Graph Analysis of Functional Connectomes of Subjects with ADHD and ASD

Benjamin Yeh
Stanford University
bentyeh@stanford.edu

Victoria Yuan
Stanford University
vyuan@stanford.edu

Abstract

The prevalence of autism spectrum disorder (ASD) and attention deficit hyperactivity disorder (ADHD) diagnoses has increased in the past decade. Commonly diagnosed in children during preschool, the disorders can cause behavioral, academic, and emotional problems for these patients. In this paper we leverage fMRI scans of patients with typically-developing (TD), ASD, and ADHD as neural connectomes. We combine structural and functional connectivity information to form spatiofunctional networks across fMRI scan time. Our paper aims to 1) investigate the neurological features underlying ASD and ADHD 2) construct predictive models that can guide clinicians in diagnosing ASD and ADHD. We compared the performance of linear supervised classification models and graph convolutional networks (GCNs). We also performed temporal network analysis to identify differences in neural circuitry across patient types.

1. Introduction

While ASD affects 1 in 68 children in the United States, ADHD affects 1 in 10. [8]. Developmental disorders such as ASD and ADHD can be represented as aberrant structures of functional brain networks. By regarding the brain as a network, nodes represent specialized regions while edges contain temporal correlations. Brain networks can thus be visualized, and their topological properties can be explored. Graph metrics such as clustering coefficients and community structure analysis of brain networks can capture neural organization and behavior. This presents an avenue in which to study the functional connectomes of ASD and ADHD, and compare them to the networks of typically developing children. A connectome is the network representation of the brain. For instance, a resting fMRI can be represented as a graph where the nodes are regions of the brain and the weighted edges are time-correlations. Investigating the neurobiology of ASD and ADHD can potentially inform the structure-function relationship of the brain and guide clinicians in diagnosing these developmental disorders.

2. Related Work

Ray et al. [10] analyze the structural and functional connectomes of typically developing, ASD, and ADHD patients using weighted, undirected networks. Graphs were generated from the resting fMRI data of 20 controls, 20 children with ADHD, and 20 high-functioning ASD children to compare the rich club organization of each patient type. A rich-club organization is a graph structure where highly connected nodes show a tendency to connect with other highly connected nodes. Since current literature correlates rich-club organization with a normal brain development, the paper focuses on analyzing the rich-club connectedness coefficient. The connectedness coefficient is the product of the network's rich-club coefficient and its connectivity index. The paper discovered that the ASD group demonstrated over-connectedness in its rich-club organizations compared to ADHD and control networks. Patients with ADHD, however, show the fewest rich-club regions, with under-connectivity in the structures.

However, the study is severely limited in its network analysis methods as it focuses solely on rich-club organization. Narrowing the network analysis to a single feature does not capture the range of motifs that may be present across controls, ASD, and ADHD patients. Additionally, the brain is a complex network yet Ray et al. neglects to investigate other network features, such as community structure and motifs.

While current literature characterizes the networks of patients with TD, ASD, and ADHD, these fMRI-generated networks can also be leveraged to create a predictive model that diagnoses ADHD. The ADHD-200 Consortium [12] is an international competition to develop predictive models for diagnosing ADHD and to identify brain features as ADHD biomarkers. While the previous literature compares the network characteristics of control patients to that of ASD patients, this paper details a range of predictive models to differentiate ADHD patients from control patients. The winning team for diagnosis prediction excelled in specificity, correctly classifying 94% of control patients. How-

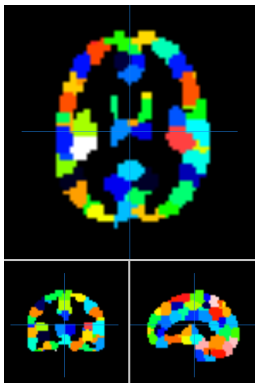

Figure 1. Sample scan labelled by CC200 ROI

ever, they were only able to correctly identify 21% of the cases with ADHD. The team's model leveraged CUR decomposition on the fMRI scans and used a gradient boosting method (GBM) for prediction. The CUR decomposition selected voxels in the brain that are typically not correlated with ADHD. Thus the model may have trained on biologically less relevant data. While the paper utilizes resting fMRI data to diagnose ADHD, it does so with poor accuracy and with little biological bases.

3. Data

We utilize publicly available fMRI datasets from the International Neuroimaging Datasharing Initiative (INDI) of control individuals, individuals with ASD, and individuals with ADHD. All fMRI data has been preprocessed to account for slice timing corrections, motion realignment, and intensity normalization using the Configurable Pipeline for the Analysis of Connectomes (C-PAC). [3] The images have also undergone functional parcellation (akin to segmentation) into 200 anatomical regions (called the CC200 ROIs, or regions of interest) using the algorithm described by Craddock, et al. [4] From INDI, we will leverage the ABIDE I Preprocessed and ADHD-200 Preprocessed fMRI data sets. ABIDE I Preprocessed contains fMRI data from 884 individuals, 408 with ASD and 476 controls; the median age of all patients is 14.7 years. While the full ADHD-200 dataset consists of 947 individuals, we were only able to access preprocessed data of 113 patients, 68 of which are controls and 45 of which were diagnosed with ADHD; their median age was 11.4 years. Each dataset is also accompanied by phenotypic data on age, sex, and symptom measures.

Undirected weighted graphs are generated from each patient using time-series correlation values as weights. For each brain, this yields a symmetric 200x200 correlation matrix. Using the correlation matrices as adjacency matrices, we constructed fully connected, undirected, weighted graphs for each subject where the nodes are different anatomical regions and weighted edges are the strength of correlation over time of measured neural activity.

4. Methods

4.1. Preprocessing

Craddock, et al. utilizes spectral clustering with normalized cut to parcellate the brain. The resulting regions are represented as numbers but lose its neurological information. For example, given region 200, it is unknown what gyri it corresponds to in the cerebral cortex. Moreover, the regions in the Craddock parcellation cross anatomical boundaries such that a CC200 ROI may overlap multiple adjacent anatomical regions of the brain.

To provide structural data, the CC200 ROIs were anatomically labelled using the Harvard-Oxford, Automated Anatomical Labelling, and the Eickhoff-Zilles brain atlases [5]. A connectivity map of brain regions was manually annotated using existing brain atlases, cortical connectivity atlases, and the thalamic cortex connectivty atlas. Using this annotation, we generated a *structural* (spatial) network with the CC200 ROIs. By combining this network with the *functional* networks from the fMRI, we created spatiofunctional networks for each patient.

4.2. Weighted node degree distribution

We performed non-parametric Mann-Whitney U tests for whether the weighted degree of any given node $i \in (1, ..., 200)$ (corresponding to the 200 CC200 ROIs) was significantly different between subjects diagnosed with ADHD versus controls, or between subjects diagnosed with ASD versus controls. Calculated p-values were adjusted using the Bonferroni correction (multiplying each p-value by 200) to account for multiple hypothesis testing. Note that many commonly used and less-stringent Type I error control procedures, such as the Benjamini-Hochberg FDR procedure, are not appropriate because of the required assumption that each of the 200 tests be independent.

4.3. Classification

Classification models were trained separately for ADHD and ASD datasets. Our supervised learning objective was to accurately determine whether an input *functional* graph corresponds to a subject who is typically-developing or diagnosed with ASD or ADHD. As baselines, we used logistic regression and support vector machines (SVMs) with a radial basis function kernel to classify nodes based on 15-dimensional embeddings generated using node2vec [6], as well as graphs based on their averaged node embeddings. Data were split 67% train / 33% test; 5-fold cross validation was performed on the train set to determine optimal regularization weights.

We also trained graph convolutional networks (GCNs), with each node represented by its timeseries, rather than correlation with other nodes. Timeseries for all nodes within a given dataset were truncated to the minimum number of time points collected for any patient within that dataset: this gave an input dimension of 78 for nodes in graphs from the ADHD-200 dataset and an input dimension of 41 for nodes in graphs from the ABIDE dataset. Edges were defined based on physical connectivity of adjacent CC200 regions in the brain. The GCN model is described in equation 1:

$$\begin{aligned} h_v^0 &= x_v \\ h_v^k &= \sigma\left(W_k \sum_{u \in N(v)} \frac{h_u^{k-1}}{|N(v)|} + B_k h_v^{k-1}\right) \\ z_v &= g(f(h_v^K)) \\ f(h_v^K) &= A_2(A_1 h_v^K + C_1) + C_2 \\ g(z) &= \begin{bmatrix} \log\left(\frac{\exp(z_1)}{\exp(z_1)+\exp(z_2)}\right) \\ \log\left(\frac{\exp(z_2)}{\exp(z_1)+\exp(z_2)}\right) \end{bmatrix} \\ \hat{y} &= \text{argmax}_i z_i \end{aligned} \tag{1}$$

where σ is the ReLU function, $v \in (1, ..., 200)$ denotes the node, K is the number of hidden layers, $k \in (1, ..., K)$ denotes the layer of a hidden representation vector h_v^k, f represents 2 fully-connected layers, and g is the log-softmax function over 2 classes.

Hyperparameters were tuned across the following ranges: 1-3 (K) hidden layers of 5-15 (d) dimensions, with a learning rate between 0.1 and 0.001. Training was performed to minimize negative log likelihood loss; a dropout ratio of 0.5 was enforced during training to limit overfitting. Models were trained on 45% of the data; optimal hyperparameters were selected based on performance on a validation set consisting of 22% of the original data. The same testing set as used for the baselines was used for final GCN model evaluation. All model evaluation was performed using the binary F1 score.

4.4. Temporal Network Analysis

Spatiofunctional temporal networks were generated for each patient by combining the structural and functional networks. A undirected, unweighted graph was generated such that nodes are the 200 Craddock anatomical regions, edges denote interactions between the regions based on the measured BOLD signal by fMRI, and each edge has a feature vector of times when the edge exists.

Given region R_i at time t_n, if R_i has a positive BOLD signal, then at time t_{n+1}, edge (R_i, R_n) is an activating edge if $R_n \in N(R_i)$, where $N(R_i)$ is the set of all neighbors of R_i from the manually annotated *structural* edges in Preprocessing (Section 4.1). If R_i has a negative BOLD signal, then an edge is labelled as an inhibiting interaction. The type of edge, activating or inhibiting, denotes the *functional* characteristic of the network. Time t_n is then appended to the temporal feature vector of the edge.

The changes in activating and inhibiting interactions across patients with ASD, ADHD, and TD were analyzed over fMRI scan time. Activated and inhibited anatomical regions were compared across patient types.

5. Results

5.1. Degree distribution

We first considered the weighted degree distribution of graphs within each dataset. We plotted the average degree distribution taken over all graphs (Figure 2). We do not see a significant difference between the degree distribution of subjects diagnosed with ADHD versus controls, or between that of subjects diagnosed with autism versus controls.

Mann-Whitney U tests for the differences between the weighted degree of ROIs from subjects diagnosed with ADHD versus controls, or between subjects diagnosed with ASD versus control, did not reveal any statistically significant ROIs (Table 1).

Dataset	Most significant ROI	p_{adj}
ADHD-200	44 - Right Calcarine Gyrus	0.46
ABIDE	123 - Right Medial Frontal Gyrus	0.17

Table 1. Most significant ROIs, based on Mann-Whitney U test of weighted degree, between subjects diagnosed with ADHD and typically-developing individuals, and between subjects diagnosed with ASD and typically-developing individuals. p_{adj}: Bonferroni-adjusted p-value. The calcarine sulcus contains the primary visual cortex. The right medial frontal gyrus is though to be important in executive decision-making [11].

5.2. Classification

To visually understand the relationship between nodes in a graph and nodes in other graphs, we plotted the first two principal components for all nodes (Figure 3). We see that nodes cluster based on human subjects rather than the corresponding CC200 ROI, suggesting that either functional dependencies in the brain are highly individual, or there may be confounding variables leading to observed batch effects. Without paired controls (e.g., same subject measured at different labs, etc.), we are unable to distinguish between the two possibilities. However, this suggested that using node aggregation (e.g., mean) methods to generate graph representations are reasonable. While there is no clear separation between controls and ADHD/ASD subjects in the PCA plots, non-linear methods proved useful.

Classification results are shown in Table 2. A simple linear logistic regression model achieves mediocre F1-scores,

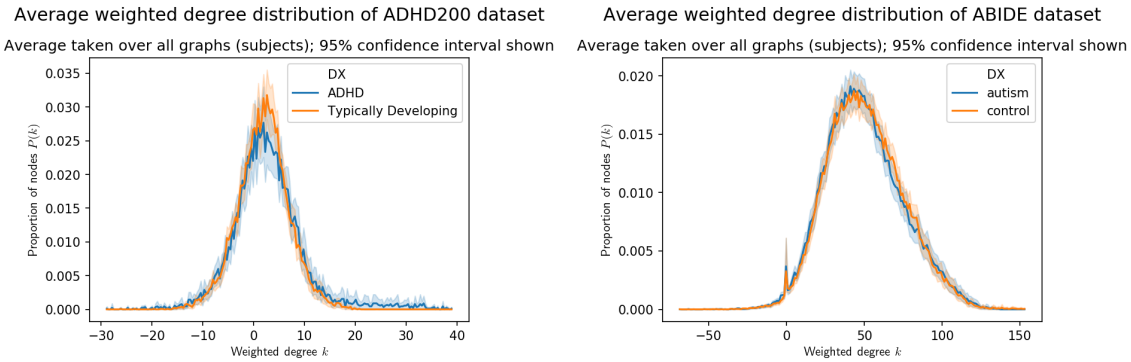

Figure 2. Average weighted degree distribution. The peak at a weighted degree of 0 in the ABIDE dataset is an artifact due to missing values in the data. Abbreviations: DX = diagnosis.

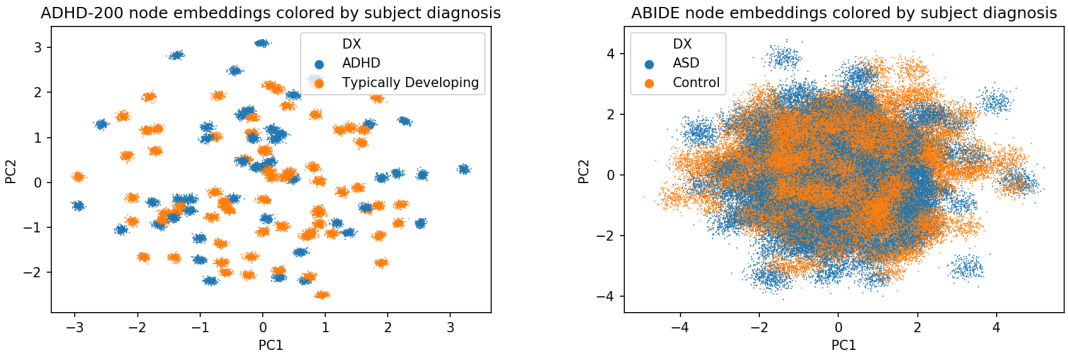

Figure 3. 15-dimensional node embeddings generated using node2vec, then plotted along the primary principal components.

which is not surprising as we were unable to identify any linear plane of separation in the PCA plots. The nonlinear RBF-based SVMs perform significantly better. On the ADHD-200 dataset, the SVM is able to perfectly classify all nodes in both the training and test set. An SVM trained on graph embeddings does not perform as well for graph classification, which is surprising because we observed earlier that graph embeddings (computed as the mean of the node embeddings) are fairly representative summaries of the nodes. A possible explanation is that the 200x larger size of the node embeddings dataset (relative to the graph embeddings dataset) allows for better training; this explanation is supported by the result that using the SVM trained on node embeddings accurate classifies graph embeddings. We do not see the same level of performance on the ABIDE dataset, which we anticipated given the larger variability observed in the node embeddings in Figure 3b.

Notably, the baseline models perform as well as the hyperparameter-optimized GCN models on the test set (Table 3). It is unclear why this is the case: for example, the similar validation and test F1 scores for the GCN model trained on the ADHD-200 dataset suggest that the models are generalizing with similar performance to that achieved on the validation set. However, we see dramatically reduced performance on the test F1 score for GCN model trained on the ABIDE dataset relative to the validation F1 score. Figure 4 provides a possible insight: the highly variable validation performance of the GCN models over 200 epochs suggests that the model predictions are not the most robust to small changes in the model parameters throughout training (Figure 4). This variability may be due to lack of clear decision boundaries and/or the relatively small size of the dataset.

5.3. Temporal Network Analysis

Networks were generated at each timepoint for each subject in the ASD and TD datasets. Degree distributions taken of subjects across the fMRI scan time showed no significant difference between patients with ASD and TD. Moreover, comparison of adjacency matrices across time revealed no significant difference in structural connectivity (Figure 5). The sharp change in mean activations and inhibitions in the respective graph is attributed to the differences in fMRI scan time. The number of timepoints varies across patients from 174 to 336. As the number of patients decreases with time, the variation between ASD and TD increases.

Because of the similar network topology across patient types, we analyzed the functional connectivity of each net-

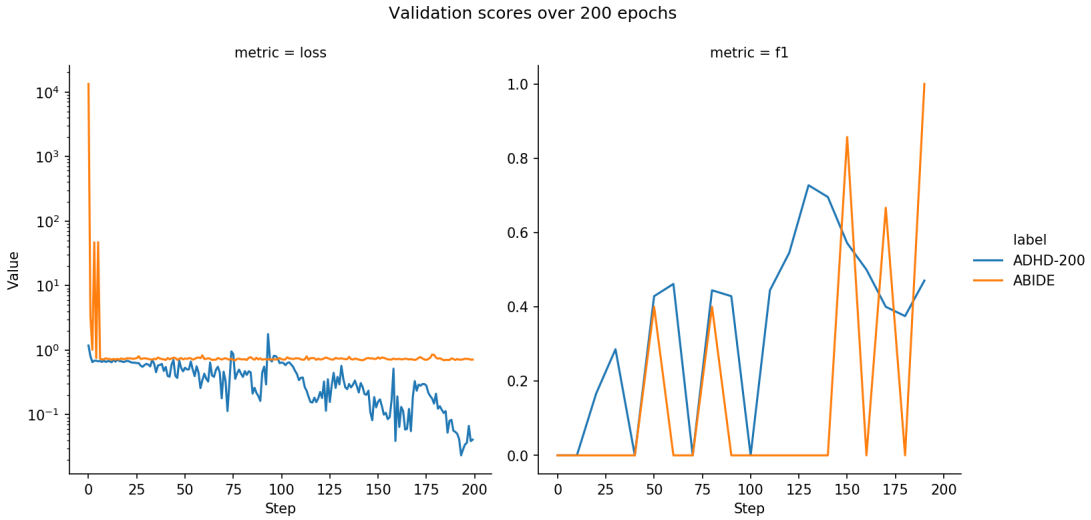

Figure 4. Training performance of GCN models over 200 epochs. Left: negative log likelihood loss on the validation set. Right: F1 scores on the validation set. The hyperparameters for these models can be found in Table 3.

Model	Dataset	Training type	Train F1	Test type	Test F1
SVM	ADHD-200	nodes	1.0	nodes	1.0
SVM*	ADHD-200	nodes	1.0	graphs	1.0
SVM	ADHD-200	graphs	0.87	graphs	0.43
Logistic regression	ADHD-200	nodes	0.60	nodes	0.60
Logistic regression*	ADHD-200	nodes	0.60	graphs	0.44
Logistic regression	ADHD-200	graphs	0.64	graphs	0.26
SVM	ABIDE	nodes	0.60	nodes	0.60
SVM*	ABIDE	nodes	0.60	graphs	0.61
SVM	ABIDE	graphs	0.35	graphs	0.27
Logistic regression	ABIDE	nodes	0.54	nodes	0.54
Logistic regression*	ABIDE	nodes	0.54	graphs	0.49
Logistic regression	ABIDE	graphs	0.56	graphs	0.49

Table 2. Baseline supervised learning results evaluated using the binary F1-score. *The Train F1 and Test F1 scores for these models are obtained by applying the corresponding trained model one row up in the table on the training and testing *graph* datasets.

work and compared the mean number of activating and inhibiting interactions between regions in each patient over time. Patients with ASD and TD show no significant difference in quantities of activation and inhibition (Figure 6). While the number of activation and inhibitions show no significant variation, we investigated the spatiofunctional networks of the patients to identify which regions of the brain are undergoing such interactions.

We found that 30 regions were differentially activated in patients with TD and 3 regions were differentially inhibited in patients with TD. Meanwhile, 17 regions were differentially activated in patients with ASD and 19 regions were differentially inhibited in patients with ASD. Most notably, the left and right superior frontal gyri were activated in patients with TD, but inhibited in patients with ASD. Patients with ASD, however, showed a significant activation of middle and inferior occipital gyri, which are associated with the visual cortex. This aligns with current literature, which suggests that patient with ASD have an overactive occipital cortex [14].

Because patients with ASD are typically reported to have difficulty with phonological and grammatical language skills, we investigated the differences in activation of the corresponding brain regions across time. The functional patterns of the left and right posterior and anterior cingulate cortex, the superior frontal gyri, and the middle frontal gyri were investigated. We found that the patterns of activation across patients with ASD were not significantly different than those of patients with TD ($p = 0.11$, two-sample t-test).

Networks were also generated at each timepoint for each subject in the ADHD and TD datasets. Similarly, comparing degree distributions and adjacency matrices showed no significant difference in structural connectivity between

Hyperparameters	Dataset	Train Loss	Val F1	Test F1
NL = 3, HD = 10, LR = 0.1	ADHD-200	0.29	0.73	0.80
NL = 2, HD = 15, LR = 0.1	ABIDE	0.70	1.00	0.40

Table 3. GCN model supervised learning results evaluated using the binary F1-score.

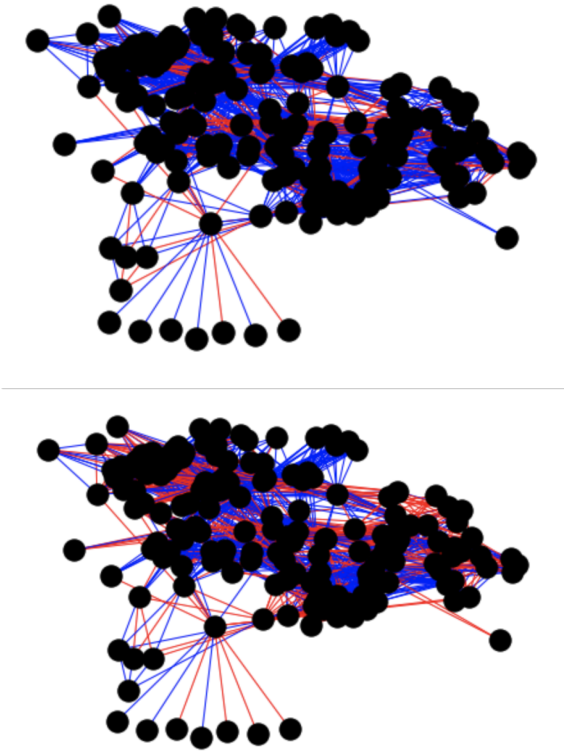

Figure 5. Network of Patient Stanford 0051162 with ASD (*left*) and Patient Stanford 0051196 with TD (*right*) at timepoint 66. Red edges indicate an activating interaction. Blue edges indicate an inhibiting interaction.

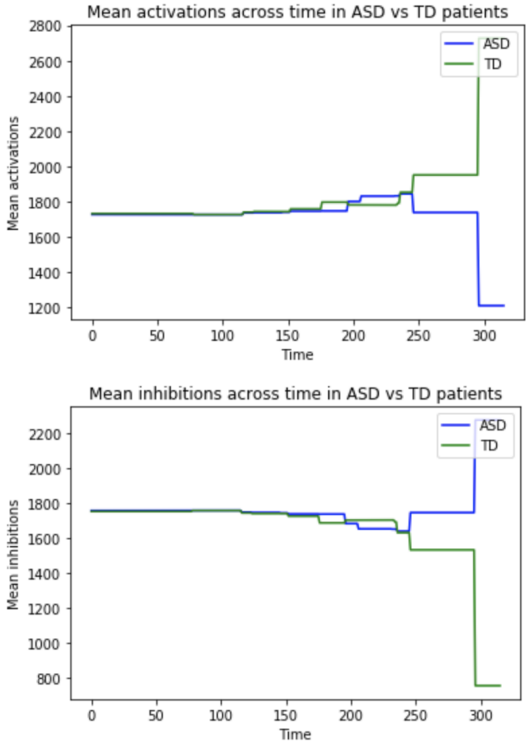

Figure 6. Mean number of activated and inhibited regions in patients with ASD (*left*) vs patients with TD (*right*)

patients with ADHD and TD (Figure 7). We then compared the functional connectivity and spatiofunctional connectivity of the networks across patient types. Patients with ADHD have an increased mean number of activations across time, compared to patients with TD (Figure 8).

While patients with TD exhibit differential inhibition in 15 regions, patients with ADHD exhibit differential activation in 11 regions over time. Most notably, the precuneus – which is responsible for processing sensory information – and rolandic operculum – which is responsible for impulse control and halting motion – were activated. The other activated regions play roles in auditory information processing, emotional formation, and learning. Patients with TD exhibited inhibition of the precentral and postcentral gyri, which control the motor cortex and the somatosensory cortex, respectively. Moreover, the cerebellum, which regulates motor movement, is differentially inhibited in patients with TD.

We compared the activation patterns of regions responsible for motor control between patients with ADHD and patients with TD. The relevant regions included the postcentral and precentral gyri, the fusiform gyri, the lingual gyri, the superior frontal gyri, the and the precuneus. Neuronal activation patterns of these regions showed a significant difference between the two patient types ($p = 4.19 * 10^{-5}$, two-sample t-test).

The spatiofunctional networks of patients with ADHD and TD were searched for a temporal motif $M(V, E)$ where $V = \{1, 2, 3, 4\}$ and $E = \{(1, 2), (2, 3), (3, 4)\}$. Given the complexity of identifying temporal motifs, we imposed a spatial limit on the search to only the motor cortex. This 4-node motif was chosen to mimic the neurological network of the limbic system, which is formed by the hippocampus, amygdala, hypothalamus, and thalamus. No significant difference in temporal motif distribution was observed between patients with ADHD and patients with TD ($p = 0.81$,

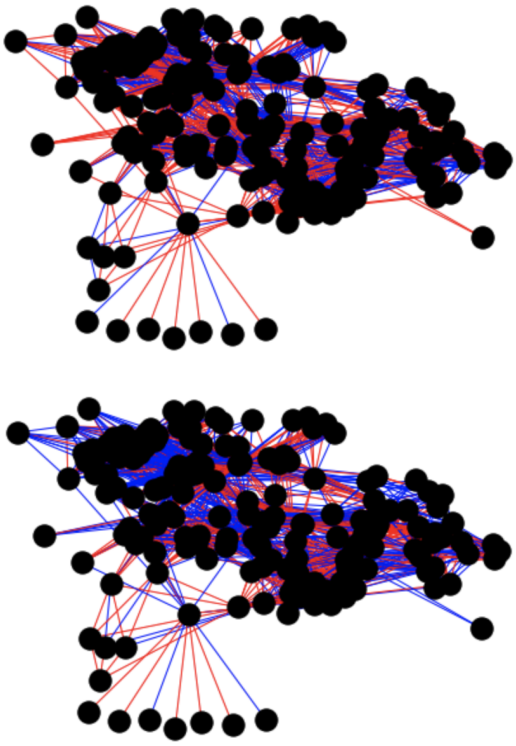

Figure 7. Network of Patient 3767334 with ADHD (*left*) and Patient 3494778 with TD (*right*) at timepoint 66. Red edges indicate an activating interaction. Blue edges indicate an inhibiting interaction.

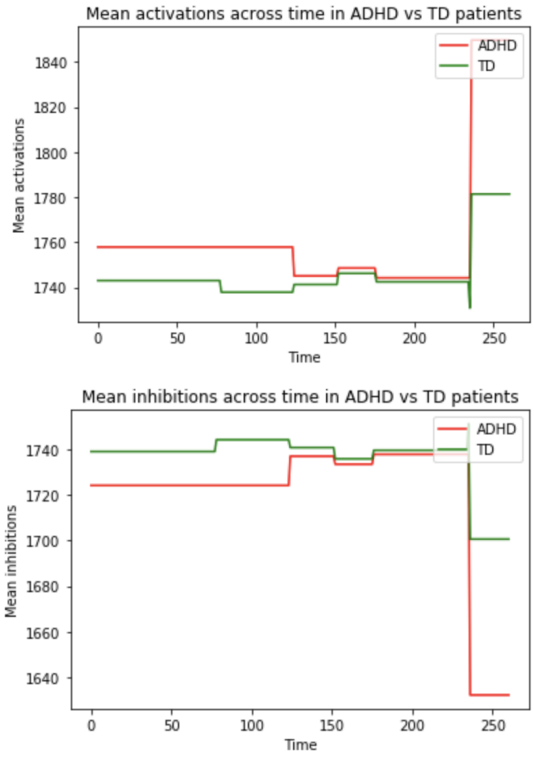

Figure 8. Mean number of activated and inhibited regions in patients with ADHD (*left*) vs patients with TD (*right*)

two-sample t-test).

6. Discussion

While we achieved good classification performance on the ADHD-200 dataset, there is clearly room for improvement on the ABIDE dataset. For example, for the GCN models, further hyperparameter tuning could be performed on the number of fully connected layers, type of activation function (e.g., sigmoid instead of ReLU), and loss function, among others. Additionally, further work could be done to interpret the models to identify what features of the functional brain networks are contributing to model predictions. Other node representations could also be explored. We chose to feed the raw timeseries data (as opposed to node embeddings derived from correlation matrices) into our GCN models because we hypothesized that GCNs were sufficiently expressive to benefit from the temporal information. Further work could look into alternative node representations, including node embeddings methods for brain regions proposed by You et al. [15].

Our analysis of the spatiofunctional networks of patients with ASD and TD focused on regions of the brain that control the motor movements for speech production. However, current research shows that children and adults exhibit difficulties with language due to differences in auditory information processing, integrating information storage and retrieval, and perceptions of human speech [14]. Children and adults with ASD, for instance, show different patterns of temporal lobe activation, which is responsible for analyzing complex auditory information like speech. The linguistic symptoms exhibited by patient with ASD are associated with higher-order cognitive functions, rather than the motor control of speaking itself. This may explain why no differences in activation patterns were observed. Future work should compare regions of the brain responsible for cognitive functions, such as Wernicke's area in the superior temporal gyrus, which controls speech comprehension. Similar comparisons of the motor cortex of patients with ADHD and TD revealed significant differences in activation patterns. Patients in ADHD have heightened activity of the right superior frontal gyrus, which is negatively correlated with motor movement termination and impulse control [7]. Our findings agree with the symptoms presented by patients with ADHD, such as fidgeting and other hyperactive movements. Moreover, further fMRI studies show dysfunction in the circuitry of the lateral prefrontal cortex, which includes the precentral and postcentral gyrus [1]. Altered functional connectivity of the precuneus is also supported by existing

research [2][13]. However, there is currently no evidence that the lingual gyri, fusiform gyri, and superior frontal gyri exhibit different behavior between patients with ADHD and TD. Though patients with ADHD have a smaller volume of superior frontal gyri, further investigations on its functional patterns must be undertaken [9]. Our analysis was also limited to the motor symptoms associated with ADHD. However, we did not consider other phenotypes, such as lack of focus, self-interested behavior, and emotional turmoil.

Across both datasets of patients with ASD and ADHD, we limited our functional analyses of brain regions. The fusiform gyri, for instance, are involved in motor skills but also social behavior. However, we investigated only the interactions of fusiform gyri with other nodes involved in movement. Given the interconnected nature of the brain, analyzing how a node interacts with its functionally distinct neighborhoods can reveal differential patterns of activation. Moreover, such patterns can guide future investigations into the neural circuitry of patients with ASD, ADHD, and TD.

7. Conclusion

In this paper, we aimed to leverage network analysis methods to better understand the biological basis of ASD and ADHD and to create a predictive model to diagnose developmental disorders based on neural connectomes. We combined structural and functional information to create spatiofunctional networks of each patient across time. Analyses of these networks demonstrate significant differential activation of the motor cortex between patients with ADHD and TD. Failure to identify significant differences in the speech cortex of patients with ASD, however, suggest that regions correlated with high-order cognitive functions should be investigated.

While our baseline and GCN models exhibit similar performance, this variability may be due to the limited dataset size. Additionally, all edges between the CC200 ROIs were hand-annotated. Increasing the number of annotations to create a more robust network may improve performance. Increasing the parcellation resolution from 200 ROIs to 400 ROIs using Craddock 400 may also improve model performance. For example, ROI 8 of CC200 corresponds to the left postcentral and left precentral gyri.

Future directions include comparing the neural regions that control different symptoms in patients with ADHD and ASD with the same regions in patients with TD. For instance, patients with ADHD and patients with ASD often demonstrate impaired social skills, which are controlled by the fusiform gyri. Additionally, given our initial analyses of resting fMRI, we believe that analyzing fMRI data taken during task performance can further emphasize differences in neurological behavior. For instance, the fMRI scans taken during listening comprehension can be investigated to show how linguistic processing differs between patients with ASD vs TD. Lastly, current fMRI studies look at autism and ADHD separately. Little is known regarding the co-occurrence of the two disorders, although the two disorders share many similar neuropsychological and behavioral symptoms. Studies show that two-thirds of individuals with ADHD show features of ASD while 50% of individuals with ASD manifest symptoms of ADHD [8]. Thus, directly comparing the spatiofunctional networks of patients with ASD and ADHD can identify shared differences in neural circuitry. This can inform the scientific community on the biological bases of the two developmental disorders, and guide future predictive models in diagnosing the two.

8. Acknowledgements

We thank the CS 224W instructors for their guidance and feedback throughout the project, as well as providing us more than sufficient credits for computational resources on Google Cloud Platform.

References

[1] George Bush. Cingulate, frontal, and parietal cortical dysfunction in attention deficit/hyperactivity disorder. *Biol Psychiatry*, (69), 2011.

[2] Castellanos F.X., Marguiles D.S., Kelly C., Uddin L.Q., Ghaffair, M., Kirsch, A., et al. Cingulate precuneus interactons: a new locus of dysfunction in adult attention-deficit/hyperactivity disorder. *Biol Psychiatry*, (63), 2008.

[3] Cameron Craddock, Sharad Sikka, Brian Cheung, Ranjeet Khanuja, Satrajit S Ghosh, Chaogan Yan, Qingyang Li, Daniel Lurie, Joshua Vogelstein, Randal Burns, Stanley Colcombe, Maarten Mennes, Clare Kelly, Adriana Di Martino, Francisco Xavier Castellanos, and Michael Milham. Towards automated analysis of connectomes: The configurable pipeline for the analysis of connectomes (c-pac). *Frontiers in Neuroinformatics*, (42), 2013.

[4] R. Cameron Craddock, G.Andrew James, Paul E. Holtzheimer, Xiaoping P. Hu, and Helen S. Mayberg. A whole brain fMRI atlas generated via spatially constrained spectral clustering. *Human Brain Mapping*, 33(8):1914–1928, aug 2012.

[5] Simon B. Eickhoff, Klaas E. Stephan, Hartmut Mohlberg, Christian Grefkes, Gereon R. Fink, Katrin Amunts, and Karl Zilles. A new spm toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *NeuroImage*, 25(4):1325 – 1335, 2005.

[6] Aditya Grover and Jure Leskovec. node2vec: Scalable Feature Learning for Networks. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '16*, pages 855–864, New York, New York, USA, 2016. ACM Press.

[7] Hu, S., Jaime S. Ide, Sheng Zhang and Chiang-shan R. Li. The right superior frontal gyrus and individual variation in proactive control of impulsive response. *Journal of Neuroscience*, (36), 2016.

[8] Yael Leitner. The Co-Occurrence of Autism and Attention Deficit Hyperactivity Disorder in Children - What Do We Know? *Frontiers in Human Neuroscience*, 8:268, apr 2014.

[9] Monuteaux, MC, et al. A preliminary study of dopamin d4 receptor genotype and structural brain alterations in adults with adhd. *Am J Med Genet B Neuropsychitr Genet*, (147B), 2008.

[10] Siddharth Ray, Meghan Miller, Sarah Karalunas, Charles Robertson, David S. Grayson, Robert P. Cary, Elizabeth Hawkey, Julia G. Painter, Daniel Kriz, Eric Fombonne, Joel T. Nigg, and Damien A. Fair. Structural and functional connectivity of the human brain in autism spectrum disorders and attention-deficit/hyperactivity disorder: A rich club-organization study. *Human Brain Mapping*, 35(12):6032–6048, dec 2014.

[11] Ardesheer Talati and Joy Hirsch. Functional specialization within the medial frontal gyrus for perceptual go/no-go decisions based on "what," "when," and "where" related information: An fmri study. *Journal of Cognitive Neuroscience*, 17(7):981–993, 2005.

[12] The ADHD-200 Consortium. The ADHD-200 Consortium: a model to advance the translational potential of neuroimaging in clinical neuroscience. *Frontiers in Systems Neuroscience*, 6:62, 2012.

[13] Uddin, L.Q., Kelly, A.M., Biswal, B.B., Marguiles, D.S., Shehzad, Z., Shaw, D., et al. Network homogeneity reveals decreased integrity of default-mode network in adhd. *Journal of Neuroscience Methods*, (169), 2008.

[14] Williams, Diane L., and Nancy J. Minshew. How the brain thinks in autism: Implications for language intervention. *American Speech-Language-Hearing Association*, (15), 2020.

[15] Hongyuan You, Adam Liska, Nathan Russell, and Payel Das. Automated brain state identification using graph embedding. In *2017 International Workshop on Pattern Recognition in Neuroimaging (PRNI)*, pages 1–5, Toronto, ON, Canada, jun 2017. IEEE.