Improving Corticostriatal Parcellation Through Multilevel Bagging with PyBASC

Aki Nikolaidis PhD (aki.nikolaidis@childmind.org)
The Child Mind Institute, 445 Park Avenue. New York, NY 10022
Joshua Vogelstein PhD (jovo@jhu.edu)
John's Hopkins University 3400 N. Charles St Baltimore, MD 21218
Pierre Bellec PhD (pierre.bellec@criugm.qc.ca)
University of Montreal, PO Box 6128 Downtown STN Montreal QC H3C 3J7 CANADA
Michael P. Milham MD PhD (michael.milham@childmind.org)
The Child Mind Institute, 445 Park Avenue. New York, NY 10022

Abstract: Understanding the functional organization the brain is a centrally important theme of human neuroscience. Ideally, these organizational maps uncover the underlying structure of the brain's functional architecture, and group-level maps are accurate representations of the individuals in the sample. Using simulated fMRI data, we demonstrate that bagging improves the ability of clustering to uncover the data's underlying structure. We show that the group-level maps become more correlated to the individual-level maps with more bootstrap aggregates, suggesting bagging improves the representativeness of the group-level solution. Using a test-retest dataset of 30 young adults, we confirm these findings. More specifically, we see bagging improves the test-retest correlation between cluster maps, and increases correlation between group-level and individual-level cluster maps, and these effects are robust to number of clusters and length of scan used. These results suggest bagging is an important method for increasing reliability and validity of functional parcellation approaches.

Keywords: fMRI, Corticostriatal circuits, basal ganglia, functional parcellation, bagging, bootstrap,

Introduction

The basal ganglia (BG) is a functionally heterogeneous structure that interacts with the cortex to produce a wide range of motor, cognitive, and affective functions[1–8]. More recently, it has become clear that interactions between the basal ganglia and cortex play a critical role in learning, cognition, and psychopathology[9–14]. Understanding the functional organization of these cortico-subcortical interactions remains a centrally important theme of human neuroscience. Robust examinations of these network interactions have been limited due to the high variability in the functional connectivity of these subcortical structures. This issue is further compounded by the low signal to noise often present in subcortical areas. BASC, Bootstrap Analysis for Stable Clusters, is a technique developed to overcome these limitations and provide a robust method for functional brain parcellation[15,16]. Recently, we've leveraged the processing power of Python-based pipelining framework, NiPype, to implement and extend this technique as an open source Python Package: PyBASC. Here, we demonstrate its ability to recover true cross-network cluster structure in simulated cortico-striatal data, and create more reproducible cortico-striatal cluster solutions in a real test-retest dataset.

Ensemble learning is a class of highly effective methods for improving model accuracy and generalizability[17,18]. The specific multi-level bootstrap aggregating (bagging) approach to clustering[19,20] implemented in PyBASC has been shown to improve clustering performance[20,21]. Bagging approaches may improve clustering because training data might not be sufficient for choosing the single best learner over a group of equally performing models[18,22]. Even in the presence of a single optimal solution, as demonstrated here, identifying this solution might still be difficult to achieve, especially in the presence of noise[18]. Cluster stability has also been demonstrated to be a good hypothesis-independent criterion for choosing the number of clusters to derive in a dataset[22]. We use both individual and group-level bagging approaches to build stability maps of clustering across individuals. This approach allows for examination of individual-level variation both with- and without group-level priors, as well as rigorous comparisons between individuals in functional organization patterns.

Methods

PyBASC Method

Originally, this method was implemented in Octave, and we have recently implemented this method in Python to take advantage of the broad range of powerful libraries available for machine learning and neuroimaging, including NiBabel, Nilearn, Scikit Learn, and NiPype. This enables PyBASC to be highly flexible and extensible in adopting new clustering algorithms and implementations, but more importantly allows for dramatic increases in computational power through the use of NiPype's easy-to-use and adjust multi-core processing capabilities. We have also made edits to the processing that make PyBASC run very efficiently with minimal memory loads. Here, we briefly address the PyBASC overall methodology. For more information, we have also explained these methods in more detail elsewhere[16]. PyBASC has a wide range of functionalities that can be customized according to user preferences. Well-performing default values are provided where applicable, but

users can choose 1) the extent of the initial feature agglomeration, 2) the number of individual and 3) group-level bootstraps, the 4) region to be parcellated and 5) the secondary region (optional) to use as the basis of cross-network clustering. Users can also define the affinity threshold and distance metric used in the clustering procedures.

Individual-level Bagging First, to reduce the dimensionality of the voxel wise time series, and to allow for better computational efficiency, we apply a hierarchical feature agglomeration from Scikit Learn with a neighboring voxel spatial constraint to each individual's data. Then, a circular block bootstrap procedure is applied to create resampled time series data from the original functional image. Each of these resampled time series are then clustered using the clustering technique of choice. Currently k-means, hierarchical agglomerative, and spectral clustering are implemented. With each application of clustering, the cluster labels are transformed into an adjacency matrix of 0s and 1s representing whether a given voxel belongs to the same cluster as another voxel, and the dimension reduction procedure is reversed to create a voxel wise adjacency matrix that puts all individual-level data into the same space. All adjacency matrices are averaged together to create an individual stability matrix (ISM), which represents a summary of all clustering assignments across all bootstraps. The ISM can be clustered to produce a final consensus cluster assignment for each individual, or it can used in group-level bagging.

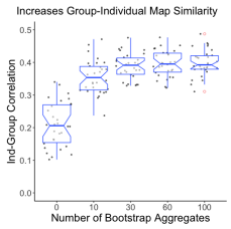

Figure 2

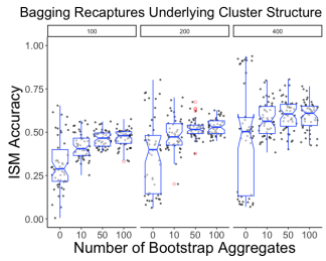

Figure 1

Group-level Bagging PyBASC applies bagging to ISMs across individuals, recreating a resampled group of ISMs from the original pool of subjects. The ISMs for each group-level bootstrap are averaged and an adjacency matrix is created through clustering. The adjacency matrix across all bootstraps are averaged to create a group stability matrix (GSM). The GSM is the final output of the group-level bagging, along with the cluster labels, and the group-level stability for each cluster. While the ISM represents a summary of the clusterings across all bootstraps for each individual, the GSM represents a summary of the clustering across all bootstraps for the group. We compare each individual's ISM to the GSM as a detailed way of assessing similarities between the individual and group-level clusterings.

Data Used and Created

Creating Individual-Level Bagging Simulations We generated 1500 simulated time series with a two clusters, each corresponding to 500 and 1000 of the time series respectively, and a signal to noise ratio (*B/E)* of 0.05/2.5, as with some higher values we found ceiling effects in the clustering. We performed individual-level clustering on these data with a simplified first level analysis of PyBASC. We repeated this analysis with and without bagging, and across a range of bootstraps to aggregate. For each combination of parameters, we repeated the analysis 50 times to calculate an estimate of the reliability of our effects (Figure 1). We use Ward's method for hierarchical agglomerative clustering because this method demonstrated superior accuracy compared to other techniques in our simulations (not presented here). This choice is also supported by recent work suggesting Ward's method to be superior to K-means and spectral clustering for both reproducibility and cluster accuracy in simulation and real test datasets[23].

Creating Group-Level Bagging Simulations To simulate the motor corticostriatal thalamocortical network, we used the same data generation algorithm above to create simulated time series with the same dimensions as the bilateral striatum and thalamus, and the Yeo Motor network[24]. We transformed these synthetic data into NIFTI images and used these synthetic MRI data in a full run of the PyBASC pipeline. We used a signal to noise ratio of 0.05/3 for our group-level data, as the group-level results are much better at capturing underlying structure in the presence of noise than the individual-level results. We clustered the simulated striatum and thalamus regions with respect to their connectivity to the Yeo Motor network to simulate corticostriatal connectivity. We repeated this analysis with and without bagging, and across a range of bootstraps to aggregate. We also varied the number of volumes in each 4D dataset, the SNR of the data, and the individual-level bootstraps as in the IBS simulation.

Assessing HNU Test Retest Data We used full runs of the PyBASC pipeline to investigate the impact of multilevel bagging on the test retest reproducibility using the HNU CORR Dataset[25]. 30 young adults had a 10 minute resting state scan acquired once every three days for a month, for a total of 10 sessions per person. To create a reference dataset for each participant, we concatenated each participant's even numbered session scans. We used PyBASC on the Reference data with cross clustering aiming to cluster the striatum and thalamus with respect to its connectivity to the Yeo motor network. We compared the group stability matrices in the Reference and Replication datasets, which were created from either 10, 20, 30, or 40 minutes of data from the odd numbered scan sessions. Broadly speaking, group cluster solutions should be representative of the individuals in the sample. To assess the extent to which our group-level parcellations were representative of each of the individuals in the sample, we also correlated each individual's ISM generated from PyBASC with the GSM generated from the

group. This allowed us to examine differences in how representative the group-level solutions were across levels of bootstrap aggregation, and length of our replication scans.

Results

Individual Bagging Simulations

Simulating the clustering approach at the individual-level demonstrated that bagging provides large improvements in the accuracy of the clustering from no bootstraps to 100 bootstraps (Figure 1). We also demonstrate that bagging considerably reduces the variability in accuracy of the cluster solution. The effect of bagging replicates across 100, 200, and 400 volumes. Most notably, with only 100 volumes, the 100 bootstrap aggregate clusters are more accurate than the 200 volume data are without bagging, and is equivalent to the 400 volume data without bagging.

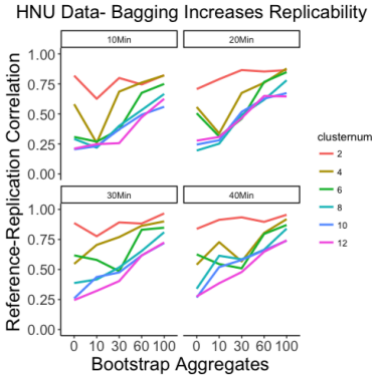

Figure 3

Group Bagging Simulations

In Figure 2, we demonstrate with group-level simulated MRI data, that with greater bootstrap aggregates, the group-level maps become more representative of each of the individual-level maps. Following the same data generation algorithm we created 30 subjects and demonstrated that multilevel bagging also increases the the coherence of the individual and group-level stability matrices, suggesting that it creates group-level clusters solutions that are more representative of the individual subjects.

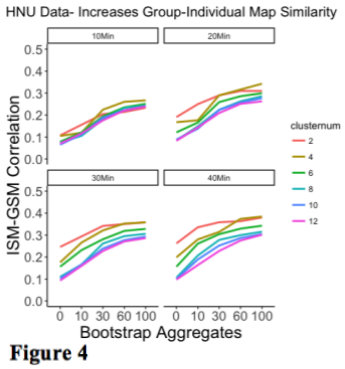

Figure 4

HNU Data

As expected, we found that having more data improves the reproducibility of our cluster assignments. Assessing the impact of bagging on the actual HNU data revealed that bagging improves clustering reproducibility (Figure 3). We found that using bootstrap aggregation increases the correlation between the group stability matrix of the reference and replication datasets, and that this effect holds across a range of volumes and cluster numbers (Figure 3). We found that as the number of clusters increase, bagging improves the reproducibility of the cluster assignments more (Figure 3), suggesting that with larger networks and more regions, using bagging will further improve the reproducibility of the functional parcellation. We also found that, that bootstrap aggregation also improves the group-individual coherence (Figure 4), with results in the HNU following along with our simulations in Figure 2.

Discussion

Producing replicable functional parcellations is a critical step for understanding the organization of brain networks. Furthermore, in areas of the brain that are typically difficult to study, such as the basal ganglia, we need advanced methods that are robust against noise. In the current work, we demonstrated that bagging, as implemented in PyBASC, offers a range of important advantages for creating an accurate functional parcellation of the brain. First, we use simulations to show that without bagging, individual-level clustering solutions have high variance and are not necessarily able to capture the underlying data structure reliably (Figure 1). With bagging however, clusters solutions are significantly more consistent, and this consistency grows with the number of bootstraps aggregated. When creating functional parcellations at the individual-level, it's very important to know the extent to which the parcellations are reliable, and our results suggest that methods such as bagging may be employed to improve reliability of the clustering solutions.

Without bagging, not only are individual-level solutions less accurate, they tend to be more variable as well. In our group-level simulations, we found that the group-level clusterings map onto the individual-level results better as a function of bootstrap aggregation. The more bootstraps aggregated, the less variability in cluster accuracy we see at the individual-level and the more accurately the group-level cluster results represent the individual-level clustering (Figure 2). Finally, we replicate these simulation results in the HNU Dataset, demonstrating that the theoretical effects seen in our simulations hold in real data as well. We show that not only does bagging improve clustering accuracy (Figure 1), but it also improves reproducibility of our findings and this effect is robust to number of clusters and length of data (10-40 minutes) (Figure 3). Just as we saw that bagging improves similarity of our group- and individual-level cluster results in simulation (Figure 2), we demonstrate that it has the same impact in real data, and this effect is also robust to number of clusters and length of data (Figure 4). Taken together, our results demonstrate that multilevel bagging provides not only better group-level estimates of functional parcellation, but also better group-level representations of individual-level parcellations. This technique may offer a unique opportunity for reliably estimating functional network organization in datasets with short scan times. Furthermore, given this technique improves reliability, this technique may offer better insight

into how individual differences in functional organization contribute to variation in cognition, behavior, and psychopathology.

Future Directions

We plan to apply this technique to a range of cortico-striatal circuits and map their maturation across childhood and adolescent development. Using PyBASC, we can develop robust parcellations of corticostriatal networks, and will be able to use these network nodes to provide reliable and detailed maps of the developmental trajectories of these circuits and their role in psychopathology, cognitive development, and learning.

Acknowledgements

We would like to thank the Child Mind Institute for providing the computational resources and support for the current work. Special thanks to Lei Ai and Steve Giavasis, for their assistance in the processing of the HNU data.

References

1. Alexander, G. E., DeLong, M. R. & Strick, P. L. Parallel Organization of Functionally Segregated Circuits Linking Basal Ganglia and Cortex. *Annu. Rev. Neurosci.* **9,** 357–381 (1986).
2. Coxon, J. P. *et al.* Functional Brain Activation Associated with Inhibitory Control Deficits in Older Adults. *Cereb. Cortex* **26,** 12–22 (2016).
3. Goble, D. J. *et al.* Brain Activity during Ankle Proprioceptive Stimulation Predicts Balance Performance in Young and Older Adults. *Journal of Neuroscience* **31,** 16344–16352 (2011).
4. Goble, D. J. *et al.* The neural basis of central proprioceptive processing in older versus younger adults: An important sensory role for right putamen. *Hum. Brain Mapp.* **33,** 895–908 (2012).
5. Graybiel, A. M. Habits, rituals, and the evaluative brain. *Annu. Rev. Neurosci.* **31,** 359–387 (2008).
6. Hikosaka, O., Nakamura, K., Sakai, K. & Nakahara, H. Central mechanisms of motor skill learning. *Curr. Opin. Neurobiol.* **12,** 217–222 (2002).
7. Hoshi, E., Tremblay, L., Feger, J., Carras, P. L. & Strick, P. L. The cerebellum communicates with the basal ganglia. *Nat. Neurosci.* **8,** 1491–1493 (2005).
8. Postuma, R. B. & Dagher, A. Basal ganglia functional connectivity based on a meta-analysis of 126 positron emission tomography and functional magnetic resonance imaging publications. *Cereb. Cortex* **16,** 1508–1521 (2006).
9. de Pasquale, F. *et al.* A cortical core for dynamic integration of functional networks in the resting human brain. *Neuron* **74,** 753–764 (2012).
10. He, B. J. & Raichle, M. E. The fMRI signal, slow cortical potential and consciousness. *Trends Cogn. Sci.* **13,** 302–309 (2009).
11. Cai, W., Chen, T., Szegletes, L., Supekar, K. & Menon, V. Aberrant Time-Varying Cross-Network Interactions in Children With Attention-Deficit/Hyperactivity Disorder and the Relation to Attention Deficits. *Biol Psychiatry Cogn Neurosci Neuroimaging* **3,** 263–273 (2018).
12. Nikolaidis, A. & Barbey, A. K. 21 Network Dynamics Theory of Human Intelligence. *The Cambridge Handbook of the* (2018).
13. Nikolaidis, A., Voss, M. W., Lee, H., Vo, L. T. K. & Kramer, A. F. Parietal plasticity after training with a complex video game is associated with individual differences in improvements in an untrained working memory task. *Front. Hum. Neurosci.* **8,** (2014).
14. Nikolaidis, A., Goatz, D., Smaragdis, P. & Kramer, A. Predicting Skill-Based Task Performance and Learning with fMRI Motor and Subcortical Network Connectivity. *2015 International Workshop on Pattern Recognition in NeuroImaging* 93–96 (2015).
15. Bellec, P., Rosa-neto, P., Lyttelton, O. C., Benali, H. & Evans, A. C. NeuroImage Multi-level bootstrap analysis of stable clusters in resting-state fMRI. *Neuroimage* **51,** 1126–1139 (2010).
16. Garcia-Garcia, M. *et al.* Detecting stable individual differences in the functional organization of the human basal ganglia. *Neuroimage* (2017). doi:10.1016/j.neuroimage.2017.07.029
17. Jain, A. K. Data clustering: 50 years beyond K-means. *Pattern Recognit. Lett.* **31,** 651–666 (2010).
18. Zhou, Z.-H. *Ensemble Methods: Foundations and Algorithms*. (CRC Press, 2012).
19. Breiman, L. Bagging predictors. *Mach. Learn.* **24,** 123–140 (1996).
20. Hoyos-Idrobo, A., Schwartz, Y., Varoquaux, G. & Thirion, B. Improving sparse recovery on structured images with bagged clustering. in *Pattern Recognition in NeuroImaging (PRNI), 2015 international workshop on* 73–76 (IEEE, 2015).
21. Dudoit, S. & Fridlyand, J. Bagging to improve the accuracy of a clustering procedure. *Bioinformatics* **19,** 1090–1099 (2003).
22. von Luxburg, U. Clustering Stability: An Overview. *Foundations and Trends® in Machine Learning* **2,** 235–274 (2010).
23. Thirion, B., Varoquaux, G., Dohmatob, E. & Poline, J. B. Which fMRI clustering gives good brain parcellations? *Front. Neurosci.* **8,** 1–13 (2014).
24. Yeo, B. T. T. *et al.* The organization of the human cerebral cortex estimated by intrinsic functional connectivity. 1125–1165 (2011).
25. HNU1. doi:10.15387/fcp_indi.corr.hnu1