through the adaptive hypergraph manifold regularization and L_1 sparse regularization.

The main contributions and advantages of this study are as follows: (A) AMR integrated graph regularization learning and adaptive hypergraph learning in BFNs, exploring high-order information among multiple brain regions, better emphasizing critical information, and reducing noise and redundancy. (B) We identified the discriminative brain regions in eMCI patients, analyzing the changes and damage in these regions, and providing a foundation for the prevention and treatment of eMCI. (C) We investigated the changes in the topological properties and analyzed the factors leading to these changes in eMCI, providing an important reference for a deeper understanding of the disease. The subsequent sections are structured in the following manner. Section 2 introduces the overview of AMR and the selected experimental data. Section 3 presents the experimental results, discriminative brain regions in eMCI, and topological properties changes. Finally, we concluded by summarizing its content, reflecting on the method's limitations, and outlining future research directions.

2 | MATERIALS AND METHODS

2.1 | Framework

Figure 1 illustrates the visual overview of the proposed method. It mainly consists of the following steps: (A) Preprocessing raw resting-state fMRI data; (B) Extracting time series from fMRI; (C) Constructing a low-order BFN according to Pearson correlation coefficients; (D) Constructing traditional hypergraph through

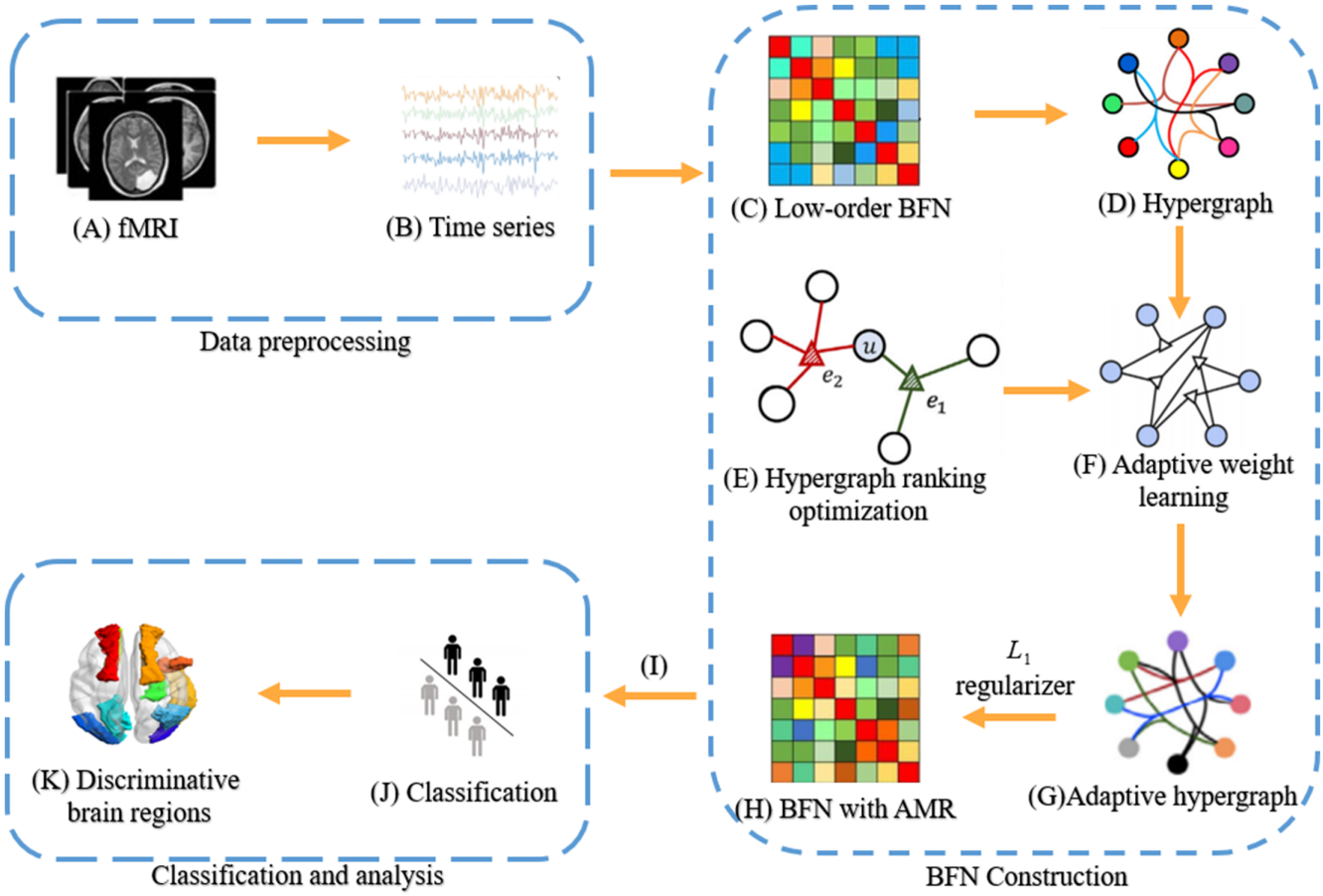

FIGURE 1 Flow of the proposed method.

the low-order BFN; (E) Improving the original hypergraph weight through hypergraph ranking optimization; (F) Refining hypergraph weight further to obtain the final adaptive hypergraph through adaptive weight learning; (G) Generating the Laplacian matrix from adaptive hypergraph and creating an adaptive hypergraph manifold regularization term; (H) Optimizing the low-order BFN to construct a BFN with AMR through the adaptive hypergraph manifold regularization and L_1 sparse regularization; (I) Extracting features and selecting discriminative features; (J) Partitioning discriminative features into training and testing sets, followed by training adopting support vector machine (SVM); (K) Visualizing discriminative brain regions based on the selected features and analyzing discriminative regions in eMCI.

2.2 | Data and preprocessing

We employed a total of 214 subjects' resting-state fMRI data from the publicly available Alzheimer's disease neuroimaging initiative (ADNI) database (https://adni.loni.usc.edu/). Fifty-one eMCI patients were taken as eMCI group, including 28 males and 23 females, aged 74.09 ± 6.60 years old. Seventy-two normal subjects were taken as the normal group, including 35 males and 38 females, aged 77.88 ± 6.47 years old. We obtained approval from the ADNI Research Ethics Committee (https://adni.loni.usc.edu/study-design/ongoing-investigations/) and ensured that all participants provided written informed consent.

The subjects were scanned by the 3.0 T Philips Achieva scanner with the following specific parameters: imaging matrix = 64 × 64, slice thickness = 3.3 mm, flip angle = 80°, repetition time (TR) = 3000 ms, and echo time (TE) = 30 ms. The raw fMRI data were preprocessed by means of statistical parameter mapping (SPM12) and the data processing assistant for resting states (DPARSF).[21] The specific steps are as follows: (A) Time point removal: The first 10 time points were discarded due to higher image noise in these initial points. (B) Image normalization: The fMRI images were converted into the Montreal Neurological Institute (MNI) standard brain space. It set the bounding box to [−90, −126, −72; 90, 90, 108], voxel size to [3 3 3], ensuring consistency between images and reducing noise interference. (C) Spatial smoothing: Gaussian kernel smoothing

was applied to the normalized fMRI images to enhance signal-to-noise ratio. (D) Bandpass filtering: A frequency range of 0.01–0.08 Hz was set to remove low-frequency drift and high-frequency noise. (E) Time series extraction: Brain regions were delineated, and time series were extracted for each region.[22]

2.3 | Hypergraph generation

We assume a time series $\boldsymbol{X}=[\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_B]\in\boldsymbol{R}^{D\times B}$, where $\boldsymbol{X}$ denotes the time series, D represents the length of the time series, and B denotes the number of brain regions. Then, we adopted the sliding window method[23] to divide the time series into different sub-time series. The time series matrix of k-th window was defined as $\boldsymbol{X}^{(k)}=[\boldsymbol{x}_1^k,\boldsymbol{x}_2^k,\cdots,\boldsymbol{x}_B^k]\in R^{L_a\times B}$, where $\boldsymbol{x}_B^k$ denotes the time series of the k-th window in the B-th brain region, L_a is the length of sliding window. A low-order BFN is constructed by calculating the Pearson correlation coefficients. The correlation coefficient matrix of the k-th window can be expressed as $\boldsymbol{P}^{(k)}\approx\boldsymbol{X}^{(k)^T}\boldsymbol{X}^{(k)}$, which can be converted into the optimized form:

$$\min_{\boldsymbol{P}^{(k)}}\|\boldsymbol{P}^{(k)}-\boldsymbol{X}^{(k)^T}\boldsymbol{X}^{(k)}\|_F^2 \tag{1}$$

where k is the window number of the time series, $\boldsymbol{P}^{(k)}$ is the coefficient matrix of the kth low-order BFN, and $\boldsymbol{X}^{(k)}$ is the time series matrix of k-th window.

The conventional graph tends to depict only pairwise relationships between nodes, while the relationships among brain regions are more complex than pairwise relationships. The hypergraph can better preserve information about the relationships among multiple brain regions.[24–26] In the hypergraph, $\boldsymbol{G}(\boldsymbol{V},\boldsymbol{E},\boldsymbol{w})$ denotes a hypergraph, where $\boldsymbol{V}$ denotes the set of nodes, $\boldsymbol{E}$ is the set of hyperedges connecting multiple brain regions, and $\boldsymbol{w}$ is the set of weight of the hyperedges. $\boldsymbol{H}$ is the association matrix, defined as:

$$\boldsymbol{H}(v,e)=\begin{cases}1, & v\in e\\ 0, & v\notin e\end{cases} \tag{2}$$

where v is a hypergraph node and e is a hyperedge. Node degree and hyperedge degree are computed for each node and hyperedge, and the formulas are shown below:

$$\begin{cases}\delta(e)=\sum_{v\in\boldsymbol{V}}\boldsymbol{H}(v,e)\\ \delta(v)=\sum_{e\in\boldsymbol{E}}w(e)\boldsymbol{H}(v,e)\end{cases} \tag{3}$$

where $w(e)$ denotes the weight of the hyperedge e, $\delta(e)$ is the node degree, and $\delta(v)$ is the hyperedge degree.

2.4 | Adaptive hypergraph learning

High-order relationships are often depicted via the time series $\boldsymbol{X}$, often resulting in fixed weight. It makes traditional hypergraph struggle to adapt to data changes, thereby failing to capture information precisely. The hypergraph is made more flexible to adapt to high-order relationships between data by adaptive hypergraph learning. The structure and weight of the hypergraph are adapted to changes in the data, which improves the ability to model complex relationships and allows for more accurate and adaptive relationships between brain regions.

The Laplacian matrix of a traditional hypergraph $\boldsymbol{L}$ is obtained, which is described as:

$$\boldsymbol{L}=\boldsymbol{I}-\boldsymbol{A} \tag{4}$$

where $\boldsymbol{A}$ is the adjacency matrix of hypergraph, $\boldsymbol{I}$ is the identity matrix.

It is necessary to assign a ranking vector $\boldsymbol{f}$ and a query vector $\boldsymbol{y}$ such that $\Omega(\boldsymbol{f})=\frac{1}{2}\boldsymbol{f}^T\boldsymbol{L}\boldsymbol{f}$. Then, $\boldsymbol{f}$ is optimized to find strong connection points by ℓ_2-norm.[27] The new ranking vector $\boldsymbol{f}^*$ is calculated accordingly:

$$\boldsymbol{f}^*=\arg\min_{\boldsymbol{f}}\Omega(\boldsymbol{f})+\theta\|\boldsymbol{f}-\boldsymbol{y}\|_2^2 \tag{5}$$

where θ is the regularization parameter, $\boldsymbol{f}$ is the ranking vector, y is the query vector. The hypergraph edge weight and structure optimization are performed after optimizing the hypergraph ranking so that the hypergraph can adaptively change rows according to the update of weight and structure.

Let $\boldsymbol{f}$ and $\boldsymbol{H}$ be fixed, define $\boldsymbol{Z}=\delta(v)^{-\frac{1}{2}}\boldsymbol{H}$ and $\boldsymbol{\rho}=\boldsymbol{Z}^T\boldsymbol{f}$. Subsequently, the weight $\boldsymbol{w}$ is optimized by adding an ℓ_2 norm regularizer,[28] which aims at optimizing and modulating the structure of the hypergraph in an efficient, robust, and automated way:

$$w^*=\arg\min_{w}\left\{\rho^T\mathrm{diag}(w)\delta(e)^{-1}\rho+\eta\|w\|_2^2\right\} \tag{6}$$

where $\boldsymbol{w}^*$ denotes the optimized weight, η is the adaptive weight regularization term parameter, $\delta(e)$ is the node degree, and theℓ_2-norm regularizer adds smoothness.

The obtained $\boldsymbol{f}^*$ and $\boldsymbol{w}^*$ will continue to optimize the structure of the hypergraph itself, after updating $\boldsymbol{f}$ and $\boldsymbol{w}$.

Finally, the adaptive hypergraph $\boldsymbol{H}^*$ is obtained, which has the objective function:

$$\boldsymbol{H}^* = \arg\min \sum \boldsymbol{f}^{*T}\left(\boldsymbol{I} - \delta(v)^{-\frac{1}{2}}\boldsymbol{H}\boldsymbol{w}^*\delta(e)^{-1}\boldsymbol{H}^T\delta(v)^{-\frac{1}{2}}\right)\boldsymbol{f}^* \tag{7}$$

The adaptive hypergraph learning is optimized by the alternating iterative algorithm to update $\boldsymbol{f}$, $\boldsymbol{w}$, and $\boldsymbol{H}$. They are optimized by solving:

$$(\boldsymbol{f}^*, \boldsymbol{w}^*, \boldsymbol{H}^*) = \arg\min_{\boldsymbol{f},\boldsymbol{w},\boldsymbol{H}} T(\boldsymbol{f}, \boldsymbol{w}, \boldsymbol{H}) \tag{8}$$

a. Update $\boldsymbol{f}^*$. Let $\boldsymbol{f}$ and $\boldsymbol{H}$ be fixed, $\boldsymbol{J} = \boldsymbol{I} - \frac{\boldsymbol{A}}{1+\theta}\boldsymbol{f}^*$ is obtained through Equation (9):

$$\boldsymbol{f}^* = \arg\min_{\boldsymbol{f}} T(\boldsymbol{f}; \boldsymbol{w}, \boldsymbol{H}) = \omega \boldsymbol{J}^{-1}\boldsymbol{y} \tag{9}$$

where $\omega = \frac{\theta}{1+\theta}$.

b. Update $\boldsymbol{w}^*$. The optimal $\boldsymbol{w}^*$ is obtained through the least mean square method.[29] It is sequentially revised upward along the negative gradient of the optimization function. $\boldsymbol{w}^*$ is computed by

$$\boldsymbol{w}^{*[t]} = \boldsymbol{w}^{*[t-1]} - \alpha_{LMS}\frac{\partial \varphi(\boldsymbol{w}^*)}{\partial \boldsymbol{w}^*}\bigg|_{\boldsymbol{w}^{*[t-1]}} \tag{10}$$

where α_{LMS} represents the learning rate, $\varphi(\boldsymbol{w}^*)$ is the optimization result of the Lagrange formula in Equation (6), and $\boldsymbol{w}$ is updated until convergence to obtain the optimal $\boldsymbol{w}^*$.

2.5 | Brain functional network with adaptive manifold regularization

The Laplacian matrix of graph reflects its intrinsic geometric structure.[30] The high-order relationships among nodes can be captured by the Laplacian matrix of the adaptive hypergraph instead of the original graph's Laplacian matrix. The Laplacian matrix of the adaptive hypergraph is calculated accordingly:

$$\boldsymbol{L}^{\boldsymbol{H}^*} = \boldsymbol{I} - \delta^*(v)^{-\frac{1}{2}}\boldsymbol{H}^*\boldsymbol{w}^*\delta^*(e)^{-1}\boldsymbol{H}^T\delta^*(v)^{-\frac{1}{2}} \tag{11}$$

where $\delta^*(v)$ and $\delta^*(e)$ are the hyperedge and node degree of the adaptive hypergraph, respectively. Inspired by Ji et al.,[19] we introduce $\boldsymbol{L}^{\boldsymbol{H}^*}$ in the form of manifold regularization, along with L_1 sparse regularization, into Equation (1) to obtain a BFN with AMR. The objective function is shown below:

$$\min_{\boldsymbol{P}^{(k)}} \| \boldsymbol{P}^{(k)} - \boldsymbol{X}^{(k)^T}\boldsymbol{X}^{(k)} \|_F^2 + \lambda \mathrm{tr}\left(\boldsymbol{P}^{(k)^T}\boldsymbol{L}^{\boldsymbol{H}^*}\boldsymbol{P}^{(k)}\right) + \gamma \| \boldsymbol{P}^{(k)} \|_1 \tag{12}$$

where $\boldsymbol{X}^{(k)}$ represents the sub-time series, $\boldsymbol{P}^{(k)}$ is the coefficient matrix of BFNs, $\boldsymbol{L}^{\boldsymbol{H}^*}$ denotes the Laplacian matrix of the adaptive hypergraph. λ and γ represent the hypergraph manifold regularization term parameter and the L_1 sparse regularization term parameter, respectively.

Equation (12) is non-differentiable, so we adopt gradient descent and the nearest-neighbor operator[31] method to solve the objective function. The gradient of the fitting term is calculated in the objective function, and then $\boldsymbol{P}^{(k)}$ is updated by the gradient descent method. The optimization formula is as follows:

$$\boldsymbol{P}^{(k)^n} = \boldsymbol{P}^{(k)^{n-1}} - \alpha_n\left[\nabla_{\boldsymbol{P}^{(k)}} f\left(\boldsymbol{X}^{(k)}, \boldsymbol{P}^{(k)^{n-1}}\right) + \lambda \boldsymbol{L}^{\boldsymbol{H}^*}\boldsymbol{P}^{(k)}\right] \tag{13}$$

where α_n represents the step size of the gradient descent, and n represents the number of updates. Calculate the nearest-neighbor operator, apply a threshold to Equation (13), and further update $\boldsymbol{P}^{(k)}$ until convergence to obtain the optimal $\boldsymbol{P}^{(k)}$. Table 1 shows the specific updating process.

3 | EXPERIMENT AND ANALYSIS

We extracted weighted clustering coefficients from BFNs as features and performed feature selection through t-test inspired by Jiao et al.[32] Subsequently, the classification purpose was accomplished through SVM. The weighted clustering coefficient features comprehensively reflected the interaction patterns between brain regions. The classification outcomes were ultimately achieved,[32] based on the distinct node clustering patterns between eMCI patients and normal subjects. SVM demonstrated excellent generalization capabilities, making them well-suited for medical data with small sample sizes.[33]

The classification performance was assessed by the tenfold cross-validation.[34] The participant's data were randomly divided into 10 equally sized subsets in tenfold cross-validation, each with a similar data distribution, for 10 independent experiments. One subset served as the

TABLE 1 Optimization process of brain functional network.

Input: The time series $\boldsymbol{X}$; The number of neighbors c; The hypergraph manifold regularization term parameter λ; The L_1 sparse regularization term parameter γ; The adaptive weight regularization term parameter η.

Output: The coefficient matrix of BFNs $\boldsymbol{P}^{(k)}$.

The KNN algorithm constructs the traditional hypergraph;
Adopt Equation (2) to obtain the association matrix $\boldsymbol{H}$;
Calculate the node degree$\delta(e)$ and hyperedge degree$\delta(v)$ of the traditional hypergraph by Equation (3);
Calculate the adjacency matrix $\boldsymbol{A}$;
Initialize the weight matrix of the hypergraph $\boldsymbol{w}$;
Define the rank vector $\boldsymbol{f}$.

While not converges
Fix other variables;
Update $\boldsymbol{f}$ by Equation (9);
Then Fix other variables;
Define $\boldsymbol{J} = \boldsymbol{I} - \frac{\boldsymbol{A}}{1+\theta}$;
Update $\boldsymbol{w}$ by Equation (10);
Then Fix other variables;
Update $\boldsymbol{H}$ by Equation (7) and Equation (8);
End while
Get adaptive hypergraph $\boldsymbol{H}^*$;
Construct adaptive hypergraph manifold regularization term by Equation (10);
Get $\boldsymbol{P}^{(k)}$ by Equation (12);
While not converges
Update $\boldsymbol{P}^{(k)}$ by equation (13).
End while

test set in each experiment, while the other nine subsets were combined to form the training set. Performance metrics were averaged over the rounds to obtain the final classification performance of the method. Accuracy (ACC), area under the curve (AUC), specificity (SPE), and sensitivity (SEN) were selected as evaluation metrics for classification performance.[35]

3.1 | Parameter selection

This section involves the selection of parameters related to AMR. The parameters included the number of neighbors c for building the hypergraph and the regularization term parameter η for constructing the adaptive hypergraph in AMR. Additionally, it included two regularization term parameters λ and γ in the method model. Parameter selection were conducted using tenfold cross-validation for comparing the classification performance.

The parameters λ and γ were fixed, and the optimal values for constructing the adaptive hypergraphs c and η were determined. The range of values for c and η were[1,3,5,7,9,11] and [0.0005, 0.0001, 0.005, 0.001, 0.05,0.01], respectively. Figure 2 (A) illustrates the classification results for different combinations of neighbors and weight parameters.

According to Figure 2A, it can be observed that ACC reached its optimum when $c = 7$ and $\eta = 0.001$. c had a

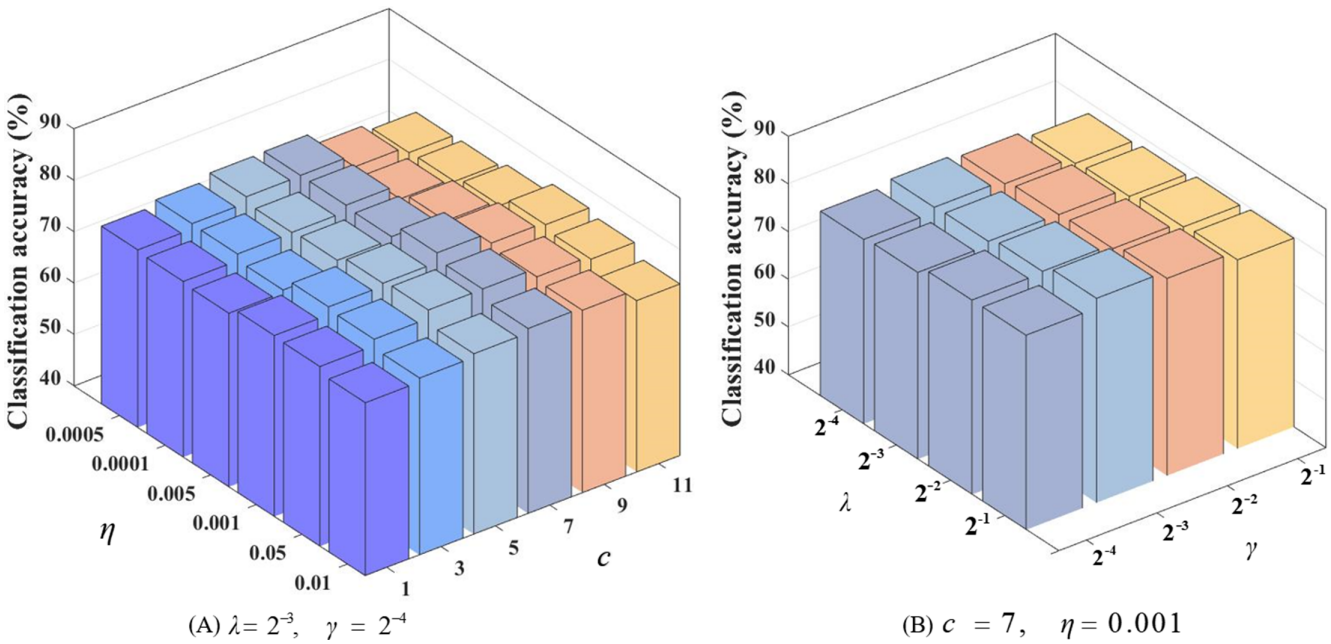

FIGURE 2 Classification accuracy for different parameter combinations.

more significant impact on classification accuracy than η under different parameter combinations. c (η) showed an increasing-then-decreasing trend in classification accuracy as its value increased when η was held constant. It may be because selecting c vertices with the closest Euclidean distance generates more hyperedges with increased node relationships, better reflecting the structural characteristics of BFNs. The nodes in the hyperedges may contain different categories when there are too many vertices, describing the overall structure of BFNs rather than local features. It may not capture complex relationships among brain regions.

Fixed $c = 7$ and $\eta = 0.001$. Varied the values of λ and γ to find the optimal values for the regularization parameters, ultimately obtaining the best classification accuracy for the method. The range of values for λ and γ was $[2^{-4}, 2^{-3}, 2^{-2}, 2^{-1}]$. Figure 2B shows the classification accuracy results for different regularization parameters.

It can be observed that there was not much fluctuation in classification accuracy under different regularization parameter combinations from the figure, indicating a certain level of stability in AMR. The classification accuracy increased with the increase of λ while a fixed value of γ. The classification accuracy showed an increasing trend followed by a decrease when fixing λ. The reason for this trend may be that a larger λ indicated a greater weight for the adaptive hypergraph manifold regularization term. It made the model more focused on the high-order relationships among multiple brain regions, that is, a preference for preserving information about high-order relationships among multiple brain regions. It may eliminate some useful information when the weight of γ was too large, leading to a decrease in classification accuracy. In summary, c was set to 7, η was set to 0.001, and the two regularization parameters λ and γ were set to 2 and 3, respectively.

3.2 | Ablation study

We conducted ablation study to discuss the impact of AMR on the classification performance. AMR contained two modules: adaptive hypergraph learning and hypergraph learning. We added modules in sequence to obtain the classification performance of methods under different modules. The classification results are presented in Table 2, with the optimal classification results highlighted in bold.

A clear improvement in classification performance was found when introducing both adaptive hypergraph learning and hypergraph learning method compared to methods without these components. The introduction of hypergraph learning allowed the method to better capture high-order information among multiple brain regions, positively contributing to enhanced classification performance. The adaptive hypergraph exhibited greater flexibility compared to traditional hypergraph. The reason may be that their weights can automatically adjust based on data characteristics, adapting more effectively to complex relationship structures. Additionally, the adaptive hypergraph can emphasize the critical information in the network more by learning the weights and structures, reducing noise and redundancy, and avoiding information loss.

TABLE 2 Classification performance of ablation study.

Hypergraph learning	Adaptive hypergraph learning	ACC (%)
×	×	74.84 ± 2.76
○	×	79.03 ± 1.88
○	○	**82.87 ± 1.88**

The standard deviation of the experimental results also reflect the size of the performance fluctuation at each fold. A smaller standard deviation signified stronger stability of the method. The introduction of hypergraph learning enhanced the stability of the method, and simultaneously introducing both hypergraph learning and adaptive learning maintained high stability. Consequently, AMR demonstrated high stability and exhibited superior classification performance.

3.3 | Contrast experiment

We evaluated the effectiveness of AMR by comparing it with five state-of-the-art methods for constructing BFNs. The baseline methods include Pearson correlation (PC),[14] sparse representation (SR),[15] graph regularization (MR),[18] sparse and hypergraph manifold regularization (SHMR),[19] and dynamic hypergraph manifold regularization (DHMR).[20] All methods adopted the same feature extraction and selection methods, used SVM for classification, and compared the final optimal results. The detailed classification performance is presented in Table 3, with the optimal results highlighted in bold.

It is evident that AMR exhibited the best classification performance, with ACC, AUC, SPE, and SEN reaching 82.87 ± 1.88%, 86.57 ± 1.06%, 84.82 ± 2.04%, and 83.35 ± 2.90%, respectively. MR demonstrated better classification performance compared to the first two methods. It indicated that focusing on the topological organizational

Method	ACC (%)	AUC (%)	SPE (%)	SEN (%)
PC	67.85 ± 3.08	70.79 ± 3.31	78.38 ± 3.70	57.32 ± 5.80
SR	64.47 ± 3.89	68.66 ± 2.40	72.44 ± 3.04	61.22 ± 4.28
MR	74.84 ± 2.76	78.73 ± 1.69	76.03 ± 1.99	71.9 ± 3.24
SHMR	79.03 ± 1.88	82.61 ± 1.23	79.53 ± 1.76	77.89 ± 3.28
DHMR	80.31 ± 2.95	84.47 ± 1.78	81.91 ± 2.52	79.58 ± 3.48
AMR	**82.87 ± 1.88**	**86.57 ± 1.06**	**84.82 ± 2.04**	**83.35 ± 2.90**

TABLE 3 Classification performance of different methods.

features of BFNs by identifying similar connectivity patterns between brain regions is more effective in distinguishing between eMCI patients and normal subjects. SHMR performed better than MR, indicating that the introduction of hypergraph instead of original graph focuses on the topological organization features of multiple brain regions. It helped to provide richer prior information. Meanwhile, the introduction of the sparse regularization term parameter can eliminate the noise and redundant values. DHMR introduced dynamic hypergraph learning based on SHMR, so that the hypergraph structure was modified according to the dynamic BFNs. It reduced the loss of the time dimension information, and achieved a classification accuracy of 80.31 ± 2.95%.

Nevertheless, AMR outperformed DHMR in terms of classification performance. It suggested that solely modifying the hypergraph structure to adapt to dynamic BFNs may not effectively capture high-order relational information among multiple brain regions. The adaptive hypergraph allowed the hypergraph to adapt even further to changes in the data based on the update of the ranking weights, with stronger generalization capabilities. AMR identified more superior relationships among brain regions and improved the distinction of relationships through hypergraph ranking optimization and adaptive weight learning. It retained the more important high-order information of brain regions in BFNs. In addition, the contrast experiments employed tenfold cross-validation, and scrutiny of the standard deviations in Table 3 assessed the fluctuation in performance across each fold. AMR performed the best in terms of standard deviation of classification performance, indicating that AMR had stronger stability.

3.4 | The most discriminative brain regions

We sought the most discriminative brain regions in eMCI classification after determining the optimal parameter combination through tenfold cross-validation. The discriminative brain regions were selected by features with p-value less than 0.05 and top 15 frequency of occurrence. The selected top 15 discriminative brain regions were visualized via the BrainNet Viewer toolbox, providing a clearer representation of their spatial locations and relationships. Figure 3 shows the visualization of discriminative brain regions, where L and R for left and right brain, respectively.

It is found from Figure 3 that the selected discriminative brain regions are mainly located in the frontal lobe and temporal lobe. On the one hand, the frontal lobe plays a crucial role in the brain, being associated with important functions such as memory, judgment, and abstract thinking.[36] The selected frontal lobe regions include the left orbital middle frontal gyrus (ORBmid.L) and the left medial superior frontal gyrus (SFGmed.L), each having functions related to decision-making and self-awareness.[36] On the other hand, the temporal lobe is primarily responsible for functions like memory, language ability, and emotional control, closely associated with various cognitive and perceptual functions.[37] The discriminative brain regions, such as the left olfactory cortex (OLF.L), right hippocampus (HIP.R), right superior temporal gyrus (HES.R), and left superior temporal gyrus (STG.L), play crucial roles in olfactory processing, memory, auditory processing, and language comprehension, respectively.[38] Literature[39] suggests that damage to HIP.R directly leads to a decline in short-term memory, which may also be affected in eMCI patients. Additionally, the selected right postcentral gyrus (PoCG.R), right peri-calcarine cortex (CAL.R), left thalamus (THA.L), and right thalamus (THA.R) are associated with sensory reception and higher cognitive functions.[40]

In particular, the selected discriminative brain regions include THA.L and THA.R. They serve as secondary brain structures, primarily functioning as a relay station for information transmission and integration.[41] The results mentioned above are generally consistent with previous pathological studies and reports on neuroimaging biomarkers.[42] Most of the selected discriminative brain regions are core members of the default mode network (DMN).[43] It includes the left posterior cingulate

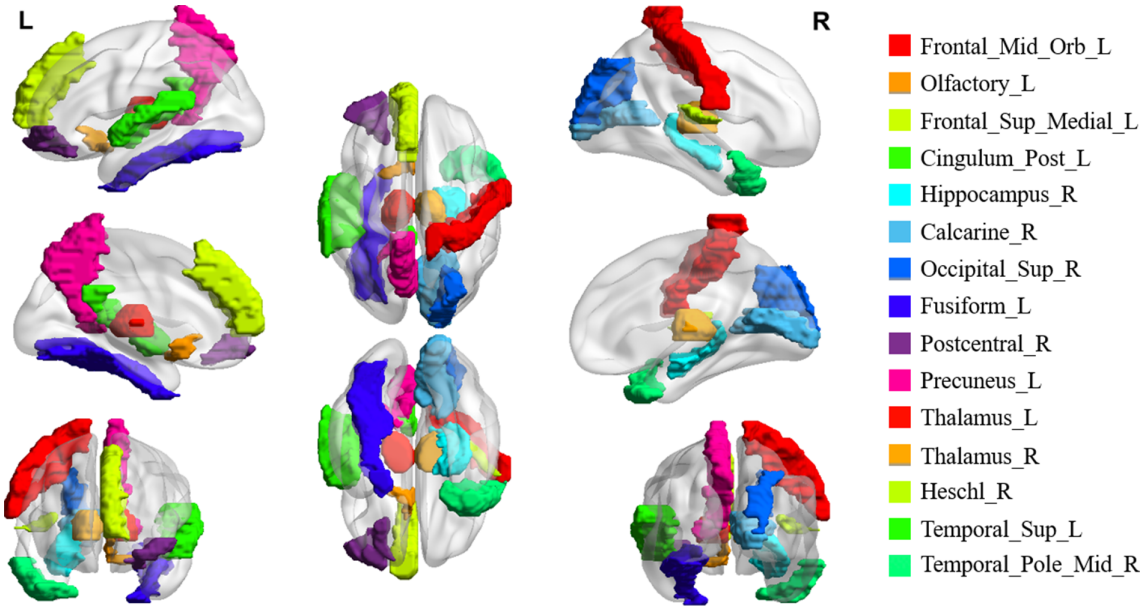

FIGURE 3 Discriminative brain regions.

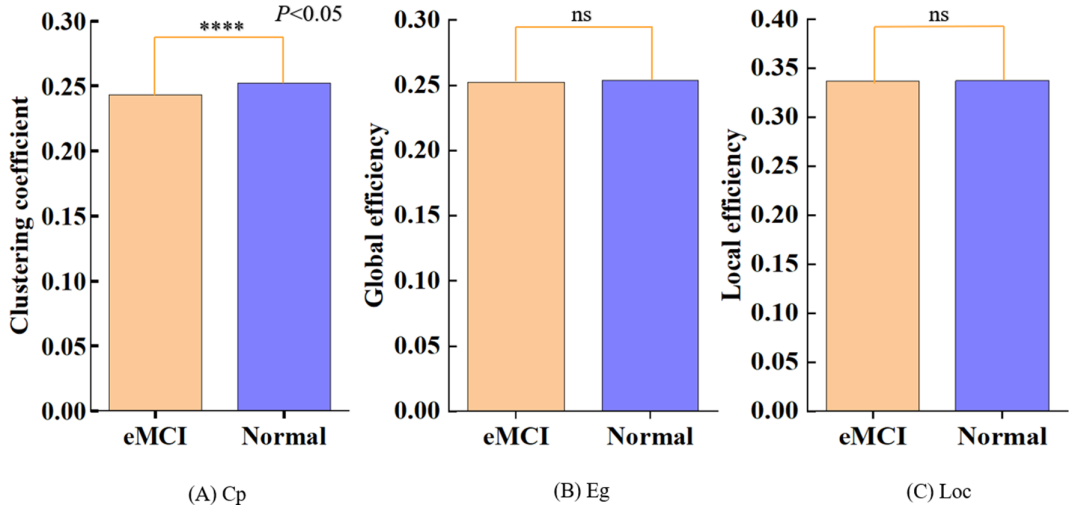

FIGURE 4 Difference of topological properties.

gyrus (PCG.L), ORBmid.L, SFGmed.L, CAL.R, and HIP. R. These findings validate the results of this study from an overall brain functional perspective and provide additional discriminative brain regions for the diagnosis of eMCI, holding significance for clinical applications.

3.5 | Topological properties analysis

Inspired by Zhang et al.,[44] we analyzed the graph theoretical topological properties of the BFN in eMCI patients and normal subjects. The Gretna toolbox was utilized to obtain the AUC for clustering coefficient (Cp), global efficiency (Eg), and local efficiency (Loc) within a threshold range. These three topological properties comprehensively described potential changes in terms of network aggregation, global information transfer efficiency, and local information transfer efficiency in eMCI patients. They possessed biologically intuitive explanations, aiding researchers in understanding the fundamental organizational principles of the BFN. We analyzed the differences of the above indicators by using two sample t-test. According to statistical research, topological properties with a p-value less than 0.05 are considered statistically significant, indicating a noticeable distinction. Figure 4 and Table 4 show the differences of topological properties.

In Figure 4, "****" represents statistically significant differences, while "ns" indicates no significant difference. Upon observing Figure 4 and Table 4, it is evident that

Topological property	eMCI patients	Normal subjects	*p*
Cp	0.243 ± 0.010	0.252 ± 0.021	0.001
Eg	0.252 ± 0.004	0.254 ± 0.005	0.285
Loc	0.337 ± 0.009	0.338 ± 0.011	0.383

TABLE 4 Comparison of topological properties.

eMCI patients exhibited a significant decrease in Cp compared to normal subjects. This decrease highlighted a reduction or weakening of connections between brain regions, leading to a decrease in local clustering. Although there was no significant difference between global efficiency and local efficiency, the values in eMCI patients showed a downward trend. This may be attributed to the decreased efficiency of information transmission due to a decline in cognitive function, although the decline was not highly significant in the early stages.

4 | DISCUSSION

Alzheimer's disease causes changes in the central nervous system of the brain and has extremely high morbidity and mortality.[2] eMCI has been focused on as a transitional stage between AD and normal aging.[45,46] Unfortunately, the pathological mechanisms of eMCI are still unclear, and its classification and identification are still a challenge. Despite eMCI being classified and identified via BFNs, the classification performance is affected due to the varying quality of brain region information in BFNs. This led to different degrees of limitations in existing methods for constructing BFNs when applied to disease classification. PC proposed by Biswal et al.[14] considered the correlation between brain regions but resulted in overly dense BFNs information. This made it challenging to discern which connections are disease-related, leading to erroneous connection information. SR proposed by Lee et al.[15] emphasized the sparsity of BFNs but overlooked many of its topological structures. MR proposed by Li et al.[18] considered topological structure information between two brain regions but did not address the complex topological structures among multiple brain regions. SHMR and DHMR did not adequately consider the differences among brain regions within different hyperedges and between hyperedges, thus limiting their classification performance.

The aforementioned methods failed to construct a manifold regularization term that better captures the high-order information of multiple brain regions in the expression of BFNs. Moreover, they neglected crucial topological structural information regarding the varying importance among brain regions. These methods solely considered low-order information between pairs of brain regions. They simplistically treated the relationships among multiple brain regions as indiscriminate high-order prior information, thus falling short in identifying discriminative brain regions reflecting eMCI changes. AMR flexibly integrated a graph regularizer representing the topological structure with adaptive hypergraph learning. It comprehensively captured the brain's topological structural information and preserved important high-order prior information among multiple brain regions. Specifically, DHMR classification performance was lower than that of AMR. It could be attributed to dynamic hypergraph overlooking the inherent improvement of hypergraph structure and weight, as well as fine-tuning multiple brain regions relationship information in individual sub-time series. AMR enhanced the original weight matrix by reordering the node and hyperedge information of each hypergraph through the hypergraph ranking optimization algorithm. The introduction of adaptive weight learning iteratively updated the ordering and weights, ultimately yielding a more information-rich adaptive hypergraph structure. The adaptive hypergraph learning placed greater emphasis on optimizing and adjusting the hypergraph structure when dealing with dynamic data, providing higher flexibility and accuracy in adapting to data variations. This made the adaptive hypergraph more suitable for preserving high-order relationships among multiple brain regions, thereby enhancing the method's classification performance compared to dynamic hypergraph.

Notably, the selected discriminative brain regions included THA.L and THA.R, while other eMCI classification methods paid less attention to them. Both THA.L and THA.R were involved in the formation of neural circuits related to memory and learning, and are closely connected with other brain regions such as HIP.R and HIP.L. The study suggested that MCI was associated with changes in cognitive function and a reduction in thalamic volume.[41] Furthermore, we found the clustering coefficient was decreased in eMCI patients, reflecting reduced modular information processing in patients' BFNs. It suggested that reduced modularity in BFNs might be a crucial feature of cognitive impairment in eMCI. In summary, AMR held potential application value in clinical diagnosis.

The proposed method has many shortcomings and limitations. Firstly, heterogeneity issues are introduced due to multisite data. We will consider incorporating

methods like federated learning[47] to address data heterogeneity in the upcoming work. Secondly, the dataset is too small, which affects the effectiveness of the proposed method. More data will be collected in the future to solve this problem. Furthermore, there may be efficiency and performance concerns due to frequent attention to irrelevant information in the construction of adaptive hypergraph. We consider introducing the attention mechanism[48] to increase the model's flexibility and focus on key information.

5 | CONCLUSIONS

We proposed an AMR method to construct a BFN and classify eMCI. It utilized a low-order BFN to construct hypergraph. The hypergraph ranking optimization and adaptive weight learning were introduced to enhance the original hypergraph structure, resulting in the adaptive hypergraph better suited to the data. High-order information was introduced into the low-order BFN adopting the Laplacian matrix as the manifold regularization term. This incorporation enriched the constructed BFN with more high-order information. The results indicated that AMR demonstrated excellent classification performance and stability on real datasets. In addition, the clustering coefficient of patients decreased, and reduced BFNs modularity may be a key feature of eMCI.

AUTHOR CONTRIBUTIONS

Xidong Fu and Zhuqing Jiao have been involved in the drafting and critical revision of the manuscript. Shengchang Shan, Chun Liu, and Yu Lu made substantial contributions to the concept and design. Xidong Fu completed the analysis and interpretation of the data.

ACKNOWLEDGMENTS

This work was supported by National Natural Science Foundation of China (Grant No. 51877013) and Jiangsu Provincial Key Research and Development Program (Grant No. BE2021636). This work was also sponsored by Qing Lan Project of Jiangsu Province.

CONFLICT OF INTEREST STATEMENT

The authors declare that there is no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available in available Alzheimer's disease neuroimaging initiative (ADNI) at https://adni.loni.usc.edu/. These data were derived from the following resources available in the public domain:—available Alzheimer's disease neuroimaging initiative (ADNI), https://adni.loni.usc.edu/.

ORCID

Zhuqing Jiao https://orcid.org/0000-0002-6547-8449

REFERENCES

1. Alzheimer's Association. 2016 Alzheimer's disease facts and figures. *Alzheimers Dement*. 2016;12(4):459-509.
2. Khazaee A, Ebrahimzadeh A, Babajani-Feremi A, Alzheimer's Disease Neuroimaging Initiative. Classification of patients with MCI and AD from healthy controls using directed graph measures of resting-state fMRI. *Behav Brain Res*. 2017;322:339-350.
3. Huang S, Li J, Sun L, et al. Learning brain connectivity of Alzheimer's disease from neuroimaging data. *Adv Neural Inf Process Syst*. 2009;22:808-816.
4. Kam TE, Zhang H, Jiao Z, Shen D. Deep learning of static and dynamic brain functional networks for early MCI detection. *IEEE Trans Med Imaging*. 2019;39(2):478-487.
5. Bi XA, Hu X, Wu H, Wang Y. Multimodal data analysis of Alzheimer's disease based on clustering evolutionary random forest. *IEEE J Biomed Health Inform*. 2020;24(10):2973-2983.
6. Zhang Y, Wang S, Phillips P, Yang J, Yuan TF. Three-dimensional eigenbrain for the detection of subjects and brain regions related with Alzheimer's disease. *J Alzheimers Dis*. 2016;50(4):1163-1179.
7. Xi ZT, Liu T, Shi H, Jiao ZQ. Hypergraph representation of multimodal brain networks for patients with end-stage renal disease associated with mild cognitive impairment. *Math Biosci Eng*. 2023;20(2):1882-1902.
8. Jann K, Gee DG, Kilroy E, et al. Functional connectivity in BOLD and CBF data: similarity and reliability of resting brain networks. *Neuroimage*. 2015;106:111-122.
9. Zhang RP, Fu XD, Song CF, Shi HF, Jiao ZQ. Characterizing topological properties of brain functional networks using multi-threshold derivative for end-stage renal disease with mild cognitive impairment. *Brain Sci*. 2023;13(8):1187.
10. Sheng Q, Zhang Y, Shi HF, Jiao ZQ. Global iterative optimization framework for predicting cognitive function statuses of patients with end-stage renal disease. *Int J Imaging Syst Technol*. 2023;33:837-852.
11. Jie B, Liu M, Shen D. Integration of temporal and spatial properties of dynamic connectivity networks for automatic diagnosis of brain disease. *Med Image Anal*. 2018;47:81-94.
12. Eyler LT, Elman JA, Hatton SN, et al. Resting state abnormalities of the default mode network in mild cognitive impairment: a systematic review and meta-analysis. *J Alzheimers Dis*. 2019;70(1):107-120.
13. Yang J, Wang F, Li Z, Yang Z, Dong X, Han Q. Constructing high-order functional networks based on hypergraph for diagnosis of autism spectrum disorders. *Front Neurosci*. 2023;17:1257982.
14. Biswal B, Zerrin Yetkin F, Haughton VM, Hyde JS. Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magn Reson Med*. 1995;34(4):537-541.
15. Lee H, Lee DS, Kang H, Kim BN, Chung MK. Sparse brain network recovery under compressed sensing. *IEEE Trans Med Imaging*. 2011;30(5):1154-1165.
16. Qiao L, Zhang H, Kim M, Teng S, Zhang L, Shen D. Estimating functional brain networks by incorporating a modularity prior. *Neuroimage*. 2016;141:399-407.
17. Wee CY, Yap PT, Shen D. Diagnosis of autism spectrum disorders using temporally distinct resting-state functional connectivity networks. *CNS Neurosci Ther*. 2016;22(3):212-219.

18. Li W, Xu X, Jiang W, Wang P, Gao X. Functional connectivity network estimation with an inter-similarity prior for mild cognitive impairment classification. *Aging (Albany NY)*. 2020; 12(17):17328-17342.
19. Ji Y, Zhang Y, Shi H, Jiao Z, Wang SH, Wang C. Constructing dynamic brain functional networks via hyper-graph manifold regularization for mild cognitive impairment classification. *Front Neurosci*. 2021;15:669345.
20. Xi Z, Song C, Zheng J, Shi H, Jiao Z. Brain functional networks with dynamic hypergraph manifold regularization for classification of end-stage renal disease associated with mild cognitive impairment. *Comput Model Eng Sci*. 2023;135(3): 2243-2266.
21. Rubinov M, Sporns O. Complex network measures of brain connectivity: uses and interpretations. *Neuroimage*. 2010;52(3): 1059-1069.
22. Nathalie TM, Brigitte L, Dimitri P, et al. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage*. 2002;15(1):273-289.
23. Lei B, Yu S, Zhao X, et al. Diagnosis of early Alzheimer's disease based on dynamic high order networks. *Brain Imaging Behav*. 2021;15:276-287.
24. Shao W, Peng Y, Zu C, Wang M, Zhang D, Alzheimer's Disease Neuroimaging Initiative. Hypergraph based multi-task feature selection for multimodal classification of Alzheimer's disease. *Comput Med Imaging Graph*. 2020;80:101663.
25. Wang M, Shao W, Huang S, Zhang DQ. Hypergraph-regularized multimodal learning by graph diffusion for imaging genetics based Alzheimer's disease diagnosis. *Med Image Anal*. 2023;89:102883.
26. Hao X, Li J, Ma M, Qin J, Zhang DQ, Liu F. Hypergraph convolutional network for longitudinal data analysis in Alzheimer's disease. *Computers in Biology Medicine*. 2024;168:107765.
27. Agarwal S, Branson K, Belongie S. Higher order learning with graphs. *Proceedings of the 23rd international conference on Machine learning*. 17–24. 2006.
28. Gao Y, Wang M, Zha ZJ, Shen J, Li X, Wu X. Visual-textual joint relevance learning for tag-based social image search. *IEEE Trans Image Process*. 2012;22(1):363-376.
29. Karantaidis G, Sarridis I, Kotropoulos C. Adaptive hypergraph learning with multi-stage optimizations for image and tag recommendation. *Signal Process Image Commun*. 2021;97:116367.
30. Zhou D, Huang J, Schölkopf B. Learning with hypergraphs: clustering, classification, and embedding. *Adv Neural Inf Process Syst*. 2006;19:1601-1608.
31. Yan CG, Cheung B, Kelly C, et al. A comprehensive assessment of regional variation in the impact of head micromovements on functional connectomics. *Neuroimage*. 2013;76:183-201.
32. Jiao Z, Xia Z, Ming X, Cheng C, Wang SH. Multi-scale feature combination of brain functional network for eMCI classification. *IEEE Access*. 2019;7:74263-74273.
33. Bi X, Xie Y, Wu H, et al. Identification of differential brain regions in MCI progression via clustering-evolutionary weighted SVM ensemble algorithm. *Front Comp Sci*. 2021;15:1-9.
34. Xu X, Li W, Mei J, et al. Feature selection and combination of information in the functional brain connectome for discrimination of mild cognitive impairment and analyses of altered brain patterns. *Front Aging Neurosci*. 2020;12:28.
35. Li Y, Liu J, Gao X, et al. Multimodal hyper-connectivity of functional networks using functionally-weighted LASSO for MCI classification. *Med Image Anal*. 2018;52(10):80-96.
36. Zhang YD, Wang SH, Sui YX, et al. Multivariate approach for Alzheimer's disease detection using stationary wavelet entropy and predator-prey particle swarm optimization. *J Alzheimers Dis*. 2018;65(3):855-869.
37. Zhang YD, Wang SH, Dong ZC. Classification of Alzheimer disease based on structural magnetic resonance imaging by kernel support vector machine decision tree. *Prog Electromagn Res*. 2014;144:171-184.
38. Wang SH. Detection of Alzheimer's disease by three-dimensional displacement field estimation in structural magnetic resonance imaging. *J Alzheimers Dis*. 2016;50(1):233-248.
39. Salvatore C, Cerasa A, Castiglioni I. MRI characterizes the progressive course of AD and predicts conversion to Alzheimer's dementia 24 months before probable diagnosis. *Front Aging Neurosci*. 2018;10:135.
40. Zhang YD, Wang SH. Detection of Alzheimer's disease by displacement field and machine learning. *PeerJ*. 2015;3:e1251.
41. Zhou Y, Si X, Chen Y, et al. Hippocampus-and thalamus-related fiber-specific white matter reductions in mild cognitive impairment. *Cereb Cortex*. 2022;32(15):3159-3174.
42. McKhann GM, Knopman DS, Chertkow H, et al. The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement*. 2011;7(3):263-269.
43. Tang C, Wei M, Sun J, Wang S, Zhang YD. CsAGP: detecting Alzheimer's disease from multimodal images via dual-transformer with cross-attention and graph pooling. *J King Saud Univ-Comput Inf*. 2023;35:101618.
44. Zhang Y, Sheng Q, Fu X, Shi H, Jiao Z. Integrated prediction framework for clinical scores of cognitive functions in ESRD patients. *Comput Intell Neurosci*. 2022;2022:1-12.
45. Wang SH, Zhang Y, Li YJ, et al. Single slice based detection for Alzheimer's disease via wavelet entropy and multilayer perceptron trained by biogeography-based optimization. *Multimed Tools Appl*. 2018;77(9):10393-10417.
46. Zhang YD. Detection of Alzheimer's disease and mild cognitive impairment based on structural volumetric MR images using 3D-DWT and WTA-KSVM trained by PSOTVAC. *Biomed Signal Process Control*. 2015;21:58-73.
47. Han T, Gong X, Feng F, Zhang J, Sun Z, Zhang Y. Privacy-preserving multi-source domain adaptation for medical data. *IEEE J Biomed Health Inform*. 2022;27(2):842-853.
48. Ji J, Ren Y, Lei M. FC–HAT: hypergraph attention network for functional brain network classification. *Inform Sci*. 2022;608: 1301-1316.

How to cite this article: Fu X, Shan S, Liu C, Lu Y, Jiao Z. Constructing brain functional networks with adaptive manifold regularization for early mild cognitive impairment. *Int J Imaging Syst Technol*. 2024;34(2):e23053. doi:10.1002/ima.23053