OPTIMIZING BRAIN CONNECTIVITY NETWORKS FOR DISEASE CLASSIFICATION USING EPIC

Gautam Prasad[1], **Shantanu H. Joshi**[2], and **Paul M. Thompson**[1]

[1]Imaging Genetics Center, Inst. for Neuroimaging and Informatics, Keck Sch. of Med. of USC, Los Angeles, CA, USA

[2]Ahmanson-Lovelace Brain Mapping Center, Department of Neurology, UCLA, Los Angeles, CA, USA

Abstract

We propose a method to adaptively select an optimal cortical segmentation for brain connectivity analysis that maximizes feature-based disease classification performance. In standard structural connectivity analysis, the cortex is typically subdivided (parcellated) into N anatomical regions. White matter fiber pathways from tractography are used to compute an $N \times N$ matrix, which represents the pairwise connectivity between those regions. We optimize this representation by sampling over the space of possible region combinations and represent each configuration as a set partition of the N anatomical regions. Each partition is assigned a score using accuracy from a support vector machine (SVM) classifier of connectivity matrices in a group of patients and controls. We then define a high-dimensional optimization problem using simulated annealing to identify an optimal partition for maximum classification accuracy. We evaluate the results separately on test data using cross-validation. Specifically, we demonstrate results on the ADNI-2 dataset, where we optimally parcellate the cortex to yield an 85% classification accuracy using connectivity information alone. We refer to our method as evolving partitions to improve connectomics (EPIC).

Index Terms

Cortical parcellation; connectivity matrix; classification; simulated annealing; partition

1. INTRODUCTION

Connectomics [1] – or the study of brain connectivity – has become popular in recent years, especially with advances in diffusion imaging and resting-state functional magnetic resonance imaging (rsfMRI), which reveal neural pathways and functional synchronization between pairs of brain regions. Brain connectivity is often characterized by determining connections among a set of N brain regions; usually the chosen regions are in the cortex. In a standard analysis of structural connectivity, tractography is applied to diffusion-weighted MRI data to extract fibers throughout the brain, and the density, number, or integrity of connections between all pairs of cortical regions can be represented as an $N \times N$ connectivity matrix for each subject in the study [2]. This representation of connectivity has been used to further our understanding of aging [3], brain development, left/right hemisphere differences

in connectivity, various diseases, psychiatric disorders, and even genetic variants associated with brain connectivity [4].

Network connectivity has largely been defined using a bottom-up approach, where one makes assumptions on the configuration of nodes, their properties, and the underlying covariance structure of their interconnections. For example, a structural connectivity network can be created by defining a set of N regions in an anatomical image [1, 5]. Alternately, a functional network may be defined based on a specific set of nodes belonging to functionally active regions in the gray matter. The choice of regions in the network may be based upon regions likely to be activated in specific cognitive tasks [6, 7] or they can be based on task-free (resting-state) oscillations of the blood oxygenation level-dependent (BOLD) signal [8].

Recently, departing from the conventional structural or functional connectivity paradigms, researchers have proposed several choices for refining network nodes in a brain connectivity analysis, including ones based on a cortical parcellation or partition, which subdivides the entire cortical surface into a set of non-overlapping regions, or patches, that jointly cover it. In [9], spectral clustering was used to compute a cortical parcellation based on functional connectivity. They demonstrated better ROI homogeneity with their new parcellation scheme, and showed better reproducibility of function connectivity when compared to anatomical atlases. However, their approach is biased to a configuration with equal sized regions. A combination of region growing and hierarchical clustering was used by [10], where coherent boundaries for functional connectivity were created. Their method relies on a set of stable seeds to generate and grow their regions. Tzourio-Mazoyer et al. defined a neurobiologically-informed cortical parcellation based on regions of interest that are known to house specific functional areas in the brain [11]. By contrast, Zalesky et al. [12] proposed a more exhaustive approach that treats each voxel as its own ROI, resulting in tens of thousands of ROIs on the cortex. An intermediate approach proposed by Wig [13] treats the elements of random parcellations of the cortex as nodes, but this approach may still fail to capture the borders of regions that make sense for capturing pathways.

Clearly, one could start by aggregating or clustering fibers into sets that have similar trajectories, and some clustering methods treat fibers as high-dimensional vectors and group them. Even so, if fiber sets were clustered into a set of bundles, the boundaries of the cortical regions they connect might not be easily inferred from the available data - the target regions for different bundles may be interleaved or overlap. Also, the meaning of connectivity may depend on the scale of the parcellation. An anatomically meaningful parcellation might cluster or bundle fibers with similar geometries or trajectories, but from a mathematical point of view there may be even better partitions that best address specific problems. It may be possible to define connectivity in a way that enhances our ability to differentiate patients from controls, predict decline, or identify genetic effects on brain connectivity.

In this paper, we formulate a top-down data-driven approach, and ask the question: what kinds of cortical parcellations lead to better detection of disease effects on brain connectivity? The answer depends upon the statistical method used to detect significant

effects of disease, and the disease itself, as different diseases may affect some connections more than others. Fronto-temporal dementia, for example, preferentially erodes connectivity of frontal and temporal lobes to the rest of the brain, whereas Alzheimer's disease typically has a more dynamic trajectory with connections lost first in temporal, limbic, and parietal association areas, and then in frontal and eventually primary cortical areas. Arguably, it makes sense to adopt a data-driven approach to estimate the optimal parcellation corresponding to a given disease effect that we intend to detect. We define the criteria to optimize the nodes in the connectivity network, by favoring a parcellation that improves the classification rate in discriminating disease. By coupling the network definition and the problem we aim to solve, our adaptive method optimizes the nodes in the network for the problem at hand. The type of networks that we study here are similar to the inference-based networks [14] that rely on the effective connectivity derived from the statistical dependencies of various nodes (regions) on each other, but whose nodes are optimally chosen to improve the classification performance of brain disease.

2. METHODS

2.1 Subject Data

Our data was collected from 87 subjects scanned as part of the ADNI-2 [15] project in which diffusion imaging was added to the standard MRI protocol. The dataset included 50 normal controls and 37 Alzheimer's disease (AD) patients. Subjects were scanned on 3-Tesla GE Medical Systems scanners, which acquired both T1-weighted 3D anatomical spoiled gradient echo (SPGR) image volumes, and diffusion-weighted images (DWI; 256×256 matrix; voxel size: $2.7 \times 2.7 \times 2.7$ mm^3; scan time = 9 min). For each subject, the DWI consisted of 41 diffusion-weighted images with $b = 1000$ s/mm^2 and 5 T2-weighted b_0 images. The ADNI data would not be considered high angular resolution, but the protocol was optimized to avoid long scan times.

2.2 Connectivity Matrices

The T1-weighted images were processed as follows. We first removed extra-cerebral tissues from the anatomical images, corrected the images for inhomogeneity, aligned them to the Colin27 template with FSL Flirt [16], and used FreeSurfer [17] for cortical extraction and labeling. The cortical segmentation was dilated with a 5 mm isotropic box kernel to ensure intersection with the white matter. For the diffusion-weighted images (DWI), we corrected for head motion and eddy current distortion in each subject by aligning the images to the average b_0 image with FSL's eddy correct tool. The images were EPI corrected with an elastic registration algorithm using mutual information that aligned the DWI images to the T1-weighted scans. We used an optimized global probabilistic tractography method based on the Hough transform [18] to generate 35,000 fibers per subject.

ALGORITHM I

Optimization of Partitions

	Inputs: *connectivity_matrices*, *labels*
1	*current_partition* = AnatomicPartition()
2	*pca_matrices* = PCA(*connectivity_matrices*, *current_partition*)
3	*current_accuracy* = LOOSVM(*pca_matrices*, *labels*)
4	*best_partition* = *current_partition*
5	*best_accuracy* = *current_accuracy*
6	**for** *i* **in** 1 **to** *max_step* **do**
7	$T = (max\_step - i + 1)/max\_step$
8	*new_partition* = GeneratePartition(*current_partition*, *T*)
9	*pca_matrices* = PCA(*connectivity_matrices*, *new_partition*)
10	*new_accuracy* = LOOSVM(*pca_matrices*, *labels*)
11	**if** P(*new_accuracy*, *current_accuracy*, *T*) > random() **then**
12	*current_partition* = *new_partition*
13	*current_accuracy* = *new_accuracy*
14	**if** *new_accuracy* > *best_accuracy* **then**
15	*best_partition* = *new_partition*
16	*best_accuracy* = *new_accuracy*

Output: *best_partition*, *best_accuracy*

Pseudocode for our simulated annealing setup to find the optimal connectivity matrix partition that maximizes the accuracy of leave-one-out cross-validated support vector machine (LOOSVM) classification of control subjects vs. Alzheimer's disease. The algorithm is initialized to the anatomical parcellation from Freesurfer and uses principal components analysis (PCA) to reduce the matrices dimension for efficient classification. The input is a set of connectivity matrices based on the anatomical segmentation and labels distinguishing the two classes and the output is the best partition and its corresponding accuracy. This optimization was on the training data for each fold and the best partition applied to the test subject.

We generated connectivity matrices for the subject by combining the fibers and dilated cortical segmentation for each subject. For all pairs of cortical regions, we find the number of fibers that connect them to create an $N \times N$ connectivity matrix (Fig. 2) where N=68 for the 34 anatomical regions segmented in each hemisphere by Freesurfer. We normalized the matrix by the number of fibers. The nodes of this matrix represent the basic building blocks of the connectivity network and our method tries to combine subsets of these regions into *super-regions* to find a better representation of brain connectivity.

2.3 Partition Representation

We represent a possible combination of the anatomical regions as a set partition of the set of N elements. A partition of a set composed of N regions is a set of nonempty subsets such that each element is in only one of the subsets. This definition means a partition can specify one to N super-regions.

Our definition of distance between partitions is borrowed from an application in computational genetics [19]. The distance between two partitions, *a* and *b*, is defined as the number of elements in *a* that need to be re-assigned to other super-regions to make it

equivalent to *b*. This convention allows us to specify our search range when computing an optimal partition.

The total number of partitions of a set of *N*-elements is the Bell number [20]. When *N*=68, the Bell number is 3.66×10^{70} making an exhaustive search of the space for the best partition for connectivity based classification infeasible. Instead we opt for an approximate search using simulated annealing to generate samples from this space.

2.4 Dimensionality Reduction

We used principal components analysis (PCA) [21] to reduce the dimensionality of the connectivity matrices specified by a partition. PCA uses orthogonal transformation to generate an orthogonal basis to span the data space. We projected our data onto a subset of these basis vectors that accounted for greater than 95% of variance from our data. For example, this projection reduces the number of points for a single subject from 4624 to around 20 points in the case of a 68×68 connectivity matrix and allows more efficient classification.

2.5 Classification

We use support vector machines (SVMs) [22] to classify the dimensionality reduced connectivity matrices when comparing the groups of healthy controls and patients with Alzheimer's disease. An SVM is a supervised machine-learning algorithm that classifies two-class data by training (tuning the free parameters of a classification function) to find the best hyperplane between the two classes. Our classification design tests the information provided by the connectivity configuration in a leave-one-out cross-validation setting. We evaluate the results based on accuracy, sensitivity, and specificity.

In our design, this cross-validation is used to find the optimal partition based on training data and is nested within another leave-one-out cross-validation that evaluates the partition on separate, disjoint test dataset.

2.6 Simulated Annealing

To find the optimal partition efficiently in such a large search space we used a generic probabilistic metaheuristic to approximate a global maximum called simulated annealing [23]. Simulated annealing is inspired by annealing from metallurgy and works by using a temperature parameter to begin a search that encompasses a large range of values and is slowly constrained as the temperature is lowered. We provide the pseudocode for finding the partition that maximizes classification accuracy using the simulated annealing algorithm in Algorithm 1.

The procedure begins with the basic anatomical connectivity matrices and evaluates their ability to classify controls vs. AD based on accuracy. This is used as the first state in the simulated annealing chain. The algorithm carries out a set number of iterations (*max_step*, set to 1000 in our experiments) where a new partition is generated and evaluated to see if it has better performance. We generate the new partition by randomly reassigning *k* anatomical regions to super-regions, where *k* is proportional to the current temperature, *T*, which

decreases proportional to the iteration count. We use the new partition to generate the corresponding connectivity matrix configuration for each subject, reduce its dimension using PCA, and evaluate its performance with leave-one-out cross-validated SVM (LOOSVM). If the new partition has a higher accuracy it is chosen as the next state, but if it is a lower accuracy it is selected based on the acceptance probability [23]

$$\frac{1}{1+e^{\Delta/T}}, \quad (1)$$

where is the difference between the current and new partition accuracies. This allows the algorithm to search a greater range by moving to points that may not improve the accuracy, but could help avoid getting trapped in a local maximum.

Throughout the algorithm we keep track of the best partition and its corresponding accuracy, and both are retained as the final output.

2.7 Experiments

We compared the accuracy, sensitivity, and specificity from classification of control subjects vs. AD patients based on their connectivity matrices derived from the anatomical segmentation from Freesurfer and the best partition from our method. Our design used leave-one-out cross-validation with PCA for dimensionality reduction on the subject matrices to represent them efficiently.

To compute statistics for each approach, we used two cross-validation runs. In the first, we used the connectivity matrices based on the 68 Freesurfer regions to represent each subject and learned an SVM based on the training subjects. We tested the resulting SVM on the left out subjects and repeated this for each fold. Our second run involved running our simulated annealing optimization on the training subjects to learn the best partition and best connectivity matrices to represent connectivity. We then tested the resulting connectivity matrix configuration on the left out subjects and repeated for each fold.

3. RESULTS

In Fig. 2 we show a sample connectivity matrix created using the Freesurfer anatomical parcellation and a visualization of the cortical regions. Its data is from an AD patient and creates a 68×68 connectivity matrix where the regions are randomly colored and displayed on the right side.

An example matrix computed from our method based on training data in one of the folds from our cross-validation setup is shown in Fig. 3. This is also from an AD subject and is 36×36 in dimension. On the right side of the figure we show the 36 super-regions that are created by grouping the original 68 regions into larger areas of the cortex.

The results of the leave-one-out cross-validated classification using the full 68×68 anatomical parcellation were 82.7% accuracy, 78.3% specificity, and 86% sensitivity. Our

method produced an optimal cortical parcellation at each fold during classification and gave results of 85.0% accuracy, 81.0% specificity, and 88% sensitivity.

4. DISCUSSION

The cross-validation results showed that our method was able to improve the accuracy, specificity, and sensitivity of the classification by choosing a different cortical segmentation and nodes for representing connectivity. As shown in Fig. 3, the configuration from our method reduced the number of distinct regions on the cortex and may emphasize cortical regions that are critical to brain connectivity in Alzheimer's disease. In contrast to the approach by [24], where features of grey matter integrity on the cortical surface are used exclusively for discriminatory purposes, our method uses classification of disease as the energy function to search for the best cortical parcellation.

Our method was able to find a segmentation of the cortex that improved the accuracy of the classification. We were able to represent, quantify, and search the space of partitions based on their classification accuracy. In contrast to the approach by Zalesky et al. [25], which focused on node selection based on both node resolution and parcellation for connectivity analysis, our method optimizes an objective function on the space of parcellations and nodes. One could potentially use different methods to select nodes: for example, one could allow a diffeomorphic flow of the cortical parcellation, and use the iterative flow to optimize the classification performance. We could also consider starting with the anatomic parcellation, and use the merge-split operation to improve the performance.

Future work could explore using different cooling schedules and acceptance probabilities or using training data to learn possible regimes that are closer to the distribution of the underlying data. Other connectivity measures such as functional connectivity could be used to determine the changes in the nodes of connectivity matrix based on multi-modal information. Here, we chose the classification rate as our objective function, but the framework allows other measures to be used such as the detection of genetic effects, age effects, sex or hemispheric differences, or changes in development.

References

1. Hagmann P, et al. Mapping the structural core of human cerebral cortex. PLoS Biology. 2008; 6(7): 1479–1493.
2. Sporns O, Chialvo DR, Kaiser M, Hilgetag CC. Organization, development and function of complex brain networks. Trends in Cognitive Sciences. 2004; 8(9):418–425. [PubMed: 15350243]
3. Meunier D, Achard S, Morcom A, Bullmore E. Age-related changes in modular organization of human brain functional networks. Neuro Image. 2009; 44(3):715–723. [PubMed: 19027073]
4. Jahanshad N, et al. Genome-wide scan of healthy human connectome discovers SPON1 gene variant influencing dementia severity. Proceedings of the National Academy of Sciences. 2013; 110(12): 4768–4773.
5. Gong G, et al. Mapping anatomical connectivity patterns of human cerebral cortex using in vivo diffusion tensor imaging tractography. Cerebral Cortex. 2009; 19(3):524–536. [PubMed: 18567609]
6. Salvador R, Suckling J, Coleman MR, Pickard JD, Menon D, Bullmore ED. Neurophysiological architecture of functional magnetic resonance images of human brain. Cerebral Cortex. 2005; 15(9): 1332–1342. [PubMed: 15635061]

7. Achard S, Salvador R, Whitcher B, Suckling J, Bullmore ED. A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs. Journal of Neuroscience. 2006; 26(1):63–72. [PubMed: 16399673]

8. Damoiseaux JS, et al. Consistent resting-state networks across healthy subjects. Proceedings of the National Academy of Sciences. 2006; 103(37):13848–13853.

9. Craddock RC, James GA, Holtzheimer PE, Hu XP, Mayberg HS. A whole brain fMRI atlas generated via spatially constrained spectral clustering. Human Brain Mapping. 2012; 33(8):1914–1928. [PubMed: 21769991]

10. Blumensath T, et al. Spatially constrained hierarchical parcellation of the brain with resting-state fMRI. Neuro Image. 2013

11. Tzourio-Mazoyer N, et al. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuro Image. 2009; 15(1):273–289. [PubMed: 11771995]

12. Zalesky A, Fornito A, Egan GF, Pantelis C, Bullmore ET. The relationship between regional and inter–regional functional connectivity deficits in schizophrenia. Human Brain Mapping. 2012; 33(11):2535–2549. [PubMed: 21922601]

13. Wig GS, Schlaggar BL, Petersen SE. Concepts and principles in the analysis of brain networks. Annals of the New York Academy of Sciences. 2011; 1224(1):126–146. [PubMed: 21486299]

14. He Y, Chen ZJ, Evans AC. Small-world anatomical networks in the human brain revealed by cortical thickness from MRI. Cerebral Cortex. 2007; 17(10):2407–2419. [PubMed: 17204824]

15. Trojanowski JQ, et al. Update on the biomarker core of the Alzheimer's Disease Neuroimaging Initiative subjects. Alzheimer's & Dementia. 2010; 6(3):230–238.

16. Jenkinson M, Bannister P, Brady M, Smith S. Improved optimization for the robust and accurate linear registration and motion correction of brain images. Neuro Image. 2002; 17(2):825–841. [PubMed: 12377157]

17. Fischl B, et al. Automatically parcellating the human cerebral cortex. Cerebral Cortex. 2004; 14(1): 11–22. [PubMed: 14654453]

18. Prasad, G.; Nir, TM.; Toga, AW.; Thompson, PM. Tractography density and network measures in Alzheimer's disease. IEEE 10th International Symposium on Biomedical Imaging; 2013. p. 692-695.

19. Gusfield D. Partition-distance: A problem and class of perfect graphs arising in clustering. Information Processing Letters. 2002; 82(3):159–164.

20. Bell, E. American Mathematical Monthly. 1934. Exponential numbers; p. 411-419.

21. Pearson K. On lines and planes of closest fit to systems of points in space. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science. 1901; 2(11):559–572.

22. Cortes C, Vapnik V. Support-vector networks. Machine Learning. 1995; 20(3):273–297.

23. Kirkpatrick S, Gelatt D Jr, Vecchi MP. Optimization by simulated annealing. Science. 1983; 220(4598):671–680. [PubMed: 17813860]

24. Nagy Z, Alexander DC, Thomas DL, Weiskopf N, Sereno MI. Using High Angular Resolution Diffusion Imaging Data to Discriminate Cortical Regions. PloS One. 2013; 8(5)

25. Zalesky A, et al. Whole-brain anatomical networks: does the choice of nodes matter? Neuro Image. 2010; 50(3):970–983. [PubMed: 20035887]
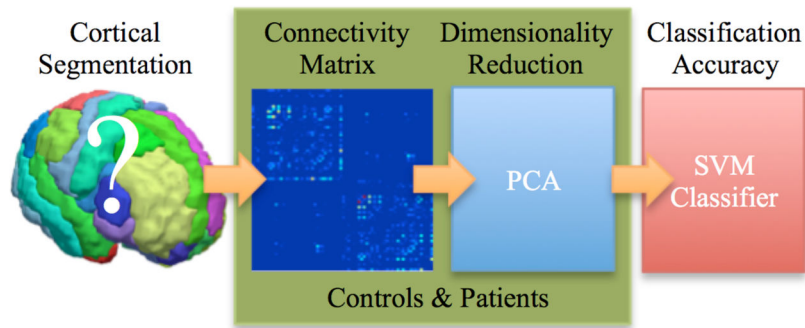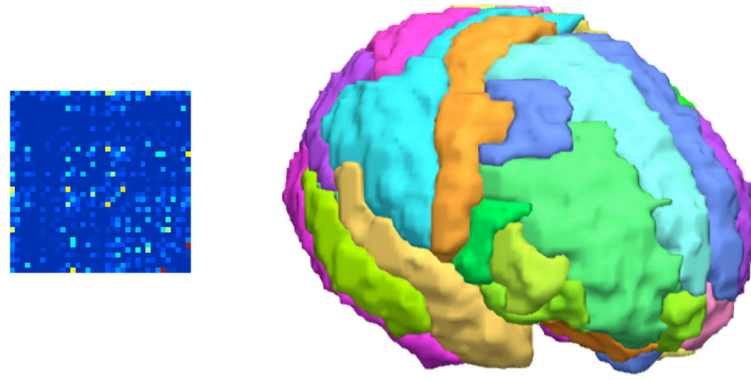
Figure 1.
An overview of our framework to evaluate the brain connectivity represented by a cortical segmentation in a cohort of controls and patients. For an arbitrary segmentation we create a connectivity matrix of the fiber connectivity between regions. We reduce the dimension of the fiber connectivity using principal components analysis (PCA) and use the features in support vector machine (SVM) classification between the two groups. The resulting accuracy is an assessment of the quality of the segmentation.
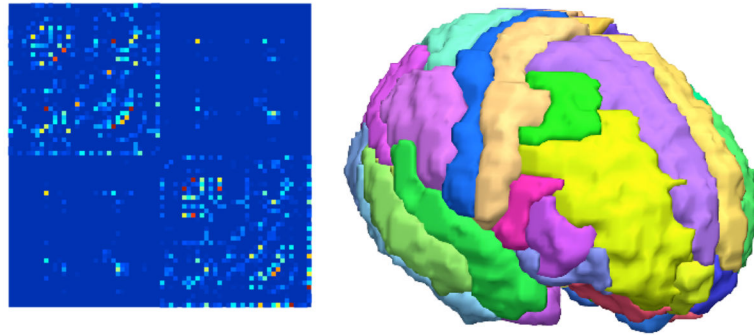
Figure 2.
The 68 × 68 connectivity matrix (left) where the rows and columns correspond to the 68 (34 in each hemisphere) anatomical regions (right) segmented with Freesurfer. Together the regions provide a base set of nodes as the building blocks of our algorithm that seeks to find the configuration of these nodes into groups of regions or super-regions that best represents brain connectivity for classifying Alzheimer's disease vs. controls.

Figure 3.
This 36×36 connectivity matrix (left) represents the optimal partition computed by our algorithm from the training subject for a single leave-one-out cross-validated support vector machine (SVM) classification fold. It is the result from the simulated annealing process for the best combination of anatomical regions for our sample. In this case it was 36 super-regions (right) that is able to represent brain connectivity in a way to maximize the accuracy of classification.