Brain Network Classification Based on Multiple Brain Network Attributes

Nan Yang

***Abstract*—Most existing brain network classification methods construct brain networks by processing the collected signals, and then classify them according to one brain network feature attribute between one or more brain regions. This kind of method often ignores other characteristic attributes of the brain network because only one attribute is considered, and the neglected features are likely to have a greater impact on the experimental results. This paper explores the different effects of various network attributes in brain network classification, and provides a new idea for future network attribute research.**

***Index Terms*—brain networks, support vector machine, MCI.**

I. INTRODUCTION

Support vector machine (SVM) has been used to classify brain networks[1-5] and achieve many results because of its superior performance. For example, Qi Xinru [10] of Beijing Jiaotong University and others based on rs-fMRI images. The method of constructing a complex network for each brain, using the network parameters in the complex network to predict Attention Deficit/Hyperactivity Disorder (ADHD); Shubhangi et al. [10] using support vector machine classification A real-time brain tumor detection system based on multi-spectral analysis was designed and developed. The results show that the system can accurately and effectively detect pathology and mark normal lateral slices. Bae et al. [11] proposed a SVM mixing ratio. Sampling method, which can determine various oversampling ratios of different minority categories and improve the accuracy of prediction. The results show that SVM effectively and efficiently solves the problem of multi-class imbalance; Zhou et al. [12] developed and designed An improved support vector machine: first, the SVM is trained for all training samples; secondly, the decision function contribution in the pair of training samples is removed. A small sample of the last remaining samples of re-training SVM, the classification for the motion picture based on the EEG, and achieved good results, it misclassification rate to 9.29%. However, other attributes in the network are often ignored in the classification process, which affects the classification effect.

II. RELATED TECHNOLOGY

Definition 1 The number of edges to which a node is connected is the degree of that node. Intuitively, the greater the degree of a node, the higher its "importance" in the network.

Manuscript received Oct 14, 2018

Nan Yang, School of Computer Science & Software Engineering, Tianjin Polytechnic University, Tianjin, 300387, China

Definition 2 Characteristic path length The minimum number of edges between nodes in a network is defined as the path length between the two points. The average of the path lengths between all nodes in the network is the characteristic path length of the network.

Definition 3 Clustering coefficient A clustering coefficient of a node. It is assumed that a node $v_i$in a network is connected to other nodes through the k_i strip, and the nodes adjacent to the k_i may form a common edge of $C_{K_i}^2$. The clustering coefficient $C(V_i)$ of the node is the ratio of the number of sides n between the k_i neighbors of v_i and the number of possible sides $C_{K_i}^2$, namely:

$$C(V_i) = \frac{n}{C_{k_i}^2} = \frac{2n}{k_i(k_i - 1)}$$

Definition 4 Clustering coefficient Local (CCL) Let$\langle M_{nn}(k) \rangle$ be the average number of edges existing between the neighbors of a node of degree k. The local clustering coefficient is:

$$C(k) = \frac{2\langle M_{nn}(k) \rangle}{k(k-1)}$$

Definition 5 Clustering coefficient Global (CCG) The global clustering coefficient C is defined as the ratio of the number of edges existing between neighbors of all nodes in the network to the maximum possible number of edges between them, namely:

$$C = [\sum_{k=0}^{k_{max}} k(k-1)P(k)C(k)]/[\langle k^2 \rangle - \langle k \rangle]$$

Where k is the degree of the node and is the second moment of the degree.

Definition 6 The Pearson correlation coefficient is used to reflect the degree of linear correlation between two variables. The formula is:

$$r = \frac{1}{n-1}\sum_{i=1}^{n}(\frac{X_i - \bar{X}}{s_X})(\frac{Y_i - \bar{Y}}{s_Y})$$

Where r reflects the degree of correlation, and its value is between -1 and 1. When it is positive, the two are positively correlated. When the value is negative, the two are negatively correlated. The larger the absolute value of r, the stronger the correlation. . It is worth mentioning that the brain function network is the only biological network discovered so far.

III. EXPERIMENTAL DESIGN

A total of 123 fMRI samples were obtained, including 51 MCI patients with 51 normal controls (Normal Control, NC).

Table 1 gives the statistics of the experimental sample information.

Category	Age	Gender（M/F）
MCI	75.2±3.0	24/27
Normal	75.5±3.2	35/37

First, the experimental data image is preprocessed to obtain a 116×116 brain correlation matrix. Then, using different thresholds to threshold the matrix and calculate the attribute values of some networks as the feature vectors of the classification, use SVM to classify the feature vectors of different thresholds, and select the threshold with the best classification effect. Finally, the optimal solution obtained in the previous step is merged into a new feature vector and classified training.

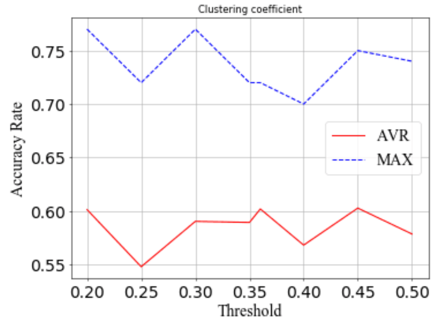
Based on the previously obtained brain network[6-9], we can calculate the parameters of the network as feature vectors of the support vector machine samples. The network parameters studied in this paper include Clustering Coefficient Global (CCG), Clustering Coefficient Local (CCL), Degree of Degree (Dgree, D), and Average Path Length (Characteristic Path Length Global, CPG), Characterization Path Local (CPL) and the original Pearson correlation matrix based on BOLD signals.

According to the literature 13 and 14, it is pointed out that the brain network has a good classification performance when the connection ratio between brain network nodes is between 25% and 75%. We will use different values to threshold the brain connection network, specifically [0.20, 0.25, 0.30, 0.35 , 0.37, 0.40, 0.45, 0.50]. Then we can get D, CC, CL three kinds of data, each data 8 groups.

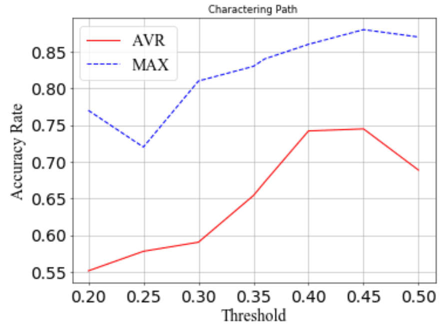
After the features are fused, the MCI [15,16]is predictively classified using a support vector machine SVM. During the experiment, 123 experimental data were randomly divided into two parts, of which the training set accounted for 1/3 of the 2/3 training set. The cross-validation method is used to adjust the model parameters to improve the classification accuracy.

IV. EXPERIMENTAL DATA ANALYSIS

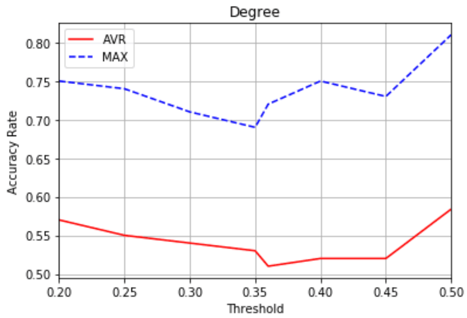
Figure 1 shows the use of clustering coefficients, characteristic path length, degrees of nodes, and the maximum and average classification accuracy of raw data at different thresholds. As shown in Fig. 1(a), the average classification accuracy of the clustering coefficient is 60.28% when the threshold is 0.45. As shown in Fig. 1(b), the average classification accuracy of characteristic path length is 74.48% when the threshold is 0.45. As shown in Fig. 1(c), the average classification accuracy of degree is 58.40% when the threshold is 0.50. As shown in Fig. 1(d), the classification accuracy distribution of a total of 50 experiments has an average classification accuracy of 69.33%. It can be seen that the classification performance of the clustering coefficient and the original data is the best. In the future experiments, we can introduce the weights according to the different classification effects in the feature fusion to improve the classification effect. Experiments show that different network attributes have different degrees of contribution to brain network classification, among which the degree of node has the least influence, the characteristic path length and clustering coefficient are obvious, which is consistent with the weakening of the small world attribute of the brain network of MCI patients previously studied. in conclusion.



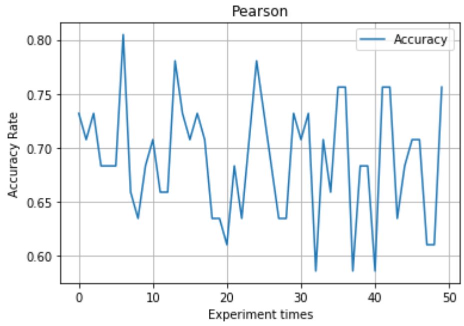
(a) Clustering coefficient



（b）Characteristic path length



（c）Degree



（d）Raw data

Figure 1 Network attribute threshold average classification accuracy rate (experiment 50 times mean)

Table 2 Comparison of relevant experimental methods

Method	Classification	Sample size (normal / patient)	Classification accuracy %
Raw data	SVM	72/51	69.33
Characteristic path	SVM	72/51	74.48
Clusteri	SVM	72/51	60.28

ng coefficient			
Degree	SVM	72/51	58.40

According to the data obtained in this experiment, the most obvious areas of MCI patients were hippocampus (Hippocampus_L, Hippocampus_R), hippocampus (ParaHippocampal_L, ParaHippocampal_R) and amygdala (Amygdala_L, Amygdala_R) [17]. It is validated that medical MCI patients have lesions in these brain regions.

V. Conclusion

This paper analyzes the prediction and classification of various network attributes of the brain network for MCI, using raw data, clustering coefficients, feature path length and degree of nodes as features. Different pairs of network attributes describe different aspects of the network and have different contributions to the classification.

At the same time, the contribution of each network attribute to classification prediction remains to be explored. In the feature fusion, some network information is often duplicated, which leads to the efficiency of classification and classification accuracy. In future research, other research should be explored. The brain network classification method solves the problem of information redundancy and further improves the accuracy and efficiency of brain network classification.

REFERENCES

[1] Cohen D. Magnetoencephalography: evidence of magnetic fields produced by alpha-rhythm currents[J]. Science, 1968, 161(3843):784.
[2] Watts D J, Strogatz S H. Collective dynamics of 'small-world' networks.[J]. Nature, 1998, 393(6684):440.
[3] Olaf S, Rolf K. Motifs in Brain Networks[J]. Plos Biology, 2004, 2(11):e369.
[4] Iturria-Medina Y, Canales-Rodríguez E J, Melie-García L, et al. Characterizing brain anatomical connections using diffusion weighted MRI and graph theory[J]. Neuroimage, 2007, 36(3):645.
[5] Iturria-Medina Y, Sotero R C, Canales-Rodríguez E J, et al. Studying the human brain anatomical network via diffusion-weighted MRI and Graph Theory[J]. Neuroimage, 2008, 40(3):1064-1076.
[6] Hagmann P, Kurant M, Gigandet X, et al. Mapping human whole-brain structural networks with diffusion MRI.[J]. Plos One, 2007, 2(7):e597.
[7] Friston K J. Functional and effective connectivity in neuroimaging: A synthesis.[J]. Brain Connectivity, 1994, 1(1):13-36.
[8] Stephan K E, Hilgetag C C, Burns G A, et al. Computational analysis of functional connectivity between areas of primate cerebral cortex[J]. Philosophical Transactions of the Royal Society of London, 2000, 355(1393):111.
[9] 冀鑫如. 基于静息状态fMRI的ADHD复杂脑网络分析[D].北京交通大学,2015.
[10] Shubhangi D C, Hiremath P S. Support vector machine (SVM) classifier for brain tumor detection[C]// International Conference on Advances in Computing, Communication and Control. ACM, 2009:444-448.
[11] Min H B, Wu T, Pan R. Mix-ratio sampling: Classifying multiclass imbalanced mouse brain images using support vector machine[J]. Expert Systems with Applications, 2010, 37(7):4955-4965.
[12] Zhou H, Xu Q, Wang Y, et al. An Improved Support Vector Machine Classifier for EEG-Based Motor Imagery Classification[M]// Advances in Neural Networks – ISNN 2009. Springer Berlin Heidelberg, 2009:267-275.
[13] Zanin M, Sousa P, Papo D, et al. Optimizing Functional Network Representation of Multivariate Time Series[J]. Scientific Reports, 2012, 2(630):630.
[14] He Y, Chen Z J, Evans A C. Small-world anatomical networks in the human brain revealed by cortical thickness from MRI.[J]. Cerebral Cortex, 2007, 17(10):2407-2419.
[15] Jie B, Jiang X, Zu C, et al. The New Graph Kernels on Connectivity Networks for Identification of MCI[M]// Machine Learning and Interpretation in Neuroimaging. Springer International Publishing, 2014.
[16] 王立鹏, 费飞, 接标,等. 基于子图选择和图核降维的脑网络分类方法[J]. 计算机科学与探索，2014, 8(10):1246-1253.
[17] Zhou Y, Jr D J, Hubner K F, et al. Abnormal connectivity in the posterior cingulate and hippocampus in early Alzheimer's disease and mild cognitive impairment.[J]. Alzheimers & Dementia the Journal of the Alzheimers Association, 2008, 4(4):265-270.