STRUCTURAL CONNECTOME VALIDATION USING FINGERPRINTING

Dmitry Petrov[⋆†] *Boris Gutman*[†] *Alexander Ivanov*[⋆] *Joshua Faskowitz*[†]
Neda Jahanshad[†] *Mikhail Belyaev*[‡⋆] *Paul Thompson*[†]

[⋆]Kharkevich Institute for Information Transmission Problems, Moscow, Russia
[†]Imaging Genetics Center, University of Southern California, Los Angeles, USA
[‡]Skolkovo Institute of Science and Technology, Moscow, Russia

ABSTRACT

In this work, we study the extent to which structural connectomes and topological derivative measures are unique to individual changes within human brains. To do so, we classify structural connectome pairs from two large longitudinal datasets as either belonging to the same individual or not. We call this modification of the traditional fingerprinting problem "pairwise classification." Our data is comprised of 227 individuals from the Alzheimers Disease Neuroimaging Initiative (ADNI) and 226 from the Parkinson's Progression Markers Initiative (PPMI). We achieve 0.99 area under the ROC curve score for features which represent either weights or network structure of the connectomes (node degrees, PageRank and local efficiency). Our approach will be a useful tool for finding meaningful brain features to be used in brain aging studies and perhaps early diagnosis classification problems.

Index Terms— machine learning, fingerprinting, DWI, structural connectomes

1. INTRODUCTION

Predictive modeling of neurodegenerative diseases using diffusion MR-based structural connectomes has become a popular sub-genre of neuroimaging [1]. The great variety of possible pre-processing approaches needed for connectome construction leads to potential challenges in downstream application of the connectomes, for example, in a classification task. Choices of e.g, non-linear registration, parcellation, or tractography, may all have a substantial impact ([2], [3]). This state of affairs presents a challenge both in terms of intrinsic connectome reliability, and the degree to which the recovered connectomes are valid, if summary, representations of true brain connectivity ([4], [5], [6], [7]).

At the same time, the performance of a particular case-control classifier may not suffice as a means of data verification due to small samples and high dimensionality. Alternative, more objective validation may be needed, such as the frequently used Intra-class Correlation Coefficient (ICC) on test-retest data ([8], [9]). But structural connectome ICC is generally low which complicates comparison of methods. Also, the parametric constraints of classic ICC [10] may not be valid and non-parameteric approachs with no assumptions on data distribution, would be more suitable.

Here, we address this issue and propose a modified fingerprinting, or "pairwise classification" approach as a way to intrinsically assess the connectome utility across time (similar attempt was recently made for functional connectomes as well [11]). For each set of connectomes C_j^i and features $f(C_j^i)$ in question, we construct all possible pairs $(C_{j_1}^{i_1}, C_{j_2}^{i_2})$, where i-indices correspond to images and j-indices correspond to subjects. We then conduct a linear classification on the pairwise differences $\|f(C_{j_1}^{i_1}) - f(C_{j_2}^{i_2})\|$ of these pairs with respect to the target variable y: $y = 1$ if $j_1 = j_2$, 0 else.

We test this pipeline on structural connectomes derived from two publicly available neuroimaging datasets: ADNI and PPMI. These datasets have scanned subjects multiple times, with at least a one year interval between pairs of scans. We achieve 0.99 ROC AUC both for features representing structure (PageRank) and weights of connectomes itself (bag of edges), suggesting that the tested data is reliable enough to distinguish subjects by the proposed approach.

It is worth noting that the fingerprinting we are conducting here for particular features does not guarantee these features will perform well for other classification tasks. For example, our target feature perfectly distinguishes classes but is incapable of distinguishing individuals. Our results suggest that some basic mappings of DWI data into feature spaces do not necessarily "glue" subjects together, and this provides some hope that we can meaningfully interpret changes within them.

2. CONNECTOME-BASED FINGERPRINTING

We propose the following pipeline for pairwise connectome classification: normalization, building connectome features, building pairwise features based on connectome features. Let's denote a set of connectomes as $\{C_j^i\}$, where j is an index of a subject and i is an index of an image .

For questions and discussion please contact to.dmitry.petrov@gmail.com

2.1. Normalizations

Topological normalization of connectivity matrices may be useful prior to any analysis, because the number of detected streamlines is known to vary from individual to individual and can also be affected by fiber tract length, volume of cortical regions and other factors ([12], [13]). There is no consensus on the best normalization approach, so we used the three following topological normalization schemes alongside with pure weights (no normalization at all) – by mean, by maximum and binary normalization with zero threshold: $a^b_{kl} = 1$ if $a_{kl} > 0, 0$ else where a_{kl} is a connectome edge.

2.2. Network features

For each connectome and each normalization we built 'bag of edges' vectors from the upper triangle of the symmetric connectivity matrix. In addition, we calculated eight network metrics for each node: weighted degrees, or strength; closeness, betweenness and eigenvector centralities; local efficiency; clustering coefficient; weighted number of triangles around node. We chose these features because they are well-described and reflect different structural properties of connectomes [14]. We also calculated PageRank for each node. Introduced in 1998 by Brin and Page [15] this metric roughly estimates probability that a person randomly clicking on links in the network will arrive at particular node.

2.3. Pairwise features

Each normalization and each set of features described above defines a mapping from connectome space to feature space $C \to f(C)$. Since our goal is to check how well this mapping separates connectomes in our new feature space, we propose various pairwise features. For each set of connectome features in question we make all possible pairs of connectome features – $(f(C^{i_1}_{j_1}), f(C^{i_2}_{j_2}))$. For each pair, we assign a binary target variable – 1 if connectomes are from the same subject ($j_1 = j_2$), 0 – if they are from different subjects ($j_1 \neq j_2$). Finally, for each pair we build a vector of three features, describing their difference $\|f(C_1) - f(C_2)\|$ according to l_1, l_2 and l_∞ norms.

2.4. Classifiers and validation

We use linear classifiers for pairwise classification: logistic regression (LR), SVM with linear kernel and stochastic gradient descent with modified Huber loss. For the linear classification, we scaled features with standard scaling and applied elastic-net regularization.

Performance of our models was measured with area under ROC curve (ROC AUC) through a two-step validation procedure. First, for each dataset, we perform hyperparameter grid search based on a 10-fold cross-validation with a fixed random state for reproducibility. Then for each dataset we evaluate the best parameters for each of the models on 100 train/test splits with fixed different random states (test size was set to 20% of data). We report the ROC AUC distribution on these 100 test splits for each combination of normalization/base features/diagnostic group in results section.

3. EXPERIMENTS

3.1. Base data

We used two datasets for our experiments. First, subjects from the second phase of the Alzheimers Disease Neuroimaging Initiative, which includes DWI scans (ADNI2). We excluded three individuals for which our preprocessing pipeline yielded connectomes with at least one zero node. This data comprises a total of 227 individuals (675 scans), mean age at baseline visit 73.1 ± 7.4, 99 females. Each individual has at least 1 brain scan and at most 6 scans. The data include 46 people with AD (111 AD scans), 80 individuals with EMCI (247 MCI scans), 40 people with LMCI (120 LMCI scans) and 61 healthy participants (160 scans). Second, we used imaging data from the Parkinsons Progression Markers Initiative (PPMI) database. From it we selected used subjects from PD (159 subjects) and healthy controls (67 controls) cohorts for whom DWI data was available. These included a total of 226 individuals (456 scans), with their mean age at the baseline visit being 61.0 ± 9.8 years, 79 were females. Each individual has at least 1 brain scan and at most 4 scans.

3.2. Network construction

Inhomogeneity corrected T1-weighted images for ADNI and PPMI data were processed with Freesufer's [16] recon-all pipeline to obtain a triangular mesh of the gray-white matter boundary registered to a shared spherical space, as well as corresponding vertex labels per subject for a cortical parcellation based on the Desikan-Killiany (DK) atlas[17]. This atlas includes 68 cortical brain regions; hence, our cortical connectivity matrices were 68×68.

In parallel, T1w images were aligned (6-dof) to the 2mm isotropic MNI template. These were used as the template to register the average b0 of the DWI images, in order to account for EPI related susceptability artifacts. DWI images were also corrected for Eddy current and motion related distortions. Tractography for ADNI data was then conducted using the distortion corrected DWI in 2mm isotropic MNI 152 space. Probabilistic streamline tractography was performed using Dipys [18] LocalTracking module and implementation of constrained spherical deconvolution (CSD) [19] with a spherical harmonics order of 6. Tracts longer than 5mm with both ends intersecting the cortical surface were retained, as per Dipys ACT. Edge weights in the original matrices are proportional to the number of streamlines detected by the algorithm.

PPMI data were processed in a slightly different fashion to account for variability in the acquisition protocols and to show our method is not dependent on any single processing scheme. Images were initially denoised with an adaptive denoising algorithm [20] and multiple DWI acquisitions from each subject were merged. DWI images were corrected for Eddy current and motion related distortions, then non-linearly epi-corrected with ANTs SyN. Tractography for PPMI data was then conducted in 2mm isotropic MNI 152 space, again using Dipy's LocalTracking module. At each voxel, the CSD was fitted recursively [19] with a spherical harmonic order of 6. Deterministic streamline tractography was seeded at two random locations in each white matter voxel. Similar to the ADNI data, tracts longer than 5mm with both ends intersecting the cortical surface were retained.

3.3. Pairwise data

For each set of connectomes described above (ADNI, PPMI) we made all possible pairs of connectomes as described in 2.3. Using this technique we obtained 227475 pairs (764 of which were labeled as 0) from ADNI2 data and 152031 pairs from PPMI data (301 of which were labeled as 1). Due to huge imbalance of classes in generated pairs, we used all samples with label 1 and equally sized random subsample of 0. Our result do not depend for different subsamples of 0s, so we report them for a random state.

3.4. Pairwise classification performance

Figure 1 shows multidimensional scaling (MDS) based on l_2-norm dissimilarity matrix of bag of edges for ADNI subjects (for PPMI data picture is essentially the same, so we omitted it). We see that in most cases images from same subject are near each other in that feature space. We also see that there is no such clear picture with diagnostic groups labels.

Figures 2,3 quantify this observation for ADNI/PPMI data in terms ROC AUC distributions depending on normalization and base features. We see that 0.99 ROC AUC can be achieved either for connectome weights itself or features that represent structure. We also see that choice of normalization greatly affects quality of pairwise classification and normalization by mean is a winner in most cases except eigenvector centrality and clustering coefficient features. It is interesting to note that for clustering coefficient and eigenvector centrality binary normalization performed better then other normalizations although it preserves somewhat less information.

Figures 3-4 shows ROC AUC distribution of pairwise classification depending on base features and diagnostic group for connectomes normalized by mean. We see that there is almost no difference in pairwise classification results in different diagnostic groups, only that interquartile spread is high for diagnostic groups with smallest number of subjects.

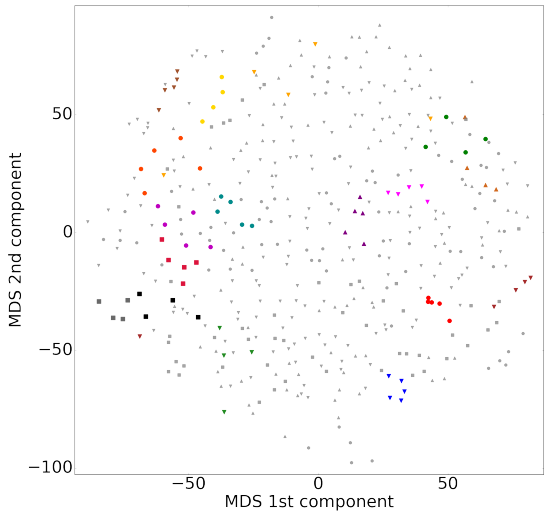

Fig. 1. Multidimensional scaling for ADNI data based on l_2 norm. Points with same colors represent same subjects. To avoid mess on the pictures we've decided to color only 17 subjects. Other subjects represented by smaller grey markers. Marker shape indicates diagnostic group: ○ – controls, △ – LMCI, ▽ – EMCI, □ – AD.

4. CONCLUSION

We have presented a method of structural connectome feature validation through pairwise classification based on them which doesn't require assumptions on distribution of these features. We tested this pipeline on ADNI and PPMI data and obtained high classification performance in terms of ROC AUC which means there are mappings of connectomes to feature spaces that at least differentiate subjects from each other.

It is worth noting that pairwise classification is not feature selection technique for classification tasks. It is possible that feature distinguish classes, but doesn't distinguish subjects at all (class label is extreme example of such feature). Our results suggest pairwise classification may be useful for validating preprocessing pipelines and particular features in terms of how much subjected-related signal they preserve. For us it is a way to verify that network constructing pipelines and features we use for another classification problems are meaningful enough to work with.

There are several limitations. First, we used only one particular preprocessing pipeline for each of two used datasets. These results may differ for other tractography algorithms and parcellation scheme choices and it would be interesting to see how exactly it affects pairwise classification.

Second, downsampling procedure which we used to obtain even sample of data pars from same/different subjects

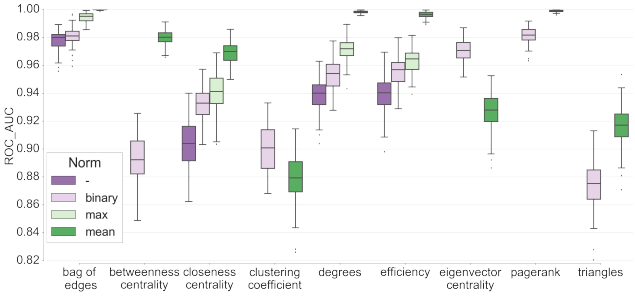

Fig. 2. ROC AUC distributions for pairwise classification on all ADNI data depending on choice of connectome normalizing scheme and base features.

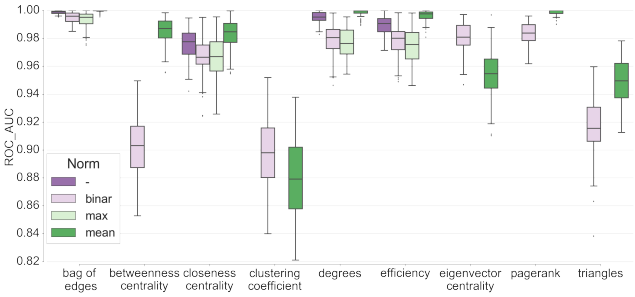

Fig. 3. ROC AUC distributions for pairwise classification on all PPMI data depending on choice of connectome normalizing scheme and base features.

may have increased our results, because as seen in MDS visualization there are cases when images from different subjects are near each other but it is highly unlikely that downsampling procedure will select them as negative example. We plan to test our pipeline on the whole pairs data as well.

Finally, images for both ADNI and PPMI data were obtained with time difference at least a year. We can't be sure that subject connectomes didn't change due to some reasons (i.e. disease progression). It may mean that normalizations which work well for pairwise classification do not reflect this changes and can be useless for disease progression prediction.

5. ACKNOWLEDGMENTS

The results of sections 2 and 3 are based on the scientific research conducted at IITP RAS and supported by the Russian Science Foundation (project 14-50-00150).

First dataset used in preparation of this article were obtained from the Alzheimers Disease Neuroimaging Initiative (ADNI) database. As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of

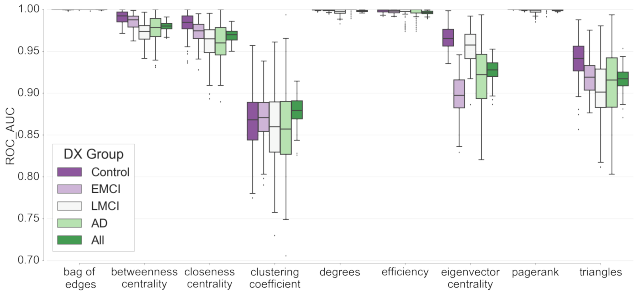

Fig. 4. ROC AUC distributions for pairwise classification on ADNI data for different diagnostic groups on different 'base' features. These results are reported for connectome normalization by mean.

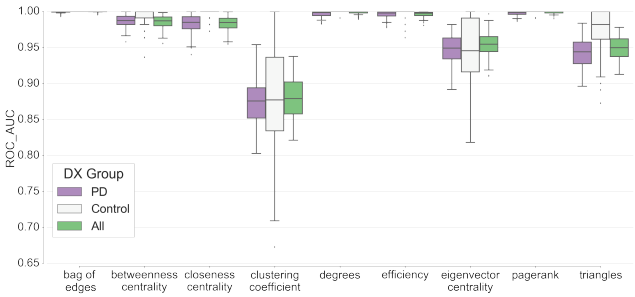

Fig. 5. ROC AUC distributions for pairwise classification on PPMI data for different diagnostic groups on different 'base' features. These results are reported for connectome normalization by mean.

this report. A complete listing of ADNI investigators as well as data acquisition protocols can be found at `adni.loni.usc.edu`. Second dataset used in the preparation of this article were obtained from the Parkinsons Progression Markers Initiative (PPMI) database. For up-to-date information on the study, visit `www.ppmi-info.org`. PPMI a public-private partnership is funded by the Michael J. Fox Foundation for Parkinsons Research and funding partners, including (list the full names of all of the PPMI funding partners found at `www.ppmi-info.org/fundingpartners`).

6. REFERENCES

[1] Mohammad R Arbabshirani, Sergey Plis, Jing Sui, and Vince D Calhoun, "Single subject prediction of brain disorders in neuroimaging: Promises and pitfalls," *NeuroImage*, 2016.

[2] Liang Zhan, Jiayu Zhou, Yalin Wang, Yan Jin, Neda Jahanshad, Gautam Prasad, Talia M Nir, Cassandra D

Leonardo, Jieping Ye, Paul M Thompson, et al., "Comparison of nine tractography algorithms for detecting abnormal structural brain networks in alzheimers disease," *Frontiers in aging neuroscience*, vol. 7, pp. 48, 2015.

[3] Marcel A de Reus and Martijn P Van den Heuvel, "The parcellation-based connectome: limitations and extensions," *Neuroimage*, vol. 80, pp. 397–404, 2013.

[4] Pierre Besson, Renaud Lopes, Xavier Leclerc, Philippe Derambure, and Louise Tyvaert, "Intra-subject reliability of the high-resolution whole-brain structural connectome," *NeuroImage*, vol. 102, pp. 283–293, 2014.

[5] Julia P Owen, Etay Ziv, Polina Bukshpun, Nicholas Pojman, Mari Wakahiro, Jeffrey I Berman, Timothy PL Roberts, Eric J Friedman, Elliott H Sherr, and Pratik Mukherjee, "Test–retest reliability of computational network measurements derived from the structural connectome of the human brain," *Brain connectivity*, vol. 3, no. 2, pp. 160–176, 2013.

[6] Leonardo Bonilha, Ezequiel Gleichgerrcht, Julius Fridriksson, Chris Rorden, Jesse L Breedlove, Travis Nesland, Walter Paulus, Gunther Helms, and Niels K Focke, "Reproducibility of the structural brain connectome derived from diffusion tensor imaging," *PloS one*, vol. 10, no. 9, pp. e0135247, 2015.

[7] Chad J Donahue, S Sotiropoulos, Saad Jbabdi, M Herandez-Ferandez, T Behrens, H Kennedy, et al., "Using diffusion tractography to predict cortical connection strength and distance: a quantitative comparison with tracers in the monkey," *J Neurosci*, vol. 36, pp. 6758–70, 2016.

[8] Xi-Nian Zuo, Clare Kelly, Jonathan S Adelstein, Donald F Klein, F Xavier Castellanos, and Michael P Milham, "Reliable intrinsic connectivity networks: test–retest evaluation using ica and dual regression approach," *Neuroimage*, vol. 49, no. 3, pp. 2163–2177, 2010.

[9] Tengda Zhao, Fei Duan, Xuhong Liao, Zhengjia Dai, Miao Cao, Yong He, and Ni Shu, "Test-retest reliability of white matter structural brain networks: a multiband diffusion mri study," *Frontiers in human neuroscience*, vol. 9, 2015.

[10] Reinhold Müller and Petra Büttner, "A critical discussion of intraclass correlation coefficients," *Statistics in medicine*, vol. 13, no. 23-24, pp. 2465–2476, 1994.

[11] Emily S. Finn, Xilin Shen, Dustin Scheinost, Monica D. Rosenberg, Jessica Huang, Marvin M. Chun, Xenophon Papademetris, and R. Todd Constable, "Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity," *Nat Neurosci*, vol. 18, no. 11, pp. 1664–1671, Nov 2015, Article.

[12] Patric Hagmann, Maciej Kurant, Xavier Gigandet, Patrick Thiran, Van J Wedeen, Reto Meuli, and Jean-Philippe Thiran, "Mapping human whole-brain structural networks with diffusion mri," *PloS one*, vol. 2, no. 7, pp. e597, 2007.

[13] Danielle S Bassett, Jesse A Brown, Vibhas Deshpande, Jean M Carlson, and Scott T Grafton, "Conserved and variable architecture of human white matter connectivity," *Neuroimage*, vol. 54, no. 2, pp. 1262–1279, 2011.

[14] Mikail Rubinov and Olaf Sporns, "Complex network measures of brain connectivity: uses and interpretations," *Neuroimage*, vol. 52, no. 3, pp. 1059–1069, 2010.

[15] Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Winograd, "The pagerank citation ranking: Bringing order to the web.," Technical Report 1999-66, Stanford InfoLab, November 1999, Previous number = SIDL-WP-1999-0120.

[16] Bruce Fischl, "Freesurfer," *Neuroimage*, vol. 62, no. 2, pp. 774–781, 2012.

[17] Rahul S Desikan, Florent Ségonne, Bruce Fischl, Brian T Quinn, Bradford C Dickerson, Deborah Blacker, Randy L Buckner, Anders M Dale, R Paul Maguire, Bradley T Hyman, et al., "An automated labeling system for subdividing the human cerebral cortex on mri scans into gyral based regions of interest," *Neuroimage*, vol. 31, no. 3, pp. 968–980, 2006.

[18] Eleftherios Garyfallidis, Matthew Brett, Bagrat Amirbekian, Ariel Rokem, Stefan Van Der Walt, Maxime Descoteaux, and Ian Nimmo-Smith, "Dipy, a library for the analysis of diffusion mri data," *Frontiers in neuroinformatics*, vol. 8, pp. 8, 2014.

[19] Chantal MW Tax, Ben Jeurissen, Sjoerd B Vos, Max A Viergever, and Alexander Leemans, "Recursive calibration of the fiber response function for spherical deconvolution of diffusion mri data," *Neuroimage*, vol. 86, pp. 67–80, 2014.

[20] Samuel St-Jean, Pierrick Coupé, and Maxime Descoteaux, "Non local spatial and angular matching: Enabling higher spatial resolution diffusion mri datasets through adaptive denoising," *Medical Image Analysis*, vol. 32, pp. 115–130, August 2016.