Developing Fingerprints to Computationally Define Functional Brain Networks and Noise

Sochat, V[1]

[1]Program in Biomedical Informatics, Stanford University

ABSTRACT

Objective: Independent component analysis can be used with functional MRI (fMRI) data to extract independent components that encompass a mix of true functional, resting state brain networks and noise. This method is growing in popularity in the field of neuroscience as a data-driven way to distinguish artifact or identify different networks to diagnose neuropsychiatric disorder, however most of this work is not automatized and is reliant on group derived templates and matching metrics that do not scale to large data. This paper aimed to develop robust spatial and temporal features to automatically characterize functional brain imaging data, and to start preliminary work exploring group differences in more detailed subnetworks extracted from the same data.

Methods: An extensive set of 246 spatial and temporal features has been developed to be used to predict 7 sets of labels indicating different types of noise and networks represented in a large set of independent components derived from fMRI data. The method employs an unsupervised learning algorithm to define functional networks at two levels, and a supervised learning algorithm to discovers characteristic features of these networks. Ten-fold cross validation and permutation testing is used to evaluate the models.

Results: Using fMRI datasets from persons with schizophrenia and matched healthy controls, this method successfully distinguishes different types of noisy components for 5 out of 7 of the manually curated standards. Specifically, the model for the standard that encompasses all noise types performs with a cross validation accuracy of .8689 and area under the curve of .9286.

Conclusion: This work demonstrates that noisy components can be computationally defined using spatial and temporal features, and, that automated methods can use these features to filter large data. Extension of this method to derive disorder specific fingerprints of functional networks will allow for the development of automated decision support systems using large, publicly available data.

1 INTRODUCTION

Understanding and subtyping of neuropsychiatric illness remains an unsolved challenge because of the heterogeneity of these diseases, and the complexity of the human brain. The World Health Organization estimates 28.47% of the total years lost to illness, disability, and premature death in the United States are due to these disorders, and that they cost Americans a total of 317.6 billions of dollars annually [3]. Neuropsychiatric disorder, in its simplest form, can be understood as aberrant brain activity that leads to noticeably different behavior and cognition that negatively impacts daily life. Regardless of the etiology of the disorder, in order to infer diagnosis and provide treatment, a comprehensive understanding of what distinguishes aberrant from normal is necessary. How might this difference be measured?

1.1 MEASURING NEUROPSYCHIATRIC DISORDER

Asking people about their thoughts and behavior directly (self-report), measuring behavior with tasks, or observational methods based on checklists (the Diagnostic and Statistical Manual of Mental Disorder 4th ed.) might give insight to a correct diagnosis, however ideally this information should come directly from the source: the human brain [2]. Aberrant function of the human brain, when understood on a computational level, will be the most robust and consistent methodology. What might this brain data look like?

The “best” data would be a recording of the firing (action potential) of every single of the brain’s approximately 100 billion neurons, but current research is limited to single neurons (single cell recording) or small groups (multi-cell recording) [24]. This task is infeasible on a large scale for the obvious reason that is invasive. The next best option is non-invasive brain imaging, such as functional magnetic resonance imaging (fMRI). fMRI allows for the measurement of brain activity on the level of the voxel, typically a 1-3 mm cube with an associated value that reflects a blood oxygen level dependent (BOLD) response of 50K-100K neurons that has been shown to be a strong measure of neural activity [17]. fMRI is not detailed or perfect, but it represents an abstraction of neural activity for small regions of the brain, and is a good way for identifying large-scale patterns of brain function.

1.2 STANDARDS FOR FUNCTIONAL NETWORKS AND NOISE

Identifying these patterns of functional networks from resting BOLD data requires some standard for what constitutes a functional brain network, and there currently exists no such standard beyond manual annotation of network by an “expert” [23]. Spatial templates and matching procedures are commonly used to identify networks of interest from single-subject data; however missing is work to define temporal and spatial features to automatically complete this task. Arguably, this gap in methodology is due to the tendency of the neuroscience community to set extremely stringent criteria on analysis parameters. In this environment, efforts to establish a standard are likely not successful due to lack of agreement about an accepted acquisition protocol, processing pipeline, and the “right” data to use. While the data is noisy and has high variance, patterns in these independent signals do exist, and an effort to break down

these established barriers and approach brain science more abstractly with large data is badly needed. Lack of a perfect "gold standard" for functional networks that might be used for a classifier should not hold back an understanding of aberrant function of the human brain. Manual annotation of networks and noise belonging to a dataset is far from a "gold standard," but it is completely feasible to distinguish components, and will allow for the beginnings of a computational understanding of brain signals. Given the current neuroinformatics landscape, the time for this type of work is now.

1.3 THE NEUROINFORMATICS LANDSCAPE

Large, publicly available databases of resting BOLD fMRI data of neuropsychiatric populations (INDI, NDAR, ABIDE, NITRC) can be utilized with established standards and methods from machine learning to discover patterns of brain function that serve as "biomarkers" of disorder. The infrastructure needed to achieve this goal are 1) an automatic method to extract functional networks and other signals, "components" of the data, 2) standards to classify noise in the data to leave only components that represent neurological signal, and 3) unsupervised approaches to infer diagnosis. This paper addresses the first two points to provide rationale for using computational fingerprints to distinguish noise from real neurological signal. Finally, the third point is briefly explored. Specifically, functional brain primary and sub-networks can be extracted with an automatic approach (Independent Component Analysis, ICA), and different types of noise can be defined by patterns of spatial and temporal features. These models can then be used as filters, leaving functional networks to be used to diagnose neuropsychiatric disorder.

2 METHODS

2.1 DATA

Resting BOLD fMRI was acquired for 53 individuals (29: Schizophrenia, 24 healthy control) with mean age 32 years (37 Male/16 Female) from the MIND Institute (New Mexico). Schizophrenia was chosen as a disorder as significant functional brain differences have been shown to exist [15].

Data were motion-corrected, spatially smoothed with a 6mm full width at half-maximum Gaussian kernel, bandpass filtered (.008 to .1 Hz) and spatially normalized into the standard Montreal Neurological Atlas Space in preparation for probabilistic Independent Component Analysis, performed with MELODIC (Multivariate Exploratory Linear Decomposition into Independent Components) Version 3.10, part of FSL (FMRIB's Software Library) [19].

2.2 INDEPENDENT COMPONENT ANALYSIS

ICA is appropriate for task-free resting BOLD fMRI data because it does not require specification of a design matrix, as is required by the commonly used general linear model (GLM). When applied to four dimensional fMRI data, the data is reshaped into an n x m matrix with n time-points and m voxels flattened into a row from a single 3D image. The data is decomposed into two new matrices, the first including temporal information (time-course components) in columns, and the second including associated, statistically independent and sparse spatial components (whole brain maps) in rows. Each row of this second matrix can be reassembled into a 3D image to visualize the map. Each time-course in the first matrix represents a pattern of signal that the particular voxels contribute over the entire functional scan [14]. The decomposition is illustrated in Figure 1. Abstractly, ICA expresses a mixed brain signal as a linear combination of statistically independent component variables. A fundamental assumption of this method is the independence of different brain signals, and that each component has a distinct spatial map that shares brain anatomy.

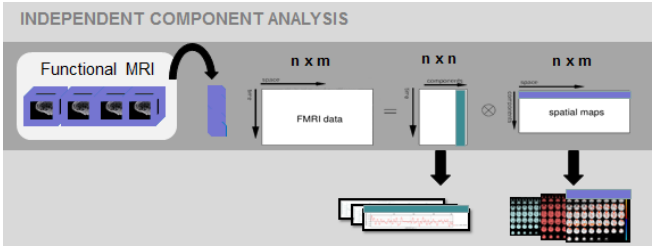

FIGURE 1: An ICA decomposition of 4D fmri data produces a timecourse and spatial maps.

Two levels of ICA were performed, first using the FastICA algorithm [11] to estimate a correct number of dimensions using the Laplace approximation to the Bayesian evidence of the model order [22][6], and second using the highest dimensionality possible given the data size (163 components). FastICA is aimed at achieving maximum degree of non-Gaussianity for all estimated source signals. There are many modifications of these algorithms [17][12] however FastICA is a solid, practical approach that is commonly used and a good choice for this analysis. The resulting data set encompasses two sets of independent components derived on two levels from the equivalent 53 datasets. Components encompass real neurological signal, physiological signal, scanner noise, and artifact.

Level 1 of Independent Component Analysis: The lower level decomposition that reveals "main" brain networks reflects a standard practice in the field, and is important for creating labels. Resulting components are interpretable by a human, and thus can be ascribed with meaningful labels to allow for the creation of a classifier. This decomposition includes 1518 components (ranging from 10-48 per individual, with an average of approximately 25 per individual).

Level 2 of Independent Component Analysis: For the higher level decomposition (representing more detailed signals, "sub-networks,") encompasses 8739 components, expert annotation is infeasible if not impossible. Therefore, it makes sense to build a model of noise using the level 1 of ICA, remove the noise from level 2 with this model, and then use unsupervised clustering to look for patterns in the filtered level 2 data.

For both levels, each component is Z-transformed to allow for comparison by subtracting the mean and dividing by the standard distribution, resulting in voxel values that are Z-scores. Each Z-score map is then thresholded to include the .05 of values in the tails of the distribution. This means that, for any two particular individuals' networks, we are not comparing the values themselves, but rather, comparable degrees of activation from the individual-specific means. This is not problematic because spatial features account for the presence of any significant activation as opposed to the Z-score itself, and temporal features are concerned with the normalized distribution of values as well, and this practice is consistently done in the current literature.

2.3 STANDARDS FOR FUNCTIONAL NETWORKS AND NOISE

The development of a simple Matlab tool allows for the annotation of components derived from ICA. The tool (**Figure 2**) displays the spatial map, associated time course and its distribution for a set of selected components, and outputs a set of labels that works seamlessly with the next stage of analysis, the derivation of a component fingerprint (**Section 2.4**). Using this tool, the entire set of 1518 components for 7 component types, 3 representing specific noise (head motion, white matter artifact and ventricles, eyeballs), 1 representing all noise types, and 3 representing real functional networks (parieto-occipital cortex, primary visual cortex, ventral primary sensorimotor cortex), is manually annotated.

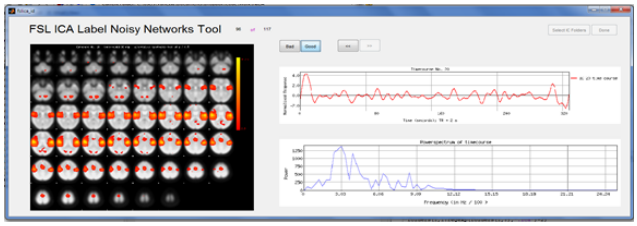

FIGURE 2: Tool to annotate ICA components

2.4 SPATIAL AND TEMPORAL FEATURES

A total of 246 spatial and temporal features (available at: http://www.vbmis.com/bmi/class/cs229/features/nica_features.xlsx) and automatic extraction methods developed based on current literature and intuition were extracted from all components for both levels of decomposition [25][8][20].

2.5 SUPERVISED METHODS TO DEFINE fMRI COMPONENTS

A supervised method, least squares linear regression with the Least Absolute Shrinkage and Selection Operator (LASSO) [21], is utilized to perform both feature selection and classification of the components extracted with ICA using the manually curated labels that define a particular noise type. This modification of least squares regression places a penalty on having more non-zero coefficients, and so it is good for finding sparse solutions. The optimal parameter alpha that controls sparsity is chosen with a grid search, and the optimal lambda is equivalently determined by choosing the value that maximizes the cross validation accuracy. This approach was chosen to identify features of different noise types with the hypothesis that each type can be defined by a small subset of the total features. The chosen features from this step for each component type defined in the main network standard compose a "functional network fingerprint." A binary classifier was chosen because, while it might not be possible to ascribe a meaningful label to every single component, it is entirely feasible to pick out a single noise type or functional network.

2.6 EVALUATION OF SELECTED FEATURES

A ten-fold cross validated receiver operator characteristic (ROC) curve is built into the classification step to evaluate specificity and sensitivity, and a 1,000 iteration permutation test that attempts the equivalent model construction with a random permutation of the labels is used to assess the significance of the cross validated accuracy. It should be noted that this approach is only used to evaluate the labels that comprise all noise types for which there is an equal proportion between the two classes (noise and real neurological signal).

2.7 UNSUPERVISED METHODS TO EXPLORE SUBNETWORKS

Each noise fingerprint will have its own model, or a set of weights applied to a particular set of features to output a class label. The model created for the labels that include all noise types is used to filter level 2 sub-network data, and this filtered data can be explored with unsupervised methods. It should be noted that the sheer number of these components and their "blob-like" nature makes evaluation of the filtered result infeasible, and so exploration of the clustering of disorder type is done knowing this limitation.

From this filtered subset of sub-networks the goal would be to use unsupervised clustering to diagnose neuropsychiatric disorder. Intuitively, sub-networks that cluster together are not necessarily good for distinguishing schizophrenia from healthy control, and so the next goal is to find clusters of sub-networks that are most pure with regard to disorder type. K-means clustering using Euclidian Distance is utilized to select clusters with 80% or greater membership of either schizophrenia patients or healthy control. K-means was performed across 49 values of K, ranging from 15 to 500 with intervals of 10. This threshold and the values of K were chosen arbitrarily. This results in a filtered subset that includes sub-networks that belong to clusters with the strongest class labels, as determined by K-means.

K-Nearest Neighbor (KNN) unsupervised clustering is then employed for each sub-network to ascribe it with a likely diagnosis based on the diagnoses of the nearest neighbors (the percentage of K nearest neighbors that have the diagnosis). Due to the previous step of selecting sub-networks with strong class membership, it was decided to set K=2 to reflect the two classes. For K-Means, since the "correct" value of K is unknown, it was decided to try a method that mimics an ensemble approach, and combine score vectors for the same individual across values of K. This is done with the idea that if a sub-network is particularly good for distinguishing schizophrenia from healthy control, it might appear as a "pure" cluster across multiple values of K, and using its multiple repetitions is akin to weighting it more highly in the final diagnosis.

This final diagnosis score, a value between 0 and 1 that represents the probability of having schizophrenia, is computed for each schizophrenia patient or healthy control based on the average of these scores, with a value of 0 representing a healthy control, and a value of 1 representing a pure schizophrenia patient. The threshold to distinguish the two classes was decided as the mean of the distribution of total scores.

3 RESULTS

3.1 NOISE AND FUNCTIONAL FINGERPRINTS

Subsets of spatial and temporal features were identified to distinguish comprehensive noise, eyeballs, head motion, white matter artifact and ventricles, parieto-occipital cortex, primary visual cortex, and ventral primary sensorimotor cortex, with best cross validation accuracies for the first five of .8689, .9834, .9808, .9675, and .9695**,** respectively. The area under the curve for the set of labels used to filter sub-networks was .9286. Complete results, including component fingerprints, selected features, and

"fingerprints" for the successful models are displayed in Figure 3.

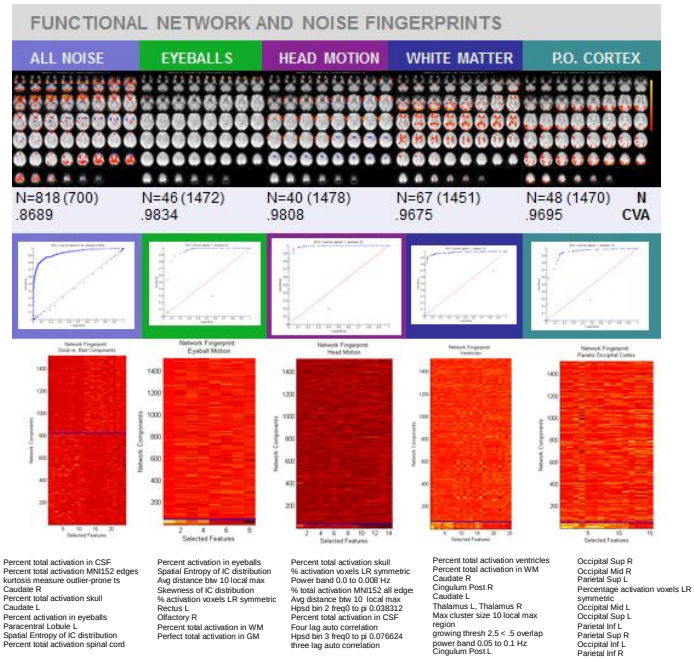

FIGURE 3: Component types, spatial maps, network counts, cross validation accuracy, ROC curve, fingerprints, and top 10 (or fewer) selected features for level 1 ICA components. These models demonstrate the computation signature of fMRI components.

3.2 SUPERVISED METHODS TO DEFINE fMRI COMPONENTS

Filtering the original 8739 components derived with higher dimensionality ICA (the "sub-network" level) with the comprehensive noise classifier resulted in a subset of 3184 components representing real neurological signal. Cluster goodness to potentially choose a particular value of K was evaluated based on mean centroid distance, and lack of a clear "best" choice further supported the decision to combine across values of K. The final calculation of accuracy for this exploratory method with k=2 was 0.5714. Adjusting the threshold of decision to slightly greater than the mean, accuracy improved up to 0.6122.

Evaluation of Pure Clusters Reveals Novel Noise Type: Across 49 values of K there were 1,838 pure clusters comprised of at least 80% of one class. A random sample of 75 of these pure clusters was visually evaluated, and surprisingly, a novel type of noise emerged in many of the samples (Figure 4). Further exploration of entire sets of clusters for a particular value of K made it apparent that this noise appeared consistently across values of K and cluster membership was predominantly schizophrenia patients.

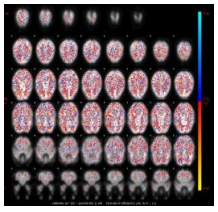

Figure 4: Novel type of noise revealed by K-means clustering of filtered sub-networks

Evaluation of Pure Clusters Reveals Anti-Hubs: A quick glance at a set of sub-networks derived with the higher level decomposition and reading current literature [7] leads one to believe that the majority of components representing real neurological signal are pieces of broken apart "whole" functional networks that might be seen at a lower level decomposition. The visual evaluation revealed clusters of one functional network type that, also surprisingly, did not look broken apart at all. This would suggest that the linear relationship between the voxels in this component is so strong that even forcing the derivation of more components does not split the group into two. Biologically, this reflects an insular, strongly connected functional network, or an "anti-hub." This is an interesting finding that deserves further investigatation, because it might be the case that the degree to which a network can be broken apart is a salient feature to distinguish disorder from healthy control.

4. DISCUSSION

It has been demonstrated that independent component analysis can be used to extract a mix of functional brain networks and noise, and that spatial and temporal features can automatically distinguish network types to allow for automated filtering of fMRI data. Further, this paper provides rationale that interesting, disorder-specific patterns exist on the level of sub-networks, and more work is needed to characterize these differences. Additionally, it was not checked (beyond counting the number of components of each type) that each individual contributed exactly one network. The counts ("N" in Figure 3) suggest that this might be the case, but what is needed is a counting of how many times we see a particular network. These frequency counts would allow for more probabilistic approaches applied to classifying the data.

This ability to automatically ascribe labels to functional networks and noise breaks down the barrier to pursuing data-drive methods for the diagnosis of neuropsychiatric disorder. This type of work is starting to be done with moderate success with structural data for which the standard is simply a standard brain anatomical template. [4].

4.1 SELECTED FEATURES AS VALIDATION

The selected spatial and temporal features serve as unofficial validation of the component. For example, it is expected to see "percent activation in eyeballs" as the mostly highly weighted feature for the eyeballs component, "percent total activation in ventricles" as the top feature for the ventricles component, and "percent activation skull" for the head motion component. From a biological standpoint, these selected features make sense. An interesting observation that has been shown in the literature is the fact that noisy components tend to be defined more-so by time-course features, while functional networks show selection of predominantly spatial features. Additionally, successful models were built for all 4 noise types, while only one of the three functional networks resulted in a successful model. It could be the case that this observation is just chance based on the standards that were created, or it could be the case that the features are not good enough to distinguish the networks. It is salient that the features were developed with the intent of classifying artifact and noise, and so further work is needed to both create more functional standards for testing and developing features that might better classify the networks once noise is removed.

4.2 NEXT STEPS FOR SUBNETWORK EXPLORATION

The unsupervised methods applied to the filtered sub-network data provide impetus for further work in this problem space. The clustering and scoring methods utilized were by no means complicated, and so an accuracy of 0.6122 is surprisingly high given this simplicity.

The discovery of a visually identifiable novel type of noise on the level of the sub-networks that was not seen on the level of the main networks speaks to the fact that the higher dimensionality ICA extracts more "detailed" independent signals that would be mixed with a stronger trend in the data at a lower level decomposition. This finding also provide rational that higher dimensionality ICA is more promising to find subtle group differences. The challenge remains, however, that manual annotation of these networks is infeasible. Although the task is daunting, developing features characteristic of sub-networks would assist in better clustering the networks to identify group differences.

Unsupervised Methods Need More Data: It was decided to derive diagnosis scores by combining across values of K in order to make up for not having enough data at any one value of K. Thus, it is clear that more data is needed.

4.3 LIMITATIONS OF STANDARDS

With the constraints that are currently set in the neuroscience community for what encompasses a gold standard, (i.e. a labeling done by many experts), it would be incredibly challenging and time consuming to entice even one expert to label a set of 1518 networks multiple times. This work was done under the guise that a careful annotation of one experienced individual would be superior to some effort using Amazon Mechanical Turk, or attempting to convince a set of experts to look at small subsets of the data. The standards used for this work in no way represent robust, widely accepted standards; however the point is to show that groups of components intelligently identified by a human being to belong to a particular group in fact can be computationally defined. To pursue this type of work the rules must be changed to allow for imperfection. The neuroscience community must acknowledge that when working with large data, the standard might not be perfect, but the large data will still allow for discovery.

5. CONCLUSION

The definition of standards and features that define different types of noise and functional networks is a move toward the goal of understanding the function of the human brain, and how this function can be aberrant across disorders. On a simple level, machine learning allows us to use our human expertise to teach a computer what encompasses a brain network. We can provide labels for the components that we do understand, and the resulting models can provide further insight to the components that we do not understand. As we develop functional "biomarkers" of disorder, we can further integrate genetic data (currently being developed by the Allen Brain Atlas), structural data, and go as far as making connections between patterns of brain function and emerging trends such as the micro-biome to answer the question of how our brain function relates to who we are. On a speculative and exciting level, logical takeaways from this work include the following:

1. Need for the development of sub-network-specific features
2. Determine frequency of each network type for different disorders (priors) to allow for probabilistic modeling
3. Understanding of which functional networks do not "break apart" between different levels of decomposition, and perhaps how the degree to which they break apart might differ between disorders.
4. Guided ICA (a "with reference approach") to bias the decomposition to add an additional constraint that incorporates prior information when updating the weights [15].

This is a prime time to be in neuroscience. We are not so far away from finding a meaningful difference in structure or function of the human brain for a particular disorder, and then querying the individuals' genome for genes expressed in that region, and then go into the blueprint of the entire machine and making a system-wide fix to actually "cure" or help some of these disorders.

REFERENCES

[1] A. Hyvärinen, R. Cristescu, and E. Oja. A fast algorithm for estimating overcomplete ICA bases for image windows. Proc. Int. Joint Conf. on Neural Networks, Washington, D.C., 1999.

[2] American Psychiatric Association. (2000). Diagnostic and statistical manual of mental disorders (4th ed., text rev.). Washington, DC.

[3] Assessing the economic cost of serious mental illness. American Journal of Psychiatry. 2008 June: 165(6):6 663-5.

[4] Bansal, R., Staib, L. H., Laine, A. F., Hao, X., Xu, D., Liu, J., Weissman, M., et al. (2012). Anatomical Brain Images Alone Can Accurately Diagnose Chronic Neuropsychiatric Illnesses. *PLoS ONE*, *7*(12), e50698. doi:10.1371/journal.pone.0050698

[5] Beckmann, C. F. (2012). Modelling with independent components. NeuroImage, 62(2), 891–901. doi:10.1016/j.neuroimage.2012.02.020

[6] C.F. Beckmann and S.M. Smith. Probabilistic Independent Component Analysis for Functional Magnetic Resonance Imaging. IEEE Transactions on Medical Imaging 23(2):137-152 2004.

[7] Ciuciu, P., Varoquaux, G., Abry, P., Sadaghiani, S., & Kleinschmidt, a. (2012). Scale-Free and Multifractal Time Dynamics of fMRI Signals during Rest and Task. Frontiers in physiology, 3(June), 186. doi:10.3389/fphys.2012.00186

[8] De Martino, F., Gentile, F., Esposito, F., Balsi, M., Di Salle, F., Goebel, R., & Formisano, E. (2007). Classification of fMRI independent components using IC-fingerprints and support vector machine classifiers. NeuroImage, 34(1), 177–94. doi:10.1016/j.neuroimage.2006.08.041

[9] Gian, R. (2010). A novel K-means based multivariate clustering of IC-fingerprints. Frontiers in Neuroscience, 4(2010), 3389–3389. doi:10.3389/conf.fnins.2010.06.00110

[10] Grosenick, L., Klingenberg, B., Knutson, B., & Taylor, J. E. (2011). A family of interpretable multivariate models for regression and classification of whole-brain fMRI data. Most, 94305(650), 1–30.

[11] Hyvärinen, A., Oja, E., 1997. A fast fixed-point algorithm for independent component analysis. Neural Comput. 9 (7), 1483–1492.

[12] Hyvärinen, A., Karhunen, J., Oja, E., 2001. Independent Component Analysis

[13] Jones, D. T., Vemuri, P., Murphy, M. C., Gunter, J. L., Senjem, M. L., Machulda, M. M., Przybelski, S. a, et al. (2012). Non-Stationarity in the "Resting Brain's" Modular Architecture. PloS one, 7(6), e39731. doi:10.1371/journal.pone.0039731

[14] Nandi, A., 1999. Blind Estimation using Higher-order Statistics.

[15] Liu, J., Ghassemi, M. M., Michael, A. M., Boutte, D., Wells, W., Perrone-Bizzozero, N., Macciardi, F., et al. (2012). An ICA with reference approach in identification of genetic variation and associated brain networks. *Frontiers in human neuroscience*, *6*(February), 21.

[16] Neuroscience, H., Calhoun, V. D., Eichele, T., & Pearlson, G. (2009). Functional brain networks in schizophrenia : a review, 3(August), 1–12.

[17] Roberts, S., Everson, R., 2001. Independent Component Analysis: Principles and Practice. Cambridge University Press.

[18] Song, AW, Diffusion modulation of the fMRI signal: early investigations on the origin of the BOLD signal, NeuroImage 62:949-52, 2012.

[19] S.M. Smith, M. Jenkinson, M.W. Woolrich, C.F. Beckmann, T.E.J. Behrens, H. Johansen-Berg, P.R. Bannister, M. De Luca, I. Drobnjak, D.E. Flitney, R. Niazy, J. Saunders, J. Vickers, Y. Zhang, N. De Stefano, J.M. Brady, and P.M. Matthews. Advances in functional and structural MR image analysis and implementation as FSL. NeuroImage, 23(S1):208-19, 2004

[20] Thomas, C. G., Harshman, R. a., & Menon, R. S. (2002). Noise Reduction in BOLD-Based fMRI Using Component Analysis. NeuroImage, 17(3), 1521–1537.

[21] Tibshirani, R. (1996). Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society*, *58*(1), 267–288.

[22] T. Minka. Automatic choice of dimensionality for PCA. Technical Report 514, MIT Media Lab Vision and Modeling Group, 2000.

[23] Wang, Z. Kelly, R. E., Alexopoulos, G. S., Gunning, F. M., Murphy, C. F., Morimoto, S. S., Kanellopoulos, D., et al. (2010). Visual inspection of independent components: defining a procedure for artifact removal from fMRI data. Journal of neuroscience methods, 189(2), 233–45. doi:10.1016/j.jneumeth.2010.03.028

[24] World Book 2001. Chicago: World Book Inc., 2001: 551.

[25] Kelly, R. E., Alexopoulos, G. S., Wang, Z., Gunning, F. M., Murphy, C. F., Morimoto, S. S., Kanellopoulos, D., et al. (2010). Visual inspection of independent components: defining a procedure for artifact removal from fMRI data. Journal of neuroscience methods, 189(2), 233–45. doi:10.1016/j.jneumeth.2010.03.028