



## OPEN ACCESS

## EDITED BY

Jared Brent Smith,  
Regenxbio Inc., United States

## REVIEWED BY

Shengchao Zhang,  
Vanderbilt University, United States  
Mohammed Alsubaie,  
Taif University, Saudi Arabia

## \*CORRESPONDENCE

Sanghoon Han  
✉ sanghoon.han@yonsei.ac.kr

†These authors have contributed equally to  
this work

RECEIVED 18 March 2024

ACCEPTED 19 November 2024

PUBLISHED 11 December 2024

## CITATION

Jeon Y, Kim J-J, Yu S, Choi J and Han S (2024)  
A fusion analytic framework for investigating  
functional brain connectivity differences using  
resting-state fMRI.  
*Front. Neurosci.* 18:1402657.  
doi: 10.3389/fnins.2024.1402657

## COPYRIGHT

© 2024 Jeon, Kim, Yu, Choi and Han. This is  
an open-access article distributed under the  
terms of the [Creative Commons Attribution  
License \(CC BY\)](#). The use, distribution or  
reproduction in other forums is permitted,  
provided the original author(s) and the  
copyright owner(s) are credited and that the  
original publication in this journal is cited, in  
accordance with accepted academic practice.  
No use, distribution or reproduction is  
permitted which does not comply with these  
terms.

# A fusion analytic framework for investigating functional brain connectivity differences using resting-state fMRI

Yeseul Jeon<sup>1†</sup>, Jeong-Jae Kim<sup>2†</sup>, SuMin Yu<sup>3</sup>, Junggu Choi<sup>4</sup> and  
Sanghoon Han<sup>2,5\*</sup>

<sup>1</sup>Department of Statistics, Texas A&M University, College Station, TX, United States, <sup>2</sup>Graduate Program in Cognitive Science, Yonsei University, Seoul, Republic of Korea, <sup>3</sup>Department of Psychology and Neuroscience, Duke University, Durham, NC, United States, <sup>4</sup>Cancer Biology, Cleveland Clinic, Lerner Research Institute, Cleveland, OH, United States, <sup>5</sup>Department of Psychology, Yonsei University, Seoul, Republic of Korea

**Introduction:** Functional magnetic resonance imaging (fMRI) data is highly complex and high-dimensional, capturing signals from regions of interest (ROIs) with intricate correlations. Analyzing such data is particularly challenging, especially in resting-state fMRI, where patterns are less identifiable without task-specific contexts. Nonetheless, interconnections among ROIs provide essential insights into brain activity and exhibit unique characteristics across groups.

**Methods:** To address these challenges, we propose an interpretable fusion analytic framework to identify and understand ROI connectivity differences between two groups, revealing their distinctive features. The framework involves three steps: first, constructing ROI-based Functional Connectivity Networks (FCNs) to manage resting-state fMRI data; second, employing a Self-Attention Deep Learning Model (Self-Attn) for binary classification to generate attention distributions encoding group-level differences; and third, utilizing a Latent Space Item-Response Model (LSIRM) to extract group-representative ROI features, visualized on group summary FCNs.

**Results:** We applied our framework to analyze four types of cognitive impairments, demonstrating their effectiveness in identifying significant ROIs that contribute to the differences between the two disease groups. The results reveal distinct connectivity patterns and unique ROI features, which differentiate cognitive impairments. Specifically, our framework highlighted group-specific differences in functional connectivity, validating its capability to capture meaningful insights from high-dimensional fMRI data.

**Discussion:** Our novel interpretable fusion analytic framework addresses the challenges of analyzing high-dimensional, resting-state fMRI data. By integrating FCNs, a Self-Attention Deep Learning Model, and LSIRM, the framework provides an innovative approach to discovering ROI connectivity disparities between groups. The attention distribution and group-representative ROI features offer interpretable insights into brain activity patterns and their variations among cognitive impairment groups. This methodology has significant potential to enhance our understanding of cognitive impairments, paving the way for more targeted therapeutic interventions.

## KEYWORDS

fMRI, ADNI, functional connectivity network, deep learning, Latent Space Item-Response Model

## 1 Introduction

Functional magnetic resonance imaging (fMRI) data, particularly resting-state fMRI (rs-fMRI), is inherently complex and high-dimensional. This complexity results in correlated matrices that capture signals from regions of interest (ROIs) in the brain at each time point. Several attempts have been made to analyze fMRI data to understand the roles of ROIs in specific tasks or symptoms (Santana et al., 2022; Wang et al., 2016). Comparing ROIs in fMRI data from different tasks has been one approach to understanding their mechanisms and identifying group differences (Li et al., 2018; Lee and Lee, 2020). However, interpreting which features of ROI connections differentiate between groups has proven challenging for previous studies (Smith et al., 2010; Gates and Molenaar, 2012). There are two main reasons for this difficulty: First, the high-dimensional and correlated nature of fMRI datasets makes it difficult to apply standard statistical models, which typically assume that the data is independent and identically distributed. The complex interactions and dependencies among ROIs in fMRI data make this independence assumption unrealistic, leading to biased or inaccurate interpretations (Smith et al., 2010, 2011). Second, identifying distinctive ROI connectivity that represents group differences is challenging because of the noise introduced by individual effects. Each fMRI dataset corresponds to an independent subject, and the inherent variability and noise in individual effects can obscure the true underlying patterns that distinguish different groups or conditions (Dubois and Adolphs, 2016).

To address these limitations, we propose a novel analytic framework that integrates a deep learning-based classification model with a statistical model, while providing visual interpretation through the functional connectivity networks (FCNs) of ROIs to offer intuitive insights. Since deep learning models are well-known for handling high-dimensional correlated structured data (Du et al., 2022), they are appropriate to apply fMRI data that exhibit complex interactions and dependencies among ROIs. In this study, we utilize a Self-Attention Deep Learning Model (Self-Attn). Self-Attn employs the self-attention mechanism (Vaswani et al., 2017), which is capable of handling correlated structured data and effectively learning adjacency connections (Chen et al., 2018; Zheng et al., 2019; Sun et al., 2020). This enables us to capture intricate connectivity patterns between ROIs in fMRI data. The self-attention mechanism provides an attention distribution for the ROIs, indicating how Self-Attn learns the relationships in the structured input data. Each row in the attention distribution reflects the likelihood of how a specific ROI relates to other ROIs. If the classification accuracy is sufficient, the output of the attention distribution for each subject's ROIs is a reliable source for identifying the ROI connections that distinguish different groups.

However, it is still challenging to identify which ROI connections differentiate groups by manually comparing these distributions. To address this, we analyze the ROIs' attention distribution using the Latent Space Item-Response Model (LSIRM) (Jeon Y. et al., 2021), a statistical network model. We interpret the attention distribution as an item-response matrix (Embretson and Reise, 2013), where ROIs represent items and subjects represent respondents. To the best of our knowledge, there has been no prior research that analyzes the ROIs' attention distribution in the context of statistical network models. Here, the

LSIRM estimates relationships between respondents and ROIs by modeling the probabilities of positive responses (connections). This estimation allows our framework to select group-representative ROIs that are consistently shown as meaningful ROIs across all individuals within each group. As a result, our framework effectively tackles the previous challenge of noise caused by individual variations. These distinctive ROI connections are then visualized on the group summary FCN.

The overview of the proposed framework is illustrated in Figure 1, and detailed aspects are presented in Figure 2. Our framework comprises three key steps. In Step 1 (Figures 2A–D), we construct FCNs for each subject's ROIs by connecting them based on their embedded positions using the mapper algorithm (Chazal and Michel, 2017). To manage the time dimension of rs-fMRI data, we first apply dimensionality reduction, and then use the mapper to construct individual FCNs. While this approach reveals the overall connectivity structure, it remains challenging to identify significant ROI connections that distinguish one group from another. In Step 2 (Figures 2E–H), we perform binary classification using Self-Attn (Vaswani et al., 2017) on the subjects' FCNs and generate an attention distribution matrix for each subject. Next, we use the coefficient of variation (CV) and mean values to build a group-level attention distribution matrix. In Step 3 (Figures 2I–J), we apply LSIRM to the group attention distribution matrix to identify group-representative ROIs. These ROIs are then mapped onto a group summary FCN to extend their connections to other ROIs.

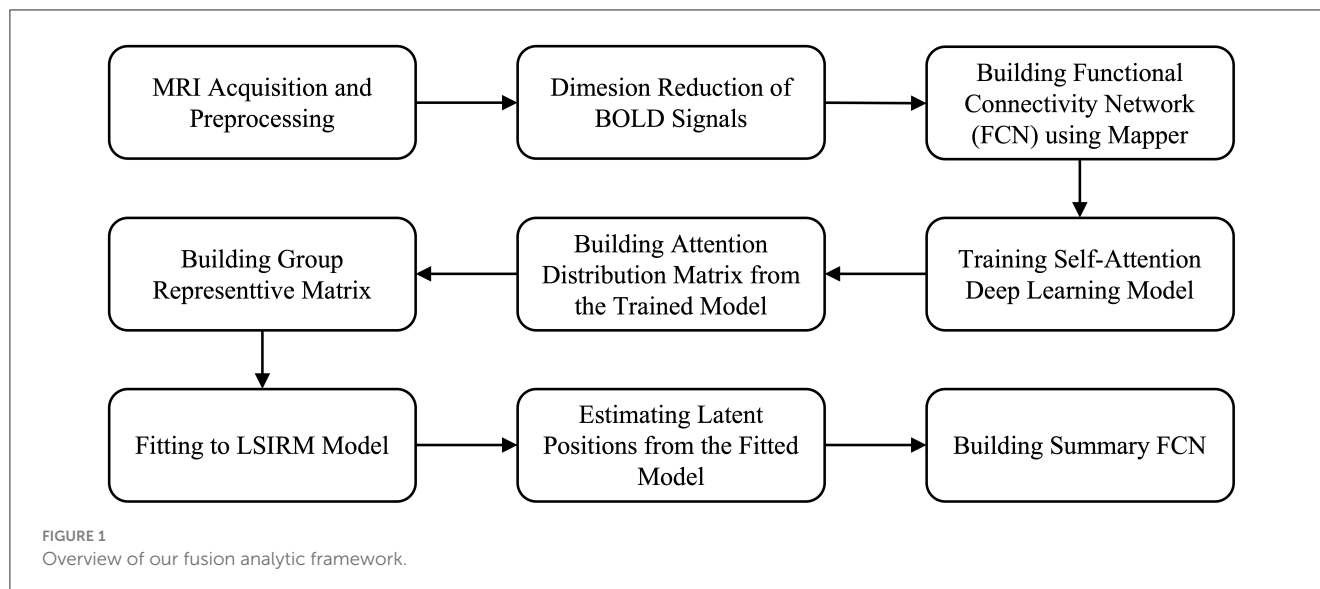
To validate the proposed framework, we applied it to classify rs-fMRI data from different stages of neurodegenerative diseases, which exhibit varying levels of cognitive impairment. We used resting brain scans from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database, which is a multisite longitudinal study widely used for biomarker exploration in Alzheimer's disease diagnosis (Jack Jr et al., 2008; Mueller et al., 2005). We aimed to uncover the distinct features of ROIs between the two groups and compare our results with previous findings from ADNI publications.

## 2 Materials and methods

Our analysis approach involves three main steps: (1) creating an FCN for each subject in each group, (2) estimating an attention distribution matrix using Self-Attn, and (3) extracting group-representative features of ROIs using LSIRM and visualizing them on the group summary FCN. In this study, we applied our analysis framework to identify the specific ROIs that differentiate four comparisons: Alzheimer's Disease (AD) vs. Mild Cognitive Impairment (MCI), AD vs. Early MCI (EMCI), AD vs. Late MCI (LMCI), and EMCI vs. LMCI. We utilized rs-fMRI data collected from AD, EMCI, MCI, and LMCI groups in the ADNI dataset.

### 2.1 ADNI study

The ADNI dataset is composed of four consecutive cohorts (ADNI1, ADNI2, ADNI-GO, and ADNI3). Participants were recruited for initial periods in the ADNI1 cohorts (October 2004). Follow-up of participants were recruited to the ADNI3 cohort period. To facilitate the preprocessing of the brain rs-fMRI data,



we selected data with the same MR parameters. Among several MR parameters, three MR parameters were applied as criteria (200 time points, TR = 3000 ms, 48 slices). After filtering based on these three conditions, 281 participants remained in the ADNI2, ADNI-GO, and ADNI3 cohorts. The ADNI1 cohort was excluded because it did not contain data that met the aforementioned conditions. As a result, we used axial rs-fMRI data from 57 AD subjects, 93 EMCI subjects, 53 LMCI subjects, and 78 MCI subjects (Table 1, Figure 2A). By focusing on these specific disease comparisons, we aim to uncover the key ROIs that exhibit distinct patterns and contribute significantly to the classification and differentiation of these cognitive impairment conditions. All data are publicly available, at <http://adni.loni.usc.edu/>.

## 2.2 MRI acquisition

The participants included in this study participated in scanning at diverse sites using 3T MRI scanners manufactured by Philips Medical Systems and Siemens Healthineers. The detailed MRI protocols of the ADNI dataset were reported on the webpage (<http://adni.loni.usc.edu/methods/mri-tool/mri-acquisition/>). In the ADNI2 and ADNI-GO cohorts, MRI scanning was performed at 26 different sites with Philips 3T MRI scanners, using synchronized scanning parameters. In the case of the ADNI3 cohort, Siemens 3T MRI scanners were used to collect fMRI data with synchronized parameters.

## 2.3 MRI preprocessing

The rs-fMRI datasets, originally formatted in Digital Imaging and Communications in Medicine (DICM), were converted to the Neuroimaging Informatics Technology Initiative (NITI) format, the standard in fMRI research. This conversion preserved all original slices across four dimensions ( $x$ ,  $y$ ,  $z$ , and time). The preprocessing steps began with slice timing correction to account for the acquisition order of slices, as required by the MRI protocol.

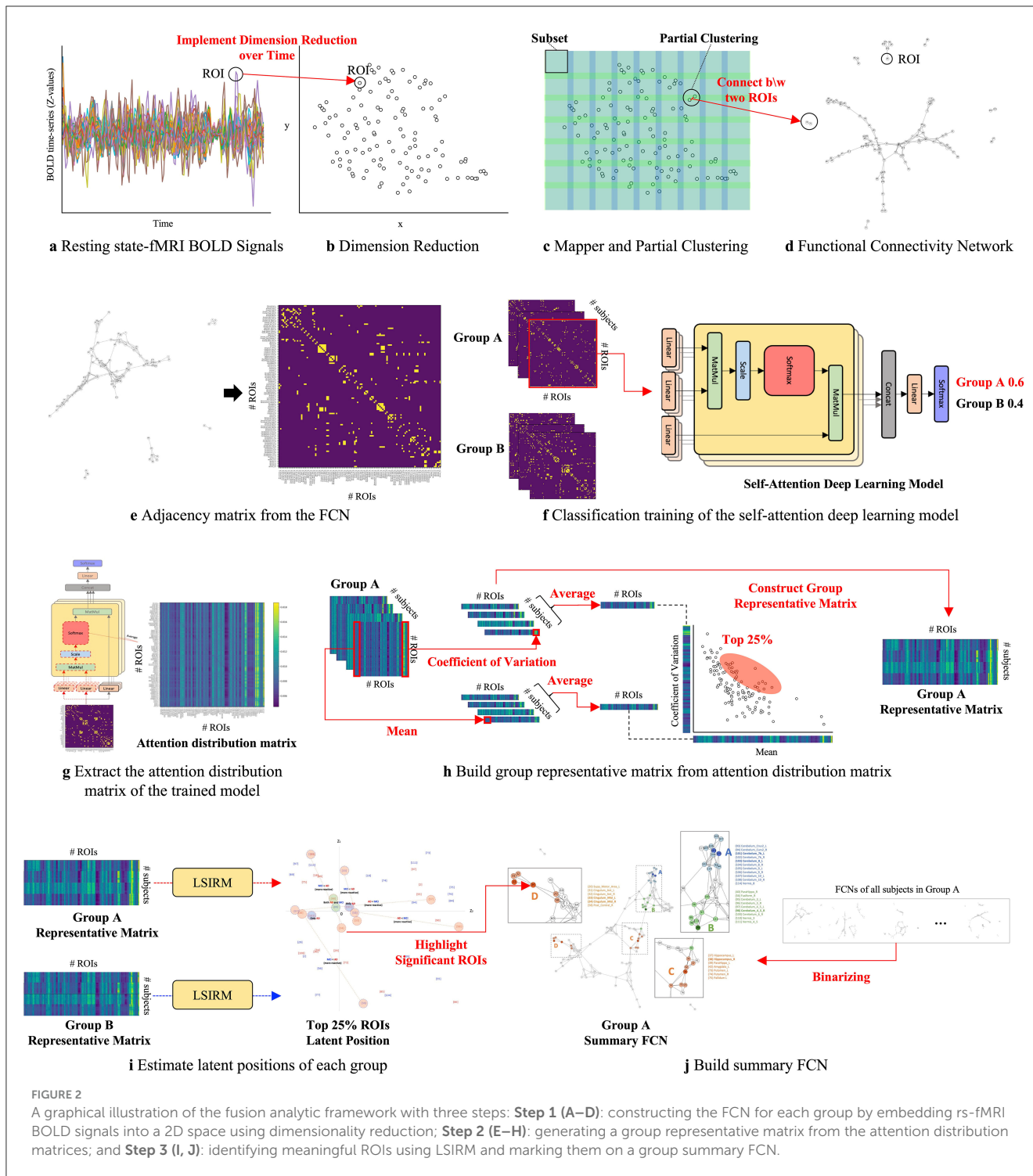
Since all rs-fMRI data were acquired using interleaved scanning (an alternating acquisition method), the slice order was applied accordingly, alternating between even- and odd-numbered slices (1 to 48). Following slice timing correction, realignment and head motion correction were performed to address misalignment or movement artifacts.

Next, the corrected rs-fMRI data were spatially normalized to a 3 mm isotropic voxel size using an EPI template, ensuring anatomical consistency across brain MR images. Smoothing was applied using a Gaussian kernel (FWHM = 6 mm, full-width at half-maximum) to optimize spatial resolution and reduce noise. To account for scanner drift and physiological fluctuations, linear trends were removed, and covariates such as white matter (WM) and cerebrospinal fluid (CSF) signals were regressed out, isolating relevant brain activation data. Temporal bandpass filtering (0.01–0.1 Hz) was then applied to minimize low- and high-frequency noise, preserving fluctuations related to intrinsic brain activity.

Finally, the Automatic Anatomical Labeling (AAL) atlas, which segments the brain into 116 regions, was used to define brain regions and extract ROI time-course data. All preprocessing steps were conducted using the Data Processing and Analysis of Brain Imaging toolbox (DPABI, Version 5.3, available at <http://rfmri.org/dpabi>), and Statistical Parametric Mapping (SPM, Version 12, available at [www.fil.ion.ucl.ac.uk/spm/software/spm12/](http://www.fil.ion.ucl.ac.uk/spm/software/spm12/)), both implemented in MATLAB R2020b (MathWorks, Natick, MA, USA) (Yan et al., 2016).

## 2.4 Functional connectivity networks

Dimension reduction methods are well-established techniques for embedding complex and structured data, including Principal Component Analysis (PCA) (Dunteman, 1989), T-distributed Stochastic Neighbor Embedding (t-SNE) (van der Maaten and Hinton, 2008), and Uniform Manifold Approximation and Projection (UMAP) (McInnes et al., 2018). Both t-SNE and UMAP model the manifold using stochastic and topological information. Respectively, t-SNE converts neighborhood distances



into conditional probabilities that represent similarity, while UMAP employs a fuzzy simplicial complex with edge weights that reflect the likelihood of connectivity.

### 2.4.1 Dimension reduction

PCA (Dunteman, 1989) is a technique that uses an orthogonal transformation to reduce high-dimensional data to low-dimensional data (main components) that are not linearly

related. The axis with the most significant variance is the first principal component, and the second greatest variance is the second principal component. This decomposition divides the sample into components that best represent the differences. On the other hand, t-SNE (van der Maaten and Hinton, 2008) is a non-linear dimension reduction method that aids in understanding the data with impact information. It is based on t-distribution, which has a heavier tail than normal distribution that helps cover up the far distribution element of high-dimensional data. The

**TABLE 1** Demographic and clinical information (mean ± standard deviation) of the studied ADNI subjects.

	Female%	Age (mean±SD)	MMSE (mean±SD)	Global CDR (mean±SD)
AD	38.6%	74.39±8.26	21.03±4.08	0.98±0.51
EMCI	47.3%	75.78±6.90	28.70±1.64	0.09±0.36
LMCI	35.8%	75.03±7.29	26.55±3.46	0.58±0.34
MCI	48.7%	74.81±8.71	27.20±2.25	0.47±0.23

t-SNE results depict the embedded points whose distances, trained by calculating the points' similarity in structure, reflect their degree of similarity. UMAP (McInnes et al., 2018) is a nonlinear dimension reduction method that models the manifold using a topological structure. Because it is based on topological space, the embedding points are close in proximity if the two data points have similar topological features. It first reorganizes the data into a fuzzy simplicial complex. This complex produces the connections based on the hyper-parameter that controls the connectivity around the data. Then, it projects the correlated structured data into a low-dimensional space based on their connection, where the connection indicates the proximity.

### 2.4.2 Mapper

Mapper is one of the techniques derived from topological data analysis, which enables the representation of the topological structure of high-dimensional data as a network. Topological data analysis simplifies the complexity of the topological space by transforming it into a network consisting of nodes and connections that capture the topological characteristics, such as points, lines, and triangles. The Mapper process involves two main steps. First, the high-dimensional topological space is mapped onto a real-valued measure space represented as a graph. This mapping function can be any real-valued function that captures the essential features of the data. For example, dimensionality reduction techniques like PCA, t-SNE, and UMAP use real-valued functions to project high-dimensional data into a lower-dimensional space such as Euclidean space. In the next step, the mapper partitions the graph into subsets of data, and each subset is clustered to define connections. This process identifies the structural relationships within the data. This process is called the *Mapper*, where each sub-cluster is treated as a node, and nodes are connected when they share similar data attributes.

The Mapper can be considered a form of partial clustering. It applies a standard clustering algorithm to subsets of the original data and examines the interactions between the resulting sub-clusters. When two non-empty subsets U and V are considered, their sub-clusters may have overlapping elements that construct a simplicial complex. The sub-clusters are referred to as vertices (or nodes), while the overlapping elements form edges in the complex. This process yields a simplicial complex consisting of points, lines, and triangles, which provides insights into the topological structure of high-dimensional data.

```

Require: Data  $\mathcal{D} = \{(FCN^{(i)}, group^{(i)})_{i=1}^N\}$ , where  $group^{(i)}$  is the group label for each subject  $i$ ,  $FCN^{(i)} \in \mathbb{R}^{116 \times 116}$ , number of attention heads  $H = 128$ 
Ensure: Attention distribution matrix set  $\{attnM^{(i)}\}_{i=1}^N$ , where  $attnM^{(i)} \in \mathbb{R}^{116 \times 116}$ 
1: Split  $\mathcal{D}$  into training and test datasets
2: Initialize Self-Attn with  $H$  heads
3: Train Self-Attn using the training data  $\{(FCN^{(i)}, group^{(i)})_{i=1}^{N_{train}}\}$ 
4: After training:
5: for each subject  $i = 1$  to  $N$  do
6:   Feed  $FCN^{(i)}$  into the trained Self-Attn
7:   Extract attention weights  $attn_h^{(i)} \in \mathbb{R}^{116 \times 116}$  for each attention head  $h = 1, 2, \dots, H$ 
8:   Compute the Attention Distribution Matrix  $attnM^{(i)} = \frac{1}{H} \sum_{h=1}^H attn_h^{(i)}$ , where  $attnM^{(i)} \in \mathbb{R}^{116 \times 116}$ 
9: end for
10: Collect all Attention Distribution Matrices:  $\{attnM^{(i)}\}_{i=1}^N$ 
11: return  $\{attnM^{(i)}\}_{i=1}^N$ 
    
```

**Algorithm 1.** Self-Attn for training attention distribution matrix.

## 2.5 Attention distribution matrix

The attention distribution matrix is obtained from Self-Attn, representing a probability distribution that captures the relationships between input elements that play a key role in performing a task. The training phase, as detailed in Algorithm 1, is shown in Figure 2F, where the FCN is trained using Self-Attn. In this phase, the model learns the relationships between ROIs from the input FCN and identifies which network of ROIs characterizes the unique features of a given group. The inference phase corresponds to Figure 2G, where the trained model takes the FCN of a subject as input and extracts values from a specific layer, generating the attention distribution matrix. This layer captures the relationships between ROIs in parallel, in the form of a probability distribution, thus capturing diverse interaction patterns among the ROIs. Finally, these relationships are averaged to produce a single matrix which is called the attention distribution matrix. This matrix explains the relationships between ROIs that contribute to classifying a subject into a specific group.

In particular, Self-Attn with multi-head self-attention can distinguish between groups but also explains the relationships among ROIs that contribute to group classification. Multi-head self-attention is one of the core mechanisms of the Transformer model. It is highly effective in learning how different elements in input data are related to each other (Vaswani et al., 2017). It is particularly useful for exploring interactions between brain regions (Zhao et al., 2022, 2024). Self-attention evaluates each element's relationships, focusing on the more important elements based on this information. This is particularly useful for capturing the interactions between ROIs and identifying relationships that distinguish groups (Velickovic et al., 2017; Lei et al., 2022; Zhang et al., 2022).

Key components of the multi-head self-attention mechanism, which plays an important role in executing self-attention, are known as Query (Q), Key (K), and Value (V). These components are used to learn the relationships between input elements. Q functions as a 'query,' capturing how a specific ROI relates to

**Require:** Data  $\mathbf{X}_{h|g,h} = \{(y_{ij})_{i=1}^{N_h}\}_{j=1}^R$  and hyperparameters  $a$  and  $b$  for priors  $\sigma_\theta^2$

**Ensure:** Posterior samples of latent positions  $\mathbf{u}$  (ROIs),  $\mathbf{v}$  (subjects),  $\beta$  (ROI difficulties), and  $\theta$  (subject abilities)

- 1: Initialize  $\mathbf{u}$ ,  $\mathbf{v}$ ,  $\beta$ ,  $\theta$  randomly
- 2: **for** each MCMC iteration **do**
- 3:   **for** each subject  $i=1$  to  $N_h$  **do**
- 4:     Sample  $\theta_i$  from its full conditional posterior:
 
$$\theta_i^* \sim \text{Normal}(0, \sigma_\theta^2)$$

where  $\sigma_\theta^2$  are based on the likelihood and prior.
- 5:     Compute the acceptance probability for  $\theta_i$ :
 
$$\alpha_{\theta_i} = \min\left(1, \frac{P(\theta_i^* | \mathbf{y}, \sigma_\theta^2, \beta, \mathbf{u}, \mathbf{v})}{P(\theta_i | \mathbf{y}, \sigma_\theta^2, \beta, \mathbf{u}, \mathbf{v})}\right)$$
- 6:     Accept or reject  $\theta_i^*$ .
- 7:   **end for**
- 8:   Sample subject-specific variance  $\sigma_\theta^2$  from its full conditional posterior:
 
$$\sigma_\theta^2 \sim \text{Inv-Gamma}\left(a + \frac{N}{2}, b + \frac{1}{2} \sum_{i=1}^N \theta_i^2\right)$$
- 9:   **for** each ROI  $j=1$  to  $R$  **do**
- 10:     Sample  $\beta_j$  from its full conditional posterior:
 
$$\beta_j^* \sim \text{Normal}(0, 1)$$
- 11:     Compute the acceptance probability for  $\beta_j$ :
 
$$\alpha_{\beta_j} = \min\left(1, \frac{P(\beta_j^* | \mathbf{y}, \theta, \mathbf{u}, \mathbf{v})}{P(\beta_j | \mathbf{y}, \theta, \mathbf{u}, \mathbf{v})}\right)$$
- 12:     Accept or reject  $\beta_j^*$ .
- 13:   **end for**
- 14:   **for** each subject  $i=1$  to  $N_h$  **do**
- 15:     Sample latent position  $\mathbf{v}_i$  from its full conditional posterior:
 
$$\mathbf{v}_i^* \sim \text{Normal}(0, 1)$$
- 16:     Accept or reject  $\mathbf{v}_i^*$  based on its acceptance probability.
- 17:   **end for**
- 18:   **for** each ROI  $j=1$  to  $R$  **do**
- 19:     Sample latent position  $\mathbf{u}_j$  from its full conditional posterior:
 
$$\mathbf{u}_j^* \sim \text{Normal}(0, 1)$$
- 20:     Accept or reject  $\mathbf{u}_j^*$  based on its acceptance probability.
- 21:   **end for**
- 22: **end for**
- 23: Return posterior samples for  $\mathbf{u}$ ,  $\mathbf{v}$ ,  $\beta$ ,  $\theta$ ,  $\sigma_\theta^2$

**Algorithm 2.** LSIRM training process using Markov chain Monte Carlo (MCMC).

other ROIs.  $\mathbf{K}$  holds the features of each ROI, encapsulating the information each ROI possesses, and is used together with  $\mathbf{Q}$  to assess the relevance between ROIs. For instance, if the dot product between  $\mathbf{Q}$  and  $\mathbf{K}$  is high, the corresponding ROIs have highly significant interactions.  $\mathbf{V}$  contains the actual information of each ROI and is propagated using the attention weights derived from  $\mathbf{Q}$  and  $\mathbf{K}$ . For example, if a particular ROI receives high attention weights, it extracts crucial information from the  $\mathbf{V}$  of other ROIs, generating more meaningful outputs.

The input consists of  $\mathbf{Q}$ ,  $\mathbf{K}$ , and  $\mathbf{V}$ , all having a dimensionality of  $d_k$ . The computation involves calculating the dot products between  $\mathbf{Q}$  and each  $\mathbf{K}$ , dividing the result of each by the square root of  $d_k$ , and then applying a softmax function to obtain weights corresponding to  $\mathbf{V}$ . If we consider a total of  $R$  ROIs, we can represent  $\mathbf{Q}$ ,  $\mathbf{K}$ , and  $\mathbf{V}$  as element of  $\mathbb{R}^{R \times d_k}$ . By applying softmax to the attention scores obtained from the dot product of  $\mathbf{Q}$  and  $\mathbf{K}$ , we derive the attention weights, denoted by  $\text{attn}(\mathbf{Q}, \mathbf{K})$ . These attention weights represent the relationships between ROIs and indicate the probability distribution of how much each ROI should focus on other ROIs. The final output is obtained by multiplying  $\mathbf{V}$  with  $\text{attn}(\mathbf{Q}, \mathbf{K})$ .

$$\text{attn}(\mathbf{Q}, \mathbf{K}) = \text{Softmax}\left(\frac{\mathbf{Q}\mathbf{K}^\top}{\sqrt{d_k}}\right) \quad (1)$$

$$\text{Attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{attn}(\mathbf{Q}, \mathbf{K})\mathbf{V}$$

Equation 1 is extended into multi-head self-attention as shown in Equation 2 to capture various aspects of ROIs, particularly in high-dimensional datasets. Here, with  $H$  heads,  $\text{head}_h$  where  $h = 1, \dots, H$ , each  $\text{head}_h$  layer consists of weight matrices  $W_h^Q$ ,  $W_h^K$ , and  $W_h^V$ , which applied to the output layer. The multi-head self-attention is formed by concatenating these  $H$  sets of  $\text{head}_h$ . The resulting output is passed through a multi-layer perceptron.

$$\text{MultiHead}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{Concat}(\text{head}_1, \dots, \text{head}_H)W^O \quad (2)$$

where  $\text{head}_h = \text{Attention}(\mathbf{Q}W_h^Q, \mathbf{K}W_h^K, \mathbf{V}W_h^V)$

In the end, we obtain the attention distribution matrix from the trained model via the inference phase (Figure 2G). For each input, the attention distribution matrix  $\text{attnM}(\mathbf{Q}, \mathbf{K})$  is obtained by averaging the attention layers across all  $H$  heads, denoted as  $\text{attn}_h$ .

$$\text{attnM}(\mathbf{Q}, \mathbf{K}) = \frac{\sum_{i=1}^H \text{attn}_i}{H} \quad (3)$$

where  $\text{attn}_i = \text{attn}(\mathbf{Q}W_i^Q, \mathbf{K}W_i^K)$

## 2.6 Group representative ROIs features using Latent Space Item-Response Model

LSIRM (Jeon M. et al., 2021) is a model that represents item-response datasets as bipartite networks, estimating interactions between items (ROIs) and respondents (subjects). In our study, we aim to estimate the latent positions of ROIs based on the

interactions between subjects and ROIs. Here, the “Interaction” is measured by the degree of value between subjects and ROIs, indicating the association between subjects and ROIs. These patterns are visualized by estimating the latent positions in space, which can be in Euclidean space, allowing for a more intuitive understanding of these associations. The original LSIRM model is designed for a binary item-response data (0 or 1) (Embretson and Reise, 2013). Therefore, we adopt a continuous version of LSIRM to apply it to group representative matrix  $\mathbf{X}_{h|g,h}$ . Each cell value  $y_{ij}$  represents the coefficient of variation of ROI  $j$  in the attention distribution of subject  $i$  from group  $h$  compared to group  $g$ . This is continuous for  $i = 1, \dots, N_h$ , and  $j = 1, \dots, R$ . Equation 4 shows the continuous version of LSIRM:

$$\mathbb{P}(y_{ij} | \Theta) \sim \text{Normal}(\theta_j + \beta_i - \|\mathbf{u}_j - \mathbf{v}_i\|, \sigma^2), \quad (4)$$

where  $\Theta$  represents  $\{\theta_j\}$ ,  $\beta = \{\beta_i\}$ ,  $\mathbf{U} = \{\mathbf{u}_j\}_{j=1}^R$ ,  $\mathbf{V} = \{\mathbf{v}_i\}_{i=1}^{N_h}$  and  $\|\mathbf{u}_j - \mathbf{v}_i\|$  denotes the Euclidean distance between subject  $i$  and ROI  $j$ . LSIRM consists of two parts, the attribute part and the interaction part. In the attribute part, there are two parameters:  $\theta_j \in \mathbb{R}$  and  $\beta_i \in \mathbb{R}$ . The parameter  $\beta_i$  represents the degree of responses for the subject  $i$ , while  $\theta_j$  represents the responses for ROI  $j$ . In the interaction part, we have the latent configurations  $\mathbf{u}_j$  and  $\mathbf{v}_i$  for each ROI  $j$  and subject  $i$ , respectively.

If subject  $i$  shows a high value in ROI  $j$ , the corresponding association is relatively strong, which is represented through their distance  $\mathbf{u}_j - \mathbf{v}_i$ . Therefore, their latent positions  $\mathbf{u}_j$  and  $\mathbf{v}_i$  become closer because of a smaller distance compared to other associations. Conversely, if subject  $i$  shows low values in ROI  $j$ , the association is weaker, and the distance between their latent positions  $\mathbf{u}_j$  and  $\mathbf{v}_i$  is comparably larger. Based on the latent positions of ROIs and subjects, we can interpret the overall relationships that are inherent in data. Algorithm 2 outlines the detailed training procedure for LSIRM.

Additionally, we can understand the significance of the latent positions of ROIs. When the latent positions of ROIs are located near the center, it indicates that most subjects respond similarly to these ROIs. This is because, when estimating the latent positions of these ROIs, the positions of all subjects are considered simultaneously, resulting in a minimization of the distance between the latent positions. Consequently, the latent positions of commonly reacted ROIs are near the center, allowing most subjects to exhibit similar reactions. This property facilitates the extraction of commonly reacted ROIs for each group  $h$  compared to group  $g$ , enabling the construction of the group representative matrix  $\mathbf{X}_{h|g,h}$ .

## 2.7 Advantages of Self-Attn and LSIRM in capturing fMRI intricacies

Self-attention (Vaswani et al., 2017) evaluates the relationships between elements in the input sequence, and assigns weights to emphasize important information from these relationships. This mechanism enables the model to consider all interactions between ROIs in a single computation. Furthermore, because it processes

all ROIs simultaneously, it is useful for capturing long-range dependencies which helps capture global information. Multi-head self-attention enables learning in parallel interactions between ROIs, allowing the model to capture diverse relationships and patterns in greater depth.

In contrast, eXtreme Gradient Boosting (XGBoost) (Chen and Guestrin, 2016) is an ensemble method based on decision trees that is focused on identifying classification rules between each data element and the target, instead of considering the relationships among elements. In other words, XGBoost classifies based on specific ROI values. However, when data patterns are unclear, such as in rs-fMRI, it is also important to understand the overall network of ROI relationships. XGBoost can have difficulty capturing these complex underlying interactions (Mørup et al., 2010; Martínez-Riño et al., 2023).

On the other hand, Multi-Layer Perceptron (MLP) (Qiu et al., 2020) primarily analyzes global features instead of local patterns, leading to the loss of information about the interactions between specific ROI features. MLP processes all input data features simultaneously, and as data passes through each layer, it gets transformed using non-linear activation functions. However, this process does not learn the interactions between the inputs, since all inputs are equally processed throughout the network. While MLP can easily capture global characteristics, they have limitations in training complex interactions between ROIs (Lai and Zhang, 2023).

TABLE 2 Performance comparison of classification accuracy using 10-fold cross-validation between Self-Attn, two previous studies (Liu et al., 2020; Wee et al., 2019), and three models [XGBoost (Chen and Guestrin, 2016; Qiu et al., 2020), and CNN (Zunair et al., 2020)].

Method		AD/MCI	EMCI/AD	LMCI/AD	EMCI/LMCI
Liu et al.		0.8890	-	-	-
Wee et al.		-	0.7920	0.6520	0.6090
XGBoost	Pearson	0.7949	0.8206	0.8511	0.7802
	Fisher	0.7460	0.6254	0.6353	0.6776
MLP	Pearson	0.8132	0.7600	0.7809	0.7457
	Fisher	0.7835	0.7533	0.7900	0.7605
	Linear	0.7461	0.7333	0.7146	0.7452
	Stochastic	0.7659	0.7600	0.6427	0.7381
	Topological	0.7819	0.7067	0.6518	0.6838
CNN	Pearson	0.8654	0.7733	0.7355	0.7667
	Fisher	0.8648	0.7390	0.7809	0.7267
	Linear	0.7747	0.7586	0.7891	0.7600
	Stochastic	0.8176	0.7657	0.7246	0.8200
	Topological	0.8192	0.7600	0.7155	0.7600
Self-Attn	Pearson	0.8659	0.8067	0.8809	0.8071
	Fisher	0.8813	0.8133	0.8718	0.8152
	Linear	0.9022	0.8467	0.8627	0.8624
	Stochastic	0.9033	<b>0.8867</b>	0.8900	<b>0.8971</b>
	Topological	<b>0.9104</b>	0.8733	<b>0.9173</b>	0.8695

The bold values indicate the models with the highest accuracy in each comparison, which were selected as the optimal models for subsequent analysis.

On the other hand, Convolutional Neural Networks (CNN) (Zunair et al., 2020) extract features by sliding small filters over the data. However, it is difficult for CNN to capture global relationships within the data through a single convolution operation. To address this, multiple layers are required, which necessitates more data and increases the training time needed to learn global relationships. While CNN is effective at learning local interactions between adjacent ROIs, they have limitations in capturing the global interactions between ROIs (Wang et al., 2018).

For these reasons, Self-Attn is particularly well-suited for learning interactions between ROIs, as it can simultaneously capture local information and global patterns. This capability makes Self-Attn effective at capturing the complex signals within rs-fMRI data. As shown in Table 2, this results in superior performance compared to other models.

Our approach uses LSIRM to visualize the interactions among ROIs, reflected in the group representative matrix derived from the attention distribution matrix. LSIRM captures these interactions by estimating the latent positions of ROIs, providing an intuitive interpretation of their relationships. Importantly, the group representative matrix is not an adjacency matrix, but a subject-by-ROI matrix, where the goal is to identify group-level features rather than individual-specific ones. Since LSIRM considers all subject-ROI interactions, it identifies common ROI features that show consistent reactions across all subjects. Therefore, by employing LSIRM, our approach can achieve a better understanding of group effects and provide a more intuitive way to interpret ROI interactions.

## 3 Results

In this study, we applied our analysis framework to identify the specific ROIs that differentiate between four disease comparisons: AD vs. MCI, AD vs. EMCI, AD vs. LMCI, and EMCI vs. LMCI. We utilized rs-fMRI data collected from AD, EMCI, MCI, and LMCI from the ADNI dataset.

### 3.1 Step 1: functional connectivity networks of each group

First, we constructed an FCN among brain regions based on their rs-fMRI Blood-Oxygen-Level-Dependent (BOLD) signals. We used AAL-116 templates to extract 116 rs-fMRI BOLD signals, representing different brain regions. Supplementary Table S1 in supporting information contains detailed information about the AAL-116 templates. Due to the high-dimensional and correlation structure of fMRI data, we implemented dimension reduction such as PCA, t-SNE, and UMAP over time to embed the high-dimensional correlated structure dataset into low two-dimensional space (Figure 2B). We empirically searched for the optimal combination of UMAP hyperparameters to enhance prediction accuracy. Based on this, we chose the number of neighbors to be 15 and set the minimum distance to 0.1, consistent with hyperparameters used in previous studies for visualizing high-dimensional data, such as genomics (Becht et al., 2019) and single-cell data (Diaz-Papkovich et al., 2021).

To mitigate the subjectivity in determining relevance among ROIs, we employed Mapper (Chazal and Michel, 2017), a partial clustering method, to identify significant connections between ROIs (represented as Figure 2C). ROIs assigned to the same cluster were considered connected. Subsequently, we generated FCNs for each set of embedded ROIs from different dimension reduction methods (represented as Figure 2D). These FCNs captured relationships and connectivity patterns within the high-dimensional correlated fMRI data, representing the data as a connectivity network.

Figure 3 and Supplementary Figures S1–S3 show each subject's rs-fMRI BOLD signals and two types of FCNs: correlation- and dimension-reduction-based FCNs obtained from dimension reduction methods corresponding to AD, MCI, EMCI, and LMCI. Through correlation- and dimension-reduction-based approaches, those different perspectives enabled a comprehensive understanding of their structural characteristics. Note that these FCNs are input for Self-Attn to classify between two diseases. One can apply other dimensionality reduction methods to obtain the embedded ROIs; however, one can select the optimal method based on the prediction accuracy provided by the Self-Attn, where the input is FCNs.

### 3.2 Step 2: attention distribution matrix from self-attention deep learning model

We focus on identifying which ROIs are important features for distinguishing groups in terms of ROIs' interactions with others. In the previous step, we applied dimension reduction techniques such as PCA, t-SNE, and UMAP to the original time-series of the ROIs to construct FCNs. Additionally, we compared these findings with two correlation matrices calculated from the original time-series data of the ROIs: (1) Pearson's  $r$  and (2) Fisher's  $z$ , both of which capture the associations among ROIs. With this data, we employed Self-Attn, using FCNs or correlation matrices as input, with the target being a binary indicator of group membership.

In the Transformer architecture,  $d_k$  is typically defined by dividing the total input dimension ( $d_{model}$ ) by the number of attention heads ( $h$ ) (Vaswani et al., 2017). However, the roles of the various attention heads vary, and not all heads contribute equally to model performance. Therefore, focusing on the most important heads can maintain model performance while reducing the risk of overfitting (Voita et al., 2019; Michel et al., 2019). To optimize classification accuracy while avoiding overfitting and underfitting, we conducted a parameter search for  $h = \{16, 32, 64, 128\}$ . By modifying  $h$ , we implicitly explored the corresponding  $d_k$  as well, specifically investigating  $d_k = \{2, 4, 8\}$  where setting  $d_k$  to a small value is practically effective specifically in self-attention (Tay et al., 2022). We validated the model through a 10-fold cross-validation, and showed that the configuration with  $h = 128$  and its corresponding  $d_k = 2$  achieved consistently robust classification performance. We provide details for parameter search in Supplementary Figure S4. We applied the model to 116 ROIs using the following settings: the batch size of 8, a dropout rate of 0.9, Adam optimizer (Kingma and Ba, 2014), the learning rate of 0.01, and cross-entropy loss.



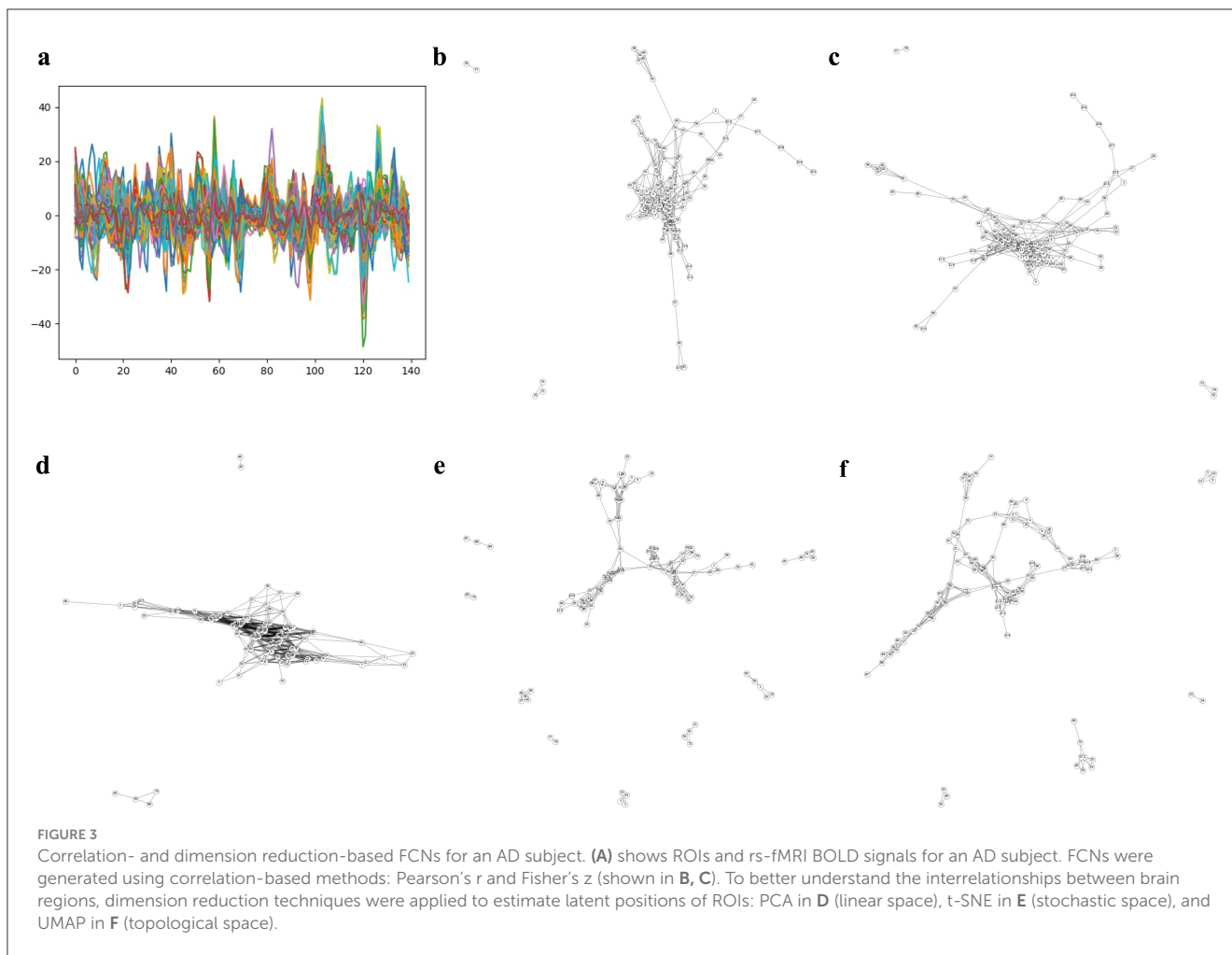


Table 2 shows the performance of Self-Attn. We compared the classification performance against recent studies (Liu et al., 2020; Wee et al., 2019) and baseline models [XGBoost (Chen and Guestrin, 2016), MLP (Qiu et al., 2020), and CNN (Zunair et al., 2020)]. Our method outperforms all other approaches across the disease group comparisons. Notably, the stochastic and topological-based FCN, which captures hidden connectivity among ROIs, achieved the highest accuracy.

Using Self-Attn, we obtained the attention distribution (Figure 2F) by each  $i$ th subject from group  $g$ . These attention distributions, denoted as  $\mathbf{A}_{(q,r)}^{(i)} \in \mathbb{R}^{116 \times 116}$ , where  $i = 1, \dots, N_g$ . Here,  $N_g$  indicates the number of subjects from each disease group  $g = \{AD, MCI, EMCI, LMCI\}$ , where  $N_{AD} = 57$ ,  $N_{MCI} = 78$ ,  $N_{EMCI} = 93$ , and  $N_{LMCI} = 53$ . These attention distributions  $\mathbf{A}_{(q,r)}^{(i)}$  reveal the features that the model focused on when classifying subjects in each disease group against the other comparison groups.

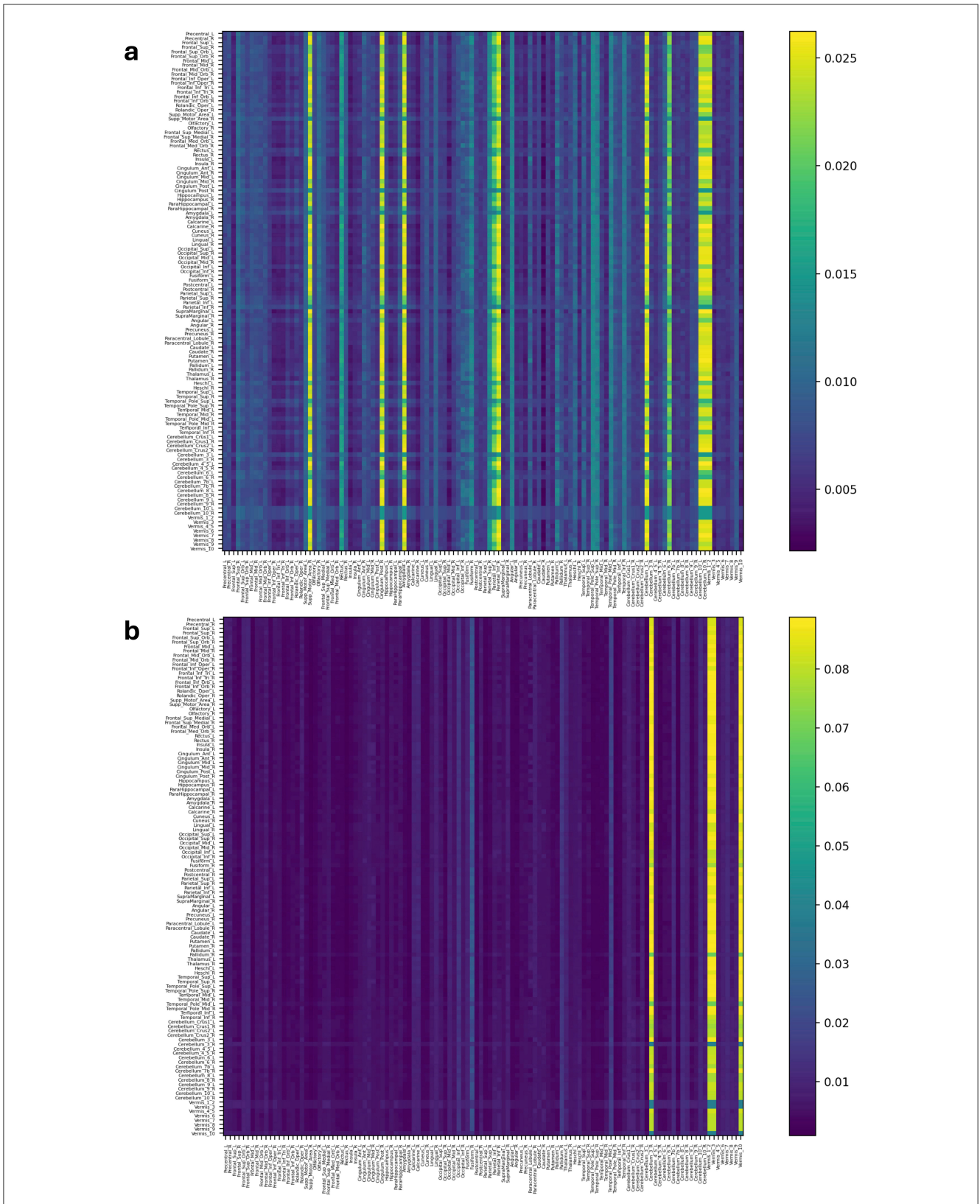
We considered this attention distribution  $\mathbf{A}_{(q,r)}^{(i)}$  as the matrix  $\mathbf{Y}_{g|h}^{(i)}$  for  $g \neq h$  and  $g, h = 1, \dots, G$ , where each row and column corresponds to ROIs, and the values indicate the significance of each ROI's contribution to the classify the subject  $i$  in group  $g$  against group  $h$  (Vig, 2019) (Figure 2G). Although the resting-state data shows low signal levels, the classification accuracy of

90% demonstrates that the attention matrices effectively distinguish between the two disease group comparisons.

Figures 4, 5 and Supplementary Figures S5, S6 represent attention matrices for four randomly selected subjects from the AD and MCI groups, respectively. These matrices are the outcomes of Self-Attn employed for AD and MCI classification, utilizing FCNs derived from topological dimension reduction techniques as inputs. ROIs with Higher attention values suggest their significance in classifying subjects. For instance, in Figure 4A, Putamen\_R shows a higher attention value, which is a consistent result in Putamen's volume in the AD group (de Jong et al., 2008). Furthermore, Angular and Paracentral areas have high values indicating that Self-Attn focuses on these regions to assign this subject to the AD category over MCI. Similarly, in Figure 4B, Supplementary Figure S5A, Caudate\_R, Caudate\_L, ParaHippocampal\_L, Cerebellum, Cingulum\_Ant\_L, and Cingulum\_Ant\_R have high values which also show significant ROIs marker in AD group (Kesslak et al., 1991; Bobinski et al., 1999; He et al., 2007; Catheline et al., 2010).

On the other hand, Figure 5A has a high value in Supp\_Motor\_Area\_R and Cingulum\_Post\_R (Lin et al., 2014) which ROIs linked to motor function and exercise (Schmahmann et al., 2001; Aggarwal et al., 2006; Bai et al., 2011).





**FIGURE 5** Attention distribution matrix of two subjects within the MCI group (A, B), generated by Self-Attn designed for AD and MCI classification, utilizing topological-based FCNs as input.

Similarly, in Figure 5B and Supplementary Figures S5A–S6B, Cerebellum, Vermis (van de Mortel et al., 2021), Thalamus\_R (Li et al., 2013; Cai et al., 2015), Hippocampus\_L, and ParaHippocampal\_R (Hämäläinen et al., 2007) emerge as significant ROIs in the MCI group. These patterns in attention distribution explain why Self-Attn places greater emphasis on classifying this subject within the MCI category rather than AD.

Based on these self-attention distributions, the classification accuracy outlined in Table 2 highlights the effectiveness of our well-trained Self-Attn. Nonetheless, individually interpreting each subject within each group can be time-consuming, and there may be potential noise originating from individual differences within each attention distribution matrix.

To identify the representative features of ROIs connections that differentiate between two groups (e.g.,  $g$  and  $h$ ), we constructed a group representative matrix  $\mathbf{X}_{g|g,h} \in \mathbb{R}^{N_g \times 116}$  from each attention distribution matrix  $\mathbf{Y}_{g|g,h}^{(i)}$  for  $i = 1, \dots, N_g$ , where each row represents a subject and each column represents an ROI (Figure 2H). The values in the  $g$ th group representative matrix  $\mathbf{X}_{g|g,h}$  represent the CV, where each column corresponds to the CV of each ROI's response to other ROIs within individual subjects' attention distributions, aggregated across all subjects. A high value for a certain ROI in  $\mathbf{X}_{g|g,h}$  indicates that this ROI shows a unique pattern in the corresponding individual attention distribution matrix, contributing to classifying that individual into a specific group. Additionally, we compiled a list of the ROIs that rank in the top 25% based on both high averaged CV and high mean values in the attention distribution matrices  $\mathbf{Y}_{g|g,h}^{(i)}$  from group  $g$ . As shown in Figure 2H, the CV is calculated for each ROI within each subject, yielding a set of CV values for ROIs across all subjects. By averaging these values, we obtain the mean CV for each ROI. The top 25% ROIs with high mean values indicate frequent interactions with other ROIs, while those with high CV values suggest non-uniform signal patterns across subjects. Table 3 displays the top 25% unique ROIs for each group, compared between two disease groups—AD and MCI. Additional details on the top 25% ROIs for other group comparisons are provided in Supplementary Tables S2–S4.

Specific ROIs can be selected based on expert knowledge such as known disease mechanisms. This approach can facilitate deeper and more focused inferences when comparing ROIs between studied diseases. However, our fusion analytical framework provides a robust solution applicable to general resting-state fMRI datasets without requiring expert knowledge. Therefore, we selected ROIs based on statistical measures, such as the mean and the coefficient of variation in the attention distribution matrices. Without applying weights to the mean and coefficient of variation, we select the first quartile, representing the top 25% in both statistical measures. Using specific percentiles is a common approach in statistical analysis to highlight significant patterns and features within a dataset (Hastie, 2009; Silverman, 2018). Based on this approach, several previous studies have selected the top 25% features to identify distinctive patterns within data (Subramanian et al., 2005; Love et al., 2014). Similarly, we selected the top 25% of ROIs ranked by their mean and coefficient of variation.

TABLE 3 Top 25% ROIs that show differences between disease group of AD and MCI.

Top	AD	MCI
Top-1	Postcentral_L	Cingulum_Mid_L
Top-2	Postcentral_R	Postcentral_L
Top-3	Temporal_Inf_L	Fusiform_R
Top-4	Supp_Motor_Area_R	Precentral_R
Top-5	Fusiform_L	Pallidum_L
Top-6	Cingulum_Mid_L	Temporal_Inf_L
Top-7	Fusiform_R	Fusiform_L
Top-8	Cerebellum_8_R	Supp_Motor_Area_R
Top-9	Cerebellum_6_L	Insula_L
Top-10	Putamen_L	Cerebellum_6_R
Top-11	Cerebellum_8_L	Cerebellum_4_5_R
Top-12	Precentral_R	Postcentral_R
Top-13	Thalamus_L	Cerebellum_4_5_L
Top-14	Temporal_Mid_R	Cerebellum_6_L
Top-15	Cerebellum_4_5_R	Putamen_R
Top-16	Insula_R	Cerebellum_8_R
Top-17	Rolandic_Oper_R	Cerebellum_Crus2_R
Top-18	Precentral_L	SupraMarginal_R
Top-19	Temporal_Inf_R	Rolandic_Oper_R
Top-20	Insula_L	Hippocampus_R
Top-21	Putamen_R	Cingulum_Mid_R
Top-22	Cerebellum_7b_R	Pallidum_R
Top-23	Temporal_Mid_L	Thalamus_L
Top-24	Hippocampus_R	Putamen_L
Top-25	Cerebellum_4_5_L	Vermis_4_5
Top-26	Cerebellum_10_R	Paracentral_Lobule_L
Top-27	Cingulum_Mid_R	Vermis_8
Top-28	Pallidum_L	Precentral_L
Top-29	Cerebellum_7b_L	Supp_Motor_Area_L

### 3.3 Step 3: group representative features using the Latent Space Item-Response Model

In Step 2, we obtain the group representative matrices  $\mathbf{X}_{h|g,h}$  for  $g \neq h$ . To capture the group representative ROIs features that commonly reacted among subjects as shown in Figure 2I, we applied LSIRM to each group representative matrix  $\mathbf{X}_{h|g,h} \in \mathbb{R}^{N_h \times R}$ . We estimated the latent positions of ROIs  $\mathbf{V} = \{\mathbf{v}_i\}$ , for  $i = 1, \dots, 116$  using Markov chain Monte Carlo (MCMC). The MCMC ran 55,000 iterations, and the first 5,000 iterations were discarded as burn-in processes. Then, from the remaining 50,000 iterations, we collected 10,000 samples with a thinning interval

of 5. We used two-dimensional Euclidean space to estimate the latent positions of ROIs. Additionally, we set 0.005 for  $\beta$  jumping rule, 0.005 for  $\theta$  jumping rule, and 0.005 for  $w_j$  and 0.003  $z_i$  jumping rules. Here, we fixed the prior  $\beta$  to follow  $N(0, 1)$ . We set  $a_\sigma = b_\sigma = 0.001$ . LSIRM takes each matrix  $X_h$  as input and provides the  $O_h$  matrix as output after the Procrustes-matching within the model. Since we calculated topics' distance on the 2-dimensional Euclidean space,  $O_h$  is of dimension  $116 \times 2$ . To overcome the identifiable issues from the invariance property, we applied oblimin rotation to the estimated topic position matrix  $O_{h\%}^*$  using the R package GPARotation (<https://cran.r-project.org/web/packages/GPARotation/index.html>).

Based on their estimated latent positions, we successfully identified ROIs that exhibited common reactions to groups. Figure 6 and Supplementary Figures S7–S9 exclusively display the latent positions of the top 25% ROIs from each group. As depicted in Figure 6, the latent positions of ROIs are visualized in Euclidean space. By comparing the latent positions of ROIs from the two groups, we identified the ROIs exhibiting distinct patterns. In this representation, red-colored numbers signify the Top 25% ROIs from the AD group, while blue-colored numbers correspond to the top 25% ROIs from the MCI group. Notably, ROIs positioned closer to the origin in the latent space suggest a heightened likelihood of shared interactions among subjects within the same group. For instance, the latent position of 98 ROI from the AD and MCI groups is both located close to the origin and marked in green, indicating their significant roles in both AD and MCI. On the other hand, the latent positions of 101 and 103 ROIs are exclusively part of the AD group's top 25%. Moreover, other numbers highlighted in orange, indicate that only one ROI group possesses latent positions near the origin. These ROIs can be interpreted as significant features that exhibit meaningful reactions exclusively compared to the other group.

In our analysis framework, therefore, we specifically focus on the ROIs that meet two criteria: being ranked in the top 25% listed in Step 2 and having latent positions near the origin. These criteria indicate that these ROIs exhibit distinct patterns among subjects and can be considered representative features of each group. This selection process helps us identify the main features that are prominent and generalized well across the groups. We visually highlight these selected ROIs on the summary FCN from each group. As shown in Figure 2J, the summary FCN for each group is obtained by averaging the connectivity of each node across all subjects within the group. A threshold of 0.2, representing the top 5% connectivity ratios among the subjects' FCNs in each group, is then applied to define the connections. We have added histogram figures showing each group's connectivity ratio in Supplementary Figure S10.

Figure 7 and Supplementary Figures S11–S13 show the differences in disease network between the two groups. The blue color indicates meaningful regions that show distinctive patterns from the attention distribution matrix compared to the other disease groups. The Orange color, on the other hand, indicates ROIs that were selected before analysis to be meaningful in both disease groups but were shown to only be meaningful in one group post-analysis. Finally, green indicates regions that were meaningful in both disease groups.

### 3.4 Interpretation of summary FCN from each group

Figure 7 and Supplementary Figure S11–S13 show the differences in disease network between the two groups. ROIs colored in blue indicate their selection as the top 25% group from the attention distribution matrix in one group, yet they do not appear as prominently significant in another group. Orange-colored ROIs indicate that they are meaningful only in one group, as revealed by the comparison between latent positions from each group. Finally, Green colored ROIs indicate that they were found to be meaningful in both disease groups. Utilizing the property of latent positions estimated from LSIRM (Jeon Y. et al., 2021), we managed to decode the structural connections among ROIs and identify ROIs that exhibited consistent significance across all subjects within each disease group. To see the overall connectivity, we merged the outcomes of LSIRM with FCNs, assigning colors to the significant ROIs and their interconnectedness with other ROIs in FCNs. The higher saturation colors indicate meaningful ROI features from LSIRM, and the same color with a lower brightness level reveals a direct connection from the meaningful ROI feature. We can regard these connections as a cluster.

#### 3.4.1 AD/MCI

Figure 7 illustrates the differences in the disease network between the AD group and MCI group, and Table 4 summarizes the meaningful ROIs. According to Figure 7A, Cluster A is comprised of Cerebellum\_7\_L (101) and Cerebellum\_8\_L (103). These two regions did not show activity in MCI, and the majority of regions that reacted in AD were connected to the Cerebellum regions. This distinction becomes evident as the AD group exhibits a diminished gray matter volume in the cerebellar anterior lobe in contrast to the non-AD group, as observed in prior research (Reiman et al., 2012).

Cluster C (orange-colored cluster) with Hippocampus\_R (38) and cluster D (orange-colored cluster) with Cingulum\_Mid\_L (33) and Cingulum\_Mid\_R (34) of Figure 7A show the cluster of regions and their direct connectivity that were more reactive in AD compared to MCI. Hippocampus\_R (38) of cluster C showed greater reactivity in AD when compared to MCI and Hippocampus (37, 38), ParaHippo (39, 40), Putamen (73, 74), Pallidum (75, 76), and Amygdala (41, 42) are densely populated in this area. We discovered that the Hippocampus (37, 38) plays an important role in expressing AD characteristics (Figure 7A Cluster C). Many studies have shown that having Hippocampus (37, 38) dysfunction affects memory (Spaniol et al., 2009; Small et al., 2011; Delbeck et al., 2003). Through our methodology, we also identified associations between the Hippocampus (37, 38), Putamen (73, 74), and ParaHippo (39, 40). These regions have been previously associated with cognitive impairment in Alzheimer's disease (Kesslak et al., 1991; Bobinski et al., 1999; de Jong et al., 2008).

Cluster F of Figure 7B shows ROIs, Postcentral\_L (57) and Postcentral\_R (58), that were more reactive in MCI compared to AD. This Postcentral (57, 58) is directly connected to Cingulum\_Mid (33, 34), Precentral (1, 2), and Paracentral\_Lob\_L



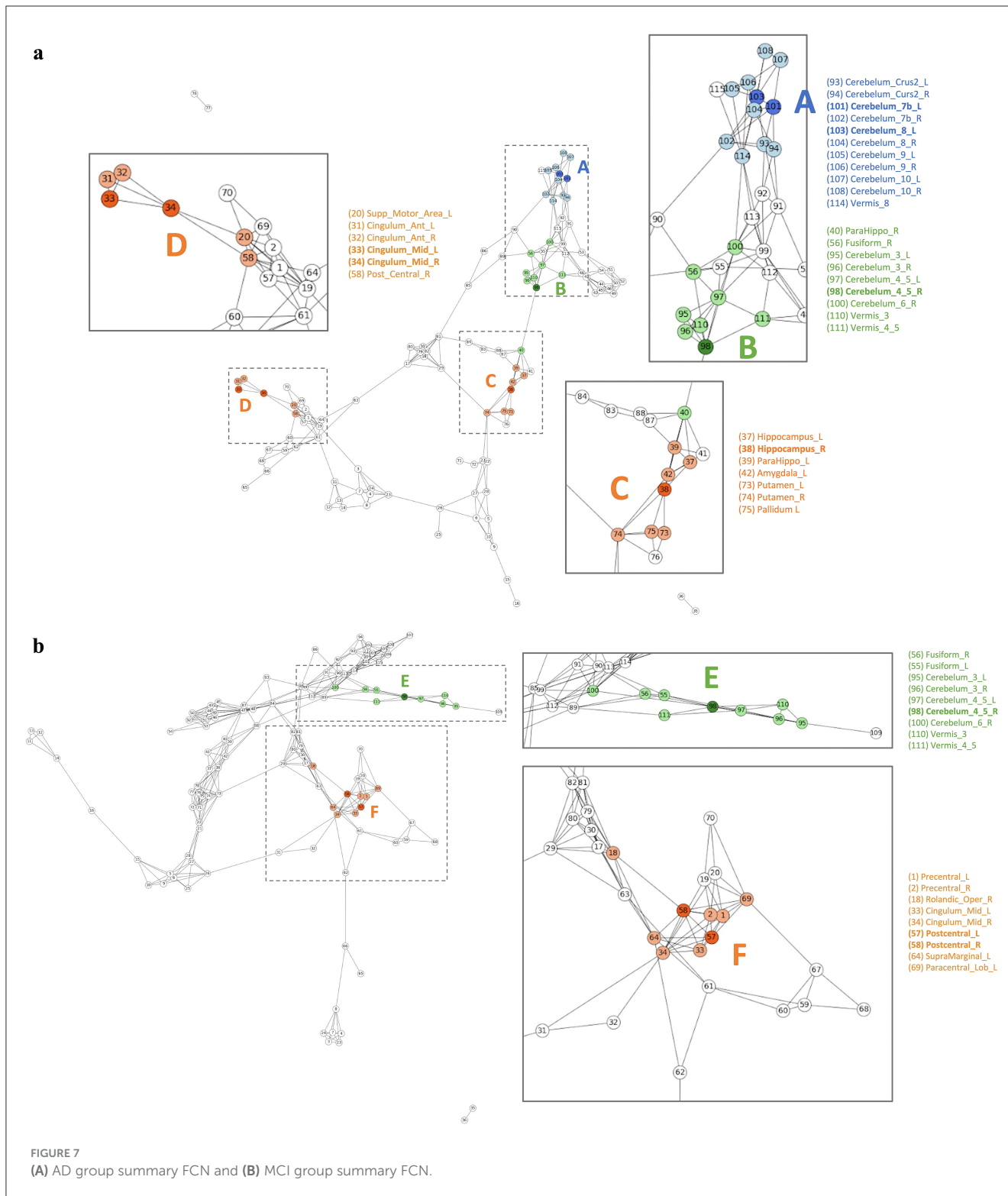


Figure S11A, which are Hippocampus\_L (37), Lingual\_R (48), Cerebellum\_4\_5\_L (97), were found to be meaningful regions not in EMCI but only in AD. The hippocampus (37, 38) in both hemispheres is directly connected. These regions are also directly connected to ParaHippo (39, 40), Putamen (73, 74), Pallidum (75, 76), and Amygdala (41, 42) and the results are similar to the results

described in Section 3.4.1. Fusiform\_R (56) and Cerebellum\_8\_L (103) are directly connected to Hippocampus\_R (38), and this is similar to cluster A and B of Supplementary Figure S11A. Hippocampus\_L (37) and Lingual\_R (48) are not only directly connected to Lingual\_L (47), but also to Calcarine (43, 44), Cuneus (45, 46), Fusiform\_L (56) and Cerebellum\_6\_L (99).

**TABLE 4** Meaningful ROIs in the comparison between the AD group and the MCI group.

ROIs	AD	MCI
Cingulum_Mid_L	▲	▽
Cingulum_Mid_R	▲	▽
Hippocampus_R	▲	▽
Postcentral_L	▽	▲
Postcentral_R	▽	▲
Cerebellum_4_5_R	▲	▲
Cerebellum_7_L	▲	–
Cerebellum_8_L	▲	–

▲ indicates regions that are active in the group, while ▽ represents regions that are less active than the other group. – indicates ROIs that are not selected as active in the group.

**TABLE 5** Meaningful ROIs in the comparison between the AD group and the EMCI group.

ROIs	AD	EMCI
Rolandic_Oper_R	▲	▽
Cingulum_Mid_L	▽	▲
Cingulum_Mid_R	▽	▲
Hippocampus_L	▲	–
Hippocampus_R	▲	▽
Lingual_R	▲	–
Fusiform_L	▲	▲
Fusiform_R	▲	▲
Pallidum_L	▽	▲
Cerebellum_Crus2_L	▽	▲
Cerebellum_4_5_L	▲	–
Cerebellum_9_R	–	▲

▲ indicates regions that are active in the group, while ▽ represents regions that are less active than the other group. – indicates ROIs that are not selected as active in the group.

Cluster D and E of [Supplementary Figure S11A](#) were more active in AD relative to EMCI and included the Hippocampus\_R (38), Rolandic\_Oper\_R (18) regions. We can see that Rolandic\_Oper\_R (18) is directly connected to Putamen (73, 74), Pallidum (75, 76) and Heschl\_L (79). Cluster F of [Supplementary Figure S11B](#) was active in EMCI but not AD, and Cerebellum\_9\_R (106) was analyzed. This region was adjacent to Cerebellum\_Crus2\_R (94), Cerebellum\_7b\_R (102), and Cerebellum\_9\_L (105). Cluster H, I, and J of [Supplementary Figure S11B](#) are regions that were more active in EMCI relative to AD, and regions Cingulum\_Mid\_L (33), Cingulum\_Mid\_R (34), Pallidum\_L (75) and Cerebellum\_Crus2\_L (93) were analyzed. Cingulum\_Mid is directly connected to Precentral (1, 2), Supp\_Motor (19, 20), Postcentral (57, 58) and Supramarginal (63, 64). These connections have recently been examined in planning and cognitive control processing ([Domic-Siede et al., 2021](#); [Cavanagh and Frank, 2014](#)).

Cluster C of [Supplementary Figure S11A](#) and Cluster G of [Supplementary Figure S11B](#) are regions that were active in

**TABLE 6** Meaningful ROIs in the comparison between the AD group and the LMCI group.

ROIs	AD	LMCI
Rolandic_Oper_R	–	▲
Putamen_L	▽	▲
Temporal_Mid_R	▲	–
Cerebellum_4_5_R	▽	▲
Cerebellum_6_L	▲	▽
Vermis_8	▽	▲

▲ indicates regions that are active in the group, while ▽ represents regions that are less active than the other group. – indicates ROIs that are not selected as active in the group.

both AD and EMCI, and correspond to Fusiform\_L (55) and Fusiform\_R (56). Fusiform (55, 56) is directly connected to Hippocampus (37, 38) and ParaHippo (39, 40). Likewise, in previous studies ([Apostolova et al., 2006](#); [Li et al., 2013](#); [Zhu et al., 2009](#)), there are connections between Hippocampus (37, 38) and ParaHippo (39, 40), Putamen (73, 74), Pallidum (75, 76), and Amygdala (41, 42).

### 3.4.3 AD/LMCI

[Supplementary Figure S12](#) illustrates the differences in disease network between the AD group and the LMCI group, and [Table 6](#) summarizes the meaningful ROIs. Cluster A of [Supplementary Figure S12A](#) was active in AD but not LMCI, and Temporal\_Mid\_R(86) was analyzed. Cluster B of [Supplementary Figure S12A](#) reacted more in AD relative to LMCI, and Cerebellum\_6\_L (99) was analyzed. Not only is this region connected with multiple Cerebellum (91, 92, 100) areas, but it is also connected to Fusiform\_L (55), Lingual (47, 48), and multiple Vermis (112, 113, 114). Cluster C of [Supplementary Figure S12B](#) was active only in LMCI and not AD, and Rolandic\_Oper\_R (18) was analyzed. This region was connected with Heschl (79, 80), Insula (29, 30) and Temporal\_Sup (81, 82). Cluster D, E, and F of [Supplementary Figure S12B](#) are regions more active in LMCI relative to AD, and regions Putamen\_L (73), Cerebellum\_4\_5\_R (98) and Vermis\_8 (114) were analyzed. Putamen (73, 74) is connected to Olfactory (21, 22), Hippocampus (37, 38), Amygdala (41, 42), Pallidum (75, 76) and Thalamus\_L (77).

### 3.4.4 EMCI/LMCI

[Supplementary Figure S13](#) illustrates the differences in disease network between the EMCI group and the LMCI group, and [Table 7](#) summarizes the meaningful ROIs. Clusters A, B, and C of [Supplementary Figure S13A](#) were regions that were only active in EMCI, and Frontal\_Inf\_Orb\_R (16), Frontal\_Med\_Orb\_R (26), and Cerebellum\_3\_R (96) were analyzed. Cerebellum\_3\_R (96) is directly connected to Cerebellum\_3\_L (95) and Vermis\_3 (110). ROIs connected to Frontal\_Inf\_Orb (15, 16) and Frontal\_Inf\_Tri (13, 14), can be grouped as Frontal regions. We can also see that they are directly connected to Putamen\_R (74). The Frontal\_Med\_Orb\_R (26) is directly connected to Frontal\_Med\_Orb\_L (25), Rectus



TABLE 7 Meaningful ROIs in the comparison between the EMCI group and the LMCI group.

ROIs	EMCI	LMCI
Frontal_Inf_Orb_R	▲	–
Frontal_Mid_Orb_R	▲	–
Putamen_L	▲	▽
Pallidum_L	▲	▽
Temporal_Mid_R	▽	▲
Cerebellum_3_R	▲	–
Cerebellum_6_L	▽	▲

▲ indicates regions that are active in the group, while ▽ represents regions that are less active than the other group. – indicates ROIs that are not selected as active in the group.

(27, 28), and Frontal\_Sup\_Orb\_R (6). Cluster D represents a region that was more active in EMCI relative to LMCI, including Putamen\_L (73), and Pallidum\_L (75) regions. The ROIs that were primarily connected to these regions can largely be defined as Caudate (71, 72), Pallidum\_R (76), Thalamus (77, 78), Hippocampus (37, 38), and Insula (29, 30). Other regions include Rolandic\_Oper\_L (17), Amygdala\_R (42), Fusiform\_L (55), and Cerebellum\_8\_L (103). [Supplementary Figure S13B](#), on the other hand, shows the FCN extracted for LMCI, which shows no significant ROIs that were significantly active only in LMCI. There is, however, cluster E, that shows ROIs more active in LMCI relative to EMCI. This cluster is comprised of ROIs connected to Temporal\_Mid\_R(86) and Cerebellum\_6\_L (99). Temporal\_Mid\_L (85) and Temporal\_Inf\_R(90) are ROIs connected with Temporal\_Mid\_R (86). ROIs connected with Cerebellum\_6\_L (99) are largely Fusiform (55, 56), Lingual (47, 48), and multiple Vermis (112, 113). Other regions include Cerebellum\_Crus1\_L (91), Cerebellum\_4\_5\_L (97), and Cerebellum\_6\_R (100).

## 4 Discussion

Our study presents a fusion analytic framework that provides interpretable distinctions in ROIs' connections between cognitive impairment groups. We construct FCNs to visualize intricate patterns within high-dimensional fMRI data and apply Self-Attn to capture hidden connectivity differences across groups. Our framework identifies group distinctive ROIs by summarizing individual-level attention distributions into group-level features. Using LSIRM, we model interactions among ROIs and estimate latent positions, providing intuitive insights into consistent group-specific responses. Finally, we highlight these distinctive ROIs within summary FCNs, offering deeper insights into the unique characteristics of each condition. Our framework can be extended to other neurodegenerative diseases by applying the same analytical approach to explore ROI connections and patterns, enabling the identification of distinctive neural characteristics across various conditions.

Furthermore, our fusion analytic framework can also be extended to multi-class classification. We can apply self-attention

model to multi-class classification data. After obtaining each corresponding attention matrices, we estimate the latent positions of ROIs using LSIRM. As mentioned in Step 3, we can extract ROIs that display distinct patterns across different classes. In future work, we aim to expand this framework to multi-class classification, where distinct ROIs are not only differentiated between two groups but also across three or more disease groups.

Our fusion analytic framework has limitations in computing time. Since our framework involves three steps, fitting the data at each stage requires time, as each step depends on the previous one. Since the attention distribution output from Self-Attn influences the group representative matrix, which serves as the input for LSIRM where Self-Attn accuracy needs to be guaranteed. These sequential processes require users to input data and analyze results at each step, which can be more time-consuming than a single-stage model. Nonetheless, this multi-step approach is valuable as it enables users to check and interpret results at each stage.

Our methodology has also uncovered significant biological insights, which have been consistently validated across multiple studies. The four key features identified are (1) Hippocampus, (2) Cingulum, (3) Fusiform, and (4) Cerebellum.

When comparing AD with MCI, we found that the Hippocampus (37, 38) plays a crucial role in expressing AD characteristics ([Figure 7A](#), Cluster C). The hippocampus is among the earliest regions to exhibit structural alterations in AD, with significant atrophy observed, particularly in subregions ([Chételat et al., 2008](#)). Numerous studies have demonstrated that dysfunction in the hippocampus (37, 38) negatively affects memory ([Delbeuck et al., 2003](#); [Spaniol et al., 2009](#); [Small et al., 2011](#)). Research has shown that disruption in connectivity between the hippocampus and parahippocampal cortex is correlated with memory dysfunction and cognitive decline ([Sun et al., 2017](#)).

Using our approach, we also identified connections between the Hippocampus (37, 38) and the Putamen (73, 74), as well as the Parahippocampal regions (39, 40), which have been associated with cognitive impairment in Alzheimer's disease ([Kesslak et al., 1991](#); [Bobinski et al., 1999](#); [de Jong et al., 2008](#)). The parahippocampal cortex has shown altered connectivity with the hippocampus in MCI and AD patients, which is linked to reduced cognitive performance. This region is essential for maintaining declarative memory, and disruptions here are associated with early cognitive decline in AD ([Liu et al., 2016](#)). The putamen, along with the thalamus, experiences significant atrophy in AD patients, with this volume reduction correlating with cognitive decline. This suggests that deep gray matter structures play a role in AD's neurodegenerative processes, extending beyond the traditional focus on the hippocampus ([de Jong et al., 2008](#)).

When compared to EMCI, the biomarkers of AD become more clearly evident ([Supplementary Figure S11A](#)). These biomarkers are detected in the connections between the Hippocampus (37, 38) and the ParaHippocampal regions (39, 40), Putamen (73, 74), Pallidum (75, 76), and Amygdala (41, 42) ([Apostolova et al., 2006](#); [Zhu et al., 2009](#); [Li et al., 2013](#)). Studies indicate that specific changes in these ROIs are significant as the condition progresses from EMCI to AD. For instance, multimodal imaging techniques have shown atrophy in the hippocampus and amygdala during the

early stages of AD, which correlates strongly with cognitive decline (Eustache et al., 2016). Additionally, functional alterations in the connectivity between the hippocampus and parahippocampal regions have been associated with disrupted memory processing and the hallmark symptoms of AD (Wei et al., 2020). Moreover, the putamen and pallidum show progressive volume loss in patients transitioning from MCI to AD, which emphasizes their role in neurodegeneration (Yi et al., 2016). Collectively, these findings underline the interconnected nature of these regions in reflecting the underlying pathology of Alzheimer's disease as it progresses from EMCI.

In terms of the Cingulum, our method provides additional insights regarding the Cingulum\_Mid (33, 34) to MCI, which has traditionally been considered one of the most important ROIs in AD (Yoshiura et al., 2002; Fellgiebel et al., 2005; Kantarci et al., 2000; Scheff and Price, 2001; Catheline et al., 2010). The Cingulum\_Mid (33, 34) is responsible for processing motor and attention-related activities (Lin et al., 2014). However, our findings reveal that the Cingulum\_Mid (33, 34) is more prominent in MCI and EMCI than in AD (Figure 7A, Supplementary Figure S11B). The Cingulum\_Mid (33, 34) is connected to the Postcentral (57, 58), Precentral (1, 2), and Paracentral\_Lob (69, 70) regions, all of which are involved in motor processing. These four regions have been investigated using fluorodeoxyglucose (FDG) positron emission tomography (PET) as relevant indicators in MCI (Xu et al., 2016). Furthermore, distinct patterns of atrophy and reduced fractional anisotropy values in these ROIs have been noted in neuroimaging studies, emphasizing their importance in predicting the progression from MCI to AD (Choo et al., 2010).

In the network comparison between EMCI and AD, our model identified a connection between the Cingulum\_Mid (33, 34) and the Supramarginal gyrus (63, 64), which is associated with motor attention-related activity (Supplementary Figure S11B). These connections have recently been examined in relation to planning and cognitive control processes (Cavanagh and Frank, 2014; Domic-Siede et al., 2021). The Cingulum\_Mid and Supramarginal gyrus have been implicated in age-related motor performance changes, highlighting their role in cognitive and motor integration (Heuninckx et al., 2005). As a result, the Cingulum\_Mid (33, 34), which has previously been underexplored, should be considered relevant to both cognitive and motor functions in MCI and EMCI. These findings support its significance in understanding the progression of cognitive decline, especially as the interaction between these regions reflects the brain's adaptive changes in response to increasing cognitive demand.

On the other hand, in both AD and MCI, the Fusiform (55, 56) emerges as a key ROI in revealing the characteristics of each disease. The Fusiform (55, 56) is well known for its role in facial recognition processing (Kanwisher et al., 1997) and has recently been studied for its genetic and epigenetic links to AD (Ma et al., 2020). Our method identified the Fusiform (55, 56) as a significant ROI in both AD and EMCI (Supplementary Figure S11), showing that it is connected to cerebellar functions. This connection is consistently observed across all network results (Figure 7, Supplementary Figure S11–S13). Research also suggests that the cerebellum and fusiform gyrus exhibit strong interconnections, which play crucial roles in cognitive functions as MCI progresses toward AD (Yao et al., 2010).

Our research highlights the increasing significance of the cerebellum in the context of cognitive impairment. Traditionally, the cerebellum was associated primarily with motor control and coordination (Schmahmann et al., 2001; Aggarwal et al., 2006). However, recent studies have shown its involvement in cognitive processing and emotional modulation (Schmahmann, 1998; Schmahmann and Sherman, 1998; Whitwell et al., 2012) and emotional modulation (Schmahmann et al., 2007). Recently, the cerebellum has also been recognized as a useful biomarker in clinical AD subjects (Russo et al., 2018). According to our results, Cerebellum\_4\_5\_R (98) plays a significant role as an ROI in both AD and MCI (Figure 7). This aligns with prior studies showing decreased activity in this region in both AD and MCI compared to the control group (Wang et al., 2011). Additionally, structural MRI studies have explored changes in the cerebellum, finding that the posterior cerebellar lobes were significantly smaller in AD patients and correlated with poorer cognitive performance, suggesting that cerebellar atrophy may contribute to AD progression (Thomann et al., 2008).

## Data availability statement

The datasets analyzed during the current study are available in the Alzheimer's Disease Neuroimaging Initiative (ADNI) repository, <https://adni.loni.usc.edu/>. We have uploaded the core part of our method, to the following URL on Github: <https://github.com/jeon9677/summary-fcn>.

## Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent from the patients/ participants or patients/participants' legal guardian/next of kin was not required to participate in this study in accordance with the national legislation and the institutional requirements.

## Author contributions

YJ: Conceptualization, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. J-JK: Conceptualization, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. SY: Investigation, Writing – review & editing. JC: Data curation, Investigation, Software, Writing – review & editing. SH: Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Writing – original draft, Writing – review & editing.

## Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated

organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnins.2024.1402657/full#supplementary-material>

## References

- Aggarwal, N. T., Wilson, R. S., Beck, T. L., Bienias, J. L., and Bennett, D. A. (2006). Motor dysfunction in mild cognitive impairment and the risk of incident alzheimer disease. *Arch. Neurol.* 63, 1763–1769. doi: 10.1001/archneur.63.12.1763
- Apostolova, L. G., Dutton, R. A., Dinov, I. D., Hayashi, K. M., Toga, A. W., Cummings, J. L., et al. (2006). Conversion of mild cognitive impairment to alzheimer disease predicted by hippocampal atrophy maps. *Arch. Neurol.* 63, 693–699. doi: 10.1001/archneur.63.5.693
- Bai, F., Liao, W., Watson, D. R., Shi, Y., Yuan, Y., Cohen, A. D., et al. (2011). Mapping the altered patterns of cerebellar resting-state function in longitudinal amnesic mild cognitive impairment patients. *J. Alzheimer's Dis.* 23, 87–99. doi: 10.3233/JAD-2010-101533
- Becht, E., McInnes, L., Healy, J., Dutertre, C.-A., Kwok, I. W., Ng, L. G., et al. (2019). Dimensionality reduction for visualizing single-cell data using UMAP. *Nat. Biotechnol.* 37, 38–44. doi: 10.1038/nbt.4314
- Bobinski, M., De Leon, M., Wegiel, J., Desanti, S., Convit, A., Saint Louis, L., et al. (1999). The histological validation of post mortem magnetic resonance imaging-determined hippocampal volume in Alzheimer's disease. *Neuroscience* 95, 721–725. doi: 10.1016/S0306-4522(99)00476-5
- Cai, S., Huang, L., Zou, J., Jing, L., Zhai, B., Ji, G., et al. (2015). Changes in thalamic connectivity in the early and late stages of amnesic mild cognitive impairment: a resting-state functional magnetic resonance study from adni. *PLoS ONE* 10:e0115573. doi: 10.1371/journal.pone.0115573
- Catheline, G., Periot, O., Amirault, M., Braun, M., Dartigues, J.-F., Auriaud, S., et al. (2010). Distinctive alterations of the cingulum bundle during aging and Alzheimer's disease. *Neurobiol. Aging* 31, 1582–1592. doi: 10.1016/j.neurobiolaging.2008.08.012
- Cavanagh, J. F., and Frank, M. J. (2014). Frontal theta as a mechanism for cognitive control. *Trends Cogn. Sci.* 18, 414–421. doi: 10.1016/j.tics.2014.04.012
- Chazal, F., and Michel, B. (2017). An introduction to topological data analysis: fundamental and practical aspects for data scientists. *arXiv [Preprint]*. arXiv:1710.04019. doi: 10.48550/arXiv.1710.04019
- Chen, S., Chen, J., Jin, Q., and Hauptmann, A. (2018). "Class-aware self-attention for audio event recognition," in *Proceedings of the 2018 ACM on International Conference on Multimedia Retrieval* (New York, NY: ACM), 28–36.
- Chen, T., and Guestrin, C. (2016). "Xgboost: a scalable tree boosting system," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (New York, NY: ACM), 785–794.
- Chételat, G., Fouquet, M., Kalpouzos, G., Denghien, I., De La Sayette, V., Viader, F., et al. (2008). Three-dimensional surface mapping of hippocampal atrophy progression from MCI to ad and over normal aging as assessed using voxel-based morphometry. *Neuropsychologia* 46, 1721–1731. doi: 10.1016/j.neuropsychologia.2007.11.037
- Choo, I. H., Lee, D. Y., Oh, J. S., Lee, J. S., Lee, D. S., Song, I. C., et al. (2010). Posterior cingulate cortex atrophy and regional cingulum disruption in mild cognitive impairment and alzheimer's disease. *Neurobiol. Aging* 31, 772–779. doi: 10.1016/j.neurobiolaging.2008.06.015
- de Jong, L. W., van der Hiele, K., Veer, I. M., Houwing, J., Westendorp, R., Bollen, E., et al. (2008). Strongly reduced volumes of putamen and thalamus in Alzheimer's disease: an mri study. *Brain* 131, 3277–3285. doi: 10.1093/brain/awn278
- Delbeck, X., Van der Linden, M., and Collette, F. (2003). Alzheimer's disease as a disconnection syndrome? *Neuropsychol. Rev.* 13, 79–92. doi: 10.1023/A:1023832305702
- Diaz-Papkovich, A., Anderson-Trocme, L., and Gravel, S. (2021). A review of umap in population genetics. *J. Hum. Genet.* 66, 85–91. doi: 10.1038/s10388-020-00851-4
- Domic-Siede, M., Irani, M., Valdés, J., Perrone-Bertolotti, M., and Ossandón, T. (2021). Theta activity from frontopolar cortex, mid-cingulate cortex and anterior cingulate cortex shows different roles in cognitive planning performance. *Neuroimage* 226:117557. doi: 10.1016/j.neuroimage.2020.117557
- Du, Y., Xu, Y., Wang, X., Liu, L., and Ma, P. (2022). Eeg temporal-spatial transformer for person identification. *Sci. Rep.* 12:14378. doi: 10.1038/s41598-022-18502-3
- Dubois, J., and Adolphs, R. (2016). Building a science of individual differences from fMRI. *Trends Cogn. Sci.* 20, 425–443. doi: 10.1016/j.tics.2016.03.014
- Dunteman, G. H. (1989). *Principal Components Analysis*. Virginia Beach, VA: SAGE.
- Embretson, S. E., and Reise, S. P. (2013). *Item Response Theory*. Abingdon: Psychology Press.
- Eustache, P., Nemmi, F., Saint-Aubert, L., Pariente, J., and Péran, P. (2016). Multimodal magnetic resonance imaging in Alzheimer's disease patients at prodromal stage. *J. Alzheimer's Dis.* 50, 1035–1050. doi: 10.3233/JAD-150353
- Fellgiebel, A., Müller, M. J., Wille, P., Dellani, P. R., Scheurich, A., Schmidt, L. G., et al. (2005). Color-coded diffusion-tensor-imaging of posterior cingulate fiber tracts in mild cognitive impairment. *Neurobiol. Aging* 26, 1193–1198. doi: 10.1016/j.neurobiolaging.2004.11.006
- Gates, K. M., and Molenaar, P. C. (2012). Group search algorithm recovers effective connectivity maps for individuals in homogeneous and heterogeneous samples. *Neuroimage* 63, 310–319. doi: 10.1016/j.neuroimage.2012.06.026
- Hämäläinen, A., Pihlajamäki, M., Tanila, H., Hänninen, T., Niskanen, E., Tervo, S., et al. (2007). Increased fmri responses during encoding in mild cognitive impairment. *Neurobiol. Aging* 28, 1889–1903. doi: 10.1016/j.neurobiolaging.2006.08.008
- Hastie, T. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York, NY: Springer.
- He, Y., Wang, L., Zang, Y., Tian, L., Zhang, X., Li, K., et al. (2007). Regional coherence changes in the early stages of Alzheimer's disease: a combined structural and resting-state functional mri study. *Neuroimage* 35, 488–500. doi: 10.1016/j.neuroimage.2006.11.042
- Heuninckx, S., Wenderoth, N., Debaere, F., Peeters, R., and Swinnen, S. P. (2005). Neural basis of aging: the penetration of cognition into action control. *J. Neurosci.* 25, 6787–6796. doi: 10.1523/JNEUROSCI.1263-05.2005
- Jack, C. R., Bernstein, M. A., Fox, N. C., Thompson, P., Alexander, G., Harvey, D., et al. (2008). The alzheimer's disease neuroimaging initiative (adni): MRI methods. *J. Magnet. Resonance Imag.* 27, 685–691. doi: 10.1002/jmri.21049
- Jeon, M., Jin, I. H., Schweinberger, M., and Baugh, S. (2021). Mapping unobserved item-responder interactions: a latent space item response model with interaction map. *Psychometrika* 86, 378–403. doi: 10.1007/s11336-021-09762-5
- Jeon, Y., Chung, D., Park, J., and Jin, I. H. (2021). Network-based trajectory analysis of topic interaction map for text mining of COVID-19 biomedical literature. *arXiv [Preprint]*. arXiv:2106.07374. doi: 10.48550/arXiv.2106.07374
- Kantarci, K., Jack, C., Xu, Y., Campeau, N., O'Brien, P., Smith, G., et al. (2000). Regional metabolic patterns in mild cognitive impairment and alzheimer's disease: a 1h MRS study. *Neurology* 55, 210–217. doi: 10.1212/WNL.55.2.210

- Kanwisher, N., McDermott, J., and Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J. Neurosci.* 17, 4302–4311. doi: 10.1523/JNEUROSCI.17-11-04302.1997
- Kesslak, J. P., Nalcioglu, O., and Cotman, C. W. (1991). Quantification of magnetic resonance scans for hippocampal and parahippocampal atrophy in Alzheimer's disease. *Neurology* 41, 51–51. doi: 10.1212/WNL.41.1.51
- Kingma, D. P., and Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv [Preprint]*. arXiv:1412.6980. doi: 10.48550/arXiv.1412.6980
- Lai, H., and Zhang, P. (2023). Few-shot object detection with local feature enhancement and feature interrelation. *Electronics* 12:4036. doi: 10.3390/electronics12194036
- Lee, T., and Lee, H. (2020). Prediction of Alzheimer's disease using blood gene expression data. *Sci. Rep.* 10:3485. doi: 10.1038/s41598-020-60595-1
- Lei, B., Zhang, Y., Liu, D., Xu, Y., Yue, G., Cao, J., et al. (2022). Longitudinal study of early mild cognitive impairment via similarity-constrained group learning and self-attention based sbi-lstm. *Knowl.-Based Syst.* 254:109466. doi: 10.1016/j.knsys.2022.109466
- Li, X., Dvornik, N. C., Zhuang, J., Ventola, P., and Duncan, J. S. (2018). "Brain biomarker interpretation in asd using deep learning and fMRI," in *Medical Image Computing and Computer Assisted Intervention-MICCAI 2018: 21st International Conference, Granada, Spain* (Cham: Springer), 206–214.
- Li, Y., He, H., Dong, H., Feng, X., Xie, G., and Zhang, L. (2013). Discriminative analysis of early-stage alzheimer's disease and normal aging with automatic segmentation technique in subcortical gray matter structures: a multicenter in vivo mri volumetric and dti study. *Acta Radiol.* 54, 1191–1200. doi: 10.1177/0284185113492971
- Lin, Y.-C., Shih, Y.-C., Tseng, W.-Y. I., Chu, Y.-H., Wu, M.-T., Chen, T.-F., et al. (2014). Cingulum correlates of cognitive functions in patients with mild cognitive impairment and early alzheimer's disease: a diffusion spectrum imaging study. *Brain Topogr.* 27, 393–402. doi: 10.1007/s10548-013-0346-2
- Liu, J., Zhang, X., Yu, C., Duan, Y., Zhuo, J., Cui, Y., et al. (2016). Impaired parahippocampus connectivity in mild cognitive impairment and alzheimer's disease. *J. Alzheimer's Dis.* 49, 1051–1064. doi: 10.3233/JAD-150727
- Liu, M., Li, F., Yan, H., Wang, K., Ma, Y., Shen, L., et al. (2020). A multi-model deep convolutional neural network for automatic hippocampus segmentation and classification in alzheimer's disease. *Neuroimage* 208:116459. doi: 10.1016/j.neuroimage.2019.116459
- Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for rna-seq data with deseq2. *Genome Biol.* 15, 1–21. doi: 10.1186/s13059-014-0550-8
- Ma, D., Fetahu, I. S., Wang, M., Fang, R., Li, J., Liu, H., et al. (2020). The fusiform gyrus exhibits an epigenetic signature for alzheimer's disease. *Clin. Epigenetics* 12, 1–16. doi: 10.1186/s13148-020-00916-3
- Martínez-Riaño, D. E., González, F., and Gómez, F. (2023). H1 persistent features of the resting-state connectome in healthy subjects. *Netw. Neurosci.* 7, 234–253. doi: 10.1162/netn\_a\_00280
- McInnes, L., Healy, J., and Melville, J. (2018). Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv [Preprint]*. arXiv:1802.03426. doi: 10.21105/joss.00861
- Michel, P., Levy, O., and Neubig, G. (2019). Are sixteen heads really better than one? *Adv. Neural Inf. Process. Syst.* 32, 14014–14024. doi: 10.5555/3454287.3455544
- Morup, M., Madsen, K., Dogonowski, A., Siebner, H., and Hansen, L. K. (2010). Infinite relational modeling of functional connectivity in resting state fmri. *Adv. Neural Inf. Process. Syst.* 23, 1750–1758. Available at: <http://papers.nips.cc/paper/4057-infinite-relational-modeling-offunctional-connectivity-in-resting-state-fmri.pdf>
- Mueller, S. G., Weiner, M. W., Thal, L. J., Petersen, R. C., Jack, C. R., Jagust, W., et al. (2005). Ways toward an early diagnosis in Alzheimer's disease: the Alzheimer's disease neuroimaging initiative (ADNI). *Alzheimer's & Dementia* 1, 55–66. doi: 10.1016/j.jalz.2005.06.003
- Qiu, S., Joshi, P. S., Miller, M. I., Xue, C., Zhou, X., Karjadi, C., et al. (2020). Development and validation of an interpretable deep learning framework for Alzheimer's disease classification. *Brain* 143, 1920–1933. doi: 10.1093/brain/awaa137
- Reiman, E. M., Quiroz, Y. T., Fleisher, A. S., Chen, K., Velez-Pardo, C., Jimenez-Del-Rio, M., et al. (2012). Brain imaging and fluid biomarker analysis in young adults at genetic risk for autosomal dominant alzheimer's disease in the presenilin 1 e280a kindred: a case-control study. *Lancet Neurol.* 11, 1048–1056. doi: 10.1016/S1474-4422(12)70228-4
- Russo, G., Sardina, D., Alongi, P., Coppola, R., Puglisi, V., Stefano, A., et al. (2018). amyloid-pet analysis based on tissue probability maps. *Physica Medica* 56, 111–112. doi: 10.1016/j.ejmp.2018.04.089
- Santana, C. P., de Carvalho, E. A., Rodrigues, I. D., Bastos, G. S., de Souza, A. D., and de Brito, L. L. (2022). rs-fMRI and machine learning for asd diagnosis: a systematic review and meta-analysis. *Sci. Rep.* 12:6030. doi: 10.1038/s41598-022-09821-6
- Scheff, S. W., and Price, D. A. (2001). Alzheimer's disease-related synapse loss in the cingulate cortex. *J. Alzheimer's Dis.* 3, 495–505. doi: 10.3233/JAD-2001-3509
- Schmahmann, J. D. (1998). Dysmetria of thought: clinical consequences of cerebellar dysfunction on cognition and affect. *Trends Cogn. Sci.* 2, 362–371. doi: 10.1016/S1364-6613(98)01218-2
- Schmahmann, J. D., Anderson, C. M., Newton, N., and Ellis, R. D. (2001). The function of the cerebellum in cognition, affect and consciousness: empirical support for the embodied mind. *Consciousn. Emot.* 2, 273–309. doi: 10.1075/ce.2.2.06sch
- Schmahmann, J. D., and Sherman, J. C. (1998). The cerebellar cognitive affective syndrome. *Brain: J. Neurol.* 121, 561–579. doi: 10.1093/brain/121.4.561
- Schmahmann, J. D., Weilburg, J. B., and Sherman, J. C. (2007). The neuropsychiatry of the cerebellum—insights from the clinic. *Cerebellum* 6, 254–267. doi: 10.1080/14734220701490995
- Silverman, B. W. (2018). *Density Estimation for Statistics and Data Analysis*. London: Routledge.
- Small, S. A., Schobel, S. A., Buxton, R. B., Witter, M. P., and Barnes, C. A. (2011). A pathophysiological framework of hippocampal dysfunction in ageing and disease. *Nature Rev. Neurosci.* 12, 585–601. doi: 10.1038/nrn3085
- Smith, C. D., Chebrolu, H., Andersen, A. H., Powell, D. A., Lovell, M. A., Xiong, S., et al. (2010). White matter diffusion alterations in normal women at risk of Alzheimer's disease. *Neurobiol. Aging* 31, 1122–1131. doi: 10.1016/j.neurobiolaging.2008.08.006
- Smith, S. M., Miller, K. L., Salimi-Khorshidi, G., Webster, M., Beckmann, C. F., Nichols, T. E., et al. (2011). Network modelling methods for fMRI. *Neuroimage* 54, 875–891. doi: 10.1016/j.neuroimage.2010.08.063
- Spaniol, J., Davidson, P. S., Kim, A. S., Han, H., Moscovitch, M., and Grady, C. L. (2009). Event-related fmri studies of episodic encoding and retrieval: meta-analyses using activation likelihood estimation. *Neuropsychologia* 47, 1765–1779. doi: 10.1016/j.neuropsychologia.2009.02.028
- Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., et al. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Nat. Acad. Sci.* 102, 15545–15550. doi: 10.1073/pnas.0506580102
- Sun, Y., Wang, Y., Liu, Z., Siegel, J., and Sarma, S. (2020). "PointGrow: Autoregressively learned point cloud generation with self-attention," in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision* (Snowmass, CO: IEEE), 61–70.
- Sun, Y., Wang, Y., Lu, J., Liu, R., Schwarz, C. G., Zhao, H., et al. (2017). Disrupted functional connectivity between perirhinal and parahippocampal cortices with hippocampal subfields in patients with mild cognitive impairment and Alzheimer's disease. *Oncotarget* 8:99112. doi: 10.18632/oncotarget.17944
- Tay, Y., Dehghani, M., Bahri, D., and Metzler, D. (2022). *Efficient Transformers: A Survey*. New York, NY: ACM.
- Thomann, P. A., Schäfer, C., Seidl, U., Dos Santos, V., Essig, M., and Schröder, J. (2008). The cerebellum in mild cognitive impairment and alzheimer's disease—a structural mri study. *J. Psychiatr. Res.* 42, 1198–1202. doi: 10.1016/j.jpsychires.2007.12.002
- van de Mortel, L. A., Thomas, R. M., van Wingen, G. A., Initiative, A. D. N., et al. (2021). Grey matter loss at different stages of cognitive decline: a role for the thalamus in developing Alzheimer's disease. *J. Alzheimer's Dis.* 83, 705–720. doi: 10.3233/JAD-210173
- van der Maaten, L., and Hinton, G. (2008). Visualizing data using t-SNE. *J. Mach. Learn. Res.* 9, 2579–2605. Available at: <https://www.jmlr.org/papers/volume9/vandermaaten08a/vandermaaten08a.pdf>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., et al. (2017). "Attention is all you need," in *Advances in neural information processing systems* (Red Hook, NY: Curran Associates, Inc.), 5998–6008.
- Velickovic, P., Cucurull, G., Casanova, A., Romero, A., Lio, P., Bengio, Y., et al. (2017). Graph attention networks. *arXiv [Preprint]*. arXiv:1710.10903. doi: 10.48550/arXiv.1710.10903
- Vig, J. (2019). A multiscale visualization of attention in the transformer model. *arXiv [Preprint]*. arXiv:1906.05714. doi: 10.18653/v1/P19-3007
- Voita, E., Talbot, D., Moiseev, F., Sennrich, R., and Titov, I. (2019). Analyzing multi-head self-attention: specialized heads do the heavy lifting, the rest can be pruned. *arXiv [Preprint]*. arXiv:1905.09418. doi: 10.18653/v1/P19-1580
- Wang, S., Wang, G., Lv, H., Wu, R., Zhao, J., and Guo, W. (2016). Abnormal regional homogeneity as potential imaging biomarker for psychosis risk syndrome: a resting-state fmri study and support vector machine analysis. *Sci. Rep.* 6:27619. doi: 10.1038/srep27619
- Wang, X., Girshick, R., Gupta, A., and He, K. (2018). "Non-local neural networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (Salt Lake City, UT: IEEE), 7794–7803.

- Wang, Z., Yan, C., Zhao, C., Qi, Z., Zhou, W., Lu, J., et al. (2011). Spatial patterns of intrinsic brain activity in mild cognitive impairment and Alzheimer's disease: a resting-state functional MRI study. *Hum. Brain Mapp.* 32, 1720–1740. doi: 10.1002/hbm.21140
- Wee, C.-Y., Liu, C., Lee, A., Poh, J. S., Ji, H., Qiu, A., et al. (2019). Cortical graph neural network for AD and MCI diagnosis and transfer learning across populations. *NeuroImage: Clin.* 23:101929. doi: 10.1016/j.nicl.2019.101929
- Wei, Y., Huang, N., Liu, Y., Zhang, X., Wang, S., and Tang, X. (2020). Hippocampal and amygdalar morphological abnormalities in Alzheimer's disease based on three Chinese MRI datasets. *Curr. Alzheimer Res.* 17, 1221–1231. doi: 10.2174/1567205018666210218150223
- Whitwell, J. L., Weigand, S. D., Boeve, B. F., Senjem, M. L., Gunter, J. L., DeJesus-Hernandez, M., et al. (2012). Neuroimaging signatures of frontotemporal dementia genetics: C9orf72, tau, progranulin and sporadics. *Brain* 135, 794–806. doi: 10.1093/brain/aws001
- Xu, L., Wu, X., Li, R., Chen, K., Long, Z., Zhang, J., et al. (2016). Prediction of progressive mild cognitive impairment by multi-modal neuroimaging biomarkers. *J. Alzheimer's Dis.* 51, 1045–1056. doi: 10.3233/JAD-151010
- Yan, C.-G., Wang, X.-D., Zuo, X.-N., and Zang, Y.-F. (2016). DPABI: data processing & analysis for (resting-state) brain imaging. *Neuroinformatics* 14, 339–351. doi: 10.1007/s12021-016-9299-4
- Yao, Z., Zhang, Y., Lin, L., Zhou, Y., Xu, C., Jiang, T., et al. (2010). Abnormal cortical networks in mild cognitive impairment and Alzheimer's disease. *PLoS Comput. Biol.* 6:e1001006. doi: 10.1371/journal.pcbi.1001006
- Yeo, B. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., et al. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J. Neurophysiol.* 106, 1125–1165. doi: 10.1152/jn.00338.2011
- Yi, H.-A., Möller, C., Dieleman, N., Bouwman, F. H., Barkhof, F., Scheltens, P., et al. (2016). Relation between subcortical grey matter atrophy and conversion from mild cognitive impairment to Alzheimer's disease. *J. Neurol. Neurosurg. Psychiatry* 87, 425–432. doi: 10.1136/jnnp-2014-309105
- Yoshiura, T., Mihara, F., Ogomori, K., Tanaka, A., Kaneko, K., and Masuda, K. (2002). Diffusion tensor in posterior cingulate gyrus: correlation with cognitive decline in Alzheimer's disease. *Neuroreport* 13, 2299–2302. doi: 10.1097/00001756-200212030-00026
- Zhang, L., Ni, H., Yu, Z., Wang, J., Qin, J., Hou, F., et al. (2020). Investigation on the alteration of brain functional network and its role in the identification of mild cognitive impairment. *Front. Neurosci.* 14:1027. doi: 10.3389/fnins.2020.558434
- Zhang, X., Shams, S. P., Yu, H., Wang, Z., and Zhang, Q. (2022). A pairwise functional connectivity similarity measure method based on few-shot learning for early MCI detection. *Front. Neurosci.* 16:1081788. doi: 10.3389/fnins.2022.1081788
- Zhao, F., Feng, F., Ye, S., Mao, Y., Chen, X., Li, Y., et al. (2024). Multi-head self-attention mechanism-based global feature learning model for ASD diagnosis. *Biomed. Signal Process. Control* 91:106090. doi: 10.1016/j.bspc.2024.106090
- Zhao, F., Li, N., Pan, H., Chen, X., Li, Y., Zhang, H., et al. (2022). Multi-view feature enhancement based on self-attention mechanism graph convolutional network for autism spectrum disorder diagnosis. *Front. Hum. Neurosci.* 16:918969. doi: 10.3389/fnhum.2022.918969
- Zheng, J., Xia, A., Shao, L., Wan, T., and Qin, Z. (2019). "Stock volatility prediction based on self-attention networks with social information," in *2019 IEEE Conference on Computational Intelligence for Financial Engineering & Economics (CIFER)* (Shenzhen: IEEE), 1–7.
- Zhu, W.-Z., Zhong, W., Wang, W., Zhan, C., Wang, C., Qi, J., et al. (2009). Quantitative MR phase-corrected imaging to investigate increased brain iron deposition of patients with Alzheimer disease. *Radiology* 253, 497–504. doi: 10.1148/radiol.2532082324
- Zunair, H., Rahman, A., Mohammed, N., and Cohen, J. P. (2020). "Uniformizing techniques to process CT scans with 3D CNNs for tuberculosis prediction," in *International Workshop on Predictive Intelligence in Medicine* (Cham: Springer), 156–168.