NeuroTree: Hierarchical Functional Brain Pathway Decoding for Mental Health Disorders

Jun-En Ding[1], Dongsheng Luo[2], Chenwei Wu[3], **Feng Liu**[1]

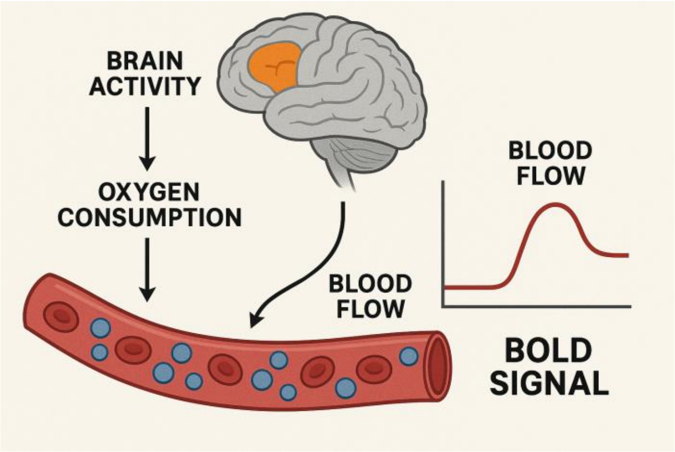
1 Stevens Institute of Technology
2 Florida International University
3 University of Michigan

Lab Website

Find Me

Background

- Analyzing functional brain networks through Functional magnetic resonance imaging (fMRI) is crucial for understanding mental disorders.



- BOLD signals reflect brain activity by measuring oxygen consumption-related blood flow changes that alter MRI signal intensity, indirectly indicating active brain regions.

ICML International Conference On Machine Learning

Challenge 1- Limitations of traditional brain network analysis in fMRI

Failed to detect high-order brain network anomalies overtime

- Ignoring the dynamic FC of brain network **changes over time**
- Unable to decode interactions in **higher-order brain activity regions**

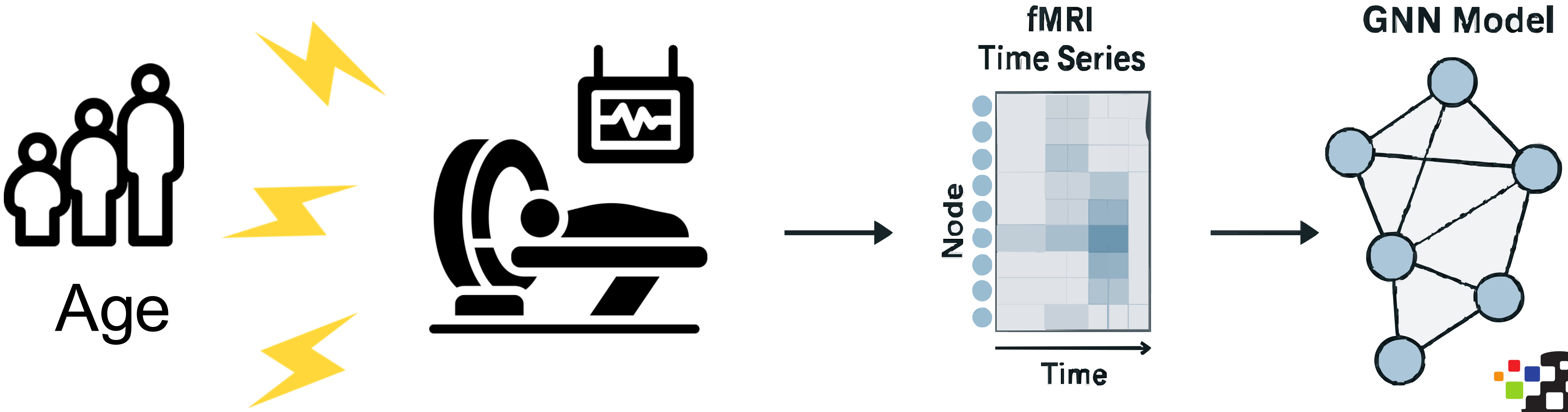
To identify disease-associated fMRI brain networks

- Existing methods cannot clearly express **brain connectivity interaction pathways**

Challenge 2- Traditional GNN-based fMRI models ignore demographics effect

Demographics impact fMRI in mental disorder patients

- Functional MRI signals exhibit **age-dependent variations**
- **Demographics** should be considered in GNN modeling of mental disorders cohorts

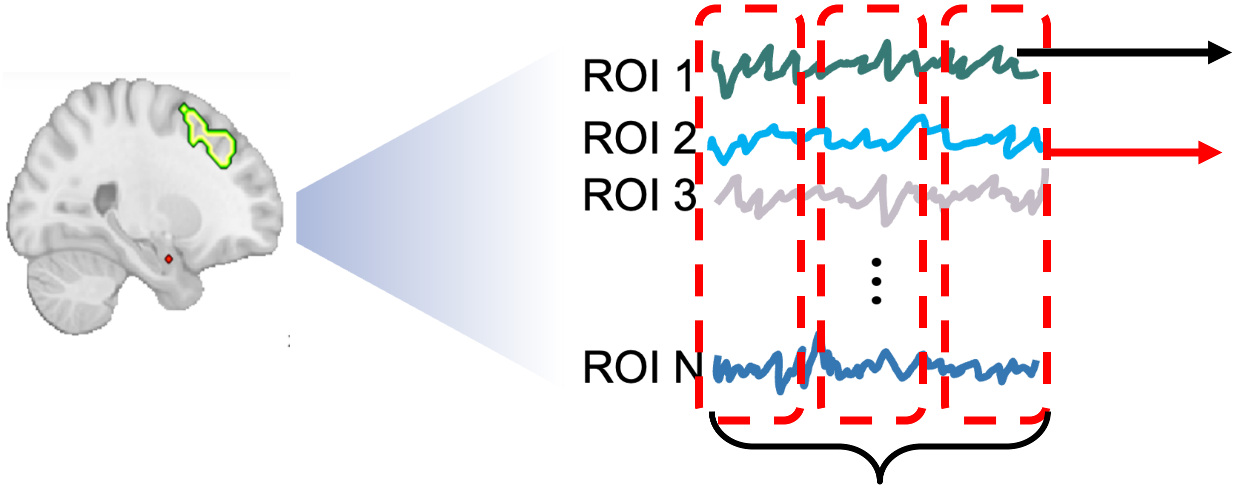


Method - K-hop ODE-Based Graph Embedding

- Neural Ordinary Differential Equation (Neural ODE)

Construct ODE incorporates demographics for dynamic BOLD fMRI signal modeling

$$\frac{dX(t)}{dt} = \eta A^d(t)X(t) + \rho \cdot \theta X(t) + Cu(t),$$



A^s : static FC matrix

$X(t)$: BOLD fMRI signals at time t

$A^d(t)$: dynamic FC matrices

$Cu(t)$: additional external stimuli

η, ρ : adjust parameters

θ : external demographics factor

Method - K-hop ODE-Based Graph Embedding

- Neural Ordinary Differential Equation (Neuro ODE)

Discrete adjacency matrices representation

$$A^d(t) = \frac{1}{\eta}\left(\frac{dX(t)}{dt}\frac{1}{X(t)} - \rho \cdot \theta\right)$$
$$\overset{\Delta t=1}{\approx} \varphi\left(\frac{X(t+1) - X(t)}{X(t)} - \rho \cdot \boxed{\theta}\right)$$

Age factor

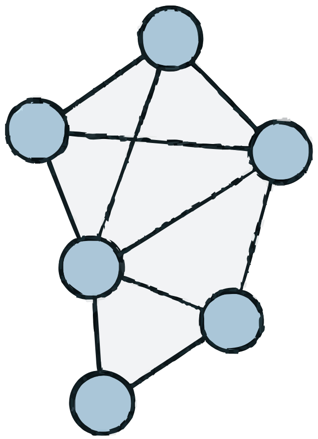
where $\varphi = \frac{1}{\eta}$ is scale factors ranging from 0 to 1.

Method - K-hop ODE-Based Graph Embedding

- Transitional GCN

$$H^{(l)} = \sigma(D^{-\frac{1}{2}} A D^{-\frac{1}{2}} H^{(l-1)} W^{(l-1)}).$$

GNN Model



- K-hop ODE-based GCN

$$H^{(l+1)}(t) = \sigma\Big(\sum_{k=0}^{K-1} \boxed{\Phi_k(t)} H^{(l)}(t) W_k^{(l)} \Big)$$

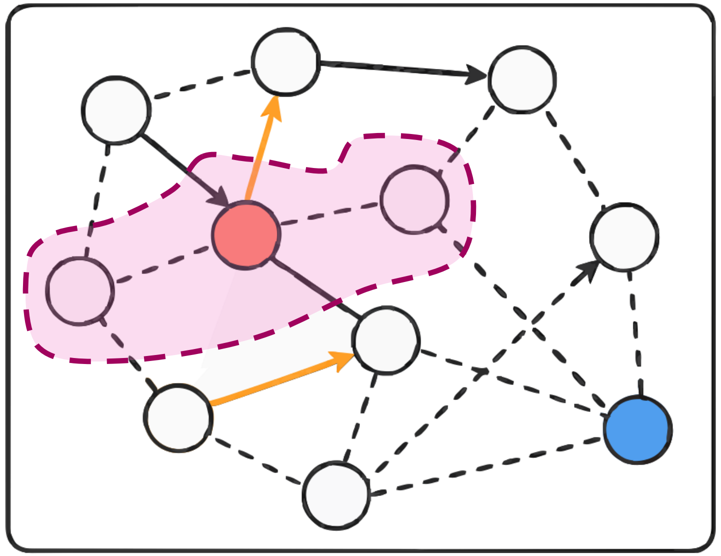
$$\Phi_k(t) = \hat{D}^{-\frac{1}{2}} \boxed{\hat{A}_k(t)} \hat{D}^{-\frac{1}{2}}$$

$$\hat{A}_k(t) = \Gamma \odot A^s \odot \underbrace{[\lambda A^d(t) + (1-\lambda)(A^d(t))^T]^k}_{\text{k-hop connectivity}}$$

Method - K-hop ODE-Based Graph Embedding

- K-hop connectivity in brain network

$$\hat{A}_k(t) = \Gamma \odot A^s \odot [\lambda A^d(t) + (1-\lambda)(A^d(t))^T]^k$$



- - - - Static A^s —— Dynamic $A^d(t)$

—— k-hop path

Method - Contrastive Masked Functional Connectivity (CMFC) Optimization

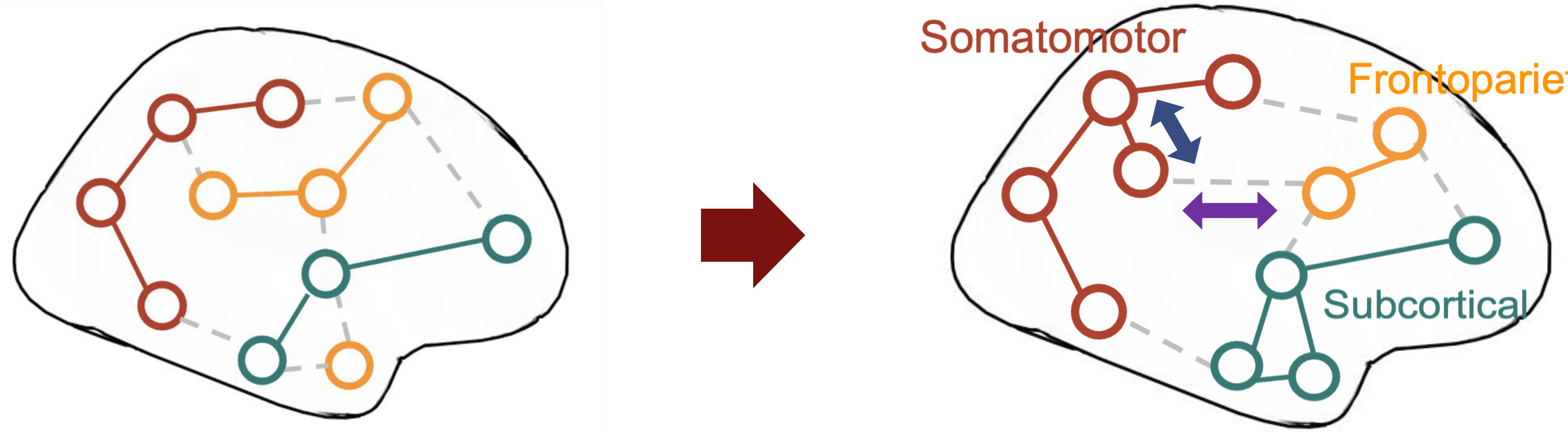
- The CMFC loss can minimize similarities while maximizing dissimilarities of brain regions

$$\mathcal{L}_{\text{pos}} = -\frac{1}{|\mathcal{A}^+|} \sum_{(i,j)\in\mathcal{A}^+} \log\left(\frac{\exp(S_{ij}(t))}{\sum_{k\in\mathcal{V}}\exp(S_{ik}(t))+\epsilon}\right),$$

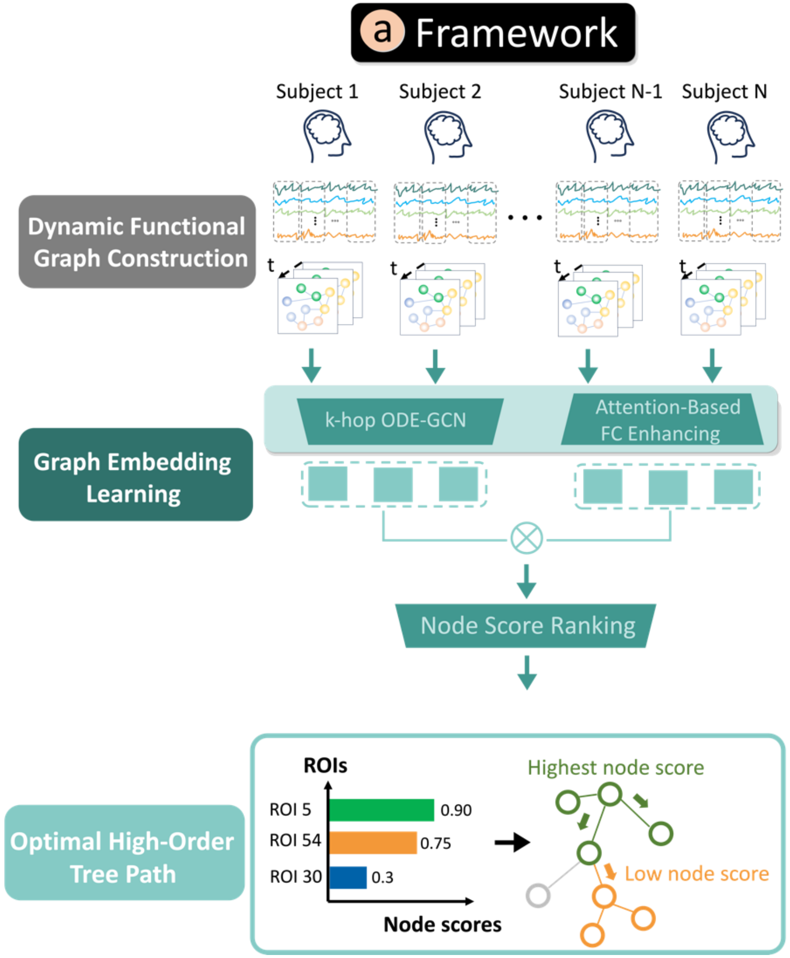
$$\mathcal{L}_{\text{neg}} = -\frac{1}{|\mathcal{A}^-|} \sum_{(i,j)\in\mathcal{A}^-} \log\left(1-\frac{\exp(S_{ij}(t))}{\sum_{k\in\mathcal{V}}\exp(S_{ik}(t))+\epsilon}\right).$$

Minimize region connectivity strength ⟷

Maximize region connectivity strength ⟷



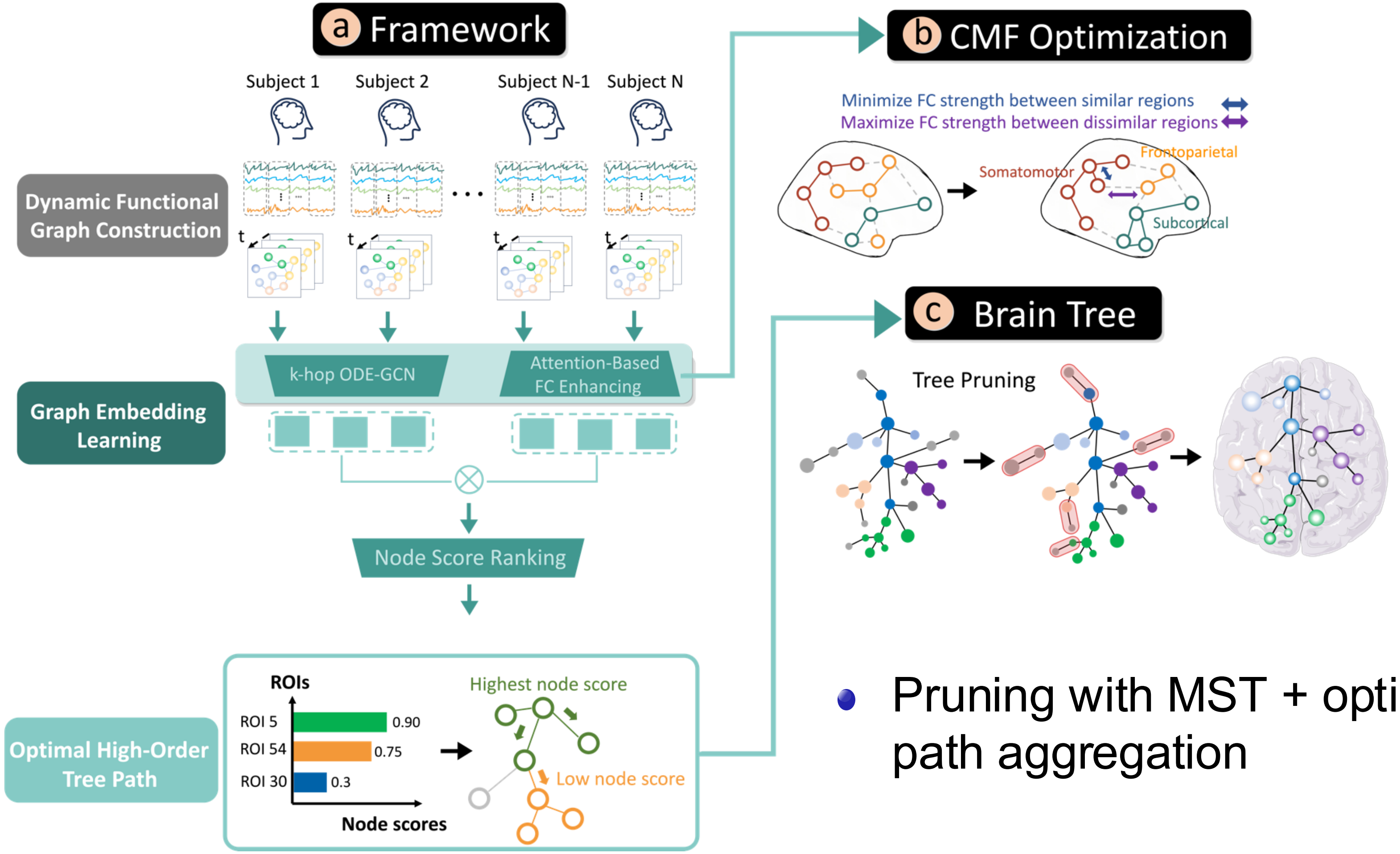
Method - Node Score Predictor



$$\mathcal{S}_i = h_i \cdot \zeta\Big(\frac{1}{|\mathcal{V}|}\sum_{j\in\mathcal{V}} Z_j(\mathbf{\Theta})^\top Z_i(\mathbf{\Theta})\Big), \quad i \in \{1, 2, \ldots, |\mathcal{V}|\}$$

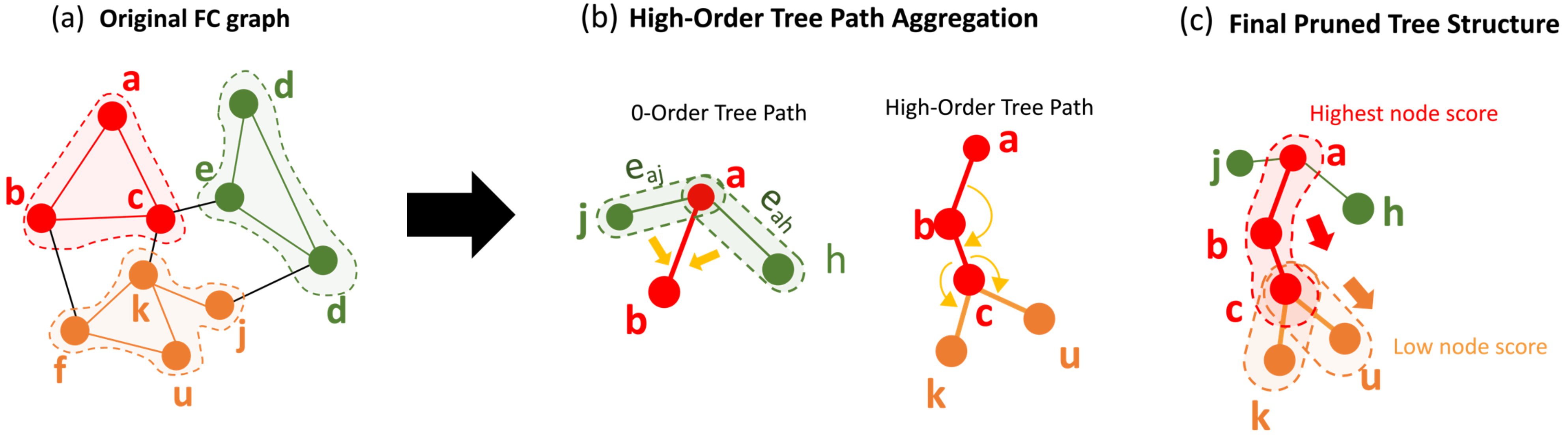
- Predicted brain regions as node score
- Reranking important node scores to assign hierarchical brain pathways

Method - Hierarchical Brain Tree Construction



- Pruning with MST + optimal weighted tree path aggregation

Method - Hierarchical Brain Tree Construction



- Direct and indirect path information aggregation with different orders
- Aggregate important node scores and weighted edges into brain pathways

$$\mathcal{W}(P) = \alpha \underbrace{\sum_{v \in P} \mathcal{S}(v; \boldsymbol{\Theta})}_{\textbf{Node Score Contribution}} + (1-\alpha) \underbrace{\sum_{s=1}^{S} \sum_{(v_i, v_j) \in E(P)} \mathcal{F}^{(s)}_{v_i v_j}}_{\textbf{High-Order FC Contribution}} .$$

Results – Mental brain disorders classification

- **Datasets: Cannabis (90 ROIs), COBRE (118 ROIs)**

Brain Network Classification

Healthy Controls Disease Cohort

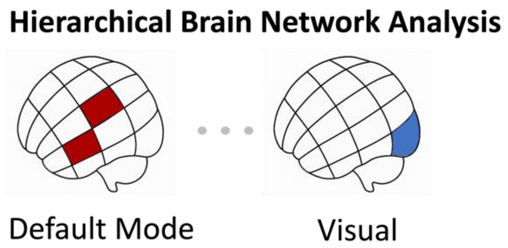
Table 1. Evaluating graph classification performance with five-fold cross-validation. We computed the most competitive baseline for each method. We compared the second-best methods denoted by blue color and calculated the improvement rate, denoted as "Improv. (%)".

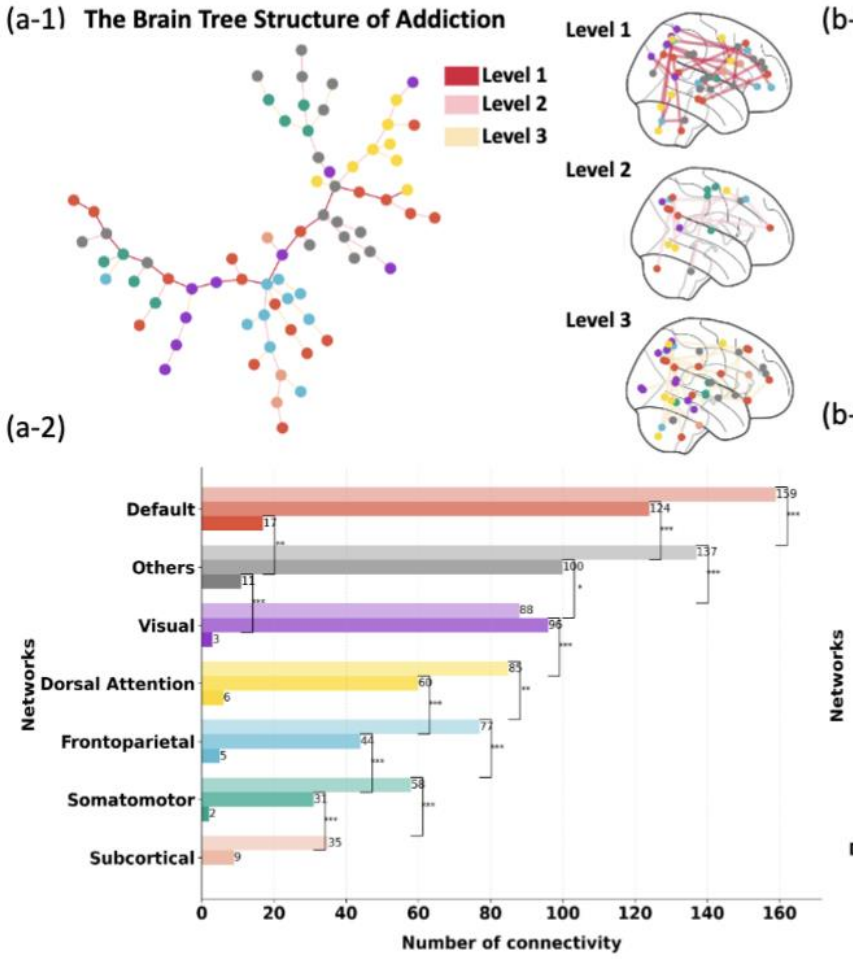
	Cannabis				COBRE			
Model	**AUC**	**Acc.**	**Prec.**	**Rec.**	**AUC**	**Acc.**	**Prec.**	**Rec.**
Pearson GCN	0.67±0.06	0.55±0.07	0.59±0.13	0.55±0.06	0.54±0.11	0.55±0.10	0.61±0.12	0.55±0.10
k-NN GCN	0.64±0.03	0.62±0.03	0.63±0.03	0.63±0.03	0.66±0.07	0.62±0.08	0.63±0.08	0.63±0.08
GAT (Veličković et al., 2017)	0.72±0.05	0.67±0.04	0.70±0.06	0.67±0.04	0.67±0.08	0.60±0.11	0.57±0.21	0.60±0.11
BrainGNN (Li et al., 2021)	0.67±0.13	0.59±0.16	0.51±0.28	0.59±0.12	0.55±0.11	0.50±0.02	0.31±0.11	0.50±0.02
BrainUSL (Zhang et al., 2023)	0.63±0.11	0.65±0.06	0.62±0.13	0.63±0.11	0.57±0.10	0.54±0.04	0.41±0.18	0.57±0.11
BrainGSL (Wen et al., 2023a)	0.59±0.11	0.65±0.02	0.67±0.17	0.65±0.02	0.55±0.12	0.51±0.04	0.45±0.11	0.51±0.04
MixHop (Abu-El-Haija et al., 2019)	0.73±0.05	0.69±0.03	0.70±0.04	0.69±0.03	0.69±0.05	0.61±0.06	0.62±0.07	0.61±0.06
GPC-GCN (Li et al., 2022b)	0.53±0.05	0.60±0.06	0.37±0.08	0.60±0.06	0.50±0.00	0.47±0.04	0.22±0.04	0.47±0.04
PathNN (Michel et al., 2023)	0.70±0.10	0.67±0.04	0.72±0.12	**0.83**±0.16	0.49±0.01	0.51±0.05	0.32±0.27	0.43±0.46
Ours (w/o θ)	0.49±0.01	0.60±0.06	0.37±0.08	0.60±0.06	0.50±0.00	0.47±0.04	0.22±0.01	0.47±0.04
Ours (w/o $\mathcal{L}_{CMFC}$)	0.74±0.08	0.73±0.05	0.73±0.04	0.73±0.05	0.69±0.10	0.63±0.10	0.64±0.10	0.63±0.10
NEUROTREE	**0.80**±0.05	**0.73**±0.04	**0.73**±0.04	0.74±0.04	**0.71**±0.10	**0.65**±0.08	**0.66**±0.08	**0.65**±0.08
Improv. (%)	8.11%	-	-	1.37%	2.89%	3.17%	3.12%	3.17%

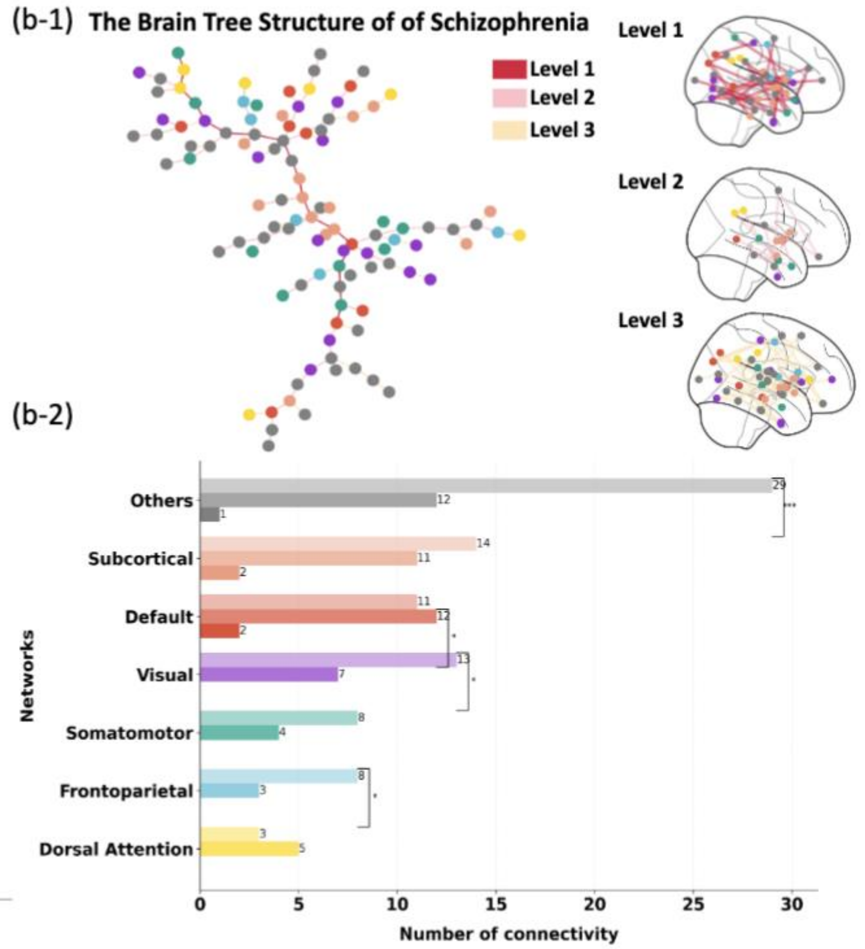
- Age-aware modeling and CMFC loss boost model robustness
- Best AUC: 0.80 (Cannabis), 0.71 (COBRE)

ICML International Conference On Machine Learning

Results – Visualization of brain tree in different brain disorders



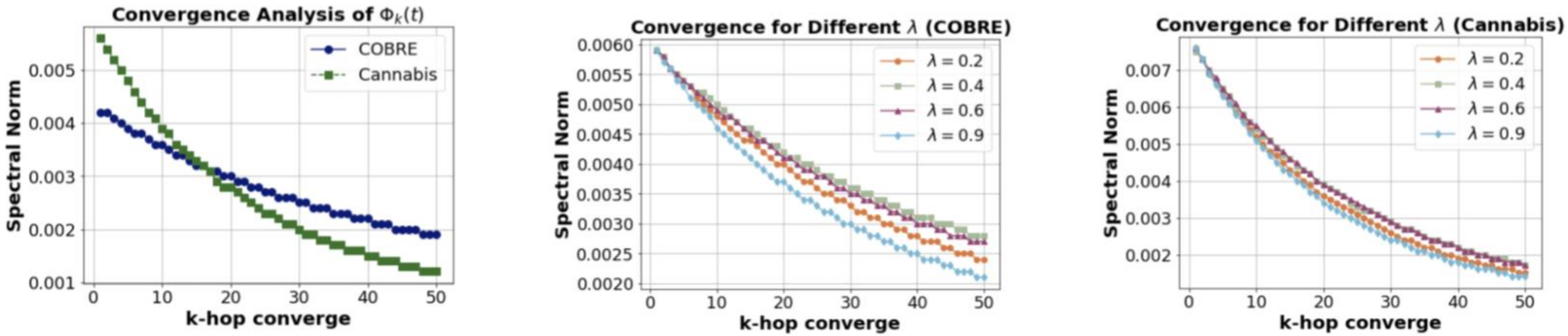




- Tree paths in addiction show DMN/VN dominance
- Schizophrenia highlights SUB and DMN

ICML
International Conference On Machine Learning

Results – Convergence Analysis

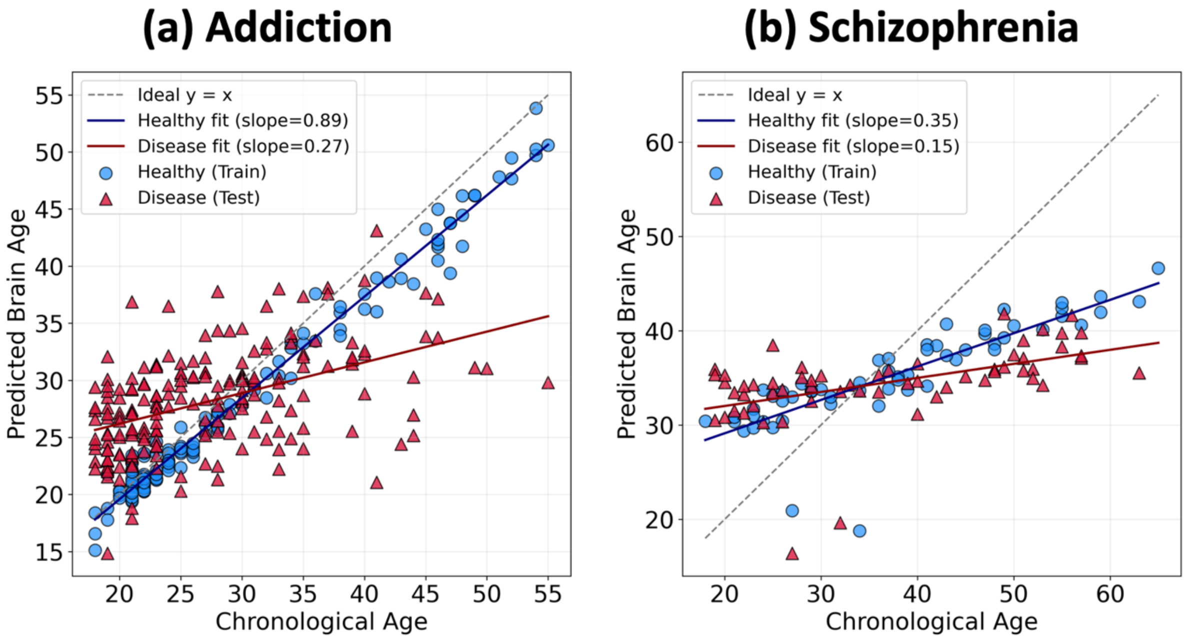


- Different mental disorders such as addiction and schizophrenia have different rates of deterioration in the brain

- Cannabis use disorder exhibits faster spectral norm convergence rates.

ICML International Conference On Machine Learning

Results – Brain age estimation



- Comparing predicted brain age from fMRI to actual age reveals insights into mental disorder severity and progression
- Younger groups show lower prediction errors, and mental disorders accelerate brain aging

Conclusion

Graph classification

- NeuroTree incorporates AGE-GCN layers to achieve SOTA graph classification

Interpretable for mental health disorders

- NeuroTree reveals disease-specific patterns (Addiction vs. Schizophrenia)
- Builds interpretable and learnable trunks and branches for hierarchical paths in tree structures

High-order brain network path learning

- NeuroTree effectively integrates high-level brain region interaction pathway features

ICML
International Conference
On Machine Learning

Thanks for your attention!

STEVENS INSTITUTE of TECHNOLOGY
THE INNOVATION UNIVERSITY®

AI in Neuroimaging & Healthcare Lab

Lab Website

Find Me

Paper: https://arxiv.org/abs/2502.18786

ICML
International Conference On Machine Learning