

A Diagnosis Model for Early Tourette Syndrome Children Based on Brain Structural Network Characteristics

Hongwei Wen^{1,2a}, Yue Liu^{3,4a}, Jieqiong Wang¹, Jishui Zhang^{3,4}, Yun Peng^{3,4*}, Huiguang He^{1,2*}

¹ State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing, China

² Research Center for Brain-inspired Intelligence, Institute of Automation, Chinese Academy of Sciences, Beijing, China

³ Department of Radiology, Beijing Children's Hospital, Capital Medical University, Beijing, China

⁴ Beijing key Lab of Magnetic Imaging Device and Technique, Beijing Children's Hospital, Capital Medical University, Beijing, China

ABSTRACT

Tourette syndrome (TS) is a childhood-onset neurobehavioral disorder characterized by the presence of multiple motor and vocal tics. Tic generation has been linked to disturbed networks of brain areas involved in planning, controlling and execution of action. The aim of our work is to select topological characteristics of structural network which were most efficient for estimating the classification models to identify early TS children. Here we employed the diffusion tensor imaging (DTI) and deterministic tractography to construct the structural networks of 44 TS children and 48 age and gender matched healthy children. We calculated four different connection matrices (fiber number, mean FA, averaged fiber length weighted and binary matrices) and then applied graph theoretical methods to extract the regional nodal characteristics of structural network. For each weighted or binary network, nodal degree, nodal efficiency and nodal betweenness were selected as features. Support Vector Machine Recursive Feature Extraction (SVM-RFE) algorithm was used to estimate the best feature subset for classification. The accuracy of 88.26% evaluated by a nested cross validation was achieved on combing best feature subset of each network characteristic. The identified discriminative brain nodes mostly located in the basal ganglia and frontal cortico-cortical networks involved in TS children which was associated with tic severity. Our study holds promise for early identification and predicting prognosis of TS children.

Keywords: Tourette syndrome, DTI, network, tractography, SVM-RFE, automatic classification, high accuracy

1. INTRODUCTION

Tourette syndrome (TS) is a childhood-onset neurobehavioral disorder characterized by chronic motor and phonic tics that usually begins between the ages of 3-8 years and occurs more frequently in males than females (ratio 3 to 1) [1]. TS is frequently associated with comorbid psychiatric disorders, such as obsessive-compulsive disorder (OCD), attention-deficit-hyperactivity disorder (ADHD) or other social and behavioral disturbances [2]. The doctors mostly need carrying out additional examination to suppress interference of other diseases. TS is typically diagnosed by observing symptoms and by evaluating the history of their onset. To date, only clinical measures have been developed to diagnose TS and the TS diagnosis remains somewhat subjective. Thus, studies of objective diagnostic method that can capture the underlying brain morphological features of an individual are of great importance.

Tic generation has been reported to link to disturbed networks of brain areas involved in planning, controlling and execution of action, particularly structural disorders in the striatum and cortico-striato-thalamo-cortical circuits [3]. Recent studies have suggested that the structural networks of the human brain can be constructed by using diffusion MRI and tractography methods [4]. Some studies used the probabilistic tractography algorithms to quantify the structural

^a These authors contributed equally to this work. * Correspondence to:

Huiguang He: Research Center for Brain-inspired Intelligence, Institute of Automation, Chinese Academy of Sciences, Beijing, 100190, China. E-mail: huiguang.he@ia.ac.cn

Yun Peng: Department of Radiology, Beijing Children's Hospital, Capital Medical University. No.56 Nanlishi Road, West District, Beijing, 100045, China. E-mail: ppengyun@yahoo.com

integration of CSTC circuits, and showed widespread structural abnormalities in CSTC white matter pathways in TS patients [5]. However, these findings are mainly obtained based on group-level statistical comparison, and thus are of limited value for individual-based disease diagnosis.

Over the past decade, many machine learning and pattern classification methods have been used for disease diagnosis [6, 7]. Moreover, many efforts have been made recently to investigate brain structural network for the diagnosis purpose [8]. However, few of them are related to TS. The aim of our work is to take a first step toward this goal and make diagnostic predictions about individuals using deterministic tractography methods and graph theoretical modeling of brain connectivity. We aimed to select the topological characteristics of structural network which were most efficient for classification and establish the effective diagnosis models for early TS children.

2. MATERIALS AND METHODS

2.1 Subjects

44 TS patients were recruited from outpatient clinics in Beijing Children's Hospital from July 2012 to May 2015 (age: 8.98 ± 3.114 years, range: 3–16 years; 11 female). All the patients met DSM-IV-TR (Diagnostic and Statistical Manual of Mental Disorders, 4th Edition, text revision) criteria for TS. We also included 48 age and gender matched health controls in our study (age: 11.00 ± 3.495 years; range: 3–17 years; 17 female). We used a clinical interview and the Children's Yale-Brown Obsessive Compulsive Scale (CY-BOCS) to diagnose OCD and used the German short version of Wender Utah rating scale (WURS-k, translated to Chinese) to diagnose ADHD. Patients fulfilling OCD criteria or other comorbidities were excluded from the study. Tic severity for all patients was rated using the Yale Global Tic Severity Scale (YGTSS) and ranged from 10 to 79 ([mean \pm SD]: 46.50 ± 18.037). The duration of TS ranged from 3 month to 5 years ([mean \pm SD]: 1.81 ± 1.423 years). For those who had course less than 1 year, TS diagnosis was made by follow-up call. After the study was approved by Beijing Children's Hospital review board, written informed consent was obtained from all the parents/guardians according to the Declaration of Helsinki.

2.2 Data acquisition

Magnetic resonance imaging was acquired using a 3.0T MR scanner (Gyrosan Interna Nova, Philips, Netherland). Head positioning was standardized using canthomeatal landmarks. The head was stabilized with foam pads to minimize head movements. Patients were instructed to suppress tics and minimize head movements during scanning as much as possible. Axial three-dimensional diffusion tensor imaging (DTI) was acquired from all the subjects. DTI was performed using the following protocol: spin-echo diffusion-weighted echo-planar imaging sequence, 2mm slice thickness, no inter-slice gap, repetition time = 4300ms, echo time = 95ms, field of view (FOV) = 255×255 mm, reconstructed image matrix = 336×336 , 30 non-collinear directions with a b value of 1000 s/mm^2 . 3D T1-weighted imaging were performed with axial three-dimensional-Fast Field Echo (3D FFE) sequence with the following parameters: repetition time (TR) = 25ms, echo time (TE) = 4.6ms, flip angle = 30° , reconstructed image matrix = 256×256 , field of view (FOV) = 200×200 mm, slice thickness = 1mm.

2.3 Preprocessing

Following DTI and T1-weighted data acquisition, we used the FMRIB's Diffusion Toolbox (FDT2.0) within FSL v4.1 (<http://www.fmrib.ox.ac.uk/fsl>) for DTI processing. For each participant, 30 DTI volumes with 1000 s/mm^2 b-value were first affinely registered to the b0 volume for correction of eddy current distortion and simple head motion. Non-brain voxels were removed using Brain Extraction Tool (BET) of FSL; a fractional intensity threshold of 0.25 was selected, resulting in a brain-extracted 4D image and a binary brain mask for each subject. We then used the eddy-corrected 4D data and corresponding brain mask to fit the diffusion tensor model at each voxel by using the FDT. Eigenvalues of diffusion tensor matrix ($\lambda_1, \lambda_2, \lambda_3$) were obtained and maps of fractional anisotropy (FA) were generated. T1-weighted images were brain extracted to exclude the non-brain tissues and reoriented with the origin set close to the anterior commissure (AC) using SPM8 package (<http://www.fil.ion.ucl.ac.uk/spm>).

2.4 Brain structural network construction

The whole brain fiber bundles linked different cortical regions form a huge complicated network. The most basic element network nodes and edges are defined as follows.

Network node definition: The entire brain is divided into multiple regions using the Automated Anatomical Labeling (AAL) atlas, where each region represents a network node[9]. Specifically, the individual FA image in native space was

co-registered to its corresponding T1-weighted image using an affine transformation. The individual structural image was then non-linearly registered to the ICBM152 template. Based on the resultant transformations in these two steps, an inverse warping transformation from the standard space to the native DTI space can be obtained. AAL template in the standard space was then inversely warped back to individual native space by applying this inverse transformation[10].

Network edge definition: All the reconstructed fiber bundles in native space were taken into account for the composition of the network edge. We constructed networks using deterministic tractography. The whole brain fiber tracking was performed via the Fiber Assignment by Continuous Tracking (FACT) algorithm[11] by seeding from the center of each voxel, with the FA threshold of 0.2 and tracking turning angular threshold of 45°. For every pair of brain nodes/regions defined above, fibers with two end-points located in their respective masks were considered to link the two nodes. After defining the network edges, both the weighted and binary network analyses were performed. For the weighted networks, we defined the number, mean FA and averaged length of the connected fibers between two regions as the weights of the network edges. For the binary networks, we considered the existence/absence of fiber bundles in which the network edges were defined as 1 if the fiber number between the two regions was larger than the threshold ($T = 3$ in our case) and as 0 otherwise. As a result, for each participant, there were four different kinds of networks, namely fiber number (FN)-weighted, FA-weighted, fiber length (FL)-weighted and binary network, each of which was represented by a symmetric 90*90 matrix (Figure 1 showed the flowchart of structural network construction).

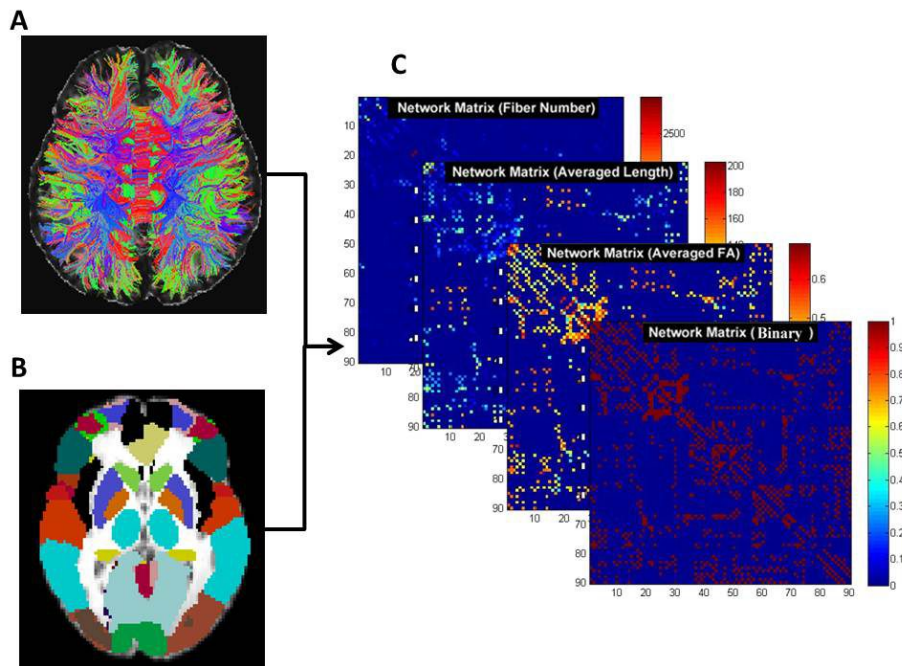


Figure 1. Flowchart for constructing brain structural networks using diffusion tractography. (A) White matter tracts reconstructed using deterministic tractography. (B) Parcellation of gray matter in diffusion space. Each color represents a node in a brain network. (C) Four different kinds of white matter connectivity maps.

2.5 Feature extraction

For each weighted or binary network, we computed three nodal (regional) characteristics of the structural networks, the nodal degree $D_{\text{nodal}}(i)$, nodal efficiency $E_{\text{nodal}}(i)$ [10] and nodal betweenness centrality $B_{\text{nodal}}(i)$ as the features. Classification between groups was undertaken using each characteristic of structural network separately in order to determine the most efficient characteristic for classification.

$$D_{\text{nodal}}(i) = \sum_{i \neq j \in G} w_{ij} \quad (1)$$

$$E_{\text{nodal}}(i) = \frac{1}{N-1} \sum_{i \neq j \in G} \frac{1}{L_{ij}} \quad (2)$$

$$B_{nodal}(i) = \sum_{s \neq i \neq t \in G} \sigma_{st}(i) \quad (3)$$

where w_{ij} is the edge weight between node i and node j in G , L_{ij} is the shortest path length between node i and node j in G , $\sigma_{st}(i)$ is the number of shortest paths from node s to node t passing through node i . $D_{nodal}(i)$ is the sum of edge w_{ij} linking to node i . $E_{nodal}(i)$ measures the average shortest path length between a given node i and all of the other nodes in the network. $B_{nodal}(i)$ measures the number of shortest paths that pass through a node.

2.6 Feature selection and classification

For each feature set, we employ a wrapper-based method for feature subset selection utilizing SVM based on recursive feature elimination which is named SVM-RFE[12]. In this algorithm, SVM is trained iteratively using selected feature subset. In each iteration, the ranking score for each feature in the selected feature subset is calculated during the SVM training process. The feature with the smallest score is eliminated in each iteration of SVM training until the classification accuracy is over a set threshold, or the number of remaining features in the selected subset is smaller than a set value. Note that SVM-RFE uses the accuracy of 10-fold cross validation (CV) to estimate the goodness of feature subset, which may avoid the overfitting problem.

Classification was then performed using the SVM algorithm sequential minimal optimization (SMO) with a radial basis function (RBF) kernel. For the SVM we needed to define two parameters including the complexity or cost constant C and the radial basis function kernel width ($\gamma > 0$). The parameter $C > 0$ determines the trade-off between margin maximization and training error minimization for the soft margin SVM. To estimate optimal values for C and γ we used a grid search in the range of $C = 2^{-2}, 2^{-1}, \dots, 2^{10}$ and $\gamma = 2^{-10}, 2^{-9}, \dots, 2^{-2}$, with 10-fold CV, which we performed for each feature set separately.

Finally we merged the best feature subset of each network characteristic at peak performance level of classification. That means that the feature vector for each subject x_i was constructed by concatenating the nodal topological characteristics for each network: $x_i = [\text{node}_j^{(FN)}, \dots, \text{node}_k^{(FN)}, \dots, \text{node}_l^{(FA)}, \dots, \text{node}_m^{(FA)}, \dots, \text{node}_n^{(FL)}, \dots, \text{node}_o^{(FL)}, \text{node}_p^{(Binary)}, \dots, \text{node}_q^{(Binary)}]$. The same feature selection and classification procedure as the single type of feature was then applied to the combining features.

In our study, a nested 10-fold cross-validation strategy was used to evaluate the classification performance. For outer CV, the data is randomly divided into 10 parts in which each class is represented in approximately the same proportions as in the full dataset. Each fold is held out in turn and the learning scheme trained on the remaining nine-tenths and the error rate is then calculated on the tenth fold. Following 10 training procedures, we calculated the average CV accuracy and considered it as the estimation of generalization [13]. The optimal SVM parameters were estimated on the training samples with another 10-fold CV, which is called inner CV. This nested CV strategy can yield an unbiased assessment of the classification method and prevent overestimation. Specifically, this nested CV procedure was repeated five times in our study to avoid any bias introduced by randomly partitioning in the cross-validation. Figure 2 shows the flow chart of the evaluation method we used for nested CV.

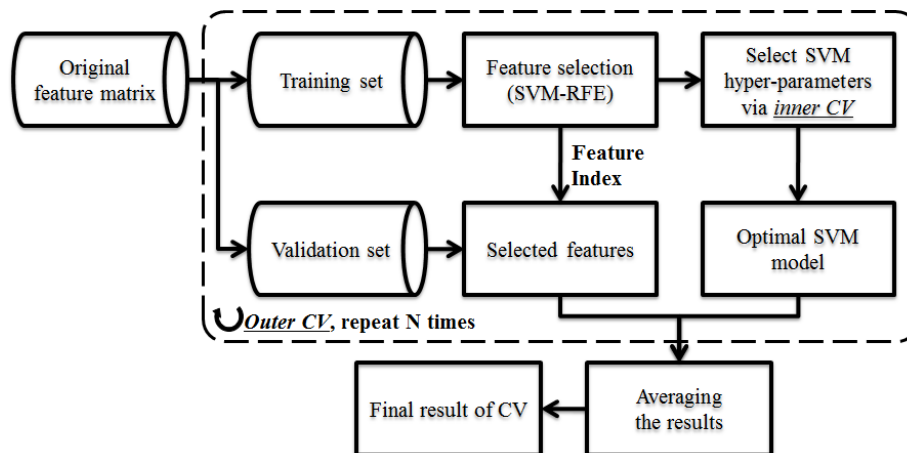


Figure 2. The flow chart of the nested CV classification method using a single type of feature.

2.7 Evaluation

The statistics we used to evaluate our classification algorithm performance are accuracy, sensitivity, specificity and the area under the curve for the receiver operated characteristic curve (AUC ROC). Accuracy is defined as $(TP+TN)/(TP+TN+FN+FP)$ where TP = True Positive, TN = True Negative, FP = False Positive and FN = False Negative. Sensitivity is defined as $TP/(TP+FN)$ and Specificity is defined as $TN/(FP+TN)$.

3. RESULTS

3.1 Classification results using a single network characteristic

To evaluate the most efficient network characteristic on the TS classification, we compared the nested CV performance with a single type of network feature. For each characteristic of structural network, the best feature subset achieving the highest accuracy for classification was shown in Table 1. For the FN-weighted, FA-weighted, FL-weighted and binary network, the highest accuracy were achieved respectively on the characteristic of FN-nodal degree (82.17%), FA-nodal efficiency (78.04%), FL-nodal betweenness centrality (71.30%), Binary-nodal degree (73.91%) (Figure 3). The optimum parameters (C and γ) were estimate using grid search.

Table 1. The classification accuracy for each characteristic of structural network.

Feature	Accuracy(%)	Sensitivity(%)	Specificity(%)	AUC(%)
FN-degree	82.17±2.12	79.55±5.57	84.58±1.86	87.37±2.67
FN-efficiency	71.74±2.55	66.82±3.45	76.25±5.02	78.73±0.81
FN-betweenness	73.48±5.25	73.64±2.59	73.33±8.12	74.70±3.78
FA-degree	74.57±1.65	72.27±4.37	76.67±2.72	75.91±1.13
FA-efficiency	78.04±3.39	75.91±1.24	80.00±6.35	82.05±1.86
FA-betweenness	65.00±4.57	55.00±7.43	74.17±4.80	66.39±5.02
FL-degree	62.39±5.25	56.36±5.88	67.92±5.63	65.45±3.94
FL-efficiency	66.96±5.47	61.82±5.88	71.67±8.01	70.94±5.52
FL-betweenness	71.30±0.60	66.36±2.49	75.83±1.86	74.08±0.98
Bi-degree	73.91±1.09	69.09±3.45	78.33±1.86	76.85±1.95
Bi-efficiency	71.96±2.60	66.82±4.43	76.67±4.01	72.85±2.77
Bi-betweenness	65.00±4.44	60.45±5.23	69.17±7.42	65.33±2.51
multi-features	88.26±1.79	82.73±2.59	93.33±1.74	94.77±1.83

Abbreviation: FN: fiber numbers weighted network. FA: mean FA values weighted network. FL: averaged fiber length weighted network. Bi: Binary network. Degree/efficiency/betweenness: nodal degree/efficiency/betweenness centrality.

3.2 Classification based on combining features

In order to evaluate the benefits of integrating multiple network characteristic on the TS classification, we compared its performance with single network characteristic. The highest accuracy reaches 88.26% on the combing feature set. This suggests that integrating multi-features is a powerful classifier. The classification results using multi-features are also listed in Table 1 and Figure 3, and the ROC curve can be found in Figure 4. We can find that the ROC curve of multi-features classifier is better than single-feature based classifiers.

3.3 The most discriminative network node regions

The current work also identifies the node regions which were most efficient characteristic for early TS classification. We summed the counts of each feature selected by our proposed method for each network characteristic. We summed the counts of each feature selected by our proposed method over the 5 rounds nested 10-fold CV. The top 19 selected features with the frequency more than 70% were provided in Table 2. We also used two-sample t-test to evaluate the variation trend of features in the selected node regions and whether each feature was significantly different in TS group compared to control group. The most discriminative node regions located mainly in bilateral angular gyrus and superior frontal gyrus, right amygdala, superior occipital gyrus and inferior frontal gyrus, left lenticular nucleus, puteman,

parahippocampal gyrus, median cingulate and paracingulate gyri. It is worth noting that multiple features in the left lenticular nucleus, putamen, median cingulate and paracingulate gyri ranked within the top 19. It indicates that these regions may be highly related to TS pathology. For the purpose of visualization, Figure 5 illustrates the most discriminative network nodes.

Table 2. The top 19 discriminative network features selected by the nested CV method as the most salient for group classification.

Region	Feature	Num	Rate	Trend	P	Abbreviation
Right amygdala	FL-efficiency	50	100%	-	0.024	AMYG.R
Left lenticular nucleus, putamen	FL-efficiency	49	98%	+	<0.001	PUT.L
Left angular gyrus	Bi-betweenness	49	98%	-	0.003	ANG.L
Right rolandic operculum	FA-degree	48	96%	+	0.024	ROL.R
Right superior frontal gyrus, dorsolateral	FA-betweenness	47	94%	-	0.024	SFGdor.R
Left parahippocampal gyrus	FL-efficiency	47	94%	+	0.037	PHG.L
Right angular gyrus	FL-efficiency	45	90%	-	N.S.	ANG.R
Right inferior frontal gyrus, opercular part	FL-degree	43	86%	+	0.019	IFGoperc.R
Right superior occipital gyrus	FL-degree	41	82%	-	N.S.	SOG.R
Right inferior frontal gyrus, orbital part	Bi-degree	41	82%	+	0.042	ORBinf.R
Right middle temporal gyrus	FA-degree	40	80%	+	N.S.	MTG.R
Left median cingulate and paracingulate gyri	Bi-efficiency	40	80%	+	0.019	DCG.L
Right middle frontal gyrus	FA-betweenness	39	78%	+	N.S.	MFG.R
Right precentral gyrus	FL-efficiency	39	78%	-	0.109	PreCG.R
Left precuneus	Bi-efficiency	38	76%	-	N.S.	PCUN.L
Left insula	FL-efficiency	37	74%	-	N.S.	INS.L
Left lenticular nucleus, putamen	Bi-efficiency	37	74%	+	0.027	PUT.L
Left median cingulate and paracingulate gyri	FN-efficiency	36	72%	+	0.023	DCG.L
Right inferior occipital gyrus	Bi-efficiency	36	72%	-	N.S.	IOG.R

Num: the counts of each parameter selected by our proposed method over the 5 rounds nested 10-fold CV. Rate: the frequency of being selected, equals Num/total times in 5 rounds nested 10-fold CV (50 times). +/-: parameter is increased/decreased in TS group compared to control group. P: p value of two sample t-test. N.S.: parameter is not significantly different between groups.

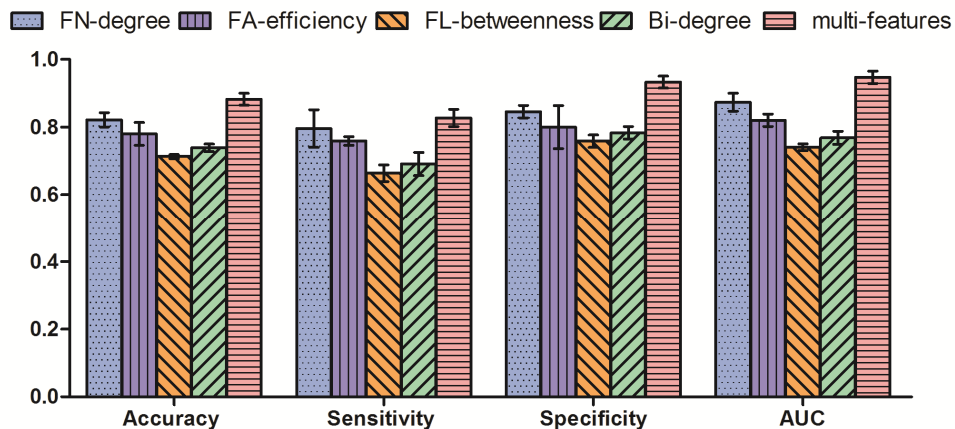


Figure 3. Comparison of performances of different methods using the single and multiple network features. For the FN-weighted, FA-weighted, FL-weighted and binary network, only the single type of characteristic achieving the highest accuracy were shown.

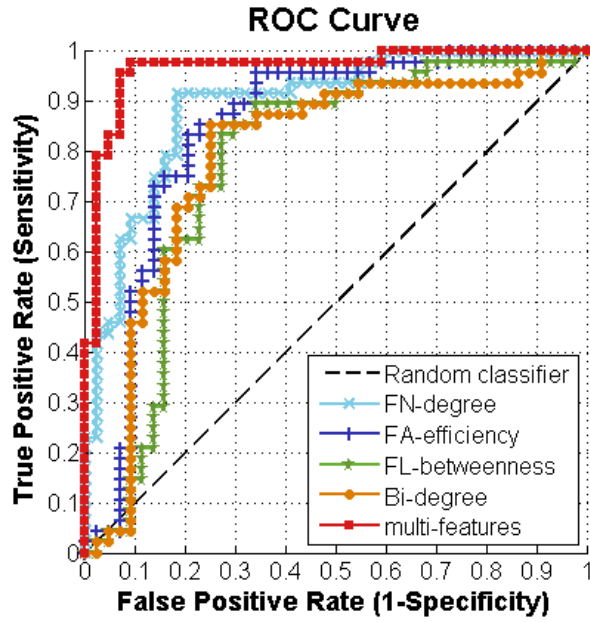


Figure 4. The ROC curve of nested CV classification on 92 subjects. For the FN-weighted, FA-weighted, FL-weighted and binary network, only the single type of characteristic achieving the highest accuracy were shown.

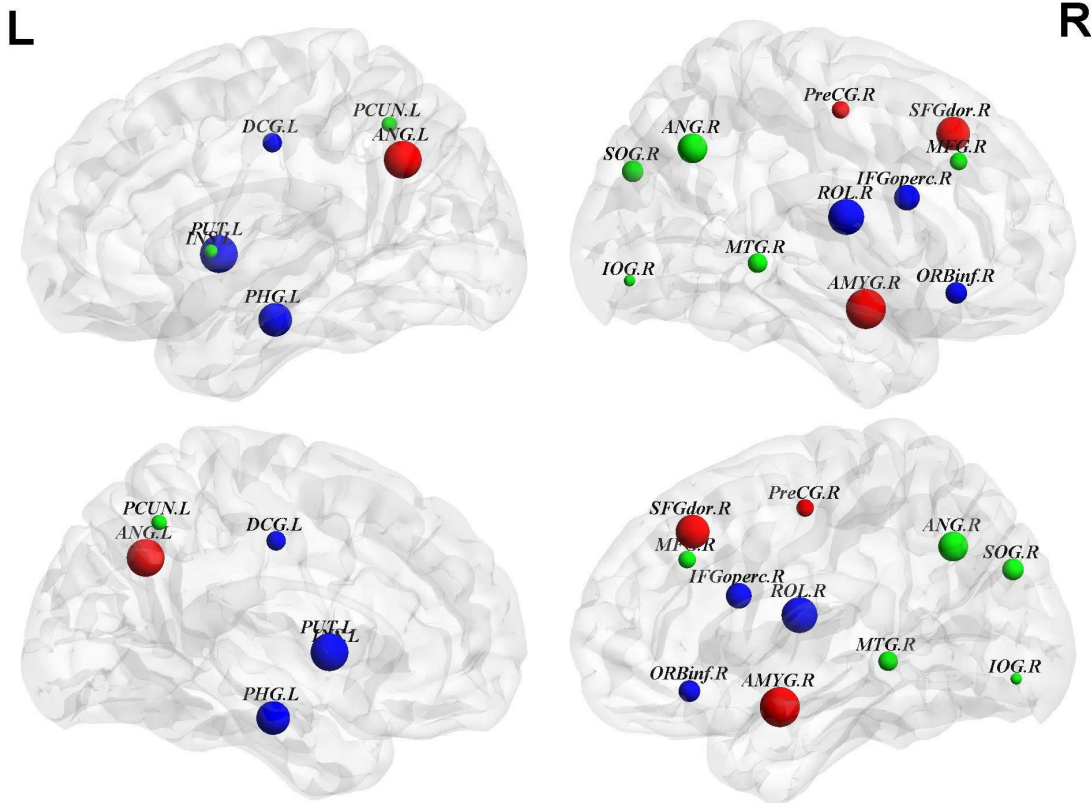


Figure 5. The most discriminative network nodes with the frequency more than 70%, the abbreviation were the same as Table 2. Blue/Red color identifies the node region has significantly higher/lower value of feature in TS compared to HC. Green color refers to the value of feature is not significantly different between groups in this node region. The node size indicates the frequency of feature being selected.

4. DISCUSSION

In this study, we used early one hundred samples with multi-modal MRI which are already quite large datasets and cost thousands of dollars and several months or years to collect. The current results show that it is possible to classify TS and healthy children with a high accuracy using an automated procedure that combines diffusion tractography, graph theoretical method and SVM. Our results achieved a highest accuracy of 88.26% better than the previous work which combined resting-state functional connectivity (RSFC) MRI and SVM for TS children classification and reached the accuracy of 74% [14].

Our study uses features extracted from multiple networks constructed from multiple parameters. The dimensionality of original features is much higher than the number of samples which cannot be directly used to train classifier. A wrapper feature selection method (SVM-RFE) is applied before training SVMs in order to prevent the overfitting problem and reduce the computational complexity. We also used a nested CV method to tune the hyper-parameters of classifiers and evaluate the performances of our method, which can yield unbiased estimation of classification method.

In addition to using a single kind of feature, our experiments show that combining multiple kinds of features can yield better classification results for early TS prediction. We used the same method to integrate features as a previous study [15] which combined several features into one single vector and then training a classifier. Wee et al. [16] also demonstrated that utilizing networks constructed from multiple DWI-derived parameters enhances classification performance.

The current work also identifies the discriminative node regions for classification which mainly located in the amygdala, basal ganglia (lenticular nucleus, putamen) and frontal cortico-cortical circuits. Our result showed the significantly reduced nodal efficiency in amygdala ranked as the most discriminative feature while the connectivity of the amygdala, for which a role in the pathophysiology of TS has been established [17]. Our result also tallied with previous related research where altered connectivity values were found in previous studies in basal ganglia and frontal cortico-cortical circuits [3]. These results support the contention that our diagnosis model may be efficient to capture the complexity of some brain disorders, and hold promise for early TS diagnosis.

5. CONCLUSION

Our study presented a classification framework for early TS prediction, which combined brain structural network, graph theoretical method and machine learning techniques. A fully automated procedure of this method is an appealing assistance for clinical TS diagnosis which is somewhat subjective and time consuming. We established an effective diagnosis model so the relevant features could be extracted and we can get a reliable prediction from the diagnosis model for the subject to be diagnosed. Our work may be useful for the early identification of individuals with TS.

ACKNOWLEDGEMENTS

This work was supported by 863 Projects (2013AA013803), National Natural Science Foundation of China (61271151, 91520202, 31271161), Youth Innovation Promotion Association CAS and Beijing Municipal Administration of Hospitals Incubating Program (PX2016035).

REFERENCES

- [1] Wolff, N., Luehr, I., Sender, J., Ehrlich, S., Schmidt-Samoa, C., Dechent, P., Roessner, V., "A DTI study on the corpus callosum of treatment-naive boys with 'pure' Tourette syndrome," *Psychiatry Res* (2015)
- [2] Stokes, A., Bawden, H.N., Camfield, P.R., Backman, J.E., Dooley, J.M., "Peer Problems in Tourettes Disorder," *Pediatrics* 87, 936-942 (1991)

- [3] Cheng, B., Braass, H., Ganos, C., Treszl, A., Biermann-Ruben, K., Hummel, F.C., Muller-Vahl, K., Schnitzler, A., Gerloff, C., Munchau, A., Thomalla, G., "Altered intrahemispheric structural connectivity in Gilles de la Tourette syndrome," *NeuroImage. Clinical* 4, 174-181 (2014)
- [4] He, Y., Evans, A., "Graph theoretical modeling of brain connectivity," *Curr Opin Neurol* 23, 341-350 (2010)
- [5] Worbe, Y., Marrakchi-Kacem, L., Lecomte, S., Valabregue, R., Poupon, F., Guevara, P., Tucholka, A., Mangin, J.F., Vidailhet, M., Lehericy, S., Hartmann, A., Poupon, C., "Altered structural connectivity of cortico-striato-pallido-thalamic networks in Gilles de la Tourette syndrome," *Brain* 138, 472-482 (2015)
- [6] Zhang, D., Shen, D., *Alzheimer's Disease Neuroimaging, I.*, "Multi-modal multi-task learning for joint prediction of multiple regression and classification variables in Alzheimer's disease," *Neuroimage* 59, 895-907 (2012)
- [7] Dai, D., Wang, J., Hua, J., He, H., "Classification of ADHD children through multimodal Magnetic Resonance Imaging," *Frontiers in Systems Neuroscience* 6, 63-63 (2012)
- [8] Jin, Y., Wee, C.Y., Shi, F., Thung, K.H., Ni, D., Yap, P.T., Shen, D., "Identification of infants at high-risk for autism spectrum disorder using multiparameter multiscale white matter connectivity networks," *Human brain mapping* (2015)
- [9] Bullmore, E., Sporns, O., "Complex brain networks: graph theoretical analysis of structural and functional systems (vol 10, pg 186, 2009)," *Nat Rev Neurosci* 10, (2009)
- [10] Liu, Y., Duan, Y.Y., He, Y., Wang, J., Xia, M.R., Yu, C.S., Dong, H.Q., Ye, J., Butzkueven, H., Li, K.C., Shu, N., "Altered Topological Organization Of White Matter Structural Networks In Patients With Neuromyelitis Optica," *Mult Scler J* 19, 666-667 (2013)
- [11] Mori, S., Crain, B.J., Chacko, V.P., van Zijl, P.C., "Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging," *Ann Neurol* 45, 265-269 (1999)
- [12] Guyon, I., Weston, J., Barnhill, S., Vapnik, V., "Gene selection for cancer classification using support vector machines," *Mach Learn* 46, 389-422 (2002)
- [13] Wilson, S.M., Ogar, J.M., Laluz, V., Growdon, M., Jang, J., Glenn, S., Miller, B.L., Weiner, M.W., Gorno-Tempini, M.L., "Automated MRI-based classification of primary progressive aphasia variants," *Neuroimage* 47, 1558-1567 (2009)
- [14] Richards, C.A., Black, K.J., "Tourette Syndrome research highlights 2014," *F1000Research* 4, 69 (2015)
- [15] Dyrba, M., Ewers, M., Wegrzyn, M., Kilimann, I., Plant, C., Oswald, A., Meindl, T., Pievani, M., Bokde, A.L.W., Fellgiebel, A., "Combining DTI and MRI for the Automated Detection of Alzheimer's Disease Using a Large European Multicenter Dataset," *Springer Berlin Heidelberg* (2012)
- [16] Wee, C.Y., Yap, P.T., Li, W., Denny, K., Browndyke, J.N., Potter, G.G., Welsh-Bohmer, K.A., Wang, L., Shen, D., "Enriched white matter connectivity networks for accurate identification of MCI patients," *Neuroimage* 54, 1812-1822 (2011)
- [17] Werner, C.J., Stocker, T., Kellermann, T., Wegener, H.P., Schneider, F., Shah, N.J., Neuner, I., "Altered amygdala functional connectivity in adult Tourette's syndrome," *European archives of psychiatry and clinical neuroscience* 260 Suppl 2, S95-99 (2010)