Medical Image Analysis 87 (2023) 102828

Contents lists available at ScienceDirect

Medical Image Analysis

journal homepage: www.elsevier.com/locate/media

Dynamic weighted hypergraph convolutional network for brain functional connectome analysis

Junqi Wang [a], Hailong Li [a,b,c,e], Gang Qu [d], Kim M. Cecil [a,e], Jonathan R. Dillman [a,c,e], Nehal A. Parikh [b,f], Lili He [a,b,c,e,*]

[a] Imaging Research Center, Department of Radiology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA
[b] Neurodevelopmental Disorders Prevention Center, Perinatal Institute, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA
[c] Artificial Intelligence Imaging Research Center, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA
[d] Department of Biomedical Engineering, Tulane University, New Orleans, LA, USA
[e] Department of Radiology, University of Cincinnati College of Medicine, Cincinnati, OH, USA
[f] Department of Pediatrics, University of Cincinnati College of Medicine, Cincinnati, OH, USA

ARTICLE INFO

Keywords:
Functional connectome
Weighted hypergraph
Dynamic hypergraph neural network
Manifold regularization

ABSTRACT

The hypergraph structure has been utilized to characterize the brain functional connectome (FC) by capturing the high order relationships among multiple brain regions of interest (ROIs) compared with a simple graph. Accordingly, hypergraph neural network (HGNN) models have emerged and provided efficient tools for hypergraph embedding learning. However, most existing HGNN models can only be applied to pre-constructed hypergraphs with a static structure during model training, which might not be a sufficient representation of the complex brain networks. In this study, we propose a dynamic weighted hypergraph convolutional network (*dw*HGCN) framework to consider a dynamic hypergraph with learnable hyperedge weights. Specifically, we generate hyperedges based on sparse representation and calculate the hyper similarity as node features. The hypergraph and node features are fed into a neural network model, where the hyperedge weights are updated adaptively during training. The *dw*HGCN facilitates the learning of brain FC features by assigning larger weights to hyperedges with higher discriminative power. The weighting strategy also improves the interpretability of the model by identifying the highly active interactions among ROIs shared by a common hyperedge. We validate the performance of the proposed model on two classification tasks with three paradigms functional magnetic resonance imaging (fMRI) data from Philadelphia Neurodevelopmental Cohort. Experimental results demonstrate the superiority of our proposed method over existing hypergraph neural networks. We believe our model can be applied to other applications in neuroimaging for its strength in representation learning and interpretation.

1. Introduction

The functional connectome (FC) derived from functional magnetic resonance imaging (fMRI) characterizes individual differences in blood-oxygen-level-dependent (BOLD) synchronization patterns across various brain regions (Gonzalez-Castillo et al., 2021). FC has served as a robust indicator to decipher brain functional variability (Finn et al., 2015; Grabner et al., 2007; Newman et al., 2003), understand brain organizations, and predict neurodevelopmental functions and brain disorders (Power et al., 2014; Wang et al., 2021; Allen et al., 2022; Yin et al., 2022; Du et al., 2021).

Prior studies have characterized FC using graph theory to define anatomical regions as a set of nodes and pairwise functional connections as edges. Graph theoretical approaches have been used to quantify the topological architectures of the FCs to reveal neurobiologically meaningful measurements (Bullmore and Sporns, 2009; Rubinov and Sporns, 2010; Bullmore and Bassett, 2011; Wang et al., 2018; Zhou et al., 2006; Honey et al., 2009; Bassett et al., 2008; Stam et al., 2006). However, the conventional graph-based approaches only considered pairwise interactions between nodes, which is insufficient to model complex relationships among multiple nodes (Eavani et al., 2015). A hypergraph is a generalization of the conventional graph, in which an edge (named as hyperedge) can connect to any number of nodes (Davison et al., 2015; Jie et al., 2016; Zu et al., 2016; Gu et al., 2017). The weights of the hyperedge reflect the strength and frequency of information flow (i.e., electrical signals) among various brain regions. Larger hyperedge weights reflect that the associated regions are more

* Corresponding author at: Imaging Research Center, Department of Radiology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA.
E-mail address: Lili.He@cchmc.org (L. He).

https://doi.org/10.1016/j.media.2023.102828
Received 14 October 2022; Received in revised form 14 April 2023; Accepted 18 April 2023
Available online 25 April 2023

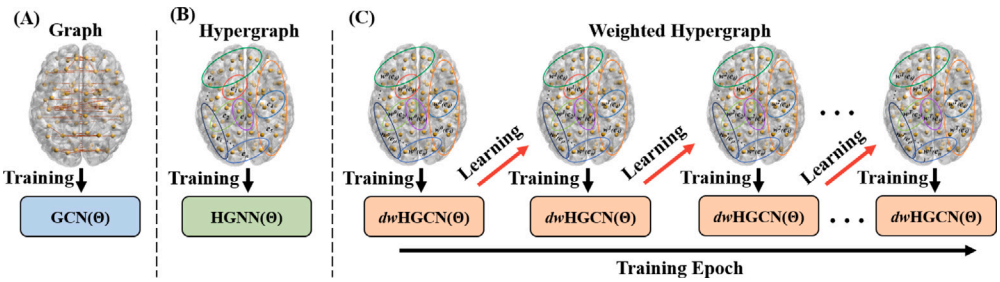

Fig. 1. Comparison of GCN, HGNN, and the proposed *dw*HGCN model. Besides the model weight Θ, our *dw*HGCN model adaptively updates the hyperedge weights during training.

likely to be co-activated during information/knowledge propagation, therefore enabling the representation of multiple interactive brain regions simultaneously in the neural process (Li et al., 2019; Ma et al., 2021; Ji et al., 2021).

To better characterize the brain FC, graph neural network (GNN) (Scarselli et al., 2008) was proposed to perform inference on data described by graphs, followed by a series of variants, such as graph attention network (Velickovic et al., 2017), where node-wise weights were considered via attention mechanism, and graph isomorphism network (Xu et al., 2019), which considered graph isomorphism test in neighbor aggregation. As the most popular variant of GNN, graph convolutional network (GCN) (Kipf and Welling, 2017) has been implemented in various neuroimaging applications with improved diagnosis outcomes. Parisot et al. (2018) applied GCN for group-level brain analysis for diagnosis of autism spectrum disorder (ASD) and Alzheimer's disease (AD) using fMRI from the Autism Brain Imaging Data Exchange (ABIDE) and Alzheimer's Disease Neuroimaging Initiative (ADNI) databases. Liu et al. (2020) used GCN models for detection of early mild cognitive impairment (EMCI) from multi-modal imaging data collected from ADNI. Xing et al. (2019) proposed a dynamic spectral GCN model to predict EMCI from healthy elderly participants with assistant task training. More recently, Yao et al. (2021) developed a multi-scale triplet GCN model for multi labels prediction, including Attention-deficit/hyperactivity disorder (ADHD), mild cognitive impairment, and cerebral small vessel disease, using functional and structural connectomes.

As a generalization of the GCN model, hypergraph neural network (HGNN) models have been proposed for representation learning of data residing on a hypergraph. HGNNs (Feng et al., 2019; Bai et al., 2021) can automatically learn the data representations by inspecting arbitrary number of neighboring nodes in irregular non-Euclidean hypergraph structure via convolutional filters and thus improved the generalizability of the model. For example, Feng et al. (2019) proposed a HGNN model for citation and visual object classifications tasks by designing a convolution operation on hypergraph structure. Shi et al. (2018) developed a hypergraph-induced convolutional network that combined hypergraph learning and graph convolutional networks with applications in visual classification. Jiang et al. (2019) created a dynamic hypergraph neural network to classify citation and social media networks, which involved hypergraph reconstruction using k-nearest neighborhood and k-means clustering. Bi et al. (2022) presented an influence hypergraph convolutional generative adversarial network (GAN) that combines traditional GCN and GAN on brain ROI-gene hypergraph for EMCI and late mild cognitive impairment classification. Madine et al. (2020) incorporated multi-kernel learning into the construction of HGNN and found applications in classifying ASD. However, most existing HGNNs neglected the hyperedge weights by assuming equal weights for each individual hyperedge. As such, the potentially distinctive features embedded in the information flow among brain FCs may not be appropriately decoded and fully utilized (Bai et al., 2021). In addition, existing models can only be applied to pre-constructed hypergraphs, which remained static during the model training. In other words, the pre-calculated graph weights were not updated during the convolutional operation and model learning process. However, keeping the weights unchanged restricted the opportunity to optimize the hyperedge weights by utilizing the outcome labels to enhance its embedded discriminative information.

In this study, to overcome the above limitations, we propose a dynamic weighted hypergraph convolutional network (*dw*HGCN) model to dynamically learn the brain FC representations (i.e., feature embeddings) by integrating weighted hypergraph learning and hypergraph neural network model training into a unified framework. As shown in Fig. 1(C), the model weights Θ is updated by minimizing model training error, and meanwhile, the weighted hypergraph $\mathcal{G}(\mathbf{w})$ is dynamically updated by learning hyperedge weights at each training epoch. The weighted hypergraph with the minimized training error is estimated to facilitate the characterization of the brain FC. Compared to HGNN models (Fig. 1(C)) that only employ a static hypergraph structure, our approach is able to assign larger weights to hyper connections with higher discriminative power automatically, and, thus, improve the characterization of the brain FC.

Moreover, we present an objective function for hypergraph classification tasks by considering the manifold regularization, which regulates the smoothness of the fMRI time series signals on the hypergraph. A signal is considered smooth if the signal values are more similar on nodes connected by stronger edges/hyperedges (Dong et al., 2019). This property is often observed in brain networks, where multiple correlated ROIs associated with a cognitive task are activated (Eavani et al., 2015). Quantitatively, we define the smoothness of signals on the hypergraph with hypergraph Laplacian quadratic form and formulated the manifold regularization term in the objective function. The enforced regularization promotes the fit between the estimated hypergraph structure and the ROI-level fMRI time series signals.

We evaluate the proposed *dw*HGCN framework on fMRI data collected from the Philadelphia Neurodevelopmental Cohort (PNC) (Satterthwaite et al., 2014). PNC is a large-scale collaborative study between the Brain Behavior Laboratory at the University of Pennsylvania and the Center for Applied Genomics at the Children's Hospital of Philadelphia, including healthy developing participants aged 8 to 22 years. We perform two classification tasks: (1) classifying participants into a child or a young adult group with resting-state fMRI, and (2) classifying participants into low or high Wide Range Achievement Test (WRAT) score groups using task-based fMRI. The experimental results demonstrate that our proposed *dw*HGCN model outperforms the peer-competing graph learning models, including the graph-based and hypergraph-based neural networks.

The contributions of this work can be summarized as follows:

- We describe a novel *dw*HGCN framework to dynamically learn the feature representation of weighted hypergraph by integrating brain FC learning and hypergraph convolutional network model training into a unified end-to-end framework. Different from the previous HGNN models (Feng et al., 2019), our *dw*HGCN can simultaneously optimize the weighted hypergraph and deep learning model for the hypergraph classification tasks.

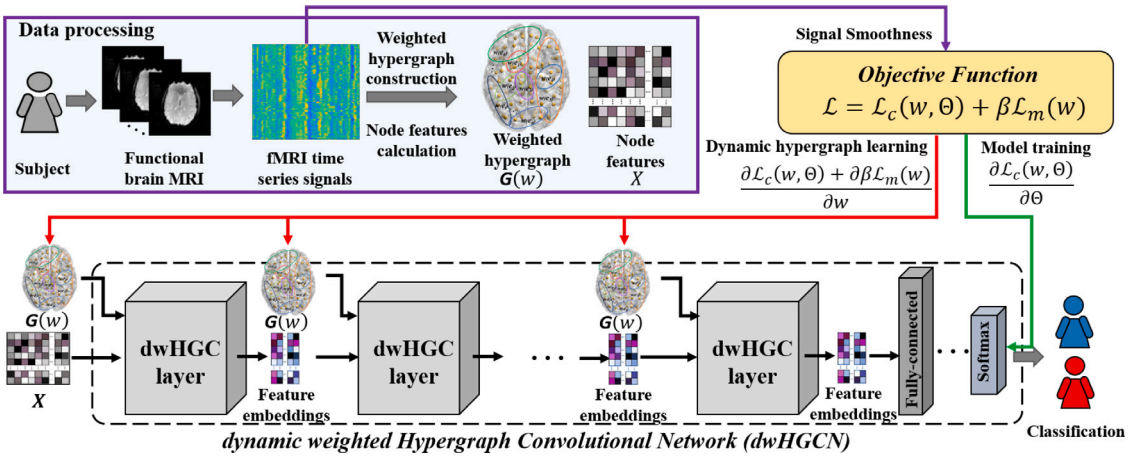

Fig. 2. The framework of the proposed *dw*HGCN model. At the first step, we construct the hypergraph and calculate the hypergraph similarity matrix as node features to perform graph classification. At the second step, we design *dw*HGC layers to simultaneously optimize both the hypergraph and the model weights during back-propagation. The updates of hypergraph are also regulated by the fMRI signals' smoothness on the estimated hypergraph. Finally, we design output layers to generate predicted labels.

- We further present an objective function for hypergraph classification tasks by integrating manifold regularization. The manifold regularization term is enforced on the hyperedge weights to control the fit between the estimated hypergraph structure and ROI-level fMRI time series signals. This is achieved by minimizing the overall smoothness across fMRI time series signals of individual ROIs (i.e., vertices) quantified by the hypergraph Laplacian quadratic form. The regulation helps generate a neurological meaningful hypergraph to represent brain FC.
- We evaluate the proposed *dw*HGCN framework with two classification tasks on a large-scale neuroimaging dataset and conduct extensive experiments. The results demonstrate the superiority of the proposed method over competing graph and hypergraph based neural network models (Kipf and Welling, 2017; Qu et al., 2021a; Qu et al., 2021b; Feng et al., 2019; Bai et al., 2021).
- Beyond classification, our proposed model provides enhanced biological interpretation of the brain FC by weighting the interactions among multiple brain ROIs indicated by their common hyperedge and thus provides more interpretable results. Potential hub regions are identified related to brain developments and learning ability during adolescence. Most of the detected regions are highly functionally relevant to the classification task and are consistent with previous reports. Therefore, our work could provide potential insights into the brain functional connectome and result in more neuroimaging applications.

2. Method

In this work, we propose a *dw*HGCN model to dynamically learn brain FC feature representation by integrating hypergraph learning and hypergraph neural network model training into a unified end-to-end framework. An overview of our study is outlined in Fig. 2. Briefly, the fMRI data from individual participants are preprocessed and parcellated into ROI-level fMRI time series signals. Using such preprocessed fMRI data, we model the brain FC as a weighted hypergraph. We calculate the hypergraph similarity matrix as node features. The proposed model consists of dynamic weighted hypergraph convolutional (*dw*HGC) layers, fully-connected layers, and a softmax layer for hypergraph classification tasks. The *dw*HGC layers perform spectral graph convolutions on weighted hypergraph and node features/feature embeddings. The proposed objective function includes a cross-entropy function $\mathcal{L}_c(\mathbf{w},\Theta)$ for the hypergraph classification and a manifold regularization term $\mathcal{L}_m(\mathbf{w})$ that is enforced on the hyperedge weights to enhance the smoothness condition quantified by hypergraph Laplacian quadratic form. In the feed-forward procedure, weighted hypergraph and node features are input into the proposed *dw*HGCN model to learn feature embeddings, which are then fed to fully connected and softmax layers for classification. In the back-propagation procedure, the *dw*HGCN utilizes the proposed objective function to simultaneously conduct dynamic hypergraph learning and model training. Different from prior static models, the weighted hypergraph is dynamically updated considering both the supervised label information and the fMRI signals' smoothness on the estimated hypergraph. The notations and definitions of the used symbols are summarized in Table 1.

Table 1
Notations and definitions.

Notation	Definition
$\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathbf{w})$	Hypergraph
$\mathcal{V} \in \mathbb{R}^N$	The vertex set of the hypergraph
$\mathcal{E} \in \mathbb{R}^M$	The hyperedge set of the hypergraph
$H \in \mathbb{R}^{N\times M}$	The incidence matrix of the hypergraph
$\mathbf{D}_v \in \mathbb{R}^{N\times N}$	The diagonal vertex degree matrix
$\mathbf{D}_e \in \mathbb{R}^{M\times M}$	The diagonal hyperedge degree matrix
$W \in \mathbb{R}^{M\times M}$	The diagonal hyperedge weight matrix
$S \in \mathbb{R}^{N\times N}$	The hypergraph similarity matrix
$L \in \mathbb{R}^{N\times N}$	The hypergraph Laplacian matrix
$F \in \mathbb{R}^{N\times P}$	The fMRI time series data
$U \in \mathbb{R}^{N\times N}$	The eigenvector matrix of L
$\mathbf{X}^{(l)} \in \mathbb{R}^{N\times k^l}$	The hypergraph signal at lth layer ($\mathbf{X}^{(0)} = S$)
$\Theta^{(l)} \in \mathbb{R}^{k^{(l)}\times k^{(l+1)}}$	The learnable weight matrix at lth layer
$\mathbf{A}^l \in \mathbb{R}^{N\times K}$	The feature embedding at the last layer

2.1. Preliminaries

We begin by introducing basic notions in hypergraph theory. A hypergraph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathbf{w})$ includes a vertex set $\mathcal{V} = \{v_1, v_2, \ldots, v_N\}$, a hyperedge set $\mathcal{E} = \{e_1, e_2, \ldots, e_M\}$, and a hyperedge weight vector $\mathbf{w} = (w(e_1), w(e_2), \ldots, w(e_M))^T \in \mathbb{R}^M$ where each hyperedge e_i is assigned with a weight $w(e_i)$. In this work, the nodes of the hypergraph represent brain ROIs and hyperedges represent high-order interactions among these ROIs. The hypergraph $\mathcal{G}$ is denoted by an incidence matrix $\mathbf{H} = [H_{ij}] \in \mathbb{R}^{N\times M}$ with entries

$$H_{ij} = \begin{cases} 1, & \text{if } v_i \in e_j \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

The vertex degree and hyperedge degree are defined as follows, respectively,

$$d(v_i) = \sum_{e_j \in \mathcal{E}} w(e_j) H_{ij} \quad \text{for } 1 \le i \le N, \tag{2}$$

and

$$\delta(e_j) = \sum_{v_i \in \mathcal{V}} H_{ij} \quad \text{for } 1 \le j \le M. \tag{3}$$

Accordingly, $\mathbf{D}_v$ and $\mathbf{D}_e$ denote the diagonal matrices of the vertex degrees and the hyperedge degrees, respectively.

As proposed in our previous work (Xiao et al., 2019), we define the similarity matrix of a hypergraph $\mathcal{G}$ to reflect the high-order interactions between vertices by

$$\mathbf{S} = \mathbf{H}\mathbf{W}\mathbf{D}_e^{-1}\mathbf{H}^T \in \mathbb{R}^{N\times N} \tag{4}$$

where

$$S_{ij} = \sum_{e_k \in \mathcal{E}} \frac{w(e_k)}{\delta(e_k)} H_{ik} H_{jk} \tag{5}$$

is the weighted summation over all hyperedges containing both vertices v_i and v_j. $\mathbf{W} = diag(\mathbf{w}) \in \mathbb{R}^{M\times M}$ denotes the diagonal matrix of the hyperedge weights, where hyperedges with larger weights are considered more informative in representing the hyper network. Similar to the simple graph setting, the normalized form of hypergraph Laplacian matrix is defined by

$$\mathbf{L} = \mathbf{I}_N - \mathbf{D}_v^{-1/2}\mathbf{H}\mathbf{W}\mathbf{D}_e^{-1}\mathbf{H}^T\mathbf{D}_v^{-1/2} \tag{6}$$

where $\mathbf{I}_N$ is the identity matrix.

2.2. Hypergraph initialization

In this subsection, we construct the incidence matrix $\mathbf{H}$ to reveal the basic structure of the hypergraph. Suppose we have fMRI time series $\mathbf{F} = [\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_P] \in \mathbb{R}^{N\times P}$ available for each subject, where N and P denote the number of ROIs and time points, respectively. For each graph signal $\mathbf{x} \in \mathbb{R}^N$ defined on the vertex set of the graph, it has been normalized with zero mean and unit Euclidean norm. Based on the previous works (Jie et al., 2016; Li et al., 2019; Gao et al., 2020), we adopt a representation based hypergraph generation method. Specifically, the incidence matrix $\mathbf{H}$ is estimated using l_1 sparse representation

$$\min_{\boldsymbol{\alpha}_i} \frac{1}{2}\|\mathbf{f}_i - \mathbf{B}_i\boldsymbol{\alpha}_i\|_2^2 + \gamma\|\boldsymbol{\alpha}_i\|_1 \tag{7}$$

where $\mathbf{B}_i = \left(\mathbf{f}_1, \dots, \mathbf{f}_{i-1}, \mathbf{f}_{i+1}, \dots, \mathbf{f}_n\right) \in \mathbb{R}^{P\times N-1}$ is the time sequential data on all the vertices except the ith one. $\boldsymbol{\alpha}_i \in \mathbb{R}^{N-1}$ is the coefficient vector (hyperedge) reflecting the neighborhood information of the remaining nodes towards centroid node i. The elements in $\boldsymbol{\alpha}_i$ with larger values indicate the stronger interactions between the corresponding ROIs and the centroid node i. In our work, ROIs with values less than 10^{-4} in their corresponding coefficient vector $\boldsymbol{\alpha}_i$ are excluded in the hyperedge e_i to filter out brain regions with insignificant or spurious interactions. γ is a regularization parameter controlling the sparsity, where the coefficient vector $\boldsymbol{\alpha}_i$ becomes dense or sparse as γ decreases or increases.

The incidence matrix $\mathbf{H}$ for each subject is characterized by concatenating the hyperedges obtained from Eq. (7) for each node, resulting in an incidence matrix $\mathbf{H} \in \mathbb{R}^{N\times N}$ for each subject. Despite the generation of the incidence matrix, the weights of the hyperedges have not been determined yet. The hyperedge weights will be updated adaptively at the *dw*HGC layers, where task-specific hypergraphs are formulated.

2.3. DwHGC layers

By initializing the diagonal matrix of hyperedge weights $\mathbf{w}$ as an identity matrix, the similarity matrix $\mathbf{S}$ and its corresponding hypergraph Laplacian matrix $\mathbf{L}$ can be obtained (Eqs. (2), (4) and (6)). It is easy to observe that the hypergraph Laplacian is positive semi-definite. We can define the hypergraph Fourier transform and its inverse for a signal $\mathbf{x}$ on hypergraph as $\hat{\mathbf{x}} = \mathbf{U}^T\mathbf{x}$ and $\mathbf{x} = \mathbf{U}\hat{\mathbf{x}}$ by performing eigendecomposition $\mathbf{L} = \mathbf{U}\Lambda\mathbf{U}^T$. Similar to the simple graph scenario, the eigenvectors are regarded as the Fourier basis and the eigenvalues are interpreted as frequencies. The spectral convolution of signal $\mathbf{x}$ and filter g can be denoted as Kipf and Welling (2017), Feng et al. (2019)

$$g * \mathbf{x} = \mathbf{U}g(\Lambda)\mathbf{U}^T\mathbf{x} \tag{8}$$

where $g(\Lambda) = diag(g(\lambda_1), \dots, g(\lambda_n))$ is a function of the eigenvalues (frequencies) of $\mathbf{L}$. The first order Chebyshev polynomial is adopted to approximate the convolution operation (Kipf and Welling, 2017) as

$$g * \mathbf{x} \approx \theta_0\mathbf{x} - \theta_1\mathbf{D}_v^{-\frac{1}{2}}\mathbf{H}\mathbf{W}\mathbf{D}_e^{-1}\mathbf{H}^T\mathbf{D}_v^{-\frac{1}{2}}\mathbf{x} \tag{9}$$

To address overfitting and reduce the number of operations, a single parameter θ is adopted to constrain the number of parameters (Kipf and Welling, 2017; Feng et al., 2019), defined as

$$\begin{cases} \theta_1 = -\frac{1}{2}\theta \\ \theta_0 = \frac{1}{2}\theta\mathbf{D}_v^{-\frac{1}{2}}\mathbf{H}\mathbf{W}\mathbf{D}_e^{-1}\mathbf{H}^T\mathbf{D}_v^{-\frac{1}{2}} \end{cases} \tag{10}$$

The defined convolution operation (8) can be further simplified to the following expression

$$g * \mathbf{x} \approx \theta\mathbf{D}_v^{-\frac{1}{2}}\mathbf{H}\mathbf{W}\mathbf{D}_e^{-1}\mathbf{H}^T\mathbf{D}_v^{-\frac{1}{2}}\mathbf{x} \tag{11}$$

Thus we can formulate the hypergraph convolution layer in the matrix form as

$$\mathbf{X}^{(l+1)} = \sigma(\mathbf{D}_v^{-\frac{1}{2}}\mathbf{H}\mathbf{W}\mathbf{D}_e^{-1}\mathbf{H}^T\mathbf{D}_v^{-\frac{1}{2}}\mathbf{X}^{(l)}\Theta^{(l)}) \tag{12}$$

where $\mathbf{X}^{(l)}$ is hypergraph embedding at lth layer with $\mathbf{X}^{(0)} = \mathbf{S}$, $\Theta^{(l)}$ is the learnable weight matrix at lth layer, and σ denotes the nonlinear activation function that we used ReLU (Rectified Linear Unit) function throughout the paper.

To emphasize the contributions of highly activated interactions among brain regions containing more information, weighted hyperedges are introduced to better characterize the hyper network. In other words, the weight matrix $\mathbf{W}$ in Eq. (12) will be updated adaptively during training. It is worth noting that the diagonal degree matrix of vertex degree is a function of $\mathbf{w}$ (i.e., $\mathbf{D}_v = diag(\mathbf{H}\mathbf{w})$) and will also be updated during training. Therefore, the weighted hypergraph convolution is derived as

$$\mathbf{X}^{(l+1)} = \sigma(\mathbf{D}_v(\mathbf{w})^{-\frac{1}{2}}\mathbf{H}\mathbf{W}\mathbf{D}_e^{-1}\mathbf{H}^T\mathbf{D}_v(\mathbf{w})^{-\frac{1}{2}}\mathbf{X}^{(l)}\Theta^{(l)}) \tag{13}$$

where both hyperedge weights $\mathbf{w}$ and filter weights Θ are learnable during training. We perform pseudoinverse on $\mathbf{D}_v$ when it becomes singular. Notably, each hyperedge e_i contains the interactions between multiple brain regions and the centroid node i. A larger weight in $\mathbf{w}$ indicates a stronger neuron activity among brain regions shared by the corresponding hyperedge. Thus, an activation map of brain sub networks can be formulated from different hyperedge weights $\mathbf{w}$.

2.4. Objective function

Apart from the label fitting with the weighted hyperedges, a neurological meaningful assumption needs to be considered that the fMRI observations (graph signals) should be smooth on the deterministic hyper network. Motivated by this, we consider smoothness concept on hypergraph and derived a manifold regularization term.

Given an arbitrary signal $\mathbf{x} = \left(x_1, x_2, \dots, x_N\right)^T \in \mathbb{R}^N$ defined on the hypergraph $\mathcal{G}$, similar to the smoothness concept on the traditional graph (Xiao et al., 2019), we can define the smoothness of $\mathbf{x}$ using the Laplacian quadratic form according to Eqs. (4) and (5) by

$$\begin{aligned} \mathbf{x}^T\mathbf{L}\mathbf{x} &= \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} S_{ij}\left(x_i - x_j\right)^2 \\ &= \frac{1}{2}\sum_{e_k\in\mathcal{E}}\sum_{v_i,v_j\in\mathcal{V}} \frac{w(e_k)H_{ik}H_{jk}}{\delta(e_k)}\left(x_i - x_j\right)^2 \end{aligned} \tag{14}$$

The signal $\mathbf{x}$ is considered to be smooth on hypergraph $\mathcal{G}$ if its value in Eq. (14) is small.

Based on the smoothness assumption, a manifold regularization term is enforced on the hyperedge weights $\mathbf{w}$ to regulate the fit between the estimated hypergraph structure and graph signals (fMRI observations $\mathbf{F}$) residing on the graph. It is achieved by minimizing the smoothness defined in Eq. (14) of all graph signals on the hypergraph, which is formulated by

$$\begin{aligned} &\min_{\mathbf{w}} \operatorname{tr}(\mathbf{F}^T\mathbf{L}\mathbf{F}) + \rho\|\mathbf{w}\|_2^2 \\ &\text{s.t. } \mathbf{w} \geq 0 \\ &\quad \|\mathbf{w}\|_1 = 1 \end{aligned} \tag{15}$$

where $\operatorname{tr}(\cdot)$ denotes trace operation, $\mathbf{w} \geq 0$ denotes all the elements in $\mathbf{w}$ is greater than or equal 0, $\mathbf{L} = \mathbf{I}_N - \mathbf{D}_v^{-1/2}\mathbf{H}\mathbf{W}\mathbf{D}_e^{-1}\mathbf{H}^T\mathbf{D}_v^{-1/2}$ is a function of $\mathbf{w}$, and ρ is a positive coefficient for the l_2 penalty. A small ρ leads to a shrinkage of the weights while a large ρ leads to similar weights assigned to all hyperedges. The two constraints are enforced to avoid negative weights and trivial solutions. Thus, the manifold loss is denoted as

$$\mathcal{L}_m(\mathbf{w}) = \operatorname{tr}(\mathbf{F}^T\mathbf{L}_{(\mathbf{w})}\mathbf{F}) + \rho\|\mathbf{w}\|_2^2 \tag{16}$$

under constraints $\mathbf{w} \geq 0$ and $\|\mathbf{w}\|_1 = 1$. Therefore, the loss function for our proposed *dw*HGCN model is formulated as

$$\mathcal{L} = \mathcal{L}_c(\mathbf{w}, \Theta) + \beta\mathcal{L}_m(\mathbf{w}) \tag{17}$$

where $\mathcal{L}_c(\mathbf{w}, \Theta)$ denotes the cross entropy loss. The gradient descent algorithm is enforce to update both learnable parameters simultaneously during training.

$$\begin{aligned} \mathbf{w}^{(l)} &= \mathbf{w}^{(l)} - r\frac{\partial\mathcal{L}}{\partial\mathbf{w}^{(l)}} \\ \Theta^{(l)} &= \Theta^{(l)} - r\frac{\partial\mathcal{L}}{\partial\Theta^{(l)}} \end{aligned} \tag{18}$$

The detailed derivation of the gradients of loss function with respect to the weights Θ and $\mathbf{w}$ can be found in the appendix.

2.5. Hyperedge interpretation strategy

Compared with most existing hypergraph-based models, our model weighs the hyper-interactions during model training. Since our model considers both the supervised label information and the signal smoothness, the learned hyperedge weights can reveal the discriminative capability of different hyperedges, which facilitates the extraction of the most meaningful biomarkers by assigning larger weights to hyperedges with higher task-specific discriminative power.

In our work, we incorporate a Gradient weighted Class Activation Mapping (Grad-CAM) (Selvaraju et al., 2017; Pope et al., 2019) method to detect class-specific contributions of the brain ROIs. Specifically, after the optimization of the weighted hypergraph and the *dw*HGCN model, we denote the feature embedding at the last *dw*HGC layer as $\mathbf{A}^l = dwHGCN_{(\Theta)}(G(\mathbf{H}, \mathbf{w}), \mathbf{X}) \in \mathbb{R}^{N\times K}$, where N is the number of nodes (i.e., brain ROIs), and K is the hidden dimension of the last layer. We calculate the gradient of each class c with respect to the weighted hypergraph embeddings at the last *dw*HGC layer. We compute the contribution coefficient of the kth filter as

$$\eta_k^c = \frac{1}{N}\sum_i^N \frac{\partial y^c}{\partial \mathbf{A}_{i,k}^l} \tag{19}$$

where y^c represents the label for class c (e.g., children or young adult group), and $\mathbf{A}_{i,k}^l$ represents the kth feature embedding element for the ith hypergraph node. The contribution vector of brain ROIs is calculated by a weighted combination of feature embedding at the last *dw*HGC layer followed by a ReLU function as in a prior study (Pope et al., 2019) by

$$ReLU(\sum_k \eta_k^c \mathbf{A}_k^l) \tag{20}$$

Table 2
Characteristics of the participants in this study. SD: standard deviation.

Age classification			
Group	Age (Mean ± SD)	Male/Female	Total
Child	10.3320 ± 0.9452	89/102	191
Young adult	19.2622 ± 0.9928	79/119	198
WRAT classification			
Group	Age (Mean ± SD)	Male/Female	Total
High score	14.6753 ± 3.1972	89/76	165
Low score	15.6257 ± 3.1791	75/108	183

where $\mathbf{A}_k^l$ is the kth feature channel of the embedding for all N nodes at the last *dw*HGC layer. In short, the contribution of brain ROIs is derived from the hypergraph incidence matrix $\mathbf{H}$, hypergraph node feature $\mathbf{X}$, optimized hyperedge weights $\mathbf{w}$, and optimized filter weight Θ of the *dw*HGCN model.

To extract the most predictive features with the highest contribution for each class, we average the gradient map across all the subjects and formed a cohort-level weight vector. The brain ROIs whose weights are two standard deviations higher than the mean value are preserved as hub regions with the most discriminative power in classifying brain FC.

3. Experimental results

3.1. Data acquisition and preprocessing

A large-scale collaborative neuroimaging dataset from PNC project (Satterthwaite et al., 2014) was utilized to evaluate the performance of the *dw*HGCN model. The PNC dataset involves over 1,400 healthy typically developing adolescents aged 8 to 22. Multiple paradigms of fMRI data, including imaging tasks of working memory (nback) and emotion identification (emoid), and resting state (rs) fMRI, are acquired on a subset of the participants. All MRI examinations were conducted on the same 3T Siemens TIM Trio scanner using an identical imaging protocol and the data are publicly and freely available at the database of Genotypes and Phenotypes (dbGaP). Participants completed a computerized neurocognitive battery (CNB) (Gur et al., 2010). During CNB assessment, the WRAT (Wilkinson and Robertson, 2006) score assessing reading skills, math skills, spelling, and comprehension was obtained. In this work, we investigated the model's performance for classifying age groups using datasets employing a resting state fMRI paradigm. We also classified WRAT score groups from task-based fMRI datasets using working memory and emotional identification paradigms.

All fMRI data were preprocessed using SPM12 (Ashburner et al., 2014), including motion correction, co-registration, spatial normalization to standard MNI space and spatial smoothing with a 3 mm FWHM Gaussian kernel. The feedback regarding head movement (captured by MoTrack, a motion tracking system) was regressed out, followed by a band-pass filter in 0.01 Hz to 0.1 Hz frequency range. Parcellation was performed to map the whole brain via a functional template (Power et al., 2011), resulting in 264 regions of interest with a sphere radius of 5 mm. The resting state fMRI images were acquired when subjects were instructed to stay awake with their eyes open, fixate on the displayed crosshair, and remain still. During the emotion identification task scan, participants were required to label sixty faces with neutral, happy, sad, angry, or fearful expressions performed by well-trained actors. During the working memory task scan, participants were required to respond to a given fractal in the 0-back setting, or respond to the fractal that was the same as the previously displayed one or two in the 1-back and 2-back settings.

For age classification, the child group consists of participants lower than 12 years old, while the young adult group consists of participants above 18 years old (Wang et al., 2020). For WRAT classification, we converted the WRAT scores into z-scores, where subjects with z-score

Table 3
Comparison of model performance on age classification task.

Category	Age classification with rs fMRI				
	Method	ACC	SEN	SPE	AUC
Graph based models	TAG	0.7649 ± 0.0218	0.6673 ± 0.0328	0.8719 ± 0.0250	0.8441 ± 0.0205
	SAGE	0.7952 ± 0.0211	0.7176 ± 0.0334	0.8699 ± 0.0238	0.8736 ± 0.0186
	CHEB	0.8051 ± 0.0212	0.7494 ± 0.0329	0.8636 ± 0.0260	0.8645 ± 0.0193
	GCN	0.8104 ± 0.0195	0.7990 ± 0.0288	0.8214 ± 0.0273	0.8622 ± 0.0191
	GAT	0.8306 ± 0.0190	0.8225 ± 0.0281	0.8386 ± 0.0264	0.8883 ± 0.0175
Hypergraph based models	HGNN	0.8692 ± 0.0170	0.8972 ± 0.0209	0.8417 ± 0.0257	0.9219 ± 0.0141
	HGSVM-L	0.8735 ± 0.0167	0.8583 ± 0.0251	**0.8893 ± 0.0233**	0.9450 ± 0.0112
	wHGNN	0.8743 ± 0.0175	0.8717 ± 0.0241	0.8770 ± 0.0247	0.9290 ± 0.0141
	HGSVM-R	0.8826 ± 0.0168	0.8805 ± 0.0251	0.8849 ± 0.0231	0.9583 ± 0.0089
	***dw*HGCN(Ours)**	**0.9236 ± 0.0133**	**0.9460 ± 0.0142**	0.8877 ± 0.0255	**0.9770 ± 0.0058**

ACC, SEN, SPE, and AUC are the abbreviations of accuracy, sensitivity, specificity, and area under the curve, respectively. We display the result for mean value ± standard derivation from 1000 bootstraps. The results with the best performance are highlighted in boldface. We use the same abbreviations and bootstrapping strategy to present the results in the following tables.

higher than 0.5 were considered high WRAT and subjects with z-score lower than −0.5 belong to low WRAT group (Xiao et al., 2019; Zille et al., 2018). We performed the Chi-Square statistic test (McHugh, 2013) and found no significant difference in distribution of males and females in both classification tasks. Detailed information of the involved subjects can be found in Table 2.

3.2. Competing methods

We compared the performance of the proposed *dw*HGCN method, which features dynamic weighted hypergraph learning with a number of competing static graph/hypergraph learning models. The details of these models were provided below:

- Graph Convolutional Networks (GCN) (Kipf and Welling, 2017): the most widely used classical graph model that utilizes mean pooling and normalized message passing to aggregate neighborhood information on a static graph. We implemented this method as a baseline approach for graph convolutional-based neural networks.
- Chebyshev convolutional network (CHEB) (Defferrard et al., 2016) : a popular variant of GCN model, where the graph filter is approximated with the second-order Chebyshev polynomial. We evaluated this method as an alternative graph convolutional-based approach that works on a static weighted graph.
- GraphSage network (SAGE) (Hamilton et al., 2017): a GNN that learns node embedding efficiently by conducting inductive representation learning on a static weighted graph using a generalized aggregation function.
- Topology adaptive graph convolutional network (TAG) (Du et al., 2017): a variant of GCN that considers the local topology of top K-hop neighbors on a static, weighted graph in the graph convolutional layer.
- Graph attention network (GAT) (Velickovic et al., 2017): a GNN that applies a self-attention strategy to the graph neural network to compute feature representations of each node in a static, weighted graph.
- Hypergraph support vector machine (HGSVM) classifiers (Xiao et al., 2019): a linear SVM classifier is trained using the weighted hypergraph similarity matrix. In this work, we extended this approach by implementing a Gaussian radial basis function (RBF) kernel-based SVM. We adopted these two approaches as traditional machine learning benchmarks for comparing with our proposed deep learning model.
- Hypergraph neural network (HGNN) (Feng et al., 2019; Bai et al., 2021): a generalized hypergraph GCN model that performs representation learning on static unweighted hypergraph structured data.
- Weighted hypergraph neural network (wHGNN): a variant of HGNN that works on a weighted hypergraph. We obtained the weighted hypergraph first from hypergraph learning (Xiao et al., 2019) and used it in the HGNN model. Compared to our approach, wHGNN can only be applied to static hypergraph.

In summary, these selected competing approaches represent the most recent advances in graph and hypergraph machine learning and deep learning domain. Our *dw*HGCN model can be viewed as a generalization of hypergraph neural networks (Feng et al., 2019; Bai et al., 2021), where only unweighted and static hypergraphs were discussed. We evaluated the wHGNN and HGSVM (Xiao et al., 2019), to explore the hyperedges that were weighted but static. HGNN model was inspired by GCN model (Kipf and Welling, 2017), so we also included the classical GCN and its several variants (Defferrard et al., 2016; Hamilton et al., 2017; Velickovic et al., 2017; Du et al., 2017) into the comparison to measure the performance of the static, unweighted or weighted graph in our brain connectome classification.

For fair comparisons of the hypergraph-based models, we used the same incidence matrix constructed in Eq. (7) and the same architecture for all competing neural network methods. For both HGSVM models, we applied the least absolute shrinkage and selection operator (LASSO) (Ranstam and Cook, 2018) for feature selection and hyperparameter optimization. We used HGSVM-L and HGSVM-R as abbreviations for linear and RBF-based HGSVM, respectively. For the classical GCN model and its variants, we calculated the Pearson's correlations among brain regions and keep the top 20% connections as the adjacency matrix. We optimized sparsity by comparing the performance with different sparsity ratios, i.e., 10%, 20%, 50%, and 100% (fully connected graph). The best performance was achieved when the sparsity value was equal to 20%. We tested the node features with an identity matrix, the connectivity matrix obtained from correlation matrix, and the original fMRI time series and report the results with the best performance.

3.3. Experimental setup and hyperparameter tuning

For both classification tasks, we randomly split the data into training, validation, and testing sets with the ratio of 80%, 10%, and 10%, respectively. The incidence matrix $\mathbf{H}$ was constructed as in Eq. (7), where the hyperparameter γ was obtained empirically from our previous work (Xiao et al., 2019). Based on the method section, the similarity matrix $\mathbf{S}$, the hyperedge degree matrix $\mathbf{D}_e$, and the vertex degree matrix $\mathbf{D}_v$ were initialized (setting $\mathbf{W} = \mathbf{I}$). We designed and optimized *dw*HGCN models by considering k = [1 − 5] *dw*HGC layers, followed by two fully connected layers, and a softmax output layer for classification label prediction.

The model was trained on training set with a maximal epoch of 150 and a learning rate of 0.001. The hyperparameters were tuned on the validation sets based on the highest area under the receiver

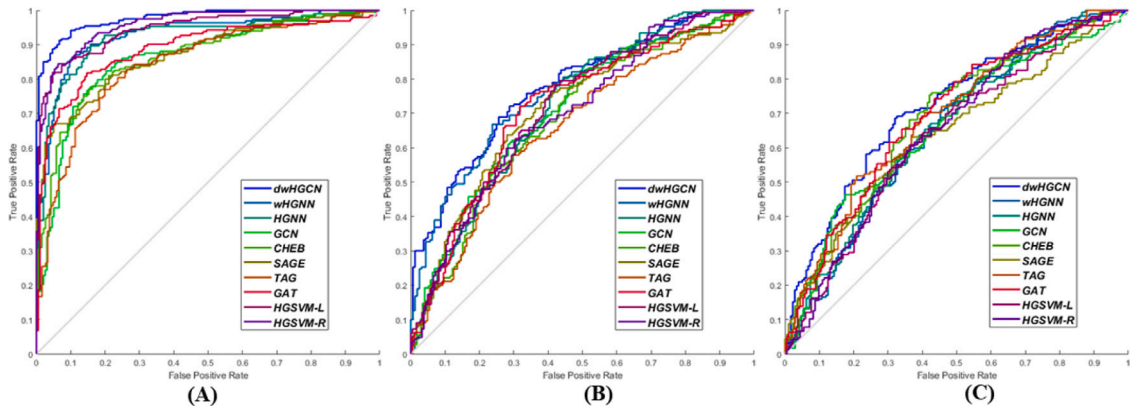

Fig. 3. Receiver operating characteristic curves of the proposed method and competing method. (A) age classification task with rs fMRI. (B), (C) WRAT classification task with nback and emoid fMRI, respectively.

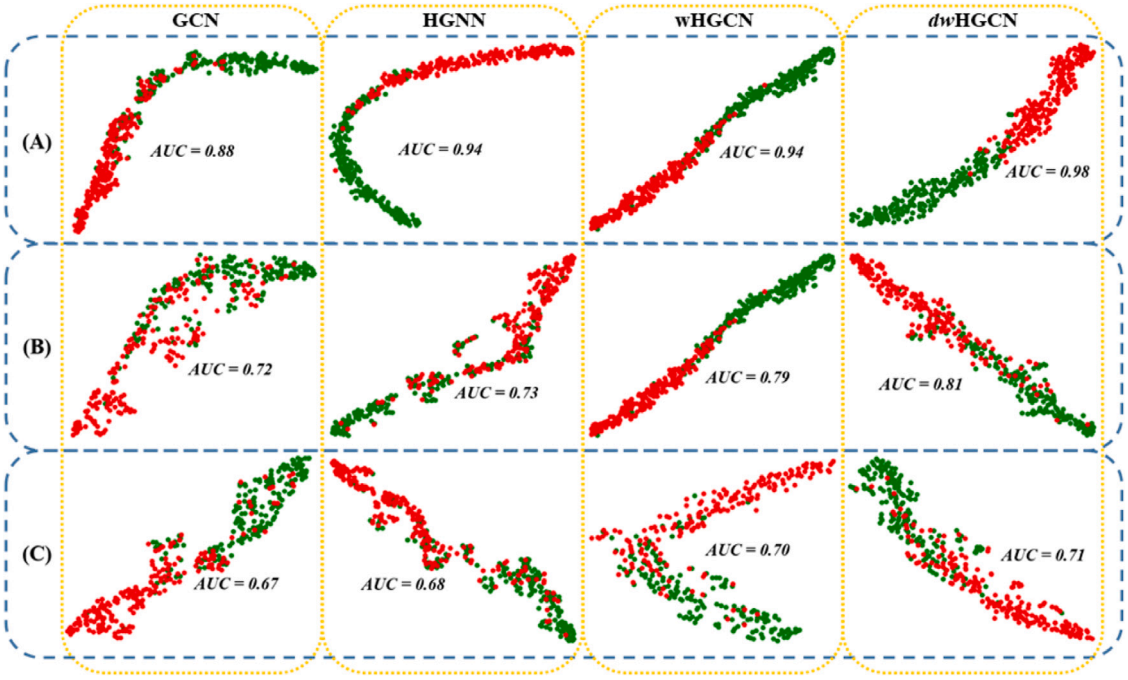

Fig. 4. T-Distributed Stochastic Neighbor embedding (t-SNE) visualization of the embedded features of the proposed method (last column) and several competing methods. (A) Age classification with rs fMRI. The green dots represent children and the red dots represent young adults. (B),(C) WRAT classification with nback and emoid fMRI. The green dots represent subjects in the low WRAT group and the red dots represent the subjects in the high WRAT group. Our proposed model achieves the best clustering quality for separating groups both visually and quantitatively.

operating characteristic curve (AUC). After optimization, the number of feature channels in the hidden layer was set to 32. We set k = 2 for age classification task and k = 3 for WRAT score classification tasks. The manifold regularization parameter β was set to 0.01. The dropout layers (drop rate $p = 0.25$) were introduced to avoid overfitting. We used accuracy, sensitivity, specificity, and AUC as criteria to evaluate the classification performance. Bootstrapping strategy (Abney, 2002) was adopted to reduce sampling bias by 1000 repeated experiments with random resampling. The 95% confidence interval was calculated based on the accuracy for all competing methods and we determined the significance based on the overlaps of the confidence intervals (if not overlapped, the results from two methods were different with confidence 95%).

3.4. Age classification

To evaluate the proposed *dw*HGCN framework on hypergraph classification task, we performed the age classification to classify subjects into children or young adults. The comparisons of the classification results with competing methods are shown in Table 3. Specifically, our *dw*HGCN model achieved the best results on both accuracy and AUC (ACC = 0.92, AUC = 0.98) and was significantly higher than all the competing methods. Compared with the second best result, our model achieved gains of 4.9% and 3.2% in accuracy and AUC, respectively. In addition, all hypergraph-based methods achieved better performance than the classical GCN and its variants, indicating the advantages of representing brain FCs with hypergraph structures. The approaches which considered hyperedge weights optimization (wHGNN and *dw*HGCN) exhibited better performance than the HGNN that only considered unweighted hypergraph in all classification tasks, indicating the importance of the hyperedge weights. The superior performance against wHGCN illustrated that the flexible learnable hyperedge weights during training further improved the characterization of the brain FC.

The receiver operating characteristic curve (ROC) plots are presented in Fig. 3.(A), where a significantly higher AUC can be observed.

Table 4
Comparison of model performance on WRAT classification task.

	WRAT classification with nback fMRI				
	Method	ACC	SEN	SPE	AUC
Graph based models	TAG	0.6084 ± 0.0253	0.6494 ± 0.0363	0.5713 ± 0.0360	0.6574 ± 0.0276
	SAGE	0.6427 ± 0.0258	0.4564 ± 0.0395	**0.8134 ± 0.0295**	0.7099 ± 0.0274
	GCN	0.6432 ± 0.0262	0.5202 ± 0.0361	0.7691 ± 0.0334	0.7094 ± 0.0279
	CHEB	0.6448 ± 0.0249	0.7623 ± 0.0305	0.5131 ± 0.0373	0.6876 ± 0.0283
	GAT	0.6967 ± 0.0246	0.7147 ± 0.0352	0.6815 ± 0.0340	0.7149 ± 0.0276
Hypergraph based models	HGSVM-L	0.6430 ± 0.0253	0.5833 ± 0.0381	0.6972 ± 0.0333	0.7168 ± 0.0271
	HGSVM-R	0.6511 ± 0.0245	0.5978 ± 0.0371	0.6995 ± 0.0332	0.6960 ± 0.0273
	HGNN	0.6531 ± 0.0238	0.6975 ± 0.0330	0.6051 ± 0.0369	0.7150 ± 0.0268
	wHGNN	0.6851 ± 0.0241	0.7314 ± 0.0317	0.6301 ± 0.0379	0.7648 ± 0.0238
	***dw*HGCN (ours)**	**0.7149 ± 0.0245**	**0.7597 ± 0.0314**	0.6614 ± 0.0375	**0.7760 ± 0.0242**
	WRAT classification with emoid fMRI				
	Method	ACC	SEN	SPE	AUC
Graph based models	GCN	0.6203 ± 0.0256	0.5591 ± 0.0373	0.6846 ± 0.0345	0.6606 ± 0.0281
	SAGE	0.6276 ± 0.0266	0.6232 ± 0.0396	0.6315 ± 0.0345	0.6410 ± 0.0302
	CHEB	0.6371 ± 0.0268	0.6235 ± 0.0375	0.6512 ± 0.0367	0.6828 ± 0.0293
	TAG	0.6428 ± 0.0254	0.6619 ± 0.0356	0.6250 ± 0.0352	0.6885 ± 0.0279
	GAT	**0.6606 ± 0.0268**	0.6185 ± 0.0390	**0.6956 ± 0.0345**	0.6892 ± 0.0291
Hypergraph based models	HGSVM-R	0.6105 ± 0.0252	0.5378 ± 0.0367	0.6772 ± 0.0356	0.6508 ± 0.0287
	HGNN	0.6123 ± 0.0260	0.6701 ± 0.0354	0.5559 ± 0.0374	0.6544 ± 0.0291
	HGSVM-L	0.6191 ± 0.0258	0.5882 ± 0.0392	0.6469 ± 0.0362	0.6369 ± 0.0287
	wHGNN	0.6208 ± 0.0259	0.6612 ± 0.0357	0.5768 ± 0.0378	0.6534 ± 0.0292
	***dw*HGCN (ours)**	0.6485 ± 0.0267	**0.7852 ± 0.0333**	0.5158 ± 0.0374	**0.7173 ± 0.0279**

Table 5
Effects of top-ranked hyperedges on model performance.

Age classification				
Model	ACC	SEN	SPE	AUC
*dw*HGCN (without top-ranked hyperedges)	0.8906 ± 0.0161	0.9134 ± 0.0198	0.8682 ± 0.0243	0.9231 ± 0.0150
***dw*HGCN** (with all hyperedges)	**0.9236 ± 0.0133**	**0.9460 ± 0.0142**	**0.8877 ± 0.0255**	**0.9770 ± 0.0058**

In summary, the proposed *dw*HGCN method outperformed the other competing graph and hypergraph based neural networks and demonstrated its advantages in facilitating the construction of the hypergraph structure with learnable weights considering both discriminative power and smoothness condition.

3.5. WRAT classification

We further evaluated the proposed method on a more challenging task to classify WRAT score groups for individuals. The experiments were based on two task based fMRI paradigms and we presented the classification performance comparison with competing methods in Table 4. Our model achieved the best performance in both accuracy and AUC (ACC = 0.71, AUC = 0.78) in nback paradigm and the second best result in accuracy and highest AUC in emoid paradigm (ACC = 0.65, AUC = 0.72). For working memory task paradigm, our model observed a significant improvement over all competing methods with gains of 1.8% and 1.1% from the second best results in accuracy and AUC. For emotion identification task paradigm, our proposed method obtained a significant improvement in accuracy over competing methods except for the GAT. The GAT exhibited a better performance in accuracy for emoid task and relatively stable performance among all three imaging paradigms due to its implicit assignments of importances to neighbor nodes during message passing, which, on the contrary, made it more challenging to interpret. The conclusions obtained from age classification demonstrating the advantages of the flexible dynamically learnable hyperedge weights still held for WRAT classification, which implied that our model can have a robust performance through multiple classification jobs.

The ROC plots are presented in Fig. 3.(B)(C), where a higher AUC over competing methods was observed. We also implemented the T-Distributed Stochastic Neighbor embedding (t-SNE) to visualize the embedded features at the last convolutional layer for four typical methods (GCN, HGNN, wHGNN, *dw*HGCN) in Fig. 4. Compared with other dimension reduction tools such as Principal Component Analysis (PCA), t-SNE is a nonlinear dimension reduction technique that better preserves the local geometry of embedded features, therefore it produces greater improvement in clustering quality with better visualization (Anowar et al., 2021; Devassy and George, 2020). The hypergraph embeddings learned from our model (last column) appeared to be more separable, compared to other peer models (1st to 3rd columns). This indicated that our proposed model was able to reveal better discriminative feature representation of the brain FC and yielded better classification performance.

3.6. Dynamic hyperedge weights estimation

To better illustrate the advantages of our *dw*HGCN framework, we explored how the weights of hyperedge were dynamically updated. As an example, we illustrated the heatmaps of hyperedges' weights along the training epochs in the age classification task. The standardized hyperedge weights after the epoch 1, 20, 40, 60, and 80 epochs are visualized in Fig. 5. In contrast to prior strategies (i.e., assigning equal weights to all hyperedges), our *dw*HGCN assigned larger weights to the hyperedges with higher task-specific discriminative power, while shrank weights of the hyperedges with less contribution to the classification task close to zero. These weights controlled the higher-order brain activation where task-specific information was more likely to interact among ROIs within hyperedges with larger weights.

We also performed the classification task to evaluate the effects of hyperedges quantitatively. Specifically, we removed the top 5 ranked hyperedge weights by setting their weights to zero. The model performance was displayed in Table 5 and Fig. 6. Reduced performance was observed when top-ranked weights were removed, demonstrating that our method was able to effectively learn weights of hyperedges, and assign larger weights to hyperedges with higher discriminative power.

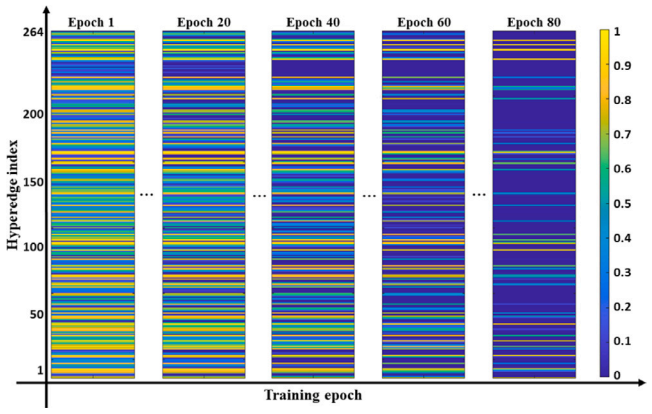

Fig. 5. The learned hyperedge weights at different training epochs (age classification task). The model tends to assign larger or smaller weights to hyperedges with higher or lower contributions for the classification.

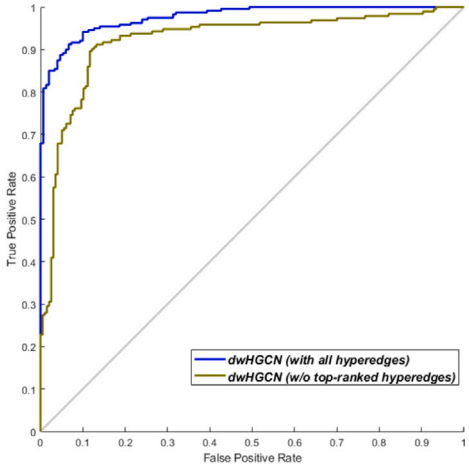

Fig. 6. Receiver operating characteristic curves of the proposed dwHGCN method with all hyperedges and with hyperedges excluding the top 5 ranked hyperedges.

3.7. Most discriminative regions

Adolescence was a crucial period between late childhood and adulthood, during which a fundamental reorganization of the brain was observed (Konrad et al., 2013). Besides improvement in classification performance, our proposed *dw*HGCN framework was able to detect highly relevant biomarkers related to brain maturation and development of intelligence reflected by the learning ability measured with WRAT. We reported the class specific hub regions identified from Grad-CAM approach (Method Section 2.5) in this subsection. We visualized the locations of the identified class-specific ROIs in Fig. 7.

For brain development, we identified the cingulate gyrus in both child and young adult groups, which has been reported to be a crucial region related to cerebral maturation regulating cognitive, social, and emotional development during puberty (Kelly et al., 2009; Bush et al., 2000). The identified anterior cingulate cortex, a hub involved in guidance and support for self-regulation and its adaptive development, has been validated as a critical component of adolescent developmental psychopathology (Lichenstein et al., 2016). Within the temporal lobe, the identified left inferior temporal gyrus was a higher order association area that showed a characteristic later maturation pattern until adulthood (Gogtay et al., 2004). The fusiform gyrus was also detected from both age groups, which was consistent with the findings that fusiform gyrus began to mature around 5–7 years and keep developing until young adulthood as measured by gray matter volume (Haist and Anzures, 2017). The detected middle temporal gyrus showed significant gray matter loss during adolescence (Xu et al., 2022), which was also identified by our previous study using the same PNC dataset (Wang et al., 2020).

For the WRAT classification task, most of the brain regions were identified within the parietal and frontal cortices, which have long been considered and reported to be crucial regions correlated with intelligence (Jung and Haier, 2007). For example, the identified inferior parietal lobule was functionally connected with cerebellar Crus I and II and formulated the cerebello-parietal component, which was reported to be highly associated with general intelligence (Van Den Heuvel et al., 2009; Yoon et al., 2017). The identified precuneus on the parietal lobe has also been reported to have a positive association with full-scale IQ (Van Den Heuvel et al., 2009). The superior frontal gyrus and middle frontal gyrus have been significantly related to working memory capability that was highly correlated with general intelligence (Colom et al., 2007). The identified medial frontal gyrus has been reported to be functionally actively activated during intelligence related studies and the finding was supported by the structurally based meta-analyses (Basten et al., 2015). Besides identified ROIs on frontal and parietal cortex, we identified anterior cingulate cortex on the limbic lobe, which has been reported to be involved in multiple cognitive control tasks and have a significant positive impact on intelligence (Hilger et al., 2017; Kozlovskiy et al., 2012). Our findings were highly consistent with the existing knowledge about brain maturation and intelligence. Most of the identified brain regions were highly correlated with their brain functions and can serve as potential biomarkers.

4. Discussion

In this study, we propose a *dw*HGCN framework and evaluated the model on a large-scale public dataset for brain FC classification. Unlike existing graph and hypergraph based neural network models (Kipf and Welling, 2017; Feng et al., 2019), we propose a novel *dw*HGC layer that is able to dynamically estimate the feature representation of the brain FC modeled with a weighted hypergraph. The *dw*HGC layer will automatically assign larger weights to hyperedges with higher discriminative power during model training, which improves the characterization of the brain FC. In the feed-forward procedure, weighted hypergraph and node features are input into the proposed *dw*HGCN model to learn weighted hypergraph embeddings, which are fed to fully connected and softmax layers for classification. In the back-propagation procedure, both the hyperedge weights and *dw*HGCN model weights are optimized simultaneously. The experimental results demonstrate that the introduction of dynamic hypergraph further enhanced the feature representation learning compared with the static weights (performance between wHGNN and *dw*HGCN in both classifications tasks).

We also propose an objective function for hypergraph classification with a manifold regularization to regulate the learnable hyperedge weights based on the assumption that the fMRI time series signals should be smooth on the hypergraph. We quantify the smoothness by the hypergraph Laplacian quadratic form and obtained a neurologically meaningful hyper network accordingly. We perform an ablation study to compare the classification performance without the dynamic hyperedge weights or the manifold regularization term in the objective function. The results (Fig. 8) demonstrate the benefits of incorporating manifold regularization and dynamic hypergraph learning with improved classification performance. The identified age-related anterior cingulate cortex has proven to be a crucial component in brain development during adolescence (Lichenstein et al., 2016). Therefore,

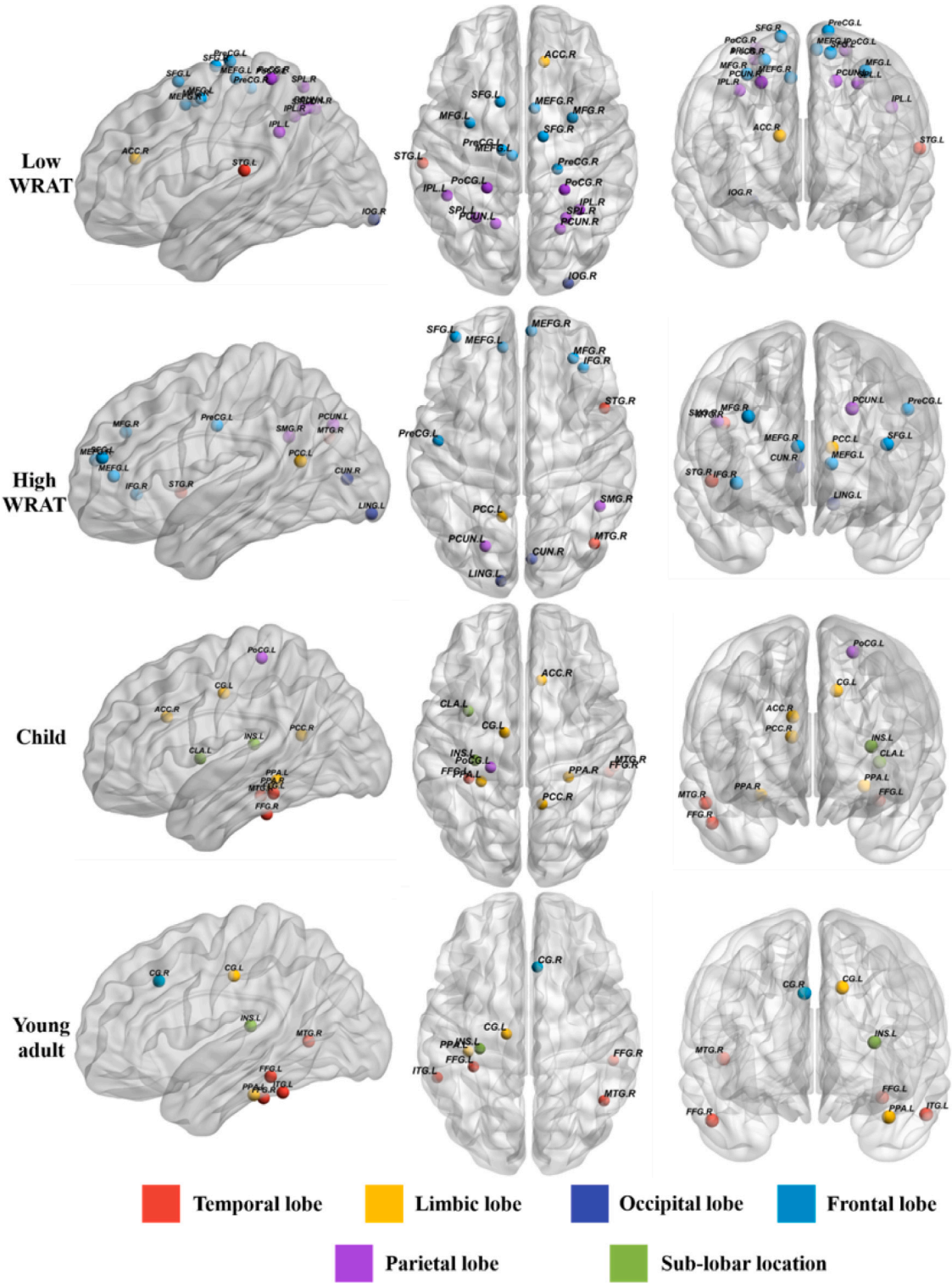

Fig. 7. Visualization of the identified class-specific ROIs. The abbreviations: R=Right, L=Left, PoCG=postcentral gyrus, PCUN=precuneus, IPL=inferior parietal lobule, SPL=superior parietal lobule, SMG=supramarginal gyrus, PreCG=precentral gyrus, SFG=superior frontal gyrus, MFG=middle frontal gyrus, MEFG=medial frontal gyrus, IFG=inferior frontal gyrus, INS=insula, CLA=claustrum, PPA=parahippocampal gyrus, ACC=anterior cingulate cortex, PCC=posterior cingulate cortex, IOG=inferior occipital gyrus, LING=lingual gyrus, CUN=cuneus, FFG=fusiform gyrus, STG=superior temporal gyrus, MTG=middle temporal gyrus, ITG=inferior temporal gyrus.

our proposed model could benefit the brain studies by uncovering meaningful findings.

We evaluate the proposed method by comparing the classification performance with several relevant graph/hypergraph based models. Our proposed method achieves the best performance for both classification tasks in three paradigms fMRI (except for emoid fMRI, we achieve the second best in accuracy and highest AUC) over all competing models. It is intriguing to observe that HGSVM-R outperformed the wHGNN model in the age classification task (Table 3). One potential reason is that the task complexity of the age classification is comparatively low. Thus, the classical machine learning models may be able to handle such classification tasks better than complex deep learning.

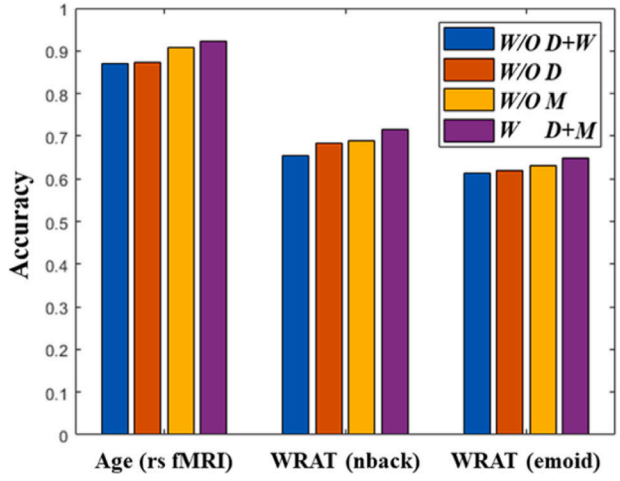

Fig. 8. Comparison of classification performance for three imaging paradigms. The legends from top to bottom represent: without dynamic update of hypergraph and hyperedge weights (*W/O D+W*); without dynamic update of hypergraph but with hyperedge weights (*W/O D*); with dynamic learning of hypergraph but without the manifold regularization (*W/O M*); with both dynamic learning of hypergraph and manifold regularization (*W D+M*).

Table 6
The number of parameters/training time (s) per epoch (batchsize 32) of various graph/hypergraph models.

	Age (rs fMRI)	WRAT (emoid)	WRAT (nback)
GCN	10.66k/10.93 s	11.71k/13.71 s	11.71k/13.86 s
CHEB	20.13k/19.11 s	22.21k/22.64 s	22.21k/20.08 s
SAGE	20.13k/12.29 s	22.21k/14.07 s	22.21k/14.05 s
TAG	39.07k/54.10 s	43.20k/62.51 s	43.20k/61.38 s
GAT	10.79k/26.49 s	11.91k/35.62 s	11.91k/36.30 s
(w)HGNN	10.66k/11.14 s	11.71k/14.93 s	11.71k/15.06 s
*dw*HGCN	10.92k/13.45 s	11.98k/18.29 s	11.98k/18.71 s

In constrast, wHGCN performed better than linear and nonlinear SVM models in the more challenging WRAT classification task, which is apparently more difficult than the age classification task (Table 4). This is consistent with prior studies (Korotcov et al., 2017; Hu et al., 2021; Liu and Lang, 2019) showing that models with higher complexity (deep learning models) tend to perform better on difficult tasks, while those lightweight models with lower complexity (e.g., SVM) may be a better choice on tasks with lower complexity. We compare the speed of the proposed method with that of the competing graph/hypergraph based neural network models and provide the number of parameters and training time for each epoch in Table 6. The *dw*HGCN has a slightly higher training time compared with regular GCN and HGNN. The major computation cost of *dw*HGCN comes from two parts: (1) The model's training requires additional gradient calculation on the hyperedge weights vector with smoothness constraint to regulate weights optimization; (2) We need to recalculate the hypergraph Laplacian matrix after dynamic estimation of the hyperedge weights.

In addition to improvement in classification performance, our *dw*HGCN framework is capable of identifying task specific brain region biomarkers. Specially, we apply the Grad-CAM method on the estimated hypergraph and calculate the gradient of each class at the last *dw*HGC layer with respect to the feature maps. We formulate a contribution vector by averaging the gradient maps among all the subjects. Our findings suggest that the most identified ROIs are highly associated with the corresponding brain functions related to brain maturation and development of intelligence during adolescence and are consistent with the existing knowledges and previous findings. For example, the identified medial frontal gyrus from WRAT classification task has been validated from both functional and anatomical neuroimaging studies (Basten et al., 2015).

Our study has some limitations. Since the focus of this paper is to demonstrate our proposed method's capability in representation learning and advantage on interpretation rather than achieving optimal classification performance, we only applied l_1 sparse representation to construct the hypergraph. To accurately reflect the higher order relationships between ROIs, the hypergraph generation is a crucial step for *dw*HGCN to output a precise decision. One potential future direction is to explore and develop more advanced hypergraph construction methods that could further improve classification performance. Another potential limitation is that we only work on a single fMRI modality for individual classification tasks. In fact, one of the benefits of hypergraph is its flexibility to represent data from multi-modality imaging, which can be achieved by simply combining the incidence matrix. In the future, we may investigate applications of the proposed model on multi-modal MRI imaging datasets.

5. Conclusion

In this paper, we propose a dynamic weighted hypergraph convolutional network model by integrating brain FC learning and neural network model training into a unified end-to-end framework. Our model considers the weighted hypergraph as inputs and generated weighted hyperedges embeddings. Unlike existing hypergraph neural network models (Feng et al., 2019; Bai et al., 2021), the proposed method is able to dynamically update hyperedge weights of a hypergraph structure, which facilitates characterization of the activation among brain regions. We validated the performance of the proposed model on two classification tasks (i.e., age and WRAT) using three paradigms of fMRI data. The experimental results demonstrate the superiority of our proposed method over traditional GCN and its variants and existing hypergraph neural networks. Apart from improvement in classification, our proposed method offers a reliable interpretation of its decisions by dynamically weighting hyperedges in the brain FC, based on which we are able to identify the hub regions related to brain development and learning ability. The identified regions are highly relevant and consistent with previous reports, indicating its potentials in many other neuroimaging applications.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

This work was supported in part by NIH, United States (R01 EB029944, R01 EB030582, R01 NS094200, and R01 NS096037) and Academic and Research Committee Awards of Cincinnati Children's Hospital Medical Center.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.media.2023.102828.

References

Abney, S., 2002. Bootstrapping. In: Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics. pp. 360–367.

Allen, E.J., St-Yves, G., Wu, Y., Breedlove, J.L., Prince, J.S., Dowdle, L.T., Nau, M., Caron, B., Pestilli, F., Charest, I., et al., 2022. A massive 7T fMRI dataset to bridge cognitive neuroscience and artificial intelligence. Nature Neurosci. 25 (1), 116–126.

Anowar, F., Sadaoui, S., Selim, B., 2021. Conceptual and empirical comparison of dimensionality reduction algorithms (pca, kpca, lda, mds, svd, lle, isomap, le, ica, t-sne). Comp. Sci. Rev. 40, 100378.

Ashburner, J., Barnes, G., Chen, C., Daunizeau, J., Flandin, G., Friston, K., Kiebel, S., Kilner, J., Litvak, V., Moran, R., et al., 2014. SPM12 Manual. Wellcome Trust Centre for Neuroimaging, London, UK.

Bai, S., Zhang, F., Torr, P.H., 2021. Hypergraph convolution and hypergraph attention. Pattern Recognit. 110, 107637.

Bassett, D.S., Bullmore, E., Verchinski, B.A., Mattay, V.S., Weinberger, D.R., Meyer-Lindenberg, A., 2008. Hierarchical organization of human cortical networks in health and schizophrenia. J. Neurosci. 28 (37), 9239–9248.

Basten, U., Hilger, K., Fiebach, C.J., 2015. Where smart brains are different: A quantitative meta-analysis of functional and structural brain imaging studies on intelligence. Intelligence 51, 10–27.

Bi, X.-a., Li, L., Wang, Z., Wang, Y., Luo, X., Xu, L., 2022. IHGC-GAN: influence hypergraph convolutional generative adversarial network for risk prediction of late mild cognitive impairment based on imaging genetic data. Brief. Bioinform. 23 (3), bbac093.

Bullmore, E.T., Bassett, D.S., 2011. Brain graphs: graphical models of the human brain connectome. Annu. Rev. Clin. Psychol. 7, 113–140.

Bullmore, E., Sporns, O., 2009. Complex brain networks: graph theoretical analysis of structural and functional systems. Nat. Rev. Neurosci. 10 (3), 186–198.

Bush, G., Luu, P., Posner, M.I., 2000. Cognitive and emotional influences in anterior cingulate cortex. Trends in Cognitive Sciences 4 (6), 215–222.

Colom, R., Jung, R.E., Haier, R.J., 2007. General intelligence and memory span: evidence for a common neuroanatomic framework. Cogn. Neuropsychol. 24 (8), 867–878.

Davison, E.N., Schlesinger, K.J., Bassett, D.S., Lynall, M.-E., Miller, M.B., Grafton, S.T., Carlson, J.M., 2015. Brain network adaptability across task states. PLoS Comput. Biol. 11 (1), e1004029.

Defferrard, M., Bresson, X., Vandergheynst, P., 2016. Convolutional neural networks on graphs with fast localized spectral filtering. In: Advances in Neural Information Processing Systems. pp. 3844–3852.

Devassy, B.M., George, S., 2020. Dimensionality reduction and visualisation of hyperspectral ink data using t-SNE. Forensic Sci. Int. 311, 110194.

Dong, X., Thanou, D., Rabbat, M., Frossard, P., 2019. Learning graphs from data: A signal representation perspective. IEEE Signal Process. Mag. 36 (3), 44–63.

Du, M., Zhang, L., Li, L., Ji, E., Han, X., Huang, G., Liang, Z., Shi, L., Yang, H., Zhang, Z., 2021. Abnormal transitions of dynamic functional connectivity states in bipolar disorder: A whole-brain resting-state fMRI study. J. Affect. Disord. 289, 7–15.

Du, J., Zhang, S., Wu, G., Moura, J.M., Kar, S., 2017. Topology adaptive graph convolutional networks. arXiv preprint arXiv:1710.10370.

Eavani, H., Satterthwaite, T.D., Filipovych, R., Gur, R.E., Gur, R.C., Davatzikos, C., 2015. Identifying sparse connectivity patterns in the brain using resting-state fMRI. Neuroimage 105, 286–299.

Feng, Y., You, H., Zhang, Z., Ji, R., Gao, Y., 2019. Hypergraph neural networks. In: Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 33. pp. 3558–3565.

Finn, E.S., Shen, X., Scheinost, D., Rosenberg, M.D., Huang, J., Chun, M.M., Papademetris, X., Constable, R.T., 2015. Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. Nature Neurosci. 18 (11), 1664–1671.

Gao, Y., Zhang, Z., Lin, H., Zhao, X., Du, S., Zou, C., 2020. Hypergraph learning: Methods and practices. IEEE Trans. Pattern Anal. Mach. Intell.

Gogtay, N., Giedd, J.N., Lusk, L., Hayashi, K.M., Greenstein, D., Vaituzis, A.C., Nugent, T.F., Herman, D.H., Clasen, L.S., Toga, A.W., et al., 2004. Dynamic mapping of human cortical development during childhood through early adulthood. Proc. Natl. Acad. Sci. USA 101 (21), 8174–8179.

Gonzalez-Castillo, J., Kam, J.W., Hoy, C.W., Bandettini, P.A., 2021. How to interpret resting-state fMRI: Ask your participants. J. Neurosci. 41 (6), 1130–1141.

Grabner, R.H., Ansari, D., Reishofer, G., Stern, E., Ebner, F., Neuper, C., 2007. Individual differences in mathematical competence predict parietal brain activation during mental calculation. Neuroimage 38 (2), 346–356.

Gu, S., Yang, M., Medaglia, J.D., Gur, R.C., Gur, R.E., Satterthwaite, T.D., Bassett, D.S., 2017. Functional hypergraph uncovers novel covariant structures over neurodevelopment. Hum. Brain Mapp. 38 (8), 3823–3835.

Gur, R.C., Richard, J., Hughett, P., Calkins, M.E., Macy, L., Bilker, W.B., Brensinger, C., Gur, R.E., 2010. A cognitive neuroscience-based computerized battery for efficient measurement of individual differences: standardization and initial construct validation. J. Neurosci. Methods 187 (2), 254–262.

Haist, F., Anzures, G., 2017. Functional development of the brain's face-processing system. Wiley Interdiscip. Rev. Cogn. Sci. 8 (1–2), e1423.

Hamilton, W., Ying, Z., Leskovec, J., 2017. Inductive representation learning on large graphs. Adv. Neural Inf. Process. Syst. 30.

Hilger, K., Ekman, M., Fiebach, C.J., Basten, U., 2017. Efficient hubs in the intelligent brain: Nodal efficiency of hub regions in the salience network is associated with general intelligence. Intelligence 60, 10–25.

Honey, C., Sporns, O., Cammoun, L., Gigandet, X., Thiran, J.P., Meuli, R., Hagmann, P., 2009. Predicting human resting-state functional connectivity from structural connectivity. Proc. Natl. Acad. Sci. 106 (6), 2035–2040.

Hu, X., Chu, L., Pei, J., Liu, W., Bian, J., 2021. Model complexity of deep learning: A survey. Knowl. Inf. Syst. 63, 2585–2619.

Ji, Y., Zhang, Y., Shi, H., Jiao, Z., Wang, S.H., Wang, C., 2021. Constructing dynamic brain functional networks via hyper-graph manifold regularization for mild cognitive impairment classification. Front. Neurosci. 15, 358.

Jiang, J., Wei, Y., Feng, Y., Cao, J., Gao, Y., 2019. Dynamic hypergraph neural networks.. In: IJCAI. pp. 2635–2641.

Jie, B., Wee, C.-Y., Shen, D., Zhang, D., 2016. Hyper-connectivity of functional networks for brain disease diagnosis. Med. Image Anal. 32, 84–100.

Jung, R.E., Haier, R.J., 2007. The Parieto–Frontal integration theory (P-FIT) of intelligence: converging neuroimaging evidence. Behav. Brain Sci. 30 (2), 135–154.

Kelly, A.C., Di Martino, A., Uddin, L.Q., Shehzad, Z., Gee, D.G., Reiss, P.T., Margulies, D.S., Castellanos, F.X., Milham, M.P., 2009. Development of anterior cingulate functional connectivity from late childhood to early adulthood. Cerebral Cortex 19 (3), 640–657.

Kipf, T.N., Welling, M., 2017. Semi-supervised classification with graph convolutional networks. In: International Conference on Learning Representations. https://openreview.net/forum?id=SJU4ayYgl.

Konrad, K., Firk, C., Uhlhaas, P.J., 2013. Brain development during adolescence: neuroscientific insights into this developmental period. Dtsch. Ärzteblatt Int. 110 (25), 425.

Korotcov, A., Tkachenko, V., Russo, D.P., Ekins, S., 2017. Comparison of deep learning with multiple machine learning methods and metrics using diverse drug discovery data sets. Mol. Pharm. 14 (12), 4462–4475.

Kozlovskiy, S.A., Nikonova, E.Y., Pyasik, M.M., Velichkovsky, B.M., et al., 2012. The cingulate cortex and human memory processes. Psychol. Russ. 5, 231.

Li, Y., Liu, J., Gao, X., Jie, B., Kim, M., Yap, P.T., Wee, C.Y., Shen, D., 2019. Multimodal hyper-connectivity of functional networks using functionally-weighted LASSO for MCI classification. Med. Image Anal. 52, 80–96.

Lichenstein, S.D., Verstynen, T., Forbes, E.E., 2016. Adolescent brain development and depression: a case for the importance of connectivity of the anterior cingulate cortex. Neurosci. Biobehav. Rev. 70, 271–287.

Liu, H., Lang, B., 2019. Machine learning and deep learning methods for intrusion detection systems: A survey. Appl. Sci. 9 (20), 4396.

Liu, J., Tan, G., Lan, W., Wang, J., 2020. Identification of early mild cognitive impairment using multi-modal data and graph convolutional networks. BMC Bioinformatics 21 (6), 1–12.

Ma, J., Wang, Y., Liu, B., Liu, W., 2021. Accurately modeling the human brain functional correlations with hypergraph Laplacian. Neurocomputing 428, 239–247.

Madine, M., Rekik, I., Werghi, N., 2020. Diagnosing autism using T1-W MRI with multi-kernel learning and hypergraph neural network. In: 2020 IEEE International Conference on Image Processing. ICIP, IEEE, pp. 438–442.

McHugh, M.L., 2013. The chi-square test of independence. Biochem. Med. 23 (2), 143–149.

Newman, S.D., Carpenter, P.A., Varma, S., Just, M.A., 2003. Frontal and parietal participation in problem solving in the tower of London: fMRI and computational modeling of planning and high-level perception. Neuropsychologia 41 (12), 1668–1682.

Parisot, S., Ktena, S.I., Ferrante, E., Lee, M., Guerrero, R., Glocker, B., Rueckert, D., 2018. Disease prediction using graph convolutional networks: application to autism spectrum disorder and Alzheimer's disease. Med. Image Anal. 48, 117–130.

Pope, P.E., Kolouri, S., Rostami, M., Martin, C.E., Hoffmann, H., 2019. Explainability methods for graph convolutional neural networks. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 10772–10781.

Power, J.D., Cohen, A.L., Nelson, S.M., Wig, G.S., Barnes, K.A., Church, J.A., Vogel, A.C., Laumann, T.O., Miezin, F.M., Schlaggar, B.L., et al., 2011. Functional network organization of the human brain. Neuron 72 (4), 665–678.

Power, J.D., Schlaggar, B.L., Petersen, S.E., 2014. Studying brain organization via spontaneous fMRI signal. Neuron 84 (4), 681–696.

Qu, G., Hu, W., Xiao, L., Wang, J., Bai, Y., Patel, B., Zhang, K., Wang, Y.-P., 2021a. Brain functional connectivity analysis via graphical deep learning. IEEE Trans. Biomed. Eng. 69 (5), 1696–1706.

Qu, G., Xiao, L., Hu, W., Wang, J., Zhang, K., Calhoun, V.D., Wang, Y.-P., 2021b. Ensemble manifold regularized multi-modal graph convolutional network for cognitive ability prediction. IEEE Trans. Biomed. Eng. 68 (12), 3564–3573.

Ranstam, J., Cook, J., 2018. LASSO regression. J. Br. Surg. 105 (10), 1348.

Rubinov, M., Sporns, O., 2010. Complex network measures of brain connectivity: uses and interpretations. Neuroimage 52 (3), 1059–1069.

Satterthwaite, T.D., Elliott, M.A., Ruparel, K., Loughead, J., Prabhakaran, K., Calkins, M.E., Hopson, R., Jackson, C., Keefe, J., Riley, M., et al., 2014. Neuroimaging of the Philadelphia neurodevelopmental cohort. Neuroimage 86, 544–553.

Scarselli, F., Gori, M., Tsoi, A.C., Hagenbuchner, M., Monfardini, G., 2008. The graph neural network model. IEEE Trans. Neural Netw. 20 (1), 61–80.

Selvaraju, R.R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., Batra, D., 2017. Grad-cam: Visual explanations from deep networks via gradient-based localization. In: Proceedings of the IEEE International Conference on Computer Vision. pp. 618–626.

Shi, H., Zhang, Y., Zhang, Z., Ma, N., Zhao, X., Gao, Y., Sun, J., 2018. Hypergraph-induced convolutional networks for visual classification. IEEE Trans. Neural Netw. Learn. Syst. 30 (10), 2963–2972.

Stam, C.J., Jones, B., Nolte, G., Breakspear, M., Scheltens, P., 2006. Small-world networks and functional connectivity in Alzheimer's disease. Cerebral Cortex 17 (1), 92–99.

Van Den Heuvel, M.P., Stam, C.J., Kahn, R.S., Pol, H.E.H., 2009. Efficiency of functional brain networks and intellectual performance. J. Neurosci. 29 (23), 7619–7624.

Velickovic, P., Cucurull, G., Casanova, A., Romero, A., Lio, P., Bengio, Y., 2017. Graph attention networks. Stat 1050, 20.

Wang, J., Calhoun, V.D., Stephen, J.M., Wilson, T.W., Wang, Y.p., 2018. Integration of network topological features and graph Fourier transform for fMRI data analysis. In: 2018 IEEE 15th International Symposium on Biomedical Imaging. ISBI 2018, IEEE, pp. 92–96.

Wang, J., Xiao, L., Hu, W., Qu, G., Wilson, T.W., Stephen, J.M., Calhoun, V.D., Wang, Y.P., 2021. Functional network estimation using multigraph learning with application to brain maturation study. Hum. Brain Mapp. 42 (9), 2880–2892.

Wang, J., Xiao, L., Wilson, T.W., Stephen, J.M., Calhoun, V.D., Wang, Y.P., 2020. Examining brain maturation during adolescence using graph Laplacian learning based Fourier transform. J. Neurosci. Methods 108649.

Wilkinson, G.S., Robertson, G.J., 2006. Wide Range Achievement Test (WRAT4). Psychological Assessment Resources, Lutz, FL.

Xiao, L., Wang, J., Kassani, P.H., Zhang, Y., Bai, Y., Stephen, J.M., Wilson, T.W., Calhoun, V.D., Wang, Y.P., 2019. Multi-hypergraph learning based brain functional connectivity analysis in fMRI data. IEEE Trans. Med. Imaging.

Xing, X., Li, Q., Wei, H., Zhang, M., Zhan, Y., Zhou, X.S., Xue, Z., Shi, F., 2019. Dynamic spectral graph convolution networks with assistant task training for early mci diagnosis. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 639–646.

Xu, K., Hu, W., Leskovec, J., Jegelka, S., 2019. How powerful are graph neural networks?. In: International Conference on Learning Representations. https://openreview.net/forum?id=ryGs6iA5Km.

Xu, J., Zhang, J., Li, J., Wang, H., Chen, J., Lyu, H., Hu, Q., 2022. Structural and functional trajectories of middle temporal gyrus sub-regions during life span: A potential biomarker of brain development and aging. Front. Aging Neurosci. 14.

Yao, D., Sui, J., Wang, M., Yang, E., Jiaerken, Y., Luo, N., Yap, P.T., Liu, M., Shen, D., 2021. A mutual multi-scale triplet graph convolutional network for classification of brain disorders using functional or structural connectivity. IEEE Trans. Med. Imaging 40 (4), 1279–1289.

Yin, W., Li, L., Wu, F.-X., 2022. Deep learning for brain disorder diagnosis based on fMRI images. Neurocomputing 469, 332–345.

Yoon, Y.B., Shin, W.G., Lee, T.Y., Hur, J.W., Cho, K.I.K., Sohn, W.S., Kim, S.G., Lee, K.H., Kwon, J.S., 2017. Brain structural networks associated with intelligence and visuomotor ability. Sci. Rep. 7 (1), 1–9.

Zhou, C., Zemanová, L., Zamora, G., Hilgetag, C.C., Kurths, J., 2006. Hierarchical organization unveiled by functional connectivity in complex brain networks. Phys. Rev. Lett. 97 (23), 238103.

Zille, P., Calhoun, V.D., Stephen, J.M., Wilson, T.W., Wang, Y.P., 2018. Fused estimation of sparse connectivity patterns from rest fMRI application to comparison of children and adult brains. IEEE Trans. Med. Imaging 37 (10), 2165–2175.

Zu, C., Gao, Y., Munsell, B., Kim, M., Peng, Z., Zhu, Y., Gao, W., Zhang, D., Shen, D., Wu, G., 2016. Identifying high order brain connectome biomarkers via learning on hypergraph. In: International Workshop on Machine Learning in Medical Imaging. Springer, pp. 1–9.