Neurodegenerative Brain Network Classification via Adaptive Diffusion with Temporal Regularization

Hyuna Cho [1] **Jaeyoon Sim** [1] **Guorong Wu** [2] **Won Hwa Kim** [1]

Abstract

Analysis of neurodegenerative diseases on brain connectomes is important in facilitating early diagnosis and predicting its onset. However, investigation of the progressive and irreversible dynamics of these diseases remains underexplored in cross-sectional studies as its diagnostic groups are considered independent. Also, as in many real-world graphs, brain networks exhibit intricate structures with both homophily and heterophily. To address these challenges, we propose **A**daptive **G**raph diffusion network with **T**emporal regularization (AGT). AGT introduces node-wise convolution to adaptively capture low (i.e., homophily) and high-frequency (i.e., heterophily) characteristics within an optimally tailored range for each node. Moreover, AGT captures sequential variations within progressive diagnostic groups with a novel temporal regularization, considering the relative feature distance between the groups in the latent space. As a result, our proposed model yields interpretable results at both node-level and group-level. The superiority of our method is validated on two neurodegenerative disease benchmarks for graph classification: Alzheimer's Disease Neuroimaging Initiative (ADNI) and Parkinson's Progression Markers Initiative (PPMI) datasets.

1. Introduction

Consider a neurodegenerative brain connectome study where each sample is represented as a graph, i.e., brain regions of interest (ROIs) correspond to nodes and connectomic features represent edges, and a subset of brain networks progressively deteriorates over time due to a disease. Unlike typical classification tasks where the groups (i.e., labels) are nominally discriminated (Gao & Ji, 2019; Zhang et al., 2018), the neurodegenerative diagnostic groups demonstrate temporal dynamics. Accurately characterizing the disease progression is essential for precise diagnosis; however, most existing cross-sectional brain network classifiers do not adequately address the evolving nature of the diseases (Wang et al., 2023; Qu et al., 2023).

Several issues, other than the temporal nature, make the analysis even more challenging. As in many real-world graphs such as social networks (Yanardag & Vishwanathan, 2015), molecule structures (Ramakrishnan et al., 2014; Axelrod et al., 2022), and traffic flows (Chen et al., 2001), brain networks are high-dimensional and sparse, i.e., only a small fraction of edges are present comprising heterogeneous structures while the number of possible edges is at the order of N^2 with N nodes. Moreover, graph components of different orders, i.e., nodes (0-simplex) and edges (1-simplex), need to be analyzed interactively. Although nodes and edges are interdependent under homophily condition, it is not universally applicable to every brain connectivity. As brain networks may exhibit heterophily, where dissimilar regions physically attach, this interplay between homophily and heterophily intertwines nodes and edges, and makes their analysis difficult.

Therefore, capturing *localized* characteristics between node and edge features becomes imperative in understanding the intricacies of these diverse relationships. Traditional Graph Neural Networks (GNNs) have shown successful results in solving various tasks on graphs. However, they do not effectively emphasize the nuances or variations between connected nodes with different properties (Kipf & Welling, 2017; Wu et al., 2019). This is because the convolution layers aggregate information from direct neighbor nodes *uniformly* across the entire nodes. By repeatedly stacking graph convolution layers, the notorious oversmoothing issue arises in which the information is excessively averaged around each node.

To handle the aforementioned issues, we propose **A**daptive **G**raph diffusion network with **T**emporal regularization (AGT) for neurodegenerative brain network classification. Overall, AGT has two key ideas: **1)** Node-variant convolution that adjusts the amount of smoothing or coarsening the

[1]Pohang University of Science and Technology (POSTECH), South Korea [2]University of North Carolina at Chapel Hill, USA. Correspondence to: Won Hwa Kim <wonhwa@postech.ac.kr>.

Proceedings of the 41st International Conference on Machine Learning, Vienna, Austria. PMLR 235, 2024. Copyright 2024 by the author(s).

information around each node within different node-wise ranges. **2)** Group-level temporal regularization that captures the sequential variations along the progressive degeneration from a healthy control group to the most deteriorated group.

With the node-variant convolution, both low-frequency (i.e., homophily) and high-frequency (i.e., heterophily) characteristics are adaptively captured within an optimal range of neighbor nodes for each node. This is realized by using a diffusion kernel (Hammond et al., 2011; Kim et al., 2012) with trainable node-wise scaling parameters. While a typical diffusion kernel identically smooths out features for every node using a pre-defined bandwidth (Xu et al., 2019a;b; Wang et al., 2021a), we design an adaptive diffusion kernel with node-wise scales that finds optimal frequency bandwidth for each node. Therefore, our method can flexibly control the amount of smoothing or highlighting node-individual features for each node by connecting indirect nodes. By performing graph convolution with the diffusion kernel, unified graph embeddings incorporating joint relationships between nodes and edges are obtained (Cho et al., 2023). Our temporal regularization controls the embedding distances between diagnostic groups in the latent space, such that the temporal dynamics along the groups are effectively captured.

In a nutshell, the main contributions of our work are summarized as follows:

- Our method addresses the challenges of analyzing intricate brain networks by introducing a node-variant convolution that adaptively captures both localized homophily and heterophily characteristics.
- Our method captures sequential variations in the progressive degeneration of brain networks, characterizing temporal features of a disease that change over time.
- Consequently, AGT yields neuroscientifically interpretable results in both brain regional analysis and inter-group analysis.

Extensive validation was performed on two independent public benchmarks: Alzheimer's Disease Neuroimaging Initiative (ADNI) and Parkinson's Progression Markers Initiative (PPMI) datasets to evaluate the generalizability and efficiency of our model.

2. Related Work

2.1. Brain Network Analysis

Recent works on cross-sectional brain network analysis lie in two-fold: predicting node values (Mao et al., 2018) and graph-level labels (Liu et al., 2023). While most of these works adopt single-modal data (Relión et al., 2019), studies on integrating multi-modal data have recently been conducted to enhance prediction (Wang et al., 2017; Zhu et al., 2018). Apart from these graph studies that use both node and edge features, graph classification methods focusing solely on edge data (Park et al., 2023) have been introduced to avoid expensive costs of acquiring medical data.

2.2. Spectral Graph Neural Networks

Spectral GNNs refer to GNNs that use spectral graph filters to analyze graph data. Conventional spectral GNNs use filters with fixed frequency bandwidths. For example, GraphWave (Donnat et al., 2018) determines the bandwidth via an analysis of variance on a wavelet basis. ChebNet (Defferrard et al., 2016) uses a Chebyshev polynomial filter, and GCN (Kipf & Welling, 2017) simplifies the Chebyshev filter with first-order approximation.

In contrast to these filters that can only capture spectral features within a predefined bandwidth, recent spectral GNNs have introduced learnable filters. These filters come with trainable scaling parameters that enable the adjustment of bandwidth magnitudes. For example, GPR-GNN (Chien et al., 2021) updates layer-wise polynomial filters via gradient descent, and AdaGNN (Dong et al., 2021) learns channel-wise parameters that capture significant frequencies in the channel level. GAT (Veličković et al., 2018) adapts how (and how much) information they get from different nodes with attention score, but it only aggregates from direct neighboring nodes. Wave-GD (Cho et al., 2023) aggregates information from both direct and indirect nodes on multi-resolution graph representations with multiple trainable scales. In this work, the number of multi-resolution graphs linearly increases as the number of used scales increases.

Without requiring many graphs at multi-resolution, Exact (Choi et al., 2022) and LSAP (Sim et al., 2024) adopt a node-wise scale learning scheme that learns node-wise bandwidths. These methods only extract low-frequency graph characteristics for the entire graphs, and thus the distinctions between two neighboring nodes with different properties may not be effectively emphasized. Unlike these approaches, our framework considers both low and high-frequency features, and these features are further used to capture spectral cross-characteristics between graph structures and node features.

3. Preliminary

3.1. Spectral Graph Wavelet Transform

An undirected graph $G = \{\mathbf{V}, \mathbf{E}, \omega\}$ with N nodes is comprised a node set $\mathbf{V}$, edges $\mathbf{E}$ and corresponding edge weights ω. The $\mathbf{E}$ and ω define a symmetric adjacency matrix $\mathbf{A} \in \mathbb{R}^{N \times N}$, whose individual element is a strength between two nodes given by the ω. Given a diagonal degree matrix $\mathbf{D}$, a graph Laplacian is defined as $\mathbf{L} = \mathbf{D} - \mathbf{A}$ and its normalized case is defined as $\tilde{\mathbf{L}} = \mathbf{D}^{-1/2}\mathbf{L}\mathbf{D}^{-1/2}$. As

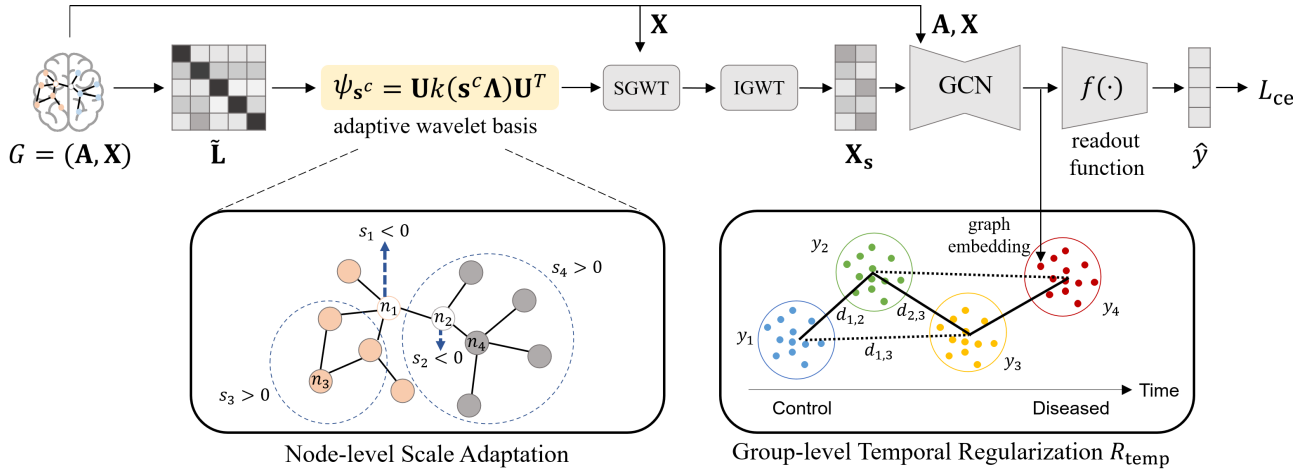


Figure 1. Overview of AGT architecture for neurodegenerative brain network classification. From a normalized graph Laplacian $\tilde{\mathbf{L}}$, adaptive wavelet basis $\psi_{\mathbf{s}^c}$ is constructed to capture localized spectral characteristics around each node. By performing graph convolution with the scale-filtered features $\mathbf{X_s}$ and the connectivity $\mathbf{A}$, the underlying cross-characteristics between the graph structure and node features are integrated as graph embedding. The temporal regularization R_{temp} captures the sequential dynamics along neurodegeneration from these graph embeddings in a mini-batch.

the $\tilde{\mathbf{L}}$ is real and positive semi-definite, it has non-negative eigenvalues $\mathbf{\Lambda} = diag(\lambda_1, ..., \lambda_N)$ and orthonormal eigenvectors $\mathbf{U} = [u_1, u_2, ..., u_N]$.

Spectral Graph Wavelet Transform (SGWT) (Hammond et al., 2011) extends the theory of wavelet transform (Mallat, 1999) to graphs via spectral graph theory. By SGWT, the graph signal x is decomposed into varying levels of granularity in the spectral space so that multiresolution graph analysis becomes available. To project the signals into the spectral space, SGWT utilizes wavelet basis $\psi_s = \mathbf{U}k(s\mathbf{\Lambda})\mathbf{U}^T$, where $k(\cdot)$ is a kernel function with a scaling parameter s. The ψ_s is a realization of a kernel function $k(\cdot)$ in the spectral domain localized with δ_n on the n-th node in the graph space, which captures graph characteristics at a fixed resolution. A graph signal x (i.e., either node features or edge weights) is projected onto the spectral domain as $W_x(s) = \psi_s \cdot x$, which yields a wavelet coefficient $W_x(s)$. Under the admissibility condition (Mallat, 1999), the original x is perfectly reconstructed via Inverse Graph Wavelet Transform (IGWT) as follows

$$x = \frac{1}{C_k} \int_0^\infty \psi_s \cdot W_x(s) \frac{ds}{s}, \tag{1}$$

with an admissibility constant $C_k = \int_0^\infty \frac{k(\lambda)^2}{\lambda} d\lambda < \infty$.

The Eq. (1) is a superposition of multi-resolution representation of x over scales $s \in [0, \infty)$. Therefore, a signal x_s in the graph space filtered at the scale s is defined as

$$x_s = \psi_s \cdot W_x(s) = \mathbf{U}k^2(s\mathbf{\Lambda})\mathbf{U}^T x. \tag{2}$$

While the raw graph data x may contain unnecessary resolutions for solving a given task, Eq. (2) allows us to extract task-relevant graph information at specific resolution(s) in the spatial domain.

3.2. Spectral Coherence between Nodes and Edges

Given a graph jointly constructed with multivariate node features $\mathbf{X} \in \mathbb{R}^{N \times F}$ and graph structures $\mathbf{A} \in \mathbb{R}^{N \times N}$, these two components should have coherent properties to explain the whole graph system. For example, given a homophily assumption that similar nodes are likely to attach, edge signals should exhibit coherence with node signals as the edges should properly explain the relationships between the nodes. Authors in (Cho et al., 2023) showed that these underlying coherence between nodes and edges, i.e., spectral coherence, can be captured in the spectral space by a dot product of wavelet coefficients as $W_{\mathbf{A}}(s) \cdot W_{\mathbf{X}}(s)$. Also, they showed that this spectral coherence at scale s can be identically obtained in the graph space as a graph convolution by performing the scale-specific IGWT (Eq. (2)) to either of the graph components as follows

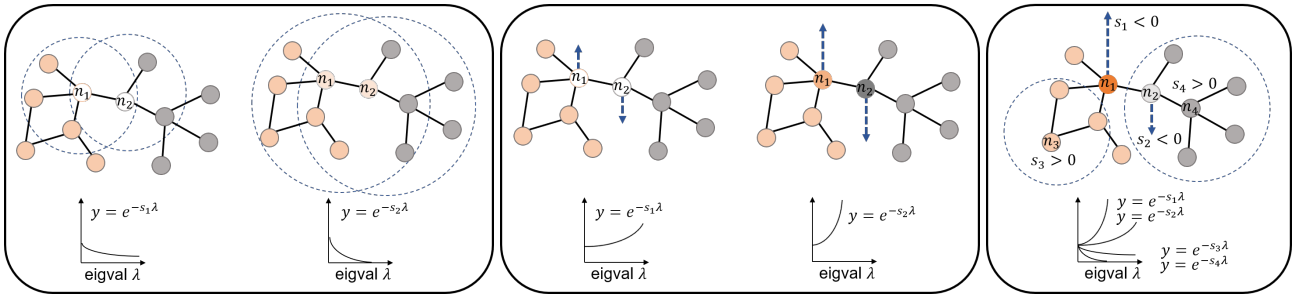
$$W_{\mathbf{A}}(s) \cdot W_{\mathbf{X}}(s) = \mathbf{A}_s\mathbf{X} = \mathbf{A}\mathbf{X}_s. \tag{3}$$

By Eq. (3), the spectral cross-characteristics of the node and edge features at a specific resolution (i.e., scale s) are captured in the spatial domain with graph convolutions.

4. Method

4.1. Graph Classification for Degenerative Disorder

A brain network is represented as a graph by considering brain regions (i.e., ROIs) as nodes and the connections between them as edges. Given a population of graphs with



(a) Low-pass global filter ($s_2 > s_1 > 0$) (b) High-pass global filter ($s_2 < s_1 < 0$) (c) Adaptive local filter

Figure 2. Comparison between global filters and an adaptive local filter. The global filter refers to the conventional filters whose range of feature aggregation is identical across the whole nodes. (a) A low-pass global filter smooths out features among neighboring nodes, (b) while a high-pass global filter accentuates the difference between a target node (e.g., n_1) and its neighboring nodes (e.g., n_2). (c) Unlike these global filters, the adaptive local filter captures optimal bandwidths for each node with trainable node-wise scaling parameters.

C number of diagnostic labels, each graph is assigned to one of these labels y_c ($c = 1, ..., C$). These labels are presented chronologically, i.e., neurodegeneration progresses from a healthy control group (y_1) to the most deteriorated state (y_C). The goal of AGT is to classify these neurodegenerative brain networks by characterizing their temporal changes.

4.2. Overview of AGT

As shown in Fig. 1, AGT contains two major components for neurodegenerative brain network classification: **1)** node-level scale adaptation and **2)** group-level temporal regularization. With trainable node-wise scale parameters, the node-level scale adaptation performs a *node-variant convolution* such that each node aggregates features from its optimal range of neighborhood at optimal graph resolutions. Such an approach redefines connectivity with an adaptive wavelet basis that connects nodes even though they are not directly connected. As this basis selectively uses both low and high-frequency spectral characteristics with learnable node-wise scales, the magnitude of smoothing or coarsening the neighbor information is optimally controlled per node. Also, by using C sets of node-level scales for different diagnostic labels, the inter-group differences in each brain region can be investigated.

In a downstream graph convolutional network (i.e., GCN in Fig. 1), these adaptively aggregated node features are further used to extract localized node-edge spectral coherence. In other words, the underlying cross-characteristics of the node and edge features are unified as a refined representation of the whole graph embedding via GCN. Using a set of graph embeddings in a mini-batch, we apply a temporal regularization that captures the sequential dynamics of the diagnostic groups along the degeneration. Finally, a downstream readout function (i.e., $f(\cdot)$ in Fig. 1) takes all embeddings and predicts their labels.

4.3. Node-level Scale Adaptation

Effect of Low and High Pass Filters. In SGWT, the choice of the kernel $k(\cdot)$ determines the graph characteristics to be captured. For example, low-pass filters smooth out local details and capture information from a wide range of neighbor nodes, while high-pass filters highlight local details and abrupt alterations within a graph. The differences between the low and high-pass filters are visualized in Fig. 2a and Fig. 2b. Given a diffusion kernel $k(s\lambda) = e^{-s\lambda}$, Fig. 2a shows the effect of the low-pass filter with $s > 0$. On the other hand, Fig. 2b shows the effect of the high-pass filter with $s < 0$, which is equal to using a filter $k'(s\lambda) = e^{s\lambda}$ with $s > 0$. As shown in these figures, the low-pass filters focus on smaller eigenvalues that carry slow varying signals, while high-pass filters pass high-frequency components corresponding to large eigenvalues.

The scale parameter s controls the bandwidth of the locality in the graph space and determines the magnitude of highlighting or smoothing local details. If the scale gets larger for the low-pass filter, features are aggregated from a wider range of neighbor nodes and the neighbor nodes (e.g., n_1 and n_2 in Fig. 2a) become more similar. On the other hand, a smaller negative scale for the high-pass filter emphasizes the difference between the nodes at a fine range such that the importance of individual nodes is highlighted.

Adaptive Wavelet Basis. Conventional graph convolutions aggregate information from a direct neighborhood uniformly across the entire nodes, which may lead to oversmoothed or undersmoothed graph representations (Wang et al., 2021b). In contrast, SGWT with a diffusion kernel allows to aggregate information from indirect nodes (Xu et al., 2019b). However, the global filters with a graph-level scale (i.e., Fig. 2a and 2b) are still not able to consider node-by-node differences in the range of feature aggregation. Therefore, we propose *adaptive wavelet basis* with an adaptive local

filter that considers node-wise optimal bandwidths. Given a set of trainable scales $\mathbf{s} = \{s_i\}_{i=1}^N$ paired with nodes $\{n_i\}_{i=1}^N$, a node-wise adaptive wavelet basis ψ_{s_i} localized at n_i on n_j is defined as

$$\psi_{s_i,i}(j) = \sum_{l=1}^{N} e^{-s_i\lambda_l} u_l^*(i) u_l(j), \tag{4}$$

where * denotes complex conjugate, u_l (l=1, ..., N) is an eigenvector, and the adaptive local filter is defined as $k(s_i\lambda) = e^{-s_i\lambda}$. That is, each scale s_i flexibly controls the bandwidth of the node n_i considering the local graph structure of n_i. Fig. 2c shows an exemplary result of this node-variant convolution with the adaptive local filter. For example, scales of n_1 and n_2 are converged to negative values and the scales of n_3 and n_4 are positive. In this case, the difference between n_1 and n_2 is magnified, while the n_3 and n_4 collect cluster-related information around themselves. Note that, each node is a registered region of interest (ROI) in the brain and the whole N nodes comprise the whole brain regions. In other words, AGT can figure out the optimal range of smoothing or coarsening local graph features at each brain region.

Class-wise Scale Adaptation. To observe class-wise differences in feature aggregation, we further constructed C number of scale sets $\{\mathbf{s}^c\}_{c=1}^C$, where each set has ROI-wise scales $\mathbf{s}^c = \{s_i^c\}_{i=1}^N$. This class- and node-wise adaptive wavelet basis $\psi_{s_i^c}$ localized at n_i on n_j is defined as

$$\psi_{s_i^c,i}(j) = \sum_{l=1}^{N} e^{-s_i^c\lambda_l} u_l^*(i) u_l(j). \tag{5}$$

Analyzing $\psi_{\mathbf{s}^c}$ enables investigating the variations in the regional range of feature aggregation among different labels.

4.4. Spectral Coherence with Node-variant Convolution

With the adaptive wavelet basis $\psi_{\mathbf{s}^c}$, spectral coherence between the graph structure and node features can be captured in a localized and adaptive way. This is done by projecting the node features $\mathbf{X}$ into the spectral space as $W_{\mathbf{X}}(\mathbf{s}) = \psi_{\mathbf{s}^c} \cdot \mathbf{X}$ and applying the scale-specific inverse transform on the wavelet coefficient $W_{\mathbf{X}}(\mathbf{s})$. As in Eq. (2), a scale-filtered node feature $\mathbf{X_s} = \psi_{\mathbf{s}^c} \cdot W_{\mathbf{X}}(\mathbf{s}) \in \mathbb{R}^{N\times F}$ is obtained via IGWT, where each element is the result of the node-variant convolution in the spatial space. According to Eq. (3), the locally adjusted spectral coherence is derived in a graph convolutional form as

$$W_{\mathbf{A}}(\mathbf{s}) \cdot W_{\mathbf{X}}(\mathbf{s}) = \mathbf{A}\mathbf{X_s}, \tag{6}$$

which contains cross-characteristics of graph structure and node features at optimal node-wise resolution (i.e., scale).

While the node-variant convolution allows feature aggregation from indirect nodes with different amounts for each node, conventional graph convolution is a *node-invariant convolution*, where all nodal features are aggregated from the equal range of their direct neighbor nodes. To take full advantage of both direct and indirect neighbor information, we fuse the given node feature $\mathbf{X}$ and the scale-filtered node feature $\mathbf{X_s}$ together with conventional graph convolutions using $\mathbf{A}$. Given a stacked node feature $\tilde{\mathbf{X}} = [\mathbf{X}, \mathbf{X_s}] \in \mathbb{R}^{N\times 2F}$, a 2-layer graph convolution yields a graph embedding e as

$$e = \sigma(\mathbf{A}\ \sigma(\mathbf{A}\tilde{\mathbf{X}}W^{(1)})\ W^{(2)}), \tag{7}$$

where $W^{(1)}$ and $W^{(2)}$ are trainable weights and $\sigma(\cdot)$ is ReLU activation. As a result, the graph embedding e comprehensively contains both the scale-filtered spectral characteristics $\mathbf{AX_s}$ and the spatial graph features $\mathbf{AX}$ aggregated from both indirect and direct edges.

4.5. Training Objective with Temporal Regularization

Group-wise Temporal Regularization. In contrast to most classification tasks where the relationship between labels is absent, neurodegenerative diseases progress sequentially, resulting in temporal dynamics between labels. Considering this degenerative characteristic of brain networks, we devised a group-level temporal regularization that extracts temporal relations between the diagnostic groups.

In a mini-batch training with B data, a set of graph embeddings $\{e_b\}_{b=1}^B$ is obtained with Eq. (7). These embeddings are categorized into C groups based on each of their labels $y_b \in \{y_1, ..., y_C\}$. From C groups of embeddings, the averaged graph representation is derived for each class as

$$\bar{e}_c = \frac{\sum_{b=1}^{B} e_b \cdot \mathbb{I}(y_b = y_c)}{\sum_{b=1}^{B} \mathbb{I}(y_b = y_c)}, \tag{8}$$

where the $\mathbb{I}(y_b = y_c)$ is an indicator function that outputs 1 if the label of the e_b is y_c and 0 otherwise.

Using $l2$-norm, the distance between the adjacent classes (i.e., c-th class and $(c+1)$-th class) is quantified as

$$d_{c,c+1} = ||\bar{e}_c - \bar{e}_{c+1}||_{l2}, \tag{9}$$

for $c = 1, ..., C-1$. To enforce the sequential relations between classes, triangular distances between three adjacent classes (i.e., $d_{c,c+1}$, $d_{c+1,c+2}$, and $d_{c,c+2}$) should be aligned in the latent feature space as

$$d_{c,c+1} + d_{c+1,c+2} = d_{c,c+2}. \tag{10}$$

Therefore, the temporal regularization R_{temp} is defined as

$$R_{\text{temp}} = \frac{1}{C-2} \sum_{c=1}^{C-2} \Big(d_{c,c+1} + d_{c+1,c+2} - d_{c,c+2} \Big). \tag{11}$$

Without this temporal regularization, the sets of graph embeddings $\{\bar{e}_c\}_{c=1}^C$ are highly likely to be arbitrarily aligned

in the feature space, posing a challenge for the model to consider their temporal relationships between classes.

Training Objective. To obtain a graph label prediction $\hat{y}$, the graph embeddings are inputted into the downstream readout function $f(\cdot)$ (i.e., MLP with ReLU). Given a true label y, a classical cross-entropy loss with ℓ_2-regularization is used as

$$L_{\text{ce}} = -\frac{1}{B}\sum_{b=1}^{B}\sum_{c=1}^{C} y_{bc} \cdot \log(\hat{y}_{bc}) + \frac{\gamma}{2B}||\mathbf{W}||_{\ell_2}^2, \quad (12)$$

where $\mathbf{W}$ is a set of trainable parameters and γ controls the strength of ℓ_2-regularization. With a hyperparameter α, the whole training loss $\mathcal{L} = L_{\text{ce}} + \alpha R_{\text{temp}}$ is minimized via backpropagation to optimize the network parameters for graph classification.

5. Experiment

In this section, we report the quantitative superiority of AGT and discuss clinical interpretations, the effect of model components, and ablation studies.

5.1. Datasets

We used two independent neurodegenerative brain network datasets: Alzheimer's Disease Neuroimaging Initiative (ADNI) and Parkinson's Progression Markers Initiative (PPMI), whose demographics are given in Table 1 and 2.

ADNI Dataset. ADNI study (Mueller et al., 2005) provides the largest public Alzheimer's Disease (AD) data with diverse biomarkers from multi-modal imaging. Structural brain networks were measured by a probabilistic tractography from Diffusion Weighted Images (DWIs) on the Destrieux atlas (Destrieux et al., 2010) with 160 regions of interest (ROIs). As node features, cortical thickness (CT) from Magnetic Resonance Imaging (MRI) and Standardized Uptake Value Ratio (SUVR) of fluorodeoxyglucose (FDG) from Positron Emission Tomography (PET) at all ROIs were measured. We performed two separate experiments on these two representative AD biomarkers. Five diagnostic labels were given: Cognitively Normal (CN), Significant Memory Concern (SMC), Early Mild Cognitive Impairment (EMCI), Late MCI (LMCI), and AD. The disease progressively deteriorates from CN to AD.

PPMI Dataset. PPMI study (Marek et al., 2011) provides public biomarkers for Parkinson's Disease (PD) progression. For a total of 195 subjects, functional MRI from individuals were parcellated by co-registration with the AAL atlas (Tzourio-Mazoyer et al., 2002) comprising 116 regions. As node features, ROI-wise Blood-Oxygen-Level-Dependent (BOLD) signals were obtained from the images. The correlations between different regions were calculated and used as connectivities. Both node and edge features were preprocessed by authors in (Xu et al., 2023). PPMI contains three labels: CN, Prodromal, and PD. Similar to the ADNI, the subjects' condition in these labels progressively deteriorates from CN to PD.

Table 1. Demographics of the ADNI dataset.

Biomarker	Category	CN	SMC	EMCI	LMCI	AD
Cortical Thickness	# of subjects	359	181	437	180	166
	Gender (M / F)	178 / 181	69 / 112	249 / 188	119 / 61	102 / 64
	Age (Mean±Std)	72.8±1.4	72.0±5.2	71.0±7.9	70.9±6.1	74.8±8.7
FDG	# of subjects	345	186	461	231	162
	Gender (M / F)	173 / 172	66 / 120	262 / 199	152 / 79	102 / 60
	Age (Mean±Std)	73.0±1.3	71.7±5.2	71.7±7.8	71.1±7.0	74.9±8.8

Table 2. Demographics of the PPMI dataset.

Biomarker	Category	CN	Prodromal	PD
BOLD	# of subjects	15	67	113
	Gender (M / F)	12 / 3	38 / 29	77 / 36
	Age (Mean±Std)	64.0±9.5	64.3±8.7	62.0±10.0

5.2. Experimental Setup

All experiments were performed with 5-fold cross-validation and the resultant accuracy, precision, recall, and specificity from all folds were averaged to avoid any biases. As the ground truth labels should not be inputted to a model during inference, we used pseudo-labels from a pretrained Exact (Choi et al., 2022) 1) to select class-wise trained scales and 2) to assign graph embeddings to diagnostic groups to calculate a validation loss. We used 5 pretrained Exact models, each trained on the same training set of individual folds used in the AGT training. We further provide implementation details of AGT in Appendix A. As baselines, we used eleven methods as follows: Linear Support Vector Machine (SVM), Multi-Layer Perceptron (MLP) with 2 layers, GCN (Kipf & Welling, 2017), GAT (Veličković et al., 2018) GDC (Gasteiger et al., 2019), GraphHeat (Xu et al., 2019b), ADC (Zhao et al., 2021), Exact (Choi et al., 2022), LSAP (Sim et al., 2024), BrainGNN (Li et al., 2021), and Brain Network Transformer (BrainNetTF) (Kan et al., 2022).

5.3. Quantitative Results

For the ADNI and PPMI datasets, the quantitative performance of AGT and baselines are compared in Table. 3. In all experiments and across all metrics, AGT consistently outperformed all baseline methods. Specifically, on the ADNI dataset with 5 labels, AGT showed an accuracy of 90.3% for the ADNI with cortical thickness (ADNI-CT) and 94.8% for the ADNI with FDG (ADNI-FDG) experiments. Compared to the second-best results, AGT gained 2.5%p and 4.4%p accuracy margins on the ADNI-CT and ADNI-FDG data, respectively. For the PPMI dataset, AGT showed

Table 3. Classification performance on the ADNI and PPMI datasets with 5-fold cross-validation. The best results are marked in bold and the second-best results are indicated by an underline.

Model	ADNI-CT			ADNI-FDG			PPMI			
	Accuracy (%)	Precision (%)	Recall (%)	Accuracy (%)	Precision (%)	Recall (%)	Accuracy (%)	Precision (%)	Recall (%)	Specificity (%)
SVM (Linear)	82.4 ± 2.7	82.2 ± 3.3	85.2 ± 2.5	85.3 ± 2.1	85.7 ± 2.7	86.9 ± 2.1	60.5 ± 10.1	30.2 ± 6.6	28.0 ± 8.2	67.5 ± 2.4
MLP (2-layers)	78.8 ± 2.2	79.2 ± 3.6	79.9 ± 2.6	87.5 ± 1.6	88.2 ± 2.4	88.2 ± 1.4	68.9 ± 3.5	36.3 ± 4.4	39.0 ± 8.6	70.0 ± 3.8
GCN	61.4 ± 3.1	59.8 ± 2.5	62.6 ± 4.4	68.8 ± 2.0	67.7 ± 2.8	69.7 ± 2.5	78.8 ± 2.1	48.1 ± 5.2	70.3 ± 4.3	82.5 ± 2.8
GAT	64.2 ± 5.5	62.7 ± 6.7	66.8 ± 4.6	69.2 ± 7.1	67.0 ± 10.6	73.6 ± 3.7	81.2 ± 2.4	51.4 ± 6.9	77.2 ± 5.5	87.0 ± 2.4
GDC	77.1 ± 4.3	76.9 ± 5.0	78.5 ± 4.4	86.2 ± 3.2	86.7 ± 3.3	87.0 ± 2.9	73.0 ± 0.7	36.5 ± 3.1	61.8 ± 9.4	79.1 ± 4.4
GraphHeat	70.9 ± 3.2	70.3 ± 3.0	71.8 ± 2.6	77.0 ± 2.4	77.5 ± 3.5	77.3 ± 1.0	79.1 ± 2.0	48.4 ± 4.7	84.5 ± 3.0	68.8 ± 6.3
ADC	82.1 ± 2.4	77.6 ± 1.9	72.8 ± 6.7	88.6 ± 2.8	70.8 ± 6.2	75.3 ± 5.3	78.8 ± 2.3	50.7 ± 8.0	66.9 ± 5.5	80.5 ± 2.5
Exact	86.2 ± 2.0	86.6 ± 1.7	86.7 ± 2.3	90.2 ± 2.7	90.7 ± 2.8	90.7 ± 2.8	79.5 ± 2.4	48.1 ± 5.4	76.6 ± 9.0	87.3 ± 1.5
LSAP	87.0 ± 2.2	86.8 ± 2.7	88.5 ± 2.7	90.4 ± 1.4	90.9 ± 1.8	91.4 ± 1.5	79.1 ± 2.0	47.5 ± 4.8	83.2 ± 3.7	72.9 ± 4.4
BrainGNN	69.3 ± 2.8	20.1 ± 0.5	23.4 ± 3.8	68.9 ± 2.4	20.3 ± 0.4	31.9 ± 13.3	69.6 ± 5.1	38.5 ± 7.9	70.5 ± 5.0	39.9 ± 10.2
BrainNetTF	87.8 ± 3.9	65.9 ± 16.7	70.6 ± 9.0	87.1 ± 5.3	66.5 ± 16.5	66.3 ± 15.6	71.3 ± 4.1	42.6 ± 10.5	76.1 ± 6.3	53.3 ± 16.9
AGT (Ours)	**90.3 ± 1.8** (+2.5)	**91.3 ± 2.4** (+4.5)	**89.9 ± 2.5** (+1.4)	**94.8 ± 1.1** (+4.4)	**94.3 ± 1.5** (+3.4)	**95.3 ± 1.4** (+3.9)	**83.6 ± 3.8** (+2.4)	**62.4 ± 4.4** (+11.0)	**87.6 ± 3.5** (+3.1)	**92.3 ± 2.2** (+5.0)

Table 4. Five brain regions with the smallest trained scales for the AD group in the ADNI dataset and the PD group in the PPMI dataset. The regional scales were averaged across 5 trained models from 5 folds.

idx	ADNI-CT						ADNI-FDG						PPMI			
	ROI	CN	SMC	EMCI	LMCI	AD	ROI	CN	SMC	EMCI	LMCI	AD	ROI	CN	Prodromal	PD
1	l s.oc.temp.med.lingual	-1.33	0.79	-0.72	0.03	-1.13	l g.oc.temp.med.lingual	-0.53	-0.23	-0.09	0.06	-1.40	l temporal.sup	-0.768	0.002	-1.410
2	r amygdala	0.01	-0.74	0.41	-0.05	-1.12	r s.precentral.sup.part	-0.73	-1.18	0.33	0.24	-1.12	l frontal.inf.oper	1.137	-0.629	-1.166
3	l s.central	0.33	0.42	-0.92	-0.49	-1.07	r thalamus.proper	0.07	1.57	-0.19	0.90	-1.12	r calcarine	-0.213	-0.115	-1.166
4	r s.front.inf	-0.64	0.67	0.51	-0.52	-0.97	l g.occipital.sup	-0.16	-0.25	0.12	-0.07	-1.04	r frontal.med.orb	-0.394	1.152	-1.110
5	r g.oc.temp.med.parahip	-0.06	0.2	-0.25	0.36	-0.93	l g.temp.sup.lateral	-0.46	-0.83	1.03	-0.29	-0.98	r temporal.sup	-0.267	-0.273	-1.077

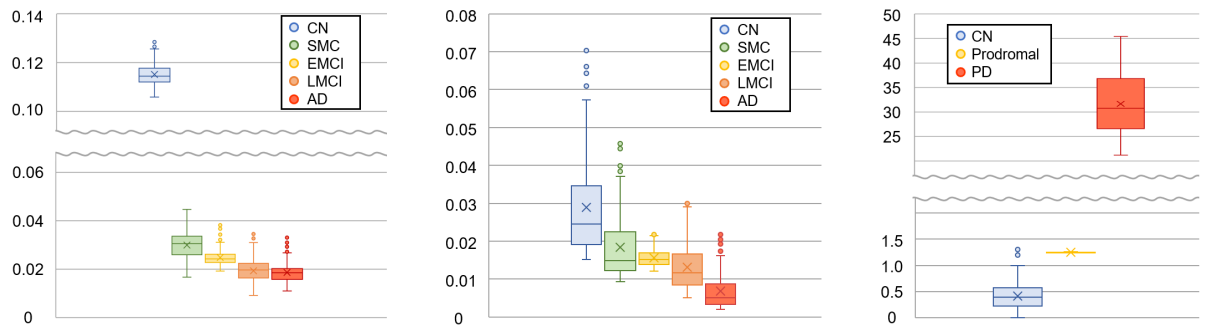
83.6% accuracy with a 2.4%p margin over the second-best result. Notably, AGT far surpassed the second-best results by 11.0%p in precision and 5.0%p in specificity, demonstrating its generalizability and superiority.

5.4. Discussions on Node-wise Significance for Brain Connectome Classification

As AGT performs a graph classification with node-wise and class-specific trainable scales, the trained model yields node-wise optimized scale values for each diagnostic class. Positive scales indicate that the graph wavelet is focusing on low-frequency information smoothed from a wide range around the corresponding ROIs, whereas negative scales indicate that the graph wavelet is focusing on high-frequency and node-individual features. In other words, the high-frequency components with smaller and negative scales focus on more node-specific details for solving the task and thus potentially highlight the importance of individual nodes. Therefore, we investigated the significance of the brain regions with small scales by analyzing clinical relationships between these regions and the disease. Table 4 shows five brain regions with the smallest scales for the most deteriorated groups (i.e., AD and PD) from the ADNI and PPMI experiments, respectively.

ROI Analysis in AD. As shown in Table 4, the identified regions in the ADNI experiments are primarily distributed in *temporal* and *occipital* lobes. Notably, the *medial occipitotemporal gyrus* exhibited the smallest scale in the ADNI-CT experiment, a region known for its vital role in encoding visual memories (Bogousslavsky et al., 1987). Another noteworthy region is the *parahippocampal gyrus*, widely recognized for its association with AD and its role in the formation of new memories (Mu & Gage, 2011). Additionally, it's worth noting that *amygdala* is identified, which plays a role in the regulation of long-term memory (McGaugh, 2004). In the ADNI-FDG experiment, *the medial occipito-temporal sulcus* and *the thalamus* were identified, both functionally connected to the hippocampus (Aggleton & Brown, 1999; Huntgeburth & Petrides, 2012), and associated with spatial memory (Aggleton et al., 2010).

ROI Analysis in PD. The identified regions for the PPMI experiment were distributed across the *temporal* and *frontal* lobes. Both the left and right *superior temporal gyrus* were identified, where neural correlations between these regions and PD have been observed in many studies (Martin et al., 2009; Péron et al., 2010; Gu et al., 2022). These regions are associated with social cognition and language comprehension (Bigler et al., 2007). Also, the *superior frontal gyrus* is implicated in the sensory system and facial expressions (Goldberg et al., 2006; Fried et al., 1998), and the *inferior frontal gyrus* plays a crucial role in language processing and motor control (Swick et al., 2008). Given that PD patients experience challenges in both motor and non-motor functions, including cognitive impairments such as deficits in attention and memory, these ROIs related to social cognitive and language processing may contribute to understanding the cognitive aspects of PD.



(a) ADNI with cortical thickness (b) ADNI with FDG (c) PPMI with BOLD

Figure 3. Comparison of distributions of the graph embedding $\bar{e}_c$ between diagnostic classes. Each box plot represents the graph embedding distribution of one diagnostic group, where the distribution is obtained from the trained AGT models from 5 folds and averaged across the folds. Temporal variations between the groups appear along with the disease progression.

Table 5. Comparison of mean accuracy with and without R_{temp}.

L_{ce}	R_{temp}	ADNI-CT	ADNI-FDG	PPMI
✓		89.2 ± 2.1	94.4 ± 1.3	82.6 ± 2.1
✓	✓	**90.3 ± 1.8 (+0.9)**	**94.8 ± 1.1 (+0.4)**	**83.6 ± 3.8 (+1.0)**

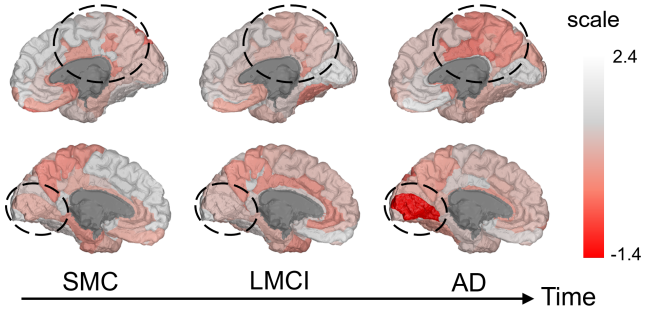


Figure 4. Visualization of the trained scales on the ADNI-FDG experiment. Top: Trained scales of right cortex ROIs. Bottom: Trained scales of left cortex ROIs. All node-wise scales are averaged across 5 folds.

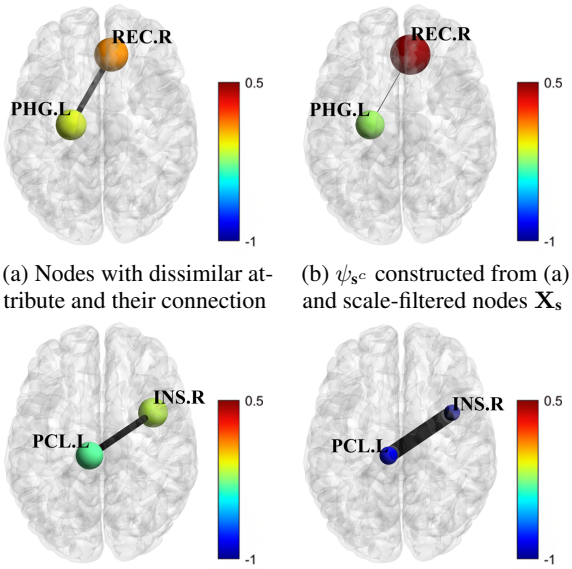
5.5. Temporal Analysis

The result of the ablation study on R_{temp} is reported in Table 5. With R_{temp}, the averaged accuracy consistently increased $\sim 1\%p$ for all experiments. Additional precision and recall results are reported in Appendix B, where the performance improvement with R_{temp} on these metrics are $\sim 2.5\%p$ and $\sim 1.2\%p$, respectively.

Temporal dynamics across degenerative classes are visualized in Fig. 3 and 4. As R_{temp} imposes sequential regularity on graph embeddings (i.e., $\bar{e}_c$ in Eq. (8)), we observed class-wise embedding distributions from a trained model. Fig. 3 shows a comparison of the embedding distributions with different diagnostic classes, revealing sequential variations along the stages of degeneration in both ADNI and PPMI datasets. In Fig. 4, regional changes in scales are observable across classes, especially for marked regions within circles. Specifically, scales of *the precuneus gyrus* and *the medial occipitotemporal gyrus* decrease at later stages in the top and bottom panels, respectively. These results indicate that the significance of high-frequency (i.e., node-individual) information is increased in these ROIs as the disease progresses. Also, these results are consistent with existing studies, as they are known to be regions where early-onset AD symptoms appear (Karas et al., 2007; Convit et al., 2000).

5.6. Effect of Node-variant Convolution

In contrast to typical graph convolution, our node-variant convolution with adaptive wavelet basis $\psi_{\mathbf{s}^c}$ aggregates features within an optimally tailored range for each node. Also, as $\psi_{\mathbf{s}^c}$ considers both high and low-frequency characteristics, it can capture both heterophily and homophily, that may coexist within a single graph. To assess the impact of node-variant convolution, we compare a sample input graph from a PD patient in the PPMI dataset with the outcome of the node-variant convolution on this graph. In the brain network of subject #3130, both heterophily and homophily coexist, which appear in Fig. 5a and 5c, respectively. The cosine similarity of BOLD signals between *the gyrus rectus* and *the parahippocampal gyrus* in Fig. 5a is -0.86, while the cosine similarity between *the paracentral lobule* and *the insula* in Fig. 5c is 0.50. Edges in Fig. 5b and 5d show the $\psi_{\mathbf{s}^c}$ calculated from a trained model, where the connectivity magnitude weakens for heterophily and strengthens for homophily. By applying node-variant convolution using $\psi_{\mathbf{s}^c}$, the difference in two node pairs is amplified. Specifically, cosine similarity is decreased to -0.87 in Fig. 5b and increased to 0.52 in Fig. 5d. These results demonstrate the flexibility of node-variant convolution in capturing localized and heterogeneous graph characteristics.



(a) Nodes with dissimilar attribute and their connection

(b) $\psi_{\mathbf{s}^c}$ constructed from (a) and scale-filtered nodes $\mathbf{X}_\mathbf{s}$

(c) Nodes with similar attribute and their connection

(d) $\psi_{\mathbf{s}^c}$ constructed from (c) and scale-filtered nodes $\mathbf{X}_\mathbf{s}$

Figure 5. (a) and (c) are from a brain network (i.e., $\mathbf{X}$ and $\mathbf{A}$) of subject #3130. Edges in (b) and (d) are the adaptive wavelet basis $\psi_{\mathbf{s}^c}$ derived from a trained model. By using $\psi_{\mathbf{s}^c}$, the node features in (b) become more discriminative from (a), while those in (d) become more similar compared to the nodes in (c).

5.7. Ablation Study on the Adaptive Local Filter

We performed an ablation study on the adaptive local filter by replacing it with a high-pass or low-pass filter. As the scales of the adaptive filter were randomly initialized within a range $[-2, 2]$ and converged within a similar range (as shown in Table 12, 13, 14), we performed six experiments by setting all scales to be fixed as one of $\{-2, -1, -0.5, 0.5, 1, 2\}$ to capture high and low-frequency graph characteristics. As shown in Table 6, the experiment with adaptive scales far outperforms the fixed scale settings as each node-wise scale is flexibly trained to aggregate information from different nodes with different strengths.

Table 6. Ablation study on the adaptive local filter for the ANDI-CT experiment. The adaptive local filter is replaced with a high-pass or a low-pass filter with a fixed scale.

scale	Accuracy (%)	Precision (%)	Recall (%)
-2 (high-pass)	72.3 ± 1.2	70.6 ± 1.9	74.7 ± 1.2
-1 (high-pass)	73.8 ± 1.7	73.1 ± 1.0	75.2 ± 2.1
-0.5 (high-pass)	74.5 ± 0.8	73.4 ± 1.6	76.6 ± 1.6
0.5 (low-pass)	74.5 ± 1.8	72.5 ± 2.4	76.0 ± 3.8
1 (low-pass)	74.5 ± 2.0	73.4 ± 2.0	75.3 ± 2.1
2 (low-pass)	74.0 ± 2.0	72.5 ± 1.9	76.4 ± 2.7
Adaptive scales	**90.3 ± 1.8** (+15.8)	**91.3 ± 2.4** (+17.9)	**89.9 ± 2.5** (+13.3)

6. Conclusion

We presented a novel wavelet-based GNN addressing challenges in analyzing the evolving dynamics of neurodegenerative diseases on brain connectomes. Our method captures the sequential variations within diagnostic groups with a group-wise temporal regularization. Also, it adaptively captures both homophily and heterophily within a graph by adjusting node-wise spectral bandwidth. As a result, our method outperformed various GNNs for brain connectome classification on two representative neurodegenerative disease benchmarks. Our framework offers clinical interpretability, showing significant potential to be applied in the analysis of various neurodegenerative diseases.

Acknowledgements

This research was supported by NRF-2022R1A2C2092336 (50%), RS-2022-II220290 (20%), RS-2022-KH127855 (10%), RS-2022-KH128705 (10%), and RS-2019-II191906 (AI Graduate Program at POSTECH, 10%).

Impact Statement

This paper presents work whose goal is to advance the field of machine learning and its application to brain network analysis to understand the progress of neurodegenerative diseases and cognitive functions. There are many potential societal consequences of our work, especially given the scope of our application, such as detecting early diagnosis of irreversible brain disorders to prevent the disease or even lead to their cure.

References

Aggleton, J. P. and Brown, M. W. Episodic memory, amnesia, and the hippocampal–anterior thalamic axis. *Behavioral and brain sciences*, 22(3):425–444, 1999.

Aggleton, J. P., O'Mara, S. M., et al. Hippocampal–anterior thalamic pathways for memory: uncovering a network of direct and indirect actions. *European Journal of Neuroscience*, 31(12):2292–2307, 2010.

Axelrod et al. GEOM: Energy-annotated molecular conformations for property prediction and molecular generation. *Scientific Data*, 9(1):185, 2022.

Bigler, E. D., Mortensen, S., et al. Superior temporal gyrus, language function, and autism. *Developmental neuropsychology*, 31(2):217–238, 2007.

Bogousslavsky, J., Miklossy, J., et al. Lingual and fusiform gyri in visual processing: a clinico-pathologic study of superior altitudinal hemianopia. *Journal of neurology, neurosurgery, and psychiatry*, 50(5):607, 1987.

Chen, C., Petty, K., et al. Freeway performance measurement system: mining loop detector data. *Transportation Research Record*, 1748(1):96–102, 2001.

Chien, E., Peng, J., et al. Adaptive universal generalized pagerank graph neural network. *International Conference on Machine Learning (ICML)*, 2021.

Cho, H., Jeong, M., et al. Multi-resolution spectral coherence for graph generation with score-based diffusion. In *Conference on Neural Information Processing Systems (NeurIPS)*, 2023.

Choi, I., Wu, G., and Kim, W. H. How much to aggregate: Learning adaptive node-wise scales on graphs for brain networks. In *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pp. 376–385. Springer, 2022.

Convit, A., De Asis, J., et al. Atrophy of the medial occipitotemporal, inferior, and middle temporal gyri in non-demented elderly predict decline to Alzheimer's disease. *Neurobiology of aging*, 21(1):19–26, 2000.

Defferrard, M., Bresson, X., and Vandergheynst, P. Convolutional neural networks on graphs with fast localized spectral filtering. *Advances in neural information processing systems (NeurIPS)*, 29, 2016.

Destrieux, C., Fischl, B., et al. Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. *Neuroimage*, 53(1):1–15, 2010.

Dong, Y., Ding, K., et al. Adagnn: Graph neural networks with adaptive frequency response filter. In *ACM International Conference on Information and Knowledge Management (CIKM)*, pp. 392–401, 2021.

Donnat, C., Zitnik, M., et al. Learning structural node embeddings via diffusion wavelets. In *ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD)*, pp. 1320–1329, 2018.

Fried, I., Wilson, C. L., et al. Electric current stimulates laughter. *Nature*, 391(6668):650–650, 1998.

Gao, H. and Ji, S. Graph U-Nets. In *International Conference on Machine Learning (ICML)*, pp. 2083–2092. PMLR, 2019.

Gasteiger, J., Weißenberger, S., and Günnemann, S. Diffusion improves graph learning. *Advances in Neural Information Processing Systems (NeurIPS)*, 32, 2019.

Goldberg, I. I., Harel, M., and Malach, R. When the brain loses its self: prefrontal inactivation during sensorimotor processing. *Neuron*, 50(2):329–339, 2006.

Gu, L., Shu, H., et al. Functional brain changes in Parkinson's disease: A whole brain ALE study. *Neurological Sciences*, 43(10):5909–5916, 2022.

Hammond, D. K., Vandergheynst, P., and Gribonval, R. Wavelets on graphs via spectral graph theory. *Applied and Computational Harmonic Analysis*, 30(2):129–150, 2011.

Huntgeburth, S. C. and Petrides, M. Morphological patterns of the collateral sulcus in the human brain. *European Journal of Neuroscience*, 35(8):1295–1311, 2012.

Kan, X., Dai, W., et al. Brain network transformer. *Advances in Neural Information Processing Systems (NeurIPS)*, 35: 25586–25599, 2022.

Karas, G., Scheltens, P., et al. Precuneus atrophy in early-onset Alzheimer's disease: a morphometric structural MRI study. *Neuroradiology*, 49:967–976, 2007.

Kim, W., Pachauri, D., et al. Wavelet based multi-scale shape features on arbitrary surfaces for cortical thickness discrimination. *Advances in Neural Information Processing Systems (NeurIPS)*, 25, 2012.

Kipf, T. N. and Welling, M. Semi-supervised classification with graph convolutional networks. *International Conference on Learning Representations (ICLR)*, 2017.

Li, X., Zhou, Y., et al. BrainGNN: Interpretable brain graph neural network for fMRI analysis. *Medical Image Analysis*, 74:102233, 2021.

Liu, L., Wen, G., et al. BrainTGL: A dynamic graph representation learning model for brain network analysis. *Computers in Biology and Medicine*, 153:106521, 2023.

Mallat, S. *A wavelet tour of signal processing*. Elsevier, 1999.

Mao, B., Huang, J., and Zhang, D. Node based row-filter convolutional neural network for brain network classification. In *Pacific Rim International Conference on Artificial Intelligence*, pp. 1069–1080. Springer, 2018.

Marek, K., Jennings, D., et al. The Parkinson progression marker initiative (PPMI). *Progress in neurobiology*, 95 (4):629–635, 2011.

Martin, W. W., Wieler, M., et al. Temporal lobe changes in early, untreated Parkinson's disease. *Movement Disorders*, 24(13):1949–1954, 2009.

McGaugh, J. L. The amygdala modulates the consolidation of memories of emotionally arousing experiences. *Annual Review of Neuroscience*, 27:1–28, 2004.

Mu, Y. and Gage, F. H. Adult hippocampal neurogenesis and its role in Alzheimer's disease. *Molecular neurodegeneration*, 6(1):1–9, 2011.

Mueller, S. G., Weiner, M. W., et al. The Alzheimer's disease neuroimaging initiative. *Neuroimaging Clinics*, 15(4):869–877, 2005.

Park, J., Hwang, Y., et al. Convolving directed graph edges via hodge laplacian for brain network analysis. In *International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI)*, pp. 789–799. Springer, 2023.

Péron, J., Le Jeune, F., et al. Subthalamic nucleus stimulation affects theory of mind network: a PET study in Parkinson's disease. *PLoS One*, 5(3):e9919, 2010.

Qu, Z., Yao, T., et al. A graph convolutional network based on univariate neurodegeneration biomarker for Alzheimer's disease diagnosis. *IEEE Journal of Translational Engineering in Health and Medicine*, 2023.

Ramakrishnan, R. et al. Quantum chemistry structures and properties of 134 kilo molecules. *Scientific data*, 1(1): 1–7, 2014.

Relión, J. D. A., Kessler, D., et al. Network classification with applications to brain connectomics. *The annals of applied statistics*, 13(3):1648, 2019.

Sim, J., Jeon, S., et al. Learning to approximate adaptive kernel convolution on graphs. *The AAAI Conference on Artificial Intelligence (AAAI)*, 2024.

Swick, D., Ashley, V., and Turken, A. U. Left inferior frontal gyrus is critical for response inhibition. *BMC neuroscience*, 9:1–11, 2008.

Tzourio-Mazoyer, N., Landeau, B., et al. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage*, 15(1):273–289, 2002.

Veličković, P., Cucurull, G., et al. Graph attention networks. *International Conference on Learning Representations (ICLR)*, 2018.

Wang, L., Huang, C., et al. Graph embedding via diffusion-wavelets-based node feature distribution characterization. In *ACM International Conference on Information & Knowledge Management (CIKM)*, pp. 3478–3482, 2021a.

Wang, M., Shao, W., et al. Hypergraph-regularized multimodal learning by graph diffusion for imaging genetics based Alzheimer's disease diagnosis. *Medical Image Analysis*, 89:102883, 2023.

Wang, Y., Wang, Y., et al. Dissecting the diffusion process in linear graph convolutional networks. *Advances in Neural Information Processing Systems (NeurIPS)*, 34: 5758–5769, 2021b.

Wang, Z., Zhu, X., et al. Multi-modal classification of neurodegenerative disease by progressive graph-based transductive learning. *Medical image analysis*, 39:218–230, 2017.

Wu, F., Souza, A., et al. Simplifying graph convolutional networks. In *International Conference on Machine Learning (ICML)*, pp. 6861–6871. PMLR, 2019.

Xu, B., Shen, H., et al. Graph wavelet neural network. *International Conference on Learning Representations (ICLR)*, 2019a.

Xu, B., Shen, H., et al. Graph convolutional networks using heat kernel for semi-supervised learning. *The International Joint Conference on Artificial Intelligence (IJCAI)*, 2019b.

Xu, J., Yang, Y., et al. Data-driven network neuroscience: On data collection and benchmark. *Conference on Neural Information Processing Systems (NeurIPS), Datasets and Benchmarks Track*, 2023.

Yanardag, P. and Vishwanathan, S. Deep graph kernels. In *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, pp. 1365–1374, 2015.

Zhang, M., Cui, Z., et al. An end-to-end deep learning architecture for graph classification. In *Proceedings of the AAAI conference on artificial intelligence (AAAI)*, volume 32, 2018.

Zhao, J., Dong, Y., et al. Adaptive diffusion in graph neural networks. *Advances in Neural Information Processing Systems (NeurIPS)*, 34:23321–23333, 2021.

Zhu, Y., Zhu, X., et al. Dynamic hyper-graph inference framework for computer-assisted diagnosis of neurodegenerative diseases. *IEEE Transactions on Medical Imaging*, 38(2):608–616, 2018.

In the appendix, we present **1)** detailed implementation setting, **2)** ablation study on the weight for the R_{temp}, **3)** ablation study on R_{temp} with additional metrics, and **4)** converged scale values for all datasets, which were not included in the main manuscript due to the page limit.

A. Detailed Implementation Setting

To train AGT, we utilized PyTorch framework with a single NVIDIA RTX 6000 Ada Generation GPU. In Table 7, we provide details of the implementation settings of AGT. We performed a grid search for all baselines and AGT to choose the best number of hidden units in $\{8, 16, 32, 64\}$, and a learning rate in $\{0.1, 0.01, 0.001, 0.0001\}$. As the temporal regularization R_{temp} requires a sufficient amount of data for measuring accurate group distance, we used the maximum size of batch comprising all training data. For the scale initialization, N number of scales were initialized randomly within a range $[-2, 2]$.

Table 7. Hyperparameters of AGT for all datasets.

Hyperparameter	ADNI-CT	ADNI-FDG	PPMI
Optimizer	Adam	Adam	Adam
Learning rate	1×10^{-2}	2×10^{-2}	1×10^{-3}
Weight for R_{temp} (i.e., α)	5×10^{-3}	5×10^{-3}	3×10^{-2}
Weight decay (i.e., γ in Eq. 12)	5×10^{-4}	5×10^{-4}	5×10^{-4}
Batch size	1058	1108	156
Number of epochs	5000	5000	5000
Initialization of scale range	[-2, 2]	[-2, 2]	[-2, 2]
Hidden dimension of GCN	8	8	16
Number of GCN layers	2	2	2
Number of $f(\cdot)$ layers	2	2	2

B. Ablation Studies

B.1. Weight for R_{temp}

In Table 8, we provide the results of the averaged accuracy with different trade-off weights (i.e., α) for R_{temp}. We performed a grid search to choose the best α, in $\{0.001, 0.001, 0.003, 0.005, 0.01, 0.015, 0.02, 0.03\}$. As demonstrated in Table 8, AGT exhibits insensitivity to the intensity of the α, with approximately $\sim 1\%p$ gap between the worst and best cases in the ADNI-CT and $\sim 2.5\%p$ for the ADNI-FDG and PPMI experiments.

Table 8. Averaged accuracy with different α configuration.

α	ADNI-CT	ADNI-FDG	PPMI
0.0001	89.80	94.51	81.03
0.001	89.72	94.51	81.54
0.003	89.87	94.72	82.09
0.005	**90.33**	**96.75**	83.08
0.010	89.95	94.58	82.56
0.015	89.72	94.66	82.05
0.02	89.50	94.58	82.05
0.03	89.80	94.44	**83.59**

B.2. Ablation study on R_{temp} with more metrics

Here we present the results for the three experiments with different loss configurations. Along with the averaged accuracy given in Table 5, additional precision and recall scores are reported in Table 9, 10, and 11 for the ADNI-CT, ADNI-FDG, and PPMI experiments, respectively. For the recall on the ADNI-CT experiment, AGT with R_{temp} showed a marginal decrease ($-0.4\%p$) compared to the AGT without R_{temp}. However, except in this case, AGT with R_{temp} consistently outperformed AGT without R_{temp} for all experiments, showing the efficiency of the temporal regularization in solving neurodegenerative graph classification task.

Table 9. Ablation studies on R_{temp} for the ADNI-CT experiment. All results are averaged across 5-folds.

L_{ce}	R_{temp}	Accuracy (%)	Precision (%)	Recall (%)
✓		89.2 ± 2.1	88.8 ± 2.6	90.3 ± 2.7
✓	✓	**90.3 ± 1.8 (+0.9)**	**91.3 ± 2.4 (+2.5)**	**89.9 ± 2.5 (-0.4)**

Table 10. Ablation studies on R_{temp} for the ADNI-FDG experiment. All results are averaged across 5-folds.

L_{ce}	R_{temp}	Accuracy (%)	Precision (%)	Recall (%)
✓		94.4 ± 1.3	94.0 ± 1.2	94.8 ± 1.0
✓	✓	**94.8 ± 1.1 (+0.4)**	**94.3 ± 1.5 (+0.3)**	**95.3 ± 1.4 (+0.5)**

Table 11. Ablation studies on R_{temp} for the PPMI experiment. All results are averaged across 5-folds.

L_{ce}	R_{temp}	Accuracy (%)	Precision (%)	Recall (%)
✓		82.6 ± 2.1	60.6 ± 7.9	86.4 ± 3.6
✓	✓	**83.6 ± 3.8 (+1.0)**	**62.4 ± 4.4 (+1.8)**	**87.6 ± 3.5 (+1.2)**

C. Converged Scales

Along with Table 4 in the main manuscript, we report the exact value of all converged scales for all experiments in Table 12, 13, 14. Specifically, Table 12 and 13 contain scales derived from the ADNI dataset parcellated on the Destrieux atlas with 160 ROIs, and Table 14 contains results derived from the PPMI dataset using the AAL atlas comprising 116 ROIs. The reported values are the average of the scales from 5 models trained with 5-fold cross-validation. Positive values indicate that a low-pass heat kernel was used for smoothing information around the corresponding ROI, while negative values indicate that a high-pass filter was used to emphasize variations at the corresponding ROI. As shown in the tables, the converged scales exhibit variations across the diagnostic classes and brain regions. These results demonstrate that training AGT on brain connectome datasets allows for both node-wise and group-wise analyses, suggesting a significant potential to be deployed for other connectome analyses of various brain disorders.

Table 12. All converged scales for the ADNI-CT experiment.

idx	ROI	CN	SMC	EMCI	LMCI	AD	idx	ROI	CN	SMC	EMCI	LMCI	AD
1	Left-G_and_S_frontomargin	0.53	0.20	-0.16	0.57	0.87	81	Right-G_and_S_cingul-Mid-Ant	0.65	-0.06	0.56	-0.04	0.26
2	Left-G_and_S_occipital_inf	1.24	0.63	0.39	0.15	0.21	82	Right-G_and_S_cingul-Mid-Post	-0.68	0.06	-0.24	0.16	-0.17
3	Left-G_and_S_paracentral	-0.89	-0.15	-0.49	0.41	-0.33	83	Right-G_cingul-Post-dorsal	-0.17	0.21	0.54	-0.13	-0.70
4	Left-G_and_S_subcentral	0.00	0.65	0.01	-0.04	-0.29	84	Right-G_cingul-Post-ventral	-0.40	0.01	-0.51	0.90	-0.10
5	Left-G_and_S_transv_frontopol	0.21	-0.03	-0.49	-0.15	0.16	85	Right-G_cuneus	-0.36	0.54	-0.85	0.32	-0.10
6	Left-G_and_S_cingul-Ant	-0.19	-0.40	0.34	0.08	0.29	86	Right-G_front_inf-Opercular	-0.59	1.08	-0.19	-0.74	-0.05
7	Left-G_and_S_cingul-Mid-Ant	-0.99	0.38	0.21	0.38	0.05	87	Right-G_front_inf-Orbital	0.11	0.01	0.24	-0.29	0.39
8	Left-G_and_S_cingul-Mid-Post	0.42	0.22	0.08	-0.01	-0.44	88	Right-G_front_inf-Triangul	-0.34	0.01	-0.76	-0.78	0.12
9	Left-G_cingul-Post-dorsal	-0.14	-0.23	-1.08	-0.66	-0.28	89	Right-G_front_middle	-1.06	-0.58	0.19	-0.30	-0.40
10	Left-G_cingul-Post-ventral	0.08	-0.78	-0.43	-0.14	1.44	90	Right-G_front_sup	0.87	-0.42	0.03	-0.12	0.86
11	Left-G_cuneus	-0.43	0.64	0.85	-0.76	0.74	91	Right-G_Ins_lg_and_S_cent_ins	-0.51	0.20	0.00	-0.18	0.33
12	Left-G_front_inf-Opercular	0.80	0.80	-0.65	-0.66	0.03	92	Right-G_insular_short	-0.84	0.08	0.92	0.40	0.01
13	Left-G_front_inf-Orbital	-0.35	0.96	-0.03	0.28	-0.03	93	Right-G_occipital_middle	-0.07	-0.81	-0.31	0.95	0.31
14	Left-G_front_inf-Triangul	-0.12	0.84	-0.51	-0.01	-0.12	94	Right-G_occipital_sup	0.26	-0.94	-0.11	0.45	0.17
15	Left-G_front_middle	0.57	0.01	-0.32	0.11	0.68	95	Right-G_oc-temp_lat-fusifor	-0.83	-0.27	-0.23	-0.32	-0.48
16	Left-G_front_sup	-0.38	0.10	-0.32	-0.35	0.87	96	Right-G_oc-temp_med-Lingual	-0.51	-0.60	0.61	0.60	0.78
17	Left-G_Ins_lg_and_S_cent_ins	1.32	0.22	0.10	-0.25	0.59	97	Right-G_oc-temp_med-Parahip	-0.06	0.20	-0.25	0.36	-0.93
18	Left-G_insular_short	0.30	-0.43	0.89	0.35	0.00	98	Right-G_orbital	-0.43	-0.32	0.26	0.26	-0.37
19	Left-G_occipital_middle	-0.44	-0.09	-0.15	-0.21	0.18	99	Right-G_pariet_inf-Angular	-0.26	0.27	-0.16	-0.47	0.20
20	Left-G_occipital_sup	1.43	0.14	-0.51	0.41	-0.26	100	Right-G_pariet_inf-Supramar	0.70	-0.61	0.35	0.37	-0.50
21	Left-G_oc-temp_lat-fusifor	-0.68	0.98	0.06	1.26	0.20	101	Right-G_parietal_sup	0.00	0.81	0.12	-0.10	-0.05
22	Left-G_oc-temp_med-Lingual	-0.80	-0.03	0.52	1.05	-0.50	102	Right-G_postcentral	-0.19	-1.03	0.25	-0.03	-0.28
23	Left-G_oc-temp_med-Parahip	0.26	-0.65	0.06	-1.17	-0.58	103	Right-G_precentral	0.81	0.44	-0.38	-0.39	0.18
24	Left-G_orbital	-0.56	0.03	-0.05	0.17	-0.46	104	Right-G_precuneus	0.15	-0.99	-0.70	-0.59	-0.10
25	Left-G_pariet_inf-Angular	0.94	-0.06	0.29	0.01	0.38	105	Right-G_rectus	-0.43	0.38	-0.12	-0.02	1.04
26	Left-G_pariet_inf-Supramar	0.39	0.09	-0.77	0.19	0.14	106	Right-G_subcallosal	0.49	-0.50	0.13	0.07	1.22
27	Left-G_parietal_sup	-0.44	-0.61	-1.02	-0.13	0.19	107	Right-G_temp_sup-G_T_transv	0.01	0.71	0.35	0.02	0.26
28	Left-G_postcentral	-0.72	0.10	0.21	-0.35	0.53	108	Right-G_temp_sup-Lateral	0.04	-0.47	-0.65	-0.26	0.66
29	Left-G_precentral	0.19	0.45	0.06	-0.60	0.90	109	Right-G_temp_sup-Plan_polar	-0.15	-0.41	-0.50	0.28	-0.37
30	Left-G_precuneus	0.81	-0.18	0.35	0.13	-0.35	110	Right-G_temp_sup-Plan_tempo	-0.49	0.29	0.63	0.06	0.28
31	Left-G_rectus	0.33	0.26	-0.33	-0.24	0.35	111	Right-G_temporal_inf	-0.59	0.75	-0.21	0.34	0.19
32	Left-G_subcallosal	-0.62	0.05	-0.08	-0.15	0.50	112	Right-G_temporal_middle	0.08	-0.70	-0.29	0.69	-0.36
33	Left-G_temp_sup-G_T_transv	0.53	0.19	-0.10	0.76	-0.41	113	Right-Lat_Fis-ant-Horizont	0.33	-0.17	0.21	-0.11	-0.84
34	Left-G_temp_sup-Lateral	-0.04	0.81	0.20	-0.25	-0.26	114	Right-Lat_Fis-ant-Vertical	-0.41	-0.66	-0.16	-0.13	-0.67
35	Left-G_temp_sup-Plan_polar	-0.43	0.06	0.55	0.55	0.11	115	Right-Lat_Fis-post	-0.20	0.36	-0.57	-0.41	-0.40
36	Left-G_temp_sup-Plan_tempo	-0.17	-0.22	0.30	-0.26	-0.06	116	Right-Pole_occipital	-0.50	0.28	0.18	-0.58	-0.45
37	Left-G_temporal_inf	0.07	-0.48	0.60	0.03	-0.43	117	Right-Pole_temporal	-1.36	-1.01	0.75	-0.28	1.00
38	Left-G_temporal_middle	-0.81	0.12	0.18	0.42	0.38	118	Right-S_calcarine	0.57	0.82	0.29	0.39	-0.42
39	Left-Lat_Fis-ant-Horizont	-0.61	-0.33	-0.39	0.28	-0.64	119	Right-S_central	0.32	0.41	-0.92	-0.49	-1.06
40	Left-Lat_Fis-ant-Vertical	0.84	-0.97	-0.08	-0.40	-0.06	120	Right-S_cingul-Marginalis	0.67	-0.10	-1.35	-0.49	-0.40
41	Left-Lat_Fis-post	-0.35	0.22	-0.19	-0.16	-0.20	121	Right-S_circular_insula_ant	-0.38	-0.04	0.13	-0.54	0.10
42	Left-Pole_occipital	0.24	0.24	0.37	0.15	0.28	122	Right-S_circular_insula_inf	-0.46	0.16	0.54	-0.63	-0.50
43	Left-Pole_temporal	-0.86	0.54	-0.21	0.17	0.03	123	Right-S_circular_insula_sup	-0.49	0.09	0.19	-0.02	-0.34
44	Left-S_calcarine	0.50	-0.77	0.39	-0.33	0.67	124	Right-S_collat_transv_ant	-0.07	-0.78	0.28	-0.02	0.28
45	Left-S_central	-0.16	-0.93	0.22	-1.25	0.25	125	Right-S_collat_transv_post	0.88	0.92	0.11	0.37	-0.25
46	Left-S_cingul-Marginalis	-0.37	0.03	0.51	0.55	0.61	126	Right-S_front_inf	-0.64	0.67	0.51	-0.52	-0.97
47	Left-S_circular_insula_ant	0.03	0.75	-0.09	-0.35	0.49	127	Right-S_front_middle	0.04	-0.55	0.69	-0.24	-0.87
48	Left-S_circular_insula_inf	0.32	0.34	1.00	-0.50	0.01	128	Right-S_front_sup	-0.27	-0.14	1.28	0.82	-0.77
49	Left-S_circular_insula_sup	0.02	-0.58	0.39	-0.27	0.17	129	Right-S_interm_prim-Jensen	-0.94	1.17	-0.15	-0.53	-0.14
50	Left-S_collat_transv_ant	0.30	-0.64	-0.30	0.67	-0.04	130	Right-S_intrapariet_and_P_trans	-1.03	-0.44	-0.41	-0.38	0.79
51	Left-S_collat_transv_post	0.12	0.30	0.08	-0.52	-0.12	131	Right-S_oc_middle_and_Lunatus	0.51	-0.42	0.37	0.21	0.36
52	Left-S_front_inf	-0.26	0.21	0.34	0.62	-0.86	132	Right-S_oc_sup_and_transversal	-0.05	-0.49	0.38	0.53	-0.04
53	Left-S_front_middle	-0.49	-0.09	-0.40	-1.24	0.29	133	Right-S_occipital_ant	-0.26	0.08	-0.60	0.03	-0.14
54	Left-S_front_sup	0.33	0.10	-0.44	-0.15	0.40	134	Right-S_oc-temp_lat	0.28	0.17	0.03	-0.49	-0.51
55	Left-S_interm_prim-Jensen	0.32	0.17	0.28	1.27	0.34	135	Right-S_oc-temp_med_and_Lingual	0.84	0.11	-0.33	-0.92	0.57
56	Left-S_intrapariet_and_P_trans	1.15	0.49	-0.11	0.11	-0.31	136	Right-S_orbital_lateral	-0.29	-0.40	1.17	-0.75	-0.43
57	Left-S_oc_middle_and_Lunatus	0.13	-0.21	-0.11	0.08	-0.35	137	Right-S_orbital_med-olfact	-0.09	-0.63	0.01	-0.48	0.51
58	Left-S_oc_sup_and_transversal	0.39	-0.30	0.41	-0.17	-0.19	138	Right-S_orbital-H_Shaped	-0.24	0.27	0.45	-1.20	-0.17
59	Left-S_occipital_ant	-0.82	-0.47	-0.52	0.62	0.33	139	Right-S_parieto_occipital	-0.12	0.38	-0.98	0.12	0.03
60	Left-S_oc-temp_lat	-0.47	-0.52	-0.05	-0.35	-0.64	140	Right-S_pericallosal	-0.60	-0.72	0.08	-0.12	0.85
61	Left-S_oc-temp_med_and_Lingual	-1.33	0.79	-0.72	0.03	-1.13	141	Right-S_postcentral	-0.19	-0.08	-0.38	-0.44	0.24
62	Left-S_orbital_lateral	-0.68	0.26	-0.48	0.86	0.11	142	Right-S_precentral-inf-part	-0.04	-0.02	0.97	-0.30	-0.04
63	Left-S_orbital_med-olfact	-0.06	0.50	-0.17	0.03	-0.53	143	Right-S_precentral-sup-part	-0.29	-0.57	-0.64	0.48	0.63
64	Left-S_orbital-H_Shaped	0.20	0.07	0.13	0.56	0.13	144	Right-S_suborbital	-0.92	0.93	-0.15	0.59	0.20
65	Left-S_parieto_occipital	0.27	0.89	-0.28	-0.08	0.32	145	Right-S_subparietal	-0.41	-0.22	0.21	0.45	-0.30
66	Left-S_pericallosal	0.04	-0.32	0.22	-0.65	0.40	146	Right-S_temporal_inf	0.24	-0.23	0.27	0.17	-0.25
67	Left-S_postcentral	1.04	-0.25	-0.38	-0.59	-0.04	147	Right-S_temporal_sup	0.33	-0.45	-0.45	-0.25	0.35
68	Left-S_precentral-inf-part	-0.29	0.31	-0.39	-0.84	-0.36	148	Right-S_temporal_transverse	-0.11	-0.96	0.42	0.70	0.46
69	Left-S_precentral-sup-part	-0.13	-0.24	0.06	-0.57	-0.24	149	Left-Amygdala	0.01	-0.74	0.41	-0.05	-1.11
70	Left-S_suborbital	-0.11	0.61	-1.36	-0.31	-0.40	150	Left-Caudate	1.27	0.10	-0.43	0.72	-0.15
71	Left-S_subparietal	0.92	0.38	-0.46	0.16	-0.04	151	Left-Hippocampus	0.28	-0.36	-0.33	-0.49	0.34
72	Left-S_temporal_inf	-0.08	-0.23	-0.66	-0.54	0.70	152	Left-Thalamus-Proper	-0.24	0.24	1.38	-0.85	0.22
73	Left-S_temporal_sup	0.93	-0.67	0.49	-0.33	-0.64	153	Left-Putamen	-0.31	0.33	-0.87	0.09	-0.74
74	Left-S_temporal_transverse	0.15	0.26	0.26	-0.01	0.97	154	Left-Pallidum	0.30	0.13	0.27	-0.55	-0.12
75	Right-G_and_S_frontomargin	-0.18	0.30	0.07	0.03	0.98	155	Right-Amygdala	-0.88	-0.50	0.66	0.62	0.84
76	Right-G_and_S_occipital_inf	0.21	-0.33	0.60	-0.75	0.57	156	Right-Caudate	0.07	-0.62	-0.55	-0.11	0.45
77	Right-G_and_S_paracentral	0.01	1.13	-0.07	-0.14	0.20	157	Right-Hippocampus	0.06	0.57	-0.08	-0.50	-0.18
78	Right-G_and_S_subcentral	0.33	-0.11	0.04	0.72	0.54	158	Right-Thalamus-Proper	-0.73	-0.02	-0.26	0.14	-0.57
79	Right-G_and_S_transv_frontopol	-0.06	0.72	0.35	0.27	-0.20	159	Right-Putamen	-0.12	-0.37	-0.28	0.55	0.49
80	Right-G_and_S_cingul-Ant	-1.12	-0.44	0.26	0.96	-0.67	160	Right-Pallidum	0.07	-0.20	0.52	0.38	0.30

Table 13. All converged scales for the ADNI-FDG experiment.

idx	ROI	CN	SMC	EMCI	LMCI	AD	idx	ROI	CN	SMC	EMCI	LMCI	AD
1	Left-G_and_S_frontomargin	0.27	-0.18	0.37	0.45	0.24	81	Right-G_and_S_cingul-Mid-Ant	-0.19	0.50	-0.28	0.21	0.02
2	Left-G_and_S_occipital_inf	0.59	0.08	0.01	-0.10	-0.48	82	Right-G_and_S_cingul-Mid-Post	-0.50	-0.50	-0.87	-0.09	-0.87
3	Left-G_and_S_paracentral	0.97	-0.65	0.21	-0.15	0.25	83	Right-G_cingul-Post-dorsal	1.03	0.59	-0.54	0.05	-0.91
4	Left-G_and_S_subcentral	0.26	-0.15	0.21	0.56	0.77	84	Right-G_cingul-Post-ventral	-0.56	-0.68	0.10	-0.60	-0.12
5	Left-G_and_S_transv_frontopol	1.25	1.27	-0.17	0.49	-0.62	85	Right-G_cuneus	0.16	0.01	-0.41	0.31	-0.44
6	Left-G_and_S_cingul-Ant	-0.55	-0.24	0.17	-0.64	-0.11	86	Right-G_front_inf-Opercular	0.11	-0.12	0.22	-0.59	0.17
7	Left-G_and_S_cingul-Mid-Ant	-0.71	0.60	-0.25	-0.63	-0.19	87	Right-G_front_inf-Orbital	-0.07	-0.79	0.14	-0.27	0.34
8	Left-G_and_S_cingul-Mid-Post	0.24	0.25	0.10	-0.68	0.68	88	Right-G_front_inf-Triangul	-0.02	-0.27	0.53	-0.93	0.26
9	Left-G_cingul-Post-dorsal	0.36	-0.28	1.10	-0.64	0.34	89	Right-G_front_middle	-0.55	0.57	-0.94	-0.01	-0.59
10	Left-G_cingul-Post-ventral	0.54	0.65	-0.01	0.33	-0.18	90	Right-G_front_sup	0.27	0.87	-0.10	0.45	0.13
11	Left-G_cuneus	-1.17	0.29	0.24	0.02	-0.24	91	Right-G_Ins_lg_and_S_cent_ins	-0.23	-0.11	-0.78	1.07	-0.43
12	Left-G_front_inf-Opercular	0.43	0.54	-0.23	-0.10	-0.18	92	Right-G_insular_short	-0.52	-0.36	-0.10	0.29	-0.69
13	Left-G_front_inf-Orbital	-0.06	-0.41	-0.11	0.19	1.22	93	Right-G_occipital_middle	0.16	-0.96	0.64	-0.36	-0.55
14	Left-G_front_inf-Triangul	0.34	0.75	-0.95	-0.64	0.02	94	Right-G_occipital_sup	-0.18	-1.02	1.03	1.19	0.08
15	Left-G_front_middle	0.35	-0.76	-0.84	-0.08	0.30	95	Right-G_oc-temp_lat-fusifor	-0.48	0.42	0.03	-0.75	-0.15
16	Left-G_front_sup	-0.98	0.88	-0.21	0.00	0.01	96	Right-G_oc-temp_med-Lingual	0.22	0.03	0.45	1.23	0.57
17	Left-G_Ins_lg_and_S_cent_ins	0.44	0.60	-0.26	-0.33	-0.18	97	Right-G_oc-temp_med-Parahip	1.18	-0.18	-0.49	0.29	-0.21
18	Left-G_insular_short	0.82	0.12	-0.03	-0.80	1.08	98	Right-G_orbital	-0.07	-0.91	-0.47	-0.47	-0.15
19	Left-G_occipital_middle	0.50	-0.35	-0.24	0.72	0.40	99	Right-G_pariet_inf-Angular	0.23	0.09	-0.43	-0.84	0.73
20	Left-G_occipital_sup	-0.16	-0.25	0.12	-0.07	-1.04	100	Right-G_pariet_inf-Supramar	0.61	0.24	0.29	0.37	-0.55
21	Left-G_oc-temp_lat-fusifor	-0.58	0.44	-0.13	0.45	-0.56	101	Right-G_parietal_sup	-0.54	-0.66	-0.54	0.04	-0.12
22	Left-G_oc-temp_med-Lingual	-0.53	-0.23	-0.09	0.06	-1.40	102	Right-G_postcentral	-0.13	-0.44	-0.38	-0.16	-0.61
23	Left-G_oc-temp_med-Parahip	0.23	-0.62	0.54	-0.19	0.22	103	Right-G_precentral	-0.01	0.15	-0.61	-0.63	0.22
24	Left-G_orbital	0.03	-0.26	0.28	0.23	0.13	104	Right-G_precuneus	0.64	-0.01	0.06	0.04	-0.71
25	Left-G_pariet_inf-Angular	-0.07	1.20	0.06	0.46	0.61	105	Right-G_rectus	-0.92	-0.45	0.30	-0.43	1.11
26	Left-G_pariet_inf-Supramar	0.50	1.01	-0.17	-0.69	-0.01	106	Right-G_subcallosal	-0.48	-0.76	0.67	-0.44	-0.71
27	Left-G_parietal_sup	-0.24	0.00	-0.39	0.39	-0.03	107	Right-G_temp_sup-G_T_transv	0.63	0.32	-0.34	0.08	0.43
28	Left-G_postcentral	-0.40	0.20	0.04	-0.51	0.79	108	Right-G_temp_sup-Lateral	-0.06	0.27	0.38	0.38	-0.04
29	Left-G_precentral	-0.90	0.52	-0.04	-0.27	0.91	109	Right-G_temp_sup-Plan_polar	0.09	0.37	0.77	0.54	0.37
30	Left-G_precuneus	-0.17	-0.74	0.46	-0.63	-0.50	110	Right-G_temp_sup-Plan_tempo	-0.29	-0.68	0.79	-0.66	0.61
31	Left-G_rectus	0.31	-0.31	1.44	0.76	0.45	111	Right-G_temporal_inf	-0.15	1.14	-0.24	0.10	0.05
32	Left-G_subcallosal	0.15	0.67	-0.53	0.42	0.79	112	Right-G_temporal_middle	-0.82	0.86	0.46	0.72	-0.76
33	Left-G_temp_sup-G_T_transv	-0.01	-0.73	0.39	0.07	-0.37	113	Right-Lat_Fis-ant-Horizont	0.12	0.58	0.42	0.94	-0.24
34	Left-G_temp_sup-Lateral	-0.46	-0.83	1.03	-0.29	-0.98	114	Right-Lat_Fis-ant-Vertical	-0.75	0.05	0.17	0.11	-0.14
35	Left-G_temp_sup-Plan_polar	-1.14	1.08	0.80	0.21	0.17	115	Right-Lat_Fis-post	-0.84	0.36	0.61	0.67	0.91
36	Left-G_temp_sup-Plan_tempo	-0.76	0.38	-0.46	0.24	0.31	116	Right-Pole_occipital	-1.29	0.68	0.40	0.23	0.06
37	Left-G_temporal_inf	-0.43	0.24	-0.19	-0.20	-0.40	117	Right-Pole_temporal	-0.09	0.25	0.10	-0.09	-0.06
38	Left-G_temporal_middle	-0.56	0.37	-0.52	0.26	-0.68	118	Right-S_calcarine	-0.65	0.27	-0.05	0.48	0.24
39	Left-Lat_Fis-ant-Horizont	-0.48	-1.07	-0.71	-0.15	-0.95	119	Right-S_central	-0.44	-0.71	0.34	-0.88	0.70
40	Left-Lat_Fis-ant-Vertical	0.78	0.38	0.75	-0.43	0.26	120	Right-S_cingul-Marginalis	0.40	-0.14	-1.07	-0.13	-0.59
41	Left-Lat_Fis-post	0.23	-0.48	0.03	-0.02	1.09	121	Right-S_circular_insula_ant	0.39	-0.83	-0.28	-0.48	0.03
42	Left-Pole_occipital	0.59	0.66	-0.02	0.71	0.12	122	Right-S_circular_insula_inf	0.11	0.26	-0.23	-0.25	-0.28
43	Left-Pole_temporal	-0.69	-0.25	0.34	-0.67	0.01	123	Right-S_circular_insula_sup	-0.19	0.02	-0.27	-0.03	-0.16
44	Left-S_calcarine	-0.47	0.13	0.16	-0.67	-0.34	124	Right-S_collat_transv_ant	0.46	0.25	1.35	-0.75	0.80
45	Left-S_central	-0.48	-0.09	-0.04	0.05	0.36	125	Right-S_collat_transv_post	0.34	0.22	-1.17	0.42	-0.07
46	Left-S_cingul-Marginalis	-0.82	-0.19	0.27	0.96	0.09	126	Right-S_front_inf	-0.77	0.40	0.00	0.72	-0.76
47	Left-S_circular_insula_ant	1.05	-0.98	-0.41	-0.74	0.57	127	Right-S_front_middle	0.04	0.16	-0.34	0.31	1.19
48	Left-S_circular_insula_inf	1.10	0.23	0.18	-0.93	0.29	128	Right-S_front_sup	-0.36	-0.08	0.44	-0.25	0.32
49	Left-S_circular_insula_sup	-0.28	-0.42	-0.79	-0.35	0.00	129	Right-S_interm_prim-Jensen	0.17	0.34	-0.32	0.62	0.06
50	Left-S_collat_transv_ant	-0.77	0.13	0.34	0.11	-0.10	130	Right-S_intrapariet_and_P_trans	0.04	-0.85	0.19	-0.69	0.50
51	Left-S_collat_transv_post	0.05	-0.40	0.29	0.08	-0.74	131	Right-S_oc_middle_and_Lunatus	0.90	-0.41	0.37	-1.37	0.09
52	Left-S_front_inf	-0.77	-0.34	0.12	-0.63	0.42	132	Right-S_oc_sup_and_transversal	1.13	0.63	-0.12	0.47	-0.82
53	Left-S_front_middle	0.05	0.45	-0.38	0.23	0.61	133	Right-S_occipital_ant	-0.13	0.11	0.22	0.97	-0.72
54	Left-S_front_sup	0.06	-0.57	0.41	-0.30	0.04	134	Right-S_oc-temp_lat	-0.24	-0.50	0.31	-0.84	-0.09
55	Left-S_interm_prim-Jensen	0.48	0.03	0.19	-0.46	-0.55	135	Right-S_oc-temp_med_and_Lingual	-0.40	0.24	0.54	0.91	0.58
56	Left-S_intrapariet_and_P_trans	0.20	0.41	0.30	-0.86	-0.69	136	Right-S_orbital_lateral	0.81	-0.18	1.04	-0.51	-0.21
57	Left-S_oc_middle_and_Lunatus	0.23	0.69	0.53	-0.61	0.09	137	Right-S_orbital_med-olfact	0.23	-0.30	-0.40	-0.26	0.57
58	Left-S_oc_sup_and_transversal	-0.93	-0.26	0.57	-0.19	0.36	138	Right-S_orbital-H_Shaped	0.65	0.24	0.15	-0.83	-0.73
59	Left-S_occipital_ant	-0.23	0.04	-0.14	0.17	0.21	139	Right-S_parieto_occipital	0.22	-0.26	-0.66	1.06	-0.22
60	Left-S_oc-temp_lat	0.62	-0.20	0.33	0.16	-0.74	140	Right-S_pericallosal	0.25	-0.58	-0.48	0.09	0.56
61	Left-S_oc-temp_med_and_Lingual	-1.18	0.64	-0.01	-0.11	0.27	141	Right-S_postcentral	0.19	0.90	-0.09	0.22	0.05
62	Left-S_orbital_lateral	-0.23	0.01	-0.35	0.33	0.60	142	Right-S_precentral-inf-part	0.24	-0.04	0.53	0.19	0.45
63	Left-S_orbital_med-olfact	0.30	0.34	-0.13	0.16	0.20	143	Right-S_precentral-sup-part	-0.73	-1.18	0.33	0.24	-1.12
64	Left-S_orbital-H_Shaped	-0.59	0.17	-0.10	0.28	-0.69	144	Right-S_suborbital	-1.26	0.36	-0.04	0.62	0.24
65	Left-S_parieto_occipital	0.52	-0.77	-0.08	-0.36	0.30	145	Right-S_subparietal	0.56	-0.63	0.57	-0.51	-0.48
66	Left-S_pericallosal	0.59	-0.31	0.51	0.37	-0.56	146	Right-S_temporal_inf	0.45	0.18	-0.14	0.09	0.40
67	Left-S_postcentral	-0.48	0.27	-0.83	0.49	1.17	147	Right-S_temporal_sup	0.10	0.16	-0.21	0.29	0.53
68	Left-S_precentral-inf-part	-0.04	-0.51	-0.37	-0.60	-0.08	148	Right-S_temporal_transverse	-0.23	-0.35	-0.39	-0.30	0.75
69	Left-S_precentral-sup-part	-0.20	-0.44	-0.05	0.69	-0.54	149	Left-Amygdala	-0.16	-0.02	0.86	-0.19	-0.60
70	Left-S_suborbital	0.71	-0.10	-0.21	-0.20	-0.30	150	Left-Caudate	-0.11	-0.29	-0.63	0.36	-0.32
71	Left-S_subparietal	-0.51	0.00	-0.32	0.64	-0.64	151	Left-Hippocampus	0.02	-0.16	-0.53	0.08	-0.23
72	Left-S_temporal_inf	0.61	-0.30	0.32	0.60	-0.11	152	Left-Thalamus-Proper	-0.24	0.26	-0.03	-0.58	0.78
73	Left-S_temporal_sup	0.07	-1.03	-0.10	0.34	0.43	153	Left-Putamen	0.34	1.27	-0.81	-0.38	-0.12
74	Left-S_temporal_transverse	-0.39	0.12	-0.02	0.61	0.22	154	Left-Pallidum	0.06	0.62	-0.06	0.15	-0.14
75	Right-G_and_S_frontomargin	0.08	0.54	0.12	0.32	0.20	155	Right-Amygdala	-1.02	-0.56	0.01	0.10	0.52
76	Right-G_and_S_occipital_inf	0.12	0.28	0.36	-0.78	0.26	156	Right-Caudate	-0.92	0.33	-0.84	0.35	-0.10
77	Right-G_and_S_paracentral	0.21	0.47	0.87	-0.17	-0.78	157	Right-Hippocampus	-0.13	-0.51	-0.23	1.05	-0.55
78	Right-G_and_S_subcentral	0.14	1.16	0.36	-0.12	-0.74	158	Right-Thalamus-Proper	0.07	1.57	-0.19	0.90	-1.11
79	Right-G_and_S_transv_frontopol	-0.34	-0.43	0.91	0.58	-0.33	159	Right-Putamen	-0.65	-0.52	-0.18	0.63	0.46
80	Right-G_and_S_cingul-Ant	0.24	0.28	-0.35	0.18	0.42	160	Right-Pallidum	-0.44	0.00	0.95	-0.52	0.95

Table 14. All converged scales for the PPMI experiment.

idx	ROI	CN	Prodromal	PD	idx	ROI	CN	Prodromal	PD
1	Precentral_L	-0.26	-0.24	-0.39	59	Parietal_Sup_L	-0.60	0.38	-0.05
2	Precentral_R	-0.64	-0.27	-0.20	60	Parietal_Sup_R	-0.20	0.21	0.18
3	Frontal_Sup_L	0.65	-0.20	-0.26	61	Parietal_Inf_L	-0.19	-0.13	-0.63
4	Frontal_Sup_R	-0.15	0.09	-0.26	62	Parietal_Inf_R	0.02	0.26	-0.50
5	Frontal_Sup_Orb_L	0.67	0.22	0.51	63	SupraMarginal_L	0.08	-0.54	0.43
6	Frontal_Sup_Orb_R	0.77	-0.16	-0.27	64	SupraMarginal_R	1.01	0.65	-0.88
7	Frontal_Mid_L	0.29	-0.65	0.36	65	Angular_L	0.11	-0.10	-0.48
8	Frontal_Mid_R	1.15	-1.02	0.40	66	Angular_R	-0.06	-0.91	-0.11
9	Frontal_Mid_Orb_L	-0.38	0.93	-0.29	67	Precuneus_L	0.54	0.74	0.09
10	Frontal_Mid_Orb_R	-0.36	-0.47	0.39	68	Precuneus_R	0.08	-0.41	-1.05
11	Frontal_Inf_Oper_L	1.14	-0.63	-1.17	69	Paracentral_Lobule_L	0.47	0.78	-0.47
12	Frontal_Inf_Oper_R	0.07	0.27	0.12	70	Paracentral_Lobule_R	-0.90	-0.43	0.56
13	Frontal_Inf_Tri_L	0.28	-0.18	-0.23	71	Caudate_L	-0.15	-0.21	1.53
14	Frontal_Inf_Tri_R	-0.93	-0.48	0.59	72	Caudate_R	0.10	-0.68	-0.03
15	Frontal_Inf_Orb_L	0.53	0.25	0.68	73	Putamen_L	-0.23	0.17	0.04
16	Frontal_Inf_Orb_R	0.21	0.80	0.70	74	Putamen_R	0.00	-0.73	0.60
17	Rolandic_Oper_L	-0.30	0.62	-0.27	75	Pallidum_L	-0.51	0.19	0.20
18	Rolandic_Oper_R	0.22	1.15	-0.46	76	Pallidum_R	0.59	0.28	0.02
19	Supp_Motor_Area_L	0.83	0.12	0.02	77	Thalamus_L	-0.21	0.43	0.47
20	Supp_Motor_Area_R	0.11	0.56	0.30	78	Thalamus_R	-0.98	-0.71	-0.39
21	Olfactory_L	-0.06	-0.73	0.05	79	Heschl_L	0.41	0.47	0.25
22	Olfactory_R	-0.06	-0.40	0.14	80	Heschl_R	0.02	0.68	1.49
23	Frontal_Sup_Medial_L	-0.25	-0.81	-0.84	81	Temporal_Sup_L	-0.77	0.00	-1.41
24	Frontal_Sup_Medial_R	-0.14	-0.19	0.52	82	Temporal_Sup_R	-0.27	-0.27	-1.08
25	Frontal_Med_Orb_L	-0.31	0.55	-0.01	83	Temporal_Pole_Sup_L	0.85	0.16	0.11
26	Frontal_Med_Orb_R	-0.39	1.15	-1.11	84	Temporal_Pole_Sup_R	0.01	0.80	1.16
27	Rectus_L	-0.38	0.37	0.56	85	Temporal_Mid_L	1.10	0.33	0.71
28	Rectus_R	-0.36	-0.04	-0.54	86	Temporal_Mid_R	0.85	-0.08	0.21
29	Insula_L	0.23	-0.39	0.15	87	Temporal_Pole_Mid_L	-0.34	0.09	0.10
30	Insula_R	-0.69	-0.25	0.54	88	Temporal_Pole_Mid_R	-0.44	-0.52	0.00
31	Cingulum_Ant_L	0.48	0.14	-0.19	89	Temporal_Inf_L	0.36	-0.82	-0.46
32	Cingulum_Ant_R	-0.63	-0.48	-0.65	90	Temporal_Inf_R	-0.49	0.49	0.00
33	Cingulum_Mid_L	-0.52	-0.07	-0.04	91	Cerebellum_Crus1_L	-0.33	-0.34	0.55
34	Cingulum_Mid_R	0.11	-0.25	0.04	92	Cerebellum_Crus1_R	-0.44	-0.04	-0.59
35	Cingulum_Post_L	-0.28	-0.73	0.05	93	Cerebellum_Crus2_L	-0.26	0.16	-0.19
36	Cingulum_Post_R	-0.79	0.30	0.00	94	Cerebellum_Crus2_R	0.02	-0.80	0.42
37	Hippocampus_L	-0.14	-0.07	-0.15	95	Cerebellum_3_L	0.24	0.89	0.66
38	Hippocampus_R	-0.11	0.66	-0.92	96	Cerebellum_3_R	-0.04	0.32	-0.08
39	ParaHippocampal_L	1.25	0.72	0.07	97	Cerebellum_4_5_L	0.69	0.25	-0.63
40	ParaHippocampal_R	-0.36	-0.50	0.38	98	Cerebellum_4_5_R	-0.44	-0.62	0.74
41	Amygdala_L	1.11	-0.25	-0.09	99	Cerebellum_6_L	0.40	-0.68	-0.63
42	Amygdala_R	0.44	0.32	0.45	100	Cerebellum_6_R	0.29	0.30	-0.54
43	Calcarine_L	1.11	-0.88	-0.56	101	Cerebellum_7b_L	-0.25	0.98	0.43
44	Calcarine_R	-0.21	-0.12	-1.17	102	Cerebellum_7b_R	0.19	-0.33	-0.10
45	Cuneus_L	-0.10	-0.59	0.19	103	Cerebellum_8_L	-0.21	-0.97	-1.03
46	Cuneus_R	0.56	0.33	-0.47	104	Cerebellum_8_R	0.17	0.28	0.29
47	Lingual_L	-0.44	0.22	0.80	105	Cerebellum_9_L	-0.53	-0.33	0.04
48	Lingual_R	0.01	0.60	0.11	106	Cerebellum_9_R	-0.09	0.37	0.55
49	Occipital_Sup_L	0.48	-0.06	0.48	107	Cerebellum_10_L	-0.27	0.51	-0.17
50	Occipital_Sup_R	0.78	-0.59	-0.60	108	Cerebellum_10_R	0.53	-0.41	0.24
51	Occipital_Mid_L	-0.08	-0.98	0.61	109	Vermis_1_2	-0.99	0.45	-0.15
52	Occipital_Mid_R	0.43	-0.36	-0.23	110	Vermis_3	-0.95	0.78	-0.12
53	Occipital_Inf_L	0.18	0.22	0.41	111	Vermis_4_5	0.06	-0.81	-0.38
54	Occipital_Inf_R	0.27	0.56	0.04	112	Vermis_6	0.40	0.45	0.54
55	Fusiform_L	-0.60	0.51	0.24	113	Vermis_7	0.42	-0.73	0.43
56	Fusiform_R	0.14	-0.15	0.10	114	Vermis_8	0.62	0.07	-0.63
57	Postcentral_L	-0.22	0.15	0.12	115	Vermis_9	-0.29	-0.46	0.44
58	Postcentral_R	0.30	0.22	0.48	116	Vermis_10	-0.31	-0.14	0.09