1 Introduction

In this paper we will be examining and analyzing the white matter structure of pre-teens and adolescent individuals with autism spectrum disorder (ASD) as collected through diffusion tensor imaging and developed with tractography. Given the pivotal nature of axon tracts on coordinating communication among brain regions (Fields, 2008), and the inherent network-like structure of these tracts, we believe that network and graph analysis may provide greater insight and understanding on the relationship between ASD and white matter structure.

More specifically, we will be examining the neural networks of 52 subjects with ASD and 43 under typical development (TD) as unique, weighted, undirected graphs, with nodes being brain regions and edges being the existence and number of neural connections between these regions. Our initial focus is on understanding various features of these regions and their interconnections, including centrality, clustering and motif analysis. In order to ground our analysis, we will be using the non-parametric Friedman test and the Wilcoxon rank-sum test to highlight any statistically significant differences between the two subject groups, tying our results back to current medical knowledge on the relationship between ASD and white matter structure.

2 Related Work

2.1 DTI and Autism

Diffusion tensor imaging (DTI) tractography is a technique used to represent white-matter neural tracts in the brain – namely, axon tracts and commissures. White matter has an active effect on learning, self-control, and mental illness, controlling the signals that neurons share (Fields, 2008).

Autism spectrum disorder (ASD) is a group of neurodevelopmental disorders with characteristics that include impaired social cognition and reciprocity, and repetitive, restricted behavior (NIMH, 2016; Lord et al., 2000).

There is evidence that a connection between ASD and abnormal white matter structure exists with studies on children and adolescents with autism finding that disruption of white matter tracts in white matter adjacent to the ventromedial prefrontal cortices, in the anterior cingulate gyri and the temporoparietal junctions may be implicated in impaired social cognition for individuals with ASD (Barnea-Goraly et al., 2004). Developmental studies of white matter in males suffering from ASD has found that there is a reduction in the structural integrity of white matter – namely, lower fractional anistrophy near the corpus callosum and in the right retrolenticular portion of the internal capsul – that may underlie the behavioral pattern observed in autism (Keller et al., 2007). Recent studies have been placing more attention on analyzing how the nature of abnormality in white matter integrity affects ASD behavioral phenomena, though most work has been kept within the boundaries of biological analysis of white matter fiber tracts (Pryweller et al., 2014; ACE, 2016). We believe that representing white matter structure as a graph and using complex network methods and measures to analyze information gather with DTI scans may help us gain further insight as to how ASD is associated with white matter formation and structure in the brain.

2.2 Brain Neural Network Analysis

The network structure of brain connectivity is a popular topic of research and exploration. The uses and interpretations of a wide array of network measures and tools have been well documented, including measures of segregation and integration, centrality, and resilience (Rubinov and Sporns, 2010). Various mapping techniques, including functional magnetic resonance imaging,
electroencephalogram, magnetoencephalography, and susceptibility tensor imaging, which capture different aspects of the functional or anatomical structure of the brain, have been analyzed with various network tools and methods (Sporns et al., 2007; Joyce et al., 2010). For instance, eigenvector centrality has been used for analyzing connectivity patterns in temporal fMRI data of the brain, with the authors concluding that the tool is an effective and computationally efficient tool for capturing intrinsic neural architecture at a voxel-wise level (Lohmann et al., 2010).

An effort to evaluate the brain architecture in autism does exist, with the Autism Imaging Data Exchange consortium aggregating and sharing over a thousand resting-state fMRI datasets of male subjects with ASD and TD (Di Martino et al., 2014). However, the structural information captured by fMRI scans is very different to that collected in DTI tractography methods – namely, while DTI tractography attempts to reveal the white matter structure of the brain, fMRI measures brain activity through changes in blood flow (Matthews and Jezzard, 2004). Notably, little work has been done in analyzing and comparing the white matter structure of the brain of individuals with ASD, despite the strong belief that myelination and white matter development in infant and teenage years can be different for people with ASD.

3 Data

We will be working with diffusion tensor imaging (DTI) data collected by the Center for Autism Research and Treatment at UCLA, obtained from the USC Multimodal Connectivity Database (Brown et al., 2012). The sample consists of 52 subjects with autism spectrum disorders (ASD) and 43 individuals under typical development (TD), all between the ages of 8 and 18. These two groups do not have statistically significant differences in their sex, age, mean and maximum head motion, or their full-scale, verbal, and performance IQ (Rudie et al., 2013). The data was collected on a Siemens 3T Trio scanning device at UCLA. After being asked to relax and keep their gaze focused on a fixation cross on a screen, T2*-weighted functional images were captured, with a TR of 3000ms, TE of 27ms, a $128 \times 128$ matrix size, 192mm FoV and $3.0 \times 3.0$mm in-plane voxel dimensions. These scans were consequently analyzed and preprocessed, with individuals with excessive motion not included in the final dataset (Rudie et al., 2013). Brain deterministic tractography, which aims to represent the neural tracts collected by the diffusion MRI images, was then performed on the scans using the Fiber Assignment by Continuous Tracking (FACT) algorithm (Mori and van Zijl, 2002), a state of the art method for fiber tracking from DTI imaging data. Fibers shorter than three voxels were discarded (no turn angle can be determined from just two voxels). The maximum turn angle of fibers propagating from voxels was of $50^\circ$ (Zalesky et al., 2010). Fibers were consequently smoothed using spline filters (Rudie et al., 2013).

The final data consists of 264 putative functional areas as defined by (Power et al., 2011), where edges between these regions correspond to the number of fibers where one endpoint finished in one region and the other endpoint in the other (Rudie et al., 2013). The data was structured in a
264 x 264 whole brain structural connectivity matrix. Each of the 94 subjects has a corresponding connectivity matrix. Figure 1 shows the graph for one of the individuals with ASD.

4 Methods

We will survey the various tools and measures we are going to be using to analyze the graphs, as well as the statistical methods we are going to be using to evaluate the statistical significance of the differences captured by these models on the two subject groups. Some of these measures are on an entire graph, while others capture information on a per-vertex basis.

4.1 Clustering

Clustering of brain networks attempts to capture the interconnectivity of groups or clusters within the network. The organization of dependencies captured by these methods may indicate the existence of segregated neural processing (Rubinov and Sporns, 2010). We will be calculating the local clustering coefficient of every node in every graph. This coefficient $c$ is for node $n$ is defined as

$$c_n = \frac{t_n}{k_n(k_n - 1)}$$

where $t_n$ is the number of triangles around a node $n$ and $k_n$ is the degree of node $n$. We will also be calculating the global clustering coefficient, defined as the average of the clustering coefficients of each node.

4.2 Transitivity

An issue with the clustering coefficient is that it may be disproportionately influenced by nodes with low degrees (Rubinov and Sporns, 2010). The transitivity does not suffer from this problem, and is defined as follows

$$T = \frac{\sum_{n \in N} \frac{2t_n}{k_n(k_n - 1)}}{\sum_{n \in N} k_n(k_n - 1)}$$

4.3 Centrality

We’ll be using a number of measures to calculate the centrality of the brain region nodes of our graphs. Each of these centrality measures captures different nuances about the importance and influence of various nodes in our graphs. We proceed with short introductions of the various centrality measures utilized in this project.

4.3.1 Closeness Centrality

The closeness $c$ of any given vertex $n_i$ is defined as

$$c_n = \frac{1}{\sum w_{n_i,n_j}}$$

where $w_{n_i,n_j}$ is the weight of the edge between nodes $n_i$ and $n_j$ (Opsahl et al., 2010).

4.3.2 Betweenness Centrality

We define the betweenness centrality of a node $n$ as

$$c(n) = \sum_{n_i \neq n_j \neq n} \frac{\sigma_{n_i,n_j}(n)}{\sigma_{n_i,n_j}}$$

where $\sigma_{n_i,n_j}$ is the number of shortest paths between $n_i$ and $n_j$, and $\sigma_{n_i,n_j}(n)$ the number of shortest paths between $n_i$ and $n_j$ that pass through $n$ (Brandes, 2001).

4.3.3 Eigenvector Centrality

The eigenvector centrality $x$ of a list of vectors $n$ with weighted adjacency matrix $A$ is defined as the eigenvector with the largest eigenvalue $\lambda$ (Langville and Meyer, 2005). Namely, it is the solution of $Ax = \lambda x$ where $\lambda$ is the largest eigenvalue.

4.3.4 Authority and Hubs

Authorities $y$ and hubs $x$ are defined as

$$x = \alpha Ay$$

$$y = \beta A^\top x$$

where $A$ is the weighted adjacency matrix and $\lambda = (\alpha \beta)^{-1}$ is the largest eigenvalue of $AA^\top$ (Kleinberg, 1999). Since our graphs are undirected, then our adjacency matrix is symmetric along the diagonal. Thus, $\lambda$ will be the singular value of $A$ and $x = y$. Therefore, the authority and hub value for every node will be the same (we will be reporting a single value for each node).

4.3.5 PageRank

For any node $n$ the PageRank value $R(n)$ is defined iteratively by

$$R(n) = \left(\frac{1 - d}{N}\right) + d \sum_{x \in \Gamma(n)} \frac{R(x) A_{n,x}}{\sum_y A_{n,y}}$$

where $A$ is a weighted adjacency matrix, $\Gamma(n)$ are the neighbors of $n$, and $d$ is a damping factor (Page et al., 1999).
4.4 Characteristic Path Length

The characteristic path length $L$ is defined as the mean averaged distance between every node and every other node. Namely

$$L = \frac{1}{n} \sum_{i \in N} L_i = \frac{1}{n} \sum_{i \in N} \frac{\sum_{j \in N, j \neq i} d_{ij}}{n - 1}$$

where $d_{ij}$ is the length of the shortest path between nodes $i$ and $j$, where, if no path exists, $d_{ij} = \infty$ (Rubinov and Sporns, 2010).

4.5 Efficiency

The efficiency of a graph is a measure of how well and efficiently information can be transmitted across the network (Rubinov and Sporns, 2010). The efficiency $E_i$ of a node $i$ is defined as

$$E_i = \frac{\sum_{j \in N, j \neq i} d_{ij}^{-1}}{n - 1}$$

The efficiency has been argued to be a more meaningful measure of integration than shortest path length since rather than being influenced by long paths it’s influenced primarily by short paths (Achard and Bullmore, 2007). The global efficiency of a network is the average efficiency of all of its nodes, as follows:

$$E = \frac{1}{n} \sum_{i \in N} E_i$$

4.6 Statistical Tests

We will be using two statistical tests to capture the statistical significance between any differences found between the two subject pools. The Friedman non-parametric test will be used on metrics that are captures on a per-node basis, while the Wilcoxon signed-rank test will be used on metrics that are captured on a per-graph basis.

4.6.1 Friedman Test

We will be testing the statistical significance of the differences in metrics calculated from the graphs of both subject pools. The Friedman test, similar to the parametric repeated measures ANOVA, is a non-parametric test used for one-way repeated measures of variance by ranks (Friedman, 1937). It tests the null hypothesis that two sets of measurements of the same individuals have the same distribution (Jones et al., 2001). We will be using the Friedman test on every average of every metric for each node to understand if there is a statistically significant difference in the information captured by that metric on the nodes or edges of the graphs of subjects with ASD and TD, with a set $p$-value of 0.05.

4.6.2 Wilcoxon Signed-Rank Test

The Wilcoxon signed-rank test is a non-parametric test used to compare two samples to test the null hypothesis that their population mean ranks differ. We’ll be using it to compare the results of whole-graph measurements for all graphs in both samples, such as transitivity or global efficiency.

The principal reason why we are using the Wilcoxon signed-rank test rather than the pared Student’s t-test is because it does not require us to assume that our distribution is normally distributed (Lowry, 2014), which is not an assumption, particularly with our relatively small number of examples, that we want to make. This makes the Wilcoxon signed-rank test a more reasonable method for our particular use. We will set our $p$-value to 0.05.
5 Initial Results

We have already run a majority of the metrics discussed previously on all of the graphs on both ASD and TD networks. The top five nodes and their corresponding average metric across all networks for each value are presented in Tables 2 and 1. Just a sampling of the top 5 nodes indicates that there seem to be some notable differences between the metrics on ASD and TD subjects, which is promising to our current research. We will proceed by calculating the $p$-score of these and the remaining metrics with the Friedman and Wilcoxon signed-rank tests in order to determine the statistical significance of these differences.

6 Further Work

Once we have all metrics calculated we will delve deeper into the graph aspects which presented greater statistical significance between the two groups – we will not know what these necessarily are until we have all of our results computer. Our ultimate goal is to tie this information back to the literature on autism detailed in the Related Work section and on the function of various brain regions to understand whether our findings and results align with current knowledge on white matter structure in individuals with ASD, or whether there are certain aspects of the structure that has not been captured by current knowledge and efforts in this field. Depending on the significance of these results, we are also planning on comparing both graphs to null models. We would be running configuration models for each of the networks, generating 1000 for each. These null models can ground our research and our understanding of the structure of white matter in the brain for both subject pools, as we can calculate the statistical significance in difference between metrics of both subject sets and the null model.

We hope that with the research performed for this project we will be able to gain greater and new insight as to the differences in structure and form of white matter in adolescents and preteens.

References

UCSD Autism Center of Excellence ACE. 2016. Diffusion tensor imaging (dti).


Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Winograd. 1999. The pagerank citation ranking: bringing order to the web.


