



ORIGINAL RESEARCH PAPER

Computer Science

RESEARCH ON BRAIN NETWORK CLASSIFICATION METHOD BASED ON INTEGRATED MODEL

KEY WORDS: brain network, subnets, integrated model, MCI

Zhang-sensen

School of computer science and technology, TianGong University , No. 399, Binshui West Road, Xiqing District, Tianjin

ABSTRACT

mild cognitive impairment (MCI) is a condition between healthy elderly people and alzheimer's disease (AD). At present, brain network analysis based on machine learning methods can help diagnose MCI. In this paper, the brain network is divided into several subnets based on the shortest path, and the feature vectors of each subnet are extracted and classified. In order to make full use of subnet information, this paper adopts integrated classification model for classification. Each base classification model can predict the classification of a subnet, and the classification results of all subnets are calculated as the classification results of brain network. In order to verify the effectiveness of this method, a brain network of 66 people was constructed and a comparative experiment was carried out. The experimental results show that the classification accuracy of the integrated classification model proposed in this paper is 19% higher than that of SVM, which effectively improves the classification accuracy.

INTRODUCTION

Mild cognitive disorder is a state between healthy elderly and Alzheimer's disease. Diagnosis and treatment of mild cognitive disorder can effectively reduce the conversion rate of Alzheimer's disease. Analysis of brain network can help diagnose mild cognitive disorder^[1]. With the development of machine learning, machine learning classification algorithm is applied to the analysis of brain network. Using classification algorithm to classify brain network has become a new way to diagnose mild cognitive impairment.

At present, many researchers at home and abroad use machine learning method to analyze the brain network. In terms of measuring the similarity of brain networks, standards and others constructed a new sub network core to measure the similarity between brain networks, and applied it to the classification of brain diseases^[2]. The sequential pattern proposed by Zhang Daoqiang can not only make full use of the weight information of the edge, but also explicitly simulate the ordinal relationship of the weighted edges in the brain connection network. The sequential pattern proposed by Zhang Daoqiang can not only make full use of the weight information of the edge, but also explicitly simulate the ordinal relationship of the weighted edges in the brain connection network^[3]; Liu Yu et al. Constructed the whole brain stratification for each experimental object Network, the brain is divided into 90, 54, 14 and 1 region. According to Pearson's correlation coefficient, connectivity between each pair of regions is calculated, and it is used as a classification feature. After dimensionality reduction, multi-core boost algorithm is used for classification^[4].

RESEARCH ON CLASSIFICATION METHOD

• Division and characteristic calculation of subnets

Given a brain network data $g(V, e)$, V_0 is any node in the brain network. The shortest path length from this node to all other nodes is calculated with this node as the starting point. The node with the shortest path length less than K between V_0 is regarded as the node set of the subnet, and the edge related to these nodes in the original network is regarded as the edge set of the subnet. In order to use the traditional classification model to classify the brain network, this paper can extract the characteristics of the subnet, including the degree of nodes, the number of nodes and clustering coefficient as the classification feature vector.

• Integrated classification model

In this paper, an integrated classification model for subnet is constructed based on decision tree. For a given brain network data set $\{G_1, G_2, \dots, G_m\}$. First of all, G_1 is divided into 116 subnets, and M group of subnets are obtained. Each group of subnets contains 116 subnets.

From each set of subnets, the same 40 subnets are randomly selected to get another m -set of subnets, each of which contains 40 subnets.

$$\{ \{G_{k_1}^1, G_{k_2}^1, \dots, G_{k_{40}}^1\}, \dots, \{G_{k_1}^m, G_{k_2}^m, \dots, G_{k_{40}}^m\} \}$$

For the above $m \times 40$ subnet, the feature calculation and feature selection are carried out to get the n -vector data set. For any K_i , we can get the sample set composed of M eigenvectors

$$X_i = \{V_{k_1}^1, V_{k_1}^2, \dots, V_{k_1}^m\}$$

The i th decision tree is constructed through the sample set X_i . Through the above process, 40 decision trees can be constructed and combined into an integrated classification model.

After the integrated classification model is built, the model is tested through the test set. The classification accuracy rate is recorded as R , the classification accuracy rate of each decision tree is recorded as R_i , four decision trees with the lowest prediction rate are selected as the elimination objects, and four subnets are selected from the rest subnets. Through the above process, the new decision tree and the original decision tree are combined into a new integrated classification model, and the test set is classified and predicted, and the classification accuracy rate is recorded as t . If $R < T$, replace the original model with the new one; if $r > t$, do not replace. Repeat the above process until the classification accuracy of the integrated model tends to be stable or meets the classification requirements.

VERIFICATION TEST

• Experiment flow

The proposed classification method firstly divides the brain network into subnets and extracts the feature attributes (feature path length, degree of nodes, clustering coefficient or graph core) of the subnets, then uses the feature selection method to reduce the dimension, and finally uses the proposed integrated model to classify.

• Data acquisition

The experimental data in this chapter are also from the ADNI database. In the experiment, we downloaded 66 brain image data from the ADNI database. Among them, 36 functional brain images of healthy elderly and 30 functional brain images of MCI patients were included. The specific information of brain image data is shown in table 1.

Table 1 Data Details

	Number	Age	Male / female
CN	36	70+5	18/18
LMCI	30	70+7	17/13

• The construction of brain network

Firstly, the brain image data are preprocessed by time correction, head motion correction, registration, segmentation, spatial standardization and smoothing. The process of data preprocessing is completed by SPM software package. Then the nodes of brain network are defined by AAL brain spectrum, and the edges between nodes are defined by Pearson correlation coefficient. The construction of brain network is completed by rest software package. The construction process of brain network is shown in figure 1:



Figure 1 Brain network construction

• Comparative analysis of experimental results

In the experimental process, we draw molecular networks under different thresholds (0.35, 0.4, 0.45, 0.5). In the comparative experiment, we extract the clustering coefficient of subnetworks, and then carry out 50 experiments with the method in this paper, and statistics the classification accuracy of the experiment. In order to extract the clustering coefficient of global brain network under the same threshold, SVM model was used to classify 50 times, and the classification results were counted. The comparison results of accuracy under different thresholds are shown in Table 2.

Table 2 Comparison Of Average Classification Accuracy

threshold	0.35	0.4	0.45	0.5
Clustering coefficient	0.54	0.53	0.62	0.55
Article method	0.81	0.78	0.79	0.69

According to Table 2, when clustering coefficient and SVM are used for classification, the best classification effect can be obtained when the threshold value is 0.45; for the classification method in this paper, when the threshold value is equal to 0.35, the classification effect is the best, and the highest classification accuracy is increased by 19% compared with the clustering coefficient. When the threshold value is equal to 0.45, SVM is used to carry out 50 experiments on the clustering coefficient; when the threshold value is equal to 0.35, the method in this paper is used to carry out 50 experiments on the clustering coefficient, and the accuracy comparison results of 50 experiments are shown in Figure 2.

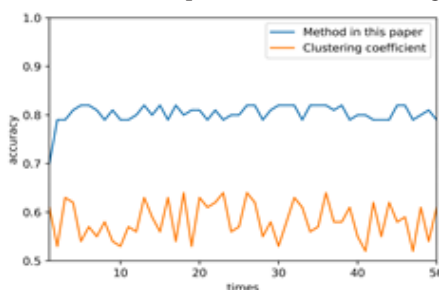


Figure 2 classification accuracy under optimal threshold

According to the analysis in Figure 4-1, under the optimal threshold conditions, the classification accuracy of the proposed classification method is higher than that of using SVM to classify the clustering coefficients; meanwhile, the accuracy variance of the proposed method is small.

SUMMARY

Based on the shortest path, this paper proposes a method to divide the brain network into several subnets. In order to capture the topological structure information of brain

network more concretely, we classify and predict each subnet, and propose an integrated classification model based on the subnet. Experimental results show that the proposed method effectively improves the classification accuracy of clustering coefficient, compared with SVM model, the accuracy of this method is improved by 19%.

REFERENCES:

- [1] Gong Li, Fang min. (2017), "early recognition and screening of cognitive impairment." Shanghai pharmaceutical, 38 (17): 3-7 + 55.
- [2] Jiebiao, Zhang Daoqiang. (2016), "New graph kernel for brain network and its application in MCI classification ." Journal of computer science, 39 (08): 1667-1680.
- [3] Zhang D, Huang J, Jie B, et al. (2018), "Ordinal pattern: a new descriptor for brain connectivity networks." IEEE transactions on medical imaging, 37(7):1711-1722.
- [4] Liu J, Li M, Lan W, et al. (2016), "Classification of alzheimer's disease using whole brain hierarchical network." IEEE/ACM transactions on computational biology and bioinformatics, 15(2):624-632.