

Novel metrics and application of nearest-neighbor feature selection for creating resting-state fMRI brain atlases

Bryan A. Dawkins¹, Trang T. Le², Alejandro A. Hernandez¹, and Brett A. McKinney^{1,3,*}

¹Department of Mathematics, University of Tulsa, Tulsa, OK 74104, USA

²Department of Biostatistics, Epidemiology and Informatics, University of Pennsylvania, Philadelphia, PA 19104

³Tandy School of Computer Science, University of Tulsa, Tulsa, OK 74104, USA.

Abstract

Resting-state functional connectivity MRI (rs-fMRI) data consists of correlation matrices, where correlations are computed between the time series from brain Regions of Interest (ROIs). There are many different parcellations of the human brain into collections of ROIs. These parcellations, or atlases, can be used in case-control studies in order to understand and accurately classify subject phenotypes. We present new metrics for nearest-neighbor distance-based feature selection at the ROI level. Using our new metrics, we apply a novel nearest-neighbor feature selection algorithm to calculate relative importance of ROIs in two existing brain atlases. We use integer programming to derive a mapping between brain atlases to determine spatially similar ROIs. With ROI importance scores and spatial similarity between atlases, we create a new brain parcellation that combines aspects of both brain atlases.

1 Background

Resting-state fMRI data exists in high dimensions and has many sources of noise, such as physiological or motion related [1]. Feature selection is typically done with the purpose of determining brain regions of interest (ROIs) that accurately discriminate between cases and controls in order to understand a particular phenotype. The data consists of pairwise ROI-ROI correlations, where each ROI is a time series measuring brain activity in a particular region or regions of the brain while a subject is not performing a task. A typical data set consists of m subject-specific correlation matrices of dimension $p \times p$, where the pairwise correlations are computed between p ROIs from a brain atlas. Nearest-neighbor distance-based feature selection in rs-fMRI data has been performed using the private evaporative cooling method, which used pairwise ROI-ROI correlations as predictors of a particular phenotype. However, nearest-neighbor feature selection algorithms have not been applied at the ROI level to assess the relative importance of ROIs for a given phenotype. To address this, we have previously proposed a new distance metric that allows us to compute the importance of individual ROIs using a nearest-neighbor distance-based approach. We use this new distance metric with a novel nearest-neighbor feature selection algorithm called Nearest-neighbor Projected Distance Regression (NPDR) in order to compute ROI importance and the corresponding pseudo P values [2]. Our analysis is done on subject rs-fMRI correlation matrices generated by two well known brain atlases [3, 4].

In order to make spatial comparisons between any pair of brain atlases, we first compute a distance matrix containing all pairwise distances between the different collections of atlas ROIs. Distances are defined based on a set dissimilarity metric that accounts for differences in voxel collections between pairs of ROIs. In a particular

coordinate system, voxels have well defined three-dimensional locations in a given brain atlas. As long as two different atlases are in the same coordinate system, we can compare voxel membership between opposing atlas ROIs. We use an integer program that defines the standard Assignment Problem (AP) to find the one-to-one mapping between the two sets of atlas ROIs [5]. The collection of all mapped ROIs gives the closest spatial analogy between the two atlases, which tells us the closest relationship between the two sets of ROIs from different atlases. The collection of unmapped ROIs gives an indication of spatial uniqueness in the two atlases, respectively. All ROIs can be further mapped to a well defined anatomical region of the brain, which allows us to point out potential targets for better understanding the phenotype of interest.

Our spatial mapping between atlases and relative importance scores for ROIs in each respective atlas provides a way to combine relevant and distinct aspects of each brain atlas into a new parcellation. This new atlas includes important ROIs that are in the optimal one-to-one mapping from the solution to the assignment problem and any important unmapped ROIs from each atlas. Spatial overlap and attribute importance can serve as an useful tool for other researchers to compare, contrast, and combine two atlases. In particular, our results show how one might choose either of the two atlases to study the phenotype of interest we are considering in this work.

2 Methods

2.1 Real rs-fMRI data

2.2 Spatial overlap between brain atlases

2.3 Relative importance of ROIs

3 Results

3.1 New brain atlas

4 Discussion

References

1. César Caballero Gaudes and Richard C. Reynolds. Methods for cleaning the BOLD fMRI signal. *NeuroImage*, 154:128–149, December 2017.
2. Trang T. Le, Bryan A. Dawkins, and Brett A. McKinney. Nearest-neighbor Projected-Distance Regression (NPDR) detects network interactions and controls for confounding and multiple testing. *Under Review*, 2019.
3. Jonathan D Power, Alexander L Cohen, Stephen M Nelson, Gagan S Wig, Kelly Anne Barnes, Jessica A Church, Alecia C Vogel, Timothy O Laumann, Fran M Miezin, Bradley L Schlaggar, and Steven E Peterson. Functional network organization of the human brain. *Neuron*, 72(4):665–678, November 2011.
4. X. Shen, F. Tokoglu, X. Papademetris, and R. T. Constable. Groupwise whole-brain parcellation from resting-state fMRI data for network node identification. *Neuroimage*, (0):403–415, November 2013.
5. David W. Pentico. Assignment problems: A golden anniversary survey. *European Journal of Operational Research*, 176:774–793, November 2007.