Graph Clustering of Functional Connectomes as a Diagnostic Tool for Autism Spectrum Disorder

Ryan Marren
Johns Hopkins University
EN.600.479 Final
rmarren1@jhu.edu

Abstract

Autism Spectrum Disorder is a neuro-developmental disorder which affects roughly 1.5% of American children. Advances in neuro-imaging technologies have lead to new hypotheses relating the presence of ASD to certain functional brain network topologies. Drawing inspiration from these hypotheses, we will leverage various graph clustering techniques build diagnostic classifiers of fMRI scans. We will then evaluate these classifiers to show that they systematically outperform chance on ASD classification tasks. Finally, we explore the efficacy of augmenting feature vectors with secondary connectome views derived from different brain atlases.

1 Introduction

1.1 Connectomes and Brain Imaging

A "functional connectome" is a term used to describe a collection of functional connections in a brain [2]. Functional connectomes can be represented in many ways. For simplicity, connectomes in this paper can be thought of as an adjacency matrix representation of a graph where nodes are brain regions of interest (ROIs) and edges measure the correlations of neural activations between these ROIs.

The connectomes used here are derived from resting state fMRI scans [4]. fMRI is an acronym for 'functional Magnetic Resonance Imaging' and is a technique to indirectly measure brain activity by measuring the levels of blood oxygenation in different regions of the brain. The spatial resolution (size in voxels of resulting images) is 1 to 5 millimeters while the temporal resolution (smallest timescale over which changes in neural activity can be reliably measured) is a few seconds.

Recently the ABIDE I data set [6] was released. This is a set of 539 fMRI scans of patients suffering from Autism Spectrum Disorder and 573 from typical controls, coming from 16 international imaging sites. This provides an opportunity to test hypotheses without the normally prohibitive steps of acquiring and pre-processing neuroimaging data.

1.2 Autism Spectrum Disorder

Autism Spectrum Disorder is characterized by disabling impairments in social interaction, repetitive behaviors, and restricted interests [1]. Due to these symptoms, it is hypothesized that ASD would be associated with global hypo-connectivity [15] and local hyper-connectivity [5] of neural circuits in a functional connectome. These hypotheses, along with the recent availability of public datasets like ABIDE, have caused a recent surge in publications attempting to classify brain disorder with machine learning methods. See [3] for an excellent review of these attempts.

29th Conference on Neural Information Processing Systems (NIPS 2016), Barcelona, Spain.

2 Previous Work

2.1 Previous attempts to classify ASD using fMRI data

There have been many recent attempts to build a diagnostic classifier for ASD using connectome representations derived from fMRI data. Dodero et al. [8] uses Graph Laplacian features run through a Kernel SVM on two groups of 79 (42+, 37-) patients and 91 (54+, 43-) patients with resulting accuracies of 61% and 74% respectively. Jin et al. [14] uses a Multi-Kernel SVM on diffusion patterns of White Matter connectomes for 80 (40+, 40-) infants, and reports accuracies of 76%. Wee et al. [19] uses ROI signal correlations from temporal clusters in conjunction with LASSO, various kernels, and a Kernel SVM on a set of 92 (45+, 47-) patients with an accuracy of 71%. Finally, Iidaka [13] uses a Probabilistic Neural Network on the ROI signal correlations of 640 (312+, 328-) children and reports accuracies of 90%.

2.2 Previous Analyses of ABIDE I

As they were released very recently, the ABIDE I Preprocessed data have seen few analyses. Vigneshwaran et al. [18] leverages 54837 dimensional feature vectors extracted using regional homogeneity (ReHo) run through a 'Projection Based Learning Metacognitive Radial Basis Function Network' and also a SVM, returning accuracies of 66.96% and 63.76% respectively (on only the male subjects). Sato et al. [17] analyzes functional connectomes using a fuzzy variant of spectral clustering to identify 5 brain regions which, with $p < 0.5$, have less entropy in ASD patients than controls. Finally, Elton et al. [10] attempts to discover relationships between seed based functional connections and categorical variables using linear regression models.

3 Methods

3.1 Data Gathering

The ADIBE I Preprocessed data-set hosted at http://preprocessed-connectomes-project.org/abide/download.html was downloaded using the available Python script. The main data derivative was the time-series extracted from the Eickhoff-Zilles atlas, which is a parcellation of the brain volume into 116 ROIs derived from post-mortem segmentation [9]. We also use the CC200 atlas as a secondary view in later experiments, which is a 200 ROI atlas derived from constrained spectral clustering [7]. The data were run through the CCS preprocessing pipeline and were further subject to bandpass and global regression filtering. Samples which were too strongly effected by noise were thrown out. The resulting Eickhoff-Zilles data-set consists of 408 116-dimensional time-series from ASD patients and 476 116-dimensional time-series from typical controls, for a total of 884 samples. The CC200 data-set is identical in dimensionality other than for the number of ROIs. Time series have different numbers of observations, with a maximum of 315 observations, a minimum of 77 observations, and an average of 191.2 observations over the data set. Before each experiment, the entire data set was randomly permuted. Then the first 684 subjects were taken as our Training Set, the next 100 as our Tuning Set, and the last 100 as our Testing Set.

3.2 Classification based on correlation matrix clustering

We will begin by building classifiers of functional connectomes using only the Eickhoff-Zilles atlas view. First, we calculate the correlation matrix of the ROIs for each sample in the Training Set. Then, we run each correlation matrix through a graph clustering algorithm to receive a list of ROI cluster labels. These lists of labels are then transformed into a sparse matrix representation of groupings. These are matrices $B^p \in \{0,1\}^{116\times 116}, p \in \{1,\ldots,664\} \subset \mathbb{Z}$ where 116 is the number of ROIs, 664 is the number of patients in the Training Set, and:

$$B^p_{ij} = \begin{cases} 1 & i \text{ has same label as } j \\ 0 & \text{otherwise} \end{cases}$$

Finally, we group these representations by condition, i.e.

$$\mathcal{A} = \{B^p \mid \text{ patient } p \text{ has ASD } \}$$

$$\mathcal{C} = \{B^p \mid \text{ patient } p \text{ is a control } \},$$

find the average ROI groupings within each diagnostic group, i.e.:

$$M_a = \frac{1}{|\mathcal{A}|} \sum_{B \in \mathcal{A}} B$$

$$M_c = \frac{1}{|\mathcal{C}|} \sum_{B \in \mathcal{C}} B,$$

and find the difference of the mean groupings i.e. $M_{a-c} = M_a - M_c$. Given a new observation, we use the correlation and clustering process above to generate a sparse grouping matrix B_{test}, then classify this matrix using the following procedure:

$$class(B_{test}) = \begin{cases} 1 & B_{test} * M_{a-c} > 0 \\ 0 & B_{test} * M_{a-c} \leq 0 \end{cases}$$

where $*$ denotes element wise matrix multiplication, 1 means the patient is classified as having ASD, and 0 is control. Note that this classification can be thought of as a perceptron with a binary threshold at 0, input nodes corresponding to the 6670 dimensional vector of pairwise group similarities (the flattened upper triangle of B_{test} excluding the diagonal) and weights corresponding to 2 times the corresponding values in M_{a-c}.

3.2.1 Randomized clustering

We first create a benchmark which uses a completely randomized clustering algorithm. E.g., we first randomly choose a value k for the number of clusters (selected from the set of possible k's used in the real clustering algorithms below) and then randomly assign labels in $\{1, \ldots, k\}$ to the ROIs. The resulting matrices M_a, M_c, M_{a-c} for this process are visualized in Figure 1. We expect the classification procedure to produce completely random results. Indeed, after running the training process 200 times and computing accuracy on the Tuning set, we get a mean accuracy of 49.9% with a distribution displayed in Figure 4.

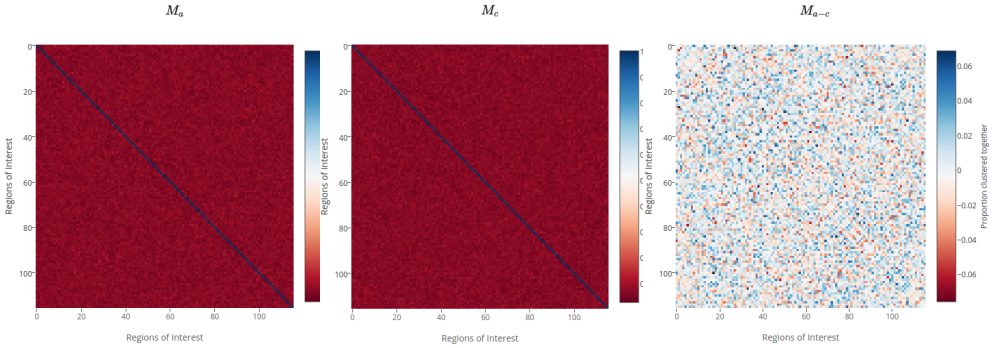

Figure 1: M_a, M_c, and M_{a-c} for Randomized Clustering on Correlation matrices.

3.2.2 Spectral Clustering

Spectral Clustering [16] is a technique to embed a matrix of affinities between samples into a low dimensional space, then perform clustering in that space (in this case, K-Means clustering). To begin, we transform our correlation matrix C into an affinity matrix A using $A = -C + 1$. We then apply a Gaussian Kernel to the matrix A with the to obtain a similarity matrix

$$S = \exp\{-\frac{A^2}{2\gamma^2}\}$$

where γ is the width of the Gaussian Kernel. We then train our model using Spectral Clustering in place of the random clustering function described above.

We first perform a hyper-parameter search over the possibilities described in Table 1.

Table 1: Hyper-Parameters for Spectral Clustering

Parameter	Description	Possibilities
n_clusts	Number of Clusters	5, 10, 15, 20, 25
γ	Gaussian Kernel Width	.25, 1, 4, 10

These parameters were chosen broadly to represent the range of parameters which give sensible results (e.g., no clusters will have only one element, and spatially close brain regions are clustered together often).

After searching over all permutations of these parameters and evaluating accuracy on the Tuning set, we found that the best parameters were n_clusts = 5, $\gamma = 1$, which produced an accuracy of 63%. A distribution of the collection of all the Tuning accuracies found during this search are in Figure 4. We now take the best model from the hyper-parameter search and apply it to the Testing Set. We receive an accuracy of 64%. The resulting matrices M_a, M_c, M_{a-c} for this process are visualized in Figure 2. A visualization of the separation of the diagnostic groups given the evidence of the classifier can be seen in Figure 5

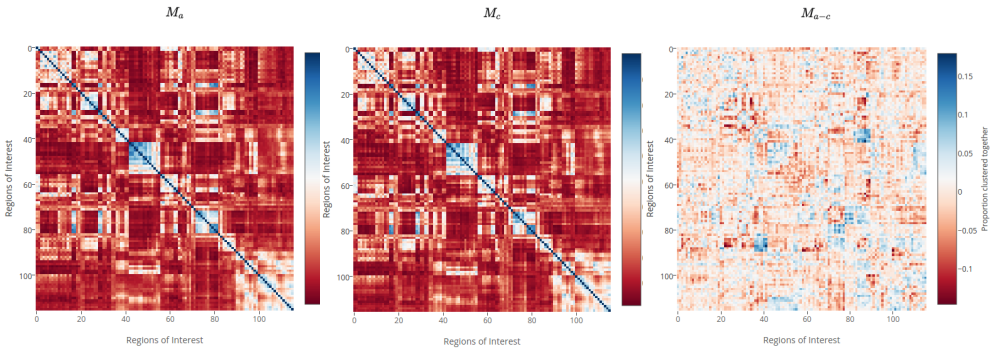

Figure 2: M_a, M_c, and M_{a-c} for Spectral Clustering on Correlation matrices.

3.2.3 Affinity Propagation Clustering

Affinity Propagation (AP) [11] is an exemplar based clustering algorithm which is based on the Belief Propagation message passing algorithm. Because AP considers each data point to be an exemplar on initialization and iteratively refines its model, it does not take a hard number of cluster centers. Rather, AP takes a 'preference' value which is essentially a prior belief of the number of clusters, which can be over-ridden by strong contrary evidence in the data. AP also takes a damping factor between 0.5 and 1 which lowers values of messages passed, encourages convergence, and prevents numerical oscillations. Since AP takes a similarity matrix, we derive one from the correlation matrix C with $S = C - 1$.

We first perform a hyper-parameter search over the possibilities described in Table 2.

Table 2: Hyper-Parameters for Affinity Propagation Clustering

Parameter	Description	Possibilities
preference	Soft Cluster Number Suggestion	-1, -0.75, -0.05, -0.25, 0
damping factor	How much to dampen each successive iteration	0.65, 0.8, 0.95

Once again, these parameters were selected to give a broad range of parameters which give sensible results.

After running through each possible permutation of the hyper-parameters, we find that the setting preference = -1 and damping = .95 had the best results on the Tuning set with an accuracy of 65%. A

distribution of all the Tuning accuracies reported the by various models in this search can be found in Figure 4.

We now take the best model discovered during the hyper-parameter search and apply it to the Testing Set. This produced an accuracy of 61%. The resulting matrices M_a, M_c, M_{a-c} for this process are visualized in Figure 3.

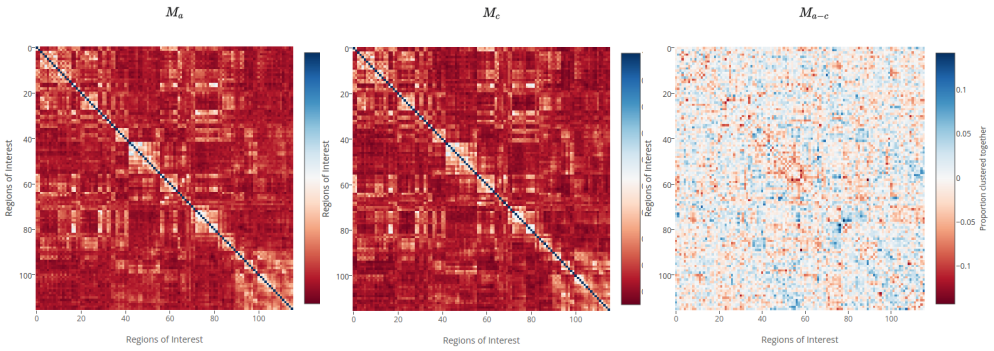

Figure 3: M_a, M_c, and M_{a-c} for Spectral Clustering on Correlation matrices.

3.3 Adding a second view

We utilize the variety of data derivatives made available by ABIDE by augmenting our feature vector with a view from the CC200 atlas.

We consider pairs of views (V_1, V_2). From these views we derive the respective sparse grouping matrices $(B_{V_1}, B_{V_2})_p$ for each patient, and from these we calculate the mean difference matrices $(M_{V_1,a-c}, M_{V_2,a-c})$ in the same way we do above. We then define the feature vectors $(f_{V_1}, f_{V_2})_p$ for patient p as the flattened upper triangular matrix of the respective matrices $(B_{V_1}, B_{V_2})_p$ excluding the diagonal. That is, for each patient we $f_{V_1,p} \in \{0,1\}^{6670}$ and $f_{V_2,p} \in \{0,1\}^{19900}$.

We then reduce the dimensionality of these feature vectors by only considering the dimensions for which the absolute value of the respective $M_{V_i,a-c}$ matrix is above 0.7. This threshold was chosen to be the lowest possible value that created feature vectors with dimensionality lower than the number of observations in the Training Set (i.e. less than 664 dimensions).

We now consider the matrices of reduced dimension feature vectors (f'_{V_1}, f'_{V_2}) and perform CCA [12] on these two feature matrices to produce subspaces (ϕ_{V_1}, ϕ_{V_2}). We then project (f'_{V_1}, f'_{V_2}) onto these subspaces to produce (P_{V_1}, P_{V_2}), the projection matrices. Finally, we stack the projections with the original high dimensional feature vectors to create features for our classifier, i.e. $\mathcal{F} = vertical\_stack(F_{V_1}, F_{V_2}, P_{V_1}, P_{V_2})$ where F_{V_1} and F_{V_2} have respective feature vectors in their columns.

Using these feature vectors we perform hyper-parameter search over the options in Table 3 by training a Perceptron Classifier and evaluating on the Tuning Set. We find that the best parameters are n_clust_1 = 5, n_clust_2 = 5, n_dims_1 = 40, and n_dims_40, with an accuracy of .64. Using these parameters, we re-run the model and replace the Perceptron Classifier with a MultiLayer Perceptron, which we re-train with the Training Set and evaluate on the Tuning Set until it seems to have converged well (e.g., accuracy above 65%). We then run our model on the Testing Set to receive an accuracy of 69%. A separation of the data for this model can be seen in Figure 6.

Table 3: Hyper-Parameters for Multi-View Spectral Clustering

Parameter	Description	Possibilities
n_clusts_1	Number of Clusters for View 1	5, 15
n_clusts_2	Number of Clusters for View 2	5, 15
n_dims_1	Number of Dimensions from ϕ_1	10, 40
n_dims_2	Number of Dimensions from ϕ_2	10, 40

4 Implications of Results

It is well known, as seen in the Previous Work section, that it is possible to use Machine Learning techniques to diagnose ASD using fMRI imaging data. All of these works; however, leverage extremely complicated models such as Multi-Kernel SVMs and 'Projection Based Learning Metacognitive Radial Basis Function Networks', to produce results which are only around 10 percentage points higher than those seen in the simple models used here (sans [13]).

With a belief that this field of connectomic diagnosis should move away from using complicated models to maximize accuracy, and focus on creating simple, robust, and understandable models (which are all the properties required for a model to be accepted and used in a clinical setting), this work was created to prove that simple models can work extremely well. In fact, we can see these results by considering results other than the raw accuracy scores:

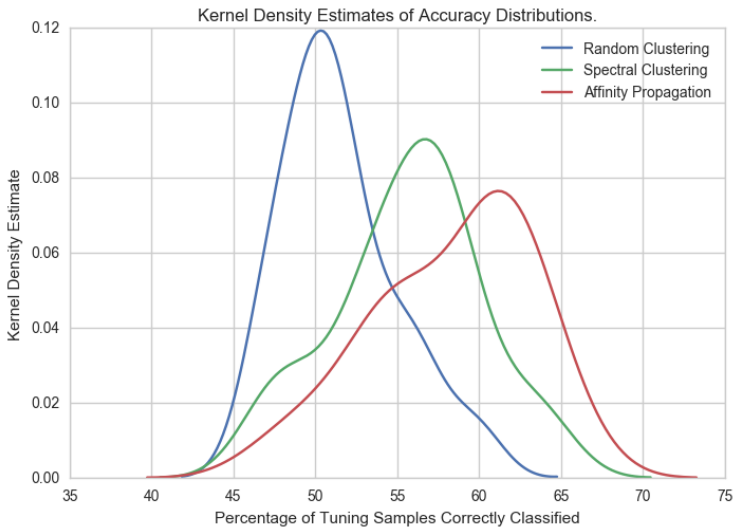

Figure 4: KDE of accuracies reported during Tuning on various models.

Figure 4 was created by collecting the accuracies of the various models encountered during the hyper-parameter searches (or for 200 random runs of the randomized clustering case) and running them them through a Kernel Density Estimator. This is an important result because for difficult classification tasks with small data sets such as this, it is possible for a 'chance' classifier to perform seemingly well. This distribution shows that the accuracies reported are not just luck, but rather because the set we are sampling models from performs systematically above chance.

Figure 5 shows an example of the diagnostic group separations which are causing these models to perform well. We see a clear separation created by a model which simply returns a linear sum of the cluster matrix.

Figure 6 shows a slightly more defined separation than Figure 5 does. Considering that this increase required the addition of a second data view, a powerful non-linear classification model, and the loss of any visual representation of what clustering structures are causing the separation, it may even be that the simple model is preferred.

5 Conclusions and Future Work

In this work we have shown with two very simple models that a linear weighted sum of the upper triangular matrix of the clustering matrix of a functional connectome can be used to systematically differentiate between patients with and without ASD. The best parameter selection of these models resulted in accuracies of 61% and 62% on the testing set, while the other models encountered during hyperparameter search were performing on average above 56% accuracy on the Tuning set. We then add complexity to this model by adding information from a second atlas view and using a non-linear

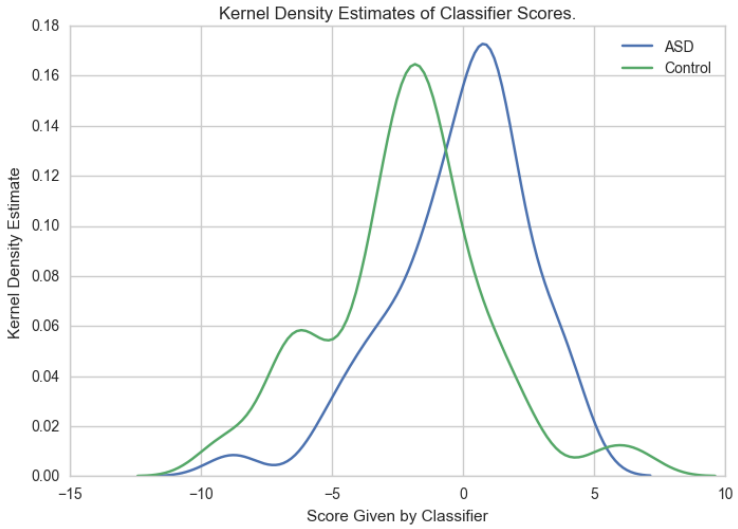

Figure 5: KDE of 'evidence' created by the element-wise multiplication of the B matrices in the Testing Set with the M_{a-c} matrix, in the Spectral Clustering case.

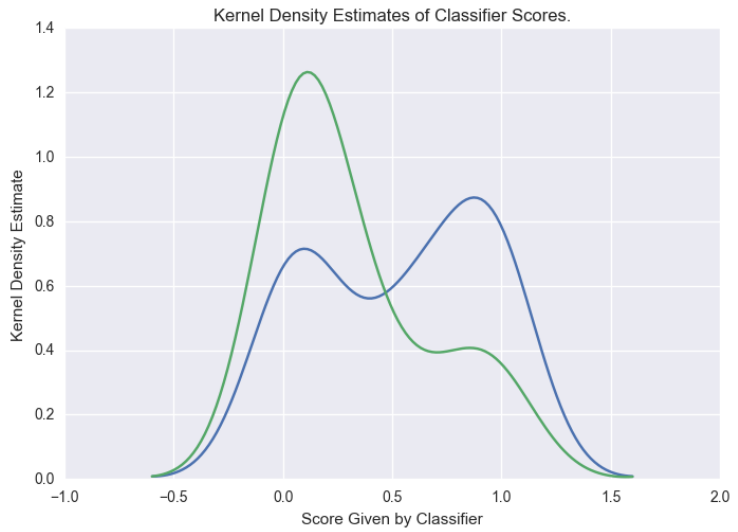

Figure 6: KDE of 'evidence' created by the MultiLayer Perceptron in the Multi-View experiment.

classifier instead of a simple weighted sum to produce a test accuracy of 69%, with other models encountered in the hyper-parameter search having a mean accuracy of 59%.

Some future directions to build on this work may include mapping the correlation matrix clustering views to a 3D representation of the atlas which produced them. This may offer some insights into exactly which regions are contributing the most to the linear sum used in the simpler classification scheme. Also, there are many data derivatives available to download through ABIDE which could make for some interesting generalized multi-view learning models. Finally, conditioning the data-set on categorical variables and studying the performance on each class could be an interesting exercise. Of particular interest is conditioning on gender, since it has been observed that ASD affects males differently than it does females [20].

References

[1] American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disor-*

ders. American Psychiatric Association, 5 2013. ISBN 0-89042-555-8. doi: 10.1176/appi.books.9780890425596. URL http://psychiatryonline.org/doi/book/10.1176/appi.books.9780890425596.

[2] Bharat B Biswal, Maarten Mennes, Xi-Nian Zuo, Suril Gohel, Clare Kelly, Steve M Smith, Christian F Beckmann, Jonathan S Adelstein, Randy L Buckner, Stan Colcombe, Anne-Marie Dogonowski, Monique Ernst, Damien Fair, Michelle Hampson, Matthew J Hoptman, James S Hyde, Vesa J Kiviniemi, Rolf Kötter, Shi-Jiang Li, Ching-Po Lin, Mark J Lowe, Clare Mackay, David J Madden, Kristoffer H Madsen, Daniel S Margulies, Helen S Mayberg, Katie McMahon, Christopher S Monk, Stewart H Mostofsky, Bonnie J Nagel, James J Pekar, Scott J Peltier, Steven E Petersen, Valentin Riedl, Serge A R B Rombouts, Bart Rypma, Bradley L Schlaggar, Sein Schmidt, Rachael D Seidler, Greg J Siegle, Christian Sorg, Gao-Jun Teng, Juha Veijola, Arno Villringer, Martin Walter, Lihong Wang, Xu-Chu Weng, Susan Whitfield-Gabrieli, Peter Williamson, Christian Windischberger, Yu-Feng Zang, Hong-Ying Zhang, F Xavier Castellanos, and Michael P Milham. Toward discovery science of human brain function. *Proceedings of the National Academy of Sciences of the United States of America*, 107(10):4734–9, 3 2010. ISSN 1091-6490. doi: 10.1073/pnas.0911855107. URL http://www.ncbi.nlm.nih.gov/pubmed/20176931http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2842060.

[3] Colin J Brown and Ghassan Hamarneh. Machine Learning on Human Connectome Data from MRI. 2016.

[4] Felicity Callard and Daniel S Margulies. The Subject at Rest: Novel conceptualizations of self and brain from cognitive neuroscience's study of the 'resting state'. *Subjectivity*, 4(3):227–257, 9 2011. ISSN 1755-6341. doi: 10.1057/sub.2011.11. URL http://link.springer.com/10.1057/sub.2011.11.

[5] Eric Courchesne and Karen Pierce. Why the frontal cortex in autism might be talking only to itself: local over-connectivity but long-distance disconnection. *Current Opinion in Neurobiology*, 15(2):225–230, 4 2005. ISSN 09594388. doi: 10.1016/j.conb.2005.03.001. URL http://www.ncbi.nlm.nih.gov/pubmed/15831407http://linkinghub.elsevier.com/retrieve/pii/S0959438805000334.

[6] P. Craddock, R.C. & Bellec. ABIDE Preprocessed. URL http://preprocessed-connectomes-project.org/abide/.

[7] R. Cameron Craddock, G.Andrew James, Paul E. Holtzheimer, Xiaoping P. Hu, and Helen S. Mayberg. A whole brain fMRI atlas generated via spatially constrained spectral clustering. *Human Brain Mapping*, 33(8):1914–1928, 8 2012. ISSN 10659471. doi: 10.1002/hbm.21333. URL http://www.ncbi.nlm.nih.gov/pubmed/21769991http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3838923http://doi.wiley.com/10.1002/hbm.21333.

[8] Luca Dodero, Ha Quang Minh, Marco San Biagio, Vittorio Murino, and Diego Sona. Kernel-based classification for brain connectivity graphs on the Riemannian manifold of positive definite matrices. In *2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI)*, pages 42–45. IEEE, 4 2015. ISBN 978-1-4799-2374-8. doi: 10.1109/ISBI.2015.7163812. URL http://ieeexplore.ieee.org/document/7163812/.

[9] Simon B. Eickhoff, Tomas Paus, Svenja Caspers, Marie-Helene Grosbras, Alan C. Evans, Karl Zilles, and Katrin Amunts. Assignment of functional activations to probabilistic cytoarchitectonic areas revisited. *NeuroImage*, 36(3):511–521, 7 2007. ISSN 10538119. doi: 10.1016/j.neuroimage.2007.03.060. URL http://www.ncbi.nlm.nih.gov/pubmed/17499520http://linkinghub.elsevier.com/retrieve/pii/S1053811907002303.

[10] Amanda Elton, Adriana Di Martino, Heather Cody Hazlett, and Wei Gao. Archival Report Neural Connectivity Evidence for a Categorical- Dimensional Hybrid Model of Autism Spectrum Disorder. 2016. doi: 10.1016/j.biopsych.2015.10.020. URL http://dx.doi.org/10.1016/j.biopsych.2015.10.020.

[11] Brendan J Frey and Delbert Dueck. Clustering by Passing Messages Between Data Points.

[12] Harold Hotelling. Relations Between Two Sets of Variates. *Biometrika*, 28(3/4):321, 12 1936. ISSN 00063444. doi: 10.2307/2333955. URL http://www.jstor.org/stable/2333955?origin=crossref.

[13] Tetsuya Iidaka. Resting state functional magnetic resonance imaging and neural network classified autism and control. *Cortex*, 63:55–67, 2015. ISSN 00109452. doi: 10.1016/j.cortex.2014.08.011.

[14] Yan Jin, Chong-Yaw Wee, Feng Shi, Kim-Han Thung, Pew-Thian Yap, Dinggang Shen, and Infant Brain Imaging Study (IBIS) Network. Identification of Infants at Risk for Autism Using Multi-parameter Hierarchical White Matter Connectomes. pages 170–177. Springer International Publishing, 2015. doi: 10.1007/978-3-319-24888-2{_}21. URL http://link.springer.com/10.1007/978-3-319-24888-2{_}21.

[15] Marcel Adam Just, Vladimir L Cherkassky, Timothy A Keller, and Nancy J Minshew. Cortical activation and synchronization during sentence comprehension in high-functioning autism: evidence of underconnectivity. *Brain*, 127:1811–1821, 2004. doi: 10.1093/brain/awh199.

[16] Ulrike Von Luxburg and Ulrike Von Luxburg. A Tutorial on Spectral Clustering. 2007. URL http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.165.9323.

[17] João Ricardo Sato, Joana Balardin, Maciel Calebe Vidal, and André Fujita. Identification of segregated regions in the functional brain connectome of autistic patients by a combination of fuzzy spectral clustering and entropy analysis. *Journal of psychiatry & neuroscience : JPN*, 41(2):124–32, 3 2016. ISSN 1488-2434. doi: 10.1503/jpn.140364. URL http://www.ncbi.nlm.nih.gov/pubmed/26505141http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4764481.

[18] S. Vigneshwaran, B. S. Mahanand, S. Suresh, and N. Sundararajan. Using regional homogeneity from functional MRI for diagnosis of ASD among males. In *2015 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8. IEEE, 7 2015. ISBN 978-1-4799-1960-4. doi: 10.1109/IJCNN.2015.7280562. URL http://ieeexplore.ieee.org/document/7280562/.

[19] Chong-Yaw Wee, Pew-Thian Yap, and Dinggang Shen. Diagnosis of Autism Spectrum Disorders Using Temporally Distinct Resting-State Functional Connectivity Networks. *CNS Neuroscience & Therapeutics*, 22(3):212–219, 3 2016. ISSN 17555930. doi: 10.1111/cns.12499. URL http://www.ncbi.nlm.nih.gov/pubmed/26821773http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4839002http://doi.wiley.com/10.1111/cns.12499.

[20] L Wing, H.K. Abramowicz, S.A. Richardson, H. Asperger, L. Bartak, M. Rutter, A. Cox, B.H. Brask, A.W.H. Buffery, J.A. Gray, J.A. Corbett, R. Harris, R.G. Robinson, M. Creak, S. Ini, M. DeMyer, M. Engler, J. Fabia, M. Drolette, J. Gould, J. Gould, C. Hutt, T.T.S. Ingram, L. Kanner, L. Kanner, L. Kanner, L. Eisenberg, E.O. Lewis, L. Lockyer, M. Rutter, V. Lotter, V. Lotter, E.E. Maccoby, C.N. Jacklin, E.E. Maccoby, C.N. Jacklin, J. Reynell, D.M. Ricks, L. Wing, H.S. Ross, G. Innes, C. Kidd, M. Rutter, M. Rutter, L. Lockyer, D.C. Taylor, C. Ounsted, L. Wing, L. Wing, L. Wing, L. Wing, J. Gould, L. Wing, J. Gould, S.R. Yeates, L.M. Brierley, L. Wing, S.R. Yeates, L.M. Brierly, and J. Gould. Sex ratios in early childhood autism and related conditions. *Psychiatry research*, 5(2):129–37, 10 1981. ISSN 0165-1781. doi: 10.1016/0165-1781(81)90043-3. URL http://www.ncbi.nlm.nih.gov/pubmed/6945608.