IDENTIFICATION OF FAULTY DTI-BASED SUB-NETWORKS IN AUTISM USING NETWORK REGULARIZED SVM

Hai Li[1], Zhong Xue[1], Timothy M. Ellmore[2], Richard E. Frye[3], Stephen T. Wong[1]*

[1]Department of Systems Medicine and Bioengineering, The Methodist Hospital Research Institute, Weill Cornell Medical College, Houston, TX; [2]Department of Neurosurgery, University of Texas Health Science Center, Houston, TX; [3]Arkansas Children's Hospital Research Institute, Little Rock, AR
E-mail: zxue@tmhs.org

ABSTRACT

Neural imaging studies of autism spectrum disorders (ASD) have consistently demonstrated deficits in connectivity. In this paper, we propose a new network regularized support vector machines (SVM) method to identify the faulty sub-networks associated with ASD using diffusion tensor imaging (DTI). After constructing the brain connectivity network of each subject using DTI, the SVM-recursive feature elimination (RFE) algorithm is adopted to identify the faulty sub-networks in order to distinguish ASD from typical developing (TD) controls. Since connections in the network are not independent of each other, their topological proximities are incorporated as the network regularization of SVM-RFE by using the graph Laplacian to obtain robust sub-networks. Experiments on both simulated and clinical datasets showed a better performance of the proposed method in faulty sub-network identification, compared with the traditional SVM-RFE method.

Index Terms— DTI, Autism, brain connectivity, SVM, network regularization.

1. INTRODUCTION

Autism spectrum disorder (ASD) is characterized by specific impairments in social interactions, communication, restricted interests, and repetitive behavior [1]. Previous studies suggest that the deficit of anatomic connectivity among different regions of the brain might be the cause of ASD [2]. To exploit these abnormal connections, substantial research interest has been conducted in the past few years [3]. However, most of these studies only examined changes in specific fiber pathways instead of examining systematic difference across the whole brain network between ASD and typically developing (TD) controls. To tackle this challenge, we present a new strategy for identification of faulty sub-networks of ASD from the whole brain network constructed by diffusion tensor imaging (DTI).

When trying to identify the group difference associated with a particular effect or certain clinical measures in a case-control study, mass-univariate statistical analysis followed with false discovery rate (FDR) is commonly used, e.g., voxel-based morphometry (VBM) [4]. However, direct extension of this method to network analysis treats each connection individually and does not consider the network's topological structure. Meanwhile, with the multiple comparison correction, the statistical power of this method is limited, especially when the data has low signal-to-noise ratio (SNR) and/or low contrast-to-noise ratio (CNR). To improve the statistical power, a new network-based statistical (NBS) method was proposed [5]. However, NBS may not handle network connections well if the connections associated with the contrast do not form large spatially connected structures.

Recently, multivariate methods were proposed to overcome the limits of univariate analysis. Wee et al. [6] utilized SVM-RFE to detect faulty regions to identify mild cognitive impairment (MCI) patients. But, one of the problems is that the results are usually noisy and lack spatial coherence. Cuingnet et al. [7] extended the traditional SVM algorithm with spatial regularization to detect the diffusion alteration associated with stroke outcome. Such extension enforces the spatial consistency directly into the SVM by using graph Laplacian regularization. The hypothesis is that spatially adjacent voxels share similar classification outputs.

In this paper, we extend the Laplacian regularization to identify faulty sub-networks in ASD. Specifically, the network topology structure is adopted to define the proximity matrix among different connections, and the graph Laplacian of the proximity matrix is used to regularize the SVM so that RFE can consistently identify the faulty connections. Compared with the mass-univariate analysis method and the NBS algorithm, the proposed network regularized SVM-RFE takes into account the intrinsic multivariate nature of the network, and at the same time, topologic structure information of the connectivity network is embedded directly into the algorithm so that more robust results can be expected even for low SNR or CNR data.

In experiments, simulated network data were first used to evaluate the algorithm, and then real data acquired in ten

ASD and 10 TD controls were used for discovering the faulty sub-network of ASD. The results showed that the proposed algorithm can effectively extract the faulty sub-network by comparing with SVM-RFE.

2. METHOD

2.1. Connectivity Network Construction from DTI

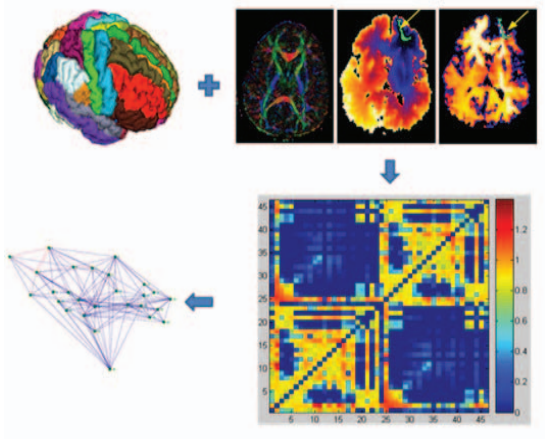


Figure 1. Construction of whole brain connectivity network.

To construct the brain connectivity network, an improved tensor-based fast marching (FM) algorithm [8] was employed to simulate the water diffusion dynamics in DTI image, and the connectivity strength among different anatomical regions of the brain was calculated based on the diffusion speed and diffusion time. Fig. 1 shows the framework of the algorithm. The JHU-DTI-MNI atlas [9] was used for brain labeling, and the connectivity strength between any two regions was defined by a combination of the fast marching results, i.e., "time map" and "velocity map", which starts the diffusion simulation in one region and terminates in another [8]. Faster diffusion between two regions indicates denser, highly oriented fibers, or stronger connectivity between them. In this way, the connectivity strengths among 46 selected anatomical regions from the cortical surface were quantified.

2.2. Faulty Sub-Network Identification Using Network Regularized SVM-RFE

Given two groups of networks our goal is to identify a stable sub-network that robustly distinguishes ASD and TD. The process involves not only group classification but also feature selection. Our work is based on the SVM-RFE framework proposed by Guyon [10]. SVM-RFE has been widely used in many applications, including gene selection, feature analysis and medical image analysis [11]. When treating each link of the network as a feature, SVM-RFE can be used to identify the sub-network associated with the group difference. However, SVM-RFE does not take into account the topological information among different connections. To overcome this disadvantage, the graph Laplacian of the proximity matrix among different connections is used to regularize SVM to improve the consistency of RFE [7]. In this section, we first introduce SVM-RFE and then describe the proposed network regularized SVM-RFE in detail.

SVM-RFE Algorithm

According to the linear SVM, the decision function is:

$$f(\mathbf{x}) = sign(\mathbf{w}\cdot\mathbf{x}+b) = sign(\sum_{i=1}^{N}\alpha_i\gamma_i < \mathbf{x}_i, \mathbf{x} > +b), \quad (1)$$

where $\mathbf{x} = (x_1, \cdots x_g, \cdots x_{|G|})$ is the vector formed by the connectivity strengths of all the network connections of a subject. $|G|$ is the size of the network or the number of links. N is the number of training samples. $\gamma_i \in \{+1, -1\}$ is the class label of each sample. $\mathbf{w} = \sum_{i=1}^{N}\alpha_i\gamma_i\mathbf{x}_i$ is the weight vector of the features. α_i and b are estimated from the training set. Based on Guyon's method [10], each component of $\mathbf{w}$ corresponds to a link in the connectivity network., If the absolute value of the gth element of vector $\mathbf{w}$, $|w_g|$, is small compared, the gth link has little influence on the group classification. If $|w_g|$ is large, the gth link will play an important role on classification. Therefore, the basic idea for RFE is that the value of $\mathbf{w}$ can be represented in original feature space and used to select important features to distinguish the two groups.

Network Regularized SVM-RFE Algorithm

In SVM-RFE, different features are assumed to be independent so the feature selection of RFE may result in unstable sub-network for low SNR data. This is due to the fact that the above mentioned SVM does not consider the relationship among features, or the topological information of the brain\network in our case. In fact, if two links connect a common anatomical region, they are considered as topologically close and are expected to contribute similarly to SVM outputs. By introducing the network neighborhood to encode the proximity of brain connections (details in the next subsection), we can use the graph Laplacian of such a proximity matrix [7] to regularize the SVM. In this way, the weighting vector $\mathbf{w}$ can reflect the topological structure of the connectivity network after training the SVM.

Therefore, after defining the graph Laplacian $\boldsymbol{L}$ of the proximity matrix $\boldsymbol{P}$, the decision function of the network regularized SVM is redefined as:

$$f(\mathbf{x}) = sign(\sum_{i=1}^{N}\alpha_i\gamma_i < \exp\{-\frac{\beta}{2}\mathbf{L}\}\mathbf{x}_i, \exp\{-\frac{\beta}{2}\mathbf{L}\}\mathbf{x} > +b), \quad (2)$$

where β controls the strength of the Laplacian regularization, and $\exp\{\}$ is the matrix exponential operation.

Network Neighborhood System and Proximity Matrix

To define the proximity matrix P of the brain connectivity network, the neighborhood system of the network and the proximity of links are first defined as shown in Fig. 2. If a link x_g connects node m and node n, its connectivity

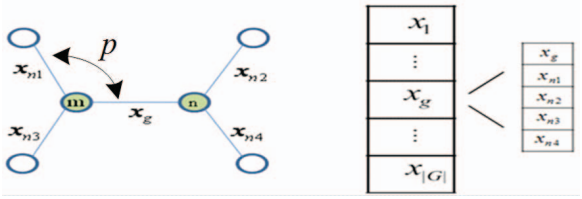
neighborhood $\Phi(x_g)$ comprises all the links directly connecting nodes *m* or *n*. In the case in Fig. 2 (a), the network neighborhood of x_g is $\Phi(x_g)=\{x_g, x_{n1}, x_{n2}, x_{n3}, x_{n4}\}$. Given two links x_n and x_g, their proximity can be defined as the average of the impacts on each other:

$$p_{ng} = p_{gn} = (a_{ng} + a_{gn})/2, \tag{3}$$

and

$$a_{ng} = \begin{cases} x_n \Big/ \sum_{x_i \in \Phi(x_g)} x_i & \text{if } x_n \in \Phi(x_g) \\ 0 & \text{otherwise} \end{cases}. \tag{4}$$

Finally, the proximity matrix $\boldsymbol{P}$ is defined as $\boldsymbol{P}=\{p_{ng}\}$. Then, the graph Laplacian $\boldsymbol{L}$ of $\boldsymbol{P}$ can be calculated and used in SVM according to Eq. (2). The proximity matrix $\boldsymbol{P}$ is defined on the mean connectivity network.



(a) (b)

Figure 2. An illustration of the connectivity neighborhood system and the proximity of the links. The nodes represent anatomical regions, and a link shows that two anatomical regions are connected according to DTI, with connection strength as *x*.

The selection of distinctive connections is based on the ranking of the connections according to the output of the network regularized SVM [10]. Since the number of connections of the faulty sub-networks is not known, we first used all the connections to train the classifier and obtained the weights **w**. This is done by using the leave-one-out strategy. Every time one subject is left out, and SVM is trained using all the others. The subject left out is then used for testing. The success rate is defined by the number of successful classifications of the left-out subjects divided by the number of experiments, and the average weight for each connection is calculated among all the experiments and ranked according to their values. The connections with less average weights are pruned from the feature sets. After pruning we can perform the leave-one-out training and testing strategy again. In this way, we iteratively reduce the number of connections involved in the SVM. The proximity matrix $\boldsymbol{P}$ is updated based on the network topology defined by the remaining connections. We continue this process until the correct classification rate does not increase.

3. RESULT

3.1 Results on Simulated Data

The proposed network regularized SVM-RFE method was applied to 40 simulated dataset (20 controls vs. 20 data with contrast). From the average brain connectivity network *M* of 10 TD controls, we simulated these connectivity data as follows:

Step 1. Randomly choose a node and then initiate a breadth-first search from the selected node to define a subset of connections as the faulty network [5]. Here, we selected 20 links.

Step 2. Gaussian noises are added to all the connections of *M* to simulate the networks of the control group. Denoting the average connection for link *g* as $\bar{x}_g$, the new connection strength will be $x_g = \bar{x}_g + \Delta_g$, where, $\Delta_g \sim N(0, \sigma_g)$ subject to Gaussian distribution.

Step 3. For the contrast group, the new connection is calculated as $x_g = \bar{x}_g + \mu_g + \Delta_g$ for those selected faulty connections, otherwise $x_g = \bar{x}_g + \Delta_g$. μ_g and Δ_g are selected based on the predefined SNR ($\text{SNR} = \bar{x}_g / \Delta_g$) and CNR ($\text{CNR} = \mu_g / \Delta_g$).

Two datasets were simulated with SNR=2, CNR=1 for group one and SNR=3, CNR=1.5 for group two. Then, we performed the following experiments for comparison: 1) FDR corrected mass-univariate analysis; 2) SVM-RFE; and 3) network regularized SVM-RFE on the simulated data.

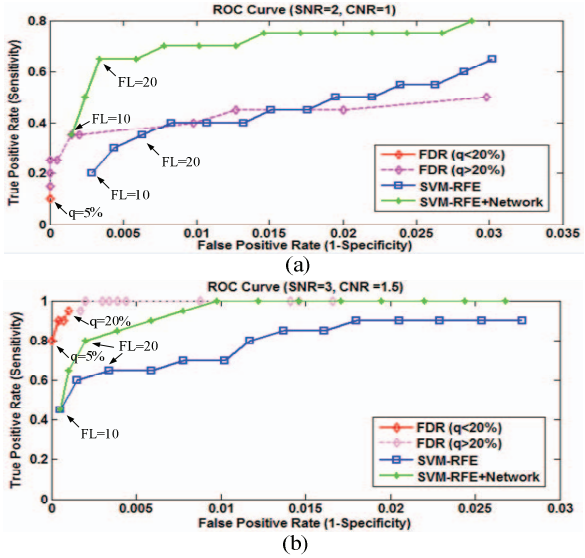


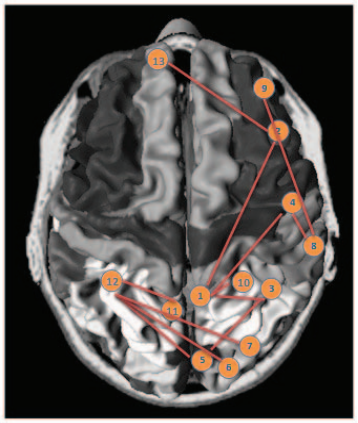
Figure 3. ROC curves for experiments on simulated data. Each point along the ROC curve for the FDR corresponds to a q-value for SVM-RFE and the network regularized SVM-RFE, or a feature level (FL), i.e. the number of connections used, for our method.

Figure 3 (a) and (b) show the ROC curves for groups 1 (lower contrast) and 2 (higher contrast), respectively. For network regularized SVM-RFE, different β value (β=1,2,3) are tested. The best performance is obtained for β=1. In the low contrast situation (Fig 3 (a)), we can see that the network regularized SVM-RFE method performs the best. For FDR, even with very high q-value threshold, it is still too stringent to identify the faulty connections. Usually, the

threshold of q-value is set as 5%-20%, so those points in the ROC curve with high q-value (>20%) are useless in practice, which are shown in dashed red line just for illustration.

For high contrast data in Fig. 3 (b), we can see that both FDR corrected mass-univariate analysis and the network regularized SVM-RFE method can get good results. For example if 10 highest ranked connections are selected, the sensitivity is 45%, and the sensitivity is 90% if 20 connections are selected (similar to q-value=5% for FDR). The results show that our algorithm performs better for low CNS and SNR, and it is comparable with FDR for high contrast simulated data.

3.2 Network Analysis of ASD



1 SPG_L	8 ITG_L
2 IFG_L	9 LFOG_L
3 AG_L	10 SMG_L
4 Ins_L	11 PrCU_R
5 Cu_L	12 SMG_R
6 SOG_L	13 RG_R
7 MOG_L	

Figure 4. Faulty sub-network identified by the proposed method.

Table 1. Comparison of Classification Accuracy (leave-one-out).

Method	Accuracy (%)
SVM-RFE	95
Network regularized SVM-RFE	100

We applied the proposed method on patients' DTI datasets. Ten high functioning ASD children (aged 7-14) and 10 matched TD subjects were used. All the subjects underwent MR imaging examinations at 3T (Philips Intera) in UT Health Science Center, Houston. Each scan consists of T1-weighted MRI and a 32-direction diffusion imaging sequence (TR/TE=8500/67 ms; FA=90 degrees; 2 mm axial slice thickness, max. b-value of 800 s/mm2). After construction of the connectivity network [8], the proposed algorithm was used to identify the faulty sub-network of ASD. The weight for Laplacian is set as β=1. The most distinguishable connections are identified and shown on Fig. 4. The study confirmed that there are widely distributed faulty connections in the whole brain, especially in left superior parietal lobule (region 1) and right supramarginal gyrus (region 12). It is interesting that most of the links in the identified faulty sub-networks are located in the visual and multimodal integration regions.

Once we got the most distinctive subset of the network, SVM was trained on these features, and a leave-one-out method was then adopted to assess the performance with respect to the generalization error. Table 1 lists the results. The accuracy of the classification with network regularized SVM-RFE method can reach 100% in our experiment, which indicates that the automatically identified faulty sub-networks can be used as a neural imaging biomarker for potential computer-assisted diagnosis of ASD.

4. CONCLUSION

Network-based analysis is increasingly important to systematically study the abnormalities of brain connectivity. In this paper, we proposed a new network regularized SVM-RFE algorithm, which embeds the topological structure of the network in identifying the distinct inter-regional connections between TD control and ASD patients. Experiments on simulated and patient data demonstrated its capability in detecting the faulty sub-network.

References

[1] J. S. Verhoeven, P. De Cock, L. Lagae, and S. Sunaert, "Neuroimaging of autism," *Neuroradiology,* vol. 52, pp. 3-14, Jan 2010.

[2] I. Wickelgren, "Neurology. Autistic brains out of synch?," *Science,* vol. 308, pp. 1856-8, Jun 24 2005.

[3] A. L. Alexander, J. E. Lee, M. Lazar, R. Boudos, M. B. DuBray, T. R. Oakes, J. N. Miller, J. Lu, E. K. Jeong, W. M. McMahon, E. D. Bigler, and J. E. Lainhart, "Diffusion tensor imaging of the corpus callosum in Autism," *Neuroimage,* vol. 34, pp. 61-73, Jan 1 2007.

[4] J. Ashburner and K. J. Friston, "Voxel-Based Morphometry - the Methods," *NeuroImage,* vol. 11, pp. 805-821, 2000.

[5] A. Zalesky, A. Fornito, and E. T. Bullmore, "Network-based statistic: Identifying differences in brain networks," *NeuroImage,* vol. 53, pp. 1197-1207, 2010.

[6] C.-Y. Wee, P.-T. Yap, W. Li, K. Denny, J. N. Browndyke, G. G. Potter, K. A. Welsh-Bohmer, L. Wang, and D. Shen, "Enriched white matter connectivity networks for accurate identification of MCI patients," *NeuroImage,* vol. 54, pp. 1812-1822, 2011.

[7] R. Cuingnet, C. Rosso, M. Chupin, S. Lehericy, D. Dormont, H. Benali, Y. Samson, and O. Colliot, "Spatial regularization of SVM for the detection of diffusion alterations associated with stroke outcome," *Medical Image Analysis,* vol. 15, pp. 729-737, 2011.

[8] H. Li, Z. Xue, K. Cui, and S. T. C. Wong, "Diffusion tensor-based fast marching for modeling human brain connectivity network," *Computerized Medical Imaging and Graphics,* vol. 35, pp. 167-178, 2011.

[9] K. Oishi, K. Zilles, K. Amunts, A. Faria, H. Jiang, X. Li, K. Akhter, K. Hua, R. Woods, A. W. Toga, G. B. Pike, P. Rosa-Neto, A. Evans, J. Zhang, H. Huang, M. I. Miller, P. C. M. van Zijl, J. Mazziotta, and S. Mori, "Human brain white matter atlas: Identification and assignment of common anatomical structures in superficial white matter," *NeuroImage,* vol. 43, pp. 447-457, 2008.

[10] G. Isabelle, W. Jason, B. Stephen, and V. Vladimir, "Gene Selection for Cancer Classification using Support Vector Machines," *Mach. Learn.,* vol. 46, pp. 389-422, 2002.

[11] Y. Mao, D. Pi, Y. Liu, and Y. Sun, "Accelerated Recursive Feature Elimination Based on Support Vector Machine for Key Variable Identification," *Chinese Journal of Chemical Engineering,* vol. 14, pp. 65-72, 2006.