

# Classification of Resting State fMRI Datasets Using Dynamic Network Clusters

Hyo Yul Byun<sup>1</sup>, James J. Lu<sup>1</sup>, Helen S. Mayberg<sup>3</sup>, Cengiz Günay<sup>1,2</sup>

<sup>1</sup>Dept. Mathematics and Computer Science, Emory University, Atlanta, GA USA

<sup>2</sup>Dept. Biology, Emory University, Atlanta, GA USA

<sup>3</sup>Dept. Psychiatry and Behavioral Sciences, Atlanta, GA USA

hbyun6@emory.edu

## Abstract

Resting state functional magnetic resonance imaging (rsfMRI) is a powerful tool for investigating intrinsic and spontaneous brain activity. The application of univariate and multivariate methods such as multi voxel pattern analysis has been instrumental in localizing neural correlates to various cognitive states and psychiatric disease. However, many existing methods of rsfMRI analysis are insufficient for investigating the true mechanism of brain activity since they make implicit assumptions that are agnostic of the temporal and spatial dynamics of brain activity. The proposed method aims to create a superior feature space for representing brain activity using k-means and to create interpretable generalizations on these features for studying group differences using support vector machine classifiers.

## Introduction<sup>1</sup>

The robust features of machine learning algorithms make them ideal for studying complex neuroimaging datasets. Resting state functional magnetic resonance imaging (rsfMRI) is a neuroimaging technique that is sensitive to spontaneous and natural correlates of brain activity. It is a popular tool used for investigating the mechanisms of the brain and its various disorders. rsfMRI scans produce a volume time series with an extremely large feature space. Such scans can contain a wealth of information, however, extracting useful information from raw scan data remains challenging. Unsupervised machine learning algorithms provide several methods to reduce high

dimensional datasets into lower dimensions that contain less redundant information. However, reducing dimensions while preserving interpretability and information relevant to classification problems can be difficult. While common univariate analysis of fMRI data is informative in revealing various correlates of the BOLD (Blood oxygenation level dependent) signal, it is insufficient for investigating the deeper layers of systems responsible for the localized correlates found from univariate analysis.

In the proposed method, the unsupervised k-means algorithm was used to find discrete and stable rsfMRI network states that appeared across time in subject scans. These clustered network states were then used to compute new feature spaces for subject rsfMRI scans. This was accomplished by calculating the relative expression of each clustered network for each subject scan. Next, supervised Support Vector Machines (SVM) classifiers were trained to classify between various subject groups within the new feature space for simulated and real rsfMRI datasets containing real subjects. The classification was performed with a rsfMRI dataset containing subject groups with major depressive disorder and healthy controls. The performance of SVM on the new feature space was examined while taking the theoretical usefulness and interpretability of the classifier's generalizations into consideration.

## Methods

The following method first decomposes many resting state fMRI scans into a finite but interpretable set of brain network states by taking advantage of clustering algorithms. Next, it quantifies the expression of these brain state networks for each subject such that their variabilities across subject classes can be studied through the use of a classifier algorithm. The method will be

<sup>1</sup> Copyright © 2014, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

utilized on an rsfMRI dataset collected for investigations into major depressive disorder.

## Subjects

This study was done with anonymized data collected from multiple studies being conducted by the Emory University School of Medicine’s Department of Psychiatry and Behavioral Sciences. The subject groups were the healthy control group (HC), major depressive disorder group (MDD), and treatment resistant depression group (TRD).

## Image Acquisition

Scans were acquired for all subjects using a 3.0 Tesla Siemens Tim Trio human MRI whole body scanner. T1 weighted anatomical scans were collected using an optimized magnetization-prepared rapid gradient-echo imaging protocol (MP-RAGE). The echo time (TE) was 5 milliseconds with a repetition time (TR) of 35. Each TR represents a frame or volume. The resulting image was a 3D matrix with dimensions 256 x 208 x 196 at 1mm isotropic resolution.

The resting state functional magnetic resonance (rsfMRI) images were T2\* weighted echo-planar images. Subjects were ordered to fixate on a crosshair with eyes open. The zSAGA sequence was used (Heberlein & Hu, 2004) in order to minimize sinus-cavity artifacts often seen in fMRI acquisitions. The parameters used were a repetition time (TR) of 2920 ms, echo time (TE) of 35 ms, and flip angle (FA) of 5 degrees. Each resulting image at each time point was a 64 x 64 x 30 dimension image. All scans had at least 140 TRs. The resulting final image format was a 4D 140 x 64 x 64 x 30 DICOM image.

The resting state functional magnetic resonance images were preprocessed using the Analysis of Functional NeuroImages (AFNI) toolkit from the NIMH (Cox, 1996) and the FMRIB Software Library (FSL) from the FMRIB Analysis Group from the University of Oxford in the UK (Smith et al., 2004).

## Regions of Interest Signal Extraction

A set of 40 regions of interest (20 ROIs unilaterally) was selected due to their relevancy in MDD from previous research. Specific ROI coordinates were defined in standard MNI space by an experienced neuroanatomist. The set of these 40 ROIs will be referred to as  $R$  containing  $\{r_1, r_2, r_3, \dots, r_{40}\}$ . For each subject, for each ROI, the BOLD signal was first averaged across all voxels for all 40 ROIs.

## Collection of Dynamic FC Networks

Dynamic functional connectivity matrices were generated and collected for all subjects. The dynamic functional connectivity of an rsfMRI scan of a subject will be a series of functional connectivity matrices across time. Given a rolling window length  $w$ , speed  $v$ , and timeseries length  $l$ ,  $\lfloor \frac{l-w}{v} \rfloor$  is the number of rolling windows available

for sampling in a timeseries of length  $l$ . Some time points at the end may be excluded from the rolling window analysis depending on the three variables. The rolling window length  $w$  and speed  $v$  are adjustable parameters. For this study, the window length  $w$  was chosen as 10 TRs and the speed/overlap  $v$  as 4 TRs. These values define the temporal resolution of the set of dynamic function connectivity matrices for a given subject. The window length and speed/overlap was chosen to minimize leftover TRs and to optimize computation time while maintaining enough spatial resolution for a single window to capture discreet cognitive states. The last four TRs were excluded from the analysis given that the rsfMRI datasets were 149 TRs in length.

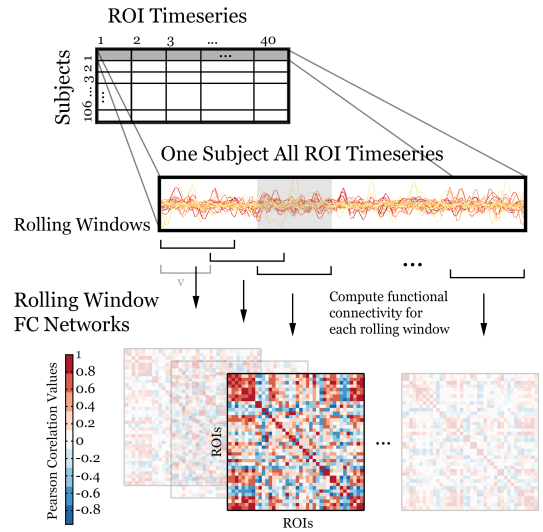


Figure 1: Dynamic Functional Connectivity Computation.

A functional connectivity matrix  $M$  was computed for each subject  $k$  for the  $t^{\text{th}}$  rolling window. Each  $M$  is a two dimensional 40 x 40 matrix. The length of each dimension represents an ROI in our ROI set  $R$ .  $M_{ij}$  is the Pearson correlation coefficient between the timeseries of ROIs  $r_i$  and  $r_j$  for one subject within a single rolling window. Thus,  $M_{ij}$  is the functional connectivity value between ROIs  $r_i$  and  $r_j$ . There will be a total of  $\lfloor \frac{l-w}{v} \rfloor$   $M$ s computed for each subject. Next, all functional connectivity matrices ( $M$ s) generated for all subjects were collated into a single 4 dimensional data structure  $C$ .  $C_{ktij}$  would refer to the functional connectivity between ROIs  $r_i$  and  $r_j$  for the  $t^{\text{th}}$  rolling window for the  $k^{\text{th}}$  subject.

The collection of dynamic functional connectivity matrices  $C$  was resized in preparation for the following analysis. The adjacency matrices representing FC networks were flattened into a single dimension, making the final dimension of  $C$  to be (1802,1600). In summary, the dataset  $C$  represents a large pool of 1,802 cognitive state network observations with a feature space of 1,600

functional connectivity values between all possible pairs of ROIs.

### k-means Clustering of Dynamic FC Windows

Next, the dataset C containing 1802 observations of windowed function connectivity networks were clustered into  $k$  clusters using  $k$ -means. In essence, this clustering step partitions all observed functional connectivity networks from all subjects and rolling windows into  $k$  clusters.  $k$  is searched for during the parameter search step with the leave-one-out cross validation (LOOCV) success percentage as the optimization criteria.

Manhattan distance was chosen over Euclidean distance due to research suggesting that Manhattan distance is a superior metric for high dimensional spaces (Aggarwal, Hinneburg, & Keim, 2010). The output of the  $k$ -means clustering algorithm will include a vertex labeling all input data points (dynamic FC network windows) to one of  $k$  clusters and cluster means (centroids) for the  $k$  clusters. For our analysis, the vertex labeling is discarded as we were only interested in taking advantage of the clustering mechanism of  $k$ -means. The other output of the  $k$ -means algorithm will be a  $k$  by 1600 matrix containing  $k$  centroids. Each centroid can be shaped to form a 40 x 40 matrix where each entry represents a functional connectivity network. The set of centroids can be thought of as an idealized clustering of network states observed in all subject datasets across time.

### Computation of Subject-Centroid Similarities

After clustering, the next step was to compute the similarities between all sets of subjects and dynamic network centroids. The Euclidean distance between the centroid networks and whole scan resting state networks for each subject was utilized. This gave relative metric for the level of expression of the clustered networks for each subject. The consequences of utilizing other distance measures were also examined.

The whole scan resting state networks was generated in a method similar to the dynamic functional connectivity matrices. One matrix was generated for each subject- this is equivalent to a single “rolling” window where window length equals the scan length. The whole scan resting state network is an average of a subject’s functional connectivity network over the entire period of a scan. Each resulting subject’s average resting state network was a 40 x 40 matrix with each entry representing the Pearson correlation coefficient between a pair of two ROIs. This matrix was also flattened to a 1,600 length vector. Now with the average resting state network for all subjects, the Euclidean distance between each subject’s network and the dynamic network centroids was computed. Since there are 106 subjects and  $k$  centroids, this resulted in a 106 by  $k$  matrix. Each entry in this matrix represents a relative metric for the level of expression of a clustered network

for each subject. The subject-centroid similarities were used as the new feature space for the SVM classifier. LIBSVM version 3.17 (Chang & Lin, 2011) was utilized for training and classification. Multi-class classification was performed using the voting method.

## Results

### Subjects

After quality control, a total of 37 healthy controls, 46 non treatment resistant major depressive disorder, and 23 treatment resistant depressive disorder subjects remained. The entire dataset consisted of 106 scans.

### Parameter Search on Subject Data

An initial parameter search was done for parameter  $k$  in  $k$ -means and  $C$  and  $\epsilon$  ranging from 0.01 to 10 in multiples of 10 in order to observe the behavior of the various parameters on the final SVM LOOCV classification percentage. The input dataset included the original subjects split into subject group labels: HC, MDD, and TRD.  $k$ -means was run 10 times in order to account for variability in the non-deterministic solutions. The SVM LOOCV thus was measured 10 times per  $(k, C, \epsilon)$  tuple. In summary, parameter selection was non-trivial.  $C$  was chosen to be 1 since performance gains appeared to be independent of overfitting as the number of centroid networks increased. The  $\epsilon$  parameter did not appear to significantly change LOOCV accuracy when being below a certain threshold. It was chosen to be 0.1.  $k$  was chosen dependent on the elbow point of the LOOCV accuracy and  $k$  curve. Euclidean, Manhattan, and cosine based distances were also used for subject-centroid networks similarities. Manhattan and Euclidean distances performed similarity, and cosine distanced decreased performance.

### Three Way Disease State Classification

The classifier achieved moderate success with a classification accuracy of up to 70.75% for a  $k$  of 29 (Figure 2).

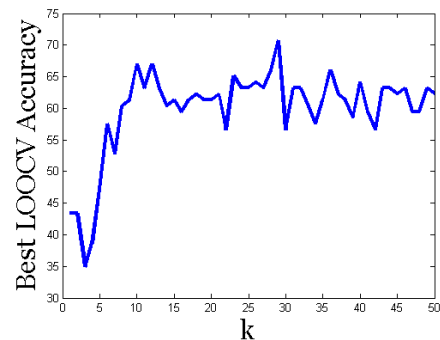


Figure 2: Best SVM LOOCV Accuracy Over 10 Iterations for Ranges of  $k$  For 3 Way Disease State Classification

However, the model created using a  $k$  of 10 (66.98% LOOCV accuracy) was chosen for discussion due to the greater simplicity of the model over the small accuracy gain attained at the cost of greater complexity. Table 1 shows the confusion matrix for the three-way classification. Next, the feature weights for all three classifiers created during the multi-way classifications were computed from the SVM models. Figure 3 shows the unordered 10 centroid networks used for the three way classification. Table 2 lists the centroids ordered by their computed feature weights for each binary classifier.

		Actual Class		
		TRD	MDD	HC
Pred	TRD	13	5	3
	MDD	6	33	9
	HC	4	8	25

Table 1: Confusion Matrix for Three Way Classification

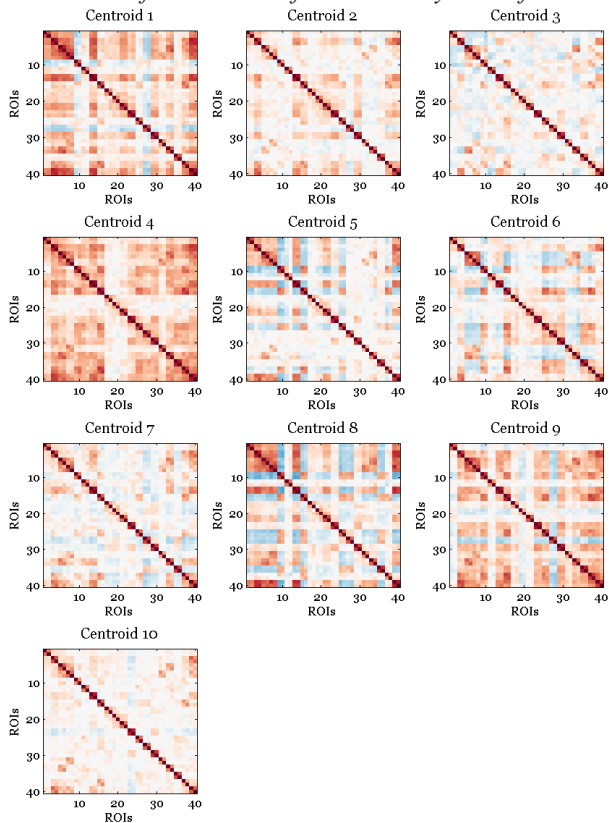


Figure 3: Unordered Centroid Networks

The proposed classifier for resting state fMRI networks classified real world data with moderate success. The  $k$ -means clustering algorithm applied to dynamic functional connectivity ROI networks successfully decomposed the subject data into an interpretable and holistic feature space. Using this new feature space, the SVM classifier was able to successfully create generalizable models. By attempting to model random data using the SVM classifier, it was shown that

TRD vs. HC		MDD vs. HC		TRD vs. MDD	
Centroid	w	Centroid	w	Centroid	w
10	-1.25	5	1.06	10	-1.07
2	-0.8	6	0.94	2	-0.85
6	0.75	9	0.88	4	0.78
7	0.54	8	-0.85	3	0.7
9	0.51	4	-0.62	5	-0.58

Table 2: Top 5 Centroid Networks Ordered by Weight Amplitude for Each Binary Classifier

the classifier was resistant to over fitting (data not shown). LOOCV ensured generalizability of the models. Furthermore, the method of feature space generation and classification escaped traditional fMRI analysis methods.

## Conclusion

A novel method was developed in order to decompose and classify rsfMRI data. The high dimensional rsfMRI data was converted into an alternate feature space with interpretability and theoretical relevance in mind. The dynamic resting state connectivity networks were collected from each subject and clustered using  $k$ -means. This decomposition method provided holistic centroid networks based on dynamic networks observed in subject scans. Next, the expression of each centroid network was computed for each subject- this became the new feature space used for the SVM classifier.

The application of the method on real data on subjects with MDD showed moderate success with 85.85% LOOCV accuracy when classifying between depressed and healthy controls and 70.75% LOOCV when classifying between patients with treatment resistant depression, non-treatment resistant depression, and healthy controls. The generalizations created by the SVM were investigated by examining feature weights from the linear SVM models.

The method discussed in this paper holds promise for its data mining ability for a diverse range of fMRI problems and may hold potential as a clinical tool in treatment selection for psychiatric illnesses. Further investigation into dynamic functional connectivity and other clustering methods is necessary in order to optimize the feature space generation step. Such insights may improve classifier accuracy.

## Acknowledgements

We would like to thank Callie McGrath, Justin Rajendra, and Ki Seung Choi, for their valuable assistance in the research summarized above. The images utilized were collected with support from Emory Healthcare, NARSAD, Woodruff Fund, the Dana Foundation, and the National Institutes of Health [R01MH080880].

## References

- Aggarwal, C. C., Hinneburg, A., & Keim, D. A. (2001). On the surprising behavior of distance metrics in high dimensional space (pp. 420-434). Springer Berlin Heidelberg.
- Chang, C. C., & Lin, C. J. (2011). LIBSVM: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology (TIST)*,2(3), 27.
- Cox, R. W. (1996). AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical research*,
- Heberlein, K. A., & Hu, X. (2004). Simultaneous acquisition of gradient-echo and asymmetric spin-echo for single-shot z-shim: Z-SAGA. *Magnetic resonance in medicine*, 51(1), 212-216.
- Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E., Johansen-Berg, H., ... & Matthews, P. M. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage*, 23, S208-S219.