Graph-Analysis of Temporally- and Spatially- Defined Brain Networks Involved in Decision-Making

Daniel Tang – danieldt@stanford.edu
Computer Science Department
Stanford University

1. Introduction

Advancements in neural recording methods have rapidly expanded the size, scope, and complexity of neural data. These data are collected with various degrees of temporal and spatial resolution, and include static anatomical data as well as time-series neural activity data. Given the complexity and scale of modern neural datasets, the application of computational methods is critical in extracting underlying functional roles and structures within the data. Network analysis is especially complementary to neurological datasets given the physical functional relationships encoded in the brain. While network analyses have been used in various fields, relatively little work has been done in systems neuroscience, with work in related fields focusing on static anatomical maps of the brain. Additionally, given the goal of systems neuroscience to understand functional activity from activity ensembles across multiple brain regions, there is a current lack of sophisticated computational tools being used to analyze multi-channel activity data. As such, the ability to analyze not only static, but activity-dependent neural data through network analysis will allow for greater understanding and interpretation of modern neuroscience datasets.

In this project, I will use network features to analyze both static and activity-based graphs of the central autonomic network (CAN), a seven-region brain network known to be involved in decision-making and homeostasis. More specifically, I will analyze the static network using several features related to clustering, centrality, and motifs to identify key functional nodes or structures. I will then apply these metrics to neural activity data during a decision-making task to evaluate whether network features of neural activity defined graphs can be used to predict the functional outcome of the physical network. To do this, I'll develop a pipeline to construct temporally-dependent networks from the activity data, calculate network features from these graphs, and use these features in a classifier model to predict behavioral output of the CAN. I hope to show that network analysis is compatible with time-series multi-channel neural data and is able to model the functional role of the brain network, thus also providing insight into particular network features, involved in decision-making.

2. Background

2.1 Central Autonomic Network

The central autonomic network (CAN) is a major internal regulation network that regulates physiological condition and drives behavior based on physiological state. As an example, this network might drive avoidance of stressful behaviors or locations based on incoming stress responses, like increased heart rate. The CAN is comprised of a number of brain regions including the medial prefrontal cortex (mPFC), anterior insula (AI), posterior insula (PI), central amygdala (CeA), hypothalamus (DMH), thalamus (VPL), and the nucleus of the solitary tract (NTS). These regions are associated with a variety of functions ranging from decision-making and stress-response, to hormone

and sympathetic nervous system regulation. Abnormal activity and network connectivity in the CAN have been associated with numerous medical conditions including generalized anxiety, addiction, and panic disorders. By using graph analysis on both anatomical and activity-based data from the CAN, I hope to better understand the key nodes or structures of the network that guide decision-making behavior. Identifying these key areas will not only improve basic science understanding of this network, but may help define interventional targets for patients afflicted with disorders related to abnormal CAN activity.

2.2 Graph Analysis in Neuroscience

Much of graph analysis in the broader neuroscience field has focused on static, anatomical data derived from neural tracing experiments. For instance, early applications of network analysis include the use of small-world network models on low-resolution tracing data from C. elegans, cats, and macaques. These datasets exhibited high-clustering and low path-length, the primary characteristic of small-world graphs. Given the higher-resolution of modern tracing methods, it was suggested that weighted graphs be integrated into the existing models to account for the varying degree of projection strength and density between regions in the brain (Basset and Bullmore, 2016). Heuvel and Sporns reported several novel measures of centrality for anatomical networks, tuned specifically for the organization of the brain (Heuvel and Sporns et al., 2013). While these methods work well for static anatomical data, relatively little work has been done in applying these measures to temporally-evolving networks derived from the neural activity. The first step in analyzing network features from such datasets would be to derive graphs from activity-data. Sich et al demonstrated the use of pair-wise cross-correlation to assign adjacency matrix weights between two regions based on their neural activity (Sich et al., 2019). This approach was used to define weighted, undirected networks for each trial within a behavioral experiment and then to identify important nodes within the network. I will draw from a combination of static network features and the activity-based network reconstruction method in analyzing the CAN network.

3. Data

I will be working with two types of data, 1) anatomical data collected from the Allen Brain Atlas and tracing literature, and 2) neural activity data recorded from the CAN during a decision-making task.

3.1 Anatomical Data

To define my anatomical network, I analyzed (n~100) viral tracing studies submitted to the Allen Atlas, and visually inferred anatomical projections based on presence of the fluorescence tracer in the downstream region (Fig. 1). I systematically went through each of the seven brain regions from the CAN and confirmed whether an anatomical edge existed between the current node and any of the other six nodes, thus defining a 7x7 unweighted adjacency matrix. While the edge weights of the network could have been defined based on the anatomical density of each connections, I elected to use a binary edge weight due to variability in the way tracing studies were performed for different regions and experiments in the Allen Atlas. As the number of tracing experiments differ from region-

to-region in the Allen Atlas, regions with very few tracing experiments (VPL and NTS) were assigned edges to other nodes based on tracing data in the literature. Using both the Allen Atlas and literature, I defined a directed graph for the CAN (Fig 2.).

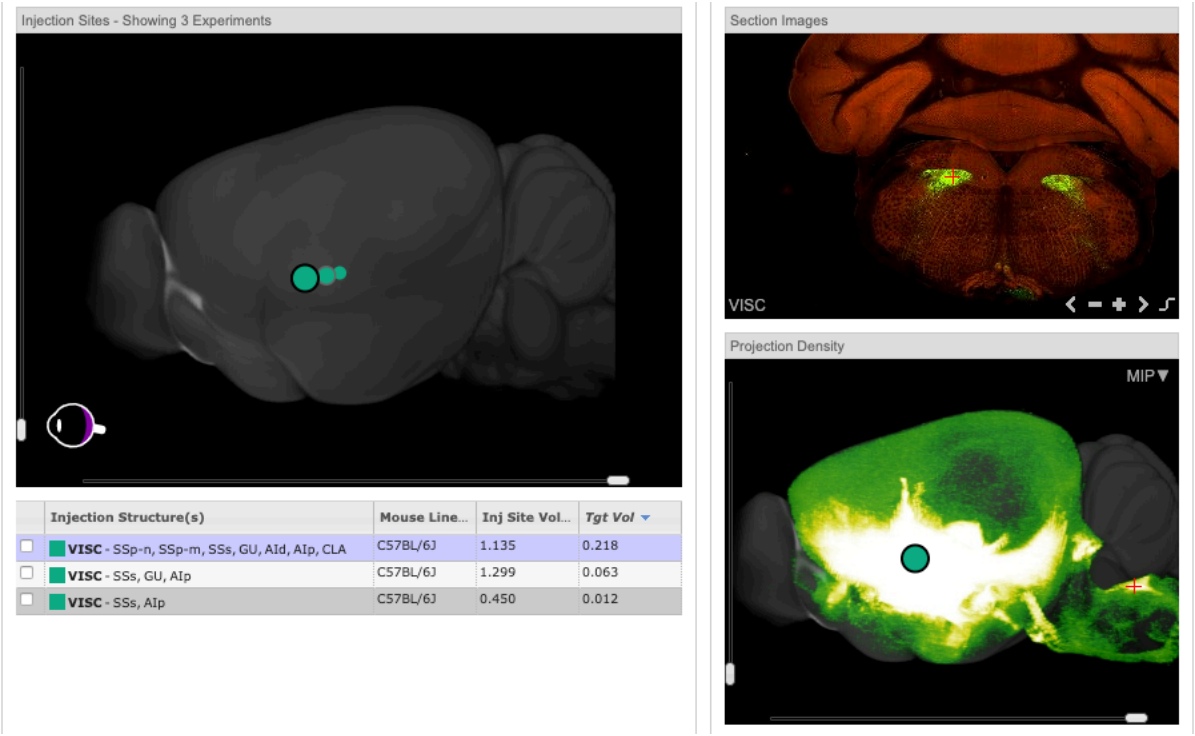


Fig 1. Example of visual inference of edge between two nodes. (Left) a list of viral tracing studies, each reporting the original site where the tracer is injected, the injection volume, and the mouse line. (Bottom Right) Visualizing the tracer throughout the brain shows projections from the target site, VISC (posterior insula), to numerous downstream targets throughout the brain. (Top Right) By examining a specific downstream region, the nucleus of the solitary tract (NTS), we are able to observe the tracer (green), and infer a directed projection or edge from posterior insula to nucleus of the solitary tract (PI -> NTS).

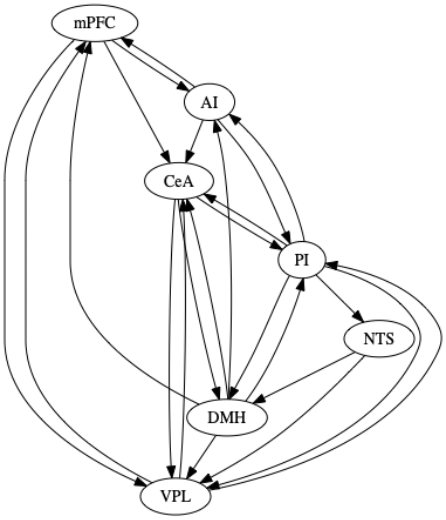


Fig 2. Anatomical network created using Allen Atlas and tracing literature

3.2 Activity Data

The neural activity-data were acquired at the Deisseroth Lab from 11 mice during a decision-making task using a fiber-photometry activity-recording setup (Kim et al, 2016). The data were recorded at 20 Hz and saved as .mat files for each experiment. Each mouse performed 3 experiments, involving 60 decisions in a "go, no-go" task, in which the mouse can either choose to perform or withhold a lever-press. For each decision, the animal is given a 5s audio cue, priming the animal that a decision is about to be made (the animal executes the decision 1s after the cue finishes). Neural activity is recorded continuously throughout each experiment and the size of the "go" (perform lever-press) and "nogo" (withhold lever-press) classes are roughly the same. The activity data were preprocessed by applying a 50 Hz low-pass butterworth filter, baseline-corrected using a moving median to account for slow changes in baseline recording over time, and finally z-scored. Following pre-processing, the continuous data for each experiment was partitioned into 60 10 second epochs, corresponding to the neural activity preceding the animal's decision outcome (5s prior to the audio cue and 5s during the cue). The data was partitioned such that each epoch captures the "decision-making" of each animal, without including any motor or choice-specific activity. Ultimately, this data constitutes a 11x180x200 matrix, in which there are 11 animals, 180 decisions per animal, and 200 sample points of neural activity per decision.

4. Methods

I used a number of node-specific and whole-graph features used previously to analyze neural fMRI data (Apostolatos A, 2018). The network features of the CAN were compared to an Erdos-Renyi random network generated using the same number of nodes and edges.

4.1 Clustering

The clustering coefficient attempts to capture functional clusters, that may be involved in specific neural processes within the larger network (Rubinov and Sporns, 2010). The local clustering coefficient c was calculated for each node n using the following equation

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

in which t_n is the number of triangles and k_n is the degree of node n. The global clustering coefficient was derived from averaging the local clustering coefficients of all nodes.

4.2 Transitivity

To account for the effect of low-degree nodes on the global clustering coefficient (Rubinov and Sporns, 2010), I also calculated the graph transitivity as defined

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

4.3 Centrality Measures

In order to capture various definitions of node importance, I used a number of centrality measures as defined below.

4.3.1 Closeness Centrality

The closeness centrality c_{n_i} for node n_i is defined as

$$c_{n_i} = \frac{1}{\sum_{n_j \in N(n_i)} w_{n_i,n_j}}$$

where w_{n_i,n_j} represents the edge weight between node n_i and its neighbor n_j (Opsahl et al., 2010).

4.3.2 Betweenness Centrality

The betweenness centrality c_n is defined as

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

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

4.3.3 Authority and Hubs

Authorities x and hubs y are defined as

$$x = \alpha A y$$
$$y = \beta A^T x$$

where A is the weighted adjacency matrix and $\lambda = (\alpha\beta)^{-1}$ is the largest eigenvalue of AA^T (Kleinberg, 1999).

4.3.4 PageRank

The PageRank value r_j for a given node j is defined as

$$r_j = \frac{1-\beta}{N} + \beta \sum_{i \to j} \frac{r_i}{d_i}$$

where β is a correction factor to account for spider-traps, N is the total number of nodes, and d_i is the degree of a predecessor node i to node j (CS224w, lecture 11).

4.4 Motif-Counting

For the last measure, I implemented a motif-counting algorithm to identify motifs in both the static and activity-based networks. I used the Exact Subgraph Enumeration method (CS224w, lecture 3), to enumerate all possible subgraphs of size k (3 for anatomical graphs, 3 and 4 for activity-based graphs), and then wrote a clustering algorithm to group isomorphic graphs together. The algorithm used a dictionary where the keys were a representative motif and the values consisted of isomorphic graphs. Unassigned graphs were compared to the representative motifs by testing for bijections based on whether the set of in- and out-degrees (total degree for activity-based networks) were the same between both motifs. New isomorphisms were added as a new key and value. In the static graph, motifs were discovered based on the anatomical. In the activity-based graphs, all possible undirected 3- and 4- node motifs were used to define the motif classes and then counts were performed for each class for each trial.

Motif analysis for the static graph was conducted by comparing counts of identified motifs to motifs from a random network. Motif analysis for the activity-derived networks were conducted by comparing the total motif counts across all go and all no-go trials for each mouse as well as comparing unique (corresponding to a unique set of nodes for a given 3- or 4- node motif) motif counts between go and no-go trials.

4.5 Activity-based Network Construction

Using the preprocessed data for each behavioral decision (7x200 matrix, corresponding to 7 channels of 10s neural data), the pair-wise cross-correlation was calculated between each channel and values normalized to zero-lag-time autocorrelations of one (xcorr, MATLAB). The maximum correlation within a lag of 250 ms was selected for thresholding, where an edge between two regions would be assigned if the cross-correlation value exceeded 2 standard deviations of a null cross-correlation matrix. The null cross-correlation matrix was derived by comparing pair-wise cross-correlations of regions of random trials. Using this network construction workflow, a network was generated for each neural activity epoch, representing 180 epochs per mouse.

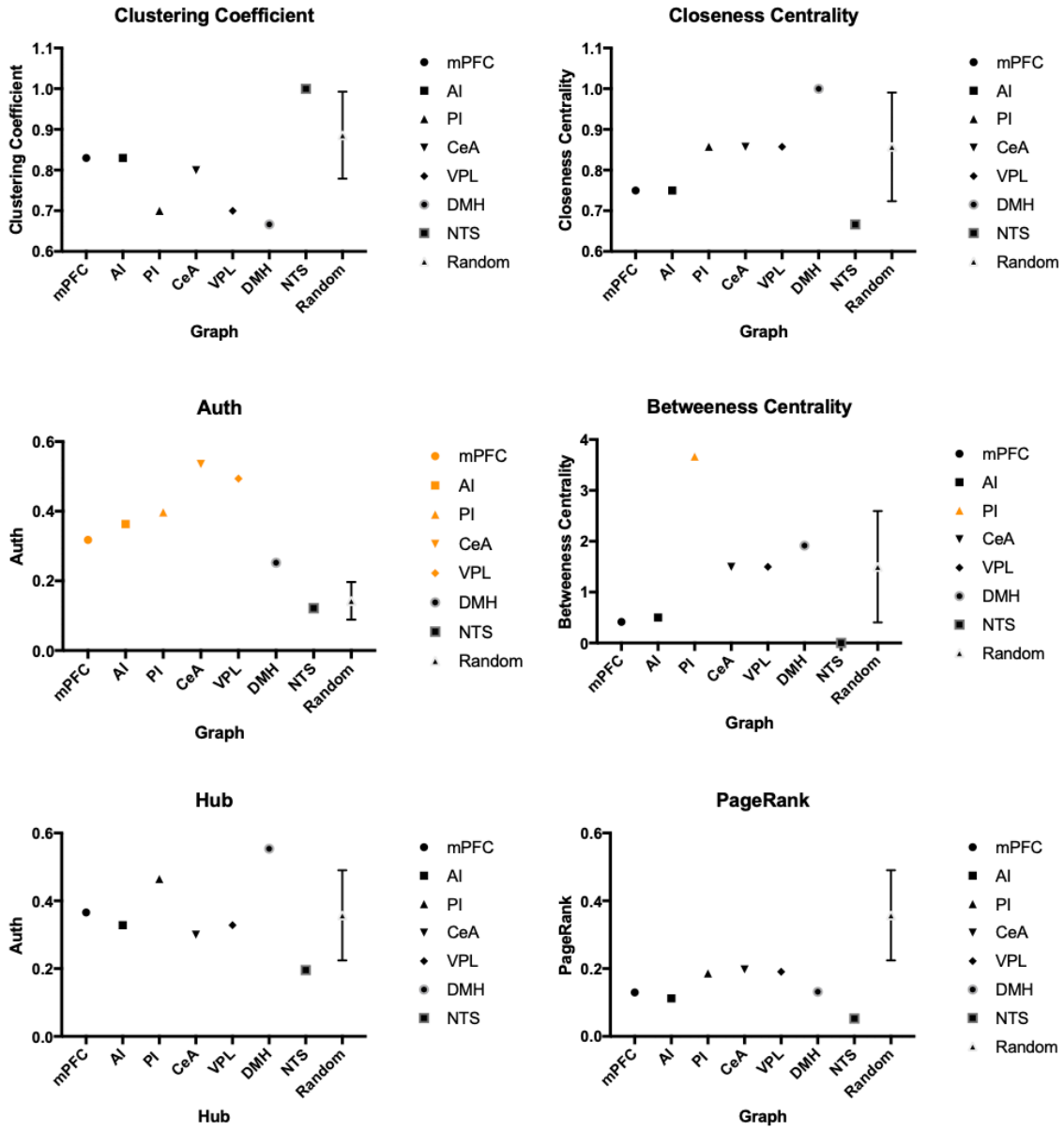
4.6 SVM Model

An SVM classifier was trained for each mouse (due to variability in neural activity across animals), using the unique motif counts and network features as input features (504 features) and the binary behavioral outcome (“go” or “no-go”) as the output class label. The neural data (180x504, trials x features) was divided into 80% training data and 20% test data. Performance metrics including precision, recall, and f1-score were reported using a weighted average across classes. Given the relatively small amount of data available, other more data-intensive models like neural networks were not considered.

5. Results

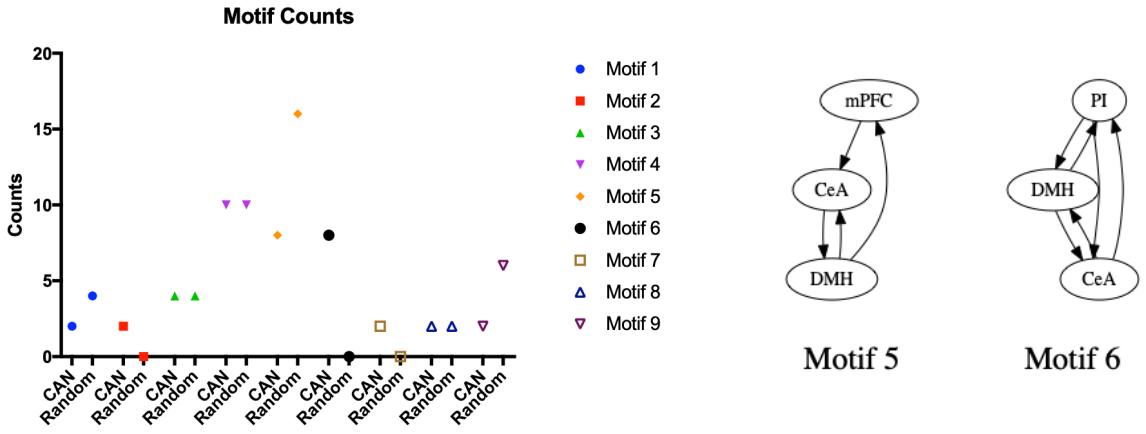
5.1 Anatomical Graph Results

5.1.1 Anatomical Graph Features



From the anatomical network features, we see that the CAN and a randomly generated graph tend to be similar across clustering coefficient, closeness centrality, hub score, and page rank. There are a number of regions that show elevated authority and betweenness centrality. Of these regions, only posterior insula displays both elevated authority and betweenness centrality.

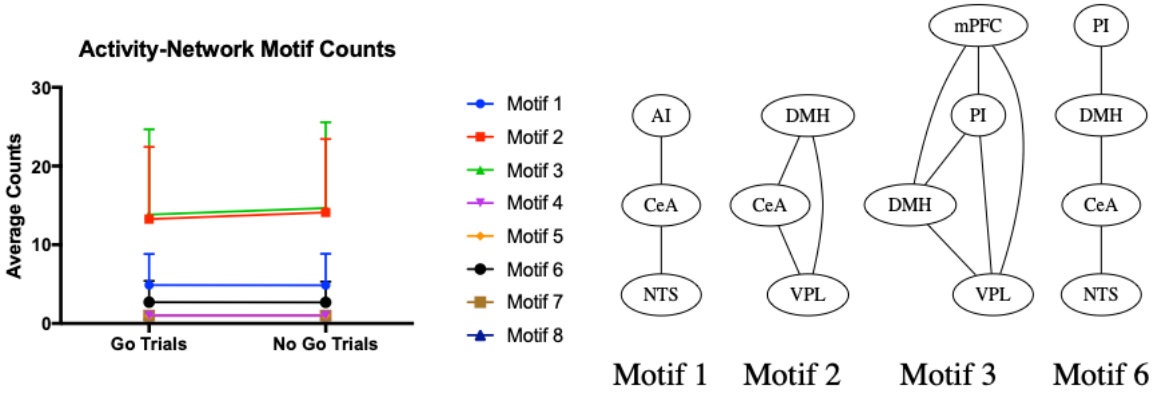
5.1.2 Anatomical Graph Motif Counts



The motif counting algorithm identified 9 different motifs in the static anatomical network. Compared to a random graph, motif 5 and 6 showed the greatest discrepancies between the two graphs. Specifically, the CAN only expressed half the count for Motif 5, what appears to be a feed-back loop between mutually connected nodes and highly elevated counts for Motif 6, three bi-directionally connected nodes.

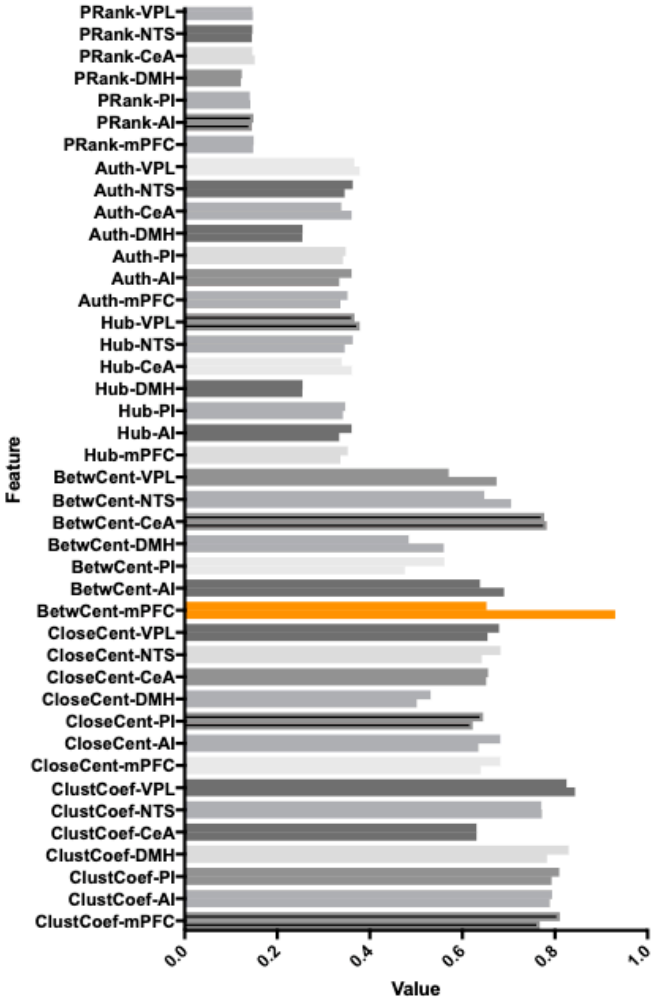
5.2 Activity-Dependent Graph Results

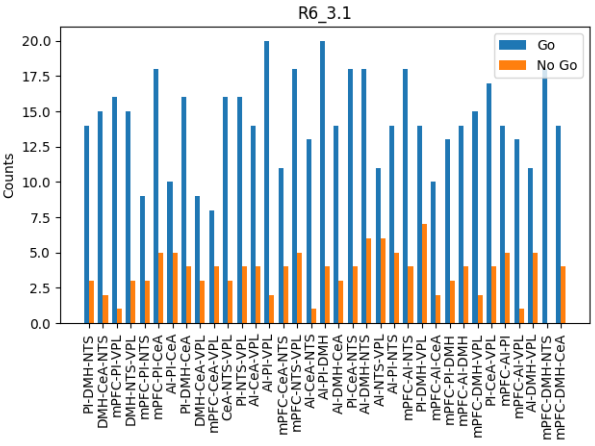
5.2.1 Motif Counts

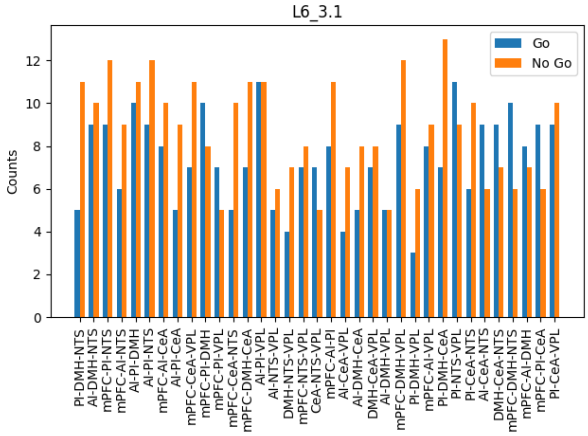


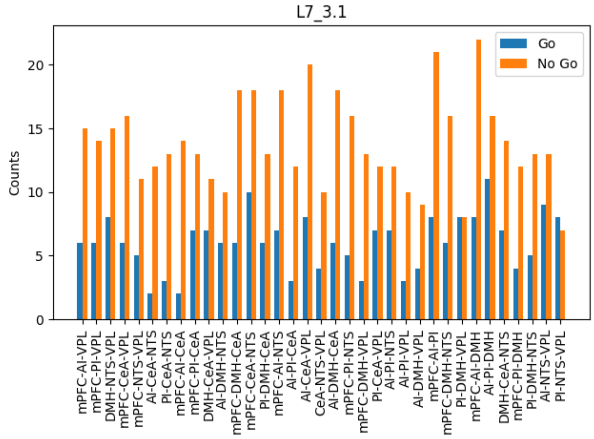
Comparing total motif counts per trial, averaged across all animals, of 3-node and 4-node motifs does not show a difference between the two decision classes. Furthermore, of the 8 possible 3- and 4- node motifs, only 4 are represented, as shown by the representative motifs 1,2,3, and 6.

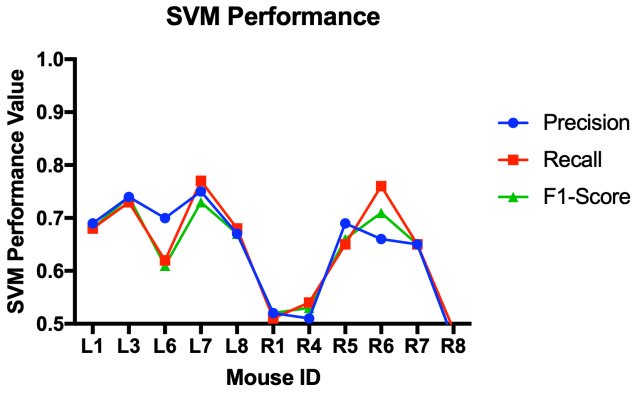
5.2.2 Network Features and SVM Performance











Network features calculated for each node of the graph indicate similar values between go and no-go trials. Betweenness centrality for the mPFC does appear to show separation between the two trial groups (top left). Though total motif counts per trial did not show a difference between go and no-go trials when counts were averaged across mice/models (5.2.1), analyzing counts of unique 3- and 4-

node motifs by specific mouse/model show differences in motifs between go and no-go trials. The models fall under 3 different regimes (right), in which mice either exhibit significantly greater network structure during go-trials (representative example in top right), similar network structure in both trial types (representative example middle right), or greater network structure during no-go trials (representative example in bottom right). Using these features as inputs into SVM models for each mouse, SVMs were able to achieve a score of 0.5-0.8 on measures of precision, recall, and f1-score. Three models performed at-chance (R1, R4, R8), while all other models perform 10-30% above chance.

6. Discussion

In applying network feature analysis to the static anatomical network of the CAN, I was able to identify a number of regions with above-random authority score, mPFC, AI, PI, CeA, and VPL, I was also able to identify a region with significant betweenness centrality, PI. Posterior Insula (PI) is a major interoceptive hub of the brain, and is thought to link brainstem and sensory nuclei, which feed in physiological information, to other cortical structures, like mPFC and AI, involved in decision-making and emotional saliency. Thus the network results seem to support existing literature that indicate PI as an integral region within this network. From the motif counting results, it appears that the CAN does not exhibit as much of Motif 5, as a random graph, this could be due to the fact that this type of feed-back motif may be redundant, and thus less utilized in nature, as the existing bi-directional relationship between two nodes of the motif preclude the need for feedback through a third node. The significant presence of Motif 6 indicates many bidirectional connections between regions, which don't appear frequently by random.

In the activity-based analysis, both network features and motif counts, when averaged across mice, did not appear to delineate between future behavioral outcome. However, upon closer inspection, it appears that there is a fair amount of variability between mice when specific motifs are analyzed that delineate go vs. no-go trials. Specifically, there appear to be three regimes within the neural activity, characterized by higher, similar, or lower network structure during the decision-making period of each subject. By training individual classifiers for each mouse, these features were able to predict future behavioral outcome with model performance between 0.6-0.8 on precision, recall, and f1-score for 8 out of 11 models. These results indicate that activity-based networks can be used to make predictions about the functional output of physical neural networks.

In conclusion, I have identified potential key nodes within the CAN through analysis of the static anatomical graph and shown the ability of activity-based network features to predict functional outcome of the CAN. Given this functional predictivity, these models can be analyzed further to identify specific regions or structures important in model prediction, thus providing insight into the functional relevance of these regions in decision-making.

7. Further Work

Additional work includes improving SVM performance by providing more data, and implementing feature elimination to reduce the risk of overfitting. Additional analysis can be done on the existing model to identify specific features that are heavily-weighted in the model prediction, thus providing biological insight into areas of interest that may play a significant role in the network during this decision-making task. Additional graph analysis can be applied, such as temporal motif counting, though such methods will require formatting for the specific time-scales of the behavior.

8. Contributions

I performed data collection and analysis for both the anatomical and neural activity data. All activity data were collected in the Deisseroth Lab at Stanford University. I wrote all code used in this project and written work for the project reports

7. Reference

Nassi JJ, Cepko CL, Born RT, Beier KT, *Neuroanatomy goes viral!*, Frontiers in Neuroanatomy, 2015

Kim C, Yang S, Pichamoorthy NJ, Young NP, Kauvar I, Jennings JH, Lerner TN, Berndt A, Lee SY, Ramakrishnan C, Davidson TJ, Inoue M, Bito H, Deisseroth K, *Simultaneous fast measurement of circuit dynamics at multiple sites across the mammalian brain*, Nature Methods, 2016

Basset S.D. and Bullmore E.T, *Small-World Brain Networks Revisited*, Sage Journals, 2016

Huevel P.v.d. M, Sporns O, *Network hubs in the human brain*, Cell Press, 2013

Sych Y, Chernysheva M, Sumanovski L, and Helmchen F, *High-density multi-fiber photometry for studying large-scale brain circuit dynamics*, Nature Methods, 2019

Apostolatos A, *Brain White Matter Network Analysis of Autism Spectrum Disorder*, CS224W Project Milestone

Rubinov M and Sporns O, *Complex network measures of brain connectivity: uses and interpretations*, Neuroimage, 2010

Opsahl T, Agneessens F, and Skvoretz J, *Node centrality in weighted networks: Generalizing degree and shortest paths*, Social networks, 2010

Brandes U, *A faster algorithm for betweenness centrality*,* Journal of mathematical sociology, 2001

Lecture 11, CS224w Fall 2019

Basset S.D and Sporns O., *Network Neuroscience*, Nature Neuroscience, 2017

Mucha J.P., Richardson T, Macon K, Porter A.M, Onnela JP, *Community Structure in Time-Dependent, Multiscale, and Multiplex Networks*, Science 2010