Research Article
Abnormal Functional Brain Network Metrics for Machine Learning Classifier in Depression Patients Identification

[1]Hao Guo, [2]Xiaohua Cao, [2]Zhifen Liu and [1]Junjie Chen
[1]College of Computer Science and Technology, Taiyuan University of Technology,
Taiyuan 030024, People's Republic of China
[2]Department of Psychiatry, First Hospital of Shanxi Medical University,
Taiyuan 030002, People's Republic of China

Abstract: Brain network is a widely used tool for identifying abnormal topological properties in whole-brain networks which has been applied to neurological disease diagnosis such as Major Depressive Disorder (MDD). But there is not any study showing that abnormal brain network topological metrics can be used in machine learning classification methods for the identification of MDD patients. In order to find an appropriate feature selection method, we hypothesize that MDD disrupts the topological organization of functional brain networks and the abnormal topological metrics could be used as effective features in constructing a classifier. Resting state functional brain networks were constructed for 26 healthy controls and 34 MDD patients by thresholding partial correlation matrices of 90 regions. The topological metrics, including global and local, were calculated using graph theory-based approaches. Non-parametric permutation tests were then used for group comparisons of topological metrics, which were used as classified features in support vector machine algorithm. Result showed that both the MDD and control groups showed small-world architecture in brain functional networks. However, some of the regions exhibited significantly abnormal nodal centralities, including the limbic system, basal ganglia and medial temporal and prefrontal regions. Support vector machine with radial basis kernel function algorithm exhibited the highest average accuracy (86.01%) with 28 features (p<0.05). Overall, the current study suggested that MDD is associated with abnormal functional brain network topological metrics and statistically significant network metrics can be successfully used in classification algorithms as features.

Keywords: Brain network, depression, feature selection, graph theory, machine learning, support vector machine

INTRODUCTION

Functional neuroimaging studies have suggested that Major Depressive Disorder (MDD) is related to some abnormal brain regions such as hippocampus, parahippocampal gyrus, precentral gyrus, caudate nucleus and so on Zhang *et al.* (2011). With studied on MDD in greater depth, we understood more about depression, but the past researches had only focused on specific brain region or connections between several brain regions (Greicius *et al.*, 2007). At present, there are few studies showed the global topological organization of whole-brain networks of MDD. Complex graph theoretical analysis provided a powerful research method for characterizing topological properties of brain networks and proved that functional brain networks of normal controls have typical features of small-world (He *et al.*, 2007). Abnormal brain network topological metrics which provided a new perspective of potential biomarkers for neurological disease diagnosis had been found in various brain diseases.

Machine learning and pattern recognition methods have recently been applied in the classification of MDD. Costafreda (2009) and Fu *et al.* (2008) constructed classifiers using structural and functional magnetic resonance imaging data and tested them with normal controls and MDD patients, respectively, reporting accuracy rates of 67 and 86%. Gong *et al.* (2011) investigated differences between the use of gray matter and white matter as classification features and used Support Vector Machine (SVM) algorithms to distinguish refractory and non-refractory depressive disorder, reporting accuracy rates of 65.22 and 76.09%, respectively.

As a widely used tool for identifying abnormal topological properties in whole-brain networks, brain network has been applied to neurological disease diagnosis such as Alzheimer's disease, schizophrenia and epilepsy. Despite the increased understanding of MDD, it remains unclear that the alterations of

topological metrics of the whole brain in MDD patients. To our knowledge, there is not any study showing that abnormal brain network topological metrics can be used in machine learning classification methods for the identification of MDD patients.

In order to find an appropriate feature selection method, we hypothesize that MDD disrupts the topological organization of functional brain networks and the abnormal topological metrics could be used as effective features in constructing a classifier. To verify this hypothesis, we collected Resting fMRI data from 34 drug-naive, first-episode Major Depressive Disorder (MDD) patients and 26 healthy controls to construct functional brain networks in a continuous sparsity range and analyze those networks using graph theoretical approaches to find the significant topological organization differences of functional brain network between MDD patients and normal controls in different sparsity. The Area Under the Curve (AUC) of each merics were calculated which were used as classification features in Support Vector Machine (SVM) algorithm in order to find a machine learning approach for the diagnosis of depression.

MATERIALS AND METHODS

Subjects: A total of 60 subjects were recruited, including 34 first-episode drug-naive MDD patients as MDD group and 26 age-and sex-matched healthy controls as control group. All subjects in the control group have no history of mental or neurological disorders according to SCID Non-Patient Edition and SCID-II. All subjects in MDD group are MDD patients according to the DSM-IV (SCID) (First *et al.*, 1997) and the 17-item Hamilton Rating Scale for Depression (HRSD) (Williams, 1988). All subjects were inpatients at the Department of Psychiatry, First Hospital of Shanxi Medical University. The control subjects were recruited from local communities. Written informed consent to participate in the study was obtained from all control subjects and from the first-degree relatives of all patients, which was approved by the Ethics Committee of the Chinese Academy of Medical Science. Data collection was performed during September 2010 to June 2011.

Data collection and preprocessing: All subjects underwent resting state functional MRI using a 3T MR scanner (Siemens Trio 3-Tesla scanner, Siemens, Erlangen, Germany). Image preprocessing was performed using DPARSF software (version 2.0; Yan Chao-Gan, State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing, China). The first 10 volumes were discarded and the remaining were slice timing- and head motion-corrected. Data of one control and one patient were discarded because their heads moved more than 1 mm of translation or 1° of rotation in any direction. After these corrections, the images were spatially normalized to the standard space of the Montreal Neurological

Institute (MNI) using an optimum 12-parameter affine transformation and nonlinear deformations to resample as 3-mm cubic voxels. Linearly detrended and band-pass filtered (0.01-0.08 Hz) were performed to reduce the effect of low frequency drifts and high-frequency noise.

Network construction and threshold selection:
Node definition: The current study divided the whole brain into 90 according to Automated Anatomical Labeling (AAL) (Tzourio-Mazoyer *et al.*, 2002) which as the international standard widely used in the field of brain imaging, with each representing a node of the network. Each node represents a brain region and the mean time series for all voxels in each region was calculated as part of the time-series of the corresponding node.

Edge definition: We computed partial correlation coefficients as edges in the network. Fist, the mean time series of each node was processed by regressing out the head motion profiles estimated in the image realignment from the mean time course. After that, any two nodes of the residuals of the regression analyses of were used to compute the partial correlation to acquire the correlation coefficient of mean time series of any two nodes, the method adopted was:

$$r_{ij} = \frac{\sum_{t=1}^{T}[x_i(t) - \overline{x_i}]g[x_j(t) - \overline{x_j}]}{\sqrt{\sum_{t=1}^{T}[x_i(t) - \overline{x_i}]^2}\sqrt{\sum_{t=1}^{T}[x_j(t) - \overline{x_j}]^2}} \quad (1)$$

where,
x_i & x_j : Respective residual time series of node i and node j
$\overline{x_i}$ & $\overline{x_j}$: Respective the mean time series of node i and node j

And then a 90×90 partial correlation matrix A of time series for each subject was obtained. At last, partial correlation matrices A were converted into binarized matrices B = [b_{ij}] according to a predefined threshold, if the value of the partial correlation between nodes i and j was larger than the threshold, the entry b_{ij} was 1 and if the value of the partial correlation between nodes i and j was smaller than the threshold, the entry b_{ij} was 0. In this article, we will use a series of successive thresholds to construct brain networks.

Threshold selection: For all correlation matrices, we must select a threshold to convert correlation matrices to binary matrices. Each threshold corresponds to a sparsity which was defined as the ratio of the number of existing edges divided by the maximum possible number of edges in a network. Instead of selecting a single threshold, we selected a threshold space according to the following criteria and construct brain networks corresponding to each sparsity in the threshold space:

- The average degree over all nodes of each network k constructed in the threshold space was larger than $2 \times \ln N$ with N = 90 here. N denotes the number of nodes
- The small-worldness scalar σ of each network constructed in the threshold space was larger than 1.1

Based on the above principles, the threshold spaces 8-32% was selected. All of the network analysis was within this threshold space and the interval was set as 1% throughout this study.

Network metrics and statistical analysis: For brain networks at each sparsity threshold in the threshold space, two global metrics (clustering coefficient and characteristic path length) and three local metrics (degree, betweenness centrality and nodal efficiency) were calculated.

More over we calculated the AUC for each metrics, which provides a summarized scalar for topological characterization of brain networks at different sparsity. The AUC metric has been used in previous brain network studies which suggested that AUC was very effective for studying topology properties of brain networks.

Non-parametric permutation tests were performed on the network metrics and their AUC of MDD group and control group in order to determine whether there existed significant group differences.

Classifier: To automatically identify the disease data, we used machine learning methods to construct a classifier. All of the metrics were selected as classification features, which were used in SVM algorithms (radial basis kernel function). Three different features selection methods were performed: AUC, peek value and both of them.

To compare classifier performance alterations with different numbers of features, we used a range of p-values as significance thresholds. We used five threshold levels: $p<0.005$, $p<0.015$, $p<0.05$, $p<0.10$ and $p<0.15$. The corresponding feature numbers were 4, 15, 28, 55 and 87.

A cross-validation method was used to evaluate the performance of models. We randomly selected 70% of samples as a training set and the remaining 30% as a test set. The accuracy rate of the test set was recorded. This procedure was repeated 200 times for each threshold. Cross-validation methods have been widely used in previous studies (Pereir *et al.*, 2009).

RESULTS

Global metrics: Result showed that comparing with random network, brain networks of both MDD and normal controls demonstrated an economic small-world

Table 1: Regions exhibiting abnormal nodal metrics in MDD patients compared to controls

Brain regions	p-value		
	Degree	Betweenness centrality	Nodal efficiency
MDD<control			
Left FFG	0.016	0.151	0.047
Left ORBinf	0.017	0.179	0.034
Right CUN	0.025	0.001	0.093
Right SFGmed	0.035	0.495	0.151
Right ORBmid	0.049	0.557	0.025
Left IFGoperc	0.050	0.295	0.141
Left CAL	0.058	0.547	0.046
Right IFGoperc	0.063	0.048	0.178
Right ORBsupmed	0.082	0.848	0.044
MDD>control			
Right HIP	0.001	0.005	0.003
Left ANG	0.008	0.284	0.011
Right PCG	0.008	0.167	0.007
Right THA	0.008	0.004	0.008
Right PUT	0.014	0.213	0.023
Right MOG	0.021	0.443	0.027
Right DCG	0.070	0.482	0.014
Left DCG	0.073	0.632	0.048

topology. Despite common small-world architecture, significant differences were found in global metrics between MDD patients and controls. Compared with normal controls, No significant differences were found in clustering coefficient, but MDD patients show significant low characteristic path length ($p<0.05$). That suggested that brain network of MDD group had more long distance connections and were closer to a randomized configuration.

Nodal metrics: Non-parametric permutation tests were performed on AUC of each nodal degree of brain networks in the threshold space of MDD group and control group. Compared with normal controls MDD patients showed significant increased nodal centrality in many brain regions including the limbic system (including the right hippocampus, right posterior cingulate gyrus, bilateral median cingulate and paracingulate gyri), part of the basal ganglia (right putamen and right thalamus) and left angular gyrus, among others.

Decreasing nodal centralities in MDD patients were predominantly located in several regions of medial occipital (right cuneus and left calcarine fissure and surrounding cortex), medial temporal (left fusiform gyrus) and prefrontal (left inferior frontal gyrus (orbital part), right superior frontal gyrus (medial), right middle frontal gyrus (orbital part), bilateral inferior frontal gyrus (opercular part) and right superior frontal gyrus (medial orbital)) regions. Most of the regions mentioned above, including the hippocampus, cingulate gyrus, lenticular nucleus and thalamus, are components of Limbic-Cortical-Striatal-Pallidal-Thalamic (LCSPT) tract (Table 1).

Classifiers: The AUC of the each metric value were selected as classification features and tested in SVM algorithms. The results revealed that when both AUC and peak value were selected as features, the classifier
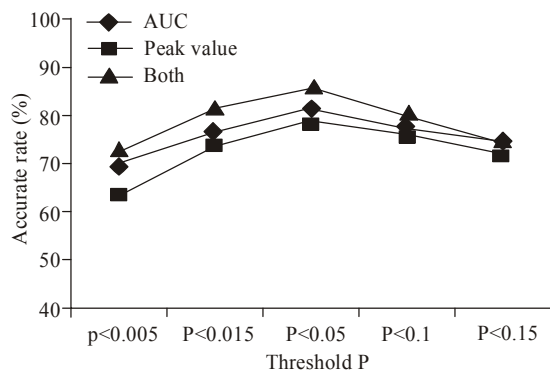
Fig. 1: Average accuracy in support vector machine algorithms with different feature selection

exhibited the best performance, with an average accuracy of five thresholds reaching 86.01% Fig. 1.

Comparing the five threshold values revealed that all the methods exhibited the highest accuracy with 28 features (p<0.05). The results suggested that different models always had an optimized feature number.

DISCUSSION

At present, studies about MDD focused on the topological organization of functional brain networks. The results of these studies reveal that compared to control group, functional brain networks in MDD group had decreased path length, implying a disturbance of the normal global integration of whole-brain networks. Moreover, many local brain regions were profoundly affected by MDD: both caudate nucleus and some of default-mode regions showed increased nodal degrees, while several regions showed decreased degrees. The result provided insights into our understanding of altered topological organization in functional brain networks of MDD.

The human brain is complex which is an interconnected and has various characteristics, such as small-world and high efficiency at a low wiring cost. In small-world networks, on the one hand, local clustering of nodes is superior to the modular information processing; on the other hand, a low characteristic path length is also needed in order to ensure global efficiency. More recently, Latora and Marchiori (2001) expanded the two measures of efficiency and cost of the small-world theory. A network that have high efficiency of information transmission must is cheap to build which is called economic small-world networks. For evaluating of functional brain networks, small world is an attractive model, because high local clustering and short characteristic path length support the two principles of functional brain network: functional separation functions and functional integration. Here, we found that whole-brain functional networks of both MDD group and control group

showed high small-world topology, by using small world model and new efficiency measures.

Despite the common small-world topology, there were significant differences between two groups in network characteristics. No significant differences in clustering coefficient of brain networks between MDD patients and controls but significant differences in characteristic path length of brain networks between MDD patients and controls. These changes of global network properties may be due to the decrease of characteristic path length in functional brain networks of MDD patients. Compared to small-world networks, random networks have low information transmission capacity and fault-tolerant ability. Therefore, the loss of some properties of the small-world network in functional brain networks of MDD group suggested that the loss of optima in functional brain networks of MDD group to further clarify the mental breakdown of organizational structure in patients with depression, reduced cognitive and emotional capacity.

We also found that degrees significantly increase in the threshold space in some brain regions of MDD group such as hippocampus, lenticular nucleus putamen and angular gyrus, which are related to emotions. Hippocampus is primarily responsible for the tasks of memory and learning, related reports indicated: if the hippocampus is damaged, the patient will lose one part or even all of memory who may be forgetful and his imagination are also affected, compared to controls (Bannerman *et al.*, 2004). Lenticular nucleus putamen is mainly responsible for movement and sensation. Angular gyrus is a brain region which connecting visual area and hearing area, if angular gyrus is damaged, the abilities of reception and expression of language will be affected (Segal *et al.*, 2010). Hippocampus, lenticular nucleus putamen and angular gyrus are parts of the limbic system which play an extremely important role in process of mood producing, memory and learning. These brain areas are also all belong to the default network, recent brain imaging studies have shown that default networks of MDD group are different from control group (Holzschneider *et al.*, 2012). Thus, our findings are consistent with previous research results. Increased nodal degree of these regions relating to emotion, suggest their strengthened roles of coordinating whole-brain networks, presumably in response to the emotional response of MDD.

While degrees decrease in the threshold space in other brain regions of MDD group including cuneus and fusiform gyrus. Some studies found that cuneus was associated with cognitive function, which was mainly responsible for self-awareness, dealing with information about oneself, episodic memory and so on. The facts proved that cuneus of first-episode drug-naive MDD patients were abnormal which was compatible with the results of our experiments (Zhu *et al.*, 2012).

We found that the classifier exhibited the best performance when both of AUC and peak value were

selected as features. AUC provides a summarized scalar for topological characterization at different sparsity, but peak value emphasizes the local characterization. If both of them were selected as features, it would intensify the contribution of the specific abnormal metrics during the classification process.

This suggests that different feature selection methods have great impact on model performance. As such, when building a classifier with brain network metric features, it is important to re-evaluate the classification algorithm.

The current study revealed that most of the classifier exhibited the highest with 28 features ($p < 0.05$). We speculated that the number of features impacted on the classifier performance rather than the p-value. Using simulation data, Hua *et al.* (2005) compared seven kinds of classifier with 0 to 200 of the sample size to find a different effect of error rate with different feature numbers, reporting highest accuracy with 28 and 30 features when the sample size was 50 for linear-SVM and LDA, in accord with the current findings.

Several important issues should be addressed in future studies. First, definition of nodes and edges are critical in brain network research and brain networks with different parcellation schemes are likely to exhibit different topological architectures (Fornito *et al.*, 2010). In the current study, functional brain networks were constructed at a regional level based on a previously published atlas. Future studies are required to determine which spatial scale is most appropriate for describing the network topology of MDD.

Second, as an exploratory analysis, the statistical results of our network metrics were not corrected. Future studies should test a large sample of MDD patients to increase statistical power.

Finally, the optimal feature number depends on a variety of factors, including sample size, classifier rules, the distribution of the sample categories, selected feature efficiency and ordering. Future studies are required to determine the optimal number of selected features for building the most effective classifier for a specific brain disease.

CONCLUSION

Constructing functional brain networks of MDD patients by the method mentioned in this article and comparing to functional brain networks of controls, we found that the MDD patients showed a decreased path length in their brain networks and properties in nodes which were related to emotion have altered as compared to controls. We could get higher classification accuracy, if we use both of peak value and AUC of the metrics as classification features to identify individuals

In summary, the present study provides evidence that MDD is associated with abnormal functional brain network topological metrics which could be used as effective features in constructing a classifier. Our research provided an initial step towards the use of brain network topological metrics. It would be clinically significant to identify individuals and predict the treatment automatically avoiding self-reported symptoms.

REFERENCES

Bannerman, D.M., J.N.P. Rawlins, S.B. McHugh, R.M.J Deacon, B.K Yee, T Bast, W.N. Zhang, H.H.J. Pothuizen and J. Feldon, 2004. Regional dissociations within the hippocampus memory and anxiety. Neurosci. Biobehav. Rev., 28: 273-283.

Costafreda, S.G., C. Chu, J. Ashburner and C.H. Fu, 2009. Prognostic and diagnostic potential of the structural neuroanatomy of depression. PLoS ONE, 7: e6353.

First, M., R. Spitzer, M. Gibbon and J. Williams, 1997. Structured Clinical Interview for DSM-IV Axis I Disorders. American Psychiatric Press, Washington DC.

Fornito, A., A. Zalesky and E.T. Bullmore, 2010. Network scaling effects in graph analytic studies of human resting-state FMRI data. Front. Syst. Neurosci., 4: 22.

Fu, C.H.Y., J. Mourao-Miranda, S.G. Costafreda, A. Khanna, A.F. Marquand, S.C. Williams and M.J. Brammer, 2008. Pattern classification of sad facial processing: Toward the development of neurobiological markers in depression. Biol. Psychiatry, 63: 656-662.

Gong, Q., Q. Wu, C. Scarpazza, S. Lui, Z. Jia, A. Marquand, X. Huang, P. McGuire and A. Mechelli, 2011. Prognostic prediction of therapeutic response in depression using high-field MR imaging. Neuroimage, 55: 1497-1503.

Greicius, M.D., B.H. Flores, V. Menon, G.H. Glover, H.B. Solvason, H. Kenna, A.L. Reiss and A.F. Schatzberg, 2007. Resting-State functional connectivity in major depression: Abnormally increased contributions from subgenual cingulate cortex and thalamus. Biol. Psychiatry, 62: 429-437

He, Y., Z.J. Chen and Evans, 2007. Small-world anatomical networks in the human brainrevealed by cortical thickness from MRI. Cereb. Cortex., 17: 2407-2419.

Holzschneider, K., T. Wolbers, B. Röder and K. Hötting, 2012. Cardiovascular fitness modulates brain activation associated with spatial learning. NeuroImage, 59: 3003-3014.

Hua, J.P., Z.X. Xiong, J. Lowey, S. Edward, R.D. Edward, 2005. Optimal number of features as a function of sample size for various classification rules. Bioinformatics, 21: 1509-1515.

Latora, V. and M. Marchiori, 2001. Efficient behavior of small-world networks, Phys. Rev. Lett., 87: 198701.

Pereir, F., T. Mitchell and M. Botvinick, 2009. Machine learning classifiers and fMRI: A tutorial overview. Neuroimage, 45(Suppl 1): 199-209.

Segal, D., M.M. Haznedar, E.A. Hazlett, J.J. Entis, R.E. Newmark, Y. Torosjan, J.S. Schneiderman, J. Friedman, K.W. Chu, C.Y. Tang, M.S. Buchsbaum and P.R. Hof, 2010. Diffusion tensor anisotropy in the cingulate gyrus in schizophrenia. NeuroImage, 50: 357-365.

Tzourio-Mazoyer, N., B. Landeau, D. Papathanassiou, F. Crivello, O. Etard, N. Delcroix, B. Mazoyer, M. Joliot,. 2002. Automated anatomical labeling of activations in SPM using a macroscopic. Neuroimage, 15: 273-289.

Williams, J.B.W., 1988. A structured interview guide for the hamilton depression rating-scale. Arch Gen Psychiatry, 45: 742-747.

Zhang, J.R., J. Wang, Q. Wu, W. Kuang, X. Huang, Y. He and Q. Gong, 2011. Disrupted brain connectivity networks in drug-naive, first-episode major depressive disorder. Biol. Psychiatry, 70(4): 334-342.

Zhu, X.L., Wang, X., J. Xiao, J. Liao, M. Zhong, W. Wang and S. Yao, 2012. Evidence of a dissociation pattern in resting-statedefault mode network connectivity in first-episode, treatment-naive major depression patients. Biol. Psychiatry, 71(7): 611-617.