



Classification of Preschoolers with Low-Functioning Autism Spectrum Disorder Using Multimodal MRI Data

Johanna Inhyang Kim¹ · Sungkyu Bang² · Jin-Ju Yang² · Heejin Kwon³ · Soomin Jang⁴ · Sungwon Roh^{1,5} · Seok Hyeon Kim^{1,5} · Mi Jung Kim⁶ · Hyun Ju Lee⁷ · Jong-Min Lee² · Bung-Nyun Kim^{4,8}

Accepted: 5 November 2021

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract

Multimodal imaging studies targeting preschoolers and low-functioning autism spectrum disorder (ASD) patients are scarce. We applied machine learning classifiers to parameters from T1-weighted MRI and DTI data of 58 children with ASD (age 3–6 years) and 48 typically developing controls (TDC). Classification performance reached an accuracy, sensitivity, and specificity of 88.8%, 93.0%, and 83.8%, respectively. The most prominent features were the cortical thickness of the right inferior occipital gyrus, mean diffusivity of the middle cerebellar peduncle, and nodal efficiency of the left posterior cingulate gyrus. Machine learning-based analysis of MRI data was useful in distinguishing low-functioning ASD preschoolers from TDCs. Combination of T1 and DTI improved classification accuracy about 10%, and large-scale multi-modal MRI studies are warranted for external validation.

Keywords Autism spectrum disorder · Machine learning · Preschool · T1-weighted magnetic resonance imaging · Diffusion tensor imaging

Introduction

Autism spectrum disorder (ASD) is a neurodevelopmental disorder characterized by social communication deficits and the presence of restricted and repetitive behaviors. Although clinical signs of ASD are detectable at an age of as early as

12 months, the global average age at the time of diagnosis is 60.48 months (range 30.90–234.57 months), according to a recent meta-analysis (van't Hof et al., 2021). The Centers for Disease Control and Prevention has estimated the median age of ASD diagnosis to be 51 months (range 38–57 months) in the US (Baio et al., 2018). The global shortage of well-qualified professionals who specialize in ASD contributes to the delay in diagnosis (Huang et al., 2020), and the current

Johanna Inhyang Kim and Sungkyu Bang have contributed equally and Shared first authorship.

✉ Jong-Min Lee
ljm@hanyang.ac.kr

✉ Bung-Nyun Kim
kbn1@snu.ac.kr

¹ Department of Psychiatry, Hanyang University Medical Center, 222-1 Wangsimni-ro, Sungdong-gu, Seoul 04763, Republic of Korea

² Department of Biomedical Engineering, Hanyang University, 222 Wangsimni-ro, Sungdong-gu, Seoul 04763, Republic of Korea

³ Department of Psychology, Yonsei University, 50 Yonsei-ro, Seodaemun-gu, Seoul 02722, Republic of Korea

⁴ Department of Psychiatry, Seoul National University College of Medicine, 101 Daehak-ro, Jongno-gu, Seoul 03080, Republic of Korea

⁵ Department of Psychiatry, Hanyang University College of Medicine, 222 Wangsimni-ro, Seongdong-gu, Seoul 04763, Republic of Korea

⁶ Department of Rehabilitation Medicine, Hanyang University College of Medicine, 222 Wangsimni-ro, Sungdong-gu, Seoul 04763, Republic of Korea

⁷ Department of Pediatrics, Hanyang University College of Medicine, 222 Wangsimni-ro, Sungdong-gu, Seoul 04763, Republic of Korea

⁸ Department of Psychiatry and Behavioral Science, Seoul National University College of Medicine, 101 Daehak-ro, Chongno-gu, Seoul 03080, Republic of Korea

clinical diagnosis process is also hindered by subjectivity and reporter-dependency (Xiao et al., 2017). Identification and implementation of early biomarkers in clinical practice may help enhance early detection and reduce misdiagnoses of ASD, which would ultimately lead to better treatment outcomes and fairer prognosis (Akshoomoff et al., 2004). The use of machine learning-based methods has been increasing applied to increase the diagnostic accuracy and speed, and most studies have used T1-weighted (T1w) magnetic resonance imaging (MRI) or functional MRI (fMRI) features as input features. A recent meta-analysis of machine learning-based studies that used T1w MRI data reported a sensitivity of 0.83 and specificity of 0.84; comparatively, a subgroup meta-analysis of fMRI data showed that the sensitivity and specificity ranged from 0.69 to 0.66 (Moon et al., 2019). However, the studies that were included in this meta-analysis were limited by high heterogeneity, risk of overfitting, and use of imaging data acquired with a single modality.

Studies on neurological underpinnings of ASD have suggested that the brain of patients with ASD is characterized by widespread abnormalities such as altered cortical anatomy, abnormal white-matter integrity, and altered brain function and connectivity (Libero et al., 2015). Because of the complexity and multilayered properties of the brain of patients with ASD, unimodal MRI studies may not be sufficient to provide a comprehensive profile of ASD neurobiology. Although multimodal classifiers can help explore multivariate dimensions and provide a rich multiparametric marker with increased accuracy (Ingalhalikar et al., 2014), such studies have been strikingly limited in number (Zhuang et al., 2019). To our knowledge, only one study has combined T1w MRI and diffusion tensor imaging (DTI) data for classification. In that study, Libero et al. (2015) used T1w MRI, DTI, and magnetic resonance spectroscopy (MRS) data and found that the highest accuracy of 91.9% was achieved with fractional anisotropy (FA), radial diffusivity (RD), and cortical thickness as key predictors. However, this study was limited by its small sample size (19 patients with ASD and 18 controls) and by targeting of only adults with high-functioning ASD (HFA), which affected the generalizability of its results to younger patients or those with low-functioning ASDs (LFAs) (Libero et al., 2015).

The young childhood age (age 3–6 years) group and the LFA group are understudied populations in ASD neuroimaging studies. This period is critical as it is close to the time of their clinical diagnosis and is a time of rapid and dynamic brain growth (Shen et al., 2016). As it precedes the age when behavioral treatments or medications influence the consolidation and adaptation of neural networks, the neural connectivity at this age may more closely reflect the emerging diagnostic features of ASD (Shen et al., 2016). The intelligence quotient (IQ) is known to be a major prognostic predictive factor, as verbal IQ was shown to be a significant predictor

of autism trajectory in a longitudinal study from the age of 2 to 15 years (Gotham et al., 2012), and pretreatment IQ levels were found to show less improvement following treatment (Ameis & Catani, 2015). ASD with an intellectual disability accounts for 31% of all ASD cases and an additional 23% function in the borderline range (Alloway, 2010). Only 11% of the participants listed in the ABIDE database, which is the most widely studied ASD open dataset, were found to have an IQ of ≤ 85 (Di Martino et al., 2014). Despite the clinical significance of this subpopulation, the LFA preschooler group has been neglected in neuroimaging studies owing to difficulties associated with cooperation during long MRI scans. Only 5 of 59 ASD-related DTI studies highlighted in a review by Ameis and Catani included ASD samples with a mean age of < 5 years (Andrews et al., 2019). Only five voxel-based morphometry (VBM) studies considered each individual's cognitive function by separating the whole sample into low-functioning autism and high-functioning autism groups (Cai et al., 2018).

The purpose of this study was twofold: (1) to classify LFA preschoolers and age- and sex-matched controls by applying machine learning methods to MRI data and (2) to examine whether multimodal data from both T1w MRI and DTI data is more effective and accurate than unimodal MRI data. We hypothesized that classification accuracy of LFA preschoolers could be achieved using MRI data, and the accuracy would be higher when data from both modalities are used than when those from a single modality.

Methods

Participants

A total of 58 individuals with ASD and 48 typically developing controls (TDCs) aged between 3 and 6 years were enrolled between October 2015 and September 2019. Patients with ASD visited the Child and Adolescent Psychiatry outpatient clinic at the Seoul National University Hospital (SNUH). For The TDC group, participants were recruited from the Seoul National University Hospital clinic and Hanyang University Medical Center (HYUMC) Pediatrics Department. From the SNUH clinic, 72 patients with ASD and 24 TDCs were recruited, among whom six patients with ASD and one TDC were excluded owing to poor image quality, and eight patients with ASD were excluded owing to missing IQ data. Among the 29 TDCs from HYUMC, four were excluded owing to poor image quality. The diagnosis of ASD was confirmed according to the criteria of the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, by board-certified child and adolescent psychiatrists, and presence of comorbidities was verified using the Kiddie Schedule for Affective Disorders and

Schizophrenia—Present and Lifetime version (K-SADS-PL Kaufman et al., 1997; Kim et al., 2004)). We also applied the Autism Diagnostic Observation Schedule (ADOS) to confirm ASD diagnosis. We only included participants with an IQ of < 70 for the analyses, as this was a definition of LFA used in several previous studies (Cai et al., 2018; Erbetta et al., 2015; Reiter et al., 2018). We also defined a subgroup of LFA with ADOS total scores of 14–23 for the sensitivity analyses, as a previous study identified a LFA group of ADOS score of this range (Lord et al., 2000). The exclusion criteria for ASD were as follows: a hereditary genetic disorder; current or past history of brain trauma, organic brain disorder, seizure, or any neurological disorder; schizophrenia or any other childhood-onset psychotic disorder; major depressive disorder or bipolar disorder; Tourette’s syndrome or a chronic motor/vocal tic disorder; obsessive–compulsive disorder; and/or a history of antipsychotic or antidepressant treatment lasting for > 1 year or within the 4-week period before the initiation of the study. The TDC group included typically developing children with no signs of developmental delay.

Written informed consent was obtained from all parents/guardians, and the children provided verbal assent to participate after receiving an explanation of the study prior to enrollment. All study protocols were approved by the Institutional Review Board of SNUH and HYUMC.

Clinical Assessment

The ADOS is a semi-structured assessment of communication, social interaction, play and stereotyped behaviors, and restricted interests. It is currently considered the gold standard for diagnosing ASD (Lord et al., 2000). The examiner administered one of four modules according to the participant’s language level and age (ranging from nonverbal young children through verbally fluent adults). The IQ of participants was assessed using the Korean Educational Developmental Institute’s Wechsler Intelligence Scale for Children (Park et al., 1996) or the Korean Leiter International Performance scale—Revised (for nonverbal participants) (Shin & Cho, 2010).

MRI Acquisition

Neuroimaging data for psychiatric disorders were collected from two sites because of the limited capacity of a single site. We acquired both T1w MRI and DTI data using two whole-body 3 T magnetic resonance systems (Siemens Magnetom Trio Tim Syngo, Germany; Philips Achieve, Best, Netherlands). Due to the difficulty in obtaining MRI data in young ASD patients, we sedated all participants during the MRI scans, with the approval of the IRB. To reduce discomfort from loud noises, earplugs, headphones, or both were

applied. The sedation and details of the MRI procedure were explained to the parents, and parents were informed to skip nap-time on the day of the procedure. Sedation was administered after an appropriate interval of fasting before sedation. Chloral hydrate was selected (dosage: 50 mg/kg; maximum dose: 1 g) due to its low complication rates and high efficacy when used in accordance with the guidelines for children’s sedation as published by the American Academy of Pediatrics (Committee on Drugs, American Academy of, 2002; Vade et al., 1995). All children were carefully monitored using pulse oximetry during the scans. The imaging session was interrupted if the child moved or woke up during scanning. After the imaging procedure was completed, discharge was permitted only when the participants fully recovered from the sedation effects.

At the SNUH, a three-dimensional (3D) T1w magnetization-prepared rapid acquisition gradient-echo (MPRAGE) scan was performed using a 3 T Siemens system with an 8-channel head coil. At the HYUMC, the T1w MRI was acquired using a 3 T Philips system equipped with a 16-channel head coil. The image parameters were as follows: 1) for the Siemens system, repetition time (TR) = 1900 ms, time to echo (TE) = 3.07 ms, inversion time (TI) = 900 ms, flip angle = 9°, voxel size = 0.9 × 0.9 × 0.9 mm³, field of view (FOV) = 230 mm², slice thickness = 0.9 mm, and slice numbers = 176; 2) for the Philips system, TR = 8.3 ms, TE = 4.6 ms, TI = 1 ms, flip angle = 8°, voxel size = 0.9 × 0.9 mm², FOV = 224 mm², slice thickness = 1, and slice numbers = 150. Subsequently, 3D-DTI images with a single-shot echo-planar imaging (EPI) sequence were acquired using the follows parameters: 1) for the Siemens system, TR = 10,000 ms, TE = 88 ms, flip angle = 0°, voxel size = 1.9 mm², field of view (FOV) = 240 mm², slice thickness = 3.5 mm, slice numbers = 50, b-values = 0 and 700 s/mm², and direction = 30; 2) for the Philips system, TR = 8192 ms, TE = 76 ms, flip angle = 90°, voxel size = 2.0 mm², FOV = 224 mm², slice thickness = 2 mm, slice numbers = 74, b-values = 0 and 800 s/mm², and direction = 30.

Multiple Feature Extraction with Multimodal MRI Data

We extracted multiple features from both T1w MRI and DTI data to fully capture the structural characteristics of the brain of patients with ASD. The features are summarized in Table S1.

Cortical Features Based on Surface Modeling

Cortical features derived from structural T1w MRI have been adopted to distinguish individuals with ASD from normal controls (Li et al., 2017; Pagnozzi et al., 2018). Based on 3D cortical surface modeling, we estimated a variety of

cortical features including cortical thickness, surface area, cortical volume, mean curvature, and gyrification index. T1w MRI was processed in several steps for bias field correction (Sled et al., 1998), brain extraction (Smith, 2002), spatial normalization (Collins, Neelin, Peters, & Evans, 1994), brain tissue segmentation (Zijdenbos et al., 2002) and surface construction for each hemisphere (MacDonald et al., 2000; Robbins et al., 2004). Two cortical surfaces, the inner and outer surface, are the boundary parameterized polygonal 3D meshes of gray/white matter and cerebrospinal fluid/gray matter tissues. After surface reconstruction, the cortical thickness was calculated as the Euclidean distance between the inner and outer cortical surfaces of linked vertices. The surface area was computed by the Voronoi area assigned to each vertex, and the area making up the surface model, which is considered as the overall degree of folding, was summed. The cortical volume was measured as the amount of gray matter volume multiplied by the cortical thickness and surface area at the vertex of the surface. The mean curvature has been considered to reflect the local complexity of cortical folding and measured at the vertices that lay within the sulcal regions on the smoothed surface. We adopted multiple mean values within given region of interests (ROIs) from 78 pre-defined cortical regions, defined by the well-received automated anatomical labeling (AAL) atlas template (Tzourio-Mazoyer et al., 2002), to obtain representative values that can be used as input variables in the ASD classifier. The gyrification index originally proposed by Zilles et al. (1988) was calculated as the 3D surface-area ratio between the outer hull surface and the cortical outer surface for each hemisphere, which means that a higher gyrification index shows a greater cortical folding. These procedures were performed using the CIVET software version 2.1.0 (Montreal Neurological Institute, McGill University; <https://github.com/aces/CIVET>), and technical details are described in a previous study (Im et al., 2006, 2008; Yang et al., 2013).

Multiple Brain Structures Based on Volume

Previous studies have attempted to examine the variations in the brain anatomy of patients with ASD, relative to that of normal controls, using a variety of brain structures (Li et al., 2017; Pagnozzi et al., 2018). We focused on gray matter, white matter, and cerebrospinal brain volume, cerebellar volume, corpus callosum area, and subcortical gray matter volumes showing that these structural changes have been reported in many previous studies. Gray matter, white matter, and cerebrospinal fluid volume were calculated by measuring the volume of voxels within a segment of brain volume. The sum of these volumes represents the brain size as an intracranial volume (Zijdenbos et al., 2002). Cerebellar volume was calculated by measuring the segmented

cerebellar volume, which was carried out using an automated procedure that sequentially combines cerebellar tissue classification, a template-based approach, and morphological operations (Lee et al., 2015). The area of the corpus callosum was calculated by segmentation of its structure in the midsagittal plane based on a Bayesian inference using sparse representation and multi-atlas voting (Park et al., 2018). Fifteen volumes of subcortical gray matter structures were automatically segmented using the FIRST (FMRIB's Integrated Registration and Segmentation Tool; <http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FIRST>) software version 6.0.1, which is a part of the FSL (FMRIB Software Library, Oxford Centre for Functional magnetic resonance imaging (MRI) of the Brain, Oxford, UK; <http://www.fmrib.ox.ac.uk/fsl>). It involves the accumbens, amygdala, caudate, hippocampus, pallidum, putamen, and thalamus for both hemispheres and brain stem (Patenaude et al., 2011).

Multi-regional Diffusion Measures

We estimated diffusion parameters such as FA and mean diffusivity (MD) from the white-matter ROIs using the Johns Hopkins University ICBM-DTI-81 White-Matter Labels Atlas (Oishi et al., 2008), which was provided in the FSL toolbox (<https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/Atlases>). DTI data were preprocessed by motion correction, eddy current distortion, re-alignment of all scans to the b0 images, and estimation of diffusion tensor matrices by fitting a tensor model using the diffusion toolbox (FDT; <https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FDT>) of the FSL package. The preprocessed volumes were aligned onto the standard template space using affine and nonlinear registration (<https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FLIRT>; <https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FNIRT>) implemented in the FSL package and then averaged within an ROI from the white-matter atlas template. Ninety-six FA and MD values of the averaged white-matter ROIs were obtained in the processing steps.

Structural Network Features Based on Graph Theory

Graph theory analysis was conducted by generating a weighted and undirected structural brain network, which was composed of nodes and edges, to evaluate abnormalities in the brain organization of patients with ASD, relative to that of controls, using the Brain Connectivity Toolbox (<https://sites.google.com/site/bctnet/>) (Rubinov & Sporns, 2010) for network parameters. It can be represented by a symmetric 90×90 matrix, the columns or rows of the matrix indicating nodes and the elements of the matrix indicating edges connected between a pair of nodes. This approach summarizes complex brain topology down to a single number (referred as a global network parameter) or a node unit with assigned edge properties (referred as a local network parameter) and

is thus attractive because of its simplicity. We calculated global network parameters including the mean degree, mean strength, clustering coefficient, characteristic path length, global efficiency, normalized clustering coefficient, normalized characteristic path length, and smallworldness, which were commonly used in previous studies for brain network analysis (Bullmore & Sporns, 2009; Li et al., 2018; Qin et al., 2018). Normalized parameters were scaled against the mean value of graphical parameters obtained from 100 matched random graphs that preserve the same number of nodes, edges, and degree sequence (Maslov & Sneppen, 2002). Additionally, three types of local network parameters were estimated to reflect particular connections relevant to ASD, which is presumed from short-range overconnectivity or long-range underconnectivity, including shortest path length, nodal efficiency (Bullmore & Sporns, 2009; Qin et al., 2018), and maximum flow (Yoo et al., 2015). We applied the procedures described in a previous study before analyzing network parameters (Qin et al., 2018). A brief summary follows. The nodes were defined according to the AAL atlas template (Tzourio-Mazoyer et al., 2002) mentioned above, which is consisted of 90 brain regions including 78 of parcellated cortical regions and 12 subcortical regions. Network edges were assigned by mean FA averaging on structurally connected fiber tracts between a pair of nodes using whole-brain fiber tracts. Whole-brain fiber tracts were constructed by selecting all white-matter voxels as a seeding region in diffusion native space using the deterministic fiber tracking algorithm (Yeh et al., 2013) with DSI Studio (<http://dsi-studio.labsolver.org>).

Harmonization of Multimodal Imaging Features

Harmonization was applied to combine MRI data obtained using different scanners and acquisition protocols from the SNUH and HYUMC using ComBat harmonization (<https://github.com/Jfortin1/ComBatHarmonization>). The ComBat harmonization is a well-established technique, which uses a multivariate linear mixed effects regression with empirical Bayes, and has been shown to improve the estimation of the model parameters in studies involving small samples, and to remove unwanted inter-site scanner variation while preserving inter-site biological variability (Fortin et al., 2017, 2018).

Based on the literature, we collected multiple features according to their homogeneity characteristics derived from cortical and volume parameters, diffusion parameters, and network parameters as a feature map and presented them in a matrix, in which rows and columns represented the imaging features and subjects, respectively. As we presumed that the magnitude of effects of the site is not constant across the different feature metrics, three metrics were identified to examine the degree to which parameters were affected

by scanner-to-scanner variation and how ComBat harmonization was performed in each setting. Finally, three of the harmonized matrices were concatenated and used as input to the classifier.

Machine Learning Classification

We performed classification of the ASD group and the TDC group using five different machine learning techniques, including support vector machine (SVM), logistic regression (LR), random forest (RF), AdaBoost, and multi-layer perceptron (MLP), which are machine learning classifiers that were frequently applied in many previous studies (Jiao et al., 2010; Libero et al., 2015; Xiao et al., 2017; Zhou et al., 2014). All of these classifiers were implemented using Python's scikit learning library (<https://github.com/scikit-learn/scikit-learn>) (Abraham et al., 2014; Pedregosa et al., 2011), to compare the performance of the classifiers based on the features acquired from T1w MRI and DTI data. Five-fold cross-validation was used to avoid overfitting on the test set and to improve generalization. In other words, the training dataset was randomly divided into five equal-sized partitions, and then four partitions were used as the training set, and the holdout partition as the test set. This procedure was repeated five times. We optimized the model parameters for each classifier based on five-fold cross-validation. For each fold, while removing features one-by-one, the highest performance points for each classifier were measured and compared. The classification performance was evaluated by comparing accuracy, sensitivity, specificity, positive predictive value and negative predictive value. We repeated this analysis after confining the LFA group to those with ADOS scores between 14 and 23 ($n = 41$).

The recursive feature elimination (RFE) method was adopted to remove unnecessary features and to evaluate the performance of the classifiers simultaneously (Qureshi et al., 2016) because the number of observations in neuroimaging studies tends to be far lesser than the number of features leading to unreliable classifier performance (Chu et al., 2012). The RFE calculates the importance of features and removes them from non-critical features one-by-one. It records the classifier performance at every moment and shows the features with the best performance. Our strategy involved identification of important features through RFE, evaluation of classification performance with them, and comparison of the high rank features with statistically significant ones.

To calculate the importance of features identified using RFE, we used the RF method for training, as it was the classifier with the highest performance. A list of selected features at the peak performance point was also recorded. In addition, we set the rank with the features selected from each

fold. Specifically, a high rank set was counted for features that appeared repeatedly in multiple folds.

An ablation study was performed by mutually removing the features of the T1w MRI and DTI, and then we compared classification performance of single modalities to that of combining features from both T1w and DTI. The three models were validated using the classification performance of the RF classifier, such as accuracy, sensitivity, and specificity, based on five-fold cross-validation.

Statistical Analysis

The characteristics of the ASD and TDC groups were compared using two-sample t-test (continuous variables) and chi-squared test (categorical variables).

We evaluated the scanner effects between the SNUH and HYUMC sites before and after applying harmonization in each feature collection of cortical and volume measures, diffusion measures, and network parameters using two-sample t-tests. We assumed that harmonized data would show no significant differences between the TDC groups of different sites if they were well corrected.

We also compared the extracted features of the T1w MRI and DTI images between the ASD and TDC groups using two-sample t-test. We compared the features that were found to be statistically different between the two groups with the

high-ranking RF classification features. The level of statistical significance was set to false discovery rate (FDR)-corrected P-value level of 0.05.

Results

The characteristics of the participants are presented in Table 1. There were no significant differences in age or sex between the ASD and TDC groups. The mean ages of participants in the ASD and TDC group were 4.1 ± 1.0 years and 4.3 ± 1.1 years, respectively. The proportion of boys was 74% in the ASD group and 67% in the TDC group. The IQ ranges in the ASD and TDC groups were 30–86 and 72–142, respectively. None were taking psychotropic medication at the time of enrollment, and most patients underwent module 1 of ADOS, with the exception of 4 ASD children who were administered module 2. The ADOS total scores ranged from 9 to 22.

Data Harmonization

To evaluate the performance of data harmonization, we compared the data distribution of three matrices for multiple features between the TDC groups of SNUH and HYUMC before and after performing harmonization, as shown in Figure S1–S3. The statistical results indicate that the differences that existed between the TDC groups of two different sites before harmonization disappeared after the harmonization.

Classifier Performance

We compared the performance results of five classifiers in Table 2: RF, SVM, LR, AdaBoost, and MLP. The RF classifier showed the best performance with an accuracy of 88.8%; its sensitivity, specificity, positive predictive value, negative predictive value, and area under curve (AUC) were 93.0%, 83.8%, 87.9%, 89.1%, and 0.86, respectively (Fig. S4). On the contrary, the LR classifier had the lowest accuracy (75.0%) and sensitivity (71.7%). The SVM classifier had the lowest specificity (78.7%). The AdaBoost and MLP

Table 1 Demographic and clinical characteristics of the ASD and TDC groups

Characteristic	ASD ($n=58$)	TDC ($n=48$)	p
Age (years), mean (SD)	4.1 (1.0)	4.3 (1.1)	0.22
Sex (male), N (%)	43 (74.1)	33 (68.8)	0.42
Intelligence quotient, mean (SD)	52.9 (15.7)	99.94 (18.3)	<0.001
ADOS score, mean (SD)			
Communication score	5.6 (1.3)		
Social interaction score	9.3 (2.3)		
Total score	14.9 (3.1)		

Bold values indicate statistical significance ($p < 0.05$)

ASD autism spectrum disorder, TDC typically developing control, SD standard deviation, ADOS autism diagnostic observation schedule

Table 2 Comparison of the classification performance

Classifier	Accuracy (%)	Sensitivity (%)	Specificity (%)	PPV	NPV
Support vector machine (linear)	78.7	83.3	72.9	79.3	78.9
Logistic regression classifier	75.0	71.7	79.1	82.1	69.8
Random forest classifier	88.8	93.0	83.8	87.9	89.1
AdaBoost classifier	85.3	86.7	83.6	87.3	84.5
Multi-layer perceptron	82.4	81.7	83.3	86.9	78.5

Best performance are presented in bold

PPV positive predictive value, NPV negative predictive value

showed fair results overall, with accuracies, sensitivities, and specificities exceeding 80%.

Table 3 shows the ranks of the features after the application of RFE and RF classification. Among these, the top 10 regions with the strongest feature importance were the thickness of the right inferior occipital gyrus, followed by the MD of the middle cerebellar peduncle and the nodal efficient values of the left posterior cingulate gyrus, which was followed by the cortical thickness of the right supplementary motor area and right temporal pole of the superior temporal gyrus, volume of the left cuneus, mean curvature of the left inferior temporal gyrus, MD of the right fornix, shortest path length, and maximum flow of the left posterior cingulate gyrus.

As a sensitivity analysis, we confined the LFA group to those with high ADOS scores of 14–23 ($n = 41$). As a result, although there was a decrease in all parameters (possibly due to reduced sample size) when using the RF classifier, performance was fair with an accuracy of 78.7%, sensitivity of 73.3%, specificity of 83.8%, positive predictive value of 80.8%, and a negative predictive value of 78.9%. Among the most predictive features, the cortical thickness of the left superior temporal gyurs, right middle temporal gyrus, left middle temporal gyrus, and the left inferior temporal gyrus were found to be predictive in the main analyses and also the sensitivity analyses.

Univariate Comparison Between ASD and TDC

Only cortical thickness and cortical volume of group comparison survived FDR correction for multiple comparisons in the cuneus, temporal, and occipital ROIs, which were included in the list of the top-ranking features identified by the RF classifier. The ASD group showed larger values of cortical thickness, cortical volume, and surface area in multiple regions and smaller values of cortical thickness of the supplementary motor area and paracentral gyrus than did the TDC group (Table S2). The mean curvature and subcortical volume showed mixed results, as some regions showed larger values in the ASD group while others did in the TDC group.

The ASD group showed higher FA values and MD values in multiple regions than did the TDC group (Table S3). In the structural network parameters, the ASD group exhibited an increase in the shortest path length and a decrease in the maximum flow, while nodal efficiency showed increases and decreases in several regions.

We compared the important features obtained by the RF classifier and the significant features obtained by the independent t-test. The cortical thickness of the right inferior occipital gyrus area was significantly higher in the ASD group than in the TDC group. The MD of the cerebellar

peduncle and nodal efficiency of the posterior cingulate gyrus decreased significantly in the ASD group.

Ablation Study

Our data consisted of features extracted from both T1w MRI and DTI. After removing the features calculated using data obtained from the other modalities, we evaluated whether the collective use of T1w MRI and DTI features could improve the classification performance. Table S4 shows the classification results of the RF classifier when T1w MRI features alone, DTI features alone, and a combination of T1w MRI and DTI features were used. The results obtained using T1w MRI features alone and DTI features alone were similar. The performance was enhanced when both T1w MRI and DTI features were used (accuracy, 88.8%). On the contrary, the results obtained using T1w MRI features alone showed the lowest accuracy (78.0%), this was slightly lower than when DTI features alone were used (78.7%).

Discussion

To our knowledge, this is the first study to apply machine learning methods to differentiate preschoolers with LFA and age-matched TDCs using both T1w and DTI parameters. As hypothesized, classification accuracy was enhanced when T1w MRI and DTI data were combined and the most prominent classification features included cortical thickness, MD, and nodal efficiency. These findings emphasize the potential of multimodal imaging for the identification of LFA preschoolers.

The top three features included cortical thickness of the right inferior occipital gyrus, MD of the middle cerebellar peduncle, and nodal efficiency of the left posterior cingulate gyrus. The inferior occipital gyrus is a part of the face-processing network, along with the fusiform gyrus and amygdala (Domes et al., 2013), and numerous studies have reported decreased activity in these areas in patients with ASD (Pierce et al., 2001, 2004). With regard to the middle cerebellar peduncle, previous studies have reported abnormal white-matter integrity in this area, as Shukla et al. revealed reduced connectivity (Shukla et al., 2010), while Sivaswamy et al. reported overconnectivity in the right middle cerebellar peduncle (Sivaswamy et al., 2010). The middle cerebellar peduncle consists of incoming fibers from various cortical areas related to not only sensory-motor control but also language, social cognition, and emotion (Crippa et al., 2016). The mean curvature of the left inferior temporal gyrus was found to be among the top-ranking classification features in our main analyses and the cortical thickness of this area was found to be predictive in the sensitivity analyses. These results are in line with a study that compared

Table 3 Rank of feature importance determined by the random forest classifier

Count	Target	Region
5	Thickness	Right inferior occipital gyrus
4	MD	Middle cerebellar peduncle
4	Nodal eff	Left posterior cingulate gyrus
3	Thickness	Right supplementary motor area
3	Thickness	Right temporal pole: superior temporal gyrus
3	GM volume	Left cuneus
3	Mean curvature	Left inferior temporal gyrus
3	MD	Right fornix (cres)/Stria terminalis
3	Shortest PL	Left posterior cingulate gyrus
3	Max. Flow	Left posterior cingulate gyrus
2	Subcortical	Right caudate
2	Thickness	Right superior frontal gyrus, orbital part
2	Thickness	Left superior occipital gyrus
2	Thickness	Left temporal pole: superior temporal gyrus
2	Thickness	Left temporal pole: middle temporal gyrus
2	GM volume	Right inferior occipital gyrus
2	Mean curvature	Left middle frontal gyrus, orbital part
2	MD	Right cingulum (hippocampus)
2	Nodal Eff	Left temporal pole: middle temporal gyrus
1	Subcortical	Left caudate
1	Subcortical	Left amygdala
1	Thickness	Right calcarine fissure and surrounding cortex
1	Thickness	Left cuneus
1	Thickness	Right middle temporal gyrus
1	Thickness	Right temporal pole: middle temporal gyrus
1	GM volume	Right inferior temporal gyrus
1	GM volume	Left calcarine fissure and surrounding cortex
1	GM volume	Right cuneus
1	GM volume	Left lingual gyrus
1	GM volume	Left superior occipital gyrus
1	GM volume	Right superior occipital gyrus
1	GM volume	Right temporal pole: superior temporal gyrus
1	GM volume	Left middle temporal gyrus
1	Surface area	Left cuneus
1	Surface area	Left superior occipital gyrus
1	Surface area	Left middle occipital gyrus
1	Surface area	Right inferior occipital gyrus
1	Mean curvature	Right precentral gyrus
1	Mean curvature	Right middle occipital gyrus
1	Mean curvature	Right inferior parietal, but supramarginal and angular gyri
1	Mean curvature	Right precuneus
1	FA	Light cingulum (cingulate gyrus)
1	FA	Right fornix (cres)/Stria terminalis
1	MD	Left posterior limb of internal capsule
1	MD	Left superior longitudinal fasciculus
1	Shortest PL	Right inferior frontal gyrus, orbital part
1	Shortest PL	Left anterior cingulate and paracingulate gyri
1	Nodal eff	Right inferior frontal gyrus, orbital part
1	Nodal eff	Left supplementary motor area
1	Nodal eff	Left anterior cingulate and paracingulate gyri
1	Max. flow	Right inferior frontal gyrus, triangular part
1	Max. flow	Left temporal pole: middle temporal gyrus

MD mean diffusivity, *Nodal Eff* nodal efficiency, *GM* gray matter, *Shortest PL* shortest path length, *Max Flow* nodal maximum flow, *FA* fractional anisotropy

children with LFA and those with HFA using whole-brain VBM-DARTEL (Cai et al., 2018). An increase in gray matter volume of the left inferior temporal gyrus was found in both the LFA and HFA groups and remained significant even after including IQ as a covariate, which suggests that this area is linked to autism neurobiology regardless of IQ levels.

Previous studies have reported findings that suggest that characteristics of the brain structure and function differ between patients with LFA and HFA. Cai et al. (2018) found that while an increase in the gray matter volume of the left inferior temporal gyrus was found in both subjects with LFA and HFA, an increase in the gray matter volume of the left middle temporal gyrus was found only in the patients with LFA and not in those with HFA. Patients with LFA have demonstrated more distinct patterns of atypical connectivity than have those with HFA in both adults and children (Gabrielsen et al., 2018; Reiter et al., 2019). Most studies have included patients with HFA alone or both patients with LFA and those with HFA but have not investigated the effect of IQ; therefore, future studies focusing on patients with LFA or stratifying the population according to IQ level are necessary to discriminate the effect of ASD depending on the effect of low IQ on brain morphology.

A strength of this study is the narrow age range of participants, which is associated with a very minor effect of brain development among patients. Classification studies using T1w MRI and DTI have been conducted in the past, but only two studies using T1w MRI data targeted the preschool age (Calderoni et al., 2012; Gori et al., 2015), while all the studies that used DTI data targeted populations over the age of 8 years (Deshpande et al., 2013; Ingalhalikar et al., 2011; Payabvash et al., 2019). Many studies suggest that brain morphological features of ASD change across age, for example, a longitudinal study on patients with ASD aged 3–35 years found that the whole-brain volume increased in young children and subsequently decreased during adolescence (Lange et al., 2015). A recent study reported that cerebral growth trajectory from the age of 2 to 13 years differed according to sex, as cerebral growth was higher in TDC boys than in those with ASD throughout childhood, whereas cerebral volumes were similar between TDC girls and those with ASD but with showed slower trajectories (Lee et al., 2020). As the brain development is dynamic in patients with ASD, studies targeting different age groups may inevitably yield different results. Future longitudinal studies are warranted to capture the developmental trajectories that are critical for understanding within-person and between-person variation in development (Lange et al., 2015).

As mentioned previously, only one study combined parameters from both T1w MRI and DTI. (Libero et al., 2015) This study combined T1w MRI, DTI, and MRS data and achieved the highest accuracy of 91.9% by including the RD for the right forceps minor, FA for the left forceps minor,

and cortical thickness for the pars opercularis aspect of the inferior frontal gyrus as the best predicting features (Libero et al., 2015). This study was limited by possible leakage as the data points included were the significant resulting values of the statistical analyses of separate neuroimaging modalities (Mateos-Perez et al., 2018). Moreover, the sample size was very small (19 patients with ASD and 18 TDCs) and targeted only adults with HFA. Although we implemented different features and classifiers, our study achieved a similar accuracy of 88.8%, and both studies clearly showed that a higher accuracy can be achieved by combining data from both T1w MRI and DTI rather than using data from a single modality. From the feature combining point of view, strengths of T1w MRI include volume and segmented regions, while those of DTI include neural tracks. It is difficult to combine them because each modality provides different characteristics. Once combined and analyzed, areas that are not covered by one modality can be supplemented by the other. From an implementation point of view, we used T1w MRI and an extraction method to extract significant features from DTI. In another study, dimension reduction was performed using principal component analysis (PCA) or independent component analysis (ICA) method using data on millions of voxels (Mateos-Perez et al., 2018). While voxel data with many sections with similar intensities can obtain effective results through PCA, our method has little effect because different features were extracted. Instead, by using the RFE method, only features that distinguish ASD from normal controls were obtained effectively. From a comparative point of view, however, as the two studies target different age ranges and IQ ranges, direct comparison is difficult. We suggest that future studies using these modalities with wider age ranges, larger sample sizes, and ASDs with different characteristics are warranted.

There were some notable limitations to our study. First, there was a significant difference in IQ between the ASD and TDC groups. Further studies should include an IQ-matched control group without an ASD diagnosis. Second, as we included only preschoolers and child participants, adolescent and adult individuals with ASD were excluded from the results. The narrow age range is advantageous as we were able to reduce the effect of age development, but generalizability to other age groups of ASD is limited. Although we used an age-matched control, rapid neuroanatomical changes occur during the 3- 6 year old period (Gilmore et al., 2018), and age-related variances could have affected our results. Adjusting variables of interest for within-group variance in age may be warranted in the future (Turesky et al., 2021). In addition, the majority of our patients were boys, but the proportion of girls was larger than that in many studies. Moreover, the sample size was relatively small to achieve sufficient statistical power, so further studies with larger sample sizes and a larger proportion of girls are needed. We

used two different tools to measure IQ, and the IQ score in ASD preschoolers may not be stable in older ages. Several studies have found cognitive score gains in ASD preschoolers, as high as 23-points, suggesting that LFA preschoolers may not meet LFA criteria at different ages (Dietz et al., 2007; Flanagan et al., 2015; Turner et al., 2006). We only used the IQ score as an indicator of LFA, but IQ may be an imprecise proxy for functional abilities in ASD (Tamm, Day & Duncan et al., 2021). Other tools like the Vineland Adaptive Behavior Scale could be implemented in the future. Data acquisition in preschoolers with ASD can be very challenging due to head motion caused by difficulty in staying still in an unfamiliar and noisy environment (Tziraki et al., 2021). Most multicenter MRI studies in ASD, like the EU-AIMS Longitudinal European Autism Project (LEAP) and Autism Brain Imaging Exchange (ABIDE) (Di Martino et al., 2014; Isaksson et al., 2018), usually target older age populations, with the exception of Enhancing NeuroImaging Genetics Through Meta-Analysis (ENIGMA), and the possibility of conducting large-scale multimodal MRI study in this population in the near future is limited. However, recent studies have focused on developing protocols for acquiring MRI scans in ASD children (Gabrielsen et al., 2018; Nordahl et al., 2016; Tziraki et al., 2021), and establishment of a well-designed protocol could increase the chances for a confirmation study of our results. Also, the MRI data were obtained from two sites using different brain imaging protocols. Although we applied robust harmonization methods, scanner effects may have biased our results. We used the AAL template and the Johns Hopkins University ICBM-DTI-81 White-Matter Labels Atlas, which were derived from adult brain MRI, to identify the parcellated brain region. Many previous studies have used the brain template, but it could induce registration errors that make it difficult to guarantee the exact match of the anatomical boundaries of regions between the brains of children and adults. Finally, our study was performed in a controlled hospital setting, with a high percentage of ASD patients (54%), therefore results may not be generalizable to a community sample due to the imbalance in classification (less than 2% of ASD). External validation of the study results in a large-scale real-world setting are warranted to increase clinical utility (Ho et al., 2020).

This was the first study to generate classification models for LFA preschoolers using multiple machine learning techniques. Implementation of both T1w MRI and DTI data yielded better results than did the use a single imaging modality. Further multimodal classification studies with larger sample sizes, wider age ranges, and patients with various levels of functioning are needed to expand these results.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10803-021-05368-z>.

Acknowledgments This research was supported by the National Research Foundation (NRF) funded by the Korean Government (MSIT) (2019M3E5D1A01069345 to J-IK and 2020M3E5D9080787 to B-NK), by the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI) and Korea Dementia Research Center (KDRC), funded by the Ministry of Health & Welfare and Ministry of Science and ICT, Republic of Korea (No. HU20C0198 to J-ML), and the Technology Innovation Program (Industrial Strategic Technology Development Program) funded by the Ministry of Trade, Industry, and Energy (20002769 to B-NK, Development of Next Generation Platform for Diagnosis and Therapeutic of Attention Deficit Hyperactivity Disorder and Intellectual Disability based on Big Data).

Author contribution All authors contributed to the study conception, design, material preparation and data collection. Analysis were performed by Sungkyu Bang and Jin-Ju Yang. The first draft of the manuscript was written by Johanna Inhyang Kim, Sungkyu Bang and Jin-Ju Yang. All authors read, commented on and approved the final manuscript. Jong-Min Lee and Bung-Nyun Kim supervised the whole process.

Declarations

Conflict of interest The authors have no conflict of interest to declare.

References

- Abraham, A., Pedregosa, F., Eickenberg, M., Gervais, P., Mueller, A., Kossaifi, J., Gramfort, A., Thirion, B., & Varoquaux, G. (2014). Machine learning for neuroimaging with scikit-learn. *Frontiers in Neuroinformatics*, 8, 14. <https://doi.org/10.3389/fninf.2014.00014>
- Akshoomoff, N., Lord, C., Lincoln, A. J., Courchesne, R. Y., Carper, R. A., Townsend, J., & Courchesne, E. (2004). Outcome classification of preschool children with autism spectrum disorders using MRI brain measures. *Journal of the American Academy of Child and Adolescent Psychiatry*, 43(3), 349–357. <https://doi.org/10.1097/00004583-200403000-00018>
- Alloway, T. P. (2010). Working memory and executive function profiles of individuals with borderline intellectual functioning. *Journal of Intellectual Disability Research*, 54(5), 448–456. <https://doi.org/10.1111/j.1365-2788.2010.01281.x>
- Ameis, S. H., & Catani, M. (2015). Altered white matter connectivity as a neural substrate for social impairment in Autism Spectrum Disorder. *Cortex*, 62, 158–181. <https://doi.org/10.1016/j.cortex.2014.10.014>
- Andrews, D. S., Lee, J. K., Solomon, M., Rogers, S. J., Amaral, D. G., & Nordahl, C. W. (2019). A diffusion-weighted imaging tract-based spatial statistics study of autism spectrum disorder in preschool-aged children. *Journal of Neurodevelopmental Disorders*, 11(1), 32. <https://doi.org/10.1186/s11689-019-9291-z>
- Baio, J., Wiggins, L., Christensen, D. L., Maenner, M. J., Daniels, J., Warren, Z., Kurzius-Spencer, M., Zahorodny, W., Rosenberg, C.R., White, T., & Dowling, N. F. (2018). Prevalence of autism spectrum disorder among children aged 8 years—Autism and developmental disabilities monitoring network, 11 sites, United States, 2014. *MMWR Surveillance Summaries*, 67(6), 1–23. <https://doi.org/10.15585/mmwr.ss6706a1>
- Bullmore, E., & Sporns, O. (2009). Complex brain networks: Graph theoretical analysis of structural and functional systems. *Nature*

- Reviews Neuroscience*, 10(3), 186–198. <https://doi.org/10.1038/nrn2575>
- Cai, J., Hu, X., Guo, K., Yang, P., Situ, M., & Huang, Y. (2018). Increased left inferior temporal gyrus was found in both low function autism and high function autism. *Front Psychiatry*, 9, 542. <https://doi.org/10.3389/fpsy.2018.00542>
- Calderoni, S., Retico, A., Biagi, L., Tancredi, R., Muratori, F., & Tosetti, M. (2012). Female children with autism spectrum disorder: An insight from mass-univariate and pattern classification analyses. *NeuroImage*, 59(2), 1013–1022. <https://doi.org/10.1016/j.neuroimage.2011.08.070>
- Chu, C., Hsu, A. L., Chou, K. H., Bandettini, P., Lin, C. P., Neuroimaging, A., & s. D. (2012). Does feature selection improve classification accuracy? Impact of sample size and feature selection on classification using anatomical magnetic resonance images. *NeuroImage*, 60(1), 59–70. <https://doi.org/10.1016/j.neuroimage.2011.11.066>
- Collins, D. L., Neelin, P., Peters, T. M., & Evans, A. C. (1994). Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. *Journal of Computer Assisted Tomography*, 18(2), 192–205.
- Committee on Drugs. American Academy of Pediatrics. (2002). Guidelines for monitoring and management of pediatric patients during and after sedation for diagnostic and therapeutic procedures: Addendum. *Pediatrics*, 110(4), 836–838. <https://doi.org/10.1542/peds.110.4.836>
- Crippa, A., Del Vecchio, G., Busti Ceccarelli, S., Nobile, M., Arrigoni, F., & Brambilla, P. (2016). Cortico-cerebellar connectivity in autism spectrum disorder: What do we know so far? *Front Psychiatry*, 7, 20. <https://doi.org/10.3389/fpsy.2016.00020>
- Deshpande, G., Libero, L. E., Sreenivasan, K. R., Deshpande, H. D., & Kana, R. K. (2013). Identification of neural connectivity signatures of autism using machine learning. *Frontiers in Human Neuroscience*, 7, 670. <https://doi.org/10.3389/fnhum.2013.00670>
- Di Martino, A., Yan, C. G., Li, Q., Denio, E., Castellanos, F. X., Alaerts, K., Anderson, J.S., Assaf, M., Bookheimer, S.Y., Dapretto, M., & Milham, M. P. (2014). The autism brain imaging data exchange: Towards a large-scale evaluation of the intrinsic brain architecture in autism. *Molecular Psychiatry*, 19(6), 659–667. <https://doi.org/10.1038/mp.2013.78>
- Dietz, C., Swinkels, S. H., Buitelaar, J. K., Van Daalen, E., & Van Engeland, H. (2007). Stability and change of IQ scores in pre-school children diagnosed with autism spectrum disorder. *European Child and Adolescent Psychiatry*, 16(6), 405–410.
- Domes, G., Heinrichs, M., Kumbier, E., Grossmann, A., Hauenstein, K., & Herpertz, S. C. (2013). Effects of intranasal oxytocin on the neural basis of face processing in autism spectrum disorder. *Biological Psychiatry*, 74(3), 164–171. <https://doi.org/10.1016/j.biopsych.2013.02.007>
- Erbetta, A., Bulgheroni, S., Contarino, V. E., Chiapparini, L., Esposito, S., Annunziata, S., & Riva, D. (2015). Low-functioning autism and nonsyndromic intellectual disability: Magnetic resonance imaging (MRI) findings. *Journal of Child Neurology*, 30(12), 1658–1664.
- Flanagan, H. E., Smith, I. M., Vaillancourt, T., Duku, E., Szatmari, P., Bryson, S., Fombonne, E., Mirenda, P., Roberts, W., Volden, J., Waddell, C., Zwaigenbaum, L., Bennett, T., Elsabbagh, M., & Georgiades, S. (2015). Stability and change in the cognitive and adaptive behavior scores of preschoolers with autism spectrum disorder. *J Autism Devel Disord*, 45, 2691–2703.
- Fortin, J. P., Cullen, N., Sheline, Y. I., Taylor, W. D., Aselcioglu, I., Cook, P. A., Adams, P., Cooper, C., Fava, M., McGrath, P.J., & Shinohara, R. T. (2018). Harmonization of cortical thickness measurements across scanners and sites. *NeuroImage*, 167, 104–120. <https://doi.org/10.1016/j.neuroimage.2017.11.024>
- Fortin, J. P., Parker, D., Tunc, B., Watanabe, T., Elliott, M. A., Ruparel, K., & Shinohara, R. T. (2017). Harmonization of multi-site diffusion tensor imaging data. *NeuroImage*, 161, 149–170. <https://doi.org/10.1016/j.neuroimage.2017.08.047>
- Gabrielsen, T. P., Anderson, J. S., Stephenson, K. G., Beck, J., King, J. B., Kellems, R., Top, D.N., Russell, N.C., Anderberg, E., Lundwall, R.A., & South, M. (2018). Functional MRI connectivity of children with autism and low verbal and cognitive performance. *Molecular Autism*, 9, 67. <https://doi.org/10.1186/s13229-018-0248-y>
- Gilmore, J. H., Knickmeyer, R. C., & Gao, W. (2018). Imaging structural and functional brain development in early childhood. *Nature Reviews Neuroscience*, 19(3), 123–137. <https://doi.org/10.1038/nrn.2018.1>
- Gori, I., Giuliano, A., Muratori, F., Saviozzi, I., Oliva, P., Tancredi, R., Cosenza, A., Tosetti, M., Calderoni, S., & Retico, A. (2015). Gray matter alterations in young children with autism spectrum disorders: Comparing morphometry at the voxel and regional level. *Journal of Neuroimaging*, 25(6), 866–874. <https://doi.org/10.1111/jon.12280>
- Gotham, K., Pickles, A., & Lord, C. (2012). Trajectories of autism severity in children using standardized ADOS scores. *Pediatrics*, 130(5), e1278–1284. <https://doi.org/10.1542/peds.2011-3668>
- Ho, S. Y., Phua, K., Wong, L., & Bin Goh, W. W. (2020). Extensions of the external validation for checking learned model interpretability and generalizability. *Patterns*. <https://doi.org/10.1016/j.patter.2020.100129>
- Huang, Z. A., Zhu, Z., Yau, C. H., & Tan, K. C. (2020). Identifying autism spectrum disorder from resting-state fMRI using deep belief network. *IEEE Transactions on Neural Networks and Learning Systems*. <https://doi.org/10.1109/TNNLS.2020.3007943>
- Im, K., Lee, J. M., Lyttelton, O., Kim, S. H., Evans, A. C., & Kim, S. I. (2008). Brain size and cortical structure in the adult human brain. *Cerebral Cortex*, 18(9), 2181–2191. <https://doi.org/10.1093/cercor/bhm244>
- Im, K., Lee, J. M., Yoon, U., Shin, Y. W., Hong, S. B., Kim, I. Y., Kwon, J.S., & Kim, S. I. (2006). Fractal dimension in human cortical surface: Multiple regression analysis with cortical thickness, sulcal depth, and folding area. *Human Brain Mapping*, 27(12), 994–1003. <https://doi.org/10.1002/hbm.20238>
- Ingalhalikar, M., Parker, D., Bloy, L., Roberts, T. P., & Verma, R. (2011). Diffusion based abnormality markers of pathology: Toward learned diagnostic prediction of ASD. *NeuroImage*, 57(3), 918–927. <https://doi.org/10.1016/j.neuroimage.2011.05.023>
- Ingalhalikar, M., Parker, W. A., Bloy, L., Roberts, T. P., & Verma, R. (2014). Creating multimodal predictors using missing data: Classifying and subtyping autism spectrum disorder. *Journal of Neuroscience Methods*, 235, 1–9. <https://doi.org/10.1016/j.jneumeth.2014.06.030>
- Isaksson, J., Tammimies, K., Neufeld, J., Cauvet, E., Lundin, K., Buitelaar, J. K., Loth, E., Murphy, D.G., Spooren, W., & group, E.-A. L. (2018). EU-AIMS Longitudinal European Autism Project (LEAP): The autism twin cohort. *Molecular Autism*, 9, 26. <https://doi.org/10.1186/s13229-018-0212-x>
- Jiao, Y., Chen, R., Ke, X. Y., Chu, K. K., Lu, Z. H., & Herskovits, E. H. (2010). Predictive models of autism spectrum disorder based on brain regional cortical thickness. *NeuroImage*, 50(2), 589–599. <https://doi.org/10.1016/j.neuroimage.2009.12.047>
- Kaufman, J., Birmaher, B., Brent, D., Rao, U., Flynn, C., Moreci, P., Williamson, D., & Ryan, N. (1997). Schedule for affective disorders and schizophrenia for school-age children-present and lifetime version (K-SADS-PL): Initial reliability and validity data. *Journal of the American Academy of Child and Adolescent Psychiatry*, 36(7), 980–988. <https://doi.org/10.1097/00004583-199707000-00021>

- Kim, Y. S., Cheon, K. A., Kim, B. N., Chang, S. A., Yoo, H. J., Kim, J. W., Cho, S. C., Seo, D. H., Bae, M. O., So, Y. K., & Leventhal, B. (2004). The reliability and validity of kiddie-schedule for affective disorders and schizophrenia-present and lifetime version- Korean version (K-SADS-PL-K). *Yonsei Medical Journal*, *45*(1), 81–89. <https://doi.org/10.3349/ymj.2004.45.1.81>
- Lange, N., Travers, B. G., Bigler, E. D., Prigge, M. B., Froehlich, A. L., Nielsen, J. A., Cariello, A.N., Zielinski, B.A., Anderson, J.S., Fletcher, P.T., & Lainhart, J. E. (2015). Longitudinal volumetric brain changes in autism spectrum disorder ages 6–35 years. *Autism Research*, *8*(1), 82–93. <https://doi.org/10.1002/aur.1427>
- Lee, D. K., Yoon, U., Kwak, K., & Lee, J. M. (2015). Automated segmentation of cerebellum using brain mask and partial volume estimation map. *Computational and Mathematical Methods in Medicine*, *2015*, 167489. <https://doi.org/10.1155/2015/167489>
- Lee, J. K., Andrews, D. S., Ozonoff, S., Solomon, M., Rogers, S., Amaral, D. G., & Nordahl, C. W. (2020). Longitudinal evaluation of cerebral growth across childhood in boys and girls with autism spectrum disorder. *Biological Psychiatry*. <https://doi.org/10.1016/j.biopsych.2020.10.014>
- Li, D., Karnath, H. O., & Xu, X. (2017). Candidate biomarkers in children with autism spectrum disorder: a review of MRI studies. *Neuroscience Bulletin*, *33*(2), 219–237. <https://doi.org/10.1007/s12264-017-0118-1>
- Li, S. J., Wang, Y., Qian, L., Liu, G., Liu, S. F., Zou, L. P., Zhang, J. S., Hu, N., Chen, X. Q., Yu, S. Y., & Ma, L. (2018). Alterations of white matter connectivity in preschool children with autism spectrum disorder. *Radiology*, *288*(1), 209–217. <https://doi.org/10.1148/radiol.2018170059>
- Liberio, L. E., DeRamus, T. P., Lahti, A. C., Deshpande, G., & Kana, R. K. (2015). Multimodal neuroimaging based classification of autism spectrum disorder using anatomical, neurochemical, and white matter correlates. *Cortex*, *66*, 46–59. <https://doi.org/10.1016/j.cortex.2015.02.008>
- Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Jr., Leventhal, B. L., DiLavore, P. C., Pickles, A., & Rutter, M. (2000). 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*(3), 205–223.
- MacDonald, D., Kabani, N., Avis, D., & Evans, A. C. (2000). Automated 3-D extraction of inner and outer surfaces of cerebral cortex from MRI. *NeuroImage*, *12*(3), 340–356. <https://doi.org/10.1006/nimg.1999.0534>
- Maslov, S., & Sneppen, K. (2002). Specificity and stability in topology of protein networks. *Science*, *296*(5569), 910–913. <https://doi.org/10.1126/science.1065103>
- Mateos-Perez, J. M., Dadar, M., Lacalle-Aurioles, M., Iturria-Medina, Y., Zeighami, Y., & Evans, A. C. (2018). Structural neuroimaging as a clinical predictor: A review of machine learning applications. *NeuroImage Clinical*, *20*, 506–522. <https://doi.org/10.1016/j.nicl.2018.08.019>
- Moon, S. J., Hwang, J., Kana, R., Torous, J., & Kim, J. W. (2019). Accuracy of machine learning algorithms for the diagnosis of autism spectrum disorder: Systematic review and meta-analysis of brain magnetic resonance imaging studies. *JMIR Mental Health*, *6*(12), e14108. <https://doi.org/10.2196/14108>
- Nordahl, C. W., Mello, M., Shen, A. M., Shen, M. D., Vismara, L. A., Li, D., Harrington, K., Tanase, C., Goodlin-Jones, B., Rogers, S., & Amaral, D. G. (2016). Methods for acquiring MRI data in children with autism spectrum disorder and intellectual impairment without the use of sedation. *Journal of Neurodevelopmental Disorders*, *8*, 20. <https://doi.org/10.1186/s11689-016-9154-9>
- Oishi, K., Zilles, K., Amunts, K., Faria, A., Jiang, H., Li, X., Akhter, K., Hua, K., Woods, R., Toga, A.W., & Mori, S. (2008). Human brain white matter atlas: Identification and assignment of common anatomical structures in superficial white matter. *NeuroImage*, *43*(3), 447–457. <https://doi.org/10.1016/j.neuroimage.2008.07.009>
- Pagnozzi, A. M., Conti, E., Calderoni, S., Fripp, J., & Rose, S. E. (2018). A systematic review of structural MRI biomarkers in autism spectrum disorder: A machine learning perspective. *International Journal of Developmental Neuroscience*, *71*, 68–82. <https://doi.org/10.1016/j.ijdevneu.2018.08.010>
- Park, G., Kwak, K., Seo, S. W., & Lee, J. M. (2018). Automatic segmentation of corpus callosum in midsagittal based on bayesian inference consisting of sparse representation error and multi-atlas voting. *Frontiers in Neuroscience*, *12*, 629. <https://doi.org/10.3389/fnins.2018.00629>
- Park, K. S., Yoon, Y. H., Park, H. J., Kwon, K. U. (1996). Development of KEDI-WISC, individual intelligence test for Korean children. In. Seoul, Republic of Korea: Korean Educational Development Institute.
- Patenaude, B., Smith, S. M., Kennedy, D. N., & Jenkinson, M. (2011). A Bayesian model of shape and appearance for subcortical brain segmentation. *NeuroImage*, *56*(3), 907–922. <https://doi.org/10.1016/j.neuroimage.2011.02.046>
- Payabvash, S., Palacios, E. M., Owen, J. P., Wang, M. B., Tavassoli, T., Gerdes, M., Brandes-Aitken, A., Cuneo, D., Marco, E. J., & Mukherjee, P. (2019). White matter connectome edge density in children with autism spectrum disorders: Potential imaging biomarkers using machine-learning models. *Brain Connect*, *9*(2), 209–220. <https://doi.org/10.1089/brain.2018.0658>
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., & Duchesnay, E. (2011). Scikit-learn: Machine learning in python. *Journal of Machine Learning Research*, *12*, 2825–2830.
- Pierce, K., Haist, F., Sedaghat, F., & Courchesne, E. (2004). The brain response to personally familiar faces in autism: Findings of fusiform activity and beyond. *Brain*, *127*(Pt 12), 2703–2716. <https://doi.org/10.1093/brain/awh289>
- Pierce, K., Muller, R. A., Ambrose, J., Allen, G., & Courchesne, E. (2001). Face processing occurs outside the fusiform “face area” in autism: Evidence from functional MRI. *Brain*, *124*(Pt 10), 2059–2073. <https://doi.org/10.1093/brain/124.10.2059>
- Qin, B., Wang, L., Zhang, Y., Cai, J., Chen, J., & Li, T. (2018). Enhanced topological network efficiency in preschool autism spectrum disorder: A diffusion tensor imaging study. *Front Psychiatry*, *9*, 278. <https://doi.org/10.3389/fpsy.2018.00278>
- Qureshi, M. N., Min, B., Jo, H. J., & Lee, B. (2016). Multiclass classification for the differential diagnosis on the ADHD subtypes using recursive feature elimination and hierarchical extreme learning machine: Structural MRI study. *PLoS ONE*, *11*(8), e0160697. <https://doi.org/10.1371/journal.pone.0160697>
- Reiter, M. A., Mash, L. E., Linke, A. C., Fong, C. H., Fishman, I., & Muller, R. A. (2019). Distinct patterns of atypical functional connectivity in lower-functioning autism. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *4*(3), 251–259. <https://doi.org/10.1016/j.bpsc.2018.08.009>
- Robbins, S., Evans, A. C., Collins, D. L., & Whitesides, S. (2004). Tuning and comparing spatial normalization methods. *Medical Image Analysis*, *8*(3), 311–323. <https://doi.org/10.1016/j.media.2004.06.009>
- Rubinov, M., & Sporns, O. (2010). Complex network measures of brain connectivity: Uses and interpretations. *NeuroImage*, *52*(3), 1059–1069. <https://doi.org/10.1016/j.neuroimage.2009.10.003>
- Shen, M. D., Li, D. D., Keown, C. L., Lee, A., Johnson, R. T., Angkustsiri, K., Rogers, S.J., Müller, R.A., Amaral, D.G., & Nordahl, C. W. (2016). Functional connectivity of the amygdala is disrupted in preschool-aged children with autism spectrum disorder. *Journal of*

- the American Academy of Child and Adolescent Psychiatry, 55(9), 817–824. <https://doi.org/10.1016/j.jaac.2016.05.020>
- Shin, M. S., & Cho, S. C. (2010). The Korean Leiter International Performance Scale—Revised (K-Leiter-R). Seoul: Hakjisa.
- Shukla, D. K., Keehn, B., Lincoln, A. J., & Muller, R. A. (2010). White matter compromise of callosal and subcortical fiber tracts in children with autism spectrum disorder: A diffusion tensor imaging study. *J Am Acad Child Adolesc Psychiatry*, 49(12), 1269–1278, 1278 e1261–1262. <https://doi.org/10.1016/j.jaac.2010.08.018>
- Sivaswamy, L., Kumar, A., Rajan, D., Behen, M., Muzik, O., Chugani, D., & Chugani, H. (2010). A diffusion tensor imaging study of the cerebellar pathways in children with autism spectrum disorder. *Journal of Child Neurology*, 25(10), 1223–1231. <https://doi.org/10.1177/0883073809358765>
- Sled, J. G., Zijdenbos, A. P., & Evans, A. C. (1998). A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Transactions on Medical Imaging*, 17(1), 87–97. <https://doi.org/10.1109/42.668698>
- Smith, S. M. (2002). Fast robust automated brain extraction. *Human Brain Mapping*, 17(3), 143–155. <https://doi.org/10.1002/hbm.10062>
- Tamm, L., Day, H. A., & Duncan, A. (2021). Comparison of adaptive functioning measures in adolescents with autism spectrum disorder without intellectual disability. *Journal of Autism and Developmental Disorders*. <https://doi.org/10.1007/s10803-021-05013-9>
- Turesky, T. K., Vanderauwera, J., & Gaab, N. (2021). Imaging the rapidly developing brain: Current challenges for MRI studies in the first five years of life. *Developmental Cognitive Neuroscience*, 47, 100893. <https://doi.org/10.1016/j.dcn.2020.100893>
- Turner, L. M., Stone, W. L., Pozdol, S. L., & Coonrod, E. E. (2006). Follow-up of children with autism spectrum disorders from age 2 to age 9. *Autism*, 10(3), 243–265.
- Tziraki, M., Garg, S., Harrison, E., Wright, N. B., Hawkes, R., Akhtar, K., Green, J., & Stivaros, S. (2021). A neuroimaging preparation protocol tailored for autism. *Autism Research*, 14(1), 65–74. <https://doi.org/10.1002/aur.2427>
- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., & Joliot, M. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage*, 15(1), 273–289. <https://doi.org/10.1006/nimg.2001.0978>
- Vade, A., Sukhani, R., Dolenga, M., & Habisohn-Schuck, C. (1995). Chloral hydrate sedation of children undergoing CT and MR imaging: Safety as judged by American Academy of Pediatrics guidelines. *AJR. American Journal of Roentgenology*, 165(4), 905–909. <https://doi.org/10.2214/ajr.165.4.7676990>
- van't Hof, M., Tisseur, C., van Berckleear-Onnes, I., van Nieuwenhuyzen, A., Daniels, A. M., Deen, M., Hoek, H. W., & Ester, W. A. (2021). Age at autism spectrum disorder diagnosis: A systematic review and meta-analysis from 2012 to 2019. *Autism*, 25(4), 862–873. <https://doi.org/10.1177/1362361320971107>
- Wook Yoo, S., Han, C. E., Shin, J. S., Won Seo, S., Na, D. L., Kaiser, M., Jeong, Y., & Seong, J. K. (2015). A network flow-based analysis of cognitive reserve in normal ageing and Alzheimer's disease. *Science and Reports*, 5, 10057. <https://doi.org/10.1038/srep10057>
- Xiao, X., Fang, H., Wu, J., Xiao, C., Xiao, T., Qian, L., Liang, F., Xiao, Z., Chu, K. K., & Ke, X. (2017). Diagnostic model generated by MRI-derived brain features in toddlers with autism spectrum disorder. *Autism Research*, 10(4), 620–630. <https://doi.org/10.1002/aur.1711>
- Yang, J. J., Yoon, U., Yun, H. J., Im, K., Choi, Y. Y., Lee, K. H., Park, H., Hough, M. G., & Lee, J. M. (2013). Prediction for human intelligence using morphometric characteristics of cortical surface: Partial least square analysis. *Neuroscience*, 246, 351–361. <https://doi.org/10.1016/j.neuroscience.2013.04.051>
- Yeh, F. C., Verstynen, T. D., Wang, Y., Fernandez-Miranda, J. C., & Tseng, W. Y. (2013). Deterministic diffusion fiber tracking improved by quantitative anisotropy. *PLoS ONE*, 8(11), e80713. <https://doi.org/10.1371/journal.pone.0080713>
- Zhou, Y., Yu, F., & Duong, T. (2014). Multiparametric MRI characterization and prediction in autism spectrum disorder using graph theory and machine learning. *PLoS ONE*, 9(6), e90405. <https://doi.org/10.1371/journal.pone.0090405>
- Zhuang, H., Liu, R., Wu, C., Meng, Z., Wang, D., Liu, D., Liu, M., & Li, Y. (2019). Multimodal classification of drug-naïve first-episode schizophrenia combining anatomical, diffusion and resting state functional resonance imaging. *Neuroscience Letters*, 705, 87–93. <https://doi.org/10.1016/j.neulet.2019.04.039>
- Zijdenbos, A. P., Forghani, R., & Evans, A. C. (2002). Automatic “pipeline” analysis of 3-D MRI data for clinical trials: Application to multiple sclerosis. *IEEE Transactions on Medical Imaging*, 21(10), 1280–1291. <https://doi.org/10.1109/TMI.2002.806283>
- Zilles, K., Armstrong, E., Schleicher, A., & Kretschmann, H. J. (1988). The human pattern of gyrification in the cerebral cortex. *Anatomy and Embryology (Berlin)*, 179(2), 173–179. <https://doi.org/10.1007/BF00304699>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.