

MADE-for-ASD: A Multi-Atlas Deep Ensemble Network for Diagnosing Autism Spectrum Disorder

Md Rakibul Hasan^{a,b,1}, Xuehan Liu^{c,1}, Tom Gedeon^a, Md Zakir Hossain^{a,c,*}

^aCurtin University, Bentley, 6102, WA, Australia

^bBRAC University, Dhaka, 1212, Bangladesh

^cAustralian National University, Canberra, 2601, ACT, Australia

Abstract

In response to the global need for efficient early diagnosis of Autism Spectrum Disorder (ASD), this paper bridges the gap between traditional, time-consuming diagnostic methods and potential automated solutions. We propose a multi-atlas deep ensemble network, *MADE-for-ASD*, that integrates multiple atlases of the brain's functional magnetic resonance imaging (fMRI) data through a weighted deep ensemble network. Our approach integrates demographic information into the prediction workflow, which enhances ASD diagnosis performance and offers a more holistic perspective on patient profiling. We experiment with the well-known publicly available ABIDE (Autism Brain Imaging Data Exchange) I dataset, consisting of resting state fMRI data from 17 different laboratories around the globe. Our proposed system achieves 75.20% accuracy on the entire dataset and 96.40% on a specific subset – both surpassing reported ASD diagnosis accuracy in ABIDE I fMRI studies. Specifically, our model improves by 4.4 percentage points over prior works on the same amount of data. The model exhibits a sensitivity of 82.90% and a specificity of 69.70% on the entire dataset, and 91.00% and 99.50%, respectively, on the specific subset. We leverage the F-score to pinpoint the top 10 ROI in ASD diagnosis, such as *precuneus* and anterior *cingulate/ventromedial*. The proposed system can potentially pave the way for more cost-effective, efficient and scalable strategies in ASD diagnosis. Codes and evaluations are publicly available at TBA.

Keywords: Autism, Neuroimaging, Computer Vision, Deep Learning, Health Computing

1. Introduction

Autism Spectrum Disorder (ASD) is a prevalent neurodevelopmental condition characterised by challenges in social and communicative abilities, as well as repetitive and hard-to-control behaviours in daily life [1]. ASD often co-occurs with various other conditions, such as intellectual impairment, seizures and anxiety. Autistic individuals display a wide range of characteristics,

*Corresponding author

Email addresses: Rakibul.Hasan@curtin.edu.au (Md Rakibul Hasan), u7094891@alumni.anu.edu.au (Xuehan Liu), Tom.Gedeon@curtin.edu.au (Tom Gedeon), Zakir.Hossain1@curtin.edu.au (Md Zakir Hossain)

¹Md Rakibul Hasan and Xuehan Liu contributed equally to this work.

from mild to severe social and communicative differences, along with restricted and repetitive behaviours and interests [2]. According to a 2022 report endorsed by the World Health Organisation² [3], approximately one in 100 children worldwide are autistic, a significant increase from the one in 160 reported in 2012 [4]. The economic impact on families of autistic individuals is substantial, making ASD a critical public health concern [5].

Diagnosis of ASD at present leverages a variety of techniques, predominantly behavioural assessments such as the Autism Diagnostic Observation Schedule (ADOS) [6] and the Autism Diagnostic Interview-Revised (ADI-R) [7]. Both ADOS and ADI-R deliver important insights into an individual’s communication, social interactions and behaviour. Nevertheless, these methods primarily depend on observing symptoms, making them largely subjective. The process can be both expensive and lengthy. Additionally, variations in symptom presentations and their severity can introduce additional complications, underscoring the need for a diagnostic method that is more efficient and objective [8]. In this context, functional magnetic resonance imaging (fMRI) of the brain could be leveraged because it provides a non-invasive means of examining brain activity. It has the potential to elucidate the intricate neurological deviations representing ASD, which can be used towards developing automated, objective, efficient and early diagnostic methods [9, 10].

The Autism Brain Imaging Data Exchange (ABIDE) consortium releases resting-state fMRI (rs-fMRI) data with T1 structural brain images and demographic information of autistic individuals and Typical Controls (TCs) [11]. ABIDE I dataset was collected from 17 international research sites, which makes it diverse and comprehensive for understanding the neurological nuances of ASD. Such heterogeneity in the data can help capture the diverse manifestations of ASD across different populations and geographical locations. ABIDE I data, therefore, can strengthen the generalisability of findings and enhance their clinical relevance and applicability. The preference for using ABIDE I data is proven by many prior works on ASD diagnosis [12, 13, 14, 9, 15, 16, 17, 18, 19, 20, 21, 22]. This paper uses ABIDE I dataset to classify between ASD and TC, hereinafter referred to as ASD diagnosis.

Machine learning (ML) and deep learning (DL) techniques have been increasingly employed to advance the understanding and diagnosis of various neurodevelopmental disorders, including ASD. ML and DL provide an objective and data-driven approach to diagnosis, thereby reducing the reliance on subjective symptom-based criteria. However, building appropriate models with high-dimensional multi-site neuroimaging data, such as ABIDE I, is challenging because of added dimensionality (i.e., different brain regions may have different cues towards ASD diagnosis) and site diversity (i.e., different data collection sites have different data collection protocols). Including other important cues, such as people’s demographic information, can potentially improve the predictive performance but make the system development more challenging while accommodating these data with the primary neuroimaging data. Addressing these challenges, we propose a novel DL-based ASD diagnosis workflow from ABIDE I brain fMRI data and corresponding demographic information. The input rs-fMRI images are processed to extract regions of interest (ROIs) according to three different atlases (brain parcellations), from where functional connectivity features are extracted. This paper’s novelty is two-fold. **(1)** It proposes *MADE-for-ASD*, consisting of a stacked sparse denoising autoencoder (SSDAE) and multi-layer perceptron (MLP), followed by a weighted ensemble learning framework. We incorporate demographic information into the model, which enhances the performance through personalised prediction. **(2)**

²<https://www.who.int/news-room/fact-sheets/detail/autism-spectrum-disorders>

This paper provides visualised insights into the most significant ROIs with a high correlation with ASD, which could further assist in understanding the neurobiological underpinnings of ASD.

2. Background and Related Work

2.1. Neuroimaging in ASD Diagnosis

Magnetic resonance imaging (MRI) is a critical tool for understanding the pathophysiology of neurological disorders such as schizophrenia and autism. MRI is valuable due to its cost-effectiveness and non-invasive nature, which have led to its widespread acceptance and application in the medical community [23]. Specifically, functional MRI (fMRI) tracks changes in blood oxygen levels over time, making it adept at inferring brain activity and drawing considerable attention from researchers studying brain dysfunctions [24]. Changes in the intensity of fMRI images throughout the acquisition period usually serve as a representation of brain activity, typically expressed as a time series. Brain disorders rarely manifest as anomalies in singular or multiple brain regions; they typically manifest as atypical connectivity among various brain regions. In this context, functional connectivity helps investigate the association of specific activities between brain regions and has found widespread use in the classification of brain disorders [25].

Several datasets are instrumental in advancing autism research by providing extensive neuroimaging data. The ABIDE I dataset aggregates resting-state fMRI data from 17 international sites [11]. ABIDE II expands on this by including additional subjects and sites, further enhancing its utility for ASD research [26]. Various works in ASD diagnosis have leveraged a variety of machine learning (ML) and deep learning (DL) models with rs-fMRI data from the ABIDE datasets [12, 13, 14, 9, 15, 16, 17, 18, 19, 20, 21, 22]. For example, Abraham et al. [12] experimented with several ML algorithms, including support vector regression and ridge regression, achieving a classification accuracy of 66.8%. However, these conventional machine learning models have limited ability to learn complex patterns from raw data and do not take advantage of deep representations. In a separate study, Heinsfeld et al. [13] embraced a deep learning technique with an autoencoder (AE) and deep neural network (DNN), reaching 70% accuracy. Eslami et al. [19] leveraged an AE with a single-layer perceptron and achieved 70.3% accuracy, while Almuqhim and Saeed [20] used an AE, achieving 70.8% classification accuracy. Although these studies have shown the proof of concept of using AEs in this task, their performance remains limited.

Few studies [27, 28, 29] utilised a subset of the ABIDE dataset. For example, Plitt et al. [29] employed a random forest classifier to predict ASD on only 179 samples and reported a classification accuracy of 95%. Chen et al. [28] selected 252 high-quality data based on criteria such as head motion, artifacts, and signal dropout and reported an accuracy of 91% using a random forest classifier.

Some studies [16, 30] used both ABIDE I and II datasets. For example, Khosla et al. [16] trained their Convolutional Neural Network (CNN) architecture with the ABIDE-I data and tested their model's performance on the ABIDE-II dataset. Aghdam et al. [30] proposed a 'mixture of experts' ensemble approach to diagnosing autistic young children (aged 5–10 years) in three dataset conditions: ABIDE I, ABIDE II, and a combination of them.

The National Database for Autism Research (NDAR) is another significant repository, offering a wide range of data types, including neuroimaging data, collected from various studies and institutions [31]. Using this dataset, Li et al. [32] proposed a multi-channel CNN model for early ASD diagnosis. Despite the variety of neuroimaging datasets available, the ABIDE I dataset remains the most widely used in ML-based ASD diagnosis [33].

2.2. Selection of Brain Atlases

Atlas-based parcellation, which divides the entire brain into spatially proximate ROIs, offers several advantages in neuroimaging studies [34]: it identifies brain regions with significant connectivity differences between groups, examines brain functional organisation, reduces data dimensionality and improves result interpretability by linking specific brain regions to conditions or phenotypes. The Automated Anatomical Labelling (AAL) atlas, with 116 ROIs, is widely used in ASD diagnosis studies [16, 35, 36, 37] due to its ability to provide precise anatomical locations essential for identifying structural abnormalities caused by ASD. The Craddock 200 atlas [38], with 200 ROIs, is based on functional connectivity and is particularly useful for analysing resting-state fMRI data [39, 9, 36, 37], as it clusters the brain into functionally homogeneous regions.

Recent studies have leveraged multiple atlases to capture diverse patterns related to ASD. For instance, Mahler et al. [37] proposed a multi-atlas framework for ASD classification using resting-state fMRI data, employing three atlases, including AAL and Craddock 200. Similarly, Deng et al. [9] utilised the AAL, Craddock 200 and Craddock 400 atlases. Their experiment indicates that Craddock 200 can outperform Craddock 400 in fMRI-based ASD diagnosis. Hence, we selected the Craddock 200 atlas, hereinafter referred to as the CC atlas.

To further enhance the analysis, we leverage the Eickhoff-Zilles (EZ) atlas with 116 ROIs, which has also been utilised in recent ASD diagnosis studies [40, 41]. Derived from cytoarchitectonic mapping, the EZ atlas combines anatomical and functional data, offering a nuanced view that bridges structural and functional aspects of brain regions [42]. By integrating the complementary strengths of AAL, CC and EZ atlases to capture different aspects of brain structure and function, our approach ensures a comprehensive analysis crucial for accurate ASD diagnosis.

In a typical fMRI-based ASD diagnosis pipeline [9, 16], atlas data is converted to functional connectivity matrices, which measures the degree of synchronised activity between different brain regions based on the time series of resting-state fMRI brain imaging data. While two of the atlases (AAL and EZ) focus on anatomical locations, they also provide a foundational framework for understanding the structural context within which functional connectivity occurs. Accordingly, functional connectivity of AAL and EZ atlases are utilised in ASD diagnosis literature [16, 9, 40].

3. Method

3.1. Task Formulation

Let us consider $X = \{X_{\text{fMRI}}, X_{\text{demog}}\}$, where X_{fMRI} and X_{demog} are the input fMRI images and demographic/phenotypic information, respectively, and $Y = \{Y_1, Y_2\}$, where Y_1 and Y_2 are ASD and TC classes, respectively. The goal is to develop a binary classifier \mathcal{F} to predict whether a sample input $x_i \in X$ can be classified as $y_i \in Y$.

3.2. Overall Framework

Figure 1 illustrates the overall framework of our ASD diagnosis system, which is fundamentally segmented into two distinct stages: the ASD/TC classification phase and the high-quality subset selection phase.

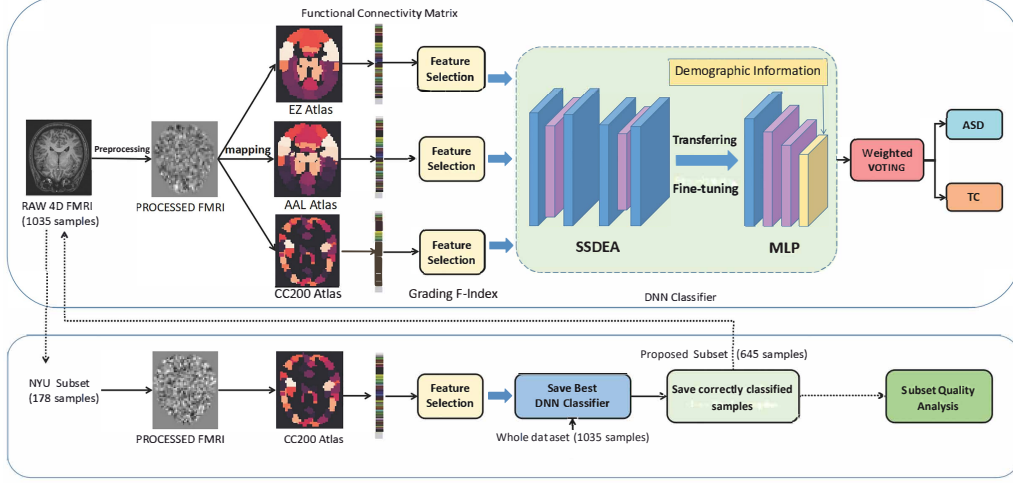


Figure 1: Overall framework of our ASD/TC classification workflow. We first calculate functional connectivity matrices on three brain atlases, followed by feature selection using F-score. To diagnose ASD, we use deep learning with a stacked sparse denoising autoencoder (SSDAE) and multi-layer perceptron (MLP), followed by a weighted ensemble. In the case of extracting high-quality data, we train the classifier with the NYU subset and predict on the whole dataset.

3.3. Data Preprocessing

As data quality control, we exclude all samples with missing fMRI time series. The missing values in the demographic data of the samples are imputed using the mean value of all available data of corresponding categories [43]. In this way, the overall distribution of the data is preserved, and the imputed values are likely to be close to the true values, assuming that the missingness is completely random. While we used the mean value to impute missing data, a more sophisticated approach, such as MICE [44], could be explored in future work.

We extract the mean time series of ROIs for each sample. We use parcellated regions as our targeted ROIs to extract voxel-level connectivity features. For each atlas, a respective connectivity matrix is formed, which is then condensed into a vector before being inputted into our model. The primary feature we use to differentiate between ASD and TC subjects is functional connectivity. We compute Pearson correlation coefficients between the time series of each pair of brain regions to produce a connectivity matrix:

$$\text{PCC}(u, v) = \frac{E(uv) - E(u)E(v)}{\sqrt{E(u^2) - E^2(u)} \sqrt{E(v^2) - E^2(v)}} \quad (1)$$

where $\text{PCC}(u, v)$ stands for the Pearson correlation coefficient between time series of two brain regions u and v , and $E(\cdot)$ is the mathematical expectation. These coefficients range from -1 to 1 ; coefficients near 1 indicate a strong positive correlation, while those near -1 indicate a strong negative correlation between the time series of two brain regions. For example, we obtain a 200×200 symmetric matrix for correlation in the case of the CC atlas because it has been divided into 200 regions. This functional connectivity matrix is used as a feature to classify subjects into ASD and TC groups.

To estimate the duplicated values in the matrix, we take the upper triangle of the symmetric matrix as the original feature representation of this subject. We then flatten the remaining triangle

by collapsing it in a one-dimension vector to retrieve a vector of features:

$$S = \frac{(N-1)N}{2} \quad (2)$$

where S is the dimension for the flattened vector, and N is the number of the regions in the atlas. For example, we get a 1-D vector with 19,900 features for each sample for the CC atlas. Following the previous computational process, for every individual subject, we secure three functional connectivity feature representations based on the respective three atlases.

3.4. Feature Selection

Using F-score [45], we rank all features in descending order to prioritise those with the highest discriminative power between ASD and TC subjects. Mathematically, the F-score is a ratio of variance between groups to variance within groups, quantifying each feature’s discriminatory power. A higher F-score indicates a larger difference in means relative to variability, suggesting better class distinction. We compute the F-score values for all features in the dataset. Let x_k represent the training vectors, with k ranging from 1 to m , n^+ being the count of positive instances, and n^- the count of negative instances. The F-score for the i^{th} feature is computed as follows:

$$F(i) = \frac{(\bar{r}_i^{(+)} - \bar{r}_i)^2 + (\bar{r}_i^{(-)} - \bar{r}_i)^2}{\frac{1}{n_+ - 1} \sum_{k=1}^{n_+} (r_{k,i}^{(+)} - \bar{r}_i^{(+)})^2 + \frac{1}{n_- - 1} \sum_{k=1}^{n_-} (r_{k,i}^{(-)} - \bar{r}_i^{(-)})^2} \quad (3)$$

where \bar{r}_i , $\bar{r}_i^{(+)}$, and $\bar{r}_i^{(-)}$ denote the average of the i^{th} feature for the complete dataset, the positive subset and the negative subset, respectively. The numerator captures the discriminatory power between the positive and negative subsets, while the denominator encapsulates the dispersion within each subset.

We retain the top 15% of ranked features. To determine this feature retention range, we conducted a series of tests with selections ranging from 5% to 50% based on different parcellations. The optimal parameters were determined based on the classifier’s average performance across different parcellations. Details of this experiment are presented in Appendix A.

The dimensions of the post-preprocessing data vary across different atlases; for example, features from the CC atlas have a dimensionality of 19,900, while the AAL and EZ atlases have a significantly smaller dimensionality of 6,670. To account for these differences and to maintain the 15% feature retention, we adopted an adaptive feature selection approach. We retained the top 1,000 ranked features for the AAL and EZ atlases and the top 3,000 ranked features for the CC atlas.

3.5. The MADE-for-ASD Model

The *MADE-for-ASD* model consists of two primary components, as illustrated in Figure 2. The initial component is an unsupervised learning stage employing a Stacked Sparse Denoising Autoencoder (SSDAE). Each layer is trained independently, capturing key variations in the data, with sparsity constraints promoting sparse, distributed representations.

The second component is a supervised learning stage using a Multi-Layer Perceptron (MLP). The parameters learned from the SSDAE are transferred to the MLP’s first two layers, followed by fine-tuning to enhance performance on the ASD classification task.

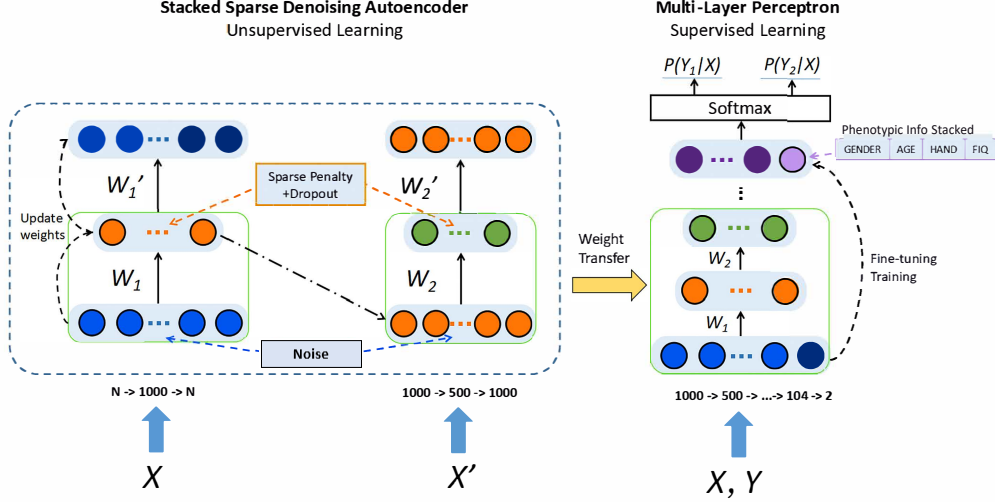


Figure 2: Overall architecture and training workflow of the deep networks of the *MADE-for-ASD* model. Knowledge from stacked sparse denoising autoencoders is transferred to a multi-layer perceptron for ASD diagnosis. Here, X refers to the input data, $Y = \{Y_1, Y_2\}$ are ASD and TC classes, and W refers to the weight parameter.

Table 1: Parameter configurations of the Autoencoders (AE1, AE2) and Multi-Layer Perceptron (MLP) of the *MADE-for-ASD* model. Here, 'n' represents the number of input features.

Parameter	AE1	AE2	MLP
Structure	n-1000-n	1000-600-1000	n-1000-500-104-2
Hidden layers	1	1	3
Learning rate	0.001	0.001	0.0005
Batch size	100	10	10
Training iteration	700	1000	200
Optimiser	GD	GD	SGD
Loss function	Cross entropy	Cross entropy	Cross entropy
Activation function	tanh	tanh	tanh
Momentum range	/	/	0.1-0.9
Dropout rate	0.5	0.5	0.3
Sparsity parameter	0.5	0.5	/
Sparse penalty	0.2	0.2	/
Noise proportion	0.3	0.1	/

3.5.1. Stacked Sparse Denoising Autoencoder

An Autoencoder (AE) consists of input, hidden and output layers, with the hidden layer containing fewer neurons. Once feature representation is obtained in an AE, it can be used to train a new AE, leading to a stacked AE with multiple layers. A Stacked Sparse Denoising Autoencoder (SSDAE) incorporates sparsity constraints and noise addition to the input data for regularisation, enhancing model robustness. The objective function includes a reconstruction

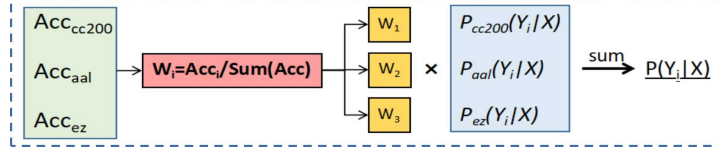


Figure 3: The multi-atlas weighted ensemble voting strategy.

loss term and a sparsity penalty term, defined as:

$$J(\Theta) = \frac{1}{m} \sum_{i=1}^m L(x^{(i)}, \hat{x}^{(i)}) + \beta \sum_{j=1}^s \text{KL}(\rho \parallel \hat{\rho}_j) \quad (4)$$

where $L(x^{(i)}, \hat{x}^{(i)})$ denotes reconstruction loss, β is the sparsity weight and $\text{KL}(\rho \parallel \hat{\rho}_j)$ is the Kullback-Leibler divergence between target sparsity ρ and average activation $\hat{\rho}_j$ of hidden unit j . The Kullback-Leibler divergence can be computed as:

$$\text{KL}(\rho \parallel \hat{\rho}_j) = \rho \log \frac{\rho}{\hat{\rho}_j} + (1 - \rho) \log \frac{1 - \rho}{1 - \hat{\rho}_j} \quad (5)$$

We employ an SSDAE with two hidden layers for unsupervised pre-training, as shown in the left part of Figure 2. Optimal model performance on the validation set is achieved using reconstruction loss (mean squared error). The input and output layers have N features, where N is the number of input features. The configurations and parameters are detailed in Table 1.

3.5.2. Transfer Learning to MLP

The SSDAE knowledge is transferred to an MLP with three hidden layers. The first two layers, with 1,000 and 500 units, inherit the SSDAE parameters, while the third layer is initialised with random weights. Four demographic features are added to the third layer as additional input. We add these into the last two layers because these features are less significant if we add them into the original input compared to the large amounts of other features. The right section of Figure 2 depicts the supervised training stage.

Fine-tuning adjusts the MLP weights to minimise prediction error in supervised tasks. The output layer consists of two units representing the likelihood of ASD or TC, using a softmax activation function to normalise the output distribution and enable the outputs to represent the corresponding probabilities belonging to a particular class. The configurations of the MLP parameters are shown in Table 1.

3.5.3. Weighed Ensemble Voting

Ensemble learning uses individual models and solves the same problem. This work adopts the bagging ensemble approach, which involves creating multiple subsets of the original data, training a model on each and combining their predictions, often through majority voting, to form a final prediction [46].

Voting can be hard or soft. In hard voting, each classifier in the ensemble votes for one class label, and the class label that gets the majority of votes is predicted. Our method is based on soft voting, as shown in Figure 3, which predicts based on the probabilities for each class label. We

assign a weight for each classifier according to their individual classification accuracy among all three classifiers:

$$w_i = \frac{\text{acc}_i}{\sum \text{acc}_i} \quad (6)$$

We then compute the sum of the products of weights and probabilities corresponding to each class across all classifiers. Subsequently, the class with the highest cumulative value is designated as the output category:

$$y = \operatorname{argmax}_{j \in Y} \sum w_i \cdot \mathbb{P}(\mathcal{F}_i(x) = j) \quad (7)$$

where $\mathbb{P}(\mathcal{F}_i(x) = j)$ is the predicted probability that instance x belongs to class j according to classifier \mathcal{F}_i .

As for the evaluation metrics, we report classification performance in terms of sensitivity and specificity in addition to accuracy because of the unbalanced nature of the dataset.

4. Experiments

4.1. Dataset

We use rs-fMRI and demographic data from ABIDE I [11], the first phase of ABIDE, with 505 autistic individuals and 530 Typical Controls (TCs). We include four key demographic features of the subjects: age (years), sex (male/female), handedness (the dominant hand; left/ambiguous/right) and full-scale IQ (overall intellectual ability) (see Table B.6 in Appendix B for their distribution in both ASD and TC classes). We also employ a subset of the T1w MRI images from the ABIDE I site, NYU Langone Medical Center, to test our proposed methodology. This subset encompasses 182 subjects, including 78 ASD and 104 TC subjects. We selected this specific subset because it contains the largest amount of data among all participatory sites in ABIDE I. Additionally, previous studies [27, 29] have reported ASD classification performance using the NYU subset, which allows us to compare our model’s performance with theirs. Refer to Table B.7 in Appendix B for details of the NYU subset having 78 autistic individuals and 104 TC subjects.

4.1.1. Preprocessing

Preprocessed ABIDE I data with four different pipelines are released under Preprocessed Connectomes Project³ [47]. The number of data varies across the resultant datasets from different pipelines. It is important to select a specific pipeline for a fair comparison with the prior works. Most of the prior works [12, 13, 22, 19, 14, 20, 21, 17, 16, 15, 9] leveraged CPAC (Configurable Pipeline for the Analysis of Connectomes) pipeline. A few studies, however, experimented with other pipelines, such as CCS (Connectome Computation System) by Dvornek et al. [18] and DPARSF (Data Processing Assistant for Resting-State fMRI) by Mahler et al. [37]. The CPAC pipeline includes several operations, such as slice timing correction, motion correction and voxel intensity normalisation. In addition, the nuisance signal was removed utilising 24 motion parameters, CompCor with five components, low-frequency drifts (linear and quadratic trends), and the global signal as regressors. Functional data underwent band-pass filtering (0.01 – 0.1 Hz) and spatial registration using a non-linear method to a template space (MNI152). We, therefore, leverage the CPAC pipeline, hereinafter referred to as the ABIDE I CPAC data.

³<http://preprocessed-connectomes-project.org/abide/>

Table 2: Classification performance of *MADE-for-ASD* model in terms of accuracy, sensitivity and specificity on the whole ABIDE I data and its subsets.

Data	# of Subjects	Accuracy (%)	Sensitivity (%)	Specificity (%)
Whole dataset	1,035	75.20	82.90	69.70
Proposed subset	645	88.71	81.82	92.50
NYU subset	175	96.40	91.00	99.50

Although Di Martino et al. [11] released the ABIDE I dataset with 1,112 samples, preprocessing and quality control reduced this number to 1,035, which is consistent with prior studies [13, 22, 19, 20]. For the NYU subset, our preprocessing and quality control reduced the number of samples from 182 to 175.

4.1.2. Subset Selection

Previous research [27, 29] demonstrated better performance with the NYU subset than the whole ABIDE I dataset. This suggests that the NYU subset may possess lower noise levels and higher data quality. Based on this, we hypothesise that using the NYU subset for training could improve the model’s ability to select high-quality data from the entire ABIDE I dataset. We, therefore, train our model on the NYU dataset using the CC atlas, save this model and repurpose it as a selector. We apply this pretrained model to classify ASD/TC instances across the whole ABIDE I dataset, which correctly classifies 645 out of 1,035 samples (around 62.3%). We separate these correctly classified samples to create a new subset, which potentially has higher quality. We refer to it as our *proposed subset* throughout this paper. Analyses on the quality of our proposed subset compared to the whole dataset and the NYU subset are presented in Appendix C.

While this subset selection aims to improve data quality, it may also lead to the selection of easier examples, potentially inflating performance metrics. Furthermore, this method may introduce exclusion criteria that could affect the generalisability of the results. A thorough analysis of the discarded subset to assess whether it excludes TC, ASD, or both is warranted, which we leave for future work.

4.2. ASD vs TC Classification

We evaluate our *MADE-for-ASD* model through experiments on different input data, ablation studies and comparative analyses with state-of-the-art methods. Following the approach of prior ASD diagnosis studies using the ABIDE I dataset [15, 17, 21, 20, 19, 13, 14, 12, 22], we employ 10-fold cross-validation to ensure robust evaluation and mitigate overfitting. This technique splits the dataset into ten equal-sized subsets, with the model being trained on nine subsets and tested on the remaining subset. This process is repeated ten times, and the classification scores are averaged across all runs, providing a cross-validated performance measure.

4.2.1. Classification Result

Table 2 reports accuracy, sensitivity and specificity on the whole ABIDE I dataset, our proposed subset and the NYU subset.

The accuracy and specificity of the whole dataset are lower than those of other subsets, which can be attributed to the noise and complexity in the whole dataset. The classifier reached its peak

Table 3: Comparison with previous studies on whole ABIDE I CPAC data in terms of classification accuracy.

Study	Model	# of Subjects	Accuracy (%)
Abraham et al. [12]	SVC	871	66.80
Heinsfeld et al. [13]	AE+DNN	1,035	70.00
Sherkatghanad et al. [22]	CNN	1,035	70.22
Eslami et al. [19]	AE+SLP	1,035	70.30
Parisot et al. [14]	GCN	871	70.40
Almuqhim and Saeed [20]	SAE+MLP	1,035	70.80
Anirudh and Thiagarajan [21]	Ensemble GCN	872	70.86
Liu et al. [17]	Extra-Trees	1,054	72.20
Wang et al. [15]	Ensemble MLP	949	74.52
Deng et al. [9]	3D-CNN	860	74.53
Ours (MADE-for-ASD)	SSDAE+MLP	1,035	75.20

Table 4: Comparison with previous studies on a subset of the ABIDE I dataset in terms of classification accuracy.

Study	Data	# of Subjects	Model	Accuracy (%)
Kong et al. [27]	NYU	182	SAE	90.30
Chen et al. [28]	Selected subset ^a	252	Random Forest	91.00
Plitt et al. [29]	NYU, USM, UCLA	178	Random Forest	95.00
Ours (MADE-for-ASD)	NYU	175	SSDAE+MLP	96.40
Ours (MADE-for-ASD)	Proposed subset	645	SSDAE+MLP	88.71

^a Subset was selected based on specific criteria of high-quality data (e.g., low head motion)

performance on the NYU subset, which can be attributed to the high quality of NYU data. This is in line with Kong et al. [27] that the classification performance is higher with the NYU subset.

4.2.2. Comparative Analysis

We compare the performance of our model trained on the whole ABIDE I dataset with the CPAC pipeline, which exhibits encouraging insights (Table 3). **(1)** Our model achieved an accuracy score of 75.20%, which is a boost of 4.4 percentage points over the best prior work [20] on 1,035 samples. **(2)** Even without considering the use of the same amount of input data, our model outperforms the 74.53% accuracy on 860 samples reported by Deng et al. [9]. **(3)** By incorporating ensemble techniques and introducing sparsity in the AE, our model outperforms other AE and DNN-based approaches [20, 19, 13] by a margin of at least 4.4 percentage points.

We further compare our work with similar studies on certain subsets of the ABIDE I dataset Table 4. Similar to the whole dataset, our proposed *MADE-for-ASD* model achieved the SOTA result on the NYU subset, showing an accuracy of 96.40%. Besides, our model’s performance on the new subset generated by our data selection achieved an accuracy of 88.71%. This proves our model’s ability to perform consistently well on different data subsets.

This result also underlines the potential influence of data quality on classification outcomes. The performance difference across subsets reinforces the need for data selection in such classification tasks. Our strategy of creating a new subset from the ABIDE I dataset has also proven to

Table 5: Classification accuracy in ablation study by removing atlases, feature selection and demographic information. Negative sign (–) refers to the removal of the corresponding component.

Component	Accuracy (%)
CC + AAL + EZ + Feature Selection + Demographic	75.20
– {AAL, EZ}	73.42
– {CC, EZ}	71.20
– {CC, AAL}	68.74
– EZ	74.10
– Feature Selection	72.80
– Demographic	73.50

be effective. The accuracy on this subset signifies the potential of using selective strategies for data preparation to boost model performance.

4.2.3. Ablation Study

As can be seen on Table 5, we carry out a series of ablation experiments to better understand the individual contributions of different components of the *MADE-for-ASD* model. These experiments include removing the ensemble voting and using single and combination of atlas data for training, abolishing the F-score-based feature selection, and omitting the sample demographic information during classification.

When we remove the voting and use only a single atlas and a combination of atlases, the classification accuracy decreases. When removing data of the AAL and EZ atlases, the accuracy drops to 73.42%. It drops further to 71.20% after removing the CC and EZ atlases. Removing the CC and AAL atlases, the accuracy reaches the lowest at 68.74%. When we only remove the EZ atlas, the accuracy is somewhat increased. This shows that integrating multiple atlases via voting contributes to the improved performance of our model. The different atlas information can complement each other, thereby enhancing the robustness of the classification task.

When we abolish the feature selection based on the F-score, it leads to a decrease in accuracy to 72.80%. The drop in accuracy demonstrates the significance of feature selection. By focusing on the most informative features and deleting the noisy features, our model can better discriminate between classes. Lastly, removing the demographic information leads to an accuracy of 73.50%. This indicates that including demographic information can aid in the classification process, further improving the decision boundaries.

Although the inclusion of demographic data leads to better performance, these data should be carefully utilised with ML-based practical ASD diagnosis systems so that the system is not biased to particular demographics. Future work should focus on quantifying and mitigating such biases in ML models predicting ASD.

While our fMRI-based approach with deep learning provides robust diagnostic performance for ASD, it is important to acknowledge the interpretability challenges associated with these methods. Traditional behaviour-based approaches offer specific insights into the symptomatic nuances of participants, which can enhance the substance of clinical reports. In contrast, our method, despite its diagnostic accuracy, may offer fewer insights into individual symptomatic specificities. This limitation is two-fold: first, there is a need for further research to improve the interpretation of fMRI data and its correlation with autism symptoms; second, the deep learning

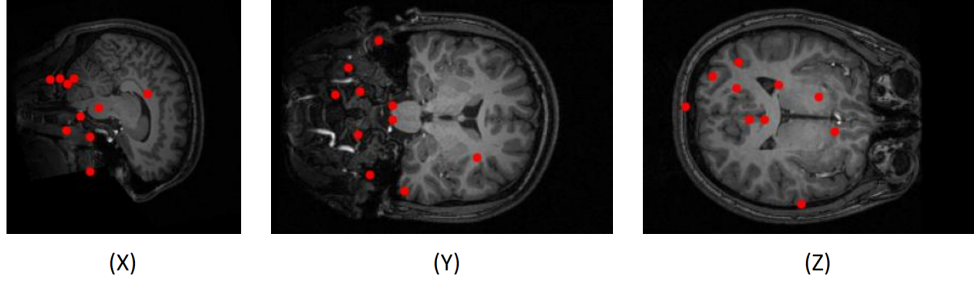


Figure 4: Top 10 most significant regions of interest towards ASD diagnosis using fMRI CC200 atlas.

model used in our study offers limited transparency, making it challenging to explain individual predictions based on input features. Future work should enhance fMRI interpretability for autism to bridge the gap between diagnostic accuracy and clinical insight.

4.3. Visualisation of Top ROIs

We utilise the F-score feature selection methodology to discern and rank the most influential features. The initial 3D images from the CC atlas are clustered based on the spatial position of ROIs. These resulting cluster centres are treated as the spatial coordinates of the ROIs. On an existing three-dimensional brain fMRI image from the dataset, we then plot the spatial coordinates of these cluster centres. We rank the appearance frequencies of the associated ROIs, thereby identifying the top 10 most frequently occurring areas of interest (Figure 4).

As can be seen, the *precuneus*, typically recognised as a central node of the default mode network (DMN) [48], has a crucial impact on ASD classification. This region is involved in self-referential thought and social cognition, which are often disrupted in autistic individuals [49]. Several studies have suggested that DMN connectivity can be associated with a neurophenotype of ASD [16]. For example, Chen et al. [28] highlighted substantial contributions from default mode and somatosensory areas toward ASD diagnosis. Similarly, Abraham et al. [12] found distinguishing connections in the DMN related to ASD diagnosis using the ABIDE dataset.

The anterior *cingulate/ventromedial* prefrontal cortex, a region with established connections to autism, was notably pronounced in the ASD classification problem [50]. This area is associated with emotional regulation and decision-making, processes that are often impaired in ASD [11]. Furthermore, anomalies in the medial prefrontal cortex node of the DMN have been shown to detect social deficits in autistic children [51]. Additionally, the left parietal cortex was stressed for ASD prediction, aligning with the lateralised activation seen in this region in autistic individuals [52]. These findings from the visualisation of top ROIs corroborate previous studies that highlight the importance of these regions in ASD pathology [16].

5. Conclusion

ASD is a neurodevelopmental disorder characterised by a spectrum of symptoms and impairments. Common features of ASD include challenges with social interaction and communication, alongside a preference for repetitive behaviours and interests, highlighting the diverse nature of this condition. This paper proposes a novel ASD diagnosis framework, *MADE-for-ASD*, involving a weighted ensemble of DNNs using multi-atlas brain fMRI data. Through the F-score-based

feature selection method, we obtain discriminative features that offer valuable visual insights into significant ROIs associated with ASD. They shed light on the interplay of different features and their respective contribution towards ASD diagnosis. This would help clinicians and researchers gain a more intuitive understanding of how different brain regions contribute to ASD. Our model consists of an SSDAE and an MLP, complemented by integrating demographic information, which significantly enhances our model’s predictive capabilities. Furthermore, our method of selecting high-quality classification subsets serves to reduce dataset noise and improve data quality. Our model achieves the SOTA accuracy on both the whole ABIDE I dataset (an improvement of 4.4 percentage points) and its subset. Such an imaging-based ASD prediction system can benefit patients, families and healthcare systems worldwide through objective, efficient, non-intrusive and early diagnosis.

Appendix A. Feature Selection using F-score

Figure A.5 illustrates our feature selection process based on F-score, showing that the classification accuracy was highest when using the top 15% of the total features. A histogram of F-scores for each atlas can reveal the decay and range of feature values, which should be investigated in future work.

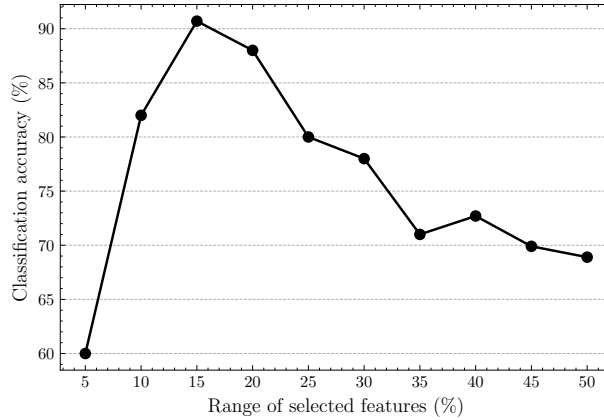


Figure A.5: ASD classification accuracy with different feature selection ranges in our F-score-based feature selection. The presented accuracy refers to the average accuracy over all three parcellations.

Appendix B. ABIDE I Dataset Statistics

The demographic details and sample distribution of the ABIDE I whole dataset and the NYU subset are presented in Table B.6 and Table B.7, respectively.

Appendix C. Subset Quality Analysis

We analyse the age distribution in the NYU subset, our selected subset and the whole dataset. As can be seen in Figure C.6, samples from younger individuals manifest a superior edge in the

Table B.6: Distribution of demographic information of the ABIDE I subjects.

Site	ASD				TC			
	Age	Sex	Hand	FIQ	Age	Sex	Hand	FIQ
CALTECH	27.4 (10.3)	15/4	0/5/14	108.2 (12.2)	28.0 (10.9)	14/4	1/3/14	114.8 (9.3)
CMU	26.4 (5.8)	11/3	1/1/12	114.5 (11.2)	26.8 (5.7)	10/3	0/1/12	114.6 (9.3)
KKI	10.0 (1.4)	16/4	1/3/16	97.9 (17.1)	10.0 (1.2)	20/8	1/3/24	112.1 (9.2)
LEUVEN	17.8 (5.0)	26/3	3/0/26	109.4 (12.6)	18.2 (5.1)	29/5	4/0/30	114.8 (12.4)
MAX_MUN	26.1 (14.9)	21/3	2/0/22	109.9 (14.2)	24.6 (8.8)	27/1	0/0/28	111.8 (9.1)
NYU	14.7 (7.1)	65/10	NaN	107.1 (16.3)	15.7 (6.2)	74/26	NaN	113.0 (13.3)
OHSU	11.4 (2.2)	12/0	1/0/11	106.0 (20.1)	10.1 (1.1)	14/0	0/0/14	115.0 (10.7)
OLIN	16.5 (3.4)	16/3	4/0/15	112.6 (17.8)	16.7 (3.6)	13/2	2/0/13	113.9 (16.0)
PITT	19.0 (7.3)	25/4	3/1/25	110.2 (14.3)	18.9 (6.6)	23/4	1/1/25	110.1 (9.2)
SBL	35.0 (10.4)	15/0	NaN	NaN	33.7 (6.6)	15/0	NaN	NaN
SDSU	14.7 (1.8)	13/1	1/0/13	111.4 (17.4)	14.2 (1.9)	16/6	3/0/19	108.1 (10.3)
STANFORD	10.0 (1.6)	15/4	3/1/15	110.7 (15.7)	10.0 (1.6)	16/4	0/2/18	112.1 (15.0)
TRINITY	16.8 (3.2)	22/0	0/0/22	108.9 (15.2)	17.1 (3.8)	25/0	0/0/25	112.5 (9.2)
UCLA	13.0 (2.5)	48/6	6/0/48	100.4 (13.4)	13.0 (1.9)	38/6	4/0/40	106.4 (11.1)
UM	13.2 (2.4)	57/9	7/8/51	105.5 (17.1)	14.8 (3.6)	56/18	9/2/63	108.2 (9.7)
USM	23.5 (8.3)	46/0	NaN	99.7 (16.4)	21.3 (8.4)	25/0	NaN	115.4 (14.8)
YALE	12.7 (3.0)	20/8	5/0/23	94.6 (21.2)	12.7 (2.8)	20/8	4/0/24	105.0 (17.1)

Age: Average (Standard Deviation), Sex: Male/Female, FIQ: Average (Standard Deviation)

Hand: Left/Ambiguous/Right; the dominant hand of samples

Table B.7: Class-wise statistics of the NYU subset.

Class	Size	Age	Gender (F/M)
ASD	78	14.54 ± 5.29	10/68
TC	104	15.87 ± 5.04	26/78

classification tasks. According to the higher classification accuracy on the NYU subset, the nature of these results suggests that age could play a pivotal role in shaping the outcomes of classification tasks. Given these findings, further exploration into the impact of age distribution on ASD classification tasks may be discussed.

Our selected subset demonstrates a reduction in noise and variance compared to the whole dataset (Figure C.7). This reduction potentially suggests higher-quality data because lower variance can enable the model to learn patterns more effectively with more accurate predictions. This aligns with findings that reduced data variability can enhance model performance [53].

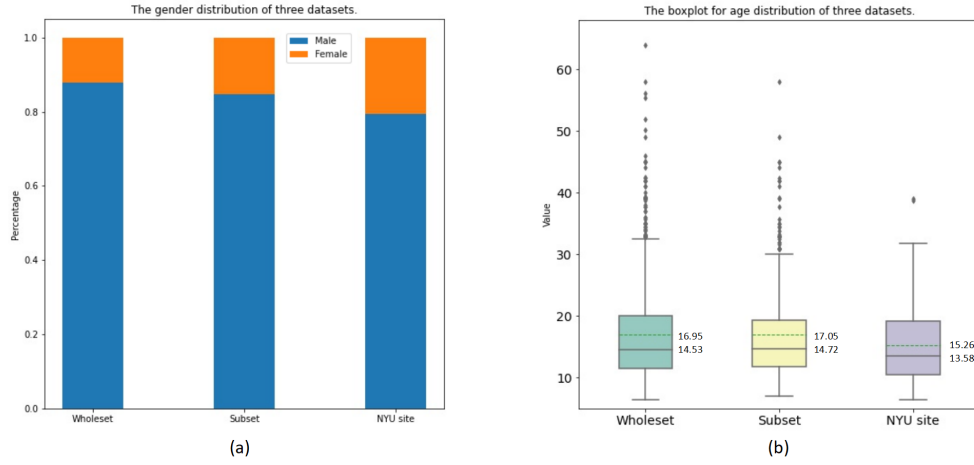


Figure C.6: Demographic information distribution on different ranges of datasets

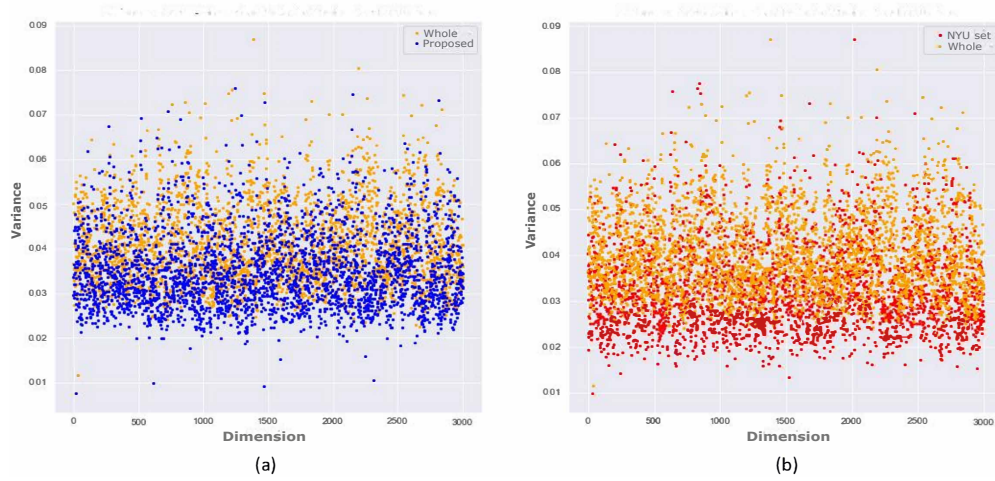


Figure C.7: Comparison of features variance between the whole dataset and (a) our proposed subset and (b) NYU subset.

References

- [1] D. G. Amaral, C. M. Schumann, C. W. Nordahl, Neuroanatomy of autism, *Trends in neurosciences* 31 (2008) 137–145.
- [2] Y. S. Kim, B. L. Leventhal, Y.-J. Koh, E. Fombonne, E. Laska, E.-C. Lim, K.-A. Cheon, S.-J. Kim, Y.-K. Kim, H. Lee, et al., Prevalence of autism spectrum disorders in a total population sample, *American Journal of Psychiatry* 168 (2011) 904–912.
- [3] J. Zeidan, E. Fombonne, J. Scorsia, A. Ibrahim, M. S. Durkin, S. Saxena, A. Yusuf, A. Shih, M. Elsabbagh, Global prevalence of autism: A systematic review update, *Autism Research* 15 (2022) 778–790. doi:10.1002/aur.2696.
- [4] M. Elsabbagh, G. Divan, Y.-J. Koh, Y. S. Kim, S. Kauchali, C. Marcin, C. Montiel-Nava, V. Patel, C. S. Paula, C. Wang, et al., Global prevalence of autism and other pervasive developmental disorders, *Autism research* 5 (2012) 160–179. doi:10.1002/aur.239.
- [5] J.-J. Ou, L.-J. Shi, G.-L. Xun, C. Chen, R.-R. Wu, X.-R. Luo, F.-Y. Zhang, J.-P. Zhao, Employment and financial burden of families with preschool children diagnosed with autism spectrum disorders in urban china: results from a descriptive study, *BMC psychiatry* 15 (2015) 1–8.
- [6] C. Lord, S. Risi, L. Lambrecht, E. H. Cook, B. L. Leventhal, P. C. DiLavore, A. Pickles, M. Rutter, The autism diagnostic observation schedule—generic: A standard measure of social and communication deficits associated with the spectrum of autism, *Journal of autism and developmental disorders* 30 (2000) 205–223.
- [7] C. Lord, M. Rutter, A. Le Couteur, Autism diagnostic interview-revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders, *Journal of autism and developmental disorders* 24 (1994) 659–685.
- [8] S. Timimi, D. Milton, V. Bovell, S. Kapp, G. Russell, Deconstructing diagnosis: Four commentaries on a diagnostic tool to assess individuals for autism spectrum disorders, *Autism (Birmingham, England)* 1 (2019).
- [9] J. Deng, M. R. Hasan, M. Mahmud, M. M. Hasan, K. A. Ahmed, M. Z. Hossain, Diagnosing autism spectrum disorder using ensemble 3D-CNN: A preliminary study, in: 2022 IEEE International Conference on Image Processing (ICIP), IEEE, 2022, pp. 3480–3484. doi:10.1109/ICIP46576.2022.9897628.
- [10] T. Iidaka, Resting state functional magnetic resonance imaging and neural network classified autism and control, *Cortex* 63 (2015) 55–67.
- [11] A. Di Martino, C.-G. Yan, Q. Li, E. Denio, F. X. Castellanos, K. Alaerts, J. S. Anderson, M. Assaf, S. Y. Bookheimer, M. Dapretto, et al., The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism, *Molecular psychiatry* 19 (2014) 659–667.
- [12] A. Abraham, M. P. Milham, A. Di Martino, R. C. Craddock, D. Samaras, B. Thirion, G. Varoquaux, Deriving reproducible biomarkers from multi-site resting-state data: An autism-based example, *NeuroImage* 147 (2017) 736–745. doi:10.1016/j.neuroimage.2016.10.045.
- [13] A. S. Heinsfeld, A. R. Franco, R. C. Craddock, A. Buchweitz, F. Meneguzzi, Identification of autism spectrum disorder using deep learning and the abide dataset, *NeuroImage: Clinical* 17 (2018) 16–23. doi:10.1016/j.nicl.2017.08.017.
- [14] S. Parisot, S. I. Ktena, E. Ferrante, M. Lee, R. Guerrero, B. Glocker, D. Rueckert, Disease prediction using graph convolutional networks: application to autism spectrum disorder and alzheimer’s disease, *Medical image analysis* 48 (2018) 117–130. doi:10.1016/j.media.2018.06.001.
- [15] Y. Wang, J. Wang, F.-X. Wu, R. Hayrat, J. Liu, AIMAF: Autism spectrum disorder identification with multi-atlas deep feature representation and ensemble learning, *Journal of neuroscience methods* 343 (2020) 108840. doi:10.1016/j.jneumeth.2020.108840.
- [16] M. Khosla, K. Jamison, A. Kucyeski, M. R. Sabuncu, Ensemble learning with 3d convolutional neural networks for functional connectome-based prediction, *NeuroImage* 199 (2019) 651–662.
- [17] Y. Liu, L. Xu, J. Li, J. Yu, X. Yu, Attentional connectivity-based prediction of autism using heterogeneous rs-fMRI data from cc200 atlas, *Experimental Neurobiology* 29 (2020) 27–37. doi:10.5607/en.2020.29.1.27.
- [18] N. C. Dvornek, P. Ventola, 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), IEEE, 2018, pp. 725–728.
- [19] T. Eslami, V. Mirjalili, A. Fong, A. R. Laird, F. Saeed, ASD-DiagNet: a hybrid learning approach for detection of autism spectrum disorder using fmri data, *Frontiers in neuroinformatics* 13 (2019) 70. doi:10.3389/fninf.2019.00070.
- [20] F. Almuqhim, F. Saeed, ASD-SAENet: a sparse autoencoder, and deep-neural network model for detecting autism spectrum disorder (asd) using fmri data, *Frontiers in Computational Neuroscience* 15 (2021) 654315. doi:10.3389/fncom.2021.654315.
- [21] R. Anirudh, 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), IEEE, 2019, pp. 3197–3201. doi:10.1109/ICASSP.2019.8683547.

- [22] Z. Sherkatghanad, M. Akhondzadeh, S. Salari, M. Zomorodi-Moghadam, M. Abdar, U. R. Acharya, R. Khosrowabadi, V. Salari, Automated detection of autism spectrum disorder using a convolutional neural network, *Frontiers in neuroscience* 13 (2020) 1325. doi:10.3389/fnins.2019.01325.
- [23] B. Crosson, A. Ford, K. M. McGregor, M. Meinzer, S. Cheshkov, X. Li, D. Walker-Batson, R. W. Briggs, Functional imaging and related techniques: an introduction for rehabilitation researchers, *Journal of rehabilitation research and development* 47 (2010) vii.
- [24] M. Greicius, Resting-state functional connectivity in neuropsychiatric disorders, *Current opinion in neurology* 21 (2008) 424–430.
- [25] Y. Du, Z. Fu, V. D. Calhoun, Classification and prediction of brain disorders using functional connectivity: promising but challenging, *Frontiers in neuroscience* 12 (2018) 525.
- [26] A. Di Martino, D. O’connor, B. Chen, K. Alaerts, J. S. Anderson, M. Assaf, J. H. Balsters, L. Baxter, A. Beggiato, S. Bernaerts, et al., Enhancing studies of the connectome in autism using the autism brain imaging data exchange ii, *Scientific data* 4 (2017) 1–15. doi:10.1038/sdata.2017.10.
- [27] Y. Kong, J. Gao, Y. Xu, Y. Pan, J. Wang, J. Liu, Classification of autism spectrum disorder by combining brain connectivity and deep neural network classifier, *Neurocomputing* 324 (2019) 63–68.
- [28] C. P. Chen, C. L. Keown, A. Jahedi, A. Nair, M. E. Pflieger, B. A. Bailey, R.-A. Müller, Diagnostic classification of intrinsic functional connectivity highlights somatosensory, default mode, and visual regions in autism, *NeuroImage: Clinical* 8 (2015) 238–245.
- [29] M. Plitt, K. A. Barnes, A. Martin, Functional connectivity classification of autism identifies highly predictive brain features but falls short of biomarker standards, *NeuroImage: Clinical* 7 (2015) 359–366.
- [30] M. A. Aghdam, A. Sharifi, M. M. Pedram, Diagnosis of autism spectrum disorders in young children based on resting-state functional magnetic resonance imaging data using convolutional neural networks, *Journal of Digital Imaging* 32 (2019) 899–918. doi:10.1007/s10278-019-00196-1.
- [31] N. Payakachat, J. M. Tilford, W. J. Ungar, National database for autism research (ndar): big data opportunities for health services research and health technology assessment, *Pharmacoeconomics* 34 (2016) 127–138. doi:10.1007/s40273-015-0331-6.
- [32] G. Li, M. Liu, Q. Sun, D. Shen, L. Wang, Early diagnosis of autism disease by multi-channel cnns, in: *Machine Learning in Medical Imaging: 9th International Workshop, MLMI 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 16, 2018, Proceedings 9*, Springer, 2018, pp. 303–309. doi:10.1007/978-3-030-00919-9_35.
- [33] M. Khodatars, A. Shoeibi, D. Sadeghi, N. Ghaasemi, M. Jafari, P. Moridian, A. Khadem, R. Alizadehsani, A. Zare, Y. Kong, A. Khosravi, S. Nahavandi, S. Hussain, U. R. Acharya, M. Berk, Deep learning for neuroimaging-based diagnosis and rehabilitation of autism spectrum disorder: A review, *Computers in Biology and Medicine* 139 (2021) 104949. doi:10.1016/j.combiomed.2021.104949.
- [34] M. P. Van Den Heuvel, H. E. H. Pol, Exploring the brain network: a review on resting-state fmri functional connectivity, *European neuropsychopharmacology* 20 (2010) 519–534.
- [35] Z. Khandan Khadem-Reza, M. A. Shahram, H. Zare, Altered resting-state functional connectivity of the brain in children with autism spectrum disorder, *Radiological Physics and Technology* 16 (2023) 284–291. doi:10.1007/s12194-023-00717-2.
- [36] F. E.-z. Bazay, A. Drissi El Maliani, Assessing the impact of preprocessing pipelines on fmri based autism spectrum disorder classification: Abide ii results, in: *International Conference on Engineering Applications of Neural Networks*, Springer, 2024, pp. 463–477. doi:10.1007/978-3-031-62495-7_35.
- [37] L. Mahler, Q. Wang, J. Steiglechner, F. Birk, S. Heczko, K. Scheffler, G. Lohmann, Pretraining is all you need: A multi-atlas enhanced transformer framework for autism spectrum disorder classification, in: *International Workshop on Machine Learning in Clinical Neuroimaging*, Springer, 2023, pp. 123–132. doi:10.1007/978-3-031-44858-4_12.
- [38] R. C. Craddock, G. James, P. E. Holtzheimer III, X. P. Hu, H. S. Mayberg, A whole brain fmri atlas generated via spatially constrained spectral clustering, *Human Brain Mapping* 33 (2012) 1914–1928. doi:10.1002/hbm.21333.
- [39] R. M. Thomas, S. Gallo, L. Cerliani, P. Zhutovsky, A. El-Gazzar, G. van Wingen, Classifying autism spectrum disorder using the temporal statistics of resting-state functional mri data with 3d convolutional neural networks, *Frontiers in Psychiatry* 11 (2020).
- [40] F. Z. Subah, K. Deb, A comprehensive study on atlas-based classification of autism spectrum disorder using functional connectivity features from resting-state functional magnetic resonance imaging, in: *Neural engineering techniques for Autism spectrum disorder*, Elsevier, 2023, pp. 269–296. doi:10.1016/B978-0-12-824421-0.00021-7.
- [41] E. Yee, Identifying neural patterns and biomarkers of asd through multi-phase resting-state functional mri analysis, in: *2023 11th International Conference on Bioinformatics and Computational Biology (ICBCB)*, 2023, pp. 147–150. doi:10.1109/ICBCB57893.2023.10246705.
- [42] S. B. Eickhoff, K. E. Stephan, H. Mohlberg, C. Grefkes, G. R. Fink, K. Amunts, K. Zilles, A new spm toolbox for

- combining probabilistic cytoarchitectonic maps and functional imaging data, *NeuroImage* 25 (2005) 1325–1335. doi:10.1016/j.neuroimage.2004.12.034.
- [43] A. R. T. Donders, G. J. Van Der Heijden, T. Stijnen, K. G. Moons, A gentle introduction to imputation of missing values, *Journal of clinical epidemiology* 59 (2006) 1087–1091.
 - [44] M. J. Azur, E. A. Stuart, C. Frangakis, P. J. Leaf, Multiple imputation by chained equations: what is it and how does it work?, *International Journal of Methods in Psychiatric Research* 20 (2011) 40–49. doi:10.1002/mpr.329.
 - [45] Y.-W. Chen, C.-J. Lin, Combining svms with various feature selection strategies, *Feature extraction: foundations and applications* (2006) 315–324.
 - [46] L. Rokach, Ensemble-based classifiers, *Artificial intelligence review* 33 (2010) 1–39.
 - [47] C. Craddock, Y. Benhajali, C. Chu, F. Chouinard, A. Evans, A. Jakab, B. S. Khundrakpam, J. D. Lewis, Q. Li, M. Milham, et al., The neuro bureau preprocessing initiative: open sharing of preprocessed neuroimaging data and derivatives, *Frontiers in Neuroinformatics* 7 (2013) 27.
 - [48] A. V. Utevsky, D. V. Smith, S. A. Huettel, Precuneus is a functional core of the default-mode network, *Journal of Neuroscience* 34 (2014) 932–940.
 - [49] V. L. Cherkassky, R. K. Kana, T. A. Keller, M. A. Just, Functional connectivity in a baseline resting-state network in autism, *NeuroReport* 17 (2006) 1687–1690. doi:10.1097/01.wnr.0000239956.45448.4c.
 - [50] T. Watanabe, N. Yahata, O. Abe, H. Kuwabara, H. Inoue, Y. Takano, N. Iwashiro, T. Natsubori, Y. Aoki, H. Takao, et al., Diminished medial prefrontal activity behind autistic social judgments of incongruent information, *PloS one* 7 (2012) e39561.
 - [51] V. Menon, Developmental pathways to functional brain networks: emerging principles, *Trends in Cognitive Sciences* 17 (2013) 627–640. doi:10.1016/j.tics.2013.09.015.
 - [52] H. Koshino, P. A. Carpenter, N. J. Minshew, V. L. Cherkassky, T. A. Keller, M. A. Just, Functional connectivity in an fmri working memory task in high-functioning autism, *Neuroimage* 24 (2005) 810–821.
 - [53] P. Domingos, A few useful things to know about machine learning, *Communications of the ACM* 55 (2012) 78–87. doi:10.1145/2347736.2347755.