Multi-threshold White Matter Structural Networks Fusion for Accurate Diagnosis of Tourette Syndrome Children

Hongwei Wen[1,2a], Yue Liu[4a], Shengpei Wang[1,2a], Zuoyong Li[5], Jishui Zhang[6], Yun Peng[4*], Huiguang He[1,2,3*]

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

[2] University of Chinese Academy of Sciences, Beijing, China

[3] Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Sciences, Beijing, China

[4] Department of Radiology, Beijing Children's Hospital, Capital Medical University, Beijing, China

[5] Fujian Provincial Key Laboratory of Information Processing and Intelligent Control (Minjiang University), Fuzhou, 350121, China

[6] Department of Neurology, Beijing Children's Hospital, Capital Medical University, Beijing, China

ABSTRACT

Tourette syndrome (TS) is a childhood-onset neurobehavioral disorder. To date, TS is still misdiagnosed due to its varied presentation and lacking of obvious clinical symptoms. Therefore, studies of objective imaging biomarkers are of great importance for early TS diagnosis. As tic generation has been linked to disturbed structural networks, and many efforts have been made recently to investigate brain functional or structural networks using machine learning methods, for the purpose of disease diagnosis. However, few studies were related to TS and some drawbacks still existed in them. Therefore, we propose a novel classification framework integrating a multi-threshold strategy and a network fusion scheme to address the preexisting drawbacks. Here we used diffusion MRI probabilistic tractography to construct the structural networks of 44 TS children and 48 healthy children. We ameliorated the similarity network fusion algorithm specially to fuse the multi-threshold structural networks. Graph theoretical analysis was then implemented, and nodal degree, nodal efficiency and nodal betweenness centrality were selected as features. Finally, support vector machine recursive feature extraction (SVM-RFE) algorithm was used for feature selection, and then optimal features are fed into SVM to automatically discriminate TS children from controls. We achieved a high accuracy of 89.13% evaluated by a nested cross validation, demonstrated the superior performance of our framework over other comparison methods. The involved discriminative regions for classification primarily located in the basal ganglia and frontal cortico-cortical networks, all highly related to the pathology of TS. Together, our study may provide potential neuroimaging biomarkers for early-stage TS diagnosis.

Keywords: Tourette syndrome, Diffusion MRI, Probabilistic tractography, Structural connectivity, Graph theoretical analysis, Similarity Network Fusion, Support vector machine

1. INTRODUCTION

Tourette syndrome (TS) is a childhood-onset neurobehavioral disorder characterized by the presence of multiple motor and vocal tics. The typical age of onset is between 5 and 7 years. Affected individuals are at increased risk for the development of various comorbid conditions, such as attention deficit hyperactivity disorder (ADHD), obsessive-compulsive disorder (OCD), school problems, depression, and anxiety [1], leading to its heterogeneous clinical expression [2]. The prevalence of TS is now estimated to range between 0.1% and 1% in school-age children. However, since tics are often so mild, especially in childhood, TS are hardly perceived and easily overlooked, the incidence in

[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

Medical Imaging 2017: Computer-Aided Diagnosis, edited by Samuel G. Armato III, Nicholas A. Petrick,
Proc. of SPIE Vol. 10134, 101341Q · © 2017 SPIE · CCC code: 1605-7422/17/$18 · doi: 10.1117/12.2251426

children has a tendency to increase in recent years [3]. To date, TS diagnosis is mainly depends on the qualitative description of symptoms. There is no reliable diagnostic test for TS, and TS is still misdiagnosed due to its varied presentation, and lacking of obvious clinical symptoms [4]. Therefore, studies of objective diagnostic method are of great importance. Moreover, if TS children can be accurately diagnosed in the early stages, treatments can be given in time to potentially slow down the progress of the disease.

The mechanism of TS is complex and not fully understood. However, neuro-pathological, electrophysiological, and neuroimaging studies have led to increased knowledge in understanding the dysfunctional neuronal processing of cortico-striato-pallido-thalamo-cortical (CSPTC) circuits in TS. Specially, structural and functional abnormalities in CSPTC circuits have been found in relation to tic generation [5]. As the human brain is a complex network, all anatomically distinct brain regions such as the cortical regions, basal ganglia and thalamus mathematically compose the nodes of brain structural network, and the white matter (WM) structural pathways connecting pairs of regions make up the links (edges) of structural network. Therefore, it is meaningful to study neurological disorders from a network perspective. The brain structural networks could provide us with an anatomical and physiological substrate of brain functions and helps us to understand how the brain structures shape functional interactions [6]. Therefore, study on structural networks has gained a lot of popularity recently, and there has grown up an urgent need to quantify the integrity of WM connections between grey matter (GM) regions in vivo, non-invasively, on a global scale, in order to depict the topological organization of brain structure [7] and apply this to the studies of neurologic disease. At present, diffusion MRI is a promising noninvasive technique for assessing the microstructural characteristics of WM tracts. Specifically, diffusion MRI tractography can reconstruct the major WM tracts faithful to the known WM anatomy, allowing mapping of the brain's structural connectivity in vivo, therefore, enhance our ability to investigate whole brain WM structures. Some previous studies using diffusion tensor imaging (DTI) showed abnormalities in the WM tracts of the somatosensory, fronto-striatal and motor pathways in TS [8, 9]. In addition, Worbe et al. [10] used the probabilistic tractography algorithms to quantify the structural integration of CSPTC circuits, and showed widespread structural abnormalities in CSPTC circuits in TS patients, which likely result from abnormal brain development in this syndrome.

Despite these previous studies have revealed disrupted structural networks in TS, these significant findings are only at group level which limited clinical application, and recent attention has turned toward integrating machine learning and neuroimaging to assist clinical diagnose. Unlike group-based comparison approaches, machine learning techniques are able to detect the fine-grained spatial discriminative patterns, which are critical for individual-based disease diagnosis [11]. There is rapidly accumulating evidence that integrating machine learning and neuroimaging measurements may be valuable for disease diagnosis, transition prediction and treatment prognosis [12].

Although many efforts have been made recently to investigate brain functional or structural networks using machine learning methods, for the purpose of disease diagnosis [13, 14], few of them are related to TS. Greene et al. [15] extracted original resting state functional connectivity (RSFC) as features, then used support vector machine (SVM) for TS classification, achieving an accuracy of 74%. Wen et al. [16] constructed four types of structural networks via deterministic tractography, and then all networks were sparse with a single fixed threshold for subsequent graph theory analysis, features fusion and SVM-based TS classification, achieving an accuracy of 88.26%. Nevertheless, these studies have three main drawbacks. Firstly, Greene et al. [15] extracted the connectivity values as features in original functional networks, which may have the noise or spurious connections, and should be sparse with proper thresholds. Secondly, Greene et al. implemented no graph theoretical approaches in their study [15]. As human brain is a complex network, the original low-level connectivity values as features could not capture high-level topological properties in human brain networks. By contrast, graph theory allows endowing networks with a great number of quantitative properties [17, 18], thus vastly enriching the set of objective descriptors of brain structure and function at neuroscientists' disposal. Thirdly, though Wen et al. implemented sparse processing and graph theory analysis [16], the selection of a single fixed threshold in their study is subjective and may be not reasonable since different noise level.

Therefore, in present study, we proposed a novel classification framework integrating the graph theory analysis, a multi-threshold strategy and a network fusion scheme to address the above drawbacks. Briefly, we first used diffusion MRI probabilistic tractography to construct multi-threshold whole brain white matter structural networks for 44 drug-naive TS children and 48 matched healthy children. These multi-threshold networks are then integrated with a network fusion scheme to capture their common and complementary information. Graph theoretical analysis was then implemented on the fused networks, and three types of nodal topological properties were selected as features for subsequent classification. After feature selection, the optimal features are fed into SVM to automatically discriminate TS children

from controls. The highly discriminative topological features of fused structural networks for classification identified by our study may be potential helpful neuroimaging biomarkers for assisting the clinical TS diagnosis.

2. MATERIALS AND METHODS

2.1 Subjects

44 TS patients were recruited from outpatient clinics in Beijing Hospital from July 2012 to May 2015 (age: 8.98±3.114 years, range: 3–16 years; 11 female). All the patients were drug-naive subjects to exclude the effects of stimulants, as previous studies have suggested that stimulants can significantly influence the structure and function of central nervous system in TS [19]. 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 healthy children 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) [20] to diagnose OCD and used the German short version of Wender Utah rating scale (WURS-k, translated to Chinese) [21] to diagnose ADHD. Patients fulfilling OCD criteria or other co-morbidities were excluded from the study. As a previous TS classification study [15] tested specifically for differences between TS children with and without ADHD, and found no significant differences in individual accuracy or in the proportion of children classified correctly. Based on similar consideration, in this study, we included children with ADHD as our previous studies [22, 23]. Tic severity for all patients was rated using the Yale Global Tic Severity Scale (YGTSS) and ranged from 10 to 79 ([mean±SD]: 46.50±18.037). The duration of TS ranged from 3 month to 5 years ([mean±SD]: 1.81±1.423 years). For subjects who had course less than 1 year, TS diagnosis was made by follow-up call and they were all finally diagnosed exactly as TS by our professional neurologists and psychiatrists. 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 (Gyroscan 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 weighted scans were acquired using the following protocol: spin-echo diffusion-weighted echo-planar imaging sequence of 30 non-collinear directions with a b value of 1000 s/mm^2, 2mm slice thickness, no inter-slice gap, repetition time = 4300ms, echo time = 95ms, field of view (FOV) =255×255mm, reconstructed image matrix = 336×336. 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) = 200×200mm, slice thickness = 1mm.

2.3 Data preprocessing

Following data acquisition, all of the 3D T1-weighted images were reoriented with the origin set close to the anterior commissure (AC). 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 1000s/mm^2 b-value were first affinely registered to the b_0 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 ($\lambda1$, $\lambda2$, $\lambda3$) were obtained and maps of fractional anisotropy (FA) were generated.

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: Briefly, individual T1-weighted images were coregistered to the b_0 images in the DTI space. Then, the transformed T1 images were then nonlinearly transformed to the ICBM152 T1 template in the MNI space. The inverse transformations were used to warp the Automated Anatomical Labeling (AAL) atlas [24] from the MNI space to the DTI native space. Of note, the nearest-neighbor interpolation method was used to preserve discrete labeling values. Using this procedure, we obtained 90 cortical and subcortical regions (45 for each hemisphere), each representing a node of the network.

Network edge definition: To define the connections (edges) between the brain regions, we performed probabilistic tractography using the FSL. Firstly, we used the bedpostx tool in FSL to run Markov Chain Monte Carlo sampling, in order to build up distributions on diffusion parameters at each voxel, which allows to model crossing fibres within each voxel of the brain. Secondly, we used the probtrackx tool in FSL to perform probabilistic tracking. Briefly, we repetitively performed 5000 sampling from the distributions of voxel-wise principal diffusion directions, each time computing a streamline through these local samples to generate a probabilistic streamline fiber. For a seed region, 5000 × n streamline fibers were sampled; n is the number of voxels in the seed region. The number of streamline fibers passing through a given region divided by 5000 × n is calculated as the connectivity probability from the seed region to the given region. In our study, each brain region in AAL template was selected as the seed region, and its connectivity probability to each of the other 89 regions was calculated. Notably, the probability from i to j is not necessarily equivalent to the probability from j to i because the tractography is dependent on the seeding location. However, these two probabilities are highly correlated across the brain regions for all subjects in our study (all Pearson $r > 0.76$, $p < 0.05$). Thus, we defined the unidirectional connectivity probability P_{ij} between region i and region j by averaging these two probabilities. Finally, to define the network edges, we computed $w_{ij} = P_{ij}$ as the weight between brain regions i and j. For each subject, a 90 × 90 symmetric weighted network was constructed.

To better remove noise and also locate more biologically meaningful features for classification, the network needs to be sparse by the proper threshold. The network sparsity is defined as the fraction of the number of edges in a graph compared to the maximum possible number of edges (e.g., a sparsity threshold of x% means x% of the topmost connections preserved). However, because there is currently no definitive way to select a single sparsity threshold, we therefore used a range of sparsity thresholds (10% < sparsity < 30% in steps of 1%). The range of thresholds was selected empirically based on previous brain structural network studies [25, 26], which has shown that the networks are small-world within these thresholds. Specifically, two brain regions were considered disconnected if the connection weight w_{ij} was below a given threshold. The flowchart for structural network construction was shown in Figure 1.

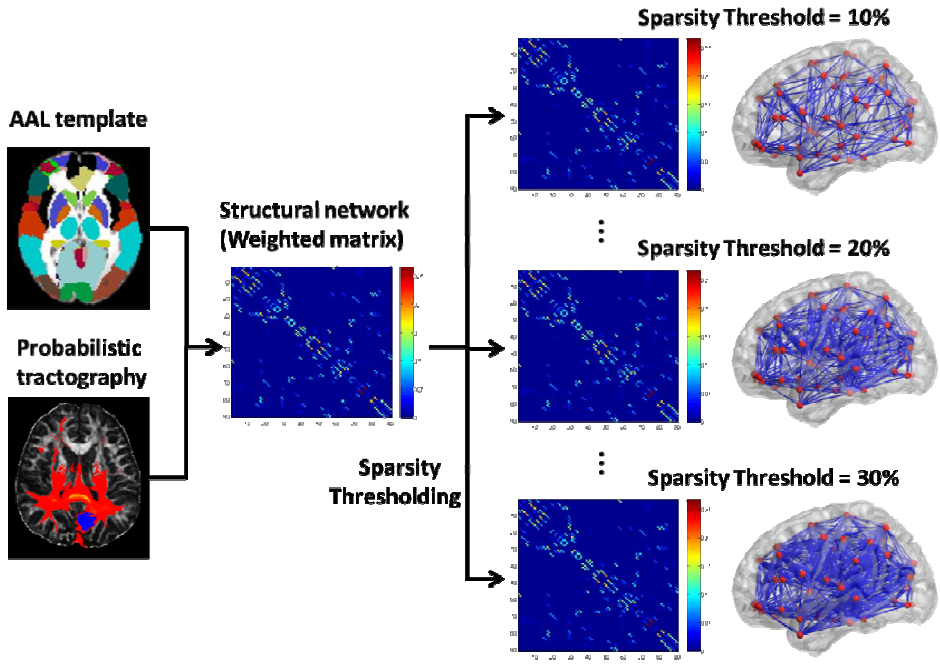

Figure 1. The flowchart of constructing the multi-threshold WM networks using diffusion probabilistic tractography.

2.5 Multi-threshold similarity networks fusion

The multi-threshold networks constructed in Section 2.4 provided complementary information for TS classification. To leverage their common and complementary information, we adopted the similarity network fusion (SNF) algorithm [27] and ameliorated it specially to merge the multi-threshold structural networks. To use SNF for our application, for each network W_i^j of subject i with sparsity threshold j, a sparse kernel matrix was constructed, which encodes the sparse, strong network connections. Let δ_u denote a set of k-nearest neighbors (the top k strongest connections) of the node u (including u itself) in W_i^j, then the sparse kernel matrix S_i^j is defined as

$$S_i^j(u,v) = \begin{cases} W_i^j(u,v) & v \in \delta_u \\ 0 & otherwise \end{cases} \tag{1}$$

Where the connection between nodes v and u existed only if v was within the k-nearest (k=27 in our case) neighbors of u. Based on these two kernel matrices, each network could be iteratively updated as follows:

$$(W_i^j)^{(m+1)} = S_i^j \times \frac{\sum_{c \neq j} (W_i^c)^{(m)}}{N-1} \times (S_i^j)^T, c = 1, \ldots, N \tag{2}$$

Where $(W_i^c)^m$ denoted the network W_i^c for subject i at the m-th iteration, and $(W_i^j)^{(m+1)}$ was the updated W_i^j after m+1 iterations. Eq. (2) implies that the connection of any two nodes in W_i^j also relies on the connections of their neighbors in other thresholded networks. Specially, if the respective neighbors of two nodes are strongly connected in other thresholded networks, the connection between them can be strengthened after the updates even though it may be weak itself, and vice versa.

The iterative process stops when it satisfied the convergence condition in Eq. (3):

$$\left\| \left(W_i^j\right)^{(m+1)} - \left(W_i^j\right)^{(m)} \right\| \leq 0.01 \tag{3}$$

After the iterative process converged, we averaged the N networks to obtain the average fused network for subject i:

$$W_i = \frac{\sum_j W_i^j}{N}, j = 1, \ldots, N \tag{4}$$

Finally, all elements in W_i were normalized to [0, 1] and the final fused network was obtained for each subject i. The fused network doesn't rely on any individual threshold and thus is less affected by noise. The SNF in our study was implemented using a MATLAB toolbox SNFmatlab_v2.1 (http://compbio.cs.toronto.edu/SNF/SNF/Software.html) [27].

2.6 Graph theory analyses

Graph theory analyses were performed on the final fused network for each subject using a graph theoretical network analysis toolbox (GRETNA). We computed three types of nodal topological properties, namely, nodal degree $D_{nodal}(i)$, nodal efficiency $E_{nodal}(i)$ [28] and nodal betweenness centrality $B_{nodal}(i)$ as the features for subsequent classification.

$$D_{\text{nodal}}(i) = \sum_{i \neq j \in G} w_{ij} \tag{5}$$

$$E_{\text{nodal}}(i) = \frac{1}{N-1} \sum_{i \neq j \in G} \frac{1}{L_{ij}} \tag{6}$$

$$B_{nodal}(i) = \sum_{s \neq i \neq t \in G} \frac{e_{sit}}{e_{st}} \tag{7}$$

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, $e_{sit}(i)$ is the number of shortest paths in the network G between node s and node t which pass through node i, $e_{st}(i)$ is

the number of shortest paths in the network G between node s and node t. $D_{nodal}(i)$ 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 fraction of all shortest paths that pass through a given node i.

Finally we merged the three types of nodal topological properties by concatenating feature vectors, namely, the feature vector for each subject x_i was: $x_i = [D_{nodal}(1), ..., D_{nodal}(90), E_{nodal}(1), ..., E_{nodal}(90), B_{nodal}(1), ..., B_{nodal}(90)]$.

2.7 Feature selection and classification

After extracting the features from each individual fused network, we employ a wrapper-based method for feature subset selection utilizing SVM based on recursive feature elimination which is named SVM-RFE [4]. 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. Of note, the feature selection was implemented on the training set, rather than entire dataset, using the accuracy of 10-fold cross validation (CV) to estimate the goodness of feature subset, which may avoid the overfitting problem.

Classification was performed using the SVM algorithm with a radial basis function (RBF) kernel. For the SVM we needed to define two parameters including the complexity or cost constant ($c>0$) 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 suitable values for C and γ we used a grid search in the range of $c=2^{-4}, 2^{-3}, ..., 2^{4}$ and $\gamma=2^{-8},2^{-7}, ...,2^{2}$, with 10-fold CV to evaluate the goodness of SVM parameters.

In this study, the same nested cross-validation strategy in our previous study [23] was used to evaluate the classification performance. 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 nested CV.

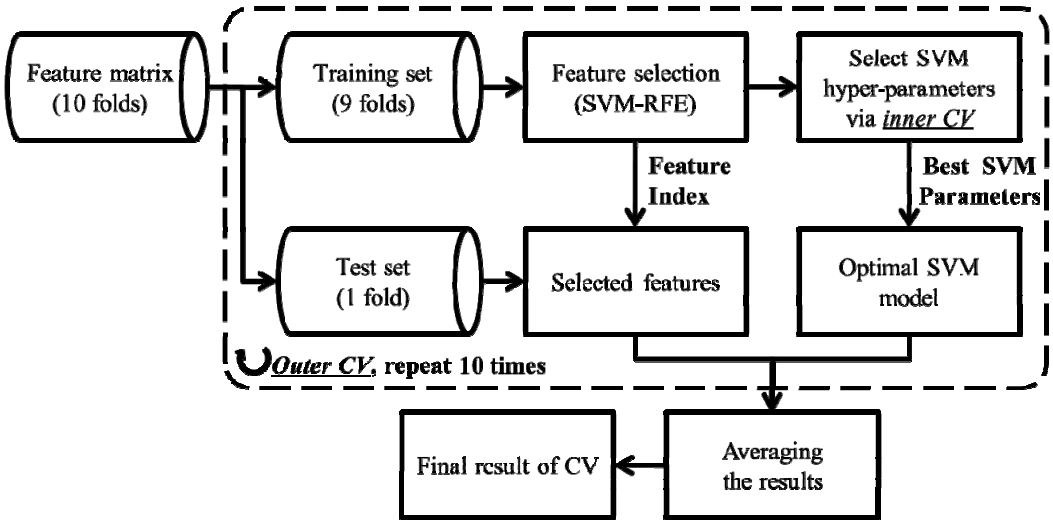

Figure 2. The flow chart of the nested CV procedure in our classification framework.

2.8 Evaluation of classification performance

Since our proposed framework had two contributions: 1) graph theory analysis and multiple thresholds processing, and 2) network fusion with SNF, to evaluate its performance, we compared it against three other methods that used:

1. The original structural connectivity (SC) matrices without graph theory analysis, 90*89/2 features for each subject.

2. The original matrices were sparse by a specific single sparsity threshold (13%) empirically based on our previous brain network study [29], 270 features (90*3 nodal properties) for each subject derived from graph theory analysis.

3. We calculated the area under the curve (AUC) for each nodal network metric [30], which can provide a threshold-independent network assessment and network fusion with different thresholds. The AUC value of each nodal property across thresholds was used as the feature for classification, 270 features for each subject.

The same feature selection (SVM-RFE), SVM-based classification, nested CV procedures were then implemented for each of these three comparison methods.

For the analysis of results, accuracy, sensitivity, specificity and the area under the curve for the receiver operated characteristic curve (AUC ROC) are shown. 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 The effects of multi-threshold similarity network fusion algorithm

To further gain the insights of our algorithm, we randomly selected three subjects from TS group and control group, respectively. Figure 3 illustrates the original SC networks and the multi-threshold networks after fusion. Compared to the original networks, our fused networks show more block-like structures with more clear layouts. Besides, the original networks look similar between groups, while our fused networks show significant difference between groups. The network connections in TS group seem to be much weaker, and this is consistent with the previous studies showing widespread structural connectivity deficits in TS [10, 31].

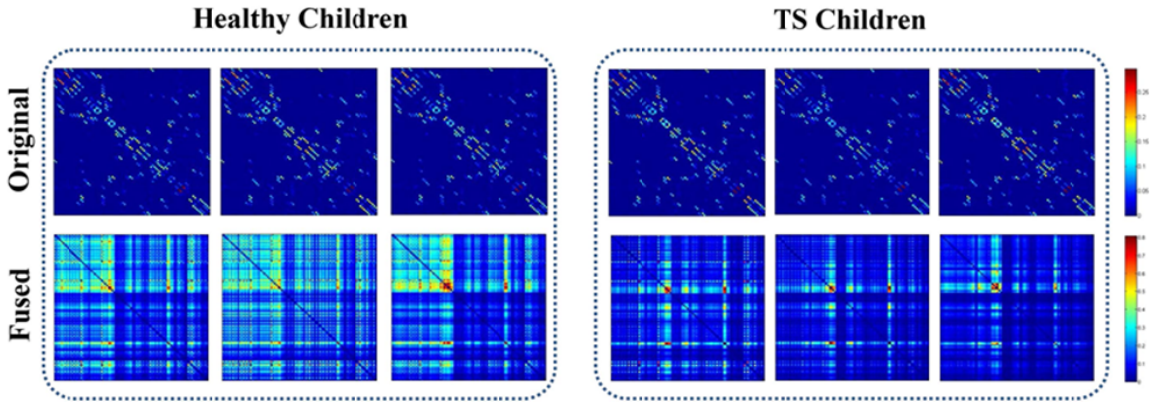

Figure 3. Original networks and fused networks of three randomly selected healthy children and TS children, respectively.

3.2 The classification performance of our proposed method

The classification performances of all methods in percentage were showed in Table 2. To exclude the possibility that the classification performance based on one-time randomly partitioning in the CV was not robust and stable, we repeated the nested CV procedure for five times. Our proposed framework achieved the highest mean (±standard deviation) accuracy (89.13±3.35%), sensitivity (86.36±5.79%), specificity (91.67±1.47%), and AUC (95.54±1.31%), indicating that the high classification power was unlikely the result of chance. Figure 4 plots the receiver operating characteristic (ROC) curves of all methods. Our proposed framework outperforms all 3 comparison methods.

Table 2. Classification performances of all methods in percentage.

Feature	ACC(%)	SEN(%)	SPE(%)	AUC(%)
Original SC	64.78±1.65	59.55±3.37	69.58±2.80	70.36±3.71
Single Threshold	61.96±4.68	57.27±7.43	66.25±2.72	67.25±3.45
AUC method	68.91±3.57	62.27±3.05	75.00±8.20	70.14±3.37
Fused Networks	**89.13±3.35**	**86.36±5.79**	**91.67±1.47**	**95.54±1.31**

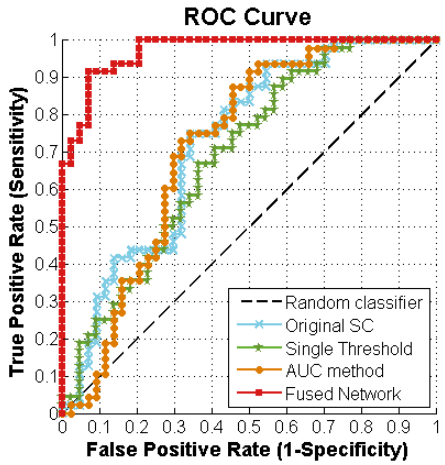

Figure 4. ROC curves of the classification results, which shows our framework is better than other comparison methods.

3.3 Regions most influential for classification

As the highest accuracy was achieved on multi-threshold fused SC networks, we also identified the most influential nodal topological properties for TS classification. We summed the counts of each feature selected by our proposed method over the five rounds nested 10-fold CV. The top 25 selected features with the frequency more than 70% were provided in Table 3. The involved node regions located mainly in frontal gyrus, putamen, supplementary motor area, hippocampus, which are discriminative for TS classification, and may be highly related to TS pathology. Of note, seven regions (eg. left superior orbital frontal gyrus and hippocampus) have multiple types of nodal topological properties as discriminative features. For the purpose of visualization, Figure 5 illustrates the discriminative nodes for classification.

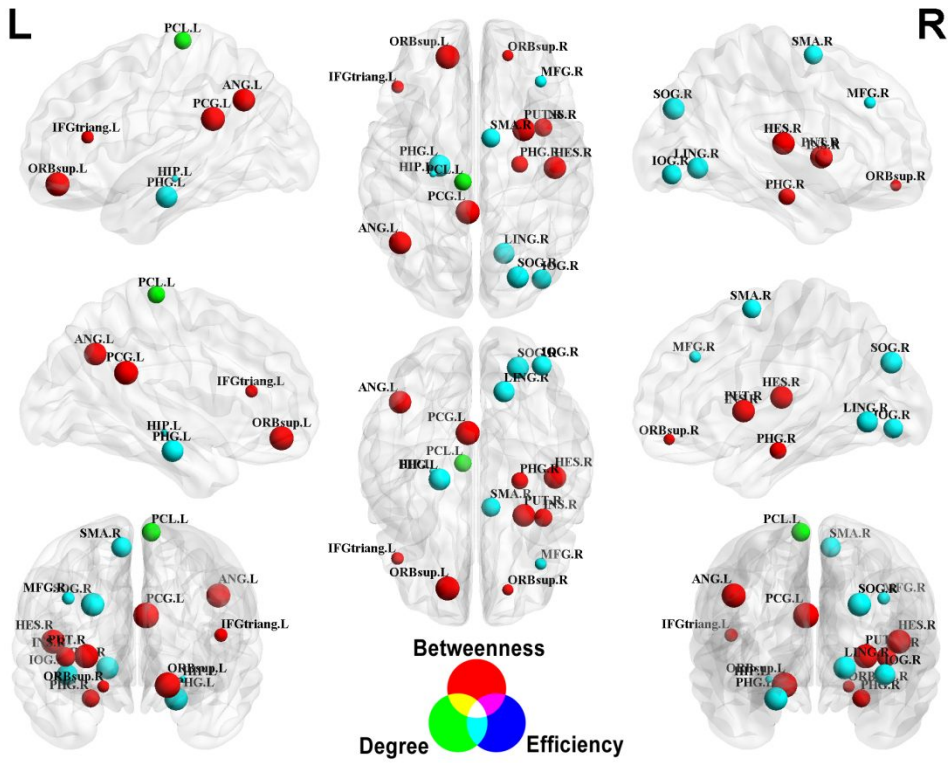

Figure 5. The brain regions with nodal topological properties as most discriminative features for classification. The brain graphs were visualized using BrainNet Viewer software. The node with RGB color indicates the type of nodal topological

properties as discriminative feature. The node sizes indicates the frequency of being selected as optimal features by the nested CV procedure, and if a node has multiple types of features, the node sizes indicates the average frequency. Abbreviation is the same as Table 1.

Table 3. The top 29 discriminative network features selected by the nested CV method as the most salient for TS classification.

Region	Feature	Count	Rate	Abbr
Left superior frontal gyrus, orbital part	Betweenness	48	96%	**ORBsup.L**
Left posterior cingulate gyrus	Betweenness	48	96%	**PCG.L**
Left angular gyrus	Betweenness	47	94%	**ANG.L**
Right lenticular nucleus, putamen	Betweenness	47	94%	**PUT.R**
Right heschl gyrus	Betweenness	47	94%	**HES.R**
Left parahippocampal gyrus	Degree	47	94%	**PHG.L**
Right superior occipital gyrus	Degree	47	94%	**SOG.R**
Right lingual gyrus	Degree	46	92%	**LING.R**
Right inferior occipital gyrus	Degree	46	92%	**IOG.R**
Left parahippocampal gyrus	Efficiency	46	92%	**PHG.L**
Right lingual gyrus	Efficiency	46	92%	**LING.R**
Right superior occipital gyrus	Efficiency	46	92%	**SOG.R**
Right supplementary motor area	Degree	45	90%	**SMA.R**
Right insula	Betweenness	44	88%	**INS.R**
Left paracentral lobule	Degree	44	88%	**PCL.L**
Right supplementary motor area	Efficiency	44	88%	**SMA.R**
Right inferior occipital gyrus	Efficiency	44	88%	**IOG.R**
Right parahippocampal gyrus	Betweenness	43	86%	**PHG.R**
Left paracentral lobule	Efficiency	43	86%	**PCL.L**
Left inferior frontal gyrus, triangular part	Betweenness	40	80%	**IFGtriang.L**
Right middle frontal gyrus	Degree	40	80%	**MFG.R**
Right superior frontal gyrus, orbital part	Betweenness	39	78%	**ORBsup.R**
Right middle frontal gyrus	Efficiency	39	78%	**MFG.R**
Left hippocampus	Degree	36	72%	**HIP.L**
Left hippocampus	Efficiency	36	72%	**HIP.L**

Count: the counts of each feature selected by our proposed method over the five rounds nested 10-fold CV. Rate: the frequency of being selected, equals Count/total times in five rounds nested 10-fold CV (50 times).

4. DISCUSSION

In the present study, we proposed a novel network-based classification framework using multi-threshold similarity network fusion algorithm, for automatically differentiating TS children from controls. Our results achieved the accuracy of 89.13%, which is better than the previous studies (74% and 88.26%, respectively) [15, 16], and may be the state-of-the-art in computer-aided TS diagnosis currently.

4.1 The advantages in feature extraction and selection

In the present study, we constructed the whole brain WM networks of both TS and healthy children, using the probabilistic tractography method to map whole-brain WM connectivity, which has advantages in tracking crossing and splitting fiber bundles compared with deterministic tractography methods [16]. Then we extracted the network topological properties as features for subsequent classification using graph theory analyses. Graph theory-based approaches model the brain as a complex network represented graphically by a collection of nodes and edges, and investigate the organizational mechanism underlying the relevant networks. In contrast to those widely used RSFC analysis methods [15], the graph-based network analyses allow us not only to visualize the overall connectivity pattern among all the brain regions but also to quantitatively characterize the global organization [32]. Therefore, the feature extraction based on graph-based techniques can probe potential mechanisms involved in TS, and may provide potential discriminative biomarkers for TS prediction or classification.

The SVM-RFE algorithm was 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. It is worth emphasizing that the feature selection method in our study was implemented on the training set rather than the full dataset. While some other studies [33-35] employed feature selection on the full dataset which may be overly optimistic, the evaluation method in our study is a better estimation of generalization.

4.2 The benefits of multi-threshold similarity network fusion

From the experimental results, our proposed method based on multi-threshold similarity network fusion has demonstrated the superior performance over other comparison methods. Unlike original similarity network fusion which enhances the sample similarity network by integrating different feature types, we ameliorated the algorithm specially to enhance the SC networks for each subject, by integrating the information provided by other topological views of the original network. At a word, during the network fusion process, each thresholded network is iteratively updated under the interaction of two networks, namely, a sparse network carrying its own important strongest connectivity information, and the average of the other networks. Compared to the multi-kernel learning (MKL) methods [23] which optimize the weights of the kernels for feature fusion, our proposed method provides a novel fusion scheme, in which the full spectrum of complementary information in multi-threshold networks can be integrated, without tuning the complex combination of weight parameters in MKL which need much computational time for the training process.

The fused network after doesn't rely on any individual threshold, and thus is less affected by noise. Moreover, it incorporates the common and complementary information from multi-threshold networks after SNF, so it can better represent the underlying ground truth of the SC networks. Superficially, the advantage of accuracy in this work (89.13%) is not obvious over the previous study which achieves an accuracy of 88.26% [16]. However, we only extracted features in one type of SC networks, while the previous study combined features in four types of SC networks for classification. In fact, when only using features in one type of SC network, the highest accuracy was 82.17% in that study [16], and the accuracy of 89.13% in this study has much advantage over it, which indeed showed the benefits of multi-threshold SNF scheme.

4.3 Most discriminative features highly related to the pathology of TS

Identification of objective imaging biomarkers is of great interest as it could, ultimately, assist clinical decisions for individual patients. With this consideration, our proposed network-based classification framework identified those topological features that are most discriminative in distinguishing between TS children and controls, thereby identifying brain regions that might be most related to TS. The involved node regions located mainly in frontal gyrus, putamen, supplementary motor area, hippocampus, which are discriminative for TS classification, all highly related to TS pathology, as the previous studies revealed [10, 16, 22, 36].

5. CONCLUSION

Our study presented a novel classification framework for early TS diagnosis, which integrated brain structural network, graph theoretical method, multi-threshold network fusion scheme and machine learning techniques. Finally, we achieved an excellent accuracy of 89.13%. A fully automated procedure of this method is an appealing assistance for clinical TS diagnosis. The current work also identified the discriminative node regions for classification, primarily located in the basal ganglia and frontal cortico-cortical networks, all highly related to the pathology of TS. Our results also suggest that the topological properties of WM structural networks could be reliable neuroimaging biomarkers for clinical TS diagnosis, as well as the monitoring of the disease progression and the treatment effects for TS children.

ACKNOWLEDGEMENTS

We thank Dr. Hao Huang at University of Pennsylvania for consultation and support on MR pulse sequences.

This work was supported by National Natural Science Foundation of China (91520202, 61271151, 31271161), Youth Innovation Promotion Association CAS and Beijing Municipal Administration of Hospitals Incubating Program (PX2016035), Beijing health system top level health technical personnel training plan (2015-3-082).

AUTHOR CONTRIBUTIONS

Designed and performed experiment: Hongwei Wen; Wrote the manuscript and prepared the figures and tables: Hongwei Wen; Conceived and designed the project: Huiguang He, Yun Peng, Yue Liu; Data Collection: Yue Liu, Jishui Zhang; Obtaining funding: Huiguang He, Yun Peng, Yue Liu. All the authors revised and approved the manuscript.

REFERENCES

[1] Felling, R.J., Singer, H.S.: Neurobiology of tourette syndrome: current status and need for further investigation. Journal of Neuroscience the Official Journal of the Society for Neuroscience 31, 12387-12395 (2011)
[2] Cavanna, A.E., Servo, S., Monaco, F., Robertson, M.M.: The behavioral spectrum of Gilles de la Tourette syndrome. The Journal of neuropsychiatry and clinical neurosciences 21, 13-23 (2009)
[3] Scharf, J.M., Miller, L.L., Gauvin, C.A., Alabiso, J., Mathews, C.A., Ben-Shlomo, Y.: Population prevalence of Tourette syndrome: a systematic review and meta-analysis. Movement disorders : official journal of the Movement Disorder Society 30, 221-228 (2015)
[4] Cavanna, A.E., Seri, S.: Tourette's syndrome. Bmj 347, 67-71 (2013)
[5] Bronfeld, M., Bar-Gad, I.: Tic disorders: what happens in the basal ganglia? Neuroscientist A Review Journal Bringing Neurobiology Neurology & Psychiatry 19, 101-108 (2012)
[6] Qi, S., Meesters, S., Nicolay, K., Bm, T.H.R., Ossenblok, P.: Structural Brain Network: What is the Effect of LiFE Optimization of Whole Brain Tractography? Frontiers in Computational Neuroscience 10, (2016)
[7] Parker, C.S., Deligianni, F., Cardoso, M.J., Daga, P., Modat, M., Dayan, M., Clark, C.A., Ourselin, S., Clayden, J.D.: Consensus between pipelines in structural brain networks. Plos One 9, e111262-e111262 (2014)
[8] Govindan, R.M., Makki, M.I., Wilson, B.J., Behen, M.E., Chugani, H.T.: Abnormal water diffusivity in corticostriatal projections in children with Tourette syndrome. Human brain mapping 31, 1665-1674 (2010)
[9] Neuner, I., Kupriyanova, Y., Stocker, T., Huang, R., Posnansky, O., Schneider, F., Tittgemeyer, M., Shah, N.J.: White-matter abnormalities in Tourette syndrome extend beyond motor pathways. Neuroimage 51, 1184-1193 (2010)
[10] 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)
[11] Liu, F., Wee, C.Y., Chen, H., Shen, D.: Inter-modality relationship constrained multi-modality multi-task feature selection for Alzheimer's Disease and mild cognitive impairment identification. Neuroimage 84, 466-475 (2014)

[12] Chi, M., Guo, S., Ning, Y., Li, J., Qi, H., Gao, M., Wang, J., Hu, X., Guo, Y., Yang, Y.: Using support vector machine to identify imaging biomarkers of major depressive disorder and anxious depression. Springer Berlin Heidelberg (2014)
[13] 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)
[14] Dai, D., He, H., Vogelstein, J.T., Hou, Z.: Accurate prediction of AD patients using cortical thickness networks. Machine Vision & Applications 24, 1445-1457 (2013)
[15] Greene, D.J., Church, J.A., Dosenbach, N.U.F., Nielsen, A.N., Adeyemo, B., Nardos, B., Petersen, S.E., Black, K.J., Schlaggar, B.L.: Multivariate pattern classification of pediatric Tourette syndrome using functional connectivity MRI. Developmental Science 19, 581–598 (2016)
[16] Wen, H., Liu, Y., Wang, J., Zhang, J., Peng, Y., He, H.: A diagnosis model for early Tourette syndrome children based on brain structural network characteristics. In: SPIE Medical Imaging, pp. 97852R-97852R-97859. International Society for Optics and Photonics, (2016)
[17] Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., Hwang, D.U.: Complex networks: Structure and dynamics. Physics Reports 424, 175-308 (2015)
[18] Papo, D., Zanin, M., Pinedapardo, J.A., Boccaletti, S., Buldú, J.M.: Functional brain networks: great expectations, hard times and the big leap forward. Philosophical Transactions of the Royal Society B Biological Sciences 369, 33 (2014)
[19] Golden, G.S.: The effect of central nervous system stimulants on Tourette syndrome. Ann Neurol 2, 69–70 (1977)
[20] Scahill, L., Riddle, M.A., McSwiggin-Hardin, M., Ort, S.I., King, R.A., Goodman, W.K., Cicchetti, D., Leckman, J.F.: Children's Yale-Brown Obsessive Compulsive Scale: reliability and validity. J Am Acad Child Adolesc Psychiatry 36, 844-852 (1997)
[21] Retz-Junginger, P., Retz, W., Blocher, D., Stieglitz, R.D., Georg, T., Supprian, T., Wender, P.H., Rosler, M.: [Reliability and validity of the Wender-Utah-Rating-Scale short form. Retrospective assessment of symptoms for attention deficit/hyperactivity disorder]. Der Nervenarzt 74, 987-993 (2003)
[22] Wen, H., Liu, Y., Wang, J., Rekik, I., Zhang, J., Zhang, Y., Tian, H., Peng, Y., He, H.: Combining tract- and atlas-based analysis reveals microstructural abnormalities in early Tourette syndrome children. Human brain mapping 37, 1903-1919 (2016)
[23] Wen, H., Liu, Y., Rekik, I., Wang, S., Chen, Z., Zhang, J., Zhang, Y., Peng, Y., He, H.: Multi-modal multiple kernel learning for accurate identification of Tourette syndrome children. Pattern Recognition 63, 601-611 (2017)
[24] Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M.: Automated Anatomical Labeling of Activations in SPM Using a Macroscopic Anatomical Parcellation of the MNI MRI Single-Subject Brain. Neuroimage 15, 273-289 (2002)
[25] Zhang, J., Wang, J., Wu, Q., Kuang, W., Huang, X., He, Y., Gong, Q.: Disrupted Brain Connectivity Networks in Drug-Naive, First-Episode Major Depressive Disorder. Biol Psychiat 70, 334-342 (2011)
[26] Korgaonkar, M.S., Fornito, A., Williams, L.M., Grieve, S.M.: Abnormal Structural Networks Characterize Major Depressive Disorder: A Connectome Analysis. Biol Psychiat 76, 567-574 (2014)
[27] Wang, B., Mezlini, A.M., Demir, F., Fiume, M., Tu, Z., Brudno, M., Haibe-Kains, B., Goldenberg, A.: Similarity network fusion for aggregating data types on a genomic scale. Nature methods 11, 333-337 (2014)
[28] 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)
[29] Wang, J., Li, T., Wang, N., Xian, J., He, H.: Graph theoretical analysis reveals the reorganization of the brain network pattern in primary open angle glaucoma patients. European radiology 1-11 (2016)
[30] Liang, X., Wang, J., Yan, C., Shu, N., Xu, K., Gong, G., He, Y.: Effects of different correlation metrics and preprocessing factors on small-world brain functional networks: a resting-state functional MRI study. Plos One 7, e32766 (2012)
[31] 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)
[32] Wang, J., Zuo, X., He, Y.: Graph-Based Network Analysis of Resting-State Functional MRI. Frontiers in Systems Neuroscience 4, 16 (2010)

[33] Haller, S., Nguyen, D., Rodriguez, C., Emch, J., Gold, G., Bartsch, A., Lovblad, K.O., Giannakopoulos, P.: Individual prediction of cognitive decline in mild cognitive impairment using support vector machine-based analysis of diffusion tensor imaging data. Journal of Alzheimer's disease : JAD 22, 315-327 (2010)
[34] O'Dwyer, L., Lamberton, F., Bokde, A.L.W., Ewers, M., Faluyi, Y.O., Tanner, C., Mazoyer, B., O'Neill, D., Bartley, M., Collins, D.R., Coughlan, T., Prvulovic, D., Hampel, H.: Using Support Vector Machines with Multiple Indices of Diffusion for Automated Classification of Mild Cognitive Impairment. Plos One 7, (2012)
[35] O'Dwyer, L., Lamberton, F., Matura, S., Scheibe, M., Miller, J., Rujescu, D., Prvulovic, D., Hampel, H.: White Matter Differences between Healthy Young ApoE4 Carriers and Non-Carriers Identified with Tractography and Support Vector Machines. Plos One 7, (2012)
[36] Wen, H., Liu, Y., Wang, J., Zhang, J., Peng, Y., He, H.: Using support vector machines with tract-based spatial statistics for automated classification of Tourette syndrome children. In: SPIE Medical Imaging, pp. 97852Q-97852Q-97859. International Society for Optics and Photonics, (2016)