Identifying brain networks in a clinically rich and naturalistic dataset using tensor decomposition

Jeff Mentch[1,2], Jian Li[3,4], Satrajit Ghosh[1,2]

[1] Speech and Hearing Bioscience and Technology, Harvard Medical School, Boston, MA

[2] McGovern Institute for Brain Research, MIT, Cambridge, MA

[3] Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Charlestown, MA

[4] Center for Neurotechnology and Neurorecovery, Massachusetts General Hospital, Boston, MA

Introduction

Much of our knowledge of cognition stems from experiments manipulating highly controlled stimuli. While this has proven fruitful, our everyday lives feature more rich and dynamic streams of information. Recently, studies using naturalistic stimuli have gained increasing attention as a more ecologically valid method of studying the brain. While they have traditionally been used to examine the shared responses across groups of individuals, less work has targeted individual differences and heterogenous clinical populations like those with autism spectrum condition (ASC), characterized by altered audiovisual perception and social communication. Naturalistic stimuli are also amenable to data-driven approaches like independent component analysis (ICA), an influential tool for identifying brain networks that can be performed either individually or collapsing temporally or spatially for group analysis (Calhoun et al., 2009). This necessary concatenation of the data into a 2D representation may potentially lose some inherent low-rank structure shared among subjects. Further, the spatial or temporal independence assumed in the ICA model may not be biologically plausible given the extensive spatial and temporal overlap between brain networks. An alternative approach to this problem without imposing independence constraint is NASCAR, a stable and robust method for identifying brain networks and their temporal dynamics across subjects using tensor decomposition (Li et al., 2021). Here we apply NASCAR to fMRI data from naturalistic stimulation and at rest in a clinically rich dataset.

Methods

A subset of Healthy Brain Network (HBN) subjects were acquired including individuals with autism spectrum condition (ASC) (n=78; ASC=37). The HBN biobank, approved by the Chesapeake Institutional Review Board, recruited individuals age 5–21 at four study sites in the New York City area and full methodological details are available in their data release (Alexander et al., 2017). fMRI data from naturalistic stimulation with a short film and at rest were preprocessed using fMRIprep 22.0.1 with CIFTI2 outputs. Motion outliers, cosine, 5 acompcor, csf, framewise displacement, and 6 motion confounds were removed and detrending was performed with Nilearn (Abraham et al., 2014) and a 4 mm FWHM Gaussian smoothing was applied using Connectome Workbench (Marcus et al., 2011). fMRI data were temporally aligned using BrainSync (Joshi et al., 2018) and mapped to a 3D tensor (grayordinate x TR x subject).

Then a Canonical Polyadic decomposition was performed using NASCAR to identify brain networks as low rank approximation of the data. Audio and visual features were extracted from the movie using Pliers and HRF convolved for reverse correlation analysis.

Results

Networks qualitatively similar to 12 known brain networks from Li et al. were observed from movie-viewing and at rest. Movie time points with the greatest mean absolute value across temporal modes were mapped back to the original stimulus and found to be engaging and notable moments of the film. Reverse correlation identified the top movie features matching the temporal mode of each component to highlight which movie features may drive each component. Qualitatively, networks encompassing visual cortex were found to be more highly correlated with visual features and networks encompassing auditory cortex were more correlated with auditory features. Only component 16 (independent t-test, p=.05) had greater contributions from controls compared to ASCs, although this trend did not survive correction for multiple comparisons.

Conclusion

Results suggest that short naturalistic and resting-state scans just a few minutes in length can be decomposed into well documented networks using tensor decomposition. Further work is needed to compare network time courses of clinical groups during naturalistic stimulation and test how stable such components are across stimuli, sessions, subjects, and clinical populations.

Reference

Abraham, A., Pedregosa, F., Eickenberg, M., Gervais, P., Mueller, A., Kossaifi, J., Gramfort, A., Thirion, B., & Varoquaux, G. (2014). Machine learning for neuroimaging with scikit-learn. *Frontiers in Neuroinformatics*, *8*, 14.

Alexander, L. M., Escalera, J., Ai, L., Andreotti, C., Febre, K., Mangone, A., Vega-Potler, N., Langer, N., Alexander, A., Kovacs, M., Litke, S., O'Hagan, B., Andersen, J., Bronstein, B., Bui, A., Bushey, M., Butler, H., Castagna, V., Camacho, N., … Milham, M. P. (2017). An open resource for transdiagnostic research in pediatric mental health and learning disorders. *Scientific Data*, *4*, 170181.

Calhoun, V. D., Liu, J., & Adali, T. (2009). A review of group ICA for fMRI data and ICA for joint inference of imaging, genetic, and ERP data. *NeuroImage*, *45*(1 Suppl), S163–S172.

Joshi, A. A., Chong, M., Li, J., Choi, S., & Leahy, R. M. (2018). Are you thinking what I'm thinking? Synchronization of resting fMRI time-series across subjects. *NeuroImage*, *172*, 740–752.

Li, J., Wisnowski, J. L., Joshi, A. A., & Leahy, R. M. (2021). Robust brain network identification from multi-subject asynchronous fMRI data. *NeuroImage*, *227*, 117615.

Marcus, D. S., Harwell, J., Olsen, T., Hodge, M., Glasser, M. F., Prior, F., Jenkinson, M.,

Laumann, T., Curtiss, S. W., & Van Essen, D. C. (2011). Informatics and data mining tools and strategies for the human connectome project. *Frontiers in Neuroinformatics*, *5*, 4.

McNamara, Q., De La Vega, A., & Yarkoni, T. (2017, August). Developing a comprehensive framework for multimodal feature extraction. In Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 1567-1574). ACM.