

Received July 18, 2018, accepted August 26, 2018, date of publication September 5, 2018, date of current version September 28, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2868813

# Multi-Center Brain Imaging Classification Using a Novel 3D CNN Approach

LIN YUAN<sup>ID</sup>, XUE WEI, HUI SHEN, LING-LI ZENG<sup>ID</sup>, AND DEWEN HU<sup>ID</sup>, (Senior Member, IEEE)

College of Artificial Intelligence, National University of Defense Technology, Changsha 410073, China

Corresponding author: Dewen Hu (dwhu@nudt.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61722313, Grant 61773391, Grant 61503397, and Grant 61420106001 and in part by the Fok Ying Tung Education Foundation under Grant 161057.

**ABSTRACT** With the development of brain imaging technology, increasing amounts of magnetic resonance imaging data are being acquired, and traditional computational analysis methods based on single sites and small samples are facing substantial challenges. Deep learning technology, which is born via artificial intelligence, has shown the powerful ability to solve the classification problem based on big data in many studies, while it has not been widely used in brain imaging classification. Herein, we utilized our proposed novel 3-D deep adding neural network to classify 6008 samples from the largest data sets in the brain imaging field collected from more than 61 centers. The proposed method utilizes multiple convolutional layers to extract gradient information in different orientations and combines spatial information at two scales via the adding operation. High accuracy (over 92.5%) was obtained with a standard fivefold cross-validation strategy, demonstrating that the proposed method can effectively handle big data classifications from multiple centers. Compared with some traditional classification methods and some deep learning architectures, the proposed method was more accurate, demonstrating its stronger power to classify data from multiple centers. Our cross-site classification results prove that the proposed method is robust when training on a data set and testing on another data set. To the best of our knowledge, this paper is the first to classify neuroimaging data on such a large scale from multiple centers with such high accuracy. With its improved performance in classification and transferable program codes, the proposed method can potentially be used in intelligent medical treatment strategies and clinical practices based on mobile terminal.

**INDEX TERMS** Artificial intelligence, artificial neural networks, image classification, machine learning, magnetic resonance imaging.

## I. INTRODUCTION

Brain imaging data, including magnetic resonance imaging (MRI) data, have been widely used in neural science and clinical applications [1]–[4]. However, with the development of brain science and neuroimaging technology, increasing amounts of neuroimaging data are being acquired, and the number of samples is unprecedented. Traditional statistical analysis methods based on single sites and small samples are becoming problematic for addressing big data from multiple centers. Large individual differences and differences between centers will make the previous methods ineffective.

Recently, deep learning has undergone unprecedented development and soared in popularity [5]–[7]. Numerous companies and research groups have successively published all kinds of deep learning models to solve image classification problems in rather large datasets (such as ImageNet [9],

which has over 100,000 images). These models (such as AlexNet [10], VGGNet [11], GoogleInceptionNet [12], ResNet [13], and GAN [14] and so on) have been very successful at image classification. Some models, such as long short term memory (LSTM) [15], can effectively address video recognition and time series classifications. Furthermore, some deep learning approaches have been used in medical practice [16]. However, all the models are based on two-dimensional (2D) images or 2D images with time series. Because the human brain has a special spatial structure that is strongly related to brain functions and activities, ignoring the spatial structure of the human brain is not advisable.

Based on the development of the deep learning algorithm, some researchers have introduced models into the neuroimaging analysis field, with some model architectures being based on three-dimensional (3D) convolutional neural net-

works (CNNs). Dolz *et al.* [18] investigated a 3D full CNN for subcortical brain structure segmentation in MRI. They tested the model on the ABIDE dataset, which includes 1112 subjects, and yielded segmentations that were highly consistent with a standard atlas-based approach. Kawahara *et al.* [19] proposed using CNNs for brain networks to predict clinical neurodevelopmental outcomes, while Wachinger *et al.* [20] proposed a deep CNN that included three convolutional layers with pooling, batch normalization (BN), non-linearities and fully connected layers with dropout for segmenting neuroanatomy. Cole *et al.* [21] utilized a CNN to predict brain age using raw T1 MRI data on a large dataset of healthy adults (2001 subjects). Chen *et al.* [22] proposed using VoxResNet, comprising 25 layers, for segmenting key brain tissues from 3D magnetic resonance images, and this approach was awarded first place in the challenge out of 37 competitors. Ktena *et al.* [23] proposed to learn a graph similarity metric using a siamese graph convolutional neural network (s-GCN) to classify neuroimaging on brain connectivity networks. They tested the proposed framework on the ABIDE database and the UK Biobank, and this approach performed better than traditional methods. Zou *et al.* [24] proposed a 3D CNN classification approach to automatic diagnose attention deficit hyperactivity disorder (ADHD) and achieved the accuracy of 69.15%, which outperformed reported classifiers in the literature.

However, previous 3D deep learning approaches were focused on brain or neuroanatomy segmentation and brain maturity analyses, and few 3D CNN models for multi-center MRI data classifications exist on a large scale. In this paper, we propose using a 3D deep learning model, termed 3D deep adding neural network (3D DANet), to solve the human brain MRI classification problem. This model contains two parts; the first part includes one 3D convolutional layer and a 3D max pooling layer, while the second part includes four 3D convolutional layers and a 3D max pooling layer. Then, the two 3D pooling layers are added together as the input for the final full connection layer. Finally, we use softmax as the classifier to generate predictive labels. In this paper, we use the subjects' gender label as the predicting label because gender is the golden standard in the neuroimaging field and does not include subjective factors. The proposed method was tested with a standard 5-fold cross-validation strategy on over 6000 MRI samples of healthy people collected from more than 61 sites in six datasets. All MRI data were preprocessed in a uniform pipeline to maximally eradicate differences between scanners. Furthermore, we also use deconvolutional computation to show differences between groups to demonstrate the effectiveness of the proposed method.

## II. MATERIALS AND METHODS

### A. DATASETS

T1-weighted MRI scans comprised all the neuroimaging data used in the study, and all the scans were from six subsets collected from more than sixty sites or scanners. Details regarding the participants in the specific samples and the

respective acquisition parameters used are outlined below. We used gender as the predicting label, and the females and males included in each dataset are also outlined below.

#### 1) 1000 FUNCTIONAL CONNECTOMES PROJECT

The 1000 Functional Connectomes Project (FCP) publishes the unrestricted public release of 1200<sup>+</sup> resting-state functional MRI (Rs-fMRI) datasets independently collected from 33 sites ([http://fcon\\_1000.projects.nitrc.org/](http://fcon_1000.projects.nitrc.org/)). All datasets were generously donated by the principal investigators from the member sites to provide the broader imaging community complete access to a large-scale functional imaging dataset. The age and sex of the participants as well as the imaging center information are provided for each of the datasets. As we used gender as the predicting label to test the proposed model, some of centers that did not provide gender information were removed.

#### 2) HUMAN CONNECTOME PROJECT

The Human Connectome Project (HCP) is an ambitious 5-year effort to characterize brain connectivity and function and their variability in healthy adults [25] to study a population of 1200 subjects including T1- and T2-weighted MRI structural neuroimaging data. The HCP samples are described in further detail in [25]. All the HCP samples used herein were downloaded from ConnectomDB (<https://db.humanconnectome.org>) [26]. At the time of this manuscript submission, the HCP had released T1-weighted MRI data from 1090 subjects, which included 592 females and 498 males. We selected all the raw T1-weighted MRI data as the input for our preprocessing pipeline.

#### 3) BRAIN GENOMICS SUPERSTRUCT PROJECT

The goal of the Brain Genomics Superstruct Project (GSP) is to enable large-scale exploration of the links between brain function, behavior, and ultimately, genetic variation [27]. MRI data from 1570 unique subjects are included in five sub-datasets, with each dataset containing high-resolution anatomical T1-weighted structural MRI scans. Imaging data, provided in the NIfTI format, were obtained from 905 female and 665 male subjects.

#### 4) NATHAN KLINE INSTITUTE

The enhanced Nathan Kline Institute-Rockland Sample (NKI-RS, [http://fcon\\_1000.projects.nitrc.org/indi/enhanced/](http://fcon_1000.projects.nitrc.org/indi/enhanced/)) [28] is an institute that aims to create a large-scale (N>1000), deeply phenotyped, community-ascertained, lifespan sample (ages 6-85 years old) comprising advanced neuroimaging and genetics data. We collected 461 healthy participants from the NKI-RS database, which included 257 females and 204 males. All approvals and procedures for collection and data sharing were approved by the NKI institutional review board, and each participant were provided written informed consent. MRI data were acquired on a 3.0 T SIMENS Trio scanner. For each subject, high-resolution T1-weighted images were acquired using

the magnetization-prepared rapid gradient echo (MPRAGE) sequence.

### 5) CONSORTIUM FOR RELIABILITY AND REPRODUCIBILITY

The Consortium for Reliability and Reproducibility (CoRR, [http://fcon\\_1000.projects.nitrc.org/indi/CoRR/html/index.html](http://fcon_1000.projects.nitrc.org/indi/CoRR/html/index.html)) aims to create an open science resource for the imaging community that facilitates the assessment of test-retest reliability and reproducibility for functional and structural connectomics. To accomplish this, they aggregated resting state fMRI (Rs-fMRI) and diffusion imaging data from laboratories worldwide, which are shared via the International Neuroimaging Data-sharing Initiative (INDI). The CoRR dataset includes 18 centers and 33 sites, 32 of which are currently available for download, comprising 1629 subjects and 3357 anatomical scans.

### 6) SOUTHWEST UNIVERSITY LONGITUDINAL IMAGING MULTIMODAL

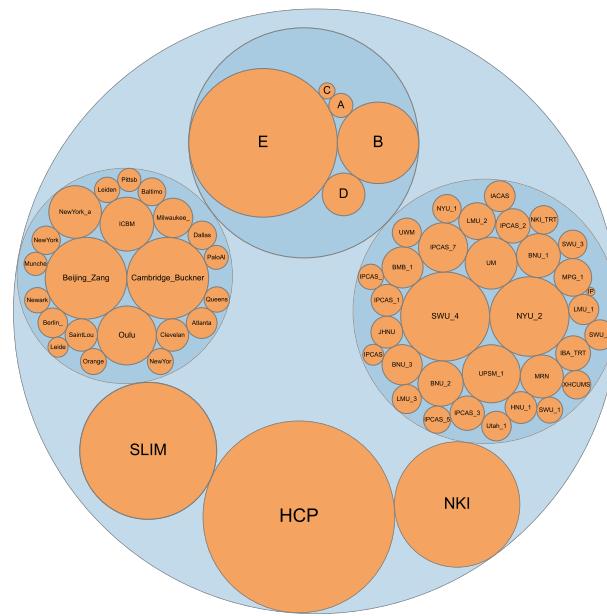
Southwest University Longitudinal Imaging Multimodal (SLIM, [http://fcon\\_1000.projects.nitrc.org/indi/retro/southwestuni\\_qiu\\_index.html](http://fcon_1000.projects.nitrc.org/indi/retro/southwestuni_qiu_index.html)) brain data comprise a long-term test-retest sample of young, healthy adults in Southwest China. The SLIM dataset aims to build a large sample with a long-term longitudinal design and a narrow age-span for the assessment of test-retest reliability and reproducibility of brain-behavior correlations as well as for the development of novel causal insights into these correlational findings. The SLIM dataset includes brain sMRI scans that provide a set of structural, diffusion and Rs-fMRI images acquired from 555 subjects (308 females and 247 males).

### 7) SUMMARY

The compositions of the datasets used herein are shown in Figure 1. Some repeated subjects (18 female subjects from the CoRR database were repeated in the NKI database; 115 female and 118 male subjects were repeated in the SLIM database), subjects missing gender information, or data that were not subjected to the data preprocessing pipeline due to pool imaging quality were removed. Therefore, the final dataset used herein totaled 6008 subjects, including 3292 females and 2716 males. Table 1 shows the details for each datasets.

### B. NEUROIMAGING DATA PREPROCESSING

To minimize the influence of different centers' machines and scanning parameters, all T1-MRI data from all datasets were preprocessed to generate normalized brain volume maps via the same preprocessing pipeline. In this study, we used the previously outlined protocol [29] to generate volumetric maps as the original input. All the structural images were preprocessed using SPM12 software (Wellcome Department of Imaging Neuroscience, University College London, UK, <http://www.fil.ion.ucl.ac.uk/spm>) and the CAT12 package (<http://dbm.neuro.uni-jena.de/cat12/>). All the T1 images were first normalized to the standard Montreal Neurological



**FIGURE 1.** The contents of the collected data used herein. The datasets comprise six subsets, including three from single centers and three from integrated centers. In total, 6008 subjects were utilized, including 1058 from the FCP, 1090 from the HCP, 1570 from the GSP, 461 from the NKI, 1274 from the CoRR and 555 from the SLIM. The large circles indicate the six subsets, and the small circles indicate the different sites in each subset.

**TABLE 1.** Subjects information for each subset used herein.

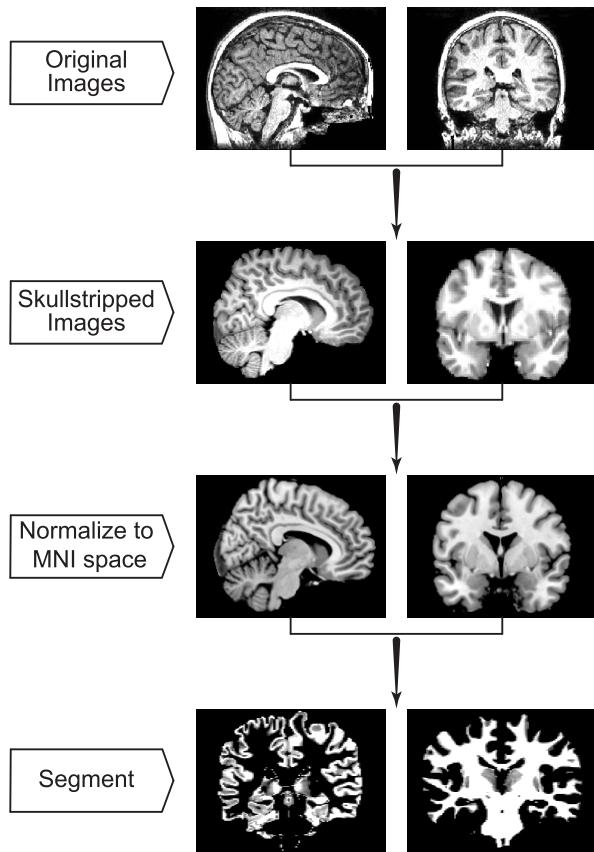
Dataset	Total number	Female/Male	Number of sites
HCP	1090	592/498	-
FCP	1058	596/462	22
GSP	1570	905/665	5
NKI	461	257/204	-
SLIM	555	308/247	-
CoRR	1274	634/640	31
Total	6008	3292/2716	>61

- means that the number of sites is not available

Institute (MNI) space, and the images were segmented into gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF) partitions using the tissue prior free segmentation routine of the “Segment Data” toolbox in CAT12. Then, using the options “Display One Slice For All Images” and “Check Sample Homogeneity Using Covariance” in “Check Data Quality”, we evaluated the segment and normalization quality. To retain the brain edge information and maintain the local differences, we did not utilize the smoothing step, which is a typical step in preprocessing MRI data. The detailed flowchart for structural MRI preprocessing is shown in Figure 2.

### C. 3D DEEP ADDING NEURAL NETWORK

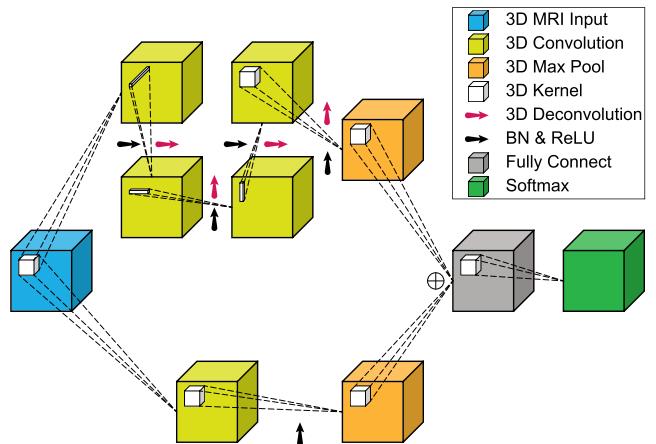
The CNN was first proposed by Lecun *et al.* [30] in 1998 and has been very actively investigated, especially recently.



**FIGURE 2.** Flowchart of the structural MRI preprocessing.

As mentioned in the *Introduction*, several different network architectures have been proposed, which have demonstrated state-of-the-art performances in many computer vision and speech recognition tasks [10]–[12]. We hypothesize that a CNN can provide the appropriate architecture to infer imaging features from processed brain MRI scans to predict a subject's gender. The 2D residual neural network (ResNet), which eases network training, was proposed by Kaiming He *et al.* [13] in 2015, and they won 1st place in the ILSVRC 2015 classification task. Some architectures based on the 3D CNN model for predicting brain maturity [21], predicting neurodevelopment [19], Alzheimer's disease classifications [31], brain segmentation [18], [20], [22] and skull stripping [32]. With the inspiration from ResNet and previous 3D CNN architectures, we propose the 3D DANet model for neuroimaging classification (Figure 3).

Our proposed 3D DANet architecture used MRI volumes of a specific size ( $z \times h \times w$ ) as inputs. Because we used CAT 12 as the preprocessing toolbox, the MRI resolutions of all subjects were normalized into  $1.5 \text{ mm} \times 1.5 \text{ mm} \times 1.5 \text{ mm}$ . Thus, in our application, the size of the input GM data was  $121 \times 145 \times 121$ . The predicted output was a single scalar representing the biological gender. A schematic illustration of the 3D DANet architecture, which includes two parts, is given in Figure 3. The bottom side contains two layers: a 3D convolutional layer (with a kernel of  $3 \times 3 \times 3$  and a stride of 2), a 3D



**FIGURE 3.** Schematic illustration of the proposed 3D DANet, which includes two branches. The bottom branch contains two layers: a 3D convolutional layer and a 3D max pooling layer. The top branch contains five layers: four 3D convolutional layers with four 3D deconvolutional layer and a 3D max pooling layer. Then, the two pooling layers are added as the input for the fully connected layer. Finally, softmax is utilized as the classifier. The blue cube indicates the 3D MRI input; the yellow cubes indicate the 3D convolutional layers with different kernels; the orange cubes indicate the 3D max pooling layers; the white cubes indicate the 3D convolutional kernels with different sizes; the gray cube indicates the fully connected layer; and the green cube indicates the softmax classifier. The red arrow indicate the 3D deconvolutional computation. The black arrow indicates the batch normalization and activation layer between the two convolutional layers.

BN layer, a rectified linear unit (ReLU), and finally, a 3D max pooling layer (with a stride of 1). The top side contains five layers: a 3D convolutional layer (with a kernel of  $3 \times 1 \times 1$ , a stride of 1 and a 3D deconvolutional layer), a 3D BN layer, a ReLU, a 3D convolutional layer (with a kernel of  $1 \times 3 \times 1$ , a stride of 1 and 3D deconvolutional computation), a 3D BN layer, a ReLU, a 3D convolutional layer (with a kernel of  $1 \times 1 \times 3$ , a stride of 1 and 3D deconvolutional computation), a 3D BN layer, a ReLU, a 3D convolutional layer (with a kernel of  $3 \times 3 \times 3$ , a stride of 1 and 3D deconvolutional computation), a 3D BN layer, a ReLU, and finally, a 3D max pooling layer (with a stride of 2). The 3D convolutional layer indicates that a convolutional calculation was performed with a 3D kernel in the 3D spatial space. The adding operation can utilize the information of different scales and can improve the performance of the proposed method. To show the detected ROIs, the proposed method involved deconvolutional computation for different convolutional layers. The kernel size of the 3D deconvolutional computation was the same as that used for the 3D convolutional layer. The BN represents the layer prior to activation, and the ReLU represents the method of activation. This function can be formulated as

$$f(x_i) = \max(0, x_i) \quad (1)$$

The number of feature channels was set to eight for all the layers. Like in standard CNNs, a fully connected layer was added at the end of the network to encode semantic information. Finally, the neurons in the final layer (the classification layer) were grouped into  $m = C$  feature maps, where  $C$  denotes the

number of classes ( $C$  typically equal 2 herein). To avoid overfitting, a dropout approach was used after the fully connected layer. The sparse softmax cross-entropy function was selected as the loss function, and the Adam algorithm was used as the optimizer. The probability score of the class  $c \in \{1, \dots, C\}$  was computed as follows:

$$p_c = \frac{\exp(y_L^c)}{\sum_{c'=1}^C \exp(y_L^{c'})} \quad (2)$$

All the results reported in Section III refer to the experiments in which the models were initialized with random truncated normally distributed parameters. Training the 3D DANet architectures with only GM input using a graphics processing unit (GPU, Nvidia GTX 1080) sped up the training. All the programs were written in TensorFlow ([www.tensorflow.org](http://www.tensorflow.org)), an open source software library, for numerical computation using data flow graphs with support for deep learning algorithms and GPU computing.

### III. RESULTS

In this section, we evaluate the performance of the proposed method, estimate the classification accuracy of the proposed method with two validation strategies, compare its performance with other methods, and show the features learned from the neural network.

#### A. 5-FOLD CROSS-VALIDATION

The classifier design can be viewed as a task of a learning model of the relationship between features and class labels in training sets. If the classifier can predict the classes of new samples satisfactorily, we can trust that it truly captures the relationship between features and classes. In this study, a standard 5-fold cross-validation strategy was used to estimate the performance of the proposed 3D DANet architecture, where all the subjects were separated into five folds randomly. One fold was used as the test set, while the others were used to train the model, and the test was repeated 5 times to ensure that every subject was tested. To remove the influence of the random separation and consider the computational time, we repeated the 5-fold cross-validation strategy five times. Based on the limitation of GPU memory and the sizes of the input images, we choose to utilize a batch size of five. The sensitivity (SS), specificity (SC), and generalization rate (GR) were defined from the results of each test result to quantify the performance of the classifier as follows:  $SS = TP/(TP + FN)$ ,  $SC = TN/(TN + FP)$ , and  $GR = (TP + TN)/(TP + FN + TN + FP)$ , where  $TP$  is the number of males correctly predicted,  $FN$  is the number of males classified as females,  $TN$  is the number of females correctly predicted, and  $FP$  is the number of females classified as males. Note that SS is the proportion of male subjects predicted correctly, SC is the proportion of female subjects predicted correctly, and GR is the overall accuracy of all subjects predicted correctly.

To remove the influence of the random separation of the datasets from the 5-fold strategy, we calculated the SS, SC

and GR of the collected datasets 5 times. The mean accuracy and standard deviation results for SS, SC, GR on each and the total dataset are shown in Table 2. We obtained an accuracy greater than 92.5% for the total dataset, which is the highest recently reported value [29], [33]. The table shows that the standard deviation is very small, indicating that the influence of random separation is negligible and that the model is robust for changing data.

**TABLE 2. Computational results of the proposed method with a standard 5-fold strategy tested five times.**

Dataset	Evaluation		
	SS	SC	GR
HCP	0.8882 ± 0.0091	0.9780 ± 0.0061	0.9370 ± 0.0043
FCP	0.8997 ± 0.0201	0.9172 ± 0.0161	0.9096 ± 0.0043
GSP	0.9454 ± 0.0157	0.9724 ± 0.0057	0.9609 ± 0.0035
NKI	0.9379 ± 0.0186	0.9105 ± 0.0206	0.9212 ± 0.0055
SLIM	0.9419 ± 0.0102	0.9524 ± 0.0131	0.9478 ± 0.0065
CoRR	0.8880 ± 0.0099	0.8769 ± 0.0182	0.8825 ± 0.0043
Total	0.9125 ± 0.0128	0.9383 ± 0.0098	0.9267 ± 0.0017

#### B. LEAVE ONE SUBSET OUT CROSS-VALIDATION

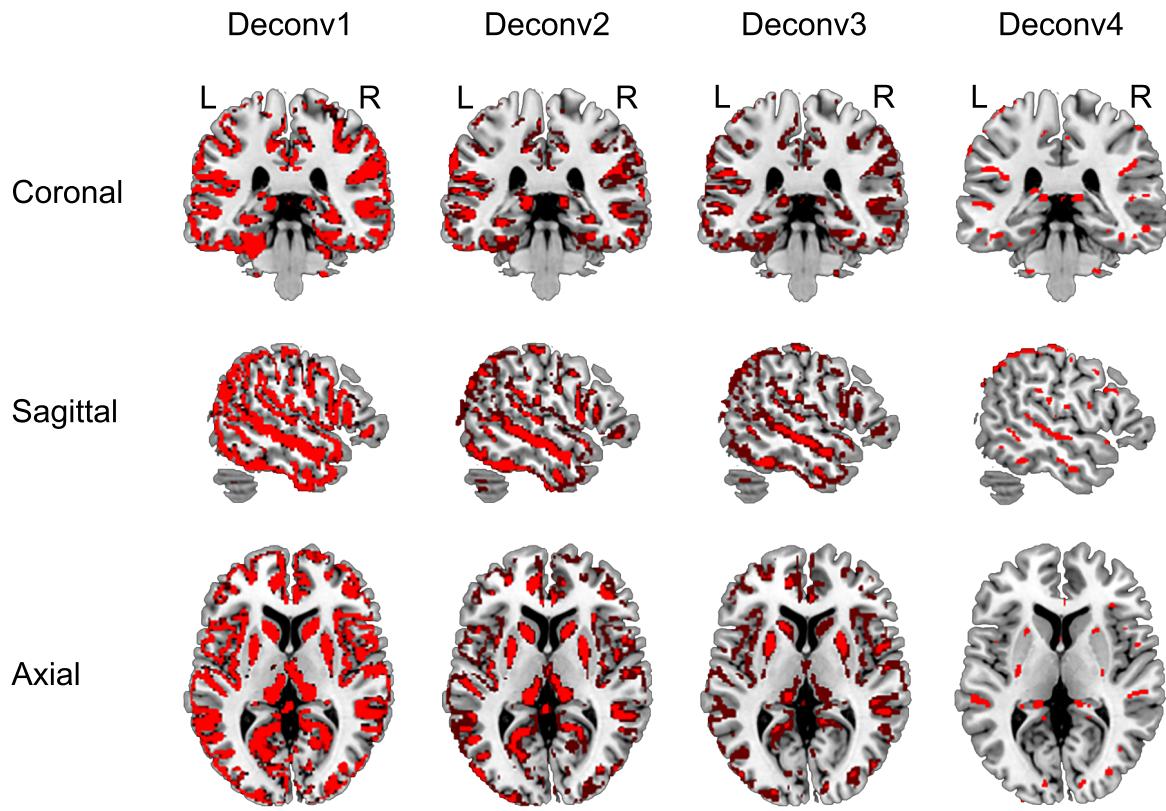
To evaluate the performance of cross-site classification, we also tested the proposed approach using the leave one subset out cross-validation (LOSO-CV) strategy, which means that the model was tested with one subset, while training was performed using the other subsets. Because six subsets were used herein, the LOSO-CV strategy was repeated six times such that all the subsets were tested in the proposed approach. The results are shown in Table 3.

**TABLE 3. Computational results of the proposed method using the LOSO-CV strategy. 'Dataset' indicates the subset that was removed for testing, and 'Total' indicates the statistical results from all the subsets.**

Dataset	Evaluation		
	SS	SC	GR
HCP	0.9598	0.8412	0.8954
FCP	0.8983	0.8842	0.8828
GSP	0.9459	0.9160	0.9046
NKI	0.8971	0.9182	0.9050
SLIM	0.9474	0.9156	0.9073
CoRR	0.8891	0.8691	0.8791
Total	0.9234	0.8879	0.9039

#### C. REGIONS OF INTERESTS

To explore the ability of 3D DANet to extract features related to the gender label, we also used deconvolution computations to show the corresponding regions of interests (ROIs), and the results are shown in Figure 4. From the figure, we can see that the sparsity of the detected ROIs increased when the depth of the neural network increased, which corresponds with our knowledge. When the network was shallower, the features learned from the network were rough and comprised nearly the entire brain. As the network deepened, the features



**FIGURE 4.** Deconvolutional mapping of the proposed method. The four columns correspond to the four convolutional layers with different sized kernels. The four columns represent the three orientation views of the ROIs. L, left; R, right.

**TABLE 4.** Detected regions of interests.

Regions	Regions
Cerebellum	Vermis
Temporal_Inf_L/R	Temporal_Mid_L/R
Temporal_Sup_R	ParaHippocampal_L/R
Frontal_Mid_L/R	Frontal_Sup_L/R
Cingulum_Post_L	Cingulum_Ant_L/R
Insula_L	Putamen_R
Precuneus_L/R	Cuneus_L/R
Occipital_Sup_L/R	Occipital_Mid_L/R
Angular_L/R	Calcarine_L/R
Supp_Motor_Area_L/R	Cingulum_Post_L
Parietal_Inf_L/R	Parietal_Sup_L/R
Frontal_Med_Orb_L/R	Frontal_Inf_Tri_L/R
Rolandic_Oper_R	Frontal_Sup_Medial_L

became sparse and some specific regions were detected. We used the AAL template to locate the detected regions, which are listed in Table 4.

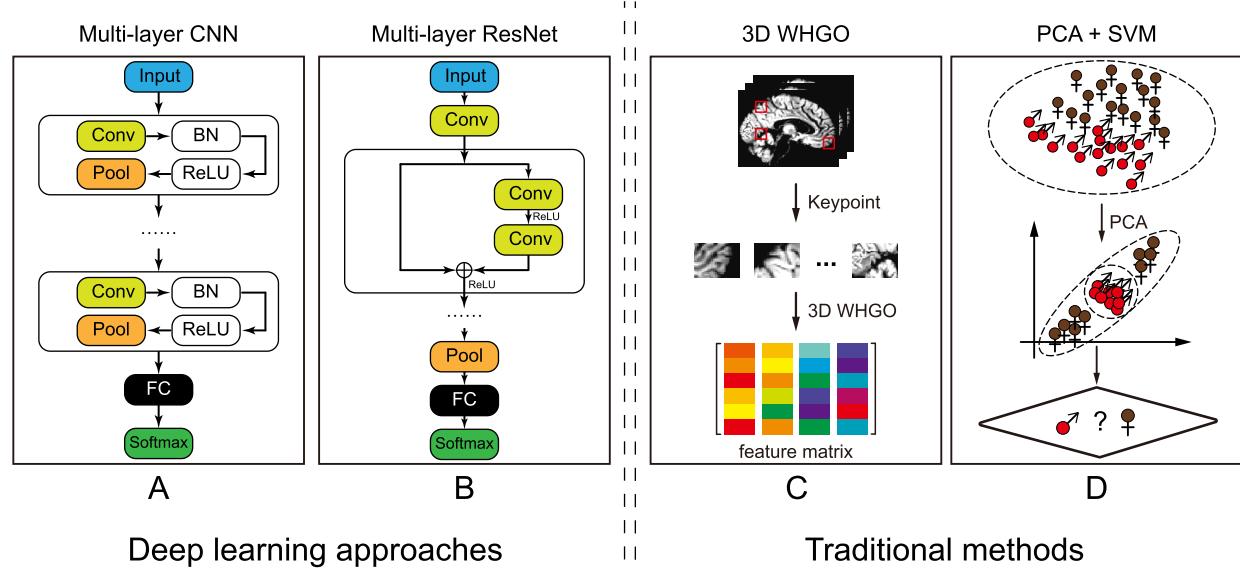
#### D. COMPARING OUR MODEL WITH OTHER METHODS

To compare the performance of our proposed model to other methods, we selected some typical traditional classification methods and deep learning approach architectures. In our previous study [29], we proposed a 3D descriptor (3D weighted histogram of gradient orientation, 3D WHGO) and

selected the PCA+SVM approach, a typical multivariable pattern analysis (MVPA) approach [33], as the comparison model. In this paper, we also compared our proposed method with the 3D WHGO method (Figure 5C) and the PCA+SVM approach (Figure 5D). For the deep learning approach, we compare our method with the typical 3D CNN (Figure 5A) and 3D ResNet (Figure 5B). Figure 5 illustrates the comparisons of the selected models.

We performed the same evaluation for each method on the total dataset and on each subset. Table 5 shows the detailed results of the model comparisons shown in Figure 5. The first column depicts the results of the 3D CNN model, and the second column shows those of the 3D ResNet approach, both of which are deep learning approaches whose number of layers correspond with the best performance. The third column depicts the results of our previous work [29], and the fourth column shows the results of the PCA+SVM approach [33], both of which are traditional methods selected for comparison with our method. Table 2 shows that the standard deviation randomly separating the datasets is sufficiently small, indicating that the influence of random separation can be ignored. Therefore, to save computational time, we tested the other models with the 5-fold strategy only once.

Table 5 shows that 3D ResNet performed better than 3D CNN, and the deep learning approaches had a stronger discriminative power than the traditional methods.



**FIGURE 5.** Illustration of some models selected for comparison with our proposed model. Deep learning approaches are shown on the left. The first approach illustrates a multiple layers 3D CNN, and the second depicts 3D ResNet model with multiple layers. Traditional methods are shown on the right. The first method is our previous 3D descriptor (3D WHGO), and the second represents PCA+SVM, a typical MVPA approach.

**TABLE 5.** Comparing results with those of other methods.

Dataset	Evaluation	Methods			
		3D CNN	3D ResNet	3D WHGO	PCA+SVM
HCP	SS	0.9719	0.9578	0.8112	0.7711
	SC	0.8767	0.9172	0.9037	0.8041
	GR	0.9202	0.9358	0.8615	0.7889
FCP	SS	0.9351	0.8918	0.6527	0.6818
	SC	0.8507	0.8926	0.9094	0.8389
	GR	0.8875	0.8922	0.7977	0.7703
GSP	SS	0.9609	0.9564	0.7624	0.7879
	SC	0.9238	0.9591	0.9348	0.8950
	GR	0.9395	0.9579	0.8618	0.8497
NKI	SS	0.9509	0.9363	0.7246	0.7647
	SC	0.8638	0.8482	0.9183	0.8482
	GR	0.9024	0.8872	0.8329	0.8113
SLIM	SS	0.9676	0.9393	0.8907	0.8381
	SC	0.9221	0.9448	0.9221	0.8571
	GR	0.9423	0.9423	0.9081	0.8486
CoRR	SS	0.9250	0.8891	0.7859	0.7500
	SC	0.7918	0.8533	0.8359	0.7965
	GR	0.8587	0.8713	0.8108	0.7732
Total	SS	0.9499	0.9267	0.7673	0.7607
	SC	0.8218	0.9092	0.9031	0.8423
	GR	0.9071	0.9171	0.8417	0.8054

The results correspond with our knowledge that deep learning approaches have a stronger ability to handle multi-center big data classifications than traditional methods. Our previous method performed better than the PCA+SVM approach, which we also demonstrated previously [29].

The comparison results of our proposed method with the other methods are shown in Figure 6. The average accuracies of each subset and the total dataset are shown as the area of the circle. The proposed method was the most accurate among all

the datasets, which further proves its discriminative power to classify neuroimaging data.

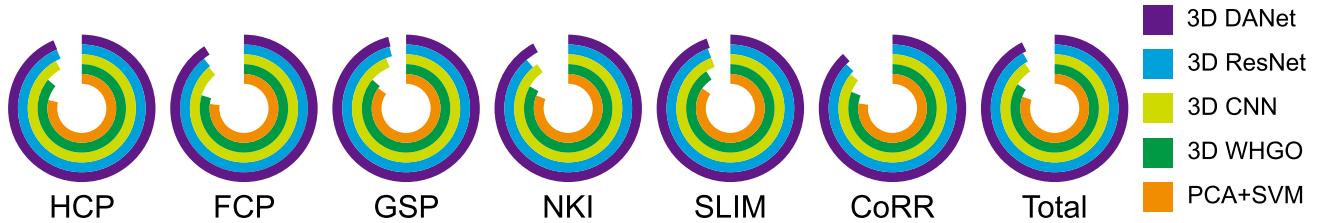
## IV. DISCUSSION

### A. ADVANTAGE OF DEEP LEARNING

With the increasing number of neuroimaging data samples, traditional analysis methods begin to lose efficacy. Some traditional statistical methods, such as principal component analysis (PCA), independent component analysis (ICA), and classifications based on feature description and feature selection approaches, such as scale invariant feature transform (SIFT) [34], 3D descriptors (e.g., 3D WHGO) [29], support vector machine (SVM) [35] and MVPA [33], [36], are incapable of handling large datasets. Because traditional methods must load all the training data to generate a model for classification, the cost of memory becomes prohibitive for completing the computation as the number of samples increases. In contrast to traditional methods, the deep learning approach has a natural advantage to address big data. The approach of partitioning large training datasets into small batches and using iterations to decrease loss-of-function by speeding up the GPU easily satisfies the need for computational speed. By performing iterations numerous times, the training dataset is utilized much more often in the deep learning approach much than in traditional methods, which is why deep learning performs better than traditional methods.

### B. DEPTH OF THE NEURAL NETWORK

Previously, researchers have proposed numerous types of deep neural network architectures [10]–[13] to classify millions of images with thousands of classes (for example,



**FIGURE 6.** Comparison of our proposed method with other methods. The different colors indicate the various methods, and the circle area indicates the accuracy; larger areas correspond with higher accuracies.

ImageNet; <http://www.image-net.org/>). Researchers have developed networks to comprise more than one hundred layers to obtain a satisfactory result. The previous classification was made to solve the multiple classes problem, and each class contained thousands of pictures. By contrast, in the neuroimaging classification field, the number of samples does not reach the level of natural scenes, and the number of neuroimaging data classes is also fewer. Furthermore, the neuroimaging dimension is much larger than the number of subjects. Due to the neuroimaging characteristics, especially the two-class classification problem described herein, the neural network should not be deeper. A shallower neural network may grasp the differences between classes, while deeper neural networks may learn the pattern uncorrelated with class labels. Based on these considerations, we herein proposed a shallower deep adding neural network to address neuroimaging classifications. He *et al.* [13] demonstrated that when deeper networks begin to converge, a degradation problem is exposed because the accuracy becomes saturated (unsurprisingly) and then rapidly degrades as the network depth increases. We believe that for the two-class neuroimaging classification problem, an effective neural network should not be deeper but rather more shallow, which is why the proposed 3D DANet has only one adding operation.

### C. KERNEL SELECTION AND SCALE COMBINATION

We choose four different sizes of convolution kernel sizes because the four kernels can yield gradient information in different orientations. The first three kernels, whose sizes were  $3 \times 1 \times 1$ ,  $1 \times 3 \times 1$  and  $1 \times 1 \times 3$ , can extract gradient information in three different orientations, while the final kernel ( $3 \times 3 \times 3$ ) can extract gradient information in a solid cube. The adding operation can combine the information from the different scales. On the top branch, 3D convolution operations are completed in the spatial space with a stride of 1, corresponding to the original 3D imaging scale. For the bottom branch, the 3D convolution operation is performed with a stride of 2 in the spatial domain, which is a smaller scale. Using two pooling layers, the results under the two scales can be altered to equal sizes and then added. Thus, the two scales are combined via the adding operation. Our previous work [29], [37] demonstrated that combining the information from different scales is effective for improving the classification performance.

### D. BIG DATA FROM MULTIPLE CENTERS

To the best of our knowledge, our work is the first to classify neuroimaging data on such a large scale and multicenter data (more than 6000 subjects from more than 60 sites) with such a high accuracy (over 92.5%). Furthermore, this is the first time that the deep learning approach has been used to solve the multicenter classification problem. With the development of neuroimaging technology, increasing amounts of MRI data are being acquired, creating large differences between scanners exist and varying image qualities. These difficulties are problematic for traditional classification methods based on single sites and small sample sizes. Our method utilizes deep learning technology to address the big data classification problem on multicenter datasets and yields satisfactory results, demonstrating that the proposed approach can handle multicenter classifications. We believe that analyses based on big data from multiple centers will increase in the future.

### E. REGIONS OF INTERESTS

We used a standard template without a gender label to obtain the deconvolutional results, which are equivalent to the features learned from the neural network. The spatial ROIs shown in Figure 4 demonstrate that the features learned from the neural network are located throughout almost entire brain. However, as the network deepens, the features become sparser, which corresponds with our knowledge. The features from the first convolutional layer are rough and comprise nearly the entire brain. However, in the fourth convolutional layer, some special regions to group differences are elucidated. Figure 4 shows that some features are located on the edge of the brain, indicating that the shape of the human brain varies, especially between females and males. Table 4 shows that some regions corresponding to group differences have been previously reported, such as the vermis [38], occipital lobe [29], cerebellum [33], cingulum cortex [39], precuneus [40] and rolandic regions [29] are reported previously. Zhang *et al.* [41] utilized functional connectivity to elucidate differences between gender, demonstrating that these differences do exist. Similar results from other previous studies demonstrated that the proposed method can potentially be applied in clinical practices and to reflect the special regions of some brain diseases.

## F. TRANSFERABILITY

We also tested the cross-site classification of the proposed method. Herein, we extended the standard  $k$ -fold cross-validation strategy and used the LOSO-CV strategy, the results of which are shown in Table 3. For the validation of each time, one center subset was left out, and the other subsets were used as the training datasets; the test subset thus had no overlap with the training datasets. All the accuracies remained high, which demonstrated that the proposed method is transferable and can effectively handle multicenter classifications. Reasonably, all the accuracies obtained using the LOSO-CV strategy were slightly lower than those obtained using the 5-fold cross-validation strategy because the distributions of the test subsets differed from those of the training datasets.

The code for our proposed model was written using TensorFlow, an open source library that is well suited for mobile deep learning, and the application thus removes the platform limitation. In the future, we plan to build a system in which the model is trained in the cloud based on big data and test the practical data on the clinical diagnosis made on mobile terminal. This may lead to significant developments in artificial intelligence (AI) and a clinical revolution.

## V. CONCLUSION

With the development of neuroimaging technology, increasing amounts of MRI data are being acquired. Herein, we proposed a 3D deep adding neural network (3D DANet) to address the brain imaging classification problem on multicenter datasets. Our method utilizes multiple convolutional kernels of varying sizes to extract gradient information in different orientations and combines spatial features at two scales via the adding operation. A standard 5-fold cross-validation strategy was used to test the performance of the proposed method. We collected a large dataset comprising 6008 healthy subjects from more than sixty scanners at six centers. A high accuracy (over 92%) in distinguish females from males was observed, which demonstrates that the proposed method can handle big data classification from multiple centers. Compared with some traditional methods and architectures of the deep learning approach, the proposed method performed better, highlighting its advantages for addressing two-class neuroimaging data classifications. With its improved performance for neuroimaging classifications and transferable program codes, the proposed method can potentially be used for intelligent medical treatment strategies and clinical practices based on mobile terminal.

## REFERENCES

- [1] F. Ge et al., "Denser growing fiber connections induce 3-hinge gyral folding," *Cerebral Cortex*, vol. 28, no. 3, pp. 1–12, 2017.
- [2] A. C. Altamura et al., "The impact of psychosis on brain anatomy in bipolar disorder: A structural MRI study," *J. Affect. Disorders*, vol. 233, pp. 100–109, Jun. 2018.
- [3] Y.-K. Kim and K.-S. Na, "Application of machine learning classification for structural brain MRI in mood disorders: Critical review from a clinical perspective," *Prog. Neuropsychopharmacol., Biol. Psychiatry*, vol. 80, pp. 71–80, Jan. 2018.
- [4] C. Tas et al., "Discriminating schizophrenia and schizo-obsessive disorder: A structural MRI study combining VBM and machine learning methods," *Neural Comput. Appl.*, vol. 29, no. 2, pp. 377–387, 2018.
- [5] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, May 2015.
- [6] G. Litjens et al., "A survey on deep learning in medical image analysis," *Med. Image Anal.*, vol. 42, pp. 60–88, Dec. 2017.
- [7] Q. Zhang, L. T. Yang, Z. Chen, and P. Li, "A survey on deep learning for big data," *Inf. Fusion*, vol. 42, pp. 146–157, Jul. 2018.
- [8] Q. Hou, M.-M. Cheng, X. Hu, A. Borji, Z. Tu, and P. H. Torr, "Deeply supervised salient object detection with short connections," in *Proc. Comput. Vis. Pattern Recognit.*, Jul. 2017, pp. 5300–5309.
- [9] O. Russakovsky et al., "ImageNet large scale visual recognition challenge," *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, Dec. 2015.
- [10] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 6, no. 6, pp. 84–90, 2017.
- [11] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *Proc. Int. Conf. Learn. Represent.*, 2015.
- [12] C. Szegedy et al., "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 1–9.
- [13] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.
- [14] I. J. Goodfellow et al., "Generative adversarial nets," in *Proc. Neural Inf. Process. Syst.*, 2014, pp. 2672–2680.
- [15] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [16] A. Esteva et al., "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. 542, pp. 115–118, Feb. 2017.
- [17] Y. Wei et al., "HCP: A flexible CNN framework for multi-label image classification," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 9, pp. 1901–1907, Jun. 2015.
- [18] J. Dolz, C. Desrosiers, and I. B. Ayed, "3D fully convolutional networks for subcortical segmentation in MRI: A large-scale study," *NeuroImage*, vol. 170, pp. 456–470, Apr. 2018.
- [19] J. Kawahara et al., "BrainNetCNN: Convolutional neural networks for brain networks; towards predicting neurodevelopment," *NeuroImage*, vol. 146, pp. 1038–1049, Feb. 2017.
- [20] C. Wachinger, M. Reuter, and T. Klein, "DeepNAT: Deep convolutional neural network for segmenting neuroanatomy," *NeuroImage*, vol. 170, pp. 434–445, Apr. 2017.
- [21] J. H. Cole et al., "Predicting brain age with deep learning from raw imaging data results in a reliable and heritable biomarker," *NeuroImage*, vol. 163, pp. 115–124, Dec. 2017.
- [22] H. Chen, Q. Dou, L. Yu, J. Qin, and P.-A. Heng, "VoxResNet: Deep voxelwise residual networks for brain segmentation from 3D MR images," *NeuroImage*, vol. 170, pp. 446–455, Apr. 2017.
- [23] S. I. Ktena et al., "Metric learning with spectral graph convolutions on brain connectivity networks," *NeuroImage*, vol. 169, pp. 431–442, Apr. 2018.
- [24] L. Zou, J. Zheng, C. Miao, M. J. McKeown, and Z. J. Wang, "3D CNN based automatic diagnosis of attention deficit hyperactivity disorder using functional and structural MRI," *IEEE Access*, vol. 5, pp. 23626–23636, 2017.
- [25] D. C. Van Essen et al., "The human connectome project: A data acquisition perspective," *NeuroImage*, vol. 62, no. 4, pp. 2222–2231, 2012.
- [26] M. R. Hodge et al., "ConnectomeDB—Sharing human brain connectivity data," *NeuroImage*, vol. 124, pp. 1102–1107, Jan. 2016.
- [27] A. J. Holmes et al., "Brain genomics superstruct project initial data release with structural, functional, and behavioral measures," *Nature ScientificData*, vol. 2, Jul. 2015, Art. no. 150031.
- [28] K. B. Nooner et al., "The NKI-Rockland sample: A model for accelerating the pace of discovery science in psychiatry," *Frontier Neurosci.*, vol. 6, p. 152, Oct. 2012.
- [29] L. Yuan, F. Chen, L. Zeng, L. Wang, and D. Hu, "Gender identification of human brain image with a novel 3D descriptor," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, vol. 15, no. 2, pp. 551–561, Mar./Apr. 2018.
- [30] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998.

- [31] S. Sarraf, D. D. DeSouza, J. Anderson, G. Tofighi, and For the Alzheimer's Disease Neuroimaging Initiative, "DeepAD: Alzheimer's disease classification via deep convolutional neural networks using MRI and fMRI," *BioRxiv*, 2016.
- [32] J. Kleesiek *et al.*, "Deep MRI brain extraction: A 3D convolutional neural network for skull stripping," *NeuroImage*, vol. 129, pp. 460–469, Apr. 2016.
- [33] L. Wang, H. Shen, F. Tang, Y. Zang, and D. Hu, "Combined structural and resting-state functional MRI analysis of sexual dimorphism in the young adult human brain: An MVPA approach," *NeuroImage*, vol. 61, no. 4, pp. 931–940, 2012.
- [34] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, 2004.
- [35] A. Ben-Hur, C. S. Ong, S. Sonnenburg, B. Scholkopf, and G. Ratsch, "Support vector machines and kernels for computational biology," *PLOS Comput. Biol.*, vol. 4, no. 10, pp. 1–10, 2008.
- [36] A. Mahmoudi, S. Takerkart, F. Regragui, D. Boussaoud, and A. Brovelli, "Multivoxel pattern analysis for fMRI data: A review," *Comput. Math. Methods Med.*, 2012.
- [37] L. Yuan, F. Chen, L. Zhou, and D. Hu, "Improve scene classification by using feature and kernel combination," *Neurocomputing*, vol. 170, pp. 213–220, Dec. 2015.
- [38] N. Raz, J. H. Dupuis, S. D. Briggs, C. McGavran, and J. D. Acker, "Differential effects of age and sex on the cerebellar hemispheres and the vermis: A prospective MR study," *Amer. J. Neuroradiol.*, vol. 19, no. 1, pp. 65–71, 1998.
- [39] K. Menzler *et al.*, "Men and women are different: Diffusion tensor imaging reveals sexual dimorphism in the microstructure of the thalamus, corpus callosum and cingulum," *NeuroImage*, vol. 54, no. 4, pp. 2557–2562, 2011.
- [40] A. N. V. Ruigrok *et al.*, "A meta-analysis of sex differences in human brain structure," *Neurosci. Biobehavioral Rev.*, vol. 39, pp. 34–50, Feb. 2014.
- [41] C. Zhang, C. C. Dougherty, S. A. Baum, T. White, and A. M. Michael, "Functional connectivity predicts gender: Evidence for gender differences in resting brain connectivity," *Hum. Brain Mapping*, vol. 39, no. 4, pp. 1765–1776, 2017.



**LIN YUAN** was born in Liaoning, China, in 1988. He received the B.Sc., M.Sc., and Ph.D. degrees from the National University of Defense Technology, China, in 2011, 2013, and 2018, respectively. From 2015 to 2016, he was a Visiting Ph.D. Student with the Department of Computer Science, University of Georgia. His research interests include computer vision, magnetic resonance imaging analysis, deep learning algorithm, pattern recognition, and machine learning.



**XUE WEI** was born in Gansu, China, in 1993. She received the B.Sc. degree from Sichuan University, China, in 2015. She is currently pursuing the master's degree with the National University of Defense Technology. Her research interests include blind source separation, deep learning algorithm, machine learning, and optical imaging data analysis.



**HUI SHEN** received the B.Sc. degree in automation and the Ph.D. degree in pattern recognition and intelligent system from the National University of Defense Technology (NUDT), Hunan, China, in 1997 and 2003, respectively. In 2013, he was a Visiting Scholar with Emory University, USA. Since 2003, he has been a Lecturer and an Associate Professor with the Automation Department, NUDT. His research interests include automation theory and application, neuroimaging analysis, and machine learning.



**LING-LI ZENG** was born in Jiangxi, China, in 1984. He received the B.Sc., M.Sc., and Ph.D. degrees from the National University of Defense Technology, China, in 2007, 2009, and 2014, respectively. From 2012 to 2013, he was a Visiting Ph.D. Student with the Harvard Medical School & Massachusetts General Hospital. Since 2014, he has been with the National University of Defense Technology. His research interests include cognitive neuroscience, image processing, and pattern recognition in neuroimaging.



**DEWEN HU** (SM'06) was born in Yueyang, Hunan, China, in 1963. He received the B.Sc. and M.Sc. degrees from Xi'an Jiaotong University, China, in 1983 and 1986, respectively, and the Ph.D. degree from the National University of Defense Technology in 1999. Since 1986, he has been with the National University of Defense Technology. From 1995 to 1996, he was a Visiting Scholar with the University of Sheffield, U.K. He was promoted to Professor in 1996. His research interests include image processing, system identification and control, neural networks, and cognitive science. He is an Action Editor of *Neural Networks* and an Associate Editor of the IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS: SYSTEMS.