

# Identification of Discriminative Subgraph Patterns in fMRI Brain Networks in Bipolar Affective Disorder

Bokai Cao<sup>1</sup>(✉), Liang Zhan<sup>2</sup>, Xiangnan Kong<sup>3</sup>, Philip S. Yu<sup>1,4</sup>,  
Nathalie Vizueta<sup>5</sup>, Lori L. Altshuler<sup>5</sup>, and Alex D. Leow<sup>6</sup>

<sup>1</sup> Department of Computer Science, University of Illinois, Chicago, IL, USA  
{caobokai,psyu}@uic.edu

<sup>2</sup> Laboratory of Neuro Imaging, Department of Neurology,  
UCLA, Los Angeles, CA, USA  
zhan.liang@gmail.com

<sup>3</sup> Department of Computer Science, Worcester Polytechnic Institute,  
Worcester, MA, USA  
xkong@wpi.edu

<sup>4</sup> Institute for Data Science, Tsinghua University, Beijing, China

<sup>5</sup> Department of Psychiatry and Behavioral Sciences, UCLA Semel Institute  
for Neuroscience and Human Behavior, Los Angeles, CA, USA  
nathalievizueta@ucla.edu, laltshuler@mednet.ucla.edu

<sup>6</sup> Department of Psychiatry, University of Illinois, Chicago, IL, USA  
alexfeuillet@gmail.com

**Abstract.** Using sophisticated graph-theoretical analyses, modern magnetic resonance imaging techniques have allowed us to model the human brain as a brain connectivity network or a *graph*. In a brain network, the nodes of the network correspond to a set of brain regions and the link or edges correspond to the functional or structural connectivity between these regions. The linkage structure in brain networks can encode valuable information about the organizational properties of the human brain as a whole. However, the complexity of such linkage information raises major challenges in the era of big data in brain informatics. Conventional approaches on brain networks primarily focus on local patterns within select brain regions or pairwise connectivity between regions. By contrast, in this study, we proposed a graph mining framework based on state-of-the-art data mining techniques. Using a statistical test based on the G-test, we validated this framework in a sample of euthymic bipolar I subjects, and identified abnormal subgraph patterns in the rsfMRI networks of these subjects relative to healthy controls.

**Keywords:** Data mining · Bipolar disorder · Brain network · Subgraph pattern · Feature selection

## 1 Introduction

To correctly diagnose and properly treat neuropsychiatric disorders, many different diagnostic tools and techniques have been developed over the last decade,

often yielding a large amount of data measurements. Especially, recent advances in neuroimaging technology have provided an efficient and noninvasive way of studying the structural and functional connectivity of the human brain, either in a normal or a diseased state [16]. This can be attributed in part to advances in magnetic resonance imaging (MRI) capabilities [13]. Functional MRI (fMRI) is a functional neuroimaging procedure that identifies localized patterns of brain activation by detecting associated changes in cerebral blood flow [3, 14, 15].

fMRI scans consist of activations of tens of thousands of voxels over time, among which a complex interaction of signals and noise is embedded [7]. Using relevant spatio-temporal brain activity tensor data from these scans, the underlying brain network, here also called a connectome [17], can be computed. The functional connectome provides a graph-theoretical viewpoint to investigate the collective pattern of functional activity across all brain regions, and has been shown to be abnormal in neuropsychiatric disorders [10]. Indeed, brain networks, both structural and functional, have been increasingly studied in recent years [1, 5, 17], with potential applications to the early detection of brain diseases [19].

To date, conventional network approaches primarily focus on local patterns at the level of brain regions [9, 24] or pairwise connectivities [25]. Ye et al. presented a kernel-based method for selecting biomarkers or brain regions from multiple heterogeneous data sources that may play more important roles than others in confirming an Alzheimer’s disease (AD) diagnosis [24]. Similarly, Huang et al. introduced a sparse composite linear discriminant analysis model for identifying disease-related brain regions in AD [9].

In contrast to detecting single brain regions as biomarkers, Zalesky et al. proposed a network-based statistic approach to identify a collection of pairwise connections, some forming subnetworks, that is abnormal in patients with schizophrenia [25]. Here, thresholding was applied to pairwise connections, thus requiring each candidate link to be statistically significant under pre-defined criteria. However, links may not be discriminative by themselves until they form a component. Moreover, only connected components present in the set of suprathreshold links were examined in [25]. However, a component may lose its significance by incorporating links that are uncorrelated with other links in the component. Therefore, a full set of candidate components need to be investigated, including those with separately insignificant links and those with a subset of significant links. In this study, we focus on integrating statistical analysis and graph mining algorithms to identify subgraph patterns that distinguish rsfMRI networks obtained from two diagnostic groups (subjects with bipolar disorder versus healthy controls).

## 2 Method

We use a subgraph mining algorithm to analyze discriminative patterns in fMRI brain networks, which are also referred to as graphs hereafter.

**Definition 1 (Binary Graph).** A binary graph is represented as  $G = (V, E)$ , where  $V = \{v_1, \dots, v_{n_v}\}$  is the set of vertices,  $E \subseteq V \times V$  is the set of deterministic edges.

Let  $\mathcal{D} = \{(G_i, y_i)\}_{i=1}^n$  denote a graph dataset. All graphs in the dataset share a given set of nodes  $V$ , which corresponds to a specific brain parcellation scheme. Each graph  $G_i$  is associated with a label  $y_i$  based on the diagnosis of this subject. Here, if a subject has bipolar disorder, the corresponding graph is labeled positive. By contrast, if a subject is in the control group, the graph is labeled negative.

## 2.1 Subgraph Patterns in Brain Networks

In brain network analysis, the ideal patterns we want to mine from the data should balance local with global graph topological information. Subgraph patterns are a desired option, which can simultaneously model the network connectivity patterns among nodes and still capture changes on a local scale [13].

**Definition 2 (Subgraph).** Let  $G' = (V', E')$  and  $G = (V, E)$  be two binary graphs.  $G'$  is a subgraph of  $G$  (denoted as  $G' \subseteq G$ ) iff  $V' \subseteq V$  and  $E' \subseteq E$ . If  $G'$  is a subgraph of  $G$ , then  $G$  is supergraph of  $G'$ .

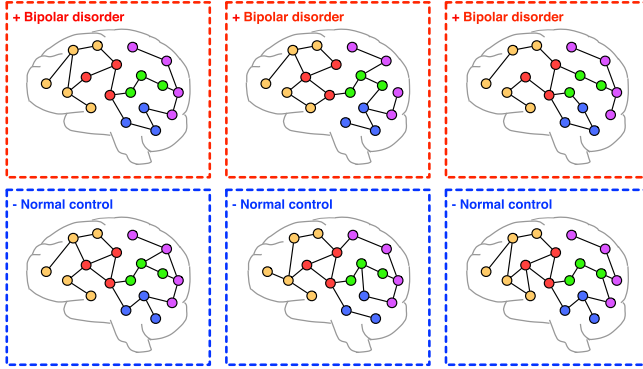
A subgraph pattern, in a brain network, represents a collection of brain regions and their connections. In other fields, mining subgraph patterns from graph data has been extensively studied by many researchers [4, 11, 18, 23]. In general, a variety of filtering criteria are proposed. A typical evaluation criterion is frequency, which aims at searching for frequently appearing subgraph patterns in a graph dataset satisfying a pre-specified value as minimum support. Most of the frequent subgraph mining approaches are unsupervised such that the discrimination power of identified subgraph patterns can not be guaranteed.

In contrast to frequent subgraph patterns, we want to mine discriminative patterns that can be used to distinguish subjects with bipolar disorder from normal controls. For example, as shown in Fig. 1, three brain regions (red nodes) may work collaboratively for normal people and abnormal connections between them can result in a diseased state (*e.g.*, bipolar disorder). Thus, it is valuable to understand which connections collectively play a significant role subserving the underlying disease mechanisms by finding discriminative subgraph patterns.

In this study, we use the G-test as the selection criterion for a subgraph pattern. It tests the null hypothesis that the frequency of a pattern in the positively-labeled (*i.e.*, bipolar disorder) graph fits its distribution in the negatively-labeled (*i.e.*, normal controls) counterparts. Rejecting the null hypothesis indicates a significant pattern with discrimination power. G-test score is defined as follows [23]:

$$t(g, \mathcal{D}) = 2m(p \cdot \ln \frac{p}{q} + (1 - p) \cdot \ln \frac{1 - p}{1 - q})$$

where  $m$  is the number of positive graphs, and  $p$  and  $q$  are the frequency of the subgraph pattern  $g$  in positive graphs and negative graphs, respectively,



**Fig. 1.** An example of discriminative subgraph patterns (connections between red nodes) in brain networks.

in the dataset  $\mathcal{D}$ . From the G-test score of a subgraph pattern, its statistical significance (*i.e.*,  $p$ -value) can be calculated using the chi-square distribution  $\chi^2$  with 1 degree of freedom in our case.

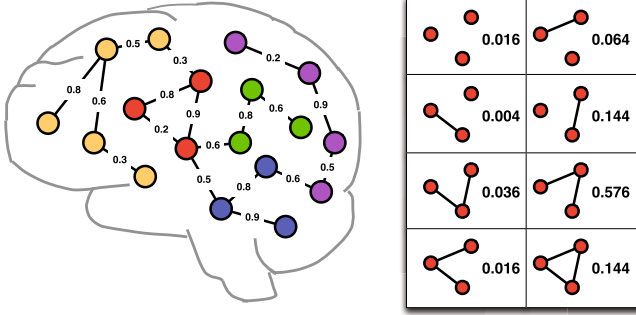
However, given that edges in fMRI brain networks are inherently weighted (correlations of time series), it is ambiguous to define whether a subgraph pattern is contained in such network, thereby difficult to determine  $p$  and  $q$ . Hence, the uncertainty information on weighted edges should be accounted for.

## 2.2 Mining fMRI Brain Networks

Conventional graph mining approaches are best suited for binary edges, where the structure of graphs is deterministic, and the binary edges represent the presence or absence of linkages between the nodes [13]. In fMRI brain network data however, there are inherently weighted edges in the graph linkage structure, as shown in Fig. 2 (left). A straightforward solution is to threshold weighted networks to yield binary networks. However, such simplification will result in potentially a great loss of information. By regarding the positive correlation of time series between two brain regions as a probability of existence for the corresponding edge, Kong et al. model fMRI brain networks as weighted graphs [12].

**Definition 3 (Weighted Graph).** A weighted graph is represented as  $\tilde{G} = (V, E, p)$ , where  $V = \{v_1, \dots, v_{n_v}\}$  is the set of vertices,  $E \subseteq V \times V$  is the set of nondeterministic edges, and  $p: E \rightarrow (0, 1]$  is a function that assigns a probability of existence to each edge in  $E$ .

For a weighted graph  $\tilde{G}(V, E, p)$ , each edge  $e \in E$  is associated with a probability  $p(e)$  indicating the likelihood of whether this edge should exist or not. It is assumed that  $p(e)$  of different edges in a weighted graph are independent from each other. Therefore, by enumerating the possible existence of all the edges in a weighted graph, we can obtain a set of binary graphs. Formally, we denote  $G$



**Fig. 2.** An example of fMRI brain networks (left) and all the possible instantiations of linkage structures between red nodes (right).

implied from  $\tilde{G}$  as  $\tilde{G} \Rightarrow G$  with a probability, as follows [12]:

$$\Pr(\tilde{G} \Rightarrow G) = \prod_{e \in E(G)} p(e) \prod_{e \in E(\tilde{G}) - E(G)} (1 - p(e)) \quad (1)$$

For example, in Fig. 2 (right), consider the three red nodes and links between them as a weighted graph. There are  $2^3 = 8$  binary graphs that can be implied with different probabilities, computed using Eq. (1).

Suppose we are given a weighted graph dataset  $\tilde{\mathcal{D}} = \{(\tilde{G}_i, y_i)\}_{i=1}^n$ . For each weighted graph  $\tilde{G}_i \in \tilde{\mathcal{D}}$ , a binary graph  $G_i$  can be implied, *i.e.*,  $\tilde{G}_i \Rightarrow G_i$ . Then a binary graph dataset  $\mathcal{D} = \{(G_i, y_i)\}_{i=1}^n$  is said to be implied from the weighted graph dataset  $\tilde{\mathcal{D}}$ , *i.e.*,  $\tilde{\mathcal{D}} \Rightarrow \mathcal{D}$ , iff  $\forall i \in \{1, \dots, n\}, \tilde{G}_i \Rightarrow G_i$ . All the possible instantiations of a weighted graph dataset  $\tilde{\mathcal{D}}$  are referred to as worlds of  $\tilde{\mathcal{D}}$ , denoted as  $\mathcal{W}(\tilde{\mathcal{D}}) = \{\mathcal{D} | \tilde{\mathcal{D}} \Rightarrow \mathcal{D}\}$ , where each world corresponds to an implied binary graph dataset  $\mathcal{D}$ . Intuitively,  $|\mathcal{W}(\tilde{\mathcal{D}})| = \prod_{i=1}^n 2^{E(\tilde{G}_i)}$ . By assuming that different weighted graphs are independent from each other, we have the probability of a binary graph dataset  $\mathcal{D} \in \mathcal{W}(\tilde{\mathcal{D}})$  being implied by  $\tilde{\mathcal{D}}$ :

$$\Pr(\tilde{\mathcal{D}} \Rightarrow \mathcal{D}) = \prod_{i=1}^n \Pr(\tilde{G}_i \Rightarrow G_i) \quad (2)$$

Thus, the expected G-test score of a subgraph pattern over the weighted fMRI brain networks can be computed as follows:

$$\text{EXP}(t(g, \tilde{\mathcal{D}})) = \sum_{\mathcal{D} \in \mathcal{W}(\tilde{\mathcal{D}})} \Pr(\tilde{\mathcal{D}} \Rightarrow \mathcal{D}) \cdot t(g, \mathcal{D})$$

By leveraging the fact that  $t(g, \mathcal{D})$  can only take a limited number of values (no more than  $n^2$ ) and the moderate-size of our brain network dataset (a small  $n$ ), we can use dynamic programming proposed in [12] to efficiently compute the expected G-test score.

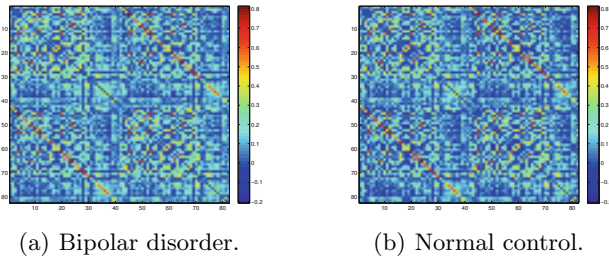
### 3 Experiments

#### 3.1 Image Data Acquisition

Our sample consists of 52 bipolar I subjects who are currently in euthymia (23 females; age =  $40.8 \pm 12.84$ ) and 45 age and gender matched healthy controls (18 females; age =  $40.15 \pm 10.77$ ). The rsfMRI scan was acquired on a 3T Siemens Trio scanner using a T2\*-weighted echo planar imaging (EPI) gradient-echo pulse sequence with integrated parallel acquisition technique (IPAT), with TR = 2 sec, TE = 25 msec, flip angle = 78, matrix = 64x64, FOV = 192 mm, in-plane voxel size = 3x3 mm, slice thickness = 3 mm, 0.75 mm gap, and 30 total interleaved slices. To allow for scanner equilibration, two TRs at the beginning of the scan were discarded. The total sequence time was 7 min and 2 sec, with 208 volumes acquired.

#### 3.2 Brain Network Construction

Functional connectomes were generated using the rsfMRI toolbox, CONN<sup>1</sup> [21]. In brief, raw EPI images were realigned, co-registered, normalized, and smoothed before analyses. Confound effects from motion artifact, white matter, and CSF were regressed out of the signal. Using the 82 labels Freesurfer-generated cortical/subcortical gray matter regions, functional brain networks were derived using pairwise BOLD signal correlations. The constructed connectivity maps are shown in Fig. 3. Here, due to the sheer amount of pairwise connections, visually it is hard to discern local differences between groups.



**Fig. 3.** Connectivity maps constructed from fMRI data. Here, the partitioning of the network is based on neuroanatomy.

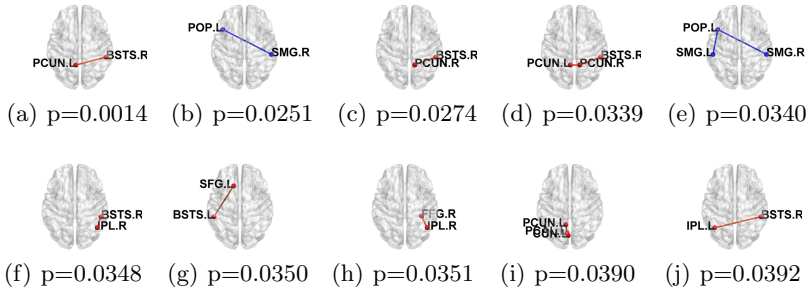
#### 3.3 Discriminative Subgraph Patterns

The most significant 10 subgraph patterns identified from these fMRI brain networks were visualized with the BrainNet Viewer<sup>2</sup> [22], as shown in Fig. 4

<sup>1</sup> <http://www.nitrc.org/projects/conn>

<sup>2</sup> <http://www.nitrc.org/projects/bnv>

where p-values are presented for each pattern. Considering our discriminative subgraph patterns were selected from a large number of candidate frequent subgraph patterns (62,776 in our dataset), multiple comparisons were accounted for by following the Benjamini-Hochberg procedure [2] to control the false discovery rate (FDR). Results indicated that these top-10 patterns would be significant with a FDR of 17.10%. Note that in contrast to the more than 60,000 subgraphs identified and tested here, conventional pairwise comparisons result in only  $\binom{82}{2} = 3,321$  comparisons.



**Fig. 4.** Discriminative subgraph patterns in the rsfMRI networks of euthymic bipolar subjects versus healthy controls. Patterns with higher frequency in bipolar subjects are shown in red (all except for b and e), while patterns with higher frequency in controls are shown in blue. See Table 1 for node abbreviations.

We observed that the patterns in Fig. 4(b) and Fig. 4(e) are shown to be more frequent in the control group while usually absent in the subjects with bipolar disorder; other patterns in Fig. 4 are present more frequently in subjects with bipolar disorder while less so in the control group. Particularly, we identified patterns in Fig. 4(d), Fig. 4(e) and Fig. 4(i) which are composed of 3 nodes, while pairwise (*i.e.*, 2 nodes) patterns have been studied before in [25]. By contrast, we investigated a full set of candidate components, including those with separately insignificant links and those with a subset of significant links.

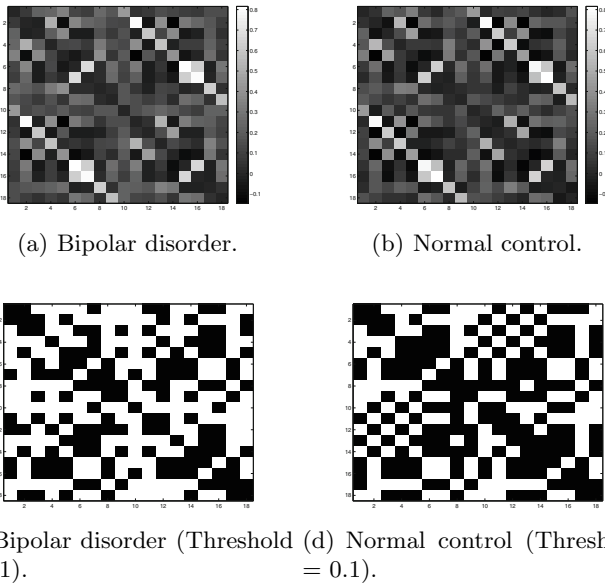
Table 1 lists the names of the nodes contributing to the top-10 discriminative subgraph patterns in Fig. 4. To better visualize the different connectivity patterns, the 18-by-18 subnetwork formed by these significant brain regions are additionally shown in Fig. 5.

## 4 Discussion

Thanks to recent advent of connectomics, global topological information can now be probed using state-of-the-art graph-theoretical analyses by modeling comprehensive patterns of brain connectivity as a network. In contrast to conventional connectome approaches that focus on local connectivity patterns among select regions-of-interest, in this study we proposed to employ sophisticated graph mining techniques to identify and quantify subgraph patterns in brain networks, and

**Table 1.** Node abbreviations.

Abbreviations Nodes	
BSTS.L(R)	left(right)-Banks of the superior temporal sulcus
PCUN.L(R)	left(right)-Precuneus
SMG.L(R)	left(right)-Supra-marginal
IPL.L(R)	left(right)-Inferior parietal
POP.L(R)	left(right)-Pars opercularis
CUN.L(R)	left(right)-Cuneus
PCAL.L(R)	left(right)-Peri-calcarine
SFG.L(R)	left(right)-Superior frontal
FFG.L(R)	left(right)-Fusiform



**Fig. 5.** 18-by-18 rsfMRI subnetwork showing the connectivity between the 18 brain regions listed in Table 1 (nodes 1-9 are regions in the left hemisphere, while nodes 10-18 in the right hemisphere; nodes are listed following the same order as in Table 1).

applied them to rsfMRI data acquired from a sample of subjects with bipolar disorder.

While we used the G-test to detect between-group differences in subgraph patterns, an alternative approach to determine statistical significance is bootstrapping by randomizing assignment of subjects into different diagnostic groups. However, such an approach would now render the two groups significantly different in gender and age, which would then need to be taken into account when interpreting subsequent results.



While the results presented here require further validation in the future, many of the identified regions that distinguished bipolar subjects from healthy controls are known regions important for mood regulation and self-referential operations in the temporal and parietal lobes. For example, as integral part of the default mode network, the inferior parietal lobule and the precuneus are known to be instrumental for the self-referential operations (thinking about self) human brains engage in at rest (but disengage during tasks) [8]. The default mode network thus is considered a task-negative network, whose function has been shown to be abnormal in both unipolar depression and bipolar disorder [6].

In the future, the technique proposed here can be easily adapted and applied to multi-modal imaging data thanks to the existence of a variety of neuroimaging techniques that characterize the brain structure and/or function from different yet complementary perspectives: diffusion tensor imaging (DTI) yields local microstructural characteristics of water diffusion; structural MRI can be used to delineate brain atrophy; fMRI records BOLD response related to neural activity; PET measures metabolic patterns [20]. Based on such a multi-modality representation, it is thus desirable to find useful patterns with rich semantics (*e.g.*, it is important to know which connectivity between brain regions is significantly altered in the context of both structure and function). Moreover, by leveraging the complementary information embedded in a multi-modality representation, better performance (*i.e.*, higher sensitivity and specificity) on disease diagnosis can be expected.

**Acknowledgments.** This work is supported in part by NSF through grants CNS-1115234.

## References

1. Ajilore, O., Zhan, L., GadElkarim, J., Zhang, A., Feusner, J.D., Yang, S., Thompson, P.M., Kumar, A., Leow, A.: Constructing the resting state structural connectome. *Frontiers in Neuroinformatics* **7** (2013)
2. Benjamini, Y., Hochberg, Y.: Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 289–300 (1995)
3. Biswal, B., Yetkin, F.Z., Haughton, V.M., Hyde, J.S.: Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magnetic Resonance in Medicine* **34**(4), 537–541 (1995)
4. Cheng, H., Lo, D., Zhou, Y., Wang, X., Yan, X.: Identifying bug signatures using discriminative graph mining. In: *ISSTA*, pp. 141–152. ACM (2009)
5. Craddock, R.C., James, G.A., Holtzheimer, P.E., Hu, X.P., Mayberg, H.S.: A whole brain fMRI atlas generated via spatially constrained spectral clustering. *Human Brain Mapping* **33**(8), 1914–1928 (2012)
6. GadElkarim, J.J., Ajilore, O., Schonfeld, D., Zhan, L., Thompson, P.M., Feusner, J.D., Kumar, A., Altshuler, L.L., Leow, A.D.: Investigating brain community structure abnormalities in bipolar disorder using path length associated community estimation. *Human Brain Mapping* **35**(5), 2253–2264 (2014)

7. Genovese, C.R., Lazar, N.A., Nichols, T.: Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *Neuroimage* **15**(4), 870–878 (2002)
8. Greicius, M.D., Krasnow, B., Reiss, A.L., Menon, V.: Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. *Proceedings of the National Academy of Sciences* **100**(1), 253–258 (2003)
9. Huang, S., Li, J., Ye, J., Wu, T., Chen, K., Fleisher, A., Reiman, E.: Identifying Alzheimer’s disease-related brain regions from multi-modality neuroimaging data using sparse composite linear discrimination analysis. In: *NIPS*, pp. 1431–1439 (2011)
10. Jie, B., Zhang, D., Gao, W., Wang, Q., Wee, C., Shen, D.: Integration of network topological and connectivity properties for neuroimaging classification. *Biomedical Engineering* **61**(2), 576 (2014)
11. Jin, N., Young, C., Wang, W.: GAIA: graph classification using evolutionary computation. In: *SIGMOD*, pp. 879–890. ACM (2010)
12. Kong, X., Ragin, A.B., Wang, X., Yu, P.S.: Discriminative feature selection for uncertain graph classification. In: *SDM*, pp. 82–93. SIAM (2013)
13. Kong, X., Yu, P.S.: Brain network analysis: a data mining perspective. *ACM SIGKDD Explorations Newsletter* **15**(2), 30–38 (2014)
14. Ogawa, S., Lee, T., Kay, A., Tank, D.: Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proceedings of the National Academy of Sciences* **87**(24), 9868–9872 (1990)
15. Ogawa, S., Lee, T.-M., Nayak, A.S., Glynn, P.: Oxygenation-sensitive contrast in magnetic resonance image of rodent brain at high magnetic fields. *Magnetic Resonance in Medicine* **14**(1), 68–78 (1990)
16. Rubinov, M., Sporns, O.: Complex network measures of brain connectivity: uses and interpretations. *Neuroimage* **52**(3), 1059–1069 (2010)
17. Sporns, O., Tononi, G., Kötter, R.: The human connectome: a structural description of the human brain. *PLoS Computational Biology* **1**(4), e42 (2005)
18. Thoma, M., Cheng, H., Gretton, A., Han, J., Kriegel, H.-P., Smola, A.J., Song, L., Philip, S.Y., Yan, X., Borgwardt, K.M.: Near-optimal supervised feature selection among frequent subgraphs. In: *SDM*, pp. 1076–1087. SIAM (2009)
19. Wang, X., Foryt, P., Ochs, R., Chung, J.-H., Wu, Y., Parrish, T., Ragin, A.B.: Abnormalities in resting-state functional connectivity in early human immunodeficiency virus infection. *Brain Connectivity* **1**(3), 207–217 (2011)
20. Wee, C.-Y., Yap, P.-T., Zhang, D., Denny, K., Browndyke, J.N., Potter, G.G., Welsh-Bohmer, K.A., Wang, L., Shen, D.: Identification of mci individuals using structural and functional connectivity networks. *Neuroimage* **59**(3), 2045–2056 (2012)
21. Whitfield-Gabrieli, S., Nieto-Castanon, A.: Conn: a functional connectivity toolbox for correlated and anticorrelated brain networks. *Brain Connectivity* **2**(3), 125–141 (2012)
22. Xia, M., Wang, J., He, Y.: Brainnet viewer: a network visualization tool for human brain connectomics. *PLoS One* **8**(7), e68910 (2013)
23. Yan, X., Cheng, H., Han, J., Yu, P.S.: Mining significant graph patterns by leap search. In: *SIGMOD*, pp. 433–444. ACM (2008)
24. Ye, J., Chen, K., Wu, T., Li, J., Zhao, Z., Patel, R., Bae, M., Janardan, R., Liu, H., Alexander, G., et al.: Heterogeneous data fusion for alzheimer’s disease study. In: *KDD*, pp. 1025–1033. ACM (2008)
25. Zalesky, A., Fornito, A., Bullmore, E.T.: Network-based statistic: identifying differences in brain networks. *Neuroimage* **53**(4), 1197–1207 (2010)