SPATIAL TEMPORAL GRAPH CONVOLUTION WITH GRAPH STRUCTURE SELF-LEARNING FOR EARLY MCI DETECTION

Yunpeng Zhao, Fugen Zhou, Bin Guo*, Bo Liu*

Image Processing Center, Beihang University, Beijing, China

ABSTRACT

Graph neural networks (GNNs) have been successfully applied to early mild cognitive impairment (EMCI) detection, with the usage of elaborately designed features constructed from blood oxygen level-dependent (BOLD) time series. However, few works explored the feasibility of using BOLD signals directly as features. Meanwhile, existing GNN-based methods primarily rely on hand-crafted explicit brain topology as the adjacency matrix, which is not optimal and ignores the implicit topological organization of the brain. In this paper, we propose a spatial temporal graph convolutional network with a novel graph structure self-learning mechanism for EMCI detection. The proposed spatial temporal graph convolution block directly exploits BOLD time series as input features, which provides an interesting view for rsfMRI-based preclinical AD diagnosis. Moreover, our model can adaptively learn the optimal topological structure and refine edge weights with the graph structure self-learning mechanism. Results on the Alzheimer’s Disease Neuroimaging Initiative (ADNI) database show that our method outperforms state-of-the-art approaches. Biomarkers consistent with previous studies can be extracted from the model, proving the reliable interpretability of our method.

Index Terms— Alzheimer’s Disease, Spatial Temporal Graph Convolution, Self-learned Graph Structure

1. INTRODUCTION

Alzheimer’s Disease (AD) is a cosmopolitan neurodegenerative disease in the brain and accounts for an estimated 60%-80% of dementia patients [1]. Progression of this irreversible and incurable disease is concomitant with gradually impaired cognitive skills, memory, and language abilities, leading to the escalation of potential social burden. Early mild cognitive impairment (EMCI) is a prodromal stage of AD with a high conversion rate [2]. Therefore, screening for EMCI is crucial for relieving the deterioration of AD.

EMCI detection is still a challenging task. In light of trivial structural differences between EMCI and normal control (NC) brains [3], recent studies seek to identify EMCI based on resting-state functional MRI (rs-fMRI), a non-invasive imaging technique that measures blood oxygen level-dependent (BOLD) signals [4]. According to related works, a traditional practice for EMCI screening with rs-fMRI is constructing features from BOLD signals and then applying machine learning methods to detect abnormal patterns. Commonly-used features include the whole-brain functional connectivity (FC) matrices, dynamic FC (dFC), and dynamic effective connectivity (dEC) extracted by group-constrained Kalman filter (gKF) algorithm [2, 3, 5]. However, existing works overlooked the rich diagnostic information recorded in the spatial temporal dependency of original BOLD signals and have not explored the feasibility of building an end-to-end model that uses BOLD time series as features directly.

In fact, in rs-fMRI, spatially segregated brain regions are functionally connected with different weights and concurrently fluctuate along the temporal dimension, making it feasible to apply the graph neural networks (GNNs) to these data for classification [6, 7]. For a GNN-based model, it is essential to construct a prior graph topological structure. Most existing methods tend to design a hand-crafted adjacency matrix that is consistent during training [8, 9]. However, these approaches only focus on learning with the explicit prior brain structure, which may be non-optimal and neglect implicit inter-region connections that contribute to EMCI detection. The ignorance of potential information from the implicit graph structure tends to yield incomplete brain structure modeling and limit the performance.

To overcome these limitations, we propose a novel spatial temporal graph convolutional network with graph structure self-learning mechanism for EMCI screening. Our contributions are three folds: 1) Our spatial temporal model can directly exploit BOLD time series as input features for EMCI detection by excavating spatial temporal dependencies in signals, which provides a new perspective for rsfMRI-based preclinical AD diagnosis. 2) To the best of our knowledge, our model is the first end-to-end GNN-based framework that can adaptively learn the optimal spatial dependency structure for EMCI detection. 3) The interpretability analysis of the self-learned graph topology could identify biomarkers consistent with existing neuroscience literature.
2. MATERIALS AND METHODOLOGY

The overall architecture of our framework is shown in Fig.1. We first construct spatial temporal graphs from BOLD time series. Thereafter spatial temporal graph convolution (ST-GC) blocks are exploited to aggregate information spatially and temporally to generate a discriminative feature map. With our graph structure self-learning mechanism, the ST-GC blocks can dynamically learn the optimal task-aware brain topological structure and layer-wise edge weights. Graph average pooling is then used for graph readout to obtain the feature vector, which is fed into a fully connected layer for graph classification.

2.1. Dataset and Preprocessing

The neuroimaging data used in our study are obtained from the Alzheimer’s Disease Neuroimaging Initiative (ADNI) database [10]. Totally 146 age-gender-matched subjects are involved. All subjects have both rs-fMRI and T1 weighted images, which are acquired by 3 Tesla scanners from different manufacturers, including Philips, Siemens, and General Electric (GE). Each subject has only one session. The demographic information of selected subjects is shown in Table 1.

| Group       | NC  | EMCI |
|-------------|-----|------|
| Gender (Male/Female) | 30/43 | 34/39 |
| Age (Mean±SD)  | 72.26±6.96 | 72.64±6.70 |

We utilize a widely-adopted DPARSF toolbox for image preprocessing with the pipeline similar to [3]: 1) slice timing; 2) head motion correction; 3) coregister T1 images to functional images; 4) regress out nuisance covariates, including mean white matter and cerebrospinal fluid signals; 5) band-pass filtering (0.01 – 0.1 Hz); 6) spatial normalization.

Time series were extracted from 116 ROIs defined by Automated Anatomical Labeling (AAL) template. For consistency, only the first 140 time points were kept (since time points were different across scanners) and further z-scored, resulting in a final matrix of size 116 × 140 for each subject.

2.2. Spatial Temporal Graph Construction

For ROI-wise average BOLD time series \( f \in \mathbb{R}^{N \times Z} \) with \( N \) ROIs and \( Z \) time points, it can naturally be represented by an undirected spatial temporal graph \( \mathcal{G} = (\mathcal{V}, \mathcal{E}) \) featuring both intra-brain connections and time series fluctuations. In the node set \( \mathcal{V} = \{v_{ti}|t = 1, ..., Z; i = 1, ..., N\} \), the \( i \)-th ROI at the \( t \)-th time point define a vertex \( v_{ti} \), whose feature \( F(v_{ti}) \) is the magnitude of the BOLD signal at \( v_{ti} \). Therefore, the feature map of the spatial temporal graph is \( F \in \mathbb{R}^{C \times N \times Z} \), where \( C = 1 \) denotes the dimension of node features. Meanwhile, the edge set \( \mathcal{E} \) describes spatial and temporal connections between the nodes. For the spatial graph, we build edges for brain regions by randomly initializing the adjacency matrix, which is described in Sec.2.4. The optimal spatial graph structure will be learned by the graph structure self-learning mechanism afterward. For the temporal graph, each ROI is linked to the same ROI in the next time point.

However, the spatial temporal dependencies are not stationary during the entire scan, and the pattern reveals considerable fluctuations. Inspired by sliding window-based dFC [3], our model is trained on short sequences temporally sampled with window size \( T \) to model such dynamics. Specifically, at each training iteration, we randomly sample a subsequence of length \( T \) from the full BOLD time series of each subject in the mini-batch. At the testing stage, we apply the trained model to slices of each testing subject, and then average the sigmoid values to produce subject-level predictions.

2.3. Spatial Temporal Feature Extraction

We define a ST-GC block consisting of a graph convolution (GC) layer followed by a specially designed temporal inception module to generate discriminative spatial temporal representations of graph-structured BOLD signals. The GC layer aggregates spatial information at each time point according to the graph topology. Then the temporal inception module learns temporal features for each brain region independently. We concatenate several ST-GC blocks to extract spatial temporal features.

Spatial Graph Convolution. Noting that the spatial dependency is graph-structured, we employ the graph convolution operation in [11] to fuse information between nodes and their neighbors at each time point. Given the adjacency matrix \( A \in \mathbb{R}^{N \times N} \) and the \( H_1 \)-channel input feature map \( X^{(1)}_{1:N \times T} \in \mathbb{R}^{H_1 \times N \times T} \), the spatial convolution at the \( t \)-th time point in the \( l \)-th ST-GC block is defined as

\[
\tilde{X}^{(l)}_{1:N \times t} = D^{-\frac{1}{2}} A D^{-\frac{1}{2}} X^{(l)}_{1:N \times t} W^{(l)},
\]

where \( \tilde{X}^{(l)}_{1:N \times t} \in \mathbb{R}^{H_2 \times N} \) denotes the \( H_2 \)-channel output feature map at time point \( t \) after the spatial graph convolution, \( D \) is the degree matrix of \( A \), and \( W^{(l)} \in \mathbb{R}^{H_1 \times H_2} \) is a trainable weight matrix. The spatial graph convolution is performed at every independent time point, through which information of brain regions is spatially communicated.
Temporal Inception Module. Different from the spatial graph, the temporal graph possesses a grid structure. Therefore, it is feasible to adopt a standard 1D convolution for the temporal graph convolution. However, the kernel size could be too large or too small to excavate long-term and short-term patterns simultaneously. To further improve the temporal representation ability, we design a 1D temporal inception module inspired by [12]. The architecture is shown in Fig.2. We perform such temporal convolution on each brain region independently. Formally, let the features after spatial graph convolution in the \(l\)-th ST-GC block \(\tilde{X}_{1:N,1:T}^{(l)} \in \mathbb{R}^{H_2 \times N \times T}\) be the input feature map, our temporal graph convolution for the \(i\)-th ROI can be represented as

\[
X_{i,1:T}^{(l+1)} = \text{ReLU}(\text{TIM}_\Omega \tilde{X}_{i,1:T}^{(l)}),
\]

where \(X_{i,1:T}^{(l+1)} \in \mathbb{R}^{H_2 \times T}\) denotes the \(H_2\)-channel output of the \(i\)-th ROI after the temporal graph convolution, \(\text{TIM}_\Omega(\cdot)\) indicates our temporal inception module with \(\Omega\) representing its trainable parameters.

![Fig. 2. Architecture of our temporal inception module.](image)

2.4. Self-learned Graph Structure for EMCI Detection

The graph structure self-learning mechanism learns the task-aware adjacency matrix adaptively to capture the optimal spatial dependencies and adjusts aggregation weights independently for each layer.

Self-learned Graph Structure. Existing GNN-based models for EMCI screening usually exploit a constant adjacency matrix and disregard the fact that the graph topology may be non-optimal. Therefore, we design a dynamic adjacency matrix that is updated during training along with the model. Since the interactions between brain regions are mutual, the binary adjacency matrix is supposed to be symmetric. To reduce the computational complexity, we leverage a trainable vector \(\Theta \in \mathbb{R}^{N \times (N+1)/2}\) to describe a lower triangular matrix, which is a compressed representation of a learnable symmetric matrix \(\tilde{A} \in \mathbb{R}^{N \times N}\). Specifically, we have \(\tilde{A}_{ij} = \tilde{A}_{ji} = \Theta_i \times (i-1)/2 + j\), where \(i,j \in [1, N]\). \(\Theta\) is randomly initialized before training. It can be easily seen that \(\tilde{A}\) is a complete graph that lacks interpretability and is computationally expensive. Therefore, we generate a sparse adjacency matrix \(A\) by a soft-thresholding operator

\[
A_{ij} = \text{ReLU}(\text{Sigmoid}(\tilde{A}_{ij}) - \text{Sigmoid}(\alpha)),
\]

where \(\text{Sigmoid}(\alpha) \in (0, 1)\) is a threshold for sparsification. To eschew arbitrary thresholding, we also make \(\alpha\) trainable and initialize it to 0. Noting that \(A_{ij} \in [0, 1], A\) can be considered as a probabilistic adjacency matrix, whose elements represent the probability to preserve the edge. After the sparsification, another problem is how to binarize \(A\). The simple thresholding approach is non-differentiable, thereby not suitable for stochastic gradient descent optimization. Therefore, we adopt the hard gumbel-softmax technique [13] to conduct discrete sampling from \(A\). Specifically, the self-learning binary adjacency matrix \(\tilde{A} \in \mathbb{R}^{N \times N}\) is constructed as

\[
\tilde{A}_{ij} = \arg\max(\text{Softmax}(\log A_{ij} + g_{ij}^1, \log(1 - A_{ij}) + g_{ij}^2) / \tau),
\]

where \(g_{ij}^1, g_{ij}^2 \sim \text{Gumbel}(0, 1)\), and \(\tau\) is the temperature set to 0.2. During back-propagating, we just abandon the \(\arg\max(\cdot)\) operation and leverage the softmax gradients to train the network.

Aggregation Weights Tuning. We assume that connections between brain regions should contribute differently to EMCI screening. Therefore, we add a trainable weight matrix on every ST-GC block to scale the importance of spatial edges. Varying from the symmetric adjacency matrix, weight matrices are non-symmetric to model directional interactions between brain regions. Moreover, elements in the weight matrix should be non-negative because graph convolution based on message passing mechanism is incompatible with negative connection weights [14]. Hence, for the spatial graph convolution in the \(l\)-th ST-GC block, we rewrite Eq.1 to the form with the self-learned graph structure, which is

\[
\hat{X}_{1:N,t}^{(l)} = D^{-\frac{1}{2}}(\tilde{A} \ast \text{ReLU}(\mathcal{M}^{(l)}))D^{-\frac{1}{2}}X_{1:N,t}^{(l)}W^{(l)},
\]

where \(\mathcal{M}^{(l)}\) is the learnable weight matrix in the \(l\)-th block which is initialized as all-ones, and * denotes the element-wise product. Ultimately, Eq.5 and Eq.2 jointly define a ST-GC block with the graph structure self-learning mechanism.

2.5. Implementation and Loss Function

For network architecture, we select the number of ST-GC blocks \(L\) to 3. Each ST-GC block has 64 channels for output. The dropout rate is set to 0.5. Our method is developed using Pytorch on a NVIDIA GeForce RTX TITAN GPU. We optimize the model using the Adam optimizer with a batch size of 16, a learning rate of 3e-4, and a weight decay of 1e-3. The window size \(T\) is set to 12 time points (36s).

For the loss function, we adopt the binary cross-entropy loss \(L_{BCE}\). To highlight effective task-specific ROI connections, we encourage the sparsity of our self-learned graph structure by adding a sparsity regularization formulated as

\[
L_{SP} = \frac{1}{N \times (N+1)/2} \sum_{i=1}^{N \times (N+1)/2} \text{Sigmoid}(\Theta_i).
\]

Ultimately, our goal is to minimize the final loss function \(\mathcal{L} = L_{BCE} + \lambda L_{SP}\), where \(\lambda\) is a hyperparameter to adjust the weight of \(L_{SP}\) and is set to 1e-4.
3. EXPERIMENTS AND RESULTS

We perform experiments using the publicly available ADNI database. The detail of data acquisition is described in Sec.2.1. Stratified 10-fold cross validation is exploited to split the training set and test set. For hyperparameter tuning, we randomly select 10% samples from the training set as the validation set in each fold. The performance of our model is measured by common evaluation metrics: accuracy (ACC), area under the curve (AUC), sensitivity (SEN), and specificity (SPE). Results are reported by the mean plus/minus standard deviation across 10 folds.

3.1. Performance Comparison with Relative Methods

In this paper, we perform a binary classification task of NC vs. EMCI. In order to demonstrate the superior performance of our proposed method, we compare our model with some related state-of-the-art approaches with a similar number of samples on the ADNI database. Results are shown in Table 2. To ensure fairness, our comparison principle is consistent with [2]. The comparison result demonstrates that our method achieves the best classification performance with 92.2% ACC, 91.7% SEN, and 92.9% SPE, which are all significantly higher than other models. The ACC of our model also achieves 94.6%.

### Table 2. Comparison with several state-of-the-art methods.

| Method       | ACC(%) | AUC(%) | SEN(%) | SPE(%) |
|--------------|--------|--------|--------|--------|
| MK-SVM[15]   | 78.3   | 77.1   | 82.1   | 74.0   |
| FSN-PFC[16]  | 82.8   | 88.2   | -      | -      |
| SF-net[17]   | 85.2   | 93.5   | 86.3   | 84.1   |
| SAC-GCN[18]  | 85.2   | 89.8   | 90.9   | 79.5   |
| MSGTNN[19]   | 44.4   | 77.4   | 89.9   | 87.0   |
| cwGAT[2]     | 72.3   | 90.9   | 96.7   | 90.4   |
| Proposed     | 92.2   | 94.6   | 91.7   | 92.9   |

All compared methods require pre-determined brain networks or other forms of features constructed from BOLD signals. In contrast, our method can learn the spatial temporal dependencies directly from BOLD time series and outperforms state-of-the-art methods.

3.2. Ablation Study

A set of ablation experiments are conducted to examine the effectiveness of different components. Specifically, we replace the self-learned graph structure with a fixed adjacency matrix in [20], replace our temporal inception module with a standard 1D convolution (kernel size=3, stride=1), remove the sparsity loss by set $\lambda = 0$, and replace layer-wise weight matrices with a learnable matrix shared across all layers. Results can be seen in Table 3. The self-learned graph structure shows a significant impact (↑ 2.1% ACC) on the classification performance, since it allows the model to learn the effective task-aware spatial topological structure between brain regions. The sparsity loss also brings improvements (↑ 0.7% ACC) because it suppresses spurious edges and benefits node feature aggregation in spatial graph convolution. The temporal inception module (↑ 0.6% ACC) possesses various kernel sizes, which can model long-term and short-term temporal dependencies simultaneously in temporal graph convolution. Compared with a single learnable weight matrix that is shared across layers, layer-wise reweighting matrices improves the accuracy (↑ 1.8% ACC). The explanation for this could be that node representations in different layers are diverse, thereby requiring varying aggregation weights.

#### Table 3. Ablation study results.

| Method                     | ACC(%) | AUC(%) | SEN(%) | SPE(%) |
|----------------------------|--------|--------|--------|--------|
| -w/o Self-learning Structure | 90.1 ± 1.9 | 92.0 ± 1.8 | 89.5 ± 2.7 | 90.3 ± 2.4 |
| -w/o Temporal Inception     | 91.6 ± 2.2 | 92.9 ± 2.5 | 90.2 ± 2.1 | 92.3 ± 2.0 |
| -w/o Sparsity Loss          | 91.5 ± 2.9 | 93.2 ± 2.8 | 90.3 ± 2.3 | 92.0 ± 2.7 |
| -w/o Layer-wise Reweighting | 90.4 ± 3.0 | 91.8 ± 2.6 | 88.9 ± 2.9 | 91.7 ± 2.4 |

3.3. Interpretability

To figure out brain regions related to EMCI, we synthesize the learned graph structure and aggregation weights, shown in Fig.3. Specifically, we generate a score vector $\mathbf{S} \in [0, 1]^N$ for brain regions by $\mathbf{S} = \sum_{l=1}^{L} \sum_{i=1}^{N} \text{Norm}(\tilde{\mathbf{A}}_l \cdot \text{ReLU}(\mathbf{M}(l)))$, where $\text{Norm}(\cdot)$ represents Min-Max scaling.

#### Fig. 3. Learned adjacency matrix and aggregation weights.

Top 10% salient ROIs are shown in Table 4. The excavated ROIs, specifically including PUN.L, IFGoperc.R, ANG.R, MFG.R, ORBinf.L, MFG.L, and IFGtriang.L, are consistent with previous finding [21], proving the credible interpretability of our model. Besides, four ROIs from the cerebellum may suggest the possible involvement of the cerebellum in cognition [22].

#### Table 4. Top 10% salient brain regions.

| ROI names     | ROI index | score | ROI names     | ROI index | score | ROI names     | ROI index | score | ROI names     | ROI index | score |
|---------------|-----------|-------|---------------|-----------|-------|---------------|-----------|-------|---------------|-----------|-------|
| PUN.L         | 67        | 1.000 | CER.6.L       | 98        | 0.628 | CER.11.L      | 107       | 0.738 |
| INS.R         | 30        | 0.925 | MFG.R         | 8         | 0.760 | ORBinf.L      | 15        | 0.722 |
| IFGoperc.R    | 12        | 0.921 | VER.3         | 110       | 0.743 | MFG.L         | 7         | 0.707 |
| ANG.R         | 66        | 0.836 | VER.9         | 115       | 0.703 | IFGtriang.R   | 14        | 0.691 |

4. CONCLUSION

This paper proposes a spatial temporal graph convolutional network with the graph structure self-learning mechanism, which directly utilize spatial temporal dependencies in BOLD time series to screen EMCI. Moreover, our graph structure self-learning mechanism successfully learns the optimal task-aware brain structure and aggregation weights. With these efforts, our model achieves better performance on the ADNI database compared with state-of-the-art methods. Further, we identify discriminative brain regions related to EMCI detection by analyzing learned graph structure and weights.
5. ACKNOWLEDGMENTS
This work was supported by the National Key R&D Program of China (2018YFA0704010 and 2018YFA0704101), the National Natural Science Foundation of China (No. 61971443), the Beijing Natural Science Foundation (L222034), and the Fundamental Research Funds for the Central Universities.

6. COMPLIANCE WITH ETHICAL STANDARDS
There are no conflicts of interest associated with this publication. Ethical approval was not required as confirmed by the license attached with the open access data.

7. REFERENCES
[1] Alzheimer’s Association, “2015 alzheimer’s disease facts and figures,” Alzheimer’s & Dementia, vol. 11, no. 3, pp. 332–384, 2015.
[2] Yang Li, Jingyu Liu, Yiqiao Jiang, Yu Liu, and Baiying Lei, “Virtual adversarial training-based deep feature aggregation network from dynamic effective connectivity for mci identification,” IEEE transactions on medical imaging, vol. 41, no. 1, pp. 237–251, 2021.
[3] Tae-Eui Kam, Han Zhang, Zhicheng Jiao, and Dinggang Shen, “Deep learning of static and dynamic brain functional networks for early mci detection,” IEEE transactions on medical imaging, vol. 39, no. 2, pp. 478–487, 2019.
[4] K A Smitha, K Akhil Raja, KM Arun, PG Rajesh, Bejoy Thomas, TR Kapilamoorthy, and Chandrasekharan Kesavadas, “Resting state fmri: A review on methods in resting state connectivity analysis and resting state networks,” The neuroradiology journal, vol. 30, no. 4, pp. 305–317, 2017.
[5] Xia-an Bi, Qing Shu, Qi Sun, and Qian Xu, “Random support vector machine cluster analysis of resting-state fmri in alzheimer’s disease,” PloS one, vol. 13, no. 3, pp. e0194479, 2018.
[6] Byung-Hoon Kim, Jong Chul Ye, and Jae-Jin Kim, “Learning dynamic graph representation of brain connectome with spatio-temporal attention,” Advances in Neural Information Processing Systems, vol. 34, pp. 4314–4327, 2021.
[7] Xuegang Song, Feng Zhou, Alejandro F Frangi, Jiwen Cao, Xiaohua Xiao, Yi Lei, Tianfu Wang, and Baiying Lei, “Multicenter and multi-channel pooling gcn for early ad diagnosis based on dual-modality fused brain network,” IEEE Transactions on Medical Imaging, 2022.
[8] Houliang Zhou, Lifang He, Yu Zhang, Li Shen, and Brian Chen, “Interpretable graph convolutional network of multimodality brain imaging for alzheimer’s disease diagnosis,” in 2022 IEEE 19th International Symposium on Biomedical Imaging (ISBI). IEEE, 2022, pp. 1–5.
[9] Xiaodian Xing, Qingfeng Li, Hao Wei, Minqing Zhang, Yiqiang Zhan, Xiang Sean Zhou, Zhong Xue, and Feng Shi, “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, 2019, pp. 639–646.
[10] Clifford R Jack Jr, Matt A Bernstein, Nick C Fox, Paul Thompson, Gene Alexander, Danielle Harvey, Bret Borowski, Paula J Britton, Jennifer L. Whitwell, Chadwick Ward, et al., “The alzheimer’s disease neuroimaging initiative (adni): Mri methods,” Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine, vol. 27, no. 4, pp. 685–691, 2008.
[11] Thomas N Kipf and Max Welling, “Semi-supervised classification with graph convolutional networks,” arXiv preprint arXiv:1609.02907, 2016.
[12] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich, “Going deeper with convolutions,” in Proceedings of the IEEE conference on computer vision and pattern recognition, 2015, pp. 1–9.
[13] Eric Jang, Shixiang Gu, and Ben Poole, “Categorical reparameterization with gumble-softmax,” in International Conference on Learning Representations (ICLR 2017). OpenReview.net, 2017.
[14] Xuan Kan, Hejie Cui, Joshua Lukemire, Ying Guo, and Carl Yang, “Fbnetgen: Task-aware gnn-based fmri analysis via functional brain network generation,” in Medical Imaging with Deep Learning, 2021.
[15] Biao Jie, Mingxia Liu, and Dinggang Shen, “Integration of temporal and spatial properties of dynamic connectivity networks for automatic diagnosis of brain disease,” Medical image analysis, vol. 47, pp. 81–94, 2018.
[16] Feng Yang, Feng Zhou, Dong Ni, Yanwu Xu, Siping Chen, Tianfu Wang, and Baiying Lei, “Fused sparse network learning for longitudinal analysis of mild cognitive impairment,” IEEE transactions on cybernetics, vol. 51, no. 1, pp. 233–246, 2019.
[17] Baiying Lei, Nina Cheng, Alejandro F Frangi, Ee-Leng Tan, Jiwen Cao, Peng Yang, Ahmed Elazab, Jie Du, Yanwu Xu, and Tianfu Wang, “Self-calibrated brain network estimation and joint non-convex multi-task learning for identification of early alzheimer’s disease,” Medical image analysis, vol. 61, pp. 101652, 2020.
[18] Xuegang Song, Feng Zhou, Alejandro F Frangi, Jiwen Cao, Xiaohua Xiao, Yi Lei, Tianfu Wang, and Baiying Lei, “Graph convolution network with similarity awareness and adaptive calibration for disease-induced deterioration prediction,” Medical Image Analysis, vol. 69, pp. 101947, 2021.
[19] Yali Qiu, Shuangzhi Yu, Yanhong Zhou, Dongdong Liu, Xuegang Song, Tianfu Wang, and Baiying Lei, “Multi-channel sparse graph transformer network for early alzheimer’s disease identification,” in 2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI). IEEE, 2021, pp. 1794–1797.
[20] Soham Gadgil, Qingyu Zhao, Adolf Pfefferbaum, Edith V Sullivan, Ehsan Adeli, and Kilian M Pohl, “Spatio-temporal graph calibration for disease-induced deterioration prediction,” in International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, 2020, pp. 528–538.
[21] Jie Xiang, Hao Guo, Rui Cao, Hong Liang, and Junjie Chen, “An abnormal resting-state functional brain network indicates progression towards alzheimer’s disease,” in Neural regeneration research, vol. 8, no. 30, pp. 2789, 2013.
[22] Heidi IL Jacobs, David A Hopkins, Helen C Mayhoffer, Emiliano Bruner, Fred W van Leeuwen, Wijnand Raaijmakers, and Jeremy D Schmahmann, “The cerebellum in alzheimer’s disease: evaluating its role in cognitive decline,” Brain, vol. 141, no. 1, pp. 37–47, 2018.