P8124 Final Project Proposal

qw2331 | Qihang Wu

Nov 20, 2022

- **Background and Objective:**
  Nowadays, neuroimaging has become a powerful tool to study the structure and function of the human brain in a non-invasive way. The researchers can then interpret findings and eventually understand the relevant neural activities or brain structural changes. Among multiple types of brain imaging techniques, functional magnetic resonance imaging (fMRI) is a skill to detect regional and time-varying changes in brain metabolism especially related to human cognition. A common way to perform the fMRI analysis is to restrict the analysis to specific spatial regions of interest (ROIs). From each ROI, the time series of blood oxygen level-dependent (BOLD) volume of each cube (i.e., voxel) are extracted and recorded to represent the neuronal activity within the ROI. Autism spectrum disorder (ASD) is one of the serious neuropsychiatric disorders all over the world, which will affect one's daily interactions and thereby impair communications with others. ASD occurring before the age of 3 will also have a negative and devastating influence throughout one's life[1].

  Despite some challenges embedded in the field of fMRI analysis such as the noises and some unknown psychological signals and the high dimensionality of the data in contrast to the sample size, the probabilistic graphical modeling, fortunately, sheds light on these issues.

  In this project, we aim to use the fMRI time series data in conjunction with graphical modeling to investigate the functional connectivity (FC) between different regions of the brain for people with ASD vs. healthy controls (i.e., to detect the existence of connections or "edges" and care less about the directionality) as we expect that the connections between some specific regions will be more evident in the patients than the healthy ones, evaluate the performance of various structural learning algorithms at both individual-level and group-level, and the community detection will be followed to explore which regions have denser connections in the ASD population.

- **Data:**
  The Autism Brain Data Exchange (ABIDE) initiative is to aggregate both functional and structural brain imaging data from various international sites to help the scientific community better understand the mechanism behind ASD. In this project, we use the fMRI data during the resting state (rs-fMRI) collected from Carnegie Mellon University (CMU). In the data, 27 subjects aged 19-40 years underwent the MRI scanning with a SIEMENS scanner, 14 of them were previously diagnosed with ASD (those individuals were medically healthy without identifiable genetic etiology for their disorder) while the rest 13 were regarded as the typical controls (TC). The main parameters involved are echo time (TE) = 30 ms, repetition time (TR) = 2000 ms, and $3.0mm^3$ voxels. Data was downloaded from ABIDE website and preprocessed using the NeuroImaging Analysis Kit (NIAK) pipeline, which parcelled the data into Dosenbach's 160 functional ROIs[2]. Within each ROI, the average BOLD signals were calculated.

To describe the data in a more quantitative way, we have $\tilde{y}_{ktp} = \{\mathbf{y}_{k1}, \mathbf{y}_{k2}, ..., \mathbf{y}_{kp}\}$ to represent the p-th ROI for the k-th subject at the corresponding t-th timepoint (s.t., $t \in \{1, .., T\}$), where $\mathbf{y_{kp}}$ is a random variable generating T samples of BOLD signals of p-th ROI for the subject k. The number of time points varies by different subjects since some samples are dropped through data preprocessing steps to remove potential artifacts or outliers. Meanwhile, the supplementary data contains some phenotypic information such as sex, age, estimated IQ, handedness, etc.

- **Analysis, Methods, and Software:**
  The general idea of the analysis is that the more similar the timecourses are between any given pair of nodes (i.e., ROIs), the more likely that there is a functional connection between those nodes. The proposed analysis for this project is shown as follows.

  - **EDA:** EDA like the phenotypic information table and other preprocessing steps may be eligible, s.t., remove the first 5 samples for each subject to ensure the stability of fMRI scanning.
  - **(optional) Feature Selection:** Since we will have 160 regions with only 27 subjects available. We would like to extract some important features based on the pairwise correlations at first to increase the computational efficiency. As suggested by Eslami et al.[3], the approach is to keep those correlations with the 1/4 largest and 1/4 smallest values and drop the rest.
  - **Structure Learning:** Pearson's correlation $\hat{\rho}_{i,j}$ (representing the correlation between any two random variables $\mathbf{y}_{ki}$ and $\mathbf{y}_{kj}$ as defined above), partial correlation, regularized inverse covariance (ICOV) estimation (i.e., graphical lasso), PC, GES (Greedy Equivalence Search), FCI (Fast Causal Inference), and LiNGAM, etc.
  - **Individual vs. Group Analysis:** An ASD individual sample will be randomly selected to compare the performance of different algorithms. To pool all subjects together within each group (ASD vs. TC), we plan to average across all subjects for each method and then vote for the top 15 most confident pairs of ROIs across all methods.
  - **Community Detection:** Based on the estimated networks obtained from the previous learning step (s.t., partial correlation), I plan to perform a community detection algorithm (e.g., Louvain, Leiden, etc) to explore the possible communities (i.e., clusters of regions) on a sub-network.
  - **Software:** Intend to use several R graphical packages including *pcalg*, *igraph*, *qgraph*, *huge*, *CliquePercolation*, etc.

References

[1] https://www.cdc.gov/ncbddd/autism/facts.html

[2] Dosenbach, N. U. F., Nardos, B., Cohen, A. L., Fair, D. A., Power, J. D., Church, J. A., Nelson, S. M., Wig, G. S., Vogel, A. C., Lessov-Schlaggar, C. N., Barnes, K. A., Dubis, J. W., Feczko, E., Coalson, R. S., Pruett, J. R., Barch, D. M., Petersen, S. E., Schlaggar, B. L. (2010). Prediction of Individual Brain Maturity Using fMRI. Science, 329(5997), 1358–1361. http://www.jstor.org/stable/41075813

[3] Eslami, T., Mirjalili, V., Fong, A., Laird, A. R., Saeed, F. (2019). ASD-DiagNet: A Hybrid Learning Approach for Detection of Autism Spectrum Disorder Using fMRI Data. Frontiers in neuroinformatics, 13, 70. https://doi.org/10.3389/fninf.2019.00070