Pattern Recognition Letters 138 (2020) 82–87

Contents lists available at ScienceDirect

Pattern Recognition Letters

journal homepage: www.elsevier.com/locate/patrec

Improved ASD classification using dynamic functional connectivity and multi-task feature selection

Jin Liu[a], Yu Sheng[a], Wei Lan[b], Rui Guo[a], Yufei Wang[a], Jianxin Wang[a,*]

[a] Hunan Provincial Key Lab on Bioinformatics, School of Computer Science and Engineering, Central South University, Changsha, China
[b] School of Computer, Electronics and Information, Guangxi University, Nanning, China

ARTICLE INFO

Article history:
Received 1 February 2020
Revised 30 May 2020
Accepted 7 July 2020
Available online 8 July 2020

MSC:
41A05
41A10
65D05
65D17

Keywords:
ASD Classification
Resting state functional MRI
Dynamic functional connectivity
Multi-task feature selection
Multi-kernel learning

ABSTRACT

Accurate diagnosis of autism spectrum disorder (ASD), which is a neurodevelopmental disorder and often accompanied by abnormal social skills, communication skills, interests and behavior patterns, has always been a challenging task in clinical practice. Recent studies have shown great potential for using fMRI data to distinguish ASD from typical control (TC). However, it has always been a challenging problem to extract which features from fMRI data and how to combine these different types of features to achieve improved ASD/TC classification performance. To address this problem, in this study we propose an improved ASD/TC classification framework based on dynamic functional connectivity (DFC) and multi-task feature selection. Our proposed ASD/TC classification framework is evaluated on 871 subjects with fMRI data from the Autism Brain Imaging Data Exchange I (ABIDE I) via a 10-fold cross validation strategy. Experimental results show that our proposed method achieves an accuracy of 76.8% and an area under the receiver operating characteristic curve (AUC) of 0.81 for ASD/TC classification. In addition, compared with some existing state-of-the-art methods, our proposed method achieves better accuracy and AUC for ASD/TC classification. Overall, our proposed ASD/TC classification framework is effective and promising for automatic diagnosis of ASD in clinical practice.

1. Introduction

Autism spectrum disorder (ASD) is a neurodevelopmental disorder and often accompanied by abnormal social skills, communication skills, interests and behavior patterns [16,21,38]. The patients with ASD are basically children from 2 to 9 years of age [30]. A recent survey from the Autism and Developmental Disabilities Monitoring (ADDM) Network in Centers for Disease Control and Prevention (CDC) show that about 1 in 59 children worldwide has ASD [4]. As a result of these symptoms, the patients with ASD have caused great economic burden for their families and society [2]. In the United States, the total annual costs for children with ASD is estimated to be between 11.5 – 60.9 billion [6]. However, since the etiology of ASD is unknown and there is no medical test (such as blood and urine tests) for the diagnosis of ASD, so far it has been a very difficult problem to accurately diagnose whether a person has ASD. Therefore, in order to treat the patients with ASD more effectively, it is urgent to develop an effective method for accurate diagnosis of ASD.

Nowadays, magnetic resonance imaging (MRI) [23] has provided a way for clinicians to examine the structural and functional changes associated with the development of diseases *in vivo* [12,13,24], and has been widely applied in clinical practice. At present, commonly used MRI mainly includes structural MRI (sMRI) and functional MRI (fMRI). Compared with sMRI, since fMRI [15] can measure the changes in hemodynamics caused by neuron activity at a series of time points for the whole brain, it has been widely applied in the research of brain dysfunction diseases. For example, in order to explore the feature representation related to ASD, [1] proposed a machine learning method based on fMRI data to perform ASD classification[36]; proposed a multi-view feature representation based on fMRI data to perform Schizophrenia classification and find the imaging markers related to Schizophrenia. Furthermore, commonly used fMRI mainly includes resting-state fMRI (rs-fMRI) and task-state fMRI (ts-fMRI). Compared with ts-fMRI, rs-fMRI attracts more attention because it is easier to acquire data. In this study we focus on rs-fMRI data.

With the rapid development of machine learning methods [5], recent studies have shown great potential for integrating fMRI data and machine learning methods to automatically diagnose brain diseases, such as Schizophrenia [36,37], Alzheimer's disease [7,20,26] and ASD [1,19,22,32]. For example, [32] used rs-fMRI data

* Corresponding author.
E-mail addresses: liujin06@csu.edu.cn (J. Liu), jxwang@mail.csu.edu.cn (J. Wang).

https://doi.org/10.1016/j.patrec.2020.07.005
0167-8655/

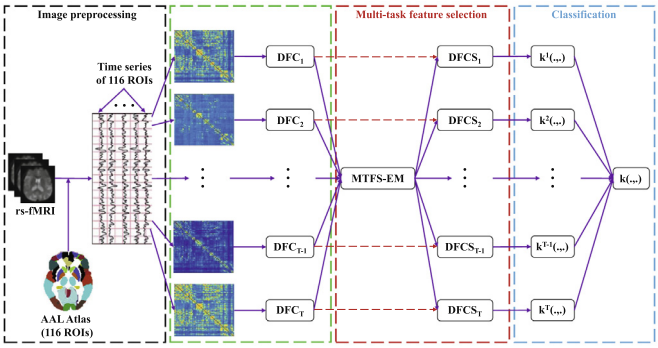

Fig. 1. Schematic diagram of our proposed ASD/TC classification framework, which includes four main steps: 1) image preprocessing, 2) dynamic functional connectivity, 3) multi-task feature selection, and 4) classification.

as the inputs of graph convolutional network (GCN) to perform ASD classification, and obtained 70.4% accuracy. Recently, following the work of [19,32] proposed an InceptionGCN method based on fMRI data to perform ASD classification task. However, although some results have been achieved by integrating fMRI data and machine learning methods for ASD classification, it has always been a challenging problem to extract which features from rs-fMRI data and how to combine these different types of features to achieve improved ASD/TC classification performance.

Based on the above analysis, to improve the performance of ASD/TC classification, in this study we propose a new ASD/TC classification framework based on dynamic functional connectivity (DFC) and multi-task feature selection as shown in Fig. 1. As can be seen from Fig. 1, our proposed ASD/TC classification framework mainly includes four steps. Firstly, we extract the time series of each brain region based on automated anatomical labeling (AAL) [35] atlas from rs-fMRI data of each subject. Secondly, we further extract the DFC by computing Pearson correlation coefficient (PCC) between paired brain regions via successive and non-overlapping time windows as feature representation of each subject. Thirdly, in order to obtain features that are more helpful from each feature set for ASD/TC classification, an improved multi-task feature selection method integrating elastic net and manifold regularization is proposed and denoted as MTFS-EM. Finally, a multi-kernel SVM learning method is adopted to perform the ASD/TC classification task by combining multiple selected feature sets. Our proposed ASD/TC classification framework is evaluated on 871 subjects (including 403 subjects with ASD and 468 TCs) with rs-fMRI data from the Autism Brain Imaging Data Exchange I (ABIDE I) via a 10-fold cross validation strategy.

2. Materials and method

2.1. Subjects and image preprocessing

The subjects involved in this study are provided by the Autism Brain Imaging Data Exchange I (ABIDE I[1]) [11]. The ABIDE I includes 539 subjects with ASD and 573 typical controls (TCs) from 17 international sites. For fair comparison, in this study we choose the same 871 subjects including 403 subjects with ASD and 468 TCs from the ABIDE I as some existing literature [1,19,32] as shown in Table 1. In Table 1, m ± std and M/F are short for mean ± standard deviation and male/female, respectively. It is worth mentioning that these 871 subjects were selected by three human experts [1].

Table 1
The demographic information of the subjects involved in this study.

	ASD (403)	TC (468)
Age (m ± std)	17.07 ± 7.95	16.84 ± 7.23
Gender (M/F)	349/54	378/90

Brain rs-fMRI data of each subject was scanned from the ABIDE I. Specifically, the rs-fMRI data of each subject had been preprocessed by different tools[2]. For fair comparison with some existing literature [1,19,32], in this study we download and use the functional preprocessed data based on automated anatomical labeling (AAL) atlas [35] with the Configurable Pipeline for the Analysis of Connectomes (CPAC) [9,10] as the original feature representation, which is composed of time series of ROIs based on AAL atlas as shown in Fig. 1. The functional preprocessing procedure of the CPAC includes slice time correction, motion correct, skull-strip, nuisance signal regression, and so on. It is worth mentioning that since the scanning parameters of different sites for rs-fMRI data are different, the number of time points of rs-fMRI data at different sites is also different, ranging from 78 to 316.

2.2. Dynamic functional connectivity

Traditionally, the feature representation of brain functional connectivity is based on the assumption that the functional connectivity between brain regions remains unchanged during the whole scanning process [1,19,32,36]. However, it is well known that brain functional connectivity change dynamically over time. In recent years, brain dynamic functional connectivity (DFC) has been widely concerned [31,33].

In this study, we extract the DFC using successive and non-overlapping time windows from original feature representation to further represent each subject as shown in Fig. 1. The extraction procedure of the DFC of each subject mainly includes two steps. Firstly, we equally segment all time series of each ROI of AAL atlas into T non-overlapping time windows. Then, we construct a DFC matrix (denoted as F_t) by computing Pearson correlation coefficient (PCC) between time series from a pair of ROIs of AAL atlas within the $t-$th time window as follows:

$$F_t(i,j) = \frac{\mathrm{cov}\left(s_i^t, s_j^t\right)}{\sigma_{s_i^t}\sigma_{s_j^t}} \tag{1}$$

where $F_t(i, j)$ denotes the connectivity weight between two ROIs i and j within the $t-$th time window, $\mathrm{cov}(\cdot,\cdot)$ denotes the covariance between two vectors, s_i^t and s_j^t denote the time series of ROIs i and j within the $t-$th time window, respectively, and $\sigma_{s_i t}$ and $\sigma_{s_j t}$ denote the standard deviation of vectors s_i^t and s_j^t, respectively. Thus, according to Eq. (1), we can obtain a set of DFC with T time windows, i.e., $F = \{F_1, F_2, \ldots, F_{T-1}, F_T\}$.

After performing the DFC procedure, we can obtain T feature sets for each subject, which are denoted as $\{DFC_1, DFC_2, \ldots, DFC_{T-1}, DFC_T\}$.

2.3. Multi-task feature selection

To obtain more discriminative features from the above T feature sets of each subject for ASD/TC classification, an improved multi-task feature selection method integrating elastic net and manifold regularization (denoted as MTFS-EM) is proposed as shown in Fig. 1. Next, we briefly introduce the MTFS-EM method as follows.

[1] http://fcon_1000.projects.nitrc.org/indi/abide/abide_I.html.

[2] http://preprocessed-connectomes-project.org/abide/.

Since elastic net [40] integrates the advantages of ridge regression [17] and Lasso regression [34], it has been widely applied in feature selection. The objective function of elastic net can be written as follows:

$$\min_{w}\left(\|y-Xw\|_2^2+\alpha\|w\|_1+\gamma\|w\|_2^2\right) \tag{2}$$

where $X=[x_1,x_2,\ldots,x_i,\ldots,x_N]^{'}\in\Re^{N\times M}$ represents all training subjects, x_i represents the i-th training subject, N is the number of training subjects, M is the number of features of each training subject, $y=[y_1,y_2,\ldots,y_i,\ldots,y_N]^{'}\in\Re^{N}$ represents the labels of all training subjects, $w\in\Re^{M}$ is the discriminant vector whose value indicates the weight of each feature, α and γ are the regularization parameters to balance the objective function, and $\|\cdot\|_1$ and $\|\cdot\|_2$ represent the ℓ_1-norm and ℓ_2-norm, respectively. However, the standard elastic net model can only be used in a single task, and cannot be directly applied to multi-task learning. To address this problem, Zou *et al.* [39] proposed a multi-task learning model based on elastic net (denoted as MTEN). Suppose there are T learning tasks, the objective function of MTEN can be written as follows:

$$\min_{W}\left(\sum_{t=1}^{T}\left\|y^t-X^tw^t\right\|_F^2+\alpha\|W\|_{2,1}+\gamma\|W\|_F^2\right) \tag{3}$$

where $W\in\Re^{M\times T}$ is the weight matrix for T tasks, and $\|\cdot\|_{2,1}$ and $\|\cdot\|_F$ represent the $\ell_{2,1}-$norm and $Frobenius(F)-$norm, respectively.

Although the above MTEN method can be used in multi-task feature selection, this method only considers the relationship between data and category labels, and ignores the mutual dependence among data. To address this problem, a manifold regularization [18] is added into the above MTEN method as follows:

$$\min_{W}\left(\sum_{t=1}^{T}\left\|y^t-X^tw^t\right\|_F^2+\alpha\|W\|_{2,1}+\gamma\|W\|_F^2+\beta\sum_{t=1}^{T}\left(X^tw^t\right)^{'}L^t\left(X^tw^t\right)\right) \tag{4}$$

where L^t is the Laplacian matrix on the $t-$th task, which is defined as follows:

$$L^t=D^t-S^t \tag{5}$$

where D^t is the diagonal matrix on the $t-$th task, which is defined as follows:

$$D_{ii}^t=\sum_{j=1}^{N}s_{ij}^t \tag{6}$$

where s_{ij}^t is the element of matrix S^t, which is a similarity matrix across different subjects x_i^t and x_j^t on the $t-$th task, and is defined as follows:

$$s_{ij}^t=\begin{cases}1 & if \text{ the labels of subjects } x_i^t \text{ and } x_j^t \text{ are the same;}\\ 0 & \text{otherwise.}\end{cases} \tag{7}$$

After performing the MTFS-EM procedure, we can obtain T selected feature sets for each subject, which are denoted as $\{DFCS_1,DFCS_2,\ldots,DFCS_{T-1},DFCS_T\}$.

2.4. Classification

Following our previous works [25,28,29], a multi-kernel SVM learning method is proposed for ASD/TC classification. The multi-kernel SVM learning method mainly contains two steps as shown in Fig. 1. Firstly, for each selected feature set of training subjects, an optimal Gaussian kernel is calculated based on the features selected by the above proposed multi-task feature selection method. Then, to combine multiple selected feature sets, a multi-kernel learning method is applied as follows:

$$k\left(x_i,x_j\right)=\sum_{t=1}^{T}\eta^t k^t\left(x_i^t,x_j^t\right) \tag{8}$$

where η^t is a non-negative weight parameter, and $k^t(x_i^t,x_j^t)$ represents the kernel function over the $t-$th selected feature set across subjects x_i^t and x_j^t as follows:

$$k^t\left(x_i^t,x_j^t\right)=\exp\left(-\frac{\left\|x_i^t-x_j^t\right\|^2}{2\sigma^2}\right) \tag{9}$$

where σ is also a non-negative weight parameter and represents the width of the Gaussian kernel.

Obviously, once η^t and σ are obtained, a standard SVM classifier can be performed for ASD/TC classification.

3. Experiments and results

3.1. Experimental settings

Our proposed ASD/TC classification framework is performed via a 10-fold cross-validation strategy [27] and repeated 50 times. Before performing our proposed framework on ASD/TC classification, all parameters in the proposed framework are set. In the dynamic functional connectivity procedure, T is set to 4. In the multi-task feature selection procedure, α, γ and β are set to a range from 0 to 1 at a step size of 0.1, respectively, i.e., $\alpha=\{0,0.1,\cdots,0.9,1\}$, $\gamma=\{0,0.1,\cdots,0.9,1\}$, and $\beta=\{0,0.1,\cdots,0.9,1\}$. Specifically, in this study only features with non-zero elements in W are selected for subsequent classification in each cross validation, and the optimal α, γ and β are determined based on the training subjects through a grid search optimization algorithm. In the classification procedure, we first create a set of 10 base Gaussian kernels with 10 different widths for each selected feature set, i.e., $\sigma=\left\{2^{-4},2^{-3},\cdots,2^4,2^5\right\}$, and then determine the optimal σ via a simple iteration strategy. Furthermore, we set a range from 0 to 1 at a step size of 0.1 for η^t, i.e., $\eta^t=\{0,0.1,\cdots,0.9,1\}, t=\{1,2,\cdots,T-1,T\}$. It's worth mentioning that the optimal η^t is also determined based on the training subjects through a grid search optimization algorithm. The SVM classifier used in the classification procedure is implemented by the LIBSVM library [8], and $C=1$.

In order to quantify the performance of our proposed ASD/TC classification framework, accuracy (ACC), sensitivity (SEN) and specificity (SPE) are calculated as follows:

$$ACC=\frac{TP+TN}{TP+TN+FP+FN}SEN=\frac{TP}{TP+FN}SPE=\frac{TN}{TN+FP} \tag{12}$$

where TP, FP, TN and FN represent the number of true positive subjects, false positive subjects, true negative subjects and false negative subjects, respectively. In addition, to further evaluate the overall performance of our proposed ASD/TC classification framework, the area under the receiver operating characteristic curve (AUC) [14] is also reported. It is worth mentioning that the higher the values of the four metrics, the better the performance of our proposed ASD/TC classification framework.

3.2. Experimental results

As can be seen from Fig. 1, in this study the classification procedure mainly includes two steps. Firstly, we find an optimal kernel matrix (i.e., $k^t(.,.)$) for each selected feature set. The specific

Table 2
Classification performance of our proposed framework for ASD/TC classification.

Kernel matrices (σ)	ACC(%)	SEN(%)	SPE(%)	AUC
$k^1(.,.)(2^1)$	68.6 ± 3.1	64.8 ± 2.5	71.5 ± 3.7	0.72 ± 2.2
$k^2(.,.)(2^{-3})$	69.4 ± 1.9	63.7 ± 3.7	73.9 ± 2.9	0.74 ± 3.5
$k^3(.,.)(2^{-2})$	65.4 ± 3.8	58.4 ± 4.3	72.4 ± 4.5	0.69 ± 4.6
$k^4(.,.)(2^4)$	67.3 ± 4.2	60.5 ± 2.1	72.7 ± 3.3	0.71 ± 3.8
$k(.,.)(-)$	***76.8*** ± 2.7	***72.5*** ± 3.2	***79.9*** ± 3.8	***0.81*** ± 3.1

method is to use SVM classifiers with different width (i.e., σ) Gaussian kernel matrices to perform ASD/TC classification. The Gaussian kernel matrix corresponding to the best classification accuracy is the optimal kernel matrix. Then, we combine T optimal kernel matrices via the above proposed multi-kernel learning method to generate a mixed kernel matrix. An SVM classifier with the mixed kernel matrix (i.e., $k(.,.)$) is used to perform the final ASD/TC classification task. Table 2 shows the classification performance obtained by SVM classifiers with the above 4 optimal kernel matrices and 1 mixed kernel matrix for ASD/TC classification. It is worth mentioning that the three parameters in the feature selection process corresponding to the results in Table 2 are $\alpha = 0.3$, $\gamma = 0.7$, and $\beta = 0.5$.

From Table 2, we can observe that our proposed framework achieve 76.8% ACC, 72.5% SEN, 79.9% SPE and 0.81 AUC for ASD/TC classification. The results show that our proposed ASD/TC classification framework is effective. In addition, we also observe that the performance (including ACC, SEN, SPE and AUC) of SVM classifier based on mixed kernel matrix (i.e., $k(.,.)$) is better than that of SVM classifiers based on other four kernel matrices (i.e., $k^1(.,.)$, $k^2(.,.)$, $k^3(.,.)$ and $k^4(.,.)$). This comparative result shows that our proposed multi-kernel SVM learning method can effectively integrate different selected feature sets. It is worth noting that the AUC value of our proposed ASD/TC classification framework is greater than 0.8, which shows that our proposed method has good robustness for ASD/TC classification.

It is worth mentioning that in our experiments, we find that the functional connectivity features selected after each feature selection includes functional connectivity features connected to left thalamus, right precuneus and left insula. This phenomenon indicates that these three ROIs play an important role in ASD/TC classification.

4. Discussion

4.1. Impact of different number of time windows

In this section, we discuss the impact of different number of time windows (i.e., T) in our proposed framework for the performance of ASD/TC classification. In order to investigate the impact of different T in our proposed framework, we have done a series of experiments based on $T = \{1, 2, \cdots, 9, 10\}$ for ASD/TC classification. Fig. 2 shows the accuracies of different T in our proposed framework for ASD/TC classification. It's worth mentioning that when $T = 1$, the functional connectivity between brain regions is calculated based on the time series of the whole scanning process.

As can be seen from Fig. 2, when $T = 4$, our proposed framework obtain the best ACC ASD/TC classification. This result shows that setting T to 4 in this study can select features that are more useful for ASD/TC classification. In addition, we also observe that when T is greater than 1, the ACC obtained by our proposed method is obviously higher than that when $T = 1$ for ASD/TC classification. This comparative result shows that using DFC as feature representation can obtain more useful features than traditional functional connectivity calculated by the whole scanning process for ASD/TC classification.

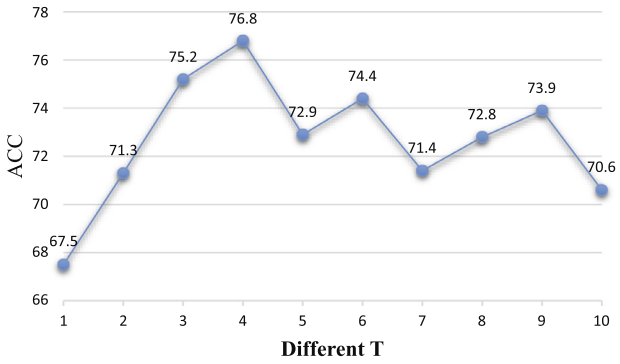

Fig. 2. Accuracies of different T in our proposed framework for ASD/TC classification.

Table 3
Comparison with other multi-task feature selection methods for ASD/TC classification.

Methods	ACC (%)	$p-$value
MTFS-L21	71.1 ± 3.4	< 0.01
M2TFS	73.4 ± 4.8	< 0.01
MTEN	74.9 ± 3.9	< 0.02
MTFS-EM	***76.8*** ± 2.7	

4.2. Comparison with other multi-task feature selection methods

In this section, we discuss the superiority of our proposed multi-task feature selection method (i.e., MTFS-EM) for ASD/TC classification. In order to investigate the superiority of MTFS-EM for ASD/TC classification, we compare our results with the three existing multi-task feature selection methods including MTFS-L21 [3], M2TFS [18] and MTEN [39]. For fair comparison, the experimental subjects used by the above three existing multi-task feature selection methods are consistent with those used in this study for each cross-validation. Table 3 shows the accuracies of different multi-task feature selection methods in our proposed framework for ASD/TC classification. In addition, we also use the two sample t-test method to evaluate the difference between our proposed multi-task feature selection method and any other multi-task feature selection method based on ACC as shown in Table 3.

As can be seen from Table 3, MTFS-EM in our proposed framework obtain the best ACC. This comparative result shows that MTFS-EM in our proposed framework can obtain more useful features than other three multi-task feature selection methods for ASD/TC classification. Furthermore, we also observe that all $p-$values are less than 0.05. The difference analysis results further show that our proposed multi-task feature selection method has significant advantages. Focusing on these four multi-task feature selection objective functions, these comparative results also show that adding reasonable regularization can obtain more reasonable fitting results.

4.3. Comparison with other existing methods

In this section, we discuss the superiority of our proposed framework for ASD/TC classification. To demonstrate the superiority of our proposed ASD/TC classification framework, we compare our results with the three recently reported methods on ASD classification, based on traditional machine learning [1] or deep learning [19,32]. Since the above three existing methods use inconsistent evaluation metrics, and their common evaluation metric only has an ACC, for fair comparison, only the ACC reported in the above three existing methods is compared. It's worth mentioning that the

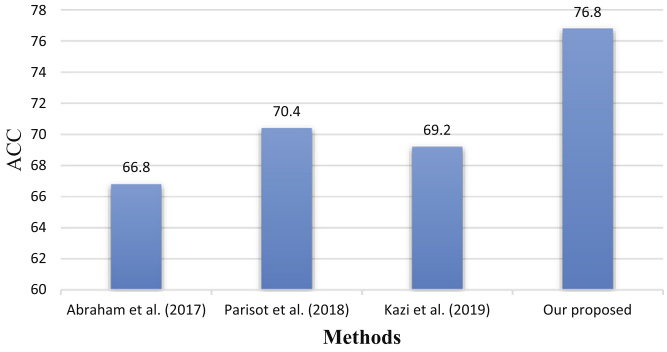

Fig. 3. Comparison with three recently reported methods for ASD/TC classification.

experimental subjects used by the three recently reported methods are consistent with those used in this study. Fig. 3 shows the comparative result for ASD/TC classification.

As can be seen from Fig. 3, the ACC of our proposed ASD/TC classification framework is the best. This comparative result further demonstrates that our proposed method is effective and has certain advantages for ASD/TC classification.

5. Conclusion

In this study, we propose an improved ASD/TC classification framework based on dynamic functional connectivity and multi-task feature selection. Experimental results demonstrate that our proposed framework is effective for ASD/TC classification. This method paves the way to discriminative imaging markers for computer-aided diagnosis of ASD.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work is supported in part by the National Natural Science Foundation of China under Grant No.61802442, No.61877059, the Natural Science Foundation of Hunan Province under Grant No.2019JJ50775, No.2018JJ2534, the 111 Project (No. B18059), and the Hunan Provincial Science and Technology Program (2018WK4001).

References

[1] A. Abraham, M.P. Milham, A. Di Martino, R.C. Craddock, D. Samaras, B. Thirion, G. Varoquaux, Deriving reproducible biomarkers from multi-site resting-state data: an autism-based example, Neuroimage 147 (2017) 736–745.
[2] D. Amendah, S.D. Grosse, G. Peacock, D.S. Mandell, The economic costs of autism: a review, Autism Spect. Disord. (2011) 1347–1360.
[3] A. Argyriou, T. Evgeniou, M. Pontil, Multi-task feature learning, in: Advances in neural information processing systems, 2007, pp. 41–48.
[4] J. Baio, L. Wiggins, D.L. Christensen, M.J. Maenner, J. Daniels, Z. Warren, M. Kurzius-Spencer, W. Zahorodny, C.R. Rosenberg, T. White, et al., Prevalence of autism spectrum disorder among children aged 8 yearsutism and developmental disabilities monitoring network, 11 sites, united states, 2014, MMWR Surveill. Summ. 67 (6) (2018) 1.
[5] C.M. Bishop, Pattern Recognition and Machine Learning, springer, 2006.
[6] A.V. Buescher, Z. Cidav, M. Knapp, D.S. Mandell, Costs of autism spectrum disorders in the united kingdom and the united states, JAMA Pediatr 168 (8) (2014) 721–728.
[7] E. Challis, P. Hurley, L. Serra, M. Bozzali, S. Oliver, M. Cercignani, Gaussian process classification of alzheimer's disease and mild cognitive impairment from resting-state fmri, Neuroimage 112 (2015) 232–243.
[8] C.-C. Chang, C.-J. Lin, Libsvm: a library for support vector machines, ACM Trans. Intell. Syst. Technol. (TIST) 2 (3) (2011) 1–27.
[9] C. Craddock, Y. Benhajali, C. Chu, F. Chouinard, A. Evans, A. Jakab, B.S. Khundrakpam, J.D. Lewis, Q. Li, M. Milham, et al., The neuro bureau preprocessing initiative: open sharing of preprocessed neuroimaging data and derivatives, Neuroinformatics 41 (2013).
[10] C. Craddock, S. Sikka, B. Cheung, R. Khanuja, S.S. Ghosh, C. Yan, Q. Li, D. Lurie, J. Vogelstein, R. Burns, et al., Towards automated analysis of connectomes: the configurable pipeline for the analysis of connectomes (c-pac), Front Neuroinform 42 (2013).
[11] A. Di Martino, C.-G. Yan, Q. Li, E. Denio, F.X. Castellanos, K. Alaerts, J.S. Anderson, M. Assaf, S.Y. Bookheimer, M. Dapretto, et al., The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism, Mol. Psychiatry 19 (6) (2014) 659–667.
[12] L. Du, K. Liu, X. Yao, S.L. Risacher, J. Han, A.J. Saykin, L. Guo, L. Shen, Detecting genetic associations with brain imaging phenotypes in alzheimer disease via a novel structured scca approach, Med. Image Anal. 61 (2020) 101656.
[13] L. Du, K. Liu, L. Zhu, X. Yao, S.L. Risacher, L. Guo, A.J. Saykin, L. Shen, A.D.N. Initiative, Identifying progressive imaging genetic patterns via multi-task sparse canonical correlation analysis: a longitudinal study of the adni cohort, Bioinformatics 35 (14) (2019) i474–i483.
[14] T. Fawcett, An introduction to roc analysis, Pattern Recognit. Lett. 27 (8) (2006) 861–874.
[15] K.J. Friston, P. Jezzard, R. Turner, Analysis of functional mri time-series, Hum. Brain Mapp. 1 (2) (1994) 153–171.
[16] U. Frith, F. Happé, Autism spectrum disorder, Curr. Biol. 15 (19) (2005) R786–R790.
[17] A.E. Hoerl, R.W. Kennard, Ridge regression: biased estimation for nonorthogonal problems, Technometrics 12 (1) (1970) 55–67.
[18] B. Jie, D. Zhang, B. Cheng, D. Shen, A.D.N. Initiative, Manifold regularized multitask feature learning for multimodality disease classification, Hum. Brain Mapp. 36 (2) (2015) 489–507.
[19] A. Kazi, S. Shekarforoush, S.A. Krishna, H. Burwinkel, G. Vivar, K. Kortüm, S.-A. Ahmadi, S. Albarqouni, N. Navab, Inceptiongcn: receptive field aware graph convolutional network for disease prediction, in: International Conference on Information Processing in Medical Imaging, Springer, 2019, pp. 73–85.
[20] A. Khazaee, A. Ebrahimzadeh, A. Babajani-Feremi, Application of advanced machine learning methods on resting-state fmri network for identification of mild cognitive impairment and alzheimer disease, Brain Imag. Behav. 10 (3) (2016) 799–817.
[21] Y. Kong, J. Gao, Y. Xu, Y. Pan, J. Wang, J. Liu, Classification of autism spectrum disorder by combining brain connectivity and deep neural network classifier, Neurocomputing 324 (2019) 63–68.
[22] S.I. Ktena, S. Parisot, E. Ferrante, M. Rajchl, M. Lee, B. Glocker, D. Rueckert, Metric learning with spectral graph convolutions on brain connectivity networks, Neuroimage 169 (2018) 431–442.
[23] Z.-P. Liang, P.C. Lauterbur, Principles of Magnetic Resonance Imaging: A Signal Processing Perspective, SPIE Optical Engineering Press, 2000.
[24] J. Liu, M. Li, W. Lan, F.-X. Wu, Y. Pan, J. Wang, Classification of alzheimer's disease using whole brain hierarchical network, IEEE/ACM Trans. Comput. Biol. Bioinf. 15 (2) (2018) 624–632.
[25] J. Liu, M. Li, Y. Pan, F.-X. Wu, X. Chen, J. Wang, Classification of schizophrenia based on individual hierarchical brain networks constructed from structural mri images, IEEE Trans. Nanobiosci. 16 (7) (2017) 600–608.
[26] J. Liu, Y. Pan, F.-X. Wu, J. Wang, Enhancing the feature representation of multi-modal mri data by combining multi-view information for mci classification, Neurocomputing 400 (2020) 322–332.
[27] J. Liu, J. Wang, H. Bin, F.-X. Wu, Y. Pan, Alzheimer disease classification based on individual hierarchical networks constructed with 3d texture features, IEEE Trans Nanobiosci. 16 (6) (2017) 428–437.
[28] J. Liu, J. Wang, Z. Tang, B. Hu, F.-X. Wu, Y. Pan, Improving alzheimer's disease classification by combining multiple measures, IEEE/ACM Trans. Comput. Biol. Bioinf. 15 (5) (2018) 1649–1659.
[29] J. Liu, X. Wang, X. Zhang, Y. Pan, X. Wang, J. Wang, Mmm: classification of schizophrenia using multi-modality multi-atlas feature representation and multi-kernel learning, Multimed. Tool. Appl. 77 (22) (2018) 29651–29667.
[30] C. Lord, S. Risi, P.S. DiLavore, C. Shulman, A. Thurm, A. Pickles, Autism from 2 to 9 years of age, Arch. Gen. Psychiatry 63 (6) (2006) 694–701.
[31] L.E. Mash, A.C. Linke, L.A. Olson, I. Fishman, T.T. Liu, R.-A. Müller, Transient states of network connectivity are atypical in autism: a dynamic functional connectivity study, Hum. Brain Mapp. 40 (8) (2019) 2377–2389.
[32] S. Parisot, S.I. Ktena, E. Ferrante, M. Lee, R. Guerrero, B. Glocker, D. Rueckert, Disease prediction using graph convolutional networks: application to autism spectrum disorder and alzheimer disease, Med. Image Anal. 48 (2018) 117–130.
[33] M.G. Preti, T.A. Bolton, D. Van De Ville, The dynamic functional connectome: state-of-the-art and perspectives, Neuroimage 160 (2017) 41–54.
[34] R. Tibshirani, Regression shrinkage and selection via the lasso, J. Roy. Stat. Soc. 58 (1) (1996) 267–288.
[35] N. Tzourio-Mazoyer, B. Landeau, D. Papathanassiou, F. Crivello, O. Etard, N. Delcroix, B. Mazoyer, M. Joliot, Automated anatomical labeling of activations in spm using a macroscopic anatomical parcellation of the mni mri single-subject brain, Neuroimage 15 (1) (2002) 273–289.
[36] Y. Xiang, J. Wang, G. Tan, F.-X. Wu, J. Liu, Schizophrenia identification using multi-view graph measures of functional brain networks, Front. Bioeng. Biotechnol. 7 (2020) 479.

[37] H. Yang, J. Liu, J. Sui, G. Pearlson, V.D. Calhoun, A hybrid machine learning method for fusing fmri and genetic data: combining both improves classification of schizophrenia, Front Hum Neurosci 4 (2010) 192.
[38] N. Zecavati, S.J. Spence, Neurometabolic disorders and dysfunction in autism spectrum disorders, Curr. Neurol. Neurosci. Rep. 9 (2) (2009) 129–136.
[39] B. Zou, V. Lampos, I. Cox, Multi-task learning improves disease models from web search, in: Proceedings of the 2018 World Wide Web Conference, 2018, pp. 87–96.
[40] H. Zou, T. Hastie, Regularization and variable selection via the elastic net, J. Roy. Stat. Soc.) 67 (2) (2005) 301–320.