

Discriminative Subnetwork Mining for Multiple Thresholded Connectivity-Networks-Based Classification of Mild Cognitive Impairment

Fei Fei, Biao Jie, Lipeng Wang, Daoqiang Zhang*

Dept. of Computer Science and Engineering,

Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China

Email: {feifei, jbiao, wanglipeng, dqzhang}@nuaa.edu.cn

Abstract—Recent studies on brain connectivity networks have suggested that many brain diseases, such as, Alzheimer’s disease (AD) and mild cognitive impairment (MCI), are related with large-scale connectivity networks, rather than individual brain regions. However, it is challenging to find those networks from the whole connectivity network due to the complexity of brain networks. In this paper, we propose a novel method to mine the discriminative subnetworks for classifying MCI patients from healthy controls (HC). Specifically, we first apply multiple thresholds to generate multiple thresholded connectivity networks, and extract a set of frequent subnetworks from each of the two groups (i.e., MCI and HC), respectively. Then, we measure the discriminative ability of those frequent subnetworks using graph-kernel-based classification method and select the most discriminative subnetworks for subsequent classification. The results on the functional connectivity networks of 12 MCI and 25 HC show that our method can obtain a competitive results compared with state-of-the-art methods on MCI classification.

I. INTRODUCTION

Alzheimer’s disease (AD), characterized by a progressive impairment of cognitive and memory functions, is one of the most prevalent neurodegenerative brain diseases in elderly people. AD is the most common form of dementia worldwide and it is predicted that AD will affect 1 in 85 people by 2050 [1]. The prodromal stage of AD is called mild cognitive impairment (MCI), which is an intermediate state of cognitive function between normal aging and dementia. Existing studies have shown that MCI subjects progress to clinical AD at an annual rate of 10-15% [2]. Some individuals with MCI remain stable or return normal over time, but more than half progress to dementia within 5 years [3]. Thus, accurate diagnosis of AD, especially MCI, is very important for possible early treatment and delay of the progression of the disease.

Existing studies show that we can obtain a better understanding of the brain disease pathology through exploring structural and functional interactions among brain regions [4]. Thus, many attempts have been made to map the structural and functional connectivity of human brain. Here, the structural connectivity can be mapped using the brain grey matter areas or cross-correlations in cortical thickness or volume across individual brains [5, 6]. The functional connectivity refers to functional association among brain regions [7]. Recently,

machine learning and pattern classification methods have been used in connectivity networks analysis of AD/MCI, including group comparison and individual classification. Existing studies have shown that AD/MCI is related with large-scale connectivity networks, not only on the single brain region. However, it is challenging to find those connectivity networks from the whole connectivity network due to the complexity of brain networks. To the best of our knowledge, few works have employed the subnetworks, especially discriminative subnetworks, for classification of brain diseases. In connectivity-networks-based studies of AD/MCI, threshold-based methods have been widely used for exploring the topological properties of functional connectivity networks [8]. In functional network analysis, it is noteworthy that there is no gold rule to determine the choice of threshold. Therefore, many studies propose multiple thresholds based methods for brain network research [9, 10]. For example, in one of these works [9], multiple thresholded functional connectivity networks are combined with multiple kernel based method for MCI classification, which achieves a better classification performance.

In this paper, we present a new method based on connectivity measures for functional connectivity-networks-based MCI classification. Here, our hypothesis is that there exist different frequent and discriminative subnetworks between MCI group and HC group. The main idea of our method is to directly mine the discriminative subnetworks from connectivity networks and then use them for subsequent classification between MCI patients from healthy controls. Specifically, we first construct a functional connectivity from r-fMRI images on each subject, and apply multiple predefined thresholds for generating multiple thresholded connectivity networks to reflect multiple level topological structure of the original connectivity network. Then, for each thresholded connectivity network, we mine two sets of the frequent subnetworks through respectively using subnetwork mining algorithm [11] on two different groups, and evaluate their respective classification ability on training data to select the most discriminative subnetworks. Finally, graph-kernel based method will be used for classifying MCI patients from health controls. Here, graph kernel [12] is used for measuring the topological similarity between subnetworks.

TABLE I
DEMOGRAPHIC AND CLINICAL INFORMATION OF THE PARTICIPANTS

Group	MCI	HC
No. of subjects (Male/Female)	6/6	9/16
Age (mean \pm SD)	75.0 \pm 8.0	72.9 \pm 7.9
Years of education (mean \pm SD)	18.0 \pm 4.1	15.8 \pm 4.1

We validate our proposed method on the functional connectivity networks of 12 MCI and 25 HC, and the experimental results show that our method outperforms the state-of-the-art connectivity-networks-based methods in the diagnosis of MCI.

II. MATERIALS

Table I gives the demographic and clinical information of the participants. All the recruited subjects were diagnosed by expert consensus panels. All the subjects were scanned using a 3T scanner with the following parameters: repetition time (TR) = 2000ms, echo time (TE) = 32ms, flip angle = 77°, acquisition matrix = 64 \times 64, FOV = 256 \times 256mm², 34 slices, 150 volumes, and voxel size = 4mm.

III. METHOD

Fig. 1 gives the framework of our proposed method, which includes four main steps: I) *preprocessing and connectivity network construction*, II) *frequent subnetwork mining*, III) *discriminative subnetwork mining*, IV) *classification*.

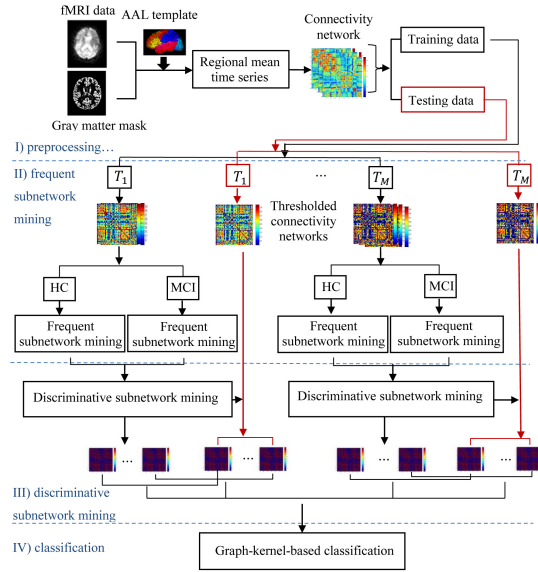


Fig. 1. The framework of the proposed method

A. Preprocessing and Connectivity Network Construction

The fMRI images were first preprocessed by applying the typical procedures of slice timing, motion correlation and spatial normalization using the Statistical Parametric Mapping software package (SPM8) (www.fil.ion.ucl.ac.uk.spm). Then the brain space of fMRI images of each subject were

parcellated into 116 regions-of-interests (ROIs) based on the Automated Anatomical Labeling (AAL) atlas [13]. The mean fMRI time series of each subject were then computed by averaging the GM-masked fMRI time series of the voxels in the ROI. The mean time series of each region was band-pass filtered (0.025-0.1Hz) based on the investigation that the fMRI dynamics of neuronal activities are most salient within this frequency range. Thus, for each subject a functional connectivity network was constructed with each ROI as a node and the Pearson correlation between ROIs as connectivity strength. Fisher's r-to-z transformation was applied on the elements of the functional connectivity network to improve the normality of the correlation coefficients. Moreover, since the functional connectivity network is intrinsically fully connected and weighted, we adopt multiple predefined values to generate multiple thresholded connectivity networks to reflect multi-level topological structure of the original connectivity network.

B. Frequent Subnetwork (subgraph) Mining

In this section, we will introduce the frequent subgraph mining algorithm to discover the most frequent subnetworks for MCI and HC, respectively. Here, in this study we adopt the well-known gSpan algorithm [11] for mining the frequent subnetworks from thresholded connectivity networks because of its time efficiency. We first start with some preliminaries, which are used to derive the gSpan algorithm [11].

Definition 1 (Labeled Undirected Graph). Let $G = (V, E, L, l)$ be a labeled undirected graph, where V is a set of nodes and $E \subseteq V \times V$ is a set of edges. $e = \{u, v\}$ indicates a edge between the nodes u and v . L is a set of labels, and l is a mapping function that assigns labels to vertices V and edges E .

Definition 2 (Subgraph). For two labeled undirected graphs $G_s = (V_s, E_s, L_s, l_s)$ and $G = (V, E, L, l)$, we say G_s is a subgraph of G if $V_s \subseteq V, E_s \subseteq E, L_s \subseteq L$ and $l_s = l$.

Definition 3 (Subgraph Frequency). Given a set of graph \mathbb{G} , the frequency of a subgraph g_s is defined as:

$$fq(g_s|\mathbb{G}) = \frac{|g_s \text{ is a subgraph of } g \text{ and } g \in \mathbb{G}|}{|\mathbb{G}|} \quad (1)$$

Definition 4 (Frequent Subgraph Mining). Given a set of labeled undirected graphs \mathbb{G} and a support parameter s where $0 < s \leq 1$, find all undirected graphs that are subgraphs in at least $s \cdot |\mathbb{G}|$ of the input graphs.

Definition 5 (Intersect-graph). Given two graphs $G_1 = (V_1, E_1, L_1, l_1)$ and $G_2 = (V_2, E_2, L_2, l_2)$, the intersect-graph $G' = (V', E')$ (denoted as $G_1 \cap G_2$) is defined as: $E' = E_1 \cap E_2$, all the nodes in E' form the nodes set V' .

gSpan algorithm first constructs a new lexicographic order among graphs, and maps each graph into a unique minimum depth-first-search (DFS) code as its canonical label. Based on this lexicographic order, gSpan algorithm utilizes the DFS strategy to mine frequent connected subgraphs efficiently. The performance of gSpan algorithm is better than other methods

in a bigger graph set with lower minimum supports. More details about gSpan algorithm see [11].

C. Discriminative Subnetwork Mining (DSM)

It is worth noting that gSpan is only used for mining the frequent subnetwork from each thresholded connectivity network, which by itself doesn't have discriminative power. Accordingly, we perform gSpan to extract two sets of frequent subnetworks from MCI group and HC group, respectively. However, some of the frequent subnetworks may still have less discriminative information for classification. Therefore, we further mine the most discriminative subnetworks according to their classification power on training subjects. Specifically, we first choose the same number of frequent subnetworks from each group, and construct multiple pairs of subnetworks across two groups. For each pair of frequent subnetworks, we utilize graph kernel proposed in [12] to measure the similarity between the training data and the frequent subnetworks and classify the training data to the class with high graph kernel value. After all, we choose those pairs of frequent subnetworks with the best classification accuracy as the most discriminative subnetworks. Finally, all the discriminative subnetworks from different threshold levels form the final discriminative subnetworks set.

Algorithm 1 summarizes the details of our proposed discriminative subnetwork mining (DSM) algorithm. Here, \mathbb{D} denotes the training set including connectivity networks of all training subjects, and MCI and HC represent the MCI and HC groups on the training set, respectively. Let G_i denote a sample in the dataset and y_i be the corresponding label, and S_1 and S_2 are two sets of frequent subnetworks mined from MCI and HC groups, respectively. Also, S_1^i (S_2^i) represents i -th subnetwork in S_1 (S_2), and G_1^i (G_2^i) is i -th intersect-graph between G and S_1^i (S_2^i). Finally, \mathbb{DS}_1 and \mathbb{DS}_2 are two sets of selected discriminative subnetworks of MCI and HC, respectively.

D. Classification

For classification of testing subject, we also compute the graph kernel between each discriminative subnetwork and the intersect-graph between testing subject and that discriminative subnetwork, and then perform graph kernel based classification. Specifically, we obtain two sets of graph kernel value, one is gotten by measuring the topological similarity between the subnetworks from MCI group and corresponding intersect-graph of the testing subject, and the other is gotten by measuring the similarity between the subnetworks from HC group and corresponding intersect-graph of the testing subject. Then we classify the testing subject to the class with the highest graph kernel value.

Algorithm 2 gives the detailed procedure of graph-kernel-based classification. Here, \mathbb{DS}_1 and \mathbb{DS}_2 represent two sets of discriminative subnetworks which are gotten via DSM algorithm, G represents a network in test subject set \mathbb{T} , and G_1^i (G_2^i) is i -th intersect-graph between G and S_1^i (S_2^i).

Algorithm 1 Discriminative Subnetwork Mining(DSM)

Input:

Training subjects $\mathbb{D}=\{\text{MCI}, \text{HC}\}=\{(G_1, y_1), \dots, (G_i, y_i), \dots, (G_N, y_N)\}$, multiple thresholds set T

Output:

Two sets of discriminative subnetworks \mathbb{DS}_1 and \mathbb{DS}_2

- 1: **for** each $t \in T$ **do**
 - 2: Threshold the connectivity networks using t
 - 3: gSpan(MCI, S_1), gSpan(HC, S_2);
 - 4: Initialize a temporary list $C = []$;
 - 5: **for** $i = 1 : n$ **do**
 - 6: **for** each $G \in \mathbb{D}$ **do**
 - 7: $G_1^i = G \cap S_1^i$, $G_2^i = G \cap S_2^i$;
 - 8: Compute the graph kernel on G_1^i , S_1^i and G_2^i , S_2^i , respectively;
 - 9: Classify G to the class with larger graph kernel value;
 - 10: Compute the accuracy c on \mathbb{D} ;
 - 11: Update list $C = [C, c]$;
 - 12: Sort S_1 , S_2 according to the C in an descending order;
 - 13: Select the top k subnetworks of S_1 (S_2) as discriminative subnetworks and add them into \mathbb{DS}_1 (\mathbb{DS}_2), respectively;
-

Algorithm 2 Graph-kernel based classification

Input:

Discriminative subnetworks set \mathbb{DS}_1 and \mathbb{DS}_2 , testing subject set \mathbb{T}

Output:

Classification accuracy acc

- 1: **for** each $G \in \mathbb{T}$ **do**
 - 2: **for** $i = 1 : k$ **do**
 - 3: $G_1^i = G \cap \mathbb{DS}_1^i$, $G_2^i = G \cap \mathbb{DS}_2^i$;
 - 4: Compute the graph kernel on G_1^i, \mathbb{DS}_1^i and G_2^i, \mathbb{DS}_2^i , respectively;
 - 5: Classify G to the class with larger graph-kernel value;
 - 6: Compute the accuracy acc on \mathbb{T} ;
-

IV. RESULTS

In the experiment, the Leave-One-Out (LOO) cross-validation strategy was adopted to evaluate the performance of our proposed method. Following the work in [9], five threshold values (i.e., $T=\{0.2, 0.3, 0.38, 0.4, 0.45\}$) were adopted to threshold the connectivity networks. The gSpan algorithm was performed with support above $s = 80\%$. The classification performance was evaluated based on classification accuracy and area under receiver operating characteristic (ROC) curve (AUC). The proposed method was compared with two connectivity-networks-based methods using graph kernel. Specifically, in the first method (denoted as Baseline), graph kernel is computed on the original thresholded connectivity networks without feature (or subnetwork) selection, while in the second method (denoted as Jie et al. [9]) graph kernel is computed on the subnetworks with selected brain regions

using a feature selection procedure. Table II gives the classification performances of all compared methods. As can be seen from Table II, our proposed method achieves the best overall classification accuracy of 97.3% with an increment of at least 5.4% from the other compared methods. Actually, only one MCI subject is misclassified by our method. Moreover, our method can also achieve best classification performance at each threshold level. On the other hand, we also compared

TABLE II
CLASSIFICATION PERFORMANCE OF DIFFERENT METHODS

	Methods	T_1	T_2	T_3	T_4	T_5	ALL
Accuracy (%)	Baseline	73.0	73.0	70.3	73.0	75.7	81.1
	Jie et al. [9]	86.5	83.8	75.7	75.7	64.9	91.9
	Proposed	89.2	97.3	97.3	91.9	91.9	97.3
AUC	Baseline	0.51	0.79	0.63	0.83	0.71	0.87
	Jie et al. [9]	0.85	0.86	0.77	0.78	0.60	0.94
	Proposed	0.90	0.96	0.96	0.86	0.94	0.96

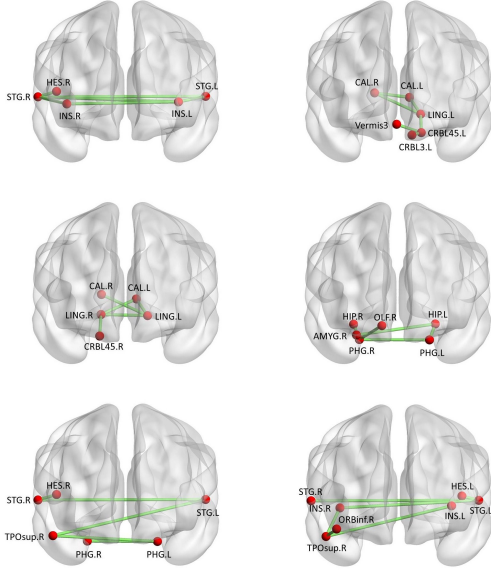


Fig. 2. The discriminative subnetworks of HC (the first column) and MCI (the second column). STG.R= Right superior temporal gyrus, STG.L= Left superior temporal gyrus, HES.R= Right transverse temporal gyri, HES.L= Left transverse temporal gyri, INS.R= Right insula, INS.L= Left insula, CAL.R= Right calcarine sulcus, CAL.L= Left calcarine sulcus, LING.R= Right lingual gyrus, LING.L= Left lingual gyrus, CRBL45.R= Right lobule IV, V of cerebellar hemisphere, CRBL45.L= Left lobule IV, V of cerebellar hemisphere, TPOsup.R= Right superior temporal pole, PHG.R= Right parahippocampal, PHG.L= Left parahippocampal, Vermis3= Lobule III of vermis, CRBL3.L= Left Lobule III of cerebellar hemisphere, HIP.R= Right hippocampus, HIP.L= Left hippocampus, OLF.R= Right olfactory cortex, AMYG.R= Right amygdala, ORBinf.R= Right orbital part of inferior frontal gyrus.

our proposed method with the recent state-of-the-art methods for connectivity-networks-based MCI classification, including Wee et al.'s method [14] and Jie et al.'s method [10]. The accuracy of 91.9% and AUC of 0.90 were reported in [14], while in [10] the accuracy of 91.9% and AUC of 0.87 were reported. Obviously, those results are inferior to the performance of our proposed method.

Fig. 2 shows the mined discriminative subnetworks of HC and MCI groups, which suggest possible disruptions in

connectivity between these regions as reported in previous studies [7]. Those pictures were visualized with the BrainNet Viewer [15] (<http://www.nitrc.org/projects/bnv/>).

V. CONCLUSION

In this paper, we have proposed a connectivity-networks-based classification framework to identify automatically MCI patients from HC. The core of the proposed method involves a discriminative subnetwork mining algorithm, which is used to discover the most discriminative subnetworks of MCI and HC. Moreover, multiple thresholds are used to threshold the functional connectivity networks to generate multiple thresholded connectivity networks with correspondingly different levels of topological structure. Finally, we combine the discriminative subnetworks from different threshold levels together and use graph-kernel-based classification method for the identification of MCI from HC. Experimental results showed the effectiveness of the proposed method. In the future, we will further evaluate our proposed method with larger dataset.

ACKNOWLEDGEMENT

This work was supported in part by the Jiangsu Natural Science Foundation for Distinguished Young Scholar (No. BK20130034), the SRFDP of Higher Education (No. 20123218110009) and the NUAA Fundamental Research Funds (No. NE2013105).

REFERENCES

- [1] Brookmeyer, R., et al., Forecasting the global burden of Alzheimer's disease. *Alzheimers Dement*, 2007. 3(3): p. 186-91.
- [2] Petersen, R.C., et al., Current concepts in mild cognitive impairment. *Arch Neurol*, 2001. 58(12): p. 1985-92.
- [3] Gauthier, S., et al., Mild cognitive impairment. *The Lancet*, 2006. 367(9518): p. 1262-1270.
- [4] Sporns, O., From simple graphs to the connectome: networks in neuroimaging. *Neuroimage*, 2012. 62(2): p. 881-6.
- [5] He, Y., Z.J. Chen, and A.C. Evans, Small-world anatomical networks in the human brain revealed by cortical thickness from MRI. *Cereb Cortex*, 2007. 17(10): p. 2407-19.
- [6] Iturria-Medina, Y., et al., Characterizing brain anatomical connections using diffusion weighted MRI and graph theory. *Neuroimage*, 2007. 36(3): p. 645-660.
- [7] Wang, J.H., et al., Disrupted functional brain connectome in individuals at risk for Alzheimer's disease. *Biol Psychiat*, 2013. 73(5): p. 472-481.
- [8] Sanz-Arigita, E.J., et al., Loss of 'small-world' networks in Alzheimer's disease: graph analysis of fMRI resting-state functional connectivity. *Plos One*, 2010. 5(11): e13788.
- [9] Jie, B., et al., Topological graph kernel on multiple thresholded functional connectivity networks for mild cognitive impairment classification. *Hum Brain Mapp*, 2013. DOI: 10.1002/hbm.22353
- [10] Jie, B., et al., Integration of network topological and connectivity properties for neuroimaging classification. *IEEE Trans Biomed Eng*, 2014. 61(2): p. 576-589.
- [11] Yan, X.F. and J.W. Han, gSpan: Graph-based substructure pattern mining. 2002 IEEE International Conference on Data Mining, Proceedings, 2002: p. 721-724.
- [12] Shervashidze, N., et al., Weisfeiler-Lehman graph kernels. *J Mach Learn Res*, 2011. 12: p. 2539-2561.
- [13] Tzourio-Mazoyer, N., et al., Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage*, 2002. 15(1): p. 273-289
- [14] Wee, C.Y., Yap, P.T., Zhang, D., Wang, L., Shen, D.: Group-constrained sparse fMRI connectivity modeling for mild cognitive impairment identification. *Brain Struct Funct*, 2013. DOI 10.1007/s00429-013-0524-8
- [15] Xia, M.R., J.H. Wang, and Y. He, BrainNet Viewer: a network visualization tool for human brain connectomics. *Plos One*, 2013. 8(7): e68910.