Graph-Theoretical Methods for Statistical Inference on MR Connectome Data

Joshua T. Vogelstein[1], John Bogovic[1], Aaron Carass[1], William R. Gray[1], Jerry L. Prince[1], Bennett Landman[2], Dzung Pham[1], Luigi Ferrucci[3], Susan M. Resnick[3], Carey E. Priebe[1], and R. Jacob Vogelstein[1,4]

[1]Johns Hopkins University, [2]Vanderbilt University, [3]National Institute on Aging, [4]Johns Hopkins University Applied Physics Laboratory

joshuav@jhu.edu

Abstract

We are developing analytical tools to perform statistical inference on the connectome. Previous work has shown that simple measures of brain connectivity (e.g. total volume of white matter) are correlated with general cognitive functions such as intelligence. Because the connectome can be represented as a large interconnected graph (in which nodes are neuroanatomical regions and synapses are bundles of white matter tracts), we hypothesize that the development of algorithms based on principles of graph theory will allow for greater prediction of performance on measures of specific cognitive functions. To test this hypothesis, we have: (i) collected multimodal MR data from a large cohort of subjects from the Baltimore Longitudinal Study of Aging (BLSA), (ii) developed and applied a high-throughput fully-automated pipeline for extracting brain-graphs from multimodal MR images, (iii) derived asymptotically optimal algorithms for graph classification, and (iv) applied these algorithms to simulations based on the BLSA data set. We show that our data processing pipeline is both efficient and robust. Furthermore, given a relatively small number of subjects, simulated classification accuracy approaches optimality. These results suggest that the developed methods may be useful for unraveling the detailed connectivity underlying many cognitive functions.

1. Background

1.1 Multimodal MR imaging

- T1-weighted images provide high-resolution anatomical data (grey matter, white matter, etc.).
- Diffusion Weighted (DW) images provide high-resolution measurement of white matter tracts [B94].
- Co-registration of T1- and DW-images allows for visualizing both the grey matter in the cortex and the fiber tracts that connect cortical regions, i.e. an MR connectome [H10].

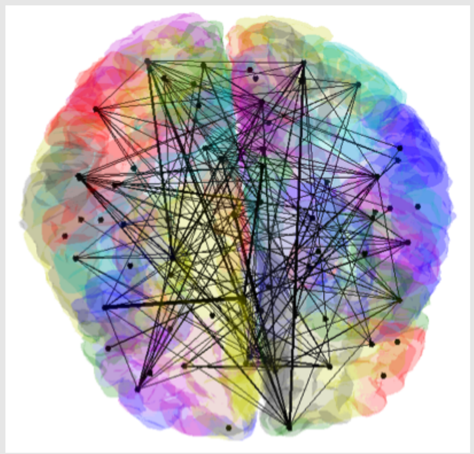
source: philips.com

1.2 Inferring cognitive properties from MR images of the brain

- Summary statistics (e.g., mean FA) from DWI have been used to relate brain connectivity data to mental properties such as intelligence, creativity, etc. [H10].
- Machine learning techniques have been used to classify mental disorders (e.g. schizophrenia) from DTI data [A10].
- Gross measures of changes in connectivity have been shown to correlate with the transition from mild cognitive impairment to Alzheimer's disease [Z08].
- However, to date, no one has developed or applied statistical tools that facilitate classifying brains based on their graph structure.

1.3 The brain-graph: a connectome

- A graph, $G = (V, A)$ is a collection of vertices, V, and an adjacency matrix, A, that describes which vertices are connected.
- For our MR connectomes, vertices correspond to cortical neuroanatomical regions.
- For our MR connectomes, edges correspond to white matter tracts.



1.4 Brain-graph classification

- A random graph describes a distribution of possible graphs, $P_\theta[G = g] \forall g \in \mathcal{G}$. For example, a uniform distribution asserts that all graphs are equally likely.
- Let Y be a binary cognitive property, such as above- or below-average intelligence.
- We aim to build a classifier that, given a brain-graph g, can correctly predict y.
- We define a parametric joint model, $P_\theta[G, Y]$ and build optimal classifiers under the model $P_\theta[G, Y]$:

$$\hat{y} = \underset{y\in\{0,1\}}{\operatorname{argmax}} P_\theta[Y|G] = \underset{y\in\{0,1\}}{\operatorname{argmax}} P_\theta[G|Y]P[Y]$$

where $P[Y]$ is the prior and $P_\theta[G|Y]$ is the likelihood.

- A very simple model is the independent edge model:

$$P_\theta[G|Y] = P_\theta[A|Y] = \prod_{ij} P_\theta[A_{ij}|Y] = \prod_{ij} \text{Bernoulli}(a_{ij}; p_{y;ij})$$

where $a_{ij} = 1$ if there is a connection from vertex i to j, and zero otherwise.

- The parameter, θ, is the collection of $p_{y;ij}$'s, each in $(0, 1)$, corresponding to the probability that $a_{ij} = 1$ when the brain-graph is in class y.

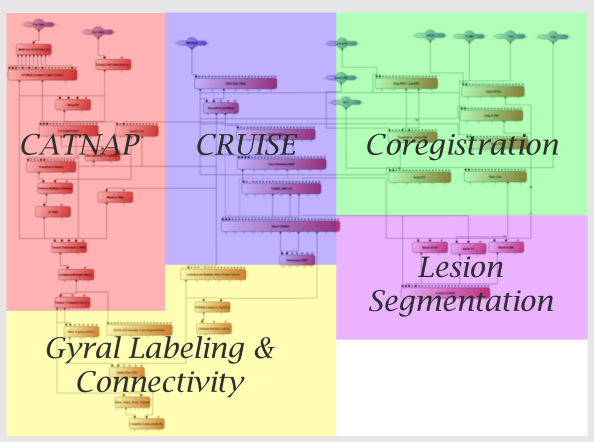
2. Methods

2.1 Data collection

- 32 total subjects, with many cognitive covariates measured in each.
- Multi-slice single-shot EPI (SENSE factor = 2.0), spin echo sequence (TR/TE=3632/100 ms) on 1.5T MR scanner.
- 30 directions, twice per subject, 256^3 voxels.
- Gradient strength of 19.5 mT/m, b-factor of 1000 s/mm^2.

2.2 Extracting brain-graphs from multimodal MRI

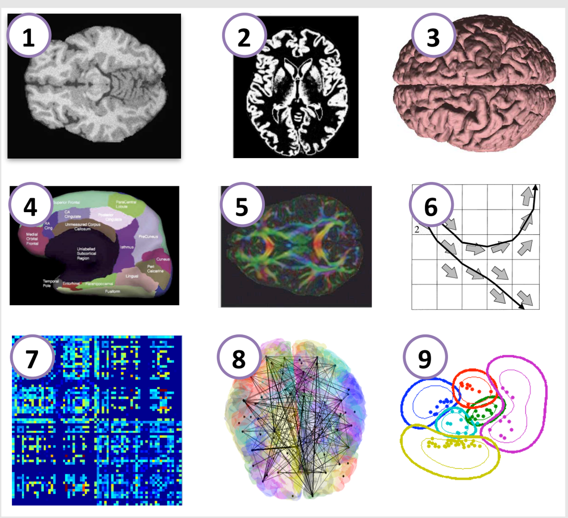
- Constructed JIST layouts to create an automated pipeline that converts T1- and DW-MR images into brain graphs [L10].
- Total processing time typically takes <2 days per brain per node on a computing cluster.
- Given a list of MR datasets, the whole process can be batched.
- Each of the 32 brain-graphs was subjectively validated for correctness.



3. Brain graphs

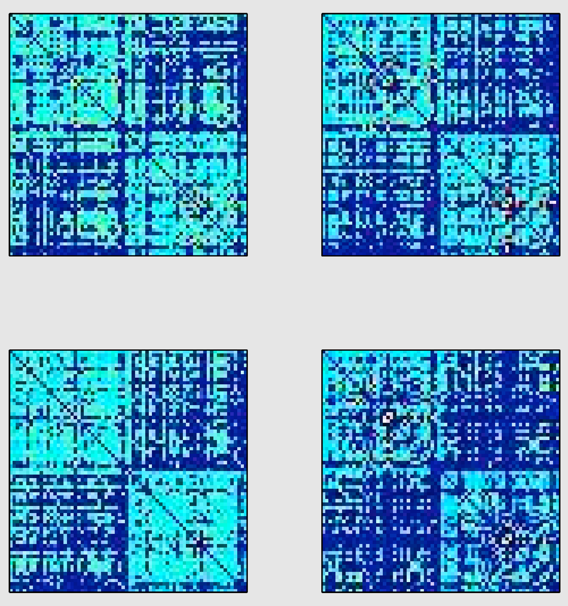
3.1 JIST pipeline and post-processing

1. Multimodal MR imaging.
2. Tissue segmentation (TOADS) [B06].
3. Cortical Surface reconstruction (CRUISE) [H04].
4. Registration, parcellation, and labeling (STAPLE) [W04,D06].
5. Tensor estimation (CATNAP) [L07].
6. Tract tracing (FACT) [B94].
7. Adjacency matrix extraction [L10].
8. Visualization.
9. Graph theoretical analysis.



3.2 Examples

- Adjacency matrices for 4 example brain-graphs shown.
- Each brain-graph has 70 vertices, corresponding to 35 cortical regions per hemisphere.
- Each voxel has a Fractional Anisotropy (FA), so each estimated tract has a mean FA.
- Color in each edge corresponds to mean of the mean FA's of all tracts connecting the two regions (lighter blue is more strongly connected).
- There are more connections within than between hemispheres.

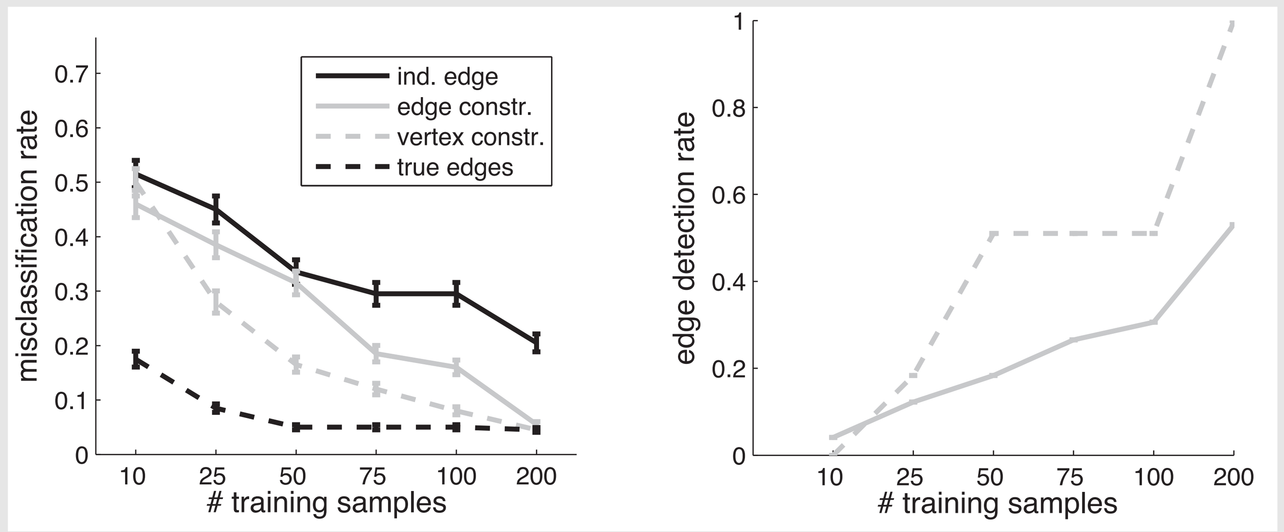


4. Results

4.1 Simulation details

- Two classes of brain-graphs simulated with parameters estimated from actual BLSA data (class 0 = male, class 1 = female).
- For class 0, we set $p_{0;ij} = \langle a_{ij}^k \rangle_{y_k=0}$, where $\langle x_k \rangle_k$ indicates the average of x over k, $a_{ij}^k = 1$ indicates that there is an edge from i to j in subject k, and we only look at subjects in class 0.
- Generated up to 200 training samples and 500 testing samples from each class.
- Training data is used to obtain maximum likelihood estimates (MLE) $\hat{p}_{y;ij} = \langle a_{ij}^k \rangle_{y_k=0,k\in\mathcal{D}_{trn}}$.
- Given these estimates, we can compute the maximum *a posteriori* class of a new brain-graph: $\hat{y} = \operatorname{argmax}_{y\in\{0,1\}} \prod_{i,j} \text{Bernoulli}(a_{ij}; \hat{p}_{y;ij})$.
- If a binary cognitive covariate is only dependent on a small set of edges, we call them the signal dependent edges. Classification performance can improve by only looking at those edges.
- Let $\mathcal{E}$ indicate the signal dependent edges, then an improved classifier is: $\hat{y} = \operatorname{argmax}_{y\in\{0,1\}} \prod_{(i,j)\in\mathcal{E}} \text{Bernoulli}(a_{ij}; \hat{p}_{y;ij})$
- If all the signal dependent edges are between a small set of vertices, $\mathcal{V}$, then a further improved classifier is: $\hat{y} = \operatorname{argmax}_{y\in\{0,1\}} \prod_{i\&j\in\mathcal{V}} \text{Bernoulli}(a_{ij}; \hat{p}_{y;ij})$.

4.2 Classification results



- Naive bayes classifier uses independent edge assumption.
- Constraining the classifier to use only important edges improves performance.
- Constraining the classifier to use only important vertices further improves performance, by utilizing graph structure.
- Both constrained classifiers achieve optimal performance after only 200 samples.
- The fraction of correctly identified signal edges increases with # of training samples for both constrained approaches.
- The vertex constraint more quickly finds correct edges by utilizing graph structure.

5. Discussion

5.1 Summary

- We can rapidly process data using our fully automated pipeline for extracting brain-graphs from multimodal MR data using TOADS, CRUISE, CATNAP and JIST.
- We derived classifiers that were optimal under some relatively restrictive assumptions.
- Simulations suggest that these classifiers perform as expected.

5.2 Next steps

- Utilize HARDI data and probabilistic tractography to extract brain-graphs from MR data.
- Run analysis on real data.

References

[A10] BA Ardenaki et al. *Diffusion tensor imaging reliably differentiates patients with schizophrenia from healthy volunteers.* Human Brain Mapping, 2010.

[B94] P Basser, et al. *Diffusion tensor imaging reliably differentiates patients with schizophrenia from healthy volunteers.* Biophysical Journal, 1994.

[B06] P Bazin & D Pham. *TOADS: topology-preserving, anatomy-driven segmentation.* Biomedical Imaging: Macro to Nano, 2006.

[B10] J Bogovic, et al. *Multi-modal structural networks: mapping of connectivity through diffusion, functional, and structural assessment of intervening pathways.* ISMRM, 2010.

[D06] RS Desikan, et al. *An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest.* NeuroImage, 2006.

[H10] P Hagmann, et al. *MR connectomics: Principles and challenges.* J Neuroscience Methods, 2010.

[H04] X. Han, et. al. *CRUISE: Cortical Reconstruction Using Implicit Surface Evolution.* NeuroImage, 2004.

[L07] B Landman, et. al. *Effects of diffusion weighting schemes on the reproducibility of DTI-derived fractional anisotropy, mean diffusivity, and principal eigenvector measurements at 1.5T.* Neuroimage, 2007.

[L10] BC Lucas, et al. *The Java Image Science Toolkit (JIST) for rapid prototyping and publishing of neuroimaging software.* Neuroinformatics, 2010.

[W04] SK Warfield, et al. *Simultaneous truth and performance level estimation (STAPLE): an algorithm for the validation of image segmentation.* IEEE Trans Med Imaging, 2004.

[Z08] Y Zhao, et al. *Abnormal connectivity in the posteroir cingulate and hippocampus in early Alzheimer's disease and mild cognitive impairment.* Alzheimer's & dimentia, 2008.

This work was supported by NIH/NINDS 1R01NS056307, NIH/NINDS 5R01NS054255, and the NSA Research Program on Applied Neuroscience.