Multiple-Network Classification of Childhood Autism using Functional Connectivity Dynamics

True Price, Chong-Yaw Wee, Wei Gao, and Dinggang Shen

Department of Radiology and Biomedical Research Imaging Center (BRIC)
The University of North Carolina at Chapel Hill, USA
jtprice@cs.unc.edu, cywee@med.unc.edu, wgao@email.unc.edu,
dinggang_shen@med.unc.edu

Abstract. Characterization of disease using stationary resting-state functional connectivity (FC) has provided important hallmarks of abnormal brain activation in many domains. Recent studies of resting-state functional magnetic resonance imaging (fMRI), however, suggest there is a considerable amount of additional knowledge to be gained by investigating the variability in FC over the course of a scan. While a few studies have begun to explore the properties of dynamic FC for characterizing disease, the analysis of dynamic FC over multiple networks at multiple time scales has yet to be fully examined. In this study, we combine dynamic connectivity features in a multi-network, multi-scale approach to evaluate the method's potential in better classifying childhood autism. Specifically, from a set of group-level intrinsic connectivity networks (ICNs), we use sliding window correlations to compute intra-network connectivity on the subject level. We derive dynamic FC features for all ICNs over a large range of window sizes and then use a multiple kernel support vector machine (MK-SVM) model to combine a subset of these features for classification. We compare the performance our multi-network, dynamic approach to the best results obtained from single-network dynamic FC features and those obtained from both single- and multi-network static FC features. Our experiments show that integrating multiple networks on different dynamic scales has a clear superiority over these existing methods.

1 Introduction

Resting-state functional connectivity (FC) has been proven to be a critical tool in understanding different disease mechanisms and has great potential to provide biomarkers for disease diagnosis and monitoring [6]. Canonical models of altered connectivity among specific regions of the brain have been proposed for a wide range of neurological diseases, including Alzheimer's Disease [8], schizophrenia [7], and autism [12]. In the past, many of these functional characterizations of mental disease have assumed that connectivity patterns in the brain do not change over the course of a resting-state fMRI scan. There is a growing consensus in the neuroimaging community, however, that FC fluctuates in a task-free environment with correspondence to cognitive state [2, 9]. These short-scale modulations in connectivity, which were latent under previous assumptions of stationarity FC, accordingly contain valuable information about functional

organization in the resting-state brain. Utilizing the novel features of FC dynamics, in turn, may allow us to build a better understanding of the effects of neurological disease on brain function.

It is only recently that dynamic FC patterns have been investigated for the purpose of characterizing disease. Ma *et al.* [14] defined group-level intrinsic connectivity networks (ICNs) in the resting-state brain and assessed differences in schizophrenic dynamic FC patterns using a sliding windows approach. Additionally, they used Markov modeling to identify abnormal modulation of brain states within the default mode network (DMN) of the schizophrenic group. Leonardi *et al.* [13] used principal components analysis over sliding windows to uncover differences in whole-brain FC variation patterns (termed "eigenconnectivites") between relapse-remitting multiple sclerosis (RRMS) patients and normal controls. This novel approach revealed altered DMN-related eigenconnectivities among RRMS patients.

In addition to dynamic changes in connectivity on a single network scale, the interaction of multiple large-scale networks has recently become an important topic in brain disease investigation [5, 11]. As the brain's functioning is a product of many concurrent neural patterns, the incorporation of multiple networks into resting-state analysis has great bearing for robust disease characterization. To date, however, there has been little research addressing the properties of dynamic connectivity in disease on a multi-network scale.

In this paper, we combine the dynamic properties of functional connectivity among multiple resting-state networks and explore the merits of such an approach in better classifying childhood autism spectrum disorders (ASDs). Using ICNs defined from group independent component analysis (ICA) on resting-state fMRI data, we investigate sliding window connectivity within individual networks on a large range of time scales. We then apply a multiple kernel support vector machine (MK-SVM) model to evaluate the combination of multiple networks on multiple scales and compare the classification results to those obtained from single-network analysis or under the assumption of stationary functional connectivity over the course of the scan.

2 Methods

Our approach to ASD classification uses multi-network combination of intra-network dynamic connectivity features. Starting with ICNs defined at the group level, we employed linear regression to recover subject-specific time courses for each network. Then, dynamic FC features were extracted for each network using sliding window correlations over a large range of window sizes (20 to 240 seconds, at step sizes of 10 seconds); stationary connectivity features were also collected. Using this wide array of feature types over all networks, we applied an iterative selection/weighting algorithm in a multiple kernel SVM (MK-SVM) model to identify optimal ICNs and window sizes for overall ASD classification. Leave-one-out cross-validation (LOOCV) was used to separate the data into the training and testing sets used in classification. An overview of the classification pipeline is illustrated in Figure 1.

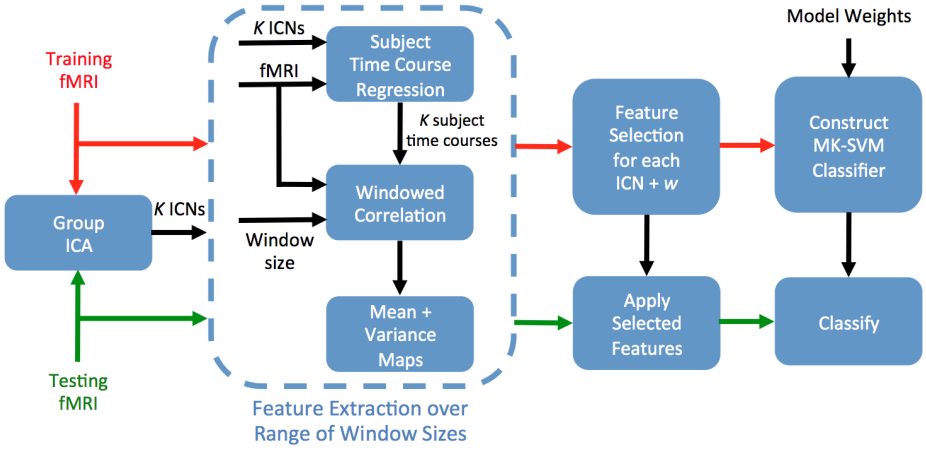

Fig. 1. Framework for the proposed classification pipeline. Subject-specific dynamic connectivity features are derived from group-level ICNs on multiple time scales (i.e. sliding window sizes). Feature selection is then applied to each ICN for every time scale. Finally, features from all ICNs at all time scales are combined in a weighted multiple kernel model.

2.1 Participant Data

Resting-state subject scans were obtained from the open-access Autism Brain Imaging Data Exchange (ABIDE) database [4]. A cohort of 60 child scans, 30 categorized as typical controls (TC) and 30 diagnosed with ASD, were selected from the NYU Langone Medical Center ABIDE site dataset. Mean group ages, in years, were 9.75 ± 1.40 for ASD patients and 9.69 ± 1.58 for controls. Subject ages ranged between 6.5 and 12 years, and individuals were selected to minimize between-group age differences (p = 0.8873). Information about participant data collection, exclusion criteria, and scan parameters for the NYU dataset is available on the ABIDE website[1].

2.2 Data Preprocessing

Initial fMRI scans were collected on a 3-Tesla Siemens Allegra scanner over six minutes taking 180 time points at a repetition time (TR) of 2s. The data were preprocessed using Data Processing Assistant for Resting-State fMRI (DPARSF) software [3]. Before preprocessing, all images had the first ten time points removed. The remaining volumes were then normalized to MNI space with a resolution of $3 \times 3 \times 3$ mm^3. Next, the images were slice timing corrected and motion corrected using the first remaining time point as a reference. White matter, CSF, global signals, and head motion were regressed out as nuisance covariates. Following this, the images underwent signal detrending and band-pass filtering (0.01-0.08Hz). Finally, motion scrubbing [16] was applied with an FD threshold of 0.5; time points with significant motion were removed from each image, along with the preceding time point and the two time points following.

[1] http://fcon_1000.projects.nitrc.org/indi/abide/

2.3 Group ICA and Recovery of Subject Time Courses

In this study, we employed group ICA to define population-based ICNs. Using FSL's MELODIC software[2] [17], all preprocessed subject images – both ASD and normal controls – were temporally concatenated and projected into a 25-dimensional subspace. Spatial ICA was performed on this data set to recover 25 statistically independent spatial maps, each representing a unique group-level functional network, with associated group-level time-courses. After recovering group-level ICNs, we followed the first step of FSL's dual regression approach [1] to define subject-specific time courses associated with each individual network. Namely, we performed linear regression to model each time point of an individual's fMRI scan as a linear sum of the group-level spatial maps.

2.4 Estimating Intra-Network Functional Connectivity

For each subject, the back-reconstruction process yielded a set of 25 time courses, each representing an underlying signal associated with a single group-level functional network. Then, for each subject, we measured the influence of a functional network i on a given voxel v as the normalized cross-correlation of their respective time courses:

$$I(i,v) = \frac{1}{N}\sum_{t=1}^{N}\frac{(T_i(t)-\overline{T_i})(T_v(t)-\overline{T_v})}{\sigma_i\sigma_v}, \tag{1}$$

where N is the length of both time courses, T_i is the subject-specific time course representing network i, T_v is the BOLD signal of voxel v, and $\overline{T}$ and σ represent the mean and standard deviation of a time course, respectively. To analyze the change in network influence over the course of the scan, we split T_i and T_v into synchronous sliding time windows and computed the correlation for each window separately. To observe the effect of time scale on dynamic functional connectivity within the networks, we repeated the experiment on a large range of window sizes, from 10 TR (20 s) to 120 TR (240 s) at intervals of 5 TR. For all window sizes, we fixed the step size between windows to 2 TR (4 s).

Because resting-state fMRI is inherently task-free, it is difficult to interpret patterns of dynamic FC in individuals, and likewise, it is impractical to directly compare changes in FC between subjects. To account for this, we took the voxel-wise mean and variance of windowed correlation values for a given window size and used these measures as comparative features. For a given voxel and a given network, the mean correlation over all time windows gives the average influence of the network within that voxel on a specific dynamic scale. Similarly, the variance gives the local stationarity of the network influence on the same scale. That is, voxels with high variance in correlation can be interpreted to experience some shift in intra-network functional connectivity over the course of the scan.

2.5 Feature Selection

As ICA may return artifactual or physiological components, we first visually inspected the set of 25 components and selected 16 as relevant to functional dynamics, discarding

[2] `http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/MELODIC`

the rest [10]. Then, prior to classification, we performed the following feature selection steps for each ICN: First, we masked each ICN to only include functionally relevant voxels, as determined by MELODIC's post-processing mixture model [17] (using a p-value cutoff of $p < 0.0001$). Next, we applied the Mann-Whitney test [15] to only select mean and variance features where one group exhibited significantly higher values than the other group (thresholded at $p < 0.05$). Using the remaining features, we further applied a logistic regression with L_1-norm regularization to ensure a small number of features were used when constructing linear kernels, which can be strongly affected by noisy features, in our MK-SVM model. Importantly, feature selection, as well as subsequent classifier training, was performed independently on each training set in the LOOCV framework.

2.6 Multiple Network Classification with MK-SVM

From our collection of ICNs, each evaluated at a range of window sizes, our next step is to integrate the features from all networks to perform combined classification. In practice, we expect that only a subset of ICNs actually contain meaningful differences in autism connectivity, and that certain window sizes will provide better disease discriminability than others for ICNs that do have important features. Therefore, we wish to develop a minimal, yet multi-network, multi-scale model of ICN connectivity that increases ASD classification over both single-network analysis and assumptions of FC stationarity. To evaluate the feasibility of such a design, we propose to use an iterative ICN selection/weighting method to find the combination of ICN features that maximizes overall LOOCV classification accuracy, using a multiple kernel SVM (MK-SVM) model [20] for classification.

Given a set of K kernels $\{\phi_k(\mathbf{x})\}$ generated from the training set $\{(\mathbf{x}_i, y_i)\}$, where $\mathbf{x}_i \in \mathbb{R}^{N\times 1}$ is a feature vector and $y_i \in \{-1, 1\}$ is a class label, MK-SVM seeks to find a maximum margin hyperplane in kernel space to separate the two classes. The primal formulation of MK-SVM seeks to solve

$$\min_{\mathbf{w}_k, b, \xi_i} \frac{1}{2}\sum_{k=1}^{K}\beta_k||\mathbf{w}_k||^2 + C\sum_{i=1}^{N}\xi_i \tag{2}$$

$$\text{s.t. } \xi \geq 0, y_i\left(\sum_{k=1}^{K}\beta_k\left(\mathbf{w}_k^T\phi_k\left(\mathbf{x}_i\right)+b\right)\right) \geq 1-\xi_i \text{ for } i = 1, \ldots, n.$$

Here, β_k is a linear weight for kernel k; $\mathbf{w}_k$ and b are the normal vector and intercept defining the hyperplane, respectively; and C is a parameter regularizing the degree of misclassification (we used a default value of $C = 1$ in our experiments). Given a test observation $\mathbf{x}$, we can then make a prediction on its class $\hat{y}$ to be

$$\hat{y} = sign\left(\sum_{k=1}^{K}\beta_k\left(\mathbf{w}_k^T\phi_k\left(\mathbf{x}\right)+b\right)\right). \tag{3}$$

In practice, we can constrain the kernel weights such that $\sum_{k=1}^{K}\beta_k = 1$ and perform a coarse search within the combinatorial space to select an optimal weight configuration.

Since we only wish to select a subset of ICNs out of a large set of network/window size combinations, we employed a greedy forward approach to select kernels one-at-a-time while simultaneously learning their associated MK-SVM weights $\{\beta_k\}$. In this approach, we first construct a linear kernel for each ICN and for each window size, forming the set of kernels $\{\phi_k(\mathbf{x})\}$. We then seek to find a subselection of K kernels that maximizes LOOCV accuracy in the MK-SVM model. Starting with the ICN that gives the best classification accuracy, we incrementally add an additional ICN to the model if its inclusion increases overall accuracy. At the same time, we use a grid search approach to select the optimal weights of selected kernels in the MK-SVM model. For each iteration, we generate all combinations of kernel weights between 0 and 1, using a step size equal to the inverse of the number of kernels in the model. We stop our iterative addition of ICNs when overall classification accuracy can no longer be improved.

3 Results and Discussion

Table 1 gives the best results of dynamic and static FC classification on both single- and multi-network approaches. For the single-network case, we report the highest LOOCV accuracy among all 16 networks using each network's features as input to a single, linear kernel SVM. Under assumptions of FC stationarity, the maximum single-network accuracy was 68%; in our experiments, this network corresponds to the salience network, which under static FC analysis has previously been shown to have a similar rate of disease discriminability children with ASD [19]. For dynamic FC analysis, we found the central executive network (CEN) to have a classification rate of 83% at a window size of 35 TR (70 s). Interestingly, it has been suggested that disruptions in high-order cognitive switching controlled by the CEN could contribute to several neurophysiological disorders, including autism [18]. Figure 2 (left) shows the salience and central executive networks recovered by ICA in our experiments. The right image in Figure 2 plots the classification performance of the two networks across all tested window sizes. We observe from the figure that the salience network has the best performance under static IC analysis, while the CEN has a peak between 30 and 40 TR. This finding is in line with our hypothesis that important FC information exists at different window sizes for different ICNs, and furthermore, it underscores the importance of considering multiple time scales when assessing dynamic FC features in mental disease.

Our results also suggest that multi-network classification, which is able to capture a larger range of disease characteristics, augments both stationary and dynamic analyses.

Table 1. Comparison of the best classification performance obtained under single- and multi-network models with and without dynamic FC features. (ACC=Accuracy, SEN=Sensitivity, SPE=Specificity, PPV=Positive Predictive Value, NPV=Negative Predictive Value)

Method	ACC (%)	SEN (%)	SPE (%)	PPV (%)	NPV (%)
Single-Network, Static FC	68	70	67	68	69
Single-Network, Dynamic FC	83	83	83	83	83
Multi-Network, Static FC	83	87	80	72	86
Multi-Network, Dynamic FC	90	87	93	93	88

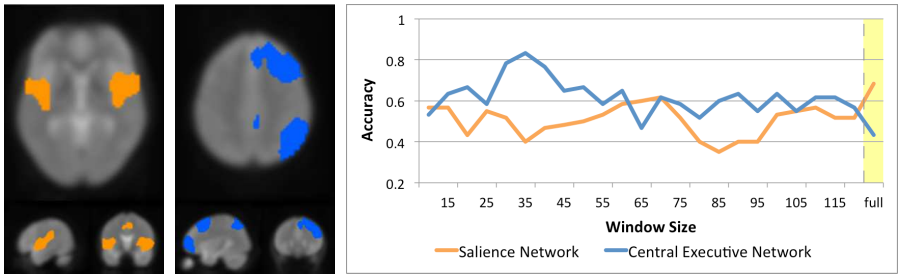

Fig. 2. Left: The salience and central executive networks, which displayed the highest classification accuracy among all 16 networks using static and dynamic connectivity features, respectively. Right: Classification accuracy of the two networks as a function of window size. Here, "full" denotes classification under stationary FC assumptions.

Moreover, compared to only using static FC features, we find that searching over multiple dynamic ranges enhances MK-SVM performance for the combination of ICNs. The multi-network, multi-scale model obtained 90% accuracy using the approach adopted here, substantially outperforming the 83% accuracy obtained by multi-network analysis under assumptions of stationary FC. This suggests that while many ICN-window size pairings may not exhibit strong disease discriminability individually, as evidenced in Figure 2 (right), there may exist important combinations of ICNs at certain time scales that can well characterize disease. Effectively, learning these modulation patterns allows for a whole-brain dynamic model of disease, where both stationary connectivity differences and associated functional compensation are captured. In other words, multiple-network dynamic FC approaches may be able to simultaneously describe both high-level aberrant connectivity and the resulting functional modulation of other brain networks in response. It will be interesting to see what multiple-network models come about in future analyses of the functional dynamics of disease.

4 Conclusion

In this study, we combined dynamic functional connectivity features from multiple networks to enhance the diagnosis of childhood autism. By using FC features over a wide range of time scales, our approach was able to substantially increase ASD classification when compared to using static FC features. Likewise, we showed that integrated network classification using a multiple kernel SVM approach has higher diagnostic potential when dynamic connectivity is considered. From our results, we conclude that incorporating different time scales for different ICNs into multi-network FC analysis provides an important, only now explored, perspective in our understanding of different disease mechanisms.

References

1. Beckmann, C., Mackay, C., Filippini, N., SM, S.: Group comparison of resting-state fmri data using multi-subject ica and dual regression. In: OBHM (2009)

2. Chang, C., Glover, G.H.: Time–frequency dynamics of resting-state brain connectivity measured with fmri. NeuroImage 50(1), 81–98 (2010)
3. Chao-Gan, Y., Yu-Feng, Z.: Dparsf: a matlab toolbox for pipeline data analysis of resting-state fmri. Front. Sys. Neurosci. 4 (2010)
4. Di Martino, A., Yan, C., Li, Q., Denio, E., Castellanos, F., Alaerts, K., Anderson, J., Assaf, M., Bookheimer, S., Dapretto, M., et al.: The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. Mol. Psychiatr. (2013)
5. Elton, A., Alcauter, S., Gao, W.: Network connectivity abnormality profile supports a categorical-dimensional hybrid model of adhd. Human Brain Mapping pp. n/a–n/a (2014)
6. Fox, M.D., Raichle, M.E.: Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. Nat. Rev. Neurosci. 8(9), 700–711 (2007)
7. Garrity, A., Pearlson, G., McKiernan, K., Lloyd, D., Kiehl, K., Calhoun, V.: Aberrant default mode functional connectivity in schizophrenia. Am. J. Psychiat. 164(3), 450–457 (2007)
8. Greicius, M.D., Srivastava, G., Reiss, A.L., Menon, V.: Default-mode network activity distinguishes alzheimer's disease from healthy aging: evidence from functional mri. P. Natl. Acad. Sci. USA 101(13), 4637–4642 (2004)
9. Hutchison, R.M., Womelsdorf, T., Allen, E.A., Bandettini, P.A., Calhoun, V.D., Corbetta, M., Penna, S.D., Duyn, J., Glover, G., Gonzalez-Castillo, J., et al.: Dynamic functional connectivity: Promises, issues, and interpretations. NeuroImage (2013)
10. Kelly Jr, R.E., Alexopoulos, G.S., Wang, Z., Gunning, F.M., Murphy, C.F., Morimoto, S.S., Kanellopoulos, D., Jia, Z., Lim, K.O., Hoptman, M.J.: Visual inspection of independent components: defining a procedure for artifact removal from fmri data. Journal of neuroscience methods 189(2), 233–245 (2010)
11. Kennedy, D.P., Adolphs, R.: The social brain in psychiatric and neurological disorders. Trends. Cogn. Sci. 16(11), 559–572 (2012)
12. Koshino, H., Carpenter, P.A., Minshew, N.J., Cherkassky, V.L., Keller, T.A., Just, M.A.: Functional connectivity in an fmri working memory task in high-functioning autism. NeuroImage 24(3), 810–821 (2005)
13. Leonardi, N., Richiardi, J., Gschwind, M., Simioni, S., Annoni, J.M., Schluep, M., Vuilleumier, P., Van De Ville, D.: Principal components of functional connectivity: A new approach to study dynamic brain connectivity during rest. NeuroImage 83, 937–950 (2013)
14. Ma, S., Calhoun, V.D., Phlypo, R., Adalı, T.: Dynamic changes of spatial functional network connectivity in healthy individuals and schizophrenia patients using independent vector analysis. NeuroImage (2014)
15. Mann, H.B., Whitney, D.R., et al.: On a test of whether one of two random variables is stochastically larger than the other. Ann. Math. Stat. 18(1), 50–60 (1947)
16. Power, J.D., Barnes, K.A., Snyder, A.Z., Schlaggar, B.L., Petersen, S.E.: Spurious but systematic correlations in functional connectivity mri networks arise from subject motion. NeuroImage 59(3), 2142–2154 (2012)
17. Smith, S.M., Jenkinson, M., Woolrich, M.W., Beckmann, C.F., Behrens, T.E., Johansen-Berg, H., Bannister, P.R., De Luca, M., Drobnjak, I., Flitney, D.E., et al.: Advances in functional and structural mr image analysis and implementation as fsl. NeuroImage 23, S208–S219 (2004)
18. Sridharan, D., Levitin, D.J., Menon, V.: A critical role for the right fronto-insular cortex in switching between central-executive and default-mode networks. P. Natl. A. Sci. 105(34), 12569–12574 (2008)
19. Uddin, L.Q., Supekar, K., Lynch, C.J., Khouzam, A., Phillips, J., Feinstein, C., Ryali, S., Menon, V.: Salience network–based classification and prediction of symptom severity in children with autism. JAMA Psychiatry 70(8), 869–879 (2013)
20. Zhang, D., Wang, Y., Zhou, L., Yuan, H., Shen, D.: Multimodal classification of alzheimer's disease and mild cognitive impairment. NeuroImage 55(3), 856–867 (2011)