

Similarity Learning with Higher-Order Graph Convolutions for Brain Network Analysis

Guixiang Ma¹, Nesreen K. Ahmed², Ted Willke², Dipanjan
Sengupta², Michael W. Cole³, Nicholas B. Turk-Browne⁴, Philip S.
Yu¹

Presenter: Arshdeep Sekhon

<https://qdata.github.io/deep2Read>

Motivation

- learn a similarity between brain networks
- Uses GCN in a Siamese framework
- incorporating higher-order proximity via random walks in graph convolutional networks
- incorporates community structure in brain nets
- 4 real world brain datasets with respect to brain health status and cognitive abilities

Method: Notations

- a multi-subject fMRI data set $\mathcal{G} = \{G_1, G_2, \dots, G_N\}$
- where $G_i = (V_i, E_i, \mathbf{A}_i)$ is the fMRI brain network of subject i ,
- V_i is the set of vertices in G_i ,
- $E_i \subset V_i \times V_i$ is the set of edges in G_i ,
- $\mathbf{A}_i \in \mathbb{R}^{m \times m}$ is the affinity matrix of G_i

Method: From fMRI data to brain connectivity networks

- A whole-brain fMRI image consists of a sequence of 3D brain image scans, where each volume consists of hundreds of thousands of voxels.
- To convert the original fMRI images to region-by-region brain networks, extract a sequence of responses from each of the regions of interest (ROI), where each ROI represents a brain region.
- compute the region-to-region brain activity correlations.
- only keep the positive correlations as the links among the brain regions.
- The final constructed network is a graph where the nodes/vertices represent brain regions and the edges are the region-to-region correlations.

Higher-order Graph Convolutional Networks.

- GCNs use spectral filterings : localized to within K neighbor nodes

$$y = g_\theta * \mathbf{x} = g_\theta(\mathbf{L})\mathbf{x} = g_\theta(\mathbf{U}\Lambda\mathbf{U}^T)\mathbf{x} = \mathbf{U}g_\theta(\Lambda)\mathbf{U}^T\mathbf{x} \quad (1)$$

$$g_\theta(\Lambda) = \sum_{k=0}^{K-1} \theta_k T_k(\hat{\Lambda}) \quad (2)$$

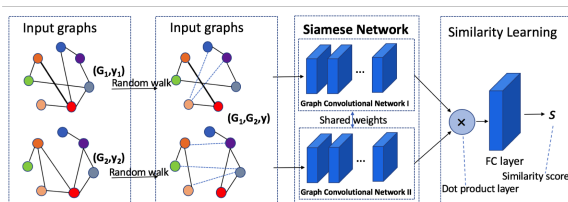
filtering operation can be written as $y = g_\theta(\mathbf{L})\mathbf{x} = \sum_{k=0}^{K-1} \theta_k T_k(\hat{\mathbf{L}})\mathbf{x}$, where $T_k(\hat{\mathbf{L}}) \in \mathbb{R}^{n \times n}$ is the Chebyshev polynomial of order k . The j^{th} output feature map of sample s is then given by

$$y_{s,j} = \sum_{i=1}^{F_{in}} g_{\theta_{i,j}}(L) x_{s,i} \in \mathbb{R}^m \quad (3)$$

build the GCN by stacking multiple convolutional layers with a non-linearity activation (ReLU) following each layer.

The framework

[t]



$$L^{hinge} = \frac{1}{N_p} \sum_{i=1}^N \sum_{j=i+1}^N \max(0, 1 - Y_{ij} s_{ij}), \quad (4)$$

where N is the total number of subjects in the training set, and $N_p = N(N - 1)/2$ is the total number of pairs from the training set.

Algorithm 1 Higher-order Siamese GCN

- Input:** $\mathcal{G} = G_1, G_2, \dots, G_n$ (training graph samples); \mathbf{y} (class labels); random walk parameters: γ (number of walks), l (walk length), w (window size)
- 1: Obtain the mean k -nn graph $\bar{G}(V, E, \bar{\mathbf{A}})$;
 - 2: Initialize a frequency matrix $\mathbf{F} \in \mathbb{R}^{m \times m}$ with 0s;
 - 3: **for** $i = 0$ to γ **do**
 - 4: $V' = \text{Shuffle}(V)$;
 - 5: **for each** $v_i \in V'$ **do**
 - 6: $W_{v_i} = \text{RandomWalk}(\bar{G}, v_i, l)$;
 - 7: Update \mathbf{F} ;
 - 8: **end for**
 - 9: **end for**
 - 10: Obtain a k -nn graph G' based on \mathbf{F} ;
 - 11: Merge the edges of G' into \bar{G} ;
 - 12: Obtain the updated adjacency matrix \mathbf{A} ;
 - 13: Prepare pairs of training samples from \mathcal{G} ;
 - 14: Initialize the parameters Θ of GCNs in Siamese network;
 - 15: **while** not converge **do**
 - 16: Perform spectral filterings according to Equation (3);
 - 17: Compute the similarity estimate s_{ij} for the input pair (G_i, G_j) ;
 - 18: Compute the loss L^{hinge} according to Equation (4) ;
 - 19: Apply stochastic gradient descent with ADAM optimizer to update Θ ;
 - 20: **end while**
-

Experiments: Data

- Autism Brain imaging Data Exchange (ABIDE)
- Human Connectome Project (HCP): Does not have class labels of cognitive traits, use three key cognitive features from the participants' behavioral data to apply K-means clustering with the three features to cluster the subjects into 2 groups
- Bipolar: fMRI data of 52 bipolar I subjects who are in euthymia and 45 healthy controls with matched demographic characteristics
- Human Immunodeficiency Virus Infection (HIV): resting-state fMRI data of 77 subjects, 56 of which are early HIV patients and the other 21 subjects are seronegative controls.

Results

Table: AUC Scores of Pair Classification (mean \pm std).

Methods	<i>ABIDE</i>	<i>HCP</i>	<i>HIV</i>	<i>Bipolar</i>
PCA	0.51 \pm 0.01	0.52 \pm 0.01	0.54 \pm 0.07	0.52 \pm 0.01
SE	0.55 \pm 0.02	0.54 \pm 0.01	0.57 \pm 0.02	0.55 \pm 0.01
S-GCN	0.78 \pm 0.29	0.81 \pm 0.36	0.61 \pm 0.25	0.74 \pm 0.19
HS-GCN	0.96 \pm 0.02	0.98 \pm 0.03	0.77 \pm 0.20	0.94 \pm 0.07

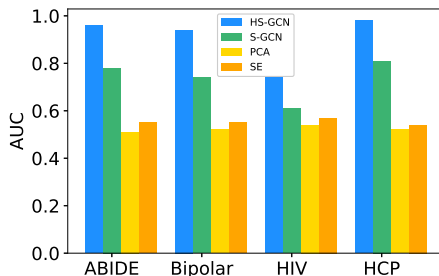
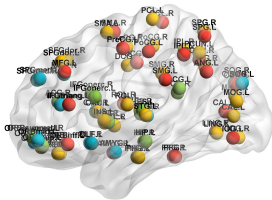
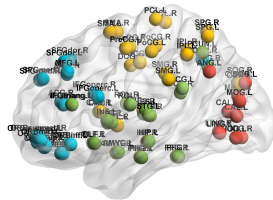


Figure: AUC Scores of Pair Classification.

Results: Visualization

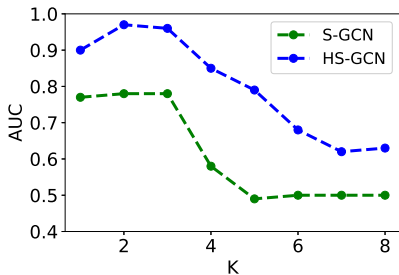


(a) Bipolar

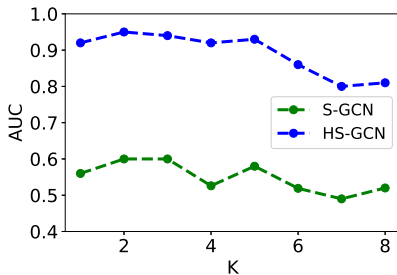


(b) Healthy

Figure: Visualization of the community structure captured by HS-GCN in healthy and bipolar disease networks. Notably this figure highlights the reduced functional connectivity as shown by decreased clustering in the bipolar network.

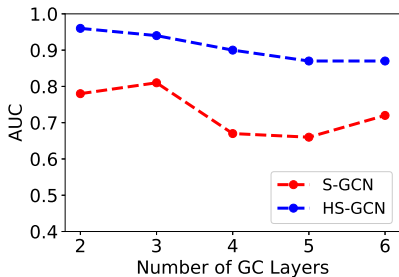


(a) *ABIDE*

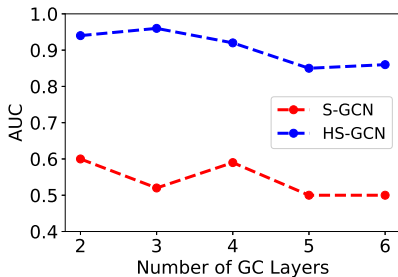


(b) *Bipolar*

Figure: Pair classification AUC of S-GCN and HS-GCN with different values for K



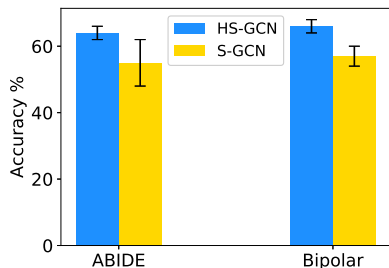
(a) *ABIDE*



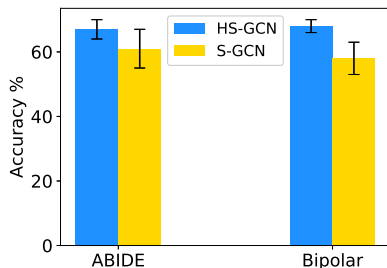
(b) *Bipolar*

Figure: Pair classification AUC of S-GCN and HS-GCN with different numbers of GC layers

Subject Classification Loss



(a) *Hinge Loss*

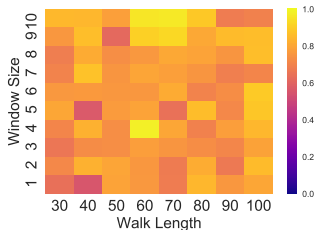


(b) *Constrained Variance Loss*

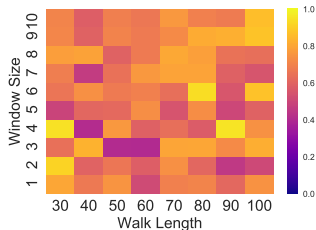
Figure: Subject classification accuracy on ABIDE and Bipolar with two different loss functions: We apply the weighted k-nearest neighbour (kNN)

$$L^{\text{convar}} = \max(0, \delta^{2+} - a) + \max(0, \delta^{2-} - a) + \max(0, m - (\sigma^+ - \sigma^-)), \quad (5)$$

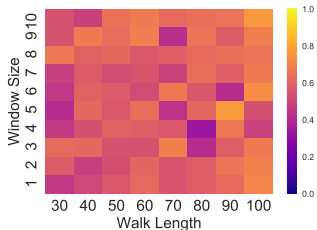
Parameter Analysis



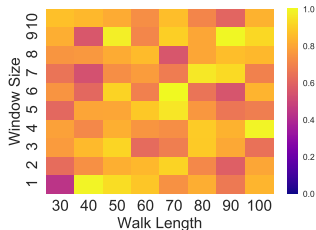
(a) *ABIDE*



(b) *Bipolar*



(c) *HIV*



(d) *HCP*

Figure: AUC scores for Pair classification with different values for Random

Conclusion

- for learning similarity among fMRI brain networks using higher-order GCNs as the twin
- working well for relatively small datasets
- Graphs structurally different per sample but uses the same modified A for all samples
- limited to *fMRI* or known graph structures