Single Volume Image Generator and Deep Learning-based ASD Classification

Md Rishad Ahmed, Yuan Zhang*, *Senior Member, IEEE,* Yi Liu, and Hongen Liao, *Senior Member, IEEE*

Abstract—Autism spectrum disorder (ASD) is an intricate neuropsychiatric brain disorder characterized by social deficits and repetitive behaviors. Deep learning approaches have been applied in clinical or behavioral identification of ASD; most erstwhile models are inadequate in their capacity to exploit the data richness. On the other hand, classification techniques often solely rely on region-based summary and/or functional connectivity analysis of functional magnetic resonance imaging (fMRI). Besides, biomedical data modeling to analyze big data related to ASD is still perplexing due to its complexity and heterogeneity. Single volume image consideration has not been previously investigated in classification purposes. By deeming these challenges, in this work, firstly, we design an image generator to generate single volume brain images from the whole-brain image by considering the voxel time point of each subject separately. Then, to classify ASD and typical control participants, we evaluate four deep learning approaches with their corresponding ensemble classifiers comprising one amended Convolutional Neural Network (CNN). Finally, to check out the data variability, we apply the proposed CNN classifier with leave-one-site-out 5-fold cross-validation across the sites and validate our findings by comparing with literature reports. We showcase our approach on large-scale multi-site brain imaging dataset (ABIDE) by considering four preprocessing pipelines, which outperforms the state-of-the-art methods. Hence, it is robust and consistent.

Index Terms—Biomedical Data Modeling, Image Generator, Convolutional Neural Network (CNN), Autism Spectrum Disorder (ASD), fMRI, ABIDE.

I. INTRODUCTION

AUTISM Spectrum Disorder (ASD) is the disturbance of the structure and functioning of the brain that cause different abnormalities such as communication difficulties, social deficits, repetitive behaviors and cognitive delays, and nonsocial features such as restricted and stereotyped behaviors, all of which have a significant impact on adaptive functioning [1]–[3]. As investigated by the Centre for Disease, Control and Prevention in the United States, the estimated ASD prevalence of 1% or higher (1 subject in 59) and it is augmented dramatically in the last decades [4]. Therefore, finding a precise biological marker to extrapolate the underlying roots of ASD pathologies is indispensable for applying effective treatment in ASD diagnosis.

*Corresponding Author: Yuan Zhang
Md Rishad Ahmed and Yuan Zhang are with the College of Electronic and Information Engineering, Southwest University, Chongqing 400715, China. e-mail: yuanzhang@swu.edu.cn.
Yi Liu is with the Department of Respiratory Medicine, Civil Aviation General Hospital, Beijing 100123, China. email: cassieliu@126.com.
Hongen Liao is with the Department of Biomedical Engineering, School of Medicine, Tsinghua University, Beijing 100084, China. e-mail: liao@tsinghua.edu.cn.

One of the significant challenges in brain disorder research is to replicate the findings through larger datasets that can reflect the heterogeneity of clinical populations. Functional magnetic resonance imaging (fMRI) has been extensively considered to perceive functional abnormalities of ASD patients which can characterize the neural pathways [5], [6]. Functional connectivity analysis has produced deep insights to see the brain abnormality connectomes between ASD and typical control (TC) either at individual or at group level characteristics. Recently, most of the machine learning techniques to study functional connectivity data rely on hand-engineered feature extraction, such as the correlation between region of interests (ROIs) and topological measurements of modularity, clustering based classification [7], segregation or integration [8]. On the other hand, brain ROIs provide the structural substrates for measuring connectivities within the individual brain and functional activation patterns of the brain. It is a common approach to analyze ASD individuals based on the expert's defined brain parcellation or data-driven strategies such as dictionary learning, clustering, and ICA (independent component analysis) [9], [10]. Both the expert-defined and data-driven ROIs strategy have several complaints such as standardization, arbitrary decision, and selection of the regions exhibiting proficient information [11]. The data-driven ROIs strategy can be biased in selecting the regions showing considerable variability across the subject, which influence the results [12]. Hence, an alternative tool or strategy is essential to overcome the above complaints and generate the volumetric brain images to see the activated regions. A single volume image generator not only generates whole brain volume images but also has the arbitrary nature of a chosen brain region scheme, which ensures the coverage of the entire brain regions of each subject.

Machine learning (ML) such as SVM (support vector machines) has been widely adopted to classify and exploit individual variation in functional connectivity of ASD [13], [14]. Recently, deep learning models with neuroimaging modalities have been effective in identifying brain disorders such as ASD, Alzheimer Disease (AD) [15], [16]. With the rapid advancement of deep learning approaches for brain disorder diagnosis, convolutional neural networks (CNN) becomes the most popular method for ASD classification [17]. However, most deep learning approaches have been focused on functional connectivity or ROIs analysis, time-series data analysis, or temporal/spatial information of fMRI [18], [19]. They also lack model transparency, i.e., how the model secures its interpretability for clinical applications, as most of the deep neural networks are not easily interpretable. The choice of the potential classification algorithm is another exception in the

connectome-based analysis of ASD. A few years ago, some of the deep learning-based ASD classification models had been focused on simple linear predictive techniques using a vectored connectivity correlation matrix [20], [21]. Additionally, ASD big data handling using deep learning techniques is still thought-provoking due to the lack of potential data mining and investigating methods from the heterogeneous, complex, and dynamic nature data to diagnose this brain disorder. Due to the heterogeneity, etiology, and severity of ASD, a more professed methodology is required to forecast and analyze the behavior and functionality of each subject. Hence motivated by the above challenges, here we focus on designing a new image generator that can generate single volume images from whole-brain fMRI and propose two novel classification architectures for classifying ASD and typical controls. The single volume brain image is the visualization of the brain regions along with specific direction and slice number, allowing the real-time voxel-periods separation of the raw fMRI data. Nevertheless, ROIs only define the interested brain regions depending on the pre-selected slice number.

The main contributions of this study are as follows:

- To the best of our knowledge for the first time, we design a single volume image generator that can produce 2D three-channel images from a functional magnetic resonance 4D NIFTI image. The main advantage of the single volume image generator is that generated 2D images represent activated brain regions for each voxel time point of the patients. It also visualizes the brain regions in axial, sagittal, and coronal axes in the form of glass brain and stat_map images.
- We incorporate four deep learning approaches with our improved CNN model to classify ASD and typical controls using generated images as the input. The advantages of our model include leveraging the voxel-2D structures of rsfMRI without possessing too many model parameters and easily interpret the complex, heterogeneous data which can be used in combination with other tools supporting clinicians to diagnose ASD with more precision.
- We propose a novel deep ensemble learning framework based on the improved CNN and the benchmark approaches to classify ASD using features extracted by VGG16 from the glass brain and stat map images. The proposed ensemble model can integrate two different types of generated images simultaneously utilizing one ensemble learning classifier individually. Thus, it overcomes the limitations of traditional machine learning models for ASD classification that often rely on ROI definitions.
- Finally, to evaluate the classifier performance and check out the data variability, we apply the proposed CNN classifier with leave-one-site-out 5-fold cross-validation across the sites and validate our findings by comparing with literature reports.

The proposed approaches with a combined loss function and generated single volume images, establish a novel benchmark model for ASD detection on the ABIDE (Autism Brain Imaging Data Exchange) database.

The leftover of the paper is organized as follows: Section II discusses the related works; Section III covers broad methodological explanation including single volume image generator and propose deep learning models. In section IV, the experiment and discussion of this method including dataset are presented and finally conclusion is drawn in Section V.

II. RELATED WORKS

The amalgamation of machine learning methods and brain imaging data permit the classification of ASD, which can assuage the significant suffering and provide safety for the patient's daily well-being. Studies on ASD classification using different imaging modalities and their analysis approaches, specifically deep learning techniques, are discussed in this section.

Study on the functional connectivity of brain networks is a sturdy utensil to understand the neurological bases of a diversity of brain disorders such as autism [22]. In the work [23], Abraham et al. employed resting-state fMRI to extract functionally-defined brain areas and support vector classifier (SVC) to compare connectivity between ASD and typical control. They considered 871 subjects from ABIDE dataset for connectome-based prediction and got 67% accuracy. Guo et al. considered multiple stacked auto-encoder (SAE) as a feature selection method from whole-brain FCP obtained by Pearson correlation of ROIs [24]. Using only UM (University of Michigan) data site, they got a classification accuracy of 86.36%. In [25], using only CCS (connectome computation system) pipeline without global signal regression data and LSTMs (long short-term memory) method for classification of individuals with ASD, Dvornek et al. achieved 68.5% accuracy.

On the other hand, time-series for several sets of regions-of-interests (ROIs) also have the potentiality to classify and see brain network connections of ASD. ROIs are usually computed using a predefined atlas or a parcellation scheme on anatomical features, functional activations, and connectivity patterns of brain [10], [26]. Dvornek et al. incorporated phenotypic data with rs-fMRI into a single LSTM based model for classifying ASD and achieved an accuracy of 70.1% [27]. They employed CCS pipeline data without global signal regression and cross-validation framework. With the development of deep learning model specifically, Convolutional Neural Networks (CNNs) have found abundant applications on 2D and 3D images which can exploit image intensities and pixel grid to decipher image segmentation and classification problems [28], [29]. Zhao et al. also evaluated an effective 3D CNN to bridge the gap between spatial overlap patterns and functional brain networks for characterizing and differentiating ASD from healthy controls [30]. They incorporated only 200 (ASD-100, HC-100) rs-fMRI data from the ABIDE dataset. The average 10-fold cross-validation accuracy was 70.5%, using a spatial overlap pattern of 144 intrinsic connectivity networks (ICN) of each subject. The CNN weight training was performed by the stochastic gradient descent (SGD) optimizer with momentum, and multinomial negative log-likelihood was used as an objective loss function. Graph convolutional neural networks (G-CNN) is another way

to identify brain patterns that can act as a neuropathological biomarker. Ktena et al. highlighted the potential of convolutional neural network models for connectome-based classification [31]. Using CPAC (configurable pipeline for the analysis of connectomes) preprocessing pipeline and Harvard-Oxford atlases in [32], Anirudh et al. achieved 70.86% classification accuracy based on the G-CNN and ensemble learning.

Table I presents the summarization of recent trending technologies in the classification of ASD. From the analysis of Table I, we notice that most of the recent research works on ASD classification used functional connectivity (FC) or ROIs data for classification purposes. Besides, the majority work belongs to either one pipeline or one site or one atlas images for classification, which is not plentiful to answer some research and clinical questions. Biomedical big data are generally preferable to improve the dependability (i.e., separate feature extraction) and solid contributions of research for the treatment of brain disorders like ASD [33], [34]. By considering these challenges, we design an image generator to generate the single volume brain images from the whole brain image using four preprocessing pipelines mentioned by ABIDE. Furthermore, we observe that extracting features from the images using deep learning models also enhance classification performance. Using the VGG16 model as a feature extractor, we also propose a novel deep ensemble learning model that can be treated as a new benchmark approach for ASD classification compared with the recent literature reports.

TABLE I: A brief summarization of the previous machine/deep learning techniques in ASD classification.

Ref.	Year	Purposes	Number of Subjects	Patterns	Methods	Accuracy
[23]	2016	Prediction	871	ROIs	SVM	67%
[24]	2017	Classification	110	FC	DNN	86.36%
[25]	2017	Identification	1112	ROIs	LSTMs	68.5%,
[27]	2018	Classification	1100	ROIs	LSTM	70.1%
[30]	2018	Differentiation	200	ICNs	3D CNN	70.5%
[31]	2017	Classification	871	ROIs	GCNN	62.9%
[32]	2019	Classification	872	ROIs	GCNN	70.86%
[35]	2019	Prediction	1112	Atlases	Ensemble CNN	72.3%
[36]	2019	Diagnosis	279	ROIs	Ensemble MVTC	72.6%
[37]	2018	Identification	1035	FC/ROIs	DE, SVM	70%
[38]	2019	Detection	1035	FC/ROIs	ASD-DiagNet	82%

III. METHODOLOGY

A. Data Preprocessing Pipelines

The ABIDE data instigated by the various preprocessing pipelines are analogous, and there is no consensus on the superlative methods [39]. The foremost difference between preprocessing pipelines is the precise algorithms and parameters used for each of the step and software implementations. Rather than being doctrinaire and preferring a single strategy, we analyze four different preprocessing approaches. The principal advantages of the four strategies are that it overcomes the controversies surrounding by bandpass filtering and global signal regression by combining with and without filtering and with and without global signal adjustment [40]. Selected strategies are Connectome Computation System (CCS), Configurable Pipeline for the Analysis of Connectomes (CPAC), Data Processing Assistant for Resting-State fMRI (DPARSF), and Neuroimaging Analysis Kit (NIAK). Table II arranges for a summary of the distinct preprocessing steps and how they vary across pipelines. As mentioned by ABIDE, functional processing was performed using only four strategies which we covered in this study; however other strategies were used for structural preprocessing and calculation of cortical measures. In this work, data are selected from the filt_global preprocessing stratagem, which is band-pass filtered (0.01-0.1 Hz) and spatially registered using a nonlinear method to MNI152 template space for each of four pipelines.

TABLE II: Overview of the basic parameters and steps of four different functional preprocessing stratagems.

	Steps	CCS	CPAC	DPARSF	NIAK
Basic Processing	Slice timing Correction	Yes	Yes	Yes	No
	Motion Re-alignment	Yes	Yes	Yes	Yes
	Intensity Normalization	Yes	Yes	No	Non-uniformity
Nuisance Signal	Motion	24-param	24-param	24-param	scrubbing and 1st principal component
Removal (Regressor)	Tissue signals	mean WM and CSF signals	CompCor (5 PCs)	mean WM and CSF signals	mean WM and CSF signals
	Motion realignment	Yes	Yes	Yes	Yes
	Low-frequency drifts	linear and quadratic trends	linear and quadratic trends	linear and quadratic trends	discrete cosine basis with a 0.01 Hz high-pass cut-off

B. Single Volume Image Generator

As we know, a 3D fMRI is the voxel image containing only one brain volume. Whereas, a 4D fMRI is a series of concatenated brain volumes over repeated time and the time is the 4th dimension representing the number of brain volume. Considering the number of brain volumes, it can give the images of activated brain regions during spontaneously fMRI acquisition, which is called a single volume image [41], [42]. The functional activation maps impact the properties of repeated fMRI scans, for example, the assessment of
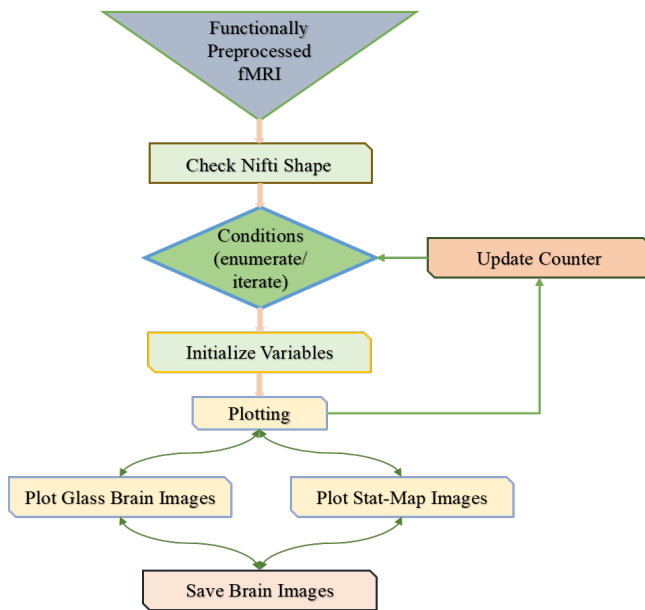
Fig. 1: Graphical representation of the proposed single volume image generator.

the relations between the symptom's intervention and brain activation patterns [43]. Therefore, there needs a tool to generate the images between the voxel periods for scrutinizing the activated brain regions. Rather than traditional analysis of functional connectivity or brain ROIs, in this work, we design an image generator to produce single volume brain images from the preprocessed whole-brain functional image. The single-volume image depicts the 2D visualization of the brain activity by considering each voxel time point. Fig. 1 represents the flow diagram to generate the single volume images by predefined displaying mode. The working principle of the single volume image generator has several folds: firstly, the generator checks out the shape of each input Nifti files, the input shape might be the 4D fMRI images. Secondly, the generator main body, which checks out the conditions to draw output depending on the enumeration and iteration counter. The enumeration and iteration segment counts the number of voxel time points during image acquisition. Then we set the corresponding parameters to demonstrate the real-time brain activations. Finally, generator plots and saves the brain images into two types of volumetric images for each ASD and TC individuals by counting the whole voxel time points for each subject. The corresponding parameters to display and save the single volume images from 4D multiple brain volumes image are shown in Table III.

In Table III, display mode selects the specific direction of the cuts following x as sagittal, y as coronal and z as axial view of the brain. On the other hand, a threshold is either a numeric value or none; none means plotted images are not threshold, and an absolute numeric value means it plots the threshold images where the value below the threshold is transparent. We choose two different threshold values: 3 and 5 in our experiment for glass brain and stat map functions, respectively. The parameter cut cords specify the number of

slices to visualize the brain images along the specific direction. Colorbar shows a vertical color bar to the right of the current axes of the plotting images, which stands "false" in our experiment. The detailed description of the plotted two types of volumetric images, and their choosing reasons are explained below. The detailed description of the plotted two types of volumetric images, and their choosing reasons are explained below.

TABLE III: Parameters weighed while plotting the volumetric images using the proposed image generator.

Data Pipelines	Plotting Functions	Parameters			
		Display Mode	Threshold	Cut Coords	Colorbar
CCS	Glass Brain	Axial (z)	3	1	False
	Stat_Map	Axial (z)	5	1	False
CPAC	Glass Brain	sagittal (x)	3	1	False
	Stat_Map	sagittal (x)	5	1	False
DPARSF	Glass Brain	Coronal (y)	3	1	False
	Stat_Map	Coronal (y)	5	1	False
NIAK	Glass Brain	Axial (z)	3	1	False
	Stat_Map	Axial (z)	5	1	False

Glass Brain and Stat_Map Images: All the neuroimaging is the part of the brain mapping. The glass brain is a 3D brain visualization that displays real-time source activity and connectivity between brain areas [44]. On the other hand, stat_map is the full name of the statistical images which plot cuts of an ROI/mask image. We prefer the glass brain and stat_map displaying mode for a single volume image because of its power of projecting high-resolution 3D model of an individual's brain, skull, and scalp tissue. Another fundamental discrepancy between general brain mapping and our generator images is the map projection. In a general brain mapping, it considers the specific brain regions (ROIs) or time series scheme to see the brain connectome. However, our proposed generator estimates the number of voxel time points for each 4D image to plot the real-time activated brain regions. The plotted images were in MNI space for all the considered pipelines to work image function accurately.

C. Proposed CNN-based Classification Model

Fig. 2 depicts the overall deep learning architecture specifically based on CNN for classification of ASD. Convolutional neural network, as a portion of the neural network is widespread for task interrelated to image classification and segmentation. CNN debilitates the limitations of traditional neural networks through a local connection, sharing weights and sampling. There exist two rudimentary operations in CNN: convolution through a kernel (weights and bias to convolve input feature map) and subsequent sampling of the convolved feature map.
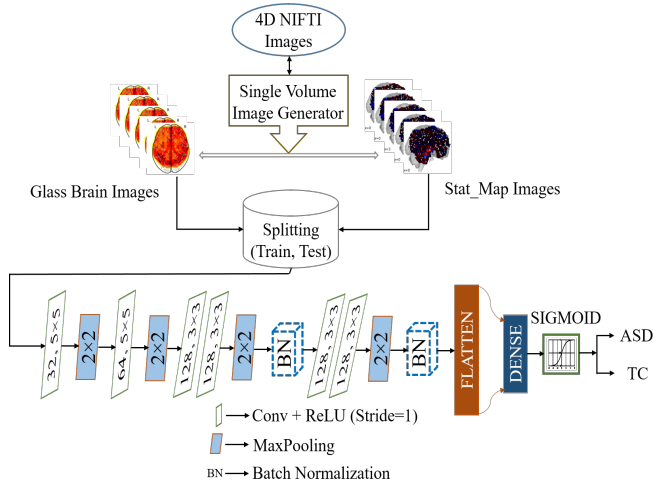
Fig. 2: Proposed CNN architecture for classifying ASD vs. TC.

From Fig. 2, the enhanced CNN architecture has six convolutional layers, four max-pooling layers, two batch normalization layers, one flatten and dropout layer with ratio 0.25, and two densely connected layers followed by a sigmoid activation function. The number of kernels and layer types are signified in each box. To conciliate the model slow training due to the layer parameter changes, we introduce the batch normalization [45] layer, which provides a much higher learning rate and flout initialization.

Loss Function: We contemplate the class-wise cross entropy loss as given below:

$$loss_i = -y_i \log(p_i) + (1 - y_i) \log(1 - p_i) \qquad (1)$$

where, y_i is the label (binary indicator) and p_i is the predicted probability of the label.

After that, we measure the higher weight (w_i) to give it in losses incurred on the training data samples.

$$w_i = \frac{\sum_{j=1}^{N} s_j}{s_i} \qquad (2)$$

where N is the total number of training samples, s_j counts the whole samples from $j = 1.....N$, and s_i is the number of per class samples. Hence the weighted binary cross entropy loss can be calculated from $equation$ (1) and (2) as,

$$loss_w = -\frac{1}{M} \sum_{i=1}^{M} w_i \times loss_i \qquad (3)$$

where M is the number of classes and $loss_w$ is the weighted loss.

Overfitting usually happens in a complex model, which is a common problem in deep learning applications. Regularization is a procedure for introducing supplementary information to prevent overfitting. Therefore, to solve the overfitting problem as well as penalize the sum of the absolute value of weights, we add L_1 regularization with $equation$ (3) to define a new loss function as given below:

$$loss_{com} = loss_w + L_1 \qquad (4)$$

where L_1 regularization is defined as,

$$L_1 = \lambda \sum_{n=1}^{k} |w_n| \qquad (5)$$

where k is the number of independent features and w is the model weights. The regularization parameter $\lambda > 0$ is manually tuned which has the effect of pushing w towards 0. After subsequent tuning during the experiment, we select the best fit value of $\lambda = 0.01$ through monitoring the model performances using Keras Callbacks function. By using EarlyStopping of Callbacks function, training could be terminated if we see the possible sign of overfitting or underfitting problem. In short, L_1 regularization is the process of regularizing the parameters that shrink the parameters to zero by reducing the number of features in the model altogether, thus avoiding the risk of overfitting.

Benchmark CNN Approaches: To analyze the generated single volume images, we executed Keras based pretrained benchmark CNN approaches as follows: DenseNet [46], ResNet [47], Xception [48], and Inception V3 [49]. For classification of ASD, the output node-set to a sigmoid function, and binary cross-entropy loss was adopted. The models are trained with a mini-batch size of 64, ADAM was used as the optimizer with learning rate 0.0004, $beta\_1 = 0.9$, and $beta\_2 = 0.999$.

D. Proposed Deep Ensemble Learning Classifier

Ensemble learning framework sometime overcomes the limitations of traditional deep learning models which often rely on ROIs based summary statistics and linear models for ASD classification [35]. In our experiments, we reconnoitered four ensemble learning strategies by combining benchmark approaches with improved CNN model. Each ensemble model was trained and tested using two types of generated images from each pipeline. Fig. 3 shows the proposed deep ensemble learning classifier's architectural overview for ASD classification, where n represents the number of classifier. Here, we adopted VGG16 [50] model, which is a Keras-based pretrained paradigm utilizing weights from the ImageNet dataset to extract features from both generated images. The feature extraction part comprises of captivating the convolutional base (consisted of only convolutional and pooling layers) of a formerly trained VGG16 model without a densely connected layer. The feature maps of a convolutional base are the presence maps of generic notions which is more reusable. Each ensemble classifier averages the predictions of the models for a specific pipeline using each one of the two types of images. The final evaluation metrics are calculated as the two-stage arithmetic mean of the specific binary class predictions for classification. The mathematics behind this is given below:

$$Score = (E_g + E_s)/2 = (Classifier(n) + Classifier(n))/2 \qquad (6)$$

where E_g and E_s denote the glass brain and stat_map images respectively and, $Classifier(n)$ is the base classifier. The base classifiers are defined as $ensemble$ $classifier$ 1 (DenseNet + Proposed CNN), $ensemble$ $classifier$ 2

(ResNet + Proposed CNN), $ensemble\ classifier\ 3$ (Xception + Proposed CNN) and $ensemble\ classifier\ 4$ (Inception V3 + Proposed CNN), respectively. The number of base classifiers is denoted as the number of ensemble classifiers, such as $n = 2 = ensemble\ classifier\ 2$, and so on. We train each ensemble classifier separately for the two types of images. For instance, we have two types (glass brain and stat map) of images for CCS pipelines; hence we train the $ensemble\ classifier\ 1$ individually to see its performance on different image features.
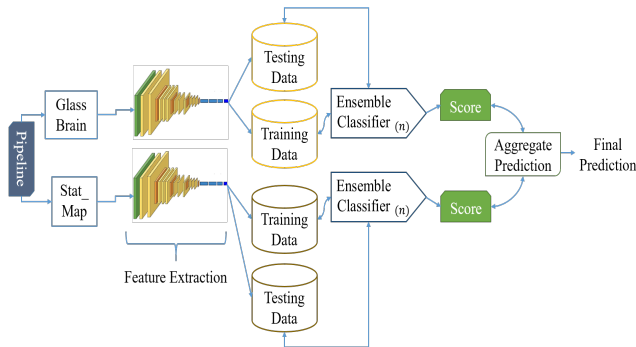
Fig. 3: Architectural overview of the proposed deep ensemble learning classifier.

The activation of the last layer was set to sigmoid function for the classification task. Each ensemble model was trained for 1000 iterations using ADAM optimizer with learning rate 0.0004, $beta\_1 = 0.9$, and $beta\_2 = 0.999$.

IV. EXPERIMENTS AND DISCUSSION

A. Participants

Resting-state fMRI allows for the investigation of the disturbance of brain networks without the added complexity of variation associated with task-related brain activation. The present study was carried out using 4D rs-fMRI data from the ABIDE-1, which is a particularly complex dataset due to its heterogeneity, vast range subjects comprised and different imaging protocols. It's also a connotation that provides previously collected multi-site ASD and matched typical control data in the scientific research community from 17 different imaging sites around the world [39]. This multisite data platform has a diagnostic procedure different from each other. Also, the scanning protocols and parameters are varied from one site to another site with controlling data quality. The fundamental benefits of the ABIDE platform is that they have available preprocessed data using different preprocessing pipelines and images from one site can be considered as one dataset. The rs-fMRI preprocessed data was downloaded through four preprocessing pipelines from the Preprocessed Connectomes Projects (PCP)[1]. We incorporated data from 529 ASD individuals and 573 typical controls (TC) for each pipeline. Table IV contains the key phenotypical information, including distribution of ASD and TC by sex and age, and the ADOS score, where Ψ means that the site did not have

this information. The 1000 Functional Connectomes Project[2] contains complete information about the data site, including full site name, corresponding laboratory/hospital, connecting projects and collaborations.

TABLE IV: Phenotypic summary of ABIDE-I 4D fMRI data for each individual site.

Site	Count		Sub. Total	Age Avg.(SD)	ADOS-ASD(SD)
	ASD	TC			
Caltech	19	19	38	28.16 (10.6)	13.1 (4.5)
CMU	14	13	27	26.6 (5.7)	13.07 (3.12)
KKI	22	33	55	10.09 (1.33)	12.55 (3.53)
Leuven-1	14	15	29	22.59 (3.55)	Ψ
Leuven-2	15	20	35	14.16 (1.42)	Ψ
MaxMun	24	33	57	26.16 (12.08)	9.5 (3.56)
NYU	79	105	184	15.25 (6.58)	11.30 (4.08)
OHSU	13	15	28	10.81 (1.87)	9.69 (3.54)
Olin	20	16	36	16.81 (3.49)	14.1 (4.04)
Pitt	30	27	57	18.90 (6.88)	12.38 (3.22)
SBL	15	15	30	34.37 (8.60)	9.25 (1.67)
SDSU	14	22	36	14.41 (1.84)	11.15 (4.26)
Stanford	20	20	40	9.96 (1.58)	11.74 (3.23)
Trinity	24	25	49	17.18 (3.64)	10.54 (2.87)
UCLA_1	41	32	73	13.16 (2.40)	10.24 (3.56)
UCLA_2	12	14	26	12.49 (1.53)	13.25 (2.67)
UM_1	55	55	110	13.40 (2.89)	Ψ
UM_2	13	22	35	15.96 (3.32)	Ψ
USM	58	43	101	22.10 (7.68)	9.03 (6.45)
Yale	28	28	56	12.71 (2.88)	11 (Ψ)

B. Visualization of the Generated Single Volume Images

The envisioned image generator, firstly ensures the shape of the functionally preprocessed images. If the checked image shape is 4D NIFTI (.nii.gz), then display the brain images according to the voxel time points of the 4D rs-fMRI with the set parameters. For instance, if an image has volume 176, the generator, therefore generates 176 brain images. Finally, the generator automatically saves the single volume images from the exposed whole volume brain images in PNG format according to the defined label. The produced glass brain and stat_map images for the first ten subjects of CCS, CPAC, DPASRF, and NIAK pipelines from both ASD and TC are shown in Fig. 4, Fig. 5, Fig. 6 and Fig. 7, respectively. In Fig. 4 - Fig. 7, row a and b represents the corresponding

[1] http://preprocessed-connectomes-project.org/abide/download.html

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

glass brain and stat_map images of ASD and TC individuals, respectively. Fig. 4 - Fig. 7, are merely presented to show the first image of each subject for both displaying modes which may not be more informative compared with other images. However, during training and testing the classification model, we considered all of the generated images for each individual.
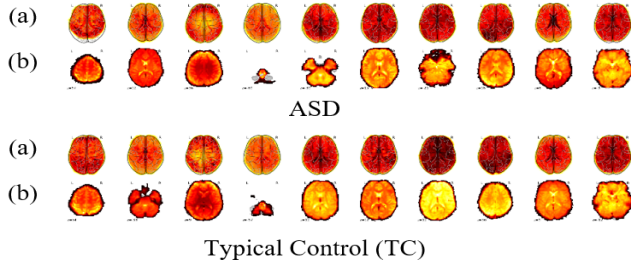

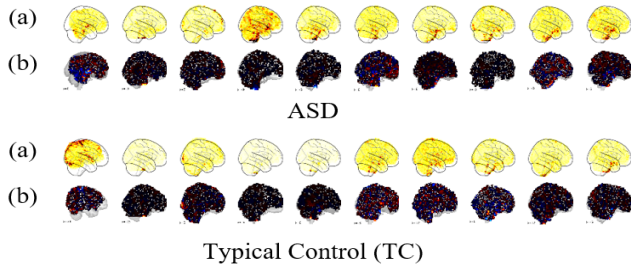Fig. 4: Axial view (z) of the CCS pipeline images.
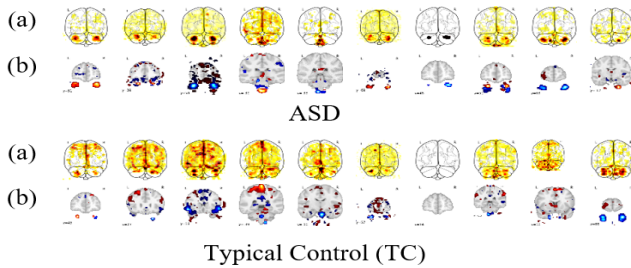

Fig. 5: Sagittal view (x) of the CPAC pipeline images.

Fig. 6: Coronal view (y) of the DPASRF pipeline images.
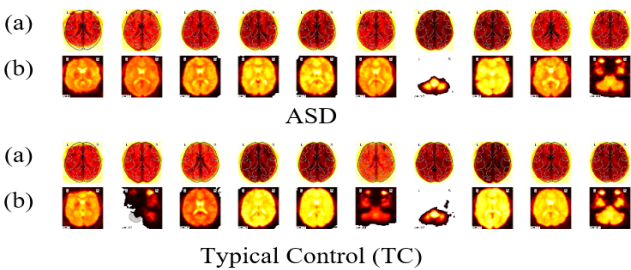
Fig. 7: Axial view (z) of the NIAK pipeline images.

From the analysis of the visualized images, Fig. 4 and Fig. 7 shows the horizontal slice that divides the brain into upper (superior) and lower (inferior) segments and is an x-z plane. On the other hand, Fig. 5 illustrates the longitudinal slice that separates the brain into left and right halves and is also an x-z plane. It is called the sagittal view because it transpires through the sagittal stitch. From Fig. 6, it depicts the frontal spheres that divide the brain into the front (posterior) and back (anterior) parts and is an x-y plane.

Finally, we process the generated images according to the model requirements because the shape of the generated images varies with different volumes for different imaging sites during image acquisition. The processing details of the input images for each model are addressed in the following subsection.

C. Model Training and Testing

In our experiment, to test the robustness of the generated images and improved model, we first divided the whole dataset into two parts: training (85%) and testing (15%) from the generated images. We reckoned every image according to their subject's image as it is either ASD or TC for binary classification. During training the first model (Fig. 2), we considered Keras based ImageDataGenerator to process the generated images according to the model input requirements. By setting up the necessary parameters such as batch size (64), rescale (1. /255), and class mode (binary), we reshaped the images in target size (160,160) of RGB images. Conversely, in the ensemble learning classifier (Fig. 3), we incorporated the VGG16 model for feature extraction. We then split the data into two parts of training (85%) and testing (15%) to feed into the classifier models. Depending on the number of voxel time points of the fMRI data, we generate multiple images for each subject. Therefore, for training and testing the deep learning models, we get an adequate number of images. Table V manifests the number of total images for each site individually according to their splitting procedure. In Table V, all the images have the same shape with different time points for different data site. For instance, the original fMRI has a shape like (61, 73, 61, voxel time points); thus, the fMRI time points are defined as the fourth dimension of the original

TABLE V: Number of generated images for each specific data site including separate training and testing set.

Data Site	Data Size	fMRI Time Points	Generated Image Number	Training Set (85%)	Testing Set (15%)
Caltech	38	145	5510	4684	826
CMU	27	235	6345	5393	952
KKI	55	151	8305	7059	1246
Leuven	64	245	15680	13328	2352
MaxMun	57	115	6555	5572	983
NYU	184	175	32200	27370	4830
OHSU	28	77	2156	1833	323
OLIN	36	205	7380	6273	1107
PITT	57	195	11115	9448	1667
SBL	30	195	5850	4972	878
SDSU	36	175	6300	5355	945
Stanford	40	175	7000	5950	1050
Trinity	49	145	7105	6039	1066
UCLA	99	115	11385	9677	1708
UM	145	295	42775	36359	6416
USM	101	235)	23735	20175	3560
Yale	56	195	10920	9282	1638
Total	1102	3037	210316	178769	31547
Mean	64.8	180.8	12371.5	10515.8	1855.7

all site

TABLE VI: Corresponding classification performances using proposed CNN model.

Pipelines	Generated Image	DenseNet (%)				ResNet (%)				Xception (%)				Inception V3 (%)				Proposed Method (%)			
		P	R	F	A	P	R	F	A	P	R	F	A	P	R	F	A	P	R	F	A
CCS	Glass Brain	79.1	78.2	78.6	80.1	78.5	79.2	78.8	79.8	80.4	77.8	79.1	82.4	80.1	78.9	79.5	80.1	77	79.5	78.3	80.2
	Stat_Map	78.2	77.8	78	79.8	78.3	80.3	79.3	81.9	77.6	79.6	78.6	80.4	78.1	79.2	78.6	79.3	**80.5**	79.3	79.9	79.3
CPAC	Glass Brain	78.9	76.9	77.9	79	79.4	78.4	78.9	79.7	78.6	76.9	77.7	82.6	78.8	77.2	78	80.8	78.2	79.5	78.9	**83**
	Stat_Map	77.9	79	78.5	80	76.9	78.8	77.8	80	79.3	78.9	79.1	79.5	78.8	77.9	78.3	79.9	79.4	77.7	78.5	78.5
DPARSF	Glass Brain	78.8	79.8	79.3	80.6	79.7	80.2	79.9	80.7	79.8	79.9	79.8	79.8	78.8	79.1	78.9	78.8	79	**81.3**	80.1	81.4
	Stat_Map	78.9	78.9	78.9	81.2	77.8	79.7	78.7	79.8	78.6	78.9	78.7	78.9	77.8	79.9	78.8	80.4	80.1	78.2	79.1	81.6
NIAK	Glass Brain	79.8	80	79.9	**81.5**	79.6	80.1	79.4	**82.1**	79.7	79.8	79.7	80.2	80.2	78.5	79.3	80.6	77.8	81	**80.9**	80.6
	Stat_Map	76.9	79.6	78.3	78.8	78.7	80.3	79.5	80.3	77.9	78.6	78.3	79.6	79.8	77.8	78.8	78.7	76.9	79.4	78.1	79.6

fMRI shape. In this work, we consider the number of voxel time points to generate 2D brain images.

On the other hand, during unit-site classification, we performed a leave-one-site-out 5-fold cross-validation approach to evaluate the performances of the proposed CNN classifier across sites. This method helps to extract more information from the images while leaving enough test samples to measure the capability of the model in classifying unobserved images. Additionally, some of the sites have a small number of samples, for example, CMU (Carnegie Mellon University), which contains only 27 samples. Therefore, performing leave-site-out 5-fold cross-validation increases the variance of cross-validation estimation.

D. CNN-based Classifier Performance Evaluation on ABIDE Dataset

Deep learning approach and ABIDE data have previously been studied to identify and analyze the ASD showing different measurement metrics. In this work, we evaluate four performance measurement metrics such as P (precision), R (recall), F (F1-score) and A (accuracy) to validate the algorithm performance. Precision is outlined as the ratio of the correctly ASD positive labeled to all ASD positive labeled, and recall is the ratio with the whole ASD positive in reality. On the other hand, F1-score considers both precision and recall measurement. Accuracy is delineated as the ratio of correctly labeled individuals to the whole number of subjects. All of the four metrics are weighed to validate the classification ability of our model for ASD and TC classification. Table VI shows the comparison between different performance measurements for benchmark and proposed CNN method using two categorized images of four pipelines.

The maximum performance is marked as bold in each section. The highest average accuracy obtain in our proposed method for CPAC glass brain images is 83%, average precision 80.5% for CCS stat_map images, average F1-score 80.9% for

DPARSF glass brain images and average specificity 81.2% for NIAK glass brain images. From the analysis of the performance comparison, it shows that the performance of the improved CNN over other methods is statistically noteworthy.

E. Deep Ensemble Learning Classifiers Performance Analysis

We performed four ensemble classifier techniques in the experiment by combining benchmark approaches with the improved CNN. We have trained each classifier separately for two different image feature's classification and shown the corresponding outputs. All the ensemble classifiers were trained using ADAM optimizer and sigmoid function for classification. The final output was taken based on the $equation$ (6) for the binary classification. Table VII represents the classification performance for the ensemble learning classifiers. The highest accuracy and other relevant measurements are marked as bold in the table for each pipeline generated images. From the analysis of the table results, the third ensemble classifier performs better than other ensemble classifiers in CCS and CPAC pipelines data, with the highest accuracy of 87%, respectively. For the other two pipelines, DPARSF and NIAK data, the second and third classifiers perform closely for the DPARSF dataset, and the first and second classifiers perform closely for the NIAK dataset, with an accuracy of 86% in both cases. The fourth ensemble classifier performs averagely in all four cases.

Ensemble learning has already been introduced for ASD classification in many literature reports but not frequently yet. In [35], Khosla et al. used stochastic parcellation and seven atlases of ABIDE dataset to classify ASD based on the 3D CNN approach. The authors also performed two ensemble learning strategies called multi-atlas ensemble (MA-Ensemble) model and stochastic parcellation ensemble strategy (SP-Ensemble) of a 3D CNN method. The average classification predictions of the MA-Ensemble model were computed using each one of the seven atlases available in PCP for ROI time series

TABLE VII: Classification performances using proposed deep ensemble learning classifiers.

Pipelines	Ensemble Classifier 1				Ensemble Classifier 2				Ensemble Classifier 3				Ensemble Classifier 4			
	P	R	F	A	P	R	F	A	P	R	F	A	P	R	F	A
CCS	82.3	79.9	81.1	**86**	85.6	84.5	85.1	**86**	**86.8**	85.2	86	**87**	86.5	85.8	**86.2**	84
CPAC	83.2	81.2	82.2	85	**86.4**	**85.3**	**85.8**	85	85.2	83.8	84.5	**87**	82.4	**86.4**	84.4	**85**
DPARSF	**86.1**	82.3	**84.2**	85	82.4	82.3	82.4	**86**	86.7	**85.9**	**86.3**	86	83.5	84.1	83.8	**85**
NIAK	83	**84.1**	83.5	**86**	81.9	83.4	82.6	**86**	**86.8**	81.8	84.2	85	84.6	84.5	84.5	84

extraction, with the highest accuracy of 71.7%. Using 30 stochastic parcellations in the SP-Ensemble model obtained a maximum efficiency of 72.3%. Additionally, in [36], Wang et al. firstly divided the data according to the subject's age and sex and found out their functional connectivity patterns for ASD classification. They proposed a sparse multi-view multitask ensemble (Sparse-MVMT-E) classification method for individualized ASD diagnosis. They considered two data sites from ABIDE, namely NYU and UM-1, and secured the highest accuracy of 72.6% and 71.4%, respectively. Compared with the results of [35] and [36],our model surpassed with a mean accuracy difference around 14% using different types of image features rather than FC or atlas data. The results comparison suggests that our proposed ensemble model classifies ASD patients more precisely.

F. Unit Site Classification Performances Analysis

As the ABIDE is a consortium of the large dataset for ASD subjects from multiple renowned institutions around the world, we performed unit-site classification to pattern out the site's variability. Leave-One-Site-Out 5-fold cross-validation has been employed to evaluate the CNN classifier performance across the sites. This process firstly excluded the data from one site from the training process and used that data as the test set to evaluate the model. Then it tested the applicability

of the model to a new, different site. Leave-One-Site cross-validation always estimates the entire 5-fold cross-validation using a single split of data folds. During the experiments, we accomplish the confidence intervals through the standard deviation (SD) of the mean accuracy. Table VIII represents the performance comparison using the proposed CNN model, including benchmark approaches and mean accuracy's standard deviation (SD) for each model. In our experiments, the highest accuracy for glass brain images is up to 88% belongs to several data sites and for stat_map images the accuracy is up to 87% for several data sites as shown in the table. Heinsfeld et al. also investigated the unit-site classification, where they achieved the highest accuracy of 68%, both for Caltech and MaxMun data site, by using patterns of functional connectivity [37]. On the other hand, Eslami et al. performed 5-fold cross-validation on each site separately using ASD-DiagNet [38]. They got the maximum accuracy of 82% for the OHSU data site without data augmentation. In both works [37] and [38], CC-200 functional parcellation brain atlas data was utilized for intra-site evaluation. The closest accuracy comparing with our model has been achieved by [38] in only one data site, while other data site performs far lower with data augmentation. The mean accuracy difference is approximately 20% between our and the comparative works. Therefore, based on the results, our proposed method, along with benchmark

TABLE VIII: Unit site classification using leave-site-out 5-fold cross-validation, compared with other methods.

Site-Out	Hensfeld et al. 2018 [37]	Eslami et al. 2019 [38]	Generated Images	DenseNet (%) (SD)	ResNet (%) (SD)	Xception (%) (SD)	Inception V3 (%) (SD)	Proposed Method (%) (SD)
Caltech	68	52.8	Glass Brain	83(0.11)	81(0.31)	83(0.03)	84(0.26)	**86(0.17)**
			Stat Map	81(0.21)	84(0.18)	82(0.13)	84(0.26)	85(0.01)
CMU	66	68.5	Glass Brain	82(0.06)	83(0.02)	84(0.23)	82(0.08)	84(0.19)
			Stat Map	83(0.13)	85(0.40)	80(0.54)	81(0.27)	85(0.01)
KKI	67	69.5	Glass Brain	83(0.09)	84(0.15)	84(0.16)	85(0.35)	**86(0.14)**
			Stat Map	82(0.04)	82(0.12)	83(0.30)	84(0.22)	85(0.01)
Leuven	65	61.3	Glass Brain	86(0.46)	85(0.26)	85(0.28)	86(0.45)	**86(0.13)**
			Stat Map	81(0.16)	84(0.14)	82(0.10)	84(0.45)	85(0.01)
MaxMun	68	48.6	Glass Brain	80(0.30)	81(0.25)	82(0.11)	80(0.32)	82(0.40)
			Stat Map	82(0.04)	80(0.38)	82(0.11)	80(0.32)	81(0.53)
NYU	66	68	Glass Brain	85(0.20)	83(0.01)	84(0.09)	86(0.27)	**88(0.22)**
			Stat Map	86(0.27)	83(0.01)	85(0.16)	84(0.12)	**86(0.08)**
OHSU	64	82	Glass Brain	79(0.62)	80(0.55)	81(0.34)	78(0.83)	82(0.56)
			Stat Map	77(1.00)	78(0.93)	82(0.15)	77(1.02)	84(0.18)
OLIN	64	65.1	Glass Brain	83(0.12)	84(0.18)	80(0.47)	84(0.27)	85(0.01)
			Stat Map	79(0.55)	82(0.15)	78(0.80)	82(0.07)	83(0.33)
PITT	66	67.8	Glass Brain	81(0.17)	83(0.01)	84(0.16)	80(0.32)	84(0.13)
			Stat Map	79(0.44)	82(0.12)	83(0.03)	79(0.45)	85(0.01)
SBL	66	51.6	Glass Brain	82(0.05)	83(0.02)	84(0.22)	80(0.44)	**86(0.19)**
			Stat Map	83(0.13)	83(0.02)	81(0.33)	82(0.07)	84(0.18)
SDSU	63	63	Glass Brain	82(0.05)	83(0.02)	82(0.13)	84(0.27)	85(0.01)
			Stat Map	83(0.12)	83(0.02)	82(0.13)	81(0.23)	84(0.16)
Stanford	66	64.2	Glass Brain	82(0.05)	84(0.17)	82(0.13)	83(0.09)	85(0.01)
			Stat Map	80(0.36)	83(0.02)	84(0.19)	82(0.06)	85(0.01)
Trinity	65	54.1	Glass Brain	85(0.39)	82(0.13)	81(0.26)	83(0.09)	85(0.01)
			Stat Map	83(0.10)	82(0.13)	83(0.03)	80(0.34)	84(0.14)
UCLA	66	73.2	Glass Brain	85(0.27)	83(0.01)	84(0.12)	86(0.36)	**87(0.20)**
			Stat Map	83(0.07)	81(0.19)	85(0.22)	83(0.06)	**86(0.10)**
UM	64	63.8	Glass Brain	84(0.14)	85(0.17)	86(0.27)	84(0.13)	**88(0.25)**
			Stat Map	84(0.14)	86(0.26)	83(0.02)	86(0.30)	**87(0.17)**
USM	64	68.2	Glass Brain	86(0.37)	85(0.21)	85(0.22)	84(0.16)	**86(0.10)**
			Stat Map	81(0.13)	84(0.11)	82(0.08)	84(0.16)	85(0.01)
Yale	64	63.6	Glass Brain	83(0.09)	82(0.12)	84(0.16)	80(0.32)	85(0.01)
			Stat Map	83(0.09)	86(0.41)	84(0.16)	81(0.19)	85(0.01)
Mean	65	63.8		82.3(0.22)	82.9(0.18)	82.8(0.20)	82.4(0.28)	84.9(0.14)

approaches, obtains the highest accuracy in most data sites and outperforms other methods on average for classifying ASD. Furthermore, the standard deviation between each data site is lower, which is an indication that the accuracy is closer to the mean accuracy. It can be concluded that the results, compared with other findings, also suggest that there has data variability (dispersion or spread) among these sites that do not exist in other sites.

V. CONCLUSION

The recent advancement of functional connectivity and brain ROIs analysis has made a conspicuous invasion into the following classification of ASD. However, it is challenging to generalize the outcomes for larger, more heterogeneous populations rather than for smaller ones. While most of the recent work investigated the functional connectivity or time series analysis of fMRI, in this study, we demonstrate a suitable image generator to harvest a stable image that can provide perceptive details on the target disease using heterogeneous neuroimaging modality. Also, we validate the generated images using two propose deep learning-based frameworks that could enhance diagnostic truthfulness, with the potential to classify and develop better treatments. Furthermore, to check out the inter-site data variability, we apply the proposed method across the sites using a leave-site-out cross-validation approach. Our significance image processing scheme and sampling, along with the precise CNN classifier ensure a trustworthy approach of ASD classification associating with the other image processing techniques. Overall, the proposed image processing scheme provides a proficient and objective way of interpreting neuroimaging applied to the deep learning model.

Future research inclinations involve expending structural preprocessing and calculation of cortical measures pipeline data for classification of ASD. Besides, it is necessary to modify the architecture to consolidate raw fMRI data as well as to analyze the correlation between brain activation regions (axial, sagittal and coronal) to perceive the neural connectivity of brain during the natural progression of autism.

ACKNOWLEDGMENT

REFERENCES

[1] E. Honey, J. Rodgers, and H. McConachie, "Measurement of restricted and repetitive behaviour in children with autism spectrum disorder: Selecting a questionnaire or interview," *Research in Autism Spectrum Disorders*, vol. 6, no. 2, pp. 757–776, 2012.

[2] M. A. Just, V. L. Cherkassky, A. Buchweitz, T. A. Keller, and T. M. Mitchell, "Identifying autism from neural representations of social interactions: Neurocognitive markers of autism," *PLOS ONE*, vol. 9, no. 12, pp. 1–22, 2014.

[3] G. Noriega, "Restricted, repetitive, and stereotypical patterns of behavior in autism—an fmri perspective," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 27, no. 6, pp. 1139–1148, 2019.

[4] J. Baio, "Prevalence of autism spectrum disorder among children aged 8 years-autism and developmental disabilities monitoring network, 11 sites, united states," *Cent. Dis. Control Prev.*, vol. 63, pp. 1–24, 2014.

[5] M. D. Kaiser, C. M. Hudac, S. Shultz, S. M. Lee, C. Cheung, A. M. Berken, B. Deen, N. B. Pitskel, D. R. Sugrue, A. C. Voos, C. A. Saulnier, P. Ventola, J. M. Wolf, A. Klin, B. C. Vander Wyk, and K. A. Pelphrey, "Neural signatures of autism," *Proceedings of the National Academy of Sciences*, vol. 107, no. 49, pp. 21 223–21 228, 2010.

[6] M. Lee, D. Y. Kim, M. K. Chung, A. L. Alexander, and R. J. Davidson, "Topological properties of the structural brain network in autism via ε-neighbor method," *IEEE Transactions on Biomedical Engineering*, vol. 65, no. 10, pp. 2323–2333, 2018.

[7] X.-a. Bi, Y. Wang, Q. Shu, Q. Sun, and Q. Xu, "Classification of autism spectrum disorder using random support vector machine cluster," *Frontiers in Genetics*, vol. 9, p. 18, 2018.

[8] Y. Kong, J. Gao, Y. Xu, Y. Pan, J. Wang, and J. Liu, "Classification of autism spectrum disorder by combining brain connectivity and deep neural network classifier," *Neurocomputing*, vol. 324, pp. 63–68, 2019.

[9] X. Bi, Y. Liu, Q. Sun, X. Luo, H. Tan, J. Chen, and N. Zeng, "The genetic-evolutionary random support vector machine cluster analysis in autism spectrum disorder," *IEEE Access*, vol. 7, pp. 30 527–30 535, 2019.

[10] B. A. Cociu, S. Das, L. Billeci, W. Jamal, K. Maharatna, S. Calderoni, A. Narzisi, and F. Muratori, "Multimodal functional and structural brain connectivity analysis in autism: A preliminary integrated approach with eeg, fmri, and dti," *IEEE Transactions on Cognitive and Developmental Systems*, vol. 10, no. 2, pp. 213–226, 2018.

[11] Z. Wang, Y. Zheng, D. C. Zhu, A. C. Bozoki, and T. Li, "Classification of alzheimer's disease, mild cognitive impairment and normal control subjects using resting-state fmri based network connectivity analysis," *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 6, pp. 1–9, 2018.

[12] B. Thirion, G. Varoquaux, E. Dohmatob, and J.-B. Poline, "Which fmri clustering gives good brain parcellations?" *Frontiers in Neuroscience*, vol. 8, p. 167, 2014.

[13] C. Wang, Z. Xiao, B. Wang, and J. Wu, "Identification of autism based on svm-rfe and stacked sparse auto-encoder," *IEEE Access*, vol. 7, pp. 118 030–118 036, 2019.

[14] Z. Yao, B. Hu, Y. Xie, F. Zheng, G. Liu, X. Chen, and W. Zheng, "Resting-state time-varying analysis reveals aberrant variations of functional connectivity in autism," *Frontiers in human neuroscience*, vol. 10, p. 463, 2016.

[15] X. Li, N. C. Dvornek, J. Zhuang, P. Ventola, and J. S. Duncan, "Brain biomarker interpretation in asd using deep learning and fmri," in *MICCAI-2018, Granada, Spain*, 2018, pp. 206–214.

[16] M. R. Ahmed, Y. Zhang, Z. Feng, B. Lo, O. T. Inan, and H. Liao, "Neuroimaging and machine learning for dementia diagnosis: Recent advancements and future prospects," *IEEE Reviews in Biomedical Engineering*, vol. 12, pp. 19–33, 2019.

[17] J. R. Sato, M. Calebe Vidal, S. de Siqueira Santos, K. Brauer Massirer, and A. Fujita, "Complex network measures in autism spectrum disorders," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 15, no. 2, pp. 581–587, 2018.

[18] Y. Zhao, Q. Dong, S. Zhang, W. Zhang, H. Chen, X. Jiang, L. Guo, X. Hu, J. Han, and T. Liu, "Automatic recognition of fmri-derived functional networks using 3-d convolutional neural networks," *IEEE Transactions on Biomedical Engineering*, vol. 65, no. 9, pp. 1975–1984, 2018.

[19] T. Iidaka, "Resting state functional magnetic resonance imaging and neural network classified autism and control," *Cortex*, vol. 63, pp. 55 – 67, 2015.

[20] P. Kassraian-Fard, C. Matthis, J. H. Balsters, M. H. Maathuis, and N. Wenderoth, "Promises, pitfalls, and basic guidelines for applying machine learning classifiers to psychiatric imaging data, with autism as an example," *Frontiers in Psychiatry*, vol. 7, p. 177, 2016.

[21] J. V. Hull, L. B. Dokovna, Z. J. Jacokes, C. M. Torgerson, A. Irimia, and J. D. Van Horn, "Resting-state functional connectivity in autism spectrum disorders: A review," *Frontiers in Psychiatry*, vol. 7, p. 205, 2017.

[22] A. Pascual-Belda, A. Díaz-Parra, and D. Moratal, "Evaluating functional connectivity alterations in autism spectrum disorder using network-based statistics," *Diagnostics*, vol. 8, no. 3, 2018.

[23] A. Abraham, M. P. Milham, A. D. Martino, R. C. Craddock, D. Samaras, B. Thirion, and G. Varoquaux, "Deriving reproducible biomarkers from multi-site resting-state data: An autism-based example," *NeuroImage*, vol. 147, pp. 736–745, 2017.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2020.2998603, IEEE Journal of Biomedical and Health Informatics

IEEE JOURNAL OF BIOMEDICAL AND HEALTH INFORMATICS, VOL. XX, NO. XX, MAY 2020
11

[24] X. Guo, K. C. Dominick, A. A. Minai, H. Li, C. A. Erickson, and L. J. Lu, "Diagnosing autism spectrum disorder from brain resting-state functional connectivity patterns using a deep neural network with a novel feature selection method," *Frontiers in Neuroscience*, vol. 11, p. 460, 2017.

[25] N. C. Dvornek, P. Ventola, K. Pelphrey, and J. S. Duncan, "Identifying autism from resting-state fmri using long short-term memory networks," in *Machine learning in medical imaging. MLMI (Workshop)*, vol. 10541, 2017, pp. 362–370.

[26] S. B. Eickhoff, B. Thirion, G. Varoquaux, and D. Bzdok, "Connectivity-based parcellation: Critique and implications." *Human brain mapping*, vol. 36 12, pp. 4771–92, 2015.

[27] N. C. Dvornek, P. Ventola, and J. S. Duncan, "Combining phenotypic and resting-state fmri data for autism classification with recurrent neural networks," in *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, 2018, pp. 725–728.

[28] S. Parisot, S. I. Ktena, E. Ferrante, M. Lee, R. G. Moreno, B. Glocker, and D. Rueckert, "Spectral graph convolutions for population-based disease prediction," in *MICCAI-2017, Quebec City, Canada*, 2017, pp. 177–185.

[29] X. Li, N. C. Dvornek, X. Papademetris, J. Zhuang, L. H. Staib, P. Ventola, and J. S. Duncan, "2-channel convolutional 3d deep neural network (2cc3d) for fmri analysis: Asd classification and feature learning," in *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), Washington, DC, USA*, 2018, pp. 1252–1255.

[30] Y. Zhao, F. Ge, S. Zhang, and T. Liu, "3d deep convolutional neural network revealed the value of brain network overlap in differentiating autism spectrum disorder from healthy controls," in *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI 2018), Granada, Spain*, 2018, pp. 172–180.

[31] S. I. Ktena, S. Parisot, E. Ferrante, M. Rajchl, M. C. H. Lee, B. Glocker, and D. Rueckert, "Distance metric learning using graph convolutional networks: Application to functional brain networks," in *MICCAI*, 2017.

[32] R. Anirudh and J. J. Thiagarajan, "Bootstrapping graph convolutional neural networks for autism spectrum disorder classification," in *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2019, pp. 3197–3201.

[33] A. Phinyomark, E. Ibanez-Marcelo, and G. Petri, "Resting-state fmri functional connectivity: Big data preprocessing pipelines and topological data analysis," *IEEE Transactions on Big Data*, vol. 3, no. 4, pp. 415–428, 2017.

[34] D. Ravì, C. Wong, F. Deligianni, M. Berthelot, J. Andreu-Perez, B. Lo, and G. Yang, "Deep learning for health informatics," *IEEE Journal of Biomedical and Health Informatics*, vol. 21, no. 1, pp. 4–21, 2017.

[35] M. Khosla, K. Jamison, A. Kuceyeski, and M. R. Sabuncu, "Ensemble learning with 3d convolutional neural networks for functional connectome-based prediction," *NeuroImage*, vol. 199, pp. 651 – 662, 2019.

[36] J. Wang, Q. Wang, H. Zhang, J. Chen, S. Wang, and D. Shen, "Sparse multiview task-centralized ensemble learning for asd diagnosis based on age- and sex-related functional connectivity patterns," *IEEE Transactions on Cybernetics*, vol. 49, no. 8, pp. 3141–3154, 2019.

[37] A. S. Heinsfeld, A. R. Franco, R. C. Craddock, A. Buchweitz, and F. Meneguzzi, "Identification of autism spectrum disorder using deep learning and the abide dataset," *NeuroImage: Clinical*, vol. 17, pp. 16 – 23, 2018.

[38] T. Eslami, V. Mirjalili, A. Fong, A. R. Laird, and F. Saeed, "Asd-diagnet: A hybrid learning approach for detection of autism spectrum disorder using fmri data," *Frontiers in Neuroinformatics*, vol. 13, p. 70, 2019.

[39] C. Craddock, Y. Benhajali, C. Carlton, C. Francois, E. Alan, A. Jakab, B. Khundrakpam, J. Lewis, l. Qingyang, M. Michael, Y. Chaogan, and P. Bellec, "The neuro bureau preprocessing initiative: open sharing of preprocessed neuroimaging data and derivatives," *Frontiers in Neuroinformatics*, vol. 7, 2013.

[40] "Abide preprocessed," http://preprocessed-connectomes-project.org/abide/, last Accessed: 2020-05-18.

[41] X. Liu, C. Chang, and J. Duyn, "Decomposition of spontaneous brain activity into distinct fmri co-activation patterns," *Frontiers in Systems Neuroscience*, vol. 7, p. 101, 2013.

[42] J. R. Sato, C. Chiann, E. Taniguchi, E. son Gomes dos Santos, P. R. Arantes, M. L. Mourão, E. A. Júnior, and P. A. Morettin, "Identifying multisubject cortical activation in functional mri: A frequency domain approach," *Journal of data science*, vol. 6, pp. 89–103, 2008.

[43] G. S. Dichter, L. Sikich, A. Song, J. Voyvodic, and J. W. Bodfish, "Functional neuroimaging of treatment effects in psychiatry: Methodological challenges and recommendations," *International Journal of Neuroscience*, vol. 122, no. 9, pp. 483–493, 2012.

[44] T. Mullen, C. Kothe, Y. M. Chi, A. Ojeda, T. Kerth, S. Makeig, G. Cauwenberghs, and T.-P. Jung, "Real-time estimation and 3d visualization of source dynamics and connectivity using wearable eeg," *Proceedings of the Fifth International Brain-Computer Interface Meeting*, 2013.

[45] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *ICML-2015, Lille, France*, vol. 37, 2015, pp. 448–456.

[46] G. Huang, Z. Liu, L. v. d. Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 2261–2269.

[47] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778.

[48] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 1800–1807.

[49] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 2818–2826.

[50] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA*, 2015.

Md Rishad Ahmed received his B.Sc. and M.Sc. degree in Applied Physics and Electronic Engineering from University of Rajshahi, Bangladesh, in 2013 and 2015, respectively. After completion of his graduation, he has served in the First Capital University of Bangladesh, Chuadanga as a Lecturer in the Department of Electrical and Electronic Engineering. Currently, he is with the College of Electronic and Information Engineering, Southwest University, China, and also with the Shandong Provincial Key Laboratory of Network Based Intelligent Computing, University of Jinan, China as a M.S. student majoring Computer Science and Technology. His research interests include biomedical image processing, machine learning, deep learning, pattern recognition, and brain disorder classification.

Yuan Zhang received the M.S. degree in communication systems and the Ph.D. degree in control theory and engineering both from Shandong University, China, in 2003 and 2012, respectively. He is currently a Professor with the College of Electronic & Information Engineering, Southwest University, China. Dr. Zhang was a visiting professor with the Computer Science Department, Georgia State University, USA, in 2014. Dr. Zhang's research interests include wearable sensing for smart health, machine learning for auxiliary diagnosis, and biomedical big data analytics. As the first author or corresponding author he has published more than 70 peer reviewed papers in international journals and conference proceedings, 1 book chapter, and 11 patents in the areas of Smart Health and Biomedical Big Data Analytics. He is a member of IEEE EMBS Wearable Biomedical Sensors and Systems Technical Committee. He is an associate editor for IEEE Reviews in Biomedical Engineering, IEEE Open Journal of Engineering in Medicine and Biology and IEEE Access. He has served as Leading Guest Editor for six special issues of IEEE, Elsevier, Springer and InderScience publications, including IEEE Internet of Things Journal special issue on Wearable Sensor Based Big Data Analysis for Smart Health and IEEE Journal of Biomedical and Health Informatics special issue on Pervasive Sensing and Machine Learning for Mental Health. He has served on the technical program committee for numerous international conferences. Dr. Zhang is a Senior Member of both IEEE and ACM.

Yi Liu received the M.S. degree in internal medicine from Capital Medical University, China, and the Ph.D. degree in internal medicine from Peking Union Medical College, China, in 2007 and 2019, respectively. She is currently an associate professor with the Peking University Health Science Center and a deputy chief physician of the Civil Aviation General Hospital. Dr. Liu's research interests include genetics of bronchial asthma, analysis of EEG signals from sleep disordered, and intelligent diagnosis of lung infections. As the first author or corresponding author, she has published 5 articles in international journals and conferences, published more than 10 articles in domestic journals, published 1 monograph, and obtained 1 patent in the field of brain fatigue monitoring. She is a member of the civil aviation technology innovation team. She is an expert member of the Wu Lien-The Public Foundation. She serves on the editorial board of the International Respiratory Journal.

Hongen Liao (M'04, SM'14) received the B.S. degree in mechanics and engineering sciences from Peking University, Beijing, China, in 1996, and the M.E. and Ph.D. degrees in precision machinery engineering from the University of Tokyo, Tokyo, Japan, in 2000 and 2003, respectively. He was a Research Fellow of the Japan Society for the Promotion of Science. Since 2004, he has been a faculty member at the Graduate School of Engineering, University of Tokyo, where he became an Associate Professor in 2007. He has been selected as a National "Thousand Talents" Distinguished Professor, National Recruitment Program of Global Experts, China, and is currently a full Professor and Vice Director in the Department of Biomedical Engineering, School of Medicine, Tsinghua University, Beijing, China. His research interests include 3-D medical image, image-guided surgery, medical robotics, computer-assisted surgery, and fusion of these techniques for minimally invasive precision diagnosis and therapy. He is the author and co-author of more than 250 peer-reviewed articles and proceedings papers, as well as 40 patents, over 290 abstracts and numerous invited lectures. He is an Associate Editor of IEEE Engineering in Medicine and Biology Society Conference, the Organization Chair of Medical Imaging and Augmented Reality Conference (MIAR) 2008, the Program Chair of the Asian Conference on Computer-Aided Surgery Conference (ACCAS) 2008 and 2009, the Tutorial co-chair of the Medical Image Computing and Computer Assisted Intervention Conference (MICCAI) 2009, the Publicity Chair of MICCAI 2010, the General Chair of MIAR 2010 and ACCAS 2012, the Program Chair of MIAR 2013, the Workshop Chair of MICCAI 2013 and 2019, and the General co-chair of MIAR 2016. He is a President of Asian Society for Computer Aided Surgery and Co-chair of Asian-Pacific Activities Working Group, Inter-national Federation for Medical and Biological Engineering (IFMBE).