

Received 27 October 2019; revised 4 January 2020 and 8 March 2020; accepted 28 March 2020.
Date of publication 3 April 2020; date of current version 24 April 2020.

Digital Object Identifier 10.1109/JTEHM.2020.2985022

The Identification of Alzheimer's Disease Using Functional Connectivity Between Activity Voxels in Resting-State fMRI Data

YUHU SHI^{id}, WEIMING ZENG^{id}, (Senior Member, IEEE), JIN DENG^{id}, WEIFANG NIE,
AND YIFEI ZHANG

Information Engineering College, Shanghai Maritime University, Shanghai 201306, China

CORRESPONDING AUTHOR: Y. SHI (syhustb2011@163.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61906117, Grant 31170952, and Grant 61701298 and in part by the Shanghai Sailing Program under Grant 19YF1419000.

ABSTRACT Background: Alzheimer's disease (AD) is a common neurodegenerative disease occurring in the elderly population. The effective and accurate classification of AD symptoms by using functional magnetic resonance imaging (fMRI) has a great significance for the clinical diagnosis and prediction of AD patients. Methods: Therefore, this paper proposes a new method for identifying AD patients from healthy subjects by using functional connectivities (FCs) between the activity voxels in the brain based on fMRI data analysis. Firstly, independent component analysis is used to detect the activity voxels in the fMRI signals of AD patients and healthy subjects; Secondly, the FCs between the common activity voxels of the two groups are calculated, and then the FCs with significant differences are further identified by statistical analysis between them; Finally, the classification of AD patients from healthy subjects is realized by using FCs with significant differences as the feature samples in support vector machine. Results: The results show that the proposed identification method can obtain higher classification accuracy, and the FCs between activity voxels within prefrontal lobe as well as those between prefrontal and parietal lobes play an important role in the prediction of AD patients. Furthermore, we also find that more brain regions and much more voxels in some regions are activity in AD group compared with health control group. Conclusion: It has a great potential value for the AD pathogenesis mechanism study.

INDEX TERMS Activity voxels, fMRI, functional connectivity, independent component analysis, support vector machine.

I. INTRODUCTION

Alzheimer's disease (AD) is a progressive neurodegenerative disease with the characters of memory loss, poor judgment, language deterioration and so on, which is one of the highest rates of disease among the elderly [1]. According to the report released by the AD International (ADI) in 2018, the current cost of the disease is about a trillion US dollars per year, and the number of people worldwide suffer from AD will increase from the current 50 million to 152 million by 2050 [2]. Therefore, the prediction and diagnosis of AD are very important at its early warning, which has become a crucial step to retard or even avoid dementia.

Many neuroimaging techniques have been used to unveil neuroanatomical and neurophysiological substrates of AD so

far [3], [4], such as positron emission tomography (PET) [5], diffusion tensor imaging (DTI) [6], electroencephalogram (EEG) [7] and functional magnetic resonance imaging (fMRI) [8]. Thereinto, fMRI is one of the most promising non-invasive techniques for early AD diagnosis [9], [10]. By using fMRI technique, some previous AD studies have demonstrated that the disrupted functional connectivities (FCs) mainly concentrated on several key brain regions of neuronal degeneration in AD, such as the anterior hippocampus [11], inferior parietal lobe [12], thalamus [13], posterior cingulate cortex [14], and prefrontal cortex [15]. These findings suggested that the FCs with significant differences in AD patients compared with healthy subjects can be served as a biomarker for the diagnosis of AD.

In addition, some other studies have tried to study AD using the measures in graph theory, for example, Binnewijzend *et al.* used voxel-wise eigenvector centrality to analyze individual whole-brain resting-state fMRI data and identified changes in brain network organization in AD patients that are related to cognition and underlying AD pathology by a permutation-based method [16]. Many methods can be used to study FCs such as elastic net regularized regression [17] and independent component analysis (ICA) [18]. Among them, ICA is a very popular blind source separation technique that has been fruitfully applied in the AD research, for example, Zhong *et al.* investigated the differences of effective connectivity of the default mode network (DMN) in AD patients and normal controls based on the identification of DMN using ICA [19].

As a simple and effective means of research, classification of healthy and unhealthy individuals for a given disease is a common practice [20]–[24]. Without exception, classification is particularly challenging in developing biomarkers for studies of AD [25]–[27]. Currently, the literatures regarding AD on automatic classification systems based on fMRI are depth and breadth [28], [29]. For example, in order to assist in the diagnosis of AD and supports the monitoring of the progression of the disease, Tripoliti *et al.* proposed a six stage method based on the features extracted from the fMRI data to classify AD patients [30]. In addition, Dai *et al.* proposed a methodological framework called multi-modal imaging and multi-level characteristics with multi-classifier (M3), to discriminate AD patients from healthy controls [31].

Many methods can be used for the classification of AD. Among them, machine learning (ML) and pattern classification techniques have played an important role in exploring the brain differences between AD patients and healthy controls [32]. As the state-of-the-art ML techniques, support vector machine (SVM) and deep learning (DL) has been widely used in brain imaging classification of AD [33], [34]. For example, Khzaee *et al.* combined graph theoretical approaches with SVM to classify AD patients from healthy subjects by studying functional brain network alteration [35]. Sarraf *et al.* outlined state-of-the-art DL-based pipelines employed to distinguish AD's MRI and fMRI data from normal healthy control data for the same age group, where the fMRI data was applied in DL for the first time to achieve medical image analysis and AD prediction [36]. Recently, Armananzas *et al.* proposed the direct use of fMRI activation voxels of brain to tackle the automatic pattern analysis of AD and healthy individuals by applying different ML techniques to the classification of fMRI data for this purpose [37]. However, as far as we know, there is no literature to study this question from the view of functional connectivity of activity voxels.

Based on this consideration, in this paper, we propose a novel identification method based on the FCs between activity voxels in the whole brain, which is used to distinguish the AD patients from the health control (HC) groups. First, the classical ICA method is used to detect the activity voxels of the AD patients and HC subjects in the resting-state fMRI

data. Second, the FCs between these activity voxels in the whole-brain are constructed, and then FCs with significant difference between AD and HC groups are identified by T-test with false discovery rate (FDR) correction, which are used as the input samples for the next classification. Finally, the machine learning techniques of SVM is adopted to classify the AD patients and HC subjects. The experimental results show that the proposed classification method in this paper can effectively distinguish AD patients and HC subjects, and can obtain higher classification accuracy.

The remainder of this paper is organized as follows: the proposed classification method will first be presented, followed by the experimental test which is designed to evaluate the performance of the classification method. Finally, the results and analysis will be presented together with interpretations and conclusions related to the advantages and limitations of this new classification method.

II. MATERIALS AND METHOD

A. OBTAINING OF COMMON ACTIVATED VOXELS

Assuming that there are a total of K subjects in the group, and all subjects have T time points and V voxels after normalization. In order to obtain the common activity voxels of subjects in the group. Firstly, we implement ICA on the single-subject level. For subject i , ICA is defined as:

$$X_i = M_i S_i, \quad (i = 1, 2, \dots, K) \quad (1)$$

where X_i represents the observed fMRI data, $S_i = (s_{i1}, s_{i2}, \dots, s_{iN_i})^T$ is an $N_i \times V$ matrix, and each row represents an independent component (IC). Each column of $M_i = (m_{i1}, m_{i2}, \dots, m_{iN_i})$ which is a $T \times N_i$ mixing matrix represents the corresponding time courses.

Next, we calculate the location set of activity voxels in the ICs of each subject at a given threshold θ , and denote it as AVL_i for subject i . Then AVL_i ($i = 1, 2, \dots, K$) can be obtained as follows:

$$AVL_i = \bigcup_{j=1}^{N_i} \left\{ v \mid \begin{array}{l} abs(s_{ij}(v)) \geq \theta, \\ v = 1, 2, \dots, V; j = 1, 2, \dots, N_i \end{array} \right\} \quad (2)$$

Finally, we can calculate the location set of common activity voxels in all subjects from AVL_i ($i = 1, 2, \dots, K$) and denote it as $CAVL$, which can be obtained according to the following formula:

$$CAVL = \bigcap_{i=1}^K AVL_i \quad (3)$$

where $CAVL$ is a $1 \times N$ row vector and N denotes the number of common activated voxels.

B. FEATURES EXTRACTION FOR CLASSIFICATION

According to the method proposed in section "A. Obtaining of common activity voxels", we firstly can obtain the location sets of common activity voxels of subjects in HC and AD groups, and denote them as $CAVL_1$ and $CAVL_2$ which include N_1 and N_2 voxels, respectively. Then the location set

of common activity voxels of the two groups can be obtained, and denotes it as $CAVL_{12}$ which include N_{12} voxels.

$$CAVL_{12} = CAVL_1 \cap CAVL_2 \quad (4)$$

Next, we calculate the FC between each pair of voxels in $CAVL_{12}$ for each subject in each group, and denote the set of FCs as FC_i for subject i which is a $1 \times N_{12}(N_{12} - 1)/2$ row vector. In this study, FC is measured by the correlation coefficient, so that FC_i can be calculated as follows:

$$FC_i = \left\{ \begin{array}{l} corr(X_i(:, CAVL_{12}(p)), X_i(:, CAVL_{12}(q))) \\ p, q = 1, \dots, N_{12}, p < q \end{array} \right\} \quad (5)$$

where $corr(\cdot)$ represents the correlation operator, X_i is a $T \times V$ matrix represents the observed fMRI data for subject i .

Assuming that HC group and AD group include K_1 and K_2 subjects, respectively, and denotes the FCs of all subjects of HC group and AD group as $FC1$ and $FC2$.

Once $FC1$ and $FC2$ are obtained, the two-sample T-test is implemented on each FC to obtain the FCs with significant difference between AD group and HC group, and denotes the index set of FCs with significant difference as $DFCI$ which can be obtained as follows:

$$DFCI = \left\{ k \mid \begin{array}{l} P_{ttest}(FC1(:, k), FC2(:, k)) < P_0, \\ k = 1, 2, \dots, N_{12}(N_{12} - 1)/2 \end{array} \right\} \quad (6)$$

where $P_{ttest}(\cdot)$ represents the P-value operator after FDR correction in T-test, P_0 represents the given level of significance which is 0.05 in our study. $FC1(:, k)$ and $FC2(:, k)$ represents the k th column of $FC1$ and $FC2$, respectively.

Finally, let $DFC1$ and $DFC2$ denotes the FCs with significant differences in HC group and AD group, which can be obtained according to the index set $DFCI$ of FCs with significant difference between HC group and AD group. Once $DFC1$ and $DFC2$ are obtained, they are used as the input samples into the process of SVM for the classification of HC subjects and AD patients. The detailed schematic diagram of the implementation steps of the proposed method was showed in Fig. 1.

III. EXPERIMENTS AND DATA PROCESSING

A. EXPERIMENTAL DATA

The resting-state fMRI datasets of sixty-seven patients with AD from 71 to 87 (mean = 77.30 ± 8.53) and seventy-six age-matched HC subjects from 71 to 87 (mean = 77.71 ± 5.71) from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database ([www.adni-info.org](http://adni-info.org)) were used to evaluate the classification performance of the proposed method in this study. Data for this study were selected based on availability of resting-state fMRI datasets for age-matched HC subjects and patients with AD.

According to ADNI protocol (<http://adni.loni.usc.edu/>), the resting-state fMRI data was acquired on 3.0 Tesla Philips scanners while subjects were instructed to keep their eyes open. The images dataset was acquired using single-shot SENSE gradient an echo planar imaging (EPI) with 48 slices,

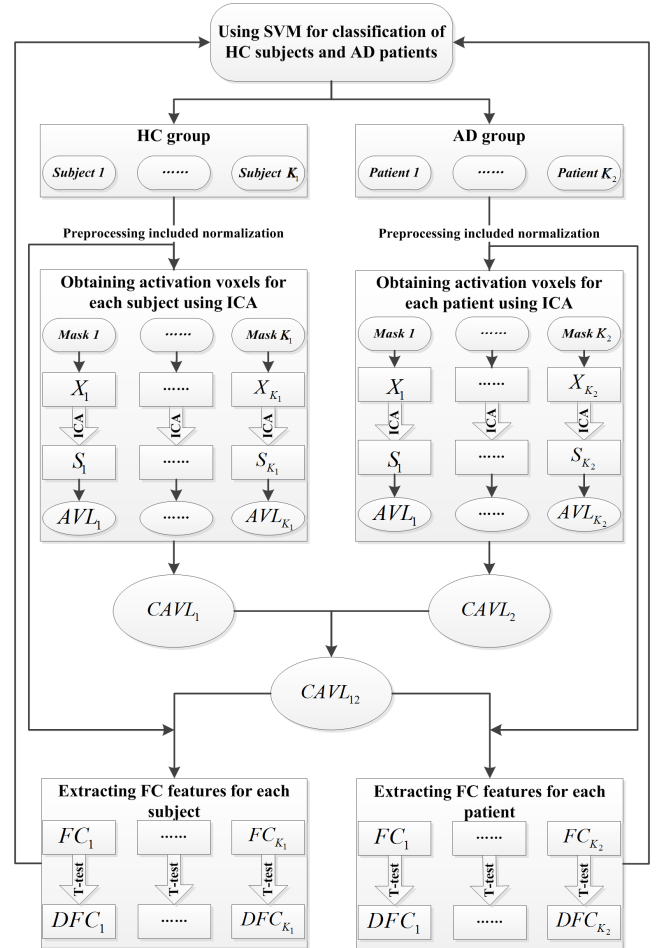


FIGURE 1. The chart flow of the proposed method.

X_i ($i = 1, 2, \dots, K_j; j = 1, 2$) and S_j ($i = 1, 2, \dots, K_j; j = 1, 2$) represent the observed fMRI data within the brain mask and the ICs obtained using ICA for each subject respectively, and AVL_i ($i = 1, 2, \dots, K_j; j = 1, 2$) represents the location set of activity voxels for each subject. $CAVL_1$ and $CAVL_2$ represent the location set of common activity voxels for HC group and AD group respectively, and $CAVL_{12}$ represents the location set of common activity voxels for HC and AD groups.

FC_i ($i = 1, 2, \dots, K_j; j = 1, 2$) represents the FCs between the common activity voxels for each subject. DFC_i ($i = 1, 2, \dots, K_j; j = 1, 2$) represents the FCs with significant differences between HC and AD groups for each subject.

providing whole-brain coverage and 140 volumes, a repetition time (TR) of 3s, an echo time (TE) of 30ms, a flip-angle (FA) of 80, and a scan resolution of 64×64 . The in-plane resolution was $3.31 \text{ mm} \times 3.31 \text{ mm}$, and the slice thickness was 3.31mm.

B. DATA PREPROCESSING

In this study, the preprocessing steps as well as all computations were running on the Matlab platform (Matlab, 2013a, Math-Works Inc., Sherborn, MA, USA). The fMRI data were preprocessed using the DPARSF software (<http://rfmri.org/DPARSF>), including removing the first 10 slices, slice timing, motion correction, spatial normalization

TABLE 1. The Statistical distribution of activity voxels of AD group and HC group in the AAL brain regions and their corresponding MNI coordinates. The “voxel number” represents the number of activity voxels in the brain region, And The “voxel ratio” represents the percentage of obtained activity voxels over the total voxels in the corresponding brain region.

AAL Regions	MNI coordinates (x, y, z)	AD group		HC group	
		voxel number	voxel ratio	voxel number	voxel ratio
Precentral_R	(41.37, -8.21, 52.09)	15	0.0150	-	-
Frontal_Sup_L	(-18.45, 34.81, 42.2)	109	0.1013	41	0.0381
Frontal_Sup_R	(21.9, 31.12, 43.82)	114	0.0984	34	0.0293
Frontal_Mid_L	(-33.43, 32.73, 35.46)	127	0.0877	75	0.0518
Frontal_Mid_R	(37.59, 33.06, 34.04)	126	0.0834	54	0.0358
Frontal_Mid_Orb_L	(-30.65, 50.43, -9.62)	-	-	17	0.0630
Frontal_Inf_Tri_L	(-45.58, 29.91, 13.99)	14	0.0193	-	-
Frontal_Inf_Orb_L	(-35.98, 30.71, -12.11)	12	0.0239	-	-
Frontal_Inf_Orb_R	(41.22, 32.23, -11.91)	17	0.0337	-	-
Supp_Motor_Area_L	(-5.32, 4.85, 61.38)	31	0.0473	-	-
Supp_Motor_Area_R	(8.62, 0.17, 61.85)	23	0.0345	-	-
Frontal_Sup_Medial_L	(-4.8, 49.17, 30.89)	91	0.1076	15	0.0177
Frontal_Sup_Medial_R	(9.1, 50.84, 30.22)	85	0.1326	13	0.0203
Frontal_Mid_Orb_L	(-5.17, 54.06, -7.4)	-	-	14	0.0622
Rectus_L	(-5.08, 37.07, -18.14)	11	0.0421	10	0.0383
ParaHippocampal_L	(-7.63, -25.36, 70.07)	24	0.0839	-	-
ParaHippocampal_R	(25.38, -15.15, -20.47)	27	0.0854	-	-
Cuneus_L	(-5.93, -80.13, 27.22)	14	0.0312	23	0.0512
Lingual_R	(16.29, -66.93, -3.87)	19	0.0280	-	-
Postcentral_R	(41.43, -25.49, 52.55)	22	0.0193	-	-
Parietal_Sup_L	(-23.45, -59.56, 58.96)	-	-	18	0.0285
Parietal_Sup_R	(26.11, -59.18, 62.06)	52	0.0804	20	0.0309
Parietal_Inf_R	(46.46, -46.29, 49.54)	25	0.0597	-	-
Precuneus_L	(-42.46, -22.63, 48.92)	32	0.0297	39	0.0361
Precuneus_R	(9.98, -56.05, 43.77)	30	0.0321	26	0.0278
Temporal_Sup_R	(58.15, -21.78, 6.8)	14	0.0145	-	-
Temporal_Pole_Sup_L	(-39.88, 15.14, -20.18)	37	0.0969	40	0.1047
Temporal_Pole_Sup_R	(48.25, 14.75, -16.86)	70	0.1750	23	0.0575

and smoothing with the Gaussian kernel set to 4 mm, band-pass filtering (0.01-0.08Hz) and detrending. Meanwhile, the whole-brain signal was removed by a multiple linear regression analysis to reduce the effect of the physiological artifacts, and the head motion as well as the cerebrospinal fluid and white matter signals was removed as nuisance covariates to reduce the effects of motion and non-neuronal blood oxygen level-dependent fluctuations. Spatial ICA was implemented using FastICA algorithm [38] in all experiments. Particularly, ICASSO [39] with 20 runs of ICA and both ‘RandInit’ and ‘Bootstrap’ was used to obtain reliable ICs, and MDL [40] was used to estimate the number of ICs. Furthermore, the location and display of these networks were assessed using the MRIcro software (<http://www.mricro.com>).

IV. RESULTS AND ANALYSIS

In this study, we mainly consider the activity voxels in the gray matter region of brain. These activity voxels are obtained with threshold $|z| \geq 2$ after z-scored the ICs of spatial ICA. **Table 1** shows the statistical distribution of activity voxels of AD and HC groups in the automated anatomical labeling (AAL) brain regions of cerebral cortex and their corresponding Montreal Neurological Institute (MNI) coordinates, which are considered as activity. Each brain region contains more than 10 voxels and accounts for more than 1% of the total voxels in the region.

It can be seen from **Table 1** that there are no activity voxels in the brain regions of Frontal_Mid_Orb_L (ORBmid.L/ ORBsupmed.L) and Parietal_Sup_L in AD

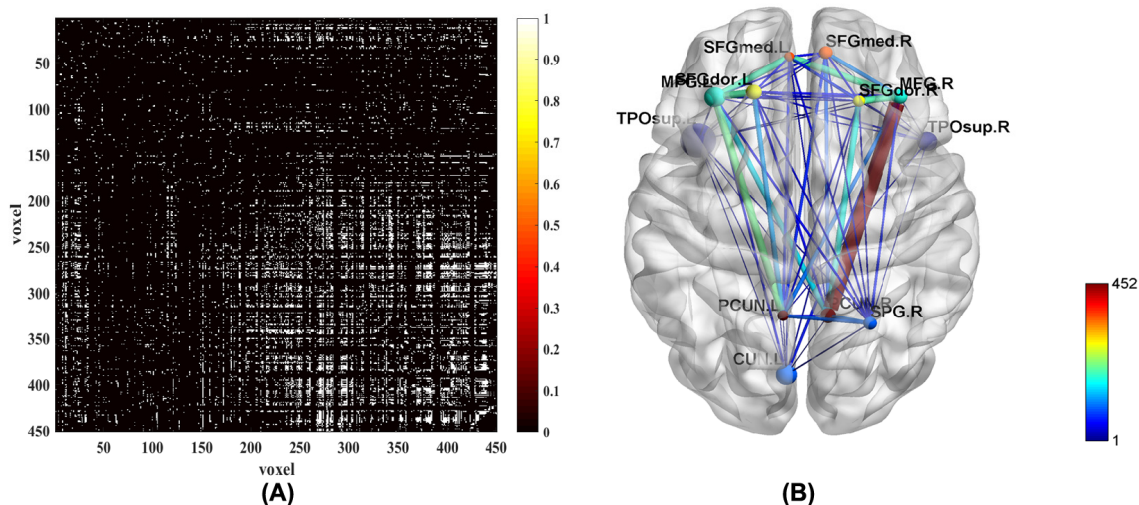


FIGURE 2. (A) the results of t-test on the FCs between the common activity voxels of AD group and HC group; (B) the spatial brain network between AAL regions corresponding to the activity voxels of FCs with significant differences.

group compared with HC group. But there are also many brain regions that are individually activated in AD group such as Precentral_R, Frontal_Inf_Tri_L, Frontal_Inf_Orb (L/R), Supp_Motor_Area (L/R), ParaHippocampal (L/R), Lingual_R, Postcentral_R, Parietal_Inf_R and Temporal_Sup_R. At the same time, it can be found that the number of activity voxels in AD group is significantly increased in the brain regions of Frontal_Sup (L/R), Frontal_Mid (L/R), Frontal_Sup_Medial (L/R), Temporal_Pole_Sup_R compared with those in HC group. In addition, the common activity voxels of AD and HC groups are mainly located in the brain regions of Frontal_Sup (L/R), Frontal_Mid (L/R), Frontal_Sup_Medial (L/R), Cuneus_L, Parietal_Sup_R, Precuneus (L/R) and Temporal_Pole_Sup (L/R).

Fig.2 (A) shows statistical test results on the FCs between the common activity voxels of AD and HC groups, which are obtained by t-test with a confidence level of 95% after FDR correction. **Fig.2 (B)** shows the spatial brain network between AAL regions corresponding to the activity voxels of FCs with significant differences, where the size of edge between different AAL regions represents the number of FCs with significant difference between the common activity voxels. The voxels number and ratio of these AAL regions and the corresponding MNI coordinates are shown in **Table 2**, which are more than 10 voxels and accounts for more than 1% of the total voxels.

It can be seen clearly from **Fig.2 (B)** that the number of FCs between Frontal_Sup_L and Frontal_Mid_L, Frontal_Sup_R/Frontal_Sup_Medial_R and Frontal_Mid_R, Frontal_Mid (L/R) and Frontal_Sup_Medial_L, Frontal_Sup_L/Frontal_Mid (L/R) and Precuneus_L, Frontal_Sup_R/Frontal_Mid (L/R) /Parietal_Sup_R and Precuneus_R are higher than those between other regions. At the same time, there is almost no FCs between Frontal_Sup_L/Frontal_Sup_Medial_R/Precuneus (L/R) and

Cuneus_L, Frontal_Sup_R/Frontal_Sup_Medial (L/R)/Cuneus_L/Precuneus (L/R)/Temporal_Pole_Sup_R and Temporal_Pole_Sup_L, Frontal_Sup_R/Parietal_Sup_R/Cuneus_L/Frontal_Sup_Medial_R/Precuneus (L/R) and Temporal_Pole_Sup_R.

These results can also be found from **Fig.3**, which shows the number of FCs with significant differences among the common activity voxels between different AAL regions in **Table 2**. We can draw the similar conclusions of **Fig.2** from **Fig.3**. Moreover, it can be further found from the figure that there are many FCs with significant differences between the common activity voxels of AD group and HC group in each AAL brain region except for the regions of Cuneus_L, Parietal_Sup_R and Temporal_Pole_Sup_R.

Fig.4 shows the classification accuracy, sensitivity and specificity of AD and HC subjects, which obtained using SVM on all FCs and FCs with significant differences between the common activity voxels of AD and HC groups. In particular, the value of parameter $C = 1$ and linear kernel function were used in the SVM, and the cross-validation was used to control the over-fitting problem [41]. Specifically, the dataset was randomly divided into halves, 80% of which were used as training and the rest 20% were used as testing, and this process was run 1000 times. The average accuracy of the 1000 times was used as the final results, and the sensitivity and specificity of these 1000 results were also given simultaneously, as shown in the figure. It can be seen clearly from the figure that the classification accuracy, sensitivity and specificity obtained using FCs with significant differences are significant higher than those obtained using all FCs between the common activity voxels of AD and HC groups, which means that the FCs with significant differences between the common activity voxels play an important role in the classification of AD and HC subjects and has a potential value for the AD neural pathogenesis research.

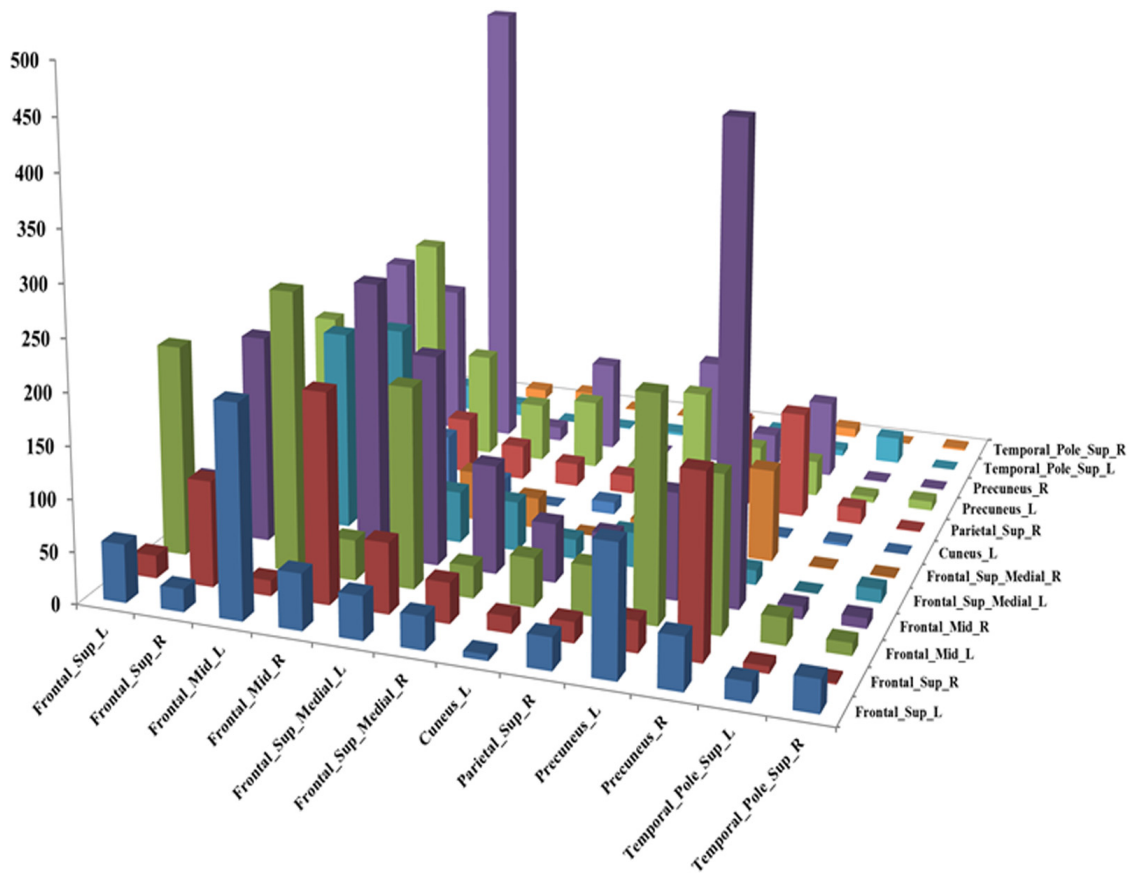


FIGURE 3. The number of FCs with significant differences between the AAL regions corresponding to the activity voxels of FCs with significant differences.

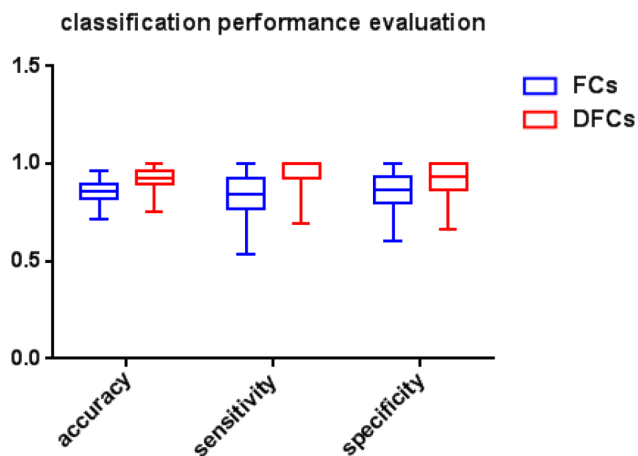


FIGURE 4. The classification accuracy, sensitivity and specificity using all FCs and FCs with significant differences (DFCs) between the common activated voxels of AD group and HC group.

In order to evaluate whether the change of this parameter affects the stability of SVM classification results, multiple thresholds are adopted for the study in this paper, and the classification results of SVM are shown in Fig.5. It can be seen from the figure that the change of threshold parameters does

not affect the robustness of SVM classification results, and the advantages of classification performance become more and more obvious with the increase of threshold parameters, especially for the cases of threshold parameter larger than 2.0. Meanwhile, in order to verify the classification effect of SVM, we compared the classification results of Linear SVM (LSVM) with a variety of classification algorithms in the revised manuscript according to the comment of reviewer, including Decision Tree (DT), Linear Discriminant (LD), Logistic Regression (LR), K-Nearest Neighbor (KNN), and Gaussian SVM (GSVM), as shown in Table 3 . It can be seen clearly that the classification effect of LSVM is better than that of other methods regard to accuracy, sensitivity and specificity.

V. DISCUSSION

Currently, the most studies are mainly focus on the activation degree of brain regions, brain networks or voxels as well as the FC differences between them in the fMRI data analysis of AD, but there are few studies to explore the difference between AD and HC groups from the perspective of activity voxel. Recently, Armananzas *et al.* used the activity voxels as features to classify AD patients from healthy subjects, and

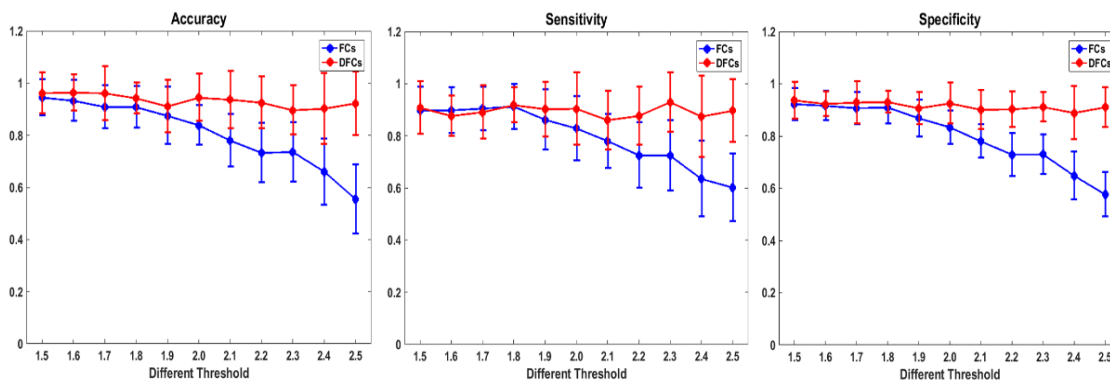


FIGURE 5. The classification accuracy, sensitivity and specificity using all FCs and FCs with significant differences (DFCs) between the common activity voxels on the cases of different threshold parameters, which ranges from 1.5 to 2.5 with a step-length of 0.1.

TABLE 2. The statistical distribution of activity voxels with FCS of significant difference in the aal brain regions and their corresponding MNI coordinates. The “voxel number” represents the number of activity voxels in the brain region, And The “voxel ratio” represents the percentage of obtained activity voxels over the total voxels in the corresponding brain region.

AAL Regions	MNI coordinates	voxel number	voxel ratio
Frontal_Sup_L	(-18.45, 34.81, 42.2)	22	0.0204
Frontal_Sup_R	(21.9, 31.12, 43.82)	17	0.0147
Frontal_Mid_L	(-33.43, 32.73, 35.46)	42	0.0290
Frontal_Mid_R	(37.59, 33.06, 34.04)	32	0.0212
Frontal_Sup_Medial_L	(-4.8, 49.17, 30.89)	10	0.0118
Frontal_Sup_Medial_R	(9.1, 50.84, 30.22)	11	0.0172
Cuneus_L	(-5.93, -80.13, 27.22)	13	0.0290
Parietal_Sup_R	(26.11, -59.18, 62.06)	11	0.0170
Precuneus_L	(-42.46, -22.63, 48.92)	14	0.0130
Precuneus_R	(9.98, -56.05, 43.77)	19	0.0203
Temporal_Pole_Sup_L	(-39.88, 15.14, -20.18)	20	0.0524
Temporal_Pole_Sup_R	(48.25, 14.75, -16.86)	11	0.0275

TABLE 3. The classification performance of multiple algorithms, including DT, LD, LR, KNN, GSVM, and LSVM.

	accuracy	sensitivity	specificity
DT	57.10%	46.15%	66.67%
LD	82.10%	76.92%	86.67%
LR	78.60%	84.62%	73.33%
KNN	78.60%	100%	60%
LSVM	92.90%	100%	86.67%
GSVM	89.30%	92.31%	86.67%

achieved a good recognition effect [37]. However, they did not explore how FCs between these activity voxels differs in healthy subjects and AD patients. Moreover, we can easily obtain the differences for different activity voxels between AD and HC groups, but whether the common activity voxels show the consistency between them, there are no relevant

researches at present. In addition, Irajia *et al.* showed that the analysis of fMRI data in the connectivity domain had more advantageous than that in the time domain, and could better explain the correlation rules of neurological diseases such AD [42].

Therefore, the main contribution of this study is to explore the distribution of activity voxels and FCs between them in AD and HC groups, and to analyze the FCs of common activity voxels between AD and HC groups in the connectivity domain, and to study their roles in distinguishing between AD patients and healthy subjects, so as to provide a certain reference basis for exploring the neuropathogenesis of AD. First, we made a statistical analysis of the distribution of activated voxels for the two groups of subjects in the AAL template, as shown in **Table 1**, in which the activity voxels were obtained by the method proposed in this paper, and elaborated on the cognitive function of the brain areas related to these activity voxels and the correlation with AD research.

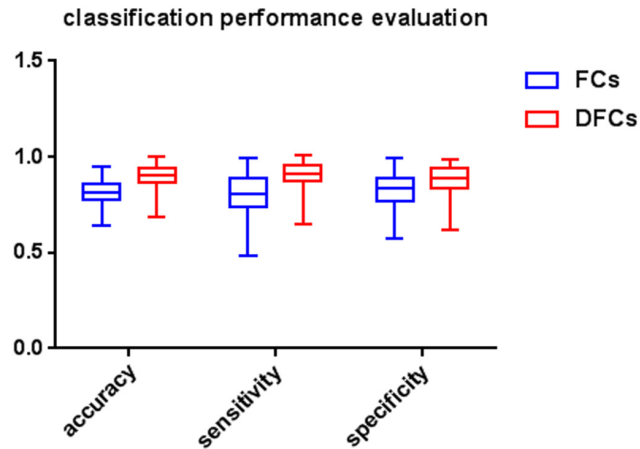


FIGURE 6. The classification accuracy, sensitivity and specificity using all FCs and DFCs between the common activated voxels of MCI and HC groups.

Although the information about brain activity in AD patients can be obtained directly by comparing the differences of brain regions or activity voxels between AD and HC groups, there is no relevant conclusion on whether the common activity voxels are consistent in AD patients and healthy subjects. Moreover, recent studies have shown that the analysis of fMRI data in the connectivity domain has more advantageous than that in the time domain, which can be better used to explain the related laws of neurological diseases such as AD [42]. Therefore, we also conducted a detailed and in-depth analysis on the FCs between the same activity voxels of the two groups of subjects in addition to directly analyze the distribution of activated voxel in the AD and HC groups as well as the related neurocognitive functions, and found the FCs with significant differences between AD and HC groups.

Then, these FCs with significant differences and the distribution of related activity voxels in the AAL brain regions were analyzed, which were consistent with previous studies about AD. In order to further verify the dissimilarity of the FCs with significant difference between AD patients and HC subjects and the significance of them in studying neuromechanism of AD, they were used as features to classify AD patients and HC subjects through SVM. SVM is a supervised machine learning algorithm derived from binary classification problem. Its advantage is that it can get much better results than other algorithms on a small sample training set, with strong generalization ability, fast classification speed and easy interpretation of results. This is because its optimizations object aims to minimize structural risk, not empirical risk, thus reducing the requirement for data size and data distribution. The results showed that these FCs have a good discrimination performance in recognition of AD patients from HC subjects compared with all FCs between the common activity voxels of the two groups, thus providing certain reference significances for the AD research at voxel level.

Many methods have been used for the study of AD classification, such as Deep Learning, AdaBoost, Support Vector Machines and Convolutional Neural Network [33]–[36], which have achieved a good classification effect, but they used different classification features, so there were different problems. For example, it is easy to ignore the boundary information between structural regions and functional regions when using 90 brain regions in AAL template to construct classification features from the fMRI data of AD, and also easy to miss the functional integration information of other brain regions on AD when specific brain functional regions are used. In addition, although recent studies have explored AD from the perspective of activated voxels [37], they have not considered the role of FCs between these activated voxels in AD prediction, and studies have shown that fMRI data analysis in the connectivity domain had more advantageous than that in time domain [42]. Therefore, this paper carried out relevant studies on AD based on the FCs between activated voxels, and achieved a good classification accuracy, which has a certain significance for the clinical diagnosis of AD.

As a classical method for detecting brain functional connectivity, ICA is a blind source separation method based on higher order statistical moments. The purpose of ICA is to decompose the observed multivariate data into the source signals which are assumed statistically independent and non-Gaussian, and it has been successfully applied to study of fMRI data and to find the underlying independent sources. Therefore, it was adopted to detect the activity voxels in this study, in which the number of ICs was estimated by MDL. When using ICA to obtain the activity voxels for each subject, the brain mask was built automatically from the data itself, and then the fMRI data within the mask was used to obtain the activity voxels. Since all subjects had the same amount of voxels after the preprocessing of spatial normalization, so that they were corresponding to the same spatial locations in the standardized space. Once the activity voxels of each subject and their corresponding positions in the standardized space are obtained, we can calculate the common activity voxels of all subjects to represent the common active areas of the corresponding group. In fact, although there are some structural differences in the brains of each subject, most areas remain the same in the standardized space, so this method allows us to preserve some individual differences while also maximizing their commonality at the activity voxel level.

In the process of obtaining activity voxels, the activity voxels in the brain region for each subject were used to preserve the individual specificity, while the common activity voxels of all subjects were used as the common activity region on behalf of the corresponding subjects in the group, such as AD group and HC group in our study. Furthermore, the common activity voxels need to be further screened through the known structural templates such as the AAL template, to determine which activity voxel is really caused by brain information. Only those voxels that belong to corresponding brain regions and reach a certain amount and proportion can

be selected as the truly meaningful activity voxels, as shown in **Table 1** of this paper. It can be seen from the table that these voxels are located in areas of the brain associated with certain cognitive functions.

This may be because the noise-induced activity voxel is less likely to occur in all ICs at once than in real neural signals. In general, the activity voxels with physiologically significant have little difference at the results of different running times, while the noise-induced pseudoactivator voxels vary greatly between subjects. Currently, there have been a number of preliminary studies in this area. For example, Armananzas *et al.* proposed the direct use of fMRI activation voxels of brain to tackle the automatic pattern analysis of AD and healthy individuals by applying different machine learning techniques to the classification of fMRI data for this purpose [37]. Therefore, the activity voxel detection method proposed in this paper can retain the useful voxels of physiological significance to the maximum extent and remove the voxels caused by noise. These voxels not only depict the differences between individuals, but also reflect the common characteristics at the group level.

In addition, it is worth noting that the classification method proposed in this paper may be related to some factors, such as the number of subjects in AD and HC groups and the number of ICs used in ICA. Different activity voxels may be obtained due to the different number of subjects or ICs, and it further leads to different FCs samples. Meanwhile, it needs to predefine a threshold parameter θ in formula (2) as the standard to determine whether a voxel is activity, and different voxels may be obtained with different threshold θ in the process of activity voxel calculation. In this study, $\theta = 2$ was selected as the threshold parameter to determine the activity voxels in the ICs after z-score transformation.

In order to illustrate the effectiveness of FCs samples selection in the proposed classification method, the FCs between the common activity voxels of AD and HC groups are used as the input samples for SVM classification. However, compared with the using of FCs with significant differences as input samples, the classification accuracy is significantly lower. Furthermore, the sample size of FCs with significant differences is far lower than that of FCs between the common activity voxels, which reduces the computational complexity of SVM in classification. In other words, the original sample contains many FCs that have no practical effect on classification. Therefore, the proposed classification method is expected to provide a new way for the clinical auxiliary diagnosis and prediction of AD.

In this paper, the fMRI data of AD and HC groups in this study were downloaded from public database ADNI (<http://adni.loni.usc.edu/>), which had the same fMRI data scanning parameters and similar demographic information. But there was no fMRI data of mild cognitive impairment (MCI) with the corresponding conditions, so only the fMRI data of AD and HC were used for this study, which was indeed a deficiency. Therefore, in order to illustrate the significance of the proposed method in the diagnosis of MCI, we used

the fMRI data of 50 patients with MCI from ADNI database (www.adni-info.org) to evaluate the performance of the proposed method, and the data description and results analysis were presented in APPENDIX.

VI. CONCLUSION

In this paper, we proposed a new method for the identification of AD patients from HC subjects by using the FCs between the activity voxels in the fMRI data. ICA was firstly used to detect the common activity voxels from the fMRI data of AD and HC groups, and then the FCs of subjects in the two groups between these voxels were obtained. Finally, two sample T-test was used to find the FCs with significant differences between AD group and HC group, and then those were further used as the feature samples of SVM for the classification of AD patients. The experimental results demonstrated that the proposed method could obtain a high classification accuracy for AD patients, and also showed that more regions and more voxels in some brain regions were activity in the brain of AD patients, which may be due to the reason of compensation mechanism in the brain, and it provided a new perspective for the study of AD pathogenesis.

According to the statistical distribution of activity voxels of AD and HC groups in **Table 1**, the results showed that more brain regions of Inferior frontal gyrus, Supplementary motor area, Parahippocampal gyrus, Lingual gyrus, Postcentral gyrus, Inferior parietal, Precuneus and Superior temporal gyrus were active in AD group compared with HC group, although there were no activity voxels in Frontal_Mid_Orb_L of Middle frontal gyrus, Frontal_Mid_Orb_L of Superior frontal gyrus, and Parietal_Sup_L of Superior parietal gyrus. They were mainly located in Prefrontal lobe, Parietal lobe, Temporal lobe, Occipital lobe and other parts of Frontal lobe. In addition, more voxels in Superior frontal gyrus, Middle frontal gyrus, Superior frontal gyrus and Superior temporal gyrus (Temporal pole) were activity in AD group, and they were mainly in Prefrontal lobe and Temporal lobe. These results were consistent with previous studies about AD, which reports significant changes in these brain regions, and the changes of FCs within each brain lobe as well as those between brain lobes in AD [43]–[46].

Moreover, the results in **Fig.4-Fig.5** showed that the identification performance using only FCs with significant difference between the activity voxels was better than that using all FCs between the activity voxels in the classification of AD patients and HC subjects. According to the distribution of the activity voxels related to FCs with significant difference shown in **Table 2**, they were mainly located in the brain regions of Superior frontal gyrus, Middle frontal gyrus, Superior frontal gyrus, Superior parietal gyrus, and Precuneus. Based on the statistics of the number of FCs with significant difference between different brain regions in **Fig.3**, we found that the FCs between activity voxels within prefrontal lobe as well as those between prefrontal and parietal lobes played an important role in the classification prediction of AD patients, which were shown in **Fig.2**. Therefore, we will further use

more AD, MCI and HC fMRI data to evaluate the effectiveness of the proposed algorithm in the future, which may be providing a new perspective for the study of AD pathogenesis mechanism.

APPENDIX

The resting-state fMRI datasets of 50 patients with MCI from 66 to 78 (mean = 72.40 ± 6.32) from ADNI database were used to evaluate the classification performance of the proposed method in this study. According to ADNI protocol, the resting-state fMRI data was acquired on 3.0 Tesla Philips scanners. The images dataset was acquired using single-shot SENSE gradient an EPI with 38 slices, providing whole-brain coverage and 160 volumes, a TR of 3s, a TE of 30ms, a FA of 90, and a scan resolution of 64×64 . The in-plane resolution was $3.39 \text{ mm} \times 3.39 \text{ mm}$, and the slice thickness was 3.39 mm.

After these MCI data have processed by the same data preprocessing steps in Section B “Data preprocessing” of III “EXPERIMENTS AND DATA PROCESSING”, the FCs and DFCs features corresponding to MCI and NC groups are obtained by using the method proposed in this paper. Then, we also used SVM with linear kernel function to classify MCI patients and HC subjects, and the classification results were shown in **Fig.s1**. **Fig.s1** showed the classification accuracy, sensitivity and specificity of MCI and HC subjects, which obtained using SVM on all FCs and DFCs between the common activity voxels of MCI and HC groups. It can be seen clearly from the figure that a better classification accuracy, sensitivity and specificity has been obtained by using DFCs between the common activity voxels of MCI and HC groups, which means that the DFCs between the common activity voxels play an important role in the classification of MCI and HC subjects and has a potential value for the MCI diagnosed.

REFERENCES

- [1] C. Reitz, C. Brayne, and R. Mayeux, “Epidemiology of Alzheimer disease,” *Nature Rev. Neurol.*, vol. 7, no. 3, pp. 137–152, 2011.
- [2] *World Alzheimer’s Report 2018: The State of the Art of Dementia Research*. New Frontiers, Alzheimer’s Disease Int., London, U.K., Sep. 2018.
- [3] X. Wang, J. Qi, Y. Yang, and P. Yang, “A survey of disease progression modeling techniques for Alzheimer’s diseases,” in *Proc. IEEE 17th Int. Conf. Ind. Informat. (INDIN)*, Jul. 2019, pp. 1237–1242.
- [4] G. Erus, M. Habes, and C. Davatzikos, “Machine learning based imaging biomarkers in large scale population studies: A neuroimaging perspective,” in *Handbook of Medical Image Computing and Computer Assisted Intervention*. New York, NY, USA: Academic, 2020, pp. 379–399.
- [5] J. C. Jack, Jr., *et al.*, “Longitudinal tau PET in ageing and Alzheimer’s disease,” *Brain*, vol. 141, no. 5, pp. 1517–1528, 2018.
- [6] K. Bruggen *et al.*, “The European DTI study on dementia—A multicenter DTI and MRI study on Alzheimer’s disease and mild cognitive impairment,” *NeuroImage*, vol. 144, pp. 305–308, Jan. 2017.
- [7] N. Houmani *et al.*, “Diagnosis of Alzheimer’s disease with electroencephalography in a differential framework,” *PLoS ONE*, vol. 13, no. 3, 2018, Art. no. e0193607.
- [8] B. Hart, I. Cribben, and M. Fiecas, “A longitudinal model for functional connectivity networks using resting-state fMRI,” *NeuroImage*, vol. 178, pp. 687–701, Sep. 2018.
- [9] H.-J. Li, X.-H. Hou, H.-H. Liu, C.-L. Yue, Y. He, and X.-N. Zuo, “Toward systems neuroscience in mild cognitive impairment and Alzheimer’s disease: A meta-analysis of 75 fMRI studies,” *Hum. Brain Mapping*, vol. 36, no. 3, pp. 1217–1232, Mar. 2015.
- [10] J. W. Vogel *et al.*, “Brain properties predict proximity to symptom onset in sporadic Alzheimer’s disease,” *Brain*, vol. 141, no. 6, pp. 1871–1883, Jun. 2018.
- [11] L. Wang *et al.*, “Changes in hippocampal connectivity in the early stages of Alzheimer’s disease: Evidence from resting state fMRI,” *NeuroImage*, vol. 31, no. 2, pp. 496–504, Jun. 2006.
- [12] Z. Wang *et al.*, “Differentially disrupted functional connectivity of the subregions of the inferior parietal lobule in Alzheimer’s disease,” *Brain Struct. Function*, vol. 220, no. 2, pp. 745–762, Mar. 2015.
- [13] Z. Wang *et al.*, “Changes in thalamus connectivity in mild cognitive impairment: Evidence from resting state fMRI,” *Eur. J. Radiol.*, vol. 81, no. 2, pp. 277–285, Feb. 2012.
- [14] M. D. Greicius, G. Srivastava, A. L. Reiss, and V. Menon, “Default-mode network activity distinguishes Alzheimer’s disease from healthy aging: Evidence from functional MRI,” *Proc. Nat. Acad. Sci. USA*, vol. 101, no. 13, pp. 4637–4642, 2004.
- [15] K. Wang *et al.*, “Altered functional connectivity in early Alzheimer’s disease: A resting-state fMRI study,” *Hum. Brain Mapping*, vol. 28, no. 10, pp. 967–978, 2007.
- [16] M. A. A. Binnewijzend *et al.*, “Brain network alterations in Alzheimer’s disease measured by eigenvector centrality in fMRI are related to cognition and CSF biomarkers,” *Hum. Brain Mapping*, vol. 35, no. 5, pp. 2383–2393, May 2014.
- [17] S. J. Teipel *et al.*, “Robust detection of impaired resting state functional connectivity networks in Alzheimer’s disease using elastic net regularized regression,” *Frontiers Aging Neurosci.*, vol. 8, p. 318, Jan. 2017.
- [18] J. S. Damoiseaux, K. E. Prater, B. L. Miller, and M. D. Greicius, “Functional connectivity tracks clinical deterioration in Alzheimer’s disease,” *Neurobiol. Aging*, vol. 33, no. 4, pp. 828.e19–828.e30, Apr. 2012.
- [19] Y. Zhong *et al.*, “Altered effective connectivity patterns of the default mode network in Alzheimer’s disease: An fMRI study,” *Neurosci. Lett.*, vol. 578, pp. 171–175, Aug. 2014.
- [20] L. Squarcina *et al.*, “Classification of first-episode psychosis in a large cohort of patients using support vector machine and multiple kernel learning techniques,” *NeuroImage*, vol. 145, pp. 238–245, Jan. 2017.
- [21] F. Liu *et al.*, “Multivariate classification of social anxiety disorder using whole brain functional connectivity,” *Brain Struct. Function*, vol. 220, no. 1, pp. 101–115, Jan. 2015.
- [22] V. M. Vergara, A. R. Mayer, E. Damaraju, K. Hutchison, and V. D. Calhoun, “The effect of preprocessing pipelines in subject classification and detection of abnormal resting state functional network connectivity using group ICA,” *NeuroImage*, vol. 145, pp. 365–376, Jan. 2017.
- [23] V. Pariyadath, E. A. Stein, and T. J. Ross, “Machine learning classification of resting state functional connectivity predicts smoking status,” *Frontiers Hum. Neurosci.*, vol. 8, p. 425, Jun. 2014.
- [24] L.-L. Zeng, H. Shen, L. Liu, and D. Hu, “Unsupervised classification of major depression using functional connectivity MRI,” *Hum. Brain Mapping*, vol. 35, no. 4, pp. 1630–1641, Apr. 2014.
- [25] F. de Vos *et al.*, “A comprehensive analysis of resting state fMRI measures to classify individual patients with Alzheimer’s disease,” *NeuroImage*, vol. 167, pp. 62–72, Feb. 2018.
- [26] S. Rathore, M. Habes, M. A. Iftikhar, A. Shacklett, and C. Davatzikos, “A review on neuroimaging-based classification studies and associated feature extraction methods for Alzheimer’s disease and its prodromal stages,” *NeuroImage*, vol. 155, pp. 530–548, Jul. 2017.
- [27] T. M. Schouten *et al.*, “Individual classification of Alzheimer’s disease with diffusion magnetic resonance imaging,” *NeuroImage*, vol. 152, pp. 476–481, May 2017.
- [28] A. Khazae, A. Ebrahimzadeh, and A. Babajani-Feremi, “Classification of patients with MCI and AD from healthy controls using directed graph measures of resting-state fMRI,” *Behav. Brain Res.*, vol. 322, pp. 339–350, Mar. 2017.
- [29] E. Challis, P. Hurley, L. Serra, M. Bozzali, S. Oliver, and M. Cercignani, “Gaussian process classification of Alzheimer’s disease and mild cognitive impairment from resting-state fMRI,” *NeuroImage*, vol. 112, pp. 232–243, May 2015.
- [30] E. E. Tripoliti, D. I. Fotiadis, M. Argyropoulou, and G. Manis, “A six stage approach for the diagnosis of the Alzheimer’s disease based on fMRI data,” *J. Biomed. Informat.*, vol. 43, no. 2, pp. 307–320, Apr. 2010.
- [31] Z. Dai *et al.*, “Discriminative analysis of early Alzheimer’s disease using multi-modal imaging and multi-level characterization with multi-classifier (M3),” *NeuroImage*, vol. 59, no. 3, pp. 2187–2195, Feb. 2012.

- [32] A. Khazaei, A. Ebrahimzadeh, and A. Babajani-Feremi, "Application of advanced machine learning methods on resting-state fMRI network for identification of mild cognitive impairment and Alzheimer's disease," *Brain Imag. Behav.*, vol. 10, no. 3, pp. 799–817, Sep. 2016.
- [33] C. Hu, R. Ju, Y. Shen, P. Zhou, and Q. Li, "Clinical decision support for Alzheimer's disease based on deep learning and brain network," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2016, pp. 1–6.
- [34] J. H. Morra, Z. Tu, L. G. Apostolova, A. E. Green, A. W. Toga, and P. M. Thompson, "Comparison of AdaBoost and support vector machines for detecting Alzheimer's disease through automated hippocampal segmentation," *IEEE Trans. Med. Imag.*, vol. 29, no. 1, pp. 30–43, Jan. 2010.
- [35] A. Khazaei, A. Ebrahimzadeh, and A. Babajani-Feremi, "Identifying patients with Alzheimer's disease using resting-state fMRI and graph theory," *Clin. Neurophysiol.*, vol. 126, no. 11, pp. 2132–2141, Nov. 2015.
- [36] S. Sarraf, J. Anderson, and G. Tofighi, "DeepAD: Alzheimer's disease classification via deep convolutional neural networks using MRI and fMRI," *BioRxiv*, Jan. 2016, Art. no. 070441.
- [37] R. Armananzas, M. Iglesias, D. A. Morales, and L. Alonso-Nanclares, "Voxel-based diagnosis of Alzheimer's disease using classifier ensembles," *IEEE J. Biomed. Health Inform.*, vol. 21, no. 3, pp. 778–784, May 2017.
- [38] Y. Shi, W. Zeng, N. Wang, and D. Chen, "A novel fMRI group data analysis method based on data-driven reference extracting from group subjects," *Comput. Methods Programs Biomed.*, vol. 122, no. 3, pp. 362–371, Dec. 2015.
- [39] J. Himberg, A. Hyvärinen, and F. Esposito, "Validating the independent components of neuroimaging time series via clustering and visualization," *NeuroImage*, vol. 22, no. 3, pp. 1214–1222, Jul. 2004.
- [40] Y.-O. Li, T. Adali, and V. D. Calhoun, "Estimating the number of independent components for functional magnetic resonance imaging data," *Hum. Brain Mapping*, vol. 28, no. 11, pp. 1251–1266, Nov. 2007.
- [41] F. Pereira, T. Mitchell, and M. Botvinick, "Machine learning classifiers and fMRI: A tutorial overview," *NeuroImage*, vol. 45, no. 1, pp. S199–S209, Mar. 2009.
- [42] A. Irajy *et al.*, "The connectivity domain: Analyzing resting state fMRI data using feature-based data-driven and model-based methods," *NeuroImage*, vol. 134, pp. 494–507, Jul. 2016.
- [43] F. Eustache *et al.*, "'In the course of time': A PET study of the cerebral substrates of autobiographical amnesia in Alzheimer's disease," *Brain*, vol. 127, no. 7, pp. 1549–1560, Jul. 2004.
- [44] G. S. Smith *et al.*, "Cholinergic modulation of the cerebral metabolic response to citalopram in Alzheimer's disease," *Brain*, vol. 132, no. 2, pp. 392–401, Feb. 2009.
- [45] R. J. Melrose, O. M. Campa, D. G. Harwood, S. Osato, M. A. Mandelkern, and D. L. Sultzer, "The neural correlates of naming and fluency deficits in Alzheimer's disease: An FDG-PET study," *Int. J. Geriatric Psychiatry*, vol. 24, no. 8, pp. 885–893, Aug. 2009.
- [46] H.-Y. Zhang *et al.*, "Resting brain connectivity: Changes during the progress of Alzheimer disease," *Radiology*, vol. 256, no. 2, pp. 598–606, 2010.

...