ACADEMIC
World Academic Union

ISSN 1746-7659, England, UK
Journal of Information and Computing Science
Vol. 15, No. 2, 2020, *pp*.141-145

ADHD Diagnosis and Recognition Based on Functional Classification

Fei Zheng, Chunzheng Cao[1]
School of Mathematics and Statistics, Nanjing University of Information Science & Technology, Nanjing 210044, China
(Received June 20, 2020, accepted July 24, 2020)

Abstract. This research starts from the lack of reliable and effective disease identification biomarkers for attention deficit hyperactivity disorder (ADHD). Based on the functional classification methods, including functional generalized linear model (FGLM), functional linear discriminant analysis (FLDA) method and functional principal component analysis (FPCA), we establish models of corpus callosum (CC) shape and give some analyses. The purpose is to verify whether the corpus callosum shape data can be used as an effective classification basis for disease discrimination and classification, and to provide a new auxiliary discriminant diagnosis idea for ADHD disease discrimination.

Keywords: ADHD, functional classification, FGLM, FLDA, FPCA.

1. Introduction

Attention deficit hyperactivity disorder (ADHD), commonly known as hyperactivity, is a common mental disorder in childhood, and its pathological causes are based on neurology [1]. It has been reported that the incidence of ADHD in school-age children in China is about 1.5%~12%, and that in foreign countries is 3%~10% [2]. Although the level of hyperactivity in children with ADHD will decrease with age, follow-up studies have found that 30% ~ 80% of children's symptoms will continue into adolescence, and still meet the diagnosis of ADHD. While 50% ~ 65% of symptoms will continue into adulthood [3]. Due to the high incidence of ADHD and its adverse effects, ADHD has become a research hotspot in many fields.

In recent years, technologies such as electroencephalography, magnetic resonance imaging, and functional magnetic resonance imaging have been used in the auxiliary diagnosis of ADHD. At the same time, the rapid development of machine learning has also made it used effectively in ADHD classification and diagnosis. Riaz et al. [4] proposed a machine learning framework based on support vector machines (SVM) for imaging data and non-imaging data to study the functional connection changes between ADHD patients and control groups. Du et al. [5] proposed a discriminative sub-network selection method, based on the kernel principal component analysis method to extract the main features from the discriminant network to classify and recognize ADHD. Sen et al. [6] studied structural MRI image features and fMRI features, and used them to train linear SVM classifiers for ADHD classification. Shao et al. [7] proposed a dual-objective ADHD classification scheme based on the L1-norm SVM model to achieve ADHD classification.

In addition, research on ADHD based on a functional data analysis framework is a new perspective. With the widespread application of big data, data tends to become more complex, quantified, diversified, heterogeneous, etc., and functional data analysis methods emerged and are widely used in many fields. The functional data classification methods can be divided into the following three categories: (1) Functional classification based on regression. The classification labels and functional predictors are connected through a regression model, then use training set to estimate parameters for classification [8, 9]. (2) Functional classification based on probability density. Functional data is firstly projected to a finite feature space, then estimate the probability density of each category with parametric or non-parametric methods. At last, predict and classify new samples based on this probability density [10, 11]. (3) Functional classification based on algorithms. Firstly, perform dimensionality reduction similar to Method 2, and then choose a non-parametric classification tool to do the classification.

Based on the above research status, we can find that attention deficit hyperactivity disorder is a medical problem that is very worth studying. Functional data analysis is a new branch of statistical data

[1] Corresponding author. *E-mail address*: caochunzheng@163.com.

DOI https://doi.org/2024-JICS-22389 | Generated on 2025-04-05 22:12:25
OPEN ACCESS

analysis, and it is also one of the popular research objects of statistics, which can find out the internal contact of the observation. Domestic research on this aspect is still relatively lacking. It is of great significance to use the FDA method to study the classification of ADHD and provide scientific auxiliary diagnosis guidance based on theoretical research.

This article is organized as follows: Section 2 mainly introduces the data source and data preprocessing. In Section 3, we briefly introduce three typical functional data classification methods, and use them to classify the corpus callosum thickness curve. Finally, in Section 4, we analyze the research results and give a summary

2. Data acquisition and processing

2.1 Data source

The MRI data set used in this article comes from the ADHD-200 global competition training data set. Due to the lack of objective biological tools in the clinic, it is impossible to provide individuals with ADHD diagnostic information to guide clinicians to make treatment decisions. In 2011, the 1000 Functional Connectomes Project team shared its database and organized a global ADHD Disease classification and discrimination contest based on MRI images. The purpose is to hold an ADHD medical image diagnosis competition to discover ways to help diagnose ADHD patients based on computer picture recognition, and accelerate the scientific community's understanding of the neural basis of ADHD.

Here we choose a data set with a sample size of 245 in the Peking University sub-data set (Peking U imaging point) (see: http://fcon_1000.projects.nitrc.org/indi/adhd200/). There are a total of 245 subjects in the data set, all of whom are children and adolescents (8-18 years old), including 143 healthy controls, with an average age of 11.42 years; 102 ADHD patients, with an average age of 12.08 years. All patients were evaluated by the scale and met the ADHD diagnostic criteria of DSM-IV. All subjects were right-handed, excluding past or existing psychiatric diseases or a history of mental disorders, learning disabilities or other psychiatric diseases as determined in DSM-IV, also all of them have no history of loss of consciousness due to head trauma. See Table 1 for specific data.

Table 1. Sample information

Diagnosis	Total number	Female	Male	Age range
Typically Developing	143	59	84	8~15（mean:11.42） year-old
ADHD-Combined	38	0	38	8~15（mean:11.56） year-old
ADHD-Inattentive	64	12	52	8~17（mean:12.39） year-old

2.2 Data processing

We perform the following processing on the CC morphological data of each research individual in the ADHD-200 data set. First, use FreeSurfer[12] (http://surfer.nmr.mgh.harvard.edu/) to process each T1-weighted MRI data, including translation correction, non-parametric non-uniform intensity normalization, and affine transformation to MNI305 Atlas, intensity standardization, skull removal, automatic cortical segmentation, etc. Keep quality control on each output image data, and intracranial volume (ICV) information can also be obtained from the output results of the FreeSurfer software package. The comparison chart before and after data preprocessing is shown in Figure 1. Then use the processed image as the input of Yuki package [13] (http://www.nitrc.org/projects/art) for corpus callosum segmentation and thickness extraction (Figure 2 left), and obtain the thickness curves of the two types of samples See Figure 2 (right).

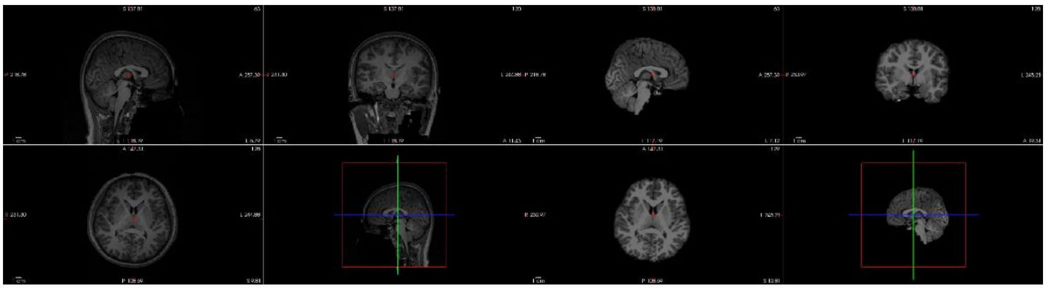

Fig 1. Data pre-processing before (left), after (right) comparison

DOI https://doi.org/2024-JICS-22389 | Generated on 2025-04-05 22:12:25
OPEN ACCESS

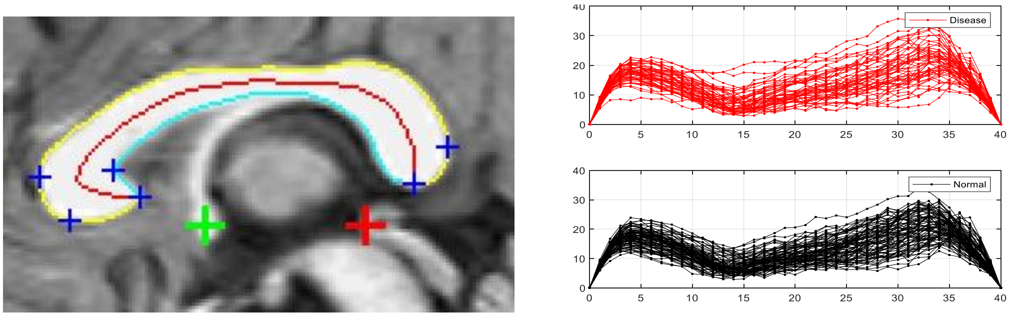

Fig 2. CC segmentation and thickness(left), thickness curves(right)

3. ADHD diagnosis and identification based on functional classification

3.1 ADHD classification and discrimination based on functional generalized linear model

Functional generalized regression models [8, 14, 15], such as functional logistic regression model, are the most popular regression-based functional classification methods. For the corpus callosum data $\{(y_i, x_i); i = 1, \ldots, n\}$, where $y_i \in \{0,1\}$ is the category label corresponding to the i-th thickness curve x_i, which can be expressed by the classification model based on functional logistic regression for

$$\pi = p(y^* = 1|x^*), \\ logit(\pi) = b_0 + \int_I x^*(t)\beta(t)dt, \tag{1}$$

where b_0 and $\beta(t)$ are intercept and coefficient functions, respectively. In model-based Bayesian classifiers, x^* is usually classified based on the maximum posterior probability $\{p(y^* = k|x^*); k = 0,1\}$, and it is easy to generalize to most classification issues.

Using 6 times of 5-fold cross-validation, the training set and test set samples are divided multiple times, and the classification accuracy of the test set samples is as follows

Table 2. Accuracy for GFLM

57.78%	61.36%	62.22%	48.89%	40.91%	53.33%	57.78%	53.33%	47.73%	54.55%
40.00%	47.73%	57.78%	55.56%	56.82%	61.36%	52.27%	53.33%	60.00%	55.56%
48.89%	64.44%	45.45%	55.56%	59.09%	60.00%	57.78%	44.44%	50.00%	56.82%

Calculated from the above table, the accuracy of classification based on FGLM is about 54.03%. In 30 experiments, the best accuracy rate was 64.44%, while the lowest accuracy rate was 40%. At the same time, the accuracy of 1/5 of the experiment is higher than 60%; and the accuracy of about 4/15 is lower than 50%. On the one hand, the above results show that the problem is challenging, on the other hand, it also shows the differences in the typicality of the samples in the sub-data set.

3.2 ADHD classification discrimination based on functional linear discriminant analysis

Different from classification methods based on regression, functional linear discriminant analysis [16] based on classical linear discriminant analysis is also a common functional classification method. The basic idea is to classify new samples based on maximizing conditional probability by using Bayesian criterion. Also suppose that the prior probability is π_k, and $\pi_0 + \pi_1 = 1$. Given the density function f_k of the k-th sample, the conditional probability of the sample to be classified x^* can be obtained by Bayesian formula as follows

$$p(y^* = k|x^*) = \frac{\pi_k f_k(x^*)}{\pi_1 f_1(x^*) + \pi_0 f_0(x^*)}. \tag{2}$$

Classify x^* based on the maximum posterior probability. If it is further assumed that the k-th sample obeys a Gaussian distribution with mean μ_k and covariance matrix Σ, then maximizing the conditional probability is equivalent to

$$\arg\max_k (L_k)\,, \tag{3}$$

where L_k is discriminant function

$$L_k = x^T \Sigma^{-1} \mu_k - \frac{\mu_k^T \Sigma^{-1} \mu_k}{2} + \log \pi_k. \tag{4}$$

Similar to Section 3.1, the calculated accuracy is about 56.95%, which is slightly higher than the result of FGLM. The reason may be that when the FLDA method is used for classification and

DOI https://doi.org/2024-JICS-22389 | Generated on 2025-04-05 22:12:25
OPEN ACCESS

discrimination, the data is projected to reduce the dimension, so that the difference between the two types of samples is enlarged and the difference within the group is reduced. This prompts us to further analyze the mean difference and variance fluctuations of the two types of samples (Figure 3,4).

Fig 3 and Fig 4 shows the mean line and the variance line of the two types of samples, in which red represents ADHD patients, and black represents normal samples. It can be seen from Figure 3 that the very similar means of the two types of samples directly lead to the low classification accuracy, and the fluctuation of the variance also brings new challenges to the classification.

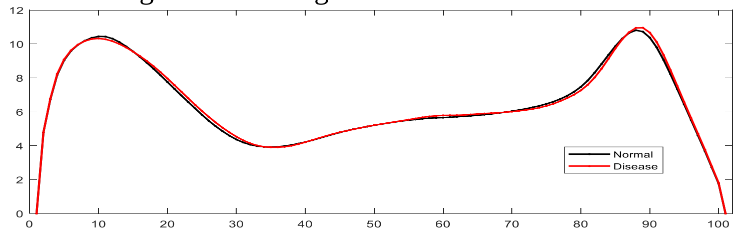

Fig 3. Mean curves for CC thickness

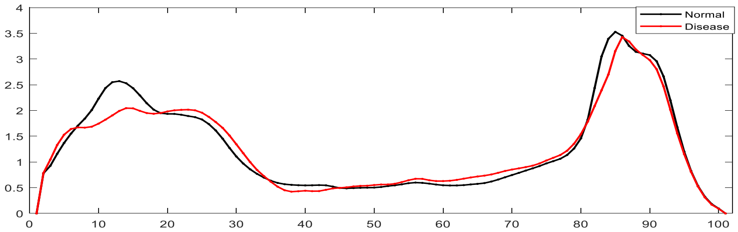

Fig 4. Variance curves for CC thickness

3.3 ADHD classification and discrimination based on functional principal component analysis

Functional principal component analysis (FPCA) [17] is another classification method based on the idea of dimensionality reduction. Suppose that the random process $X(t)$ has expectation $E\{X(t)\} = \mu(t)$ and the covariance function $cov\{X(t), X(s)\} = G(t,s)$. Expand $G(t,s)$ orthogonally, then any thickness curve has the following Karhunen-Loève expansion

$$X(t) = \mu(t) + \sum_m \varepsilon_m \phi_m(t), \quad 0 \le t \le T, \tag{5}$$

where ϕ_m is the characteristic function corresponding to the characteristic value λ_m,

$$\varepsilon_m = \int_0^T (X(t) - \mu(t))\phi_m(t)dt,$$

are uncorrelated random variables, and $E(\varepsilon_m)$=0, $E(\varepsilon_m^2) = \lambda_m$,$\sum \lambda_m < \infty$. ε_m is also called the principal component score of functional principal components FPCs）ϕ_m. Then, based on the empirical covariance to calculate the m-th eigenvector $\widehat{\boldsymbol{\phi}}_m$ and the corresponding eigenvalue λ_m, the score of the i-th thickness curve on the m-th functional principal component can be obtained as follows

$$\hat{\boldsymbol{\varepsilon}}_{\boldsymbol{im}} = \sum_{k=1}^{S} \big(X_i(s_k) - \hat{\mu}(s_k)\big)\widehat{\boldsymbol{\phi}}_m(s_k). \tag{6}$$

Thus, the FPC score can be used to predict a single sample

$$\widehat{X_i}(t) = \hat{\mu}(t) + \sum_{m=1}^{M} \hat{\boldsymbol{\varepsilon}}_{\boldsymbol{im}} \widehat{\boldsymbol{\phi}}_m(t), \quad 0 \le t \le T. \tag{7}$$

Similarly, the accuracy rate is about 54.84%, and the highest accuracy rate is 68.63%.

Table 3. Accuracy for FPCA

49.02%	54.90%	54.90%	54.90%	52.94%	60.78%	58.82%	47.06%	58.82%	54.90%
52.94%	60.78%	54.90%	49.02%	52.94%	47.06%	50.98%	68.63%	60.78%	50.98%
50.98%	60.78%	60.78%	49.02%	49.02%	60.78%	49.02%	58.82%	50.98%	58.82%

4. Conclusion and discussion

In this article, we use the shape of corpus callosum as a biomarker to diagnose ADHD. We use MRI images to compare the shape and structure of the CC between children with ADHD and normal children under the framework of functional data analysis (FDA), with the help of functional generalized linear models, functional linear discriminant analysis and functional principal component analysis. Then establish classification models for the diagnosis of ADHD based on these. We hope they can realize scientific

DOI https://doi.org/2024-JICS-22389 | Generated on 2025-04-05 22:12:25

OPEN ACCESS

diagnosis and provide auxiliary diagnosis methods to reduce misdiagnosis caused by excessively high subjective diagnosis standards.

It can be seen from the results table in Section 3 that the classification accuracy rate is not high. We analyze the reasons as follows:

(1) Loss of information. In this problem, the original data is a three-dimensional image, and this article only uses its two-dimensional information on the median sagittal plane, which leads to the loss of part of the information. In addition, new errors will be introduced in the process of transforming shape data into thickness curves, leading to new information loss;

(2) Weak recognizability. Because the difference between ADHD patients and healthy individuals is not completely reflected in the shape of the corpus callosum, the classification based on the shape (thickness curve) is not recognizable, and the personal characteristics such as age, gender, and IQ should also be considered when classifying;

(3) Local differences are weakened. It can be seen from Fig 3 that the difference between the two types of samples is only reflected in the local position, but the traditional method will make the local difference weakened by the global similarity. Therefore, in the classification model, it is necessary to accurately identify and amplify the local difference to improve the performance of the classifier;

(4) Limitations of traditional functional classification methods. Based on the above analysis, the traditional classification method needs to be improved. A good idea is to introduce a loss function including both mean and variance to fully explore and use local difference.

References

[1] C. Huang, S. Martin, H. Zhu. Clustering high-dimensional landmark-based two-dimensional shape data. Journal of the American Statistical Association, 2015, 110(511): 946-961.

[2] T. G. Dias, S. P. Iyer, S. D. Carpenter, et al. Characterizing heterogeneity in children with and without ADHD based on reward system connectivity. Dev Cogn Neurosci, 2015, 11: 155-174.

[3] Y. Li. A study on cognitive characteristics and psychosocial development characteristics of children with attention deficit hyperactivity disorder. Shandong University, 2007.

[4] A. Riaz, M. Asad, E. Alonso, and G. Slabaugh. Fusion of fMRI and nonimaging data for ADHD classification, Computerized Med. Imag. Graph., 2018(65):115-128.

[5] J. Du, L. Wang, B. Jie, and D. Zhang. Network-based classification of ADHD patients using discriminative subnetwork selection and graph kernel PCA. Comput. Med. Imag. Graph., 2016(52): 82-88.

[6] B. Sen, N. C. Borle, R. Greiner, and M. R. G. Brown. A general prediction model for the detection of ADHD and autism using structural and functional MRI. PLoS One, 2018, 13 (4): e0194856.

[7] L. Shao, Y. Xu, and D. Fu. Classification of ADHD with bi-objective optimization. Biomed. Informat., 2018 , (84): 164-170.

[8] H. Müller, U. Stadtmüller. Generalized functional linear models. The Annals of Statistics, 2005, 33(2): 774-805.

[9] J. Goldsmith, J. Bobb, C. M. Crainiceanu, et al. Penalized functional regression. Journal of Computational and Graphical Statistics, 2011, 20(4): 830-851.

[10] P. Hall, P. B. Presnell. A functional data-analytic approach to signal discrimination. Technometrics, 2001, 43(1): 1-9.

[11] F. Ferraty, P. Vieu. Curves discrimination: a nonparametric functional approach. Computational Statistics & Data Analysis, 2003, 44(1-2): 161-173.

[12] A. M. Dale, B. Fischl, M. I. Sereno. Cortical surface-based analysis: I. segmentation and surface reconstruction. NeuroImage, 1999(9): 179-194.

[13] Ardekani, B. A, Yuki Module of the Automatic Registration Toolbox (ART) for Corpus Callosum Segmentation, http://www.nitrc.org/projects/art.

[14] G. M. James. Generalized linear models with functional predictors. Journal of the Royal Statistical Society: Series B., 2002, 64: 411-432.

[15] H. Müller. Functional modelling and classication of longitudinal data. Scandinavian Journal of Statistics. 2005, 32: 223-240.

[16] G. M. James, T. J. Hastie. Functional linear discriminant analysis for irregularly sampled curves. Journal of the Royal Statistical Society: Series B., 2001, 63: 533-550.

[17] F. Yao, H. Müller, J. L. Wang. Functional data analysis for sparse longitudinal data. Journal of the American Statistical Association, 2005, 100, 577-590.

DOI https://doi.org/2024-JICS-22389 | Generated on 2025-04-05 22:12:25

OPEN ACCESS