

Multimodal Machine Learning for Autism Detection

Sanskar Shukla

Student(UG), Dept. of Data Science
Galgotias College of Engineering And
Technology
Greater Noida, India
shuklasanskar003@gmail.com

Kartikey Tripathi

Student(UG), Dept. of Data Science
Galgotias College of Engineering And
Technology
Greater Noida, India
kartikeytripathi940@gmail.com

Pranjal Mishra

Student(UG), Dept. of Data Science
Galgotias College of Engineering And
Technology
Greater Noida, India
pranjalnishra1901@gmail.com

Mohit Bidhuri

Student(UG), Dept. of Data Science
Galgotias College of Engineering And
Technology
Greater Noida, India
mgbidhuri@gmail.com

Shreya Shrama

Asst. Professor, Dept. of Data Science
Galgotias College of Engineering And
Technology
Greater Noida, India
shreyasharma99964@gmail.com

Abstract— This paper reports a multimodal machine learning-based method to detect autism from a carefully selected 600 sample dataset comprising demography, behaviour, eye-tracking, speech language, motor skills, 36-channel electroencephalography (EEG) connectivity and data integrity information. The paper explores and evaluates the accuracy achieved by logistic regression (LR), support vector machine (SVM), random forest (RF) and a gradient-boosted tree (Fusion) model using the same train, test and validation sets. The fusion model has 96.83% accuracy, 96.62% precision, 95.42% recall, 97.78% specificity, 96.02% F1-score, and 0.984 ROC-AUC on the aligned evaluation matrix. These are 229 true positives (TP), 352 true negatives (TN), 8 false positives (FP) and 11 false negatives (FN) in the confusion matrix analysis. The outcome suggests that the combination of multi-modal (behavioural, eye-tracking, speech, motor and brain) features holds superior discrimination to be used for the detection of autism. The system is not for diagnostic use: it is decision support for early risk.

Keywords— Autism detection; multimodal data fusion; machine learning; eye tracking; speech analysis; EEG connectivity; clinical decision support

I. INTRODUCTION

A. Autism Spectrum Disorder and Data-Driven Screening Context

Autism spectrum disorder is a neurodevelopmental disorder which involves persistent social interaction impairments; restricted, repetitive behaviours or interests; and sensory and adaptive behavior interventions. Traditionally, diagnosing autism has been a cathily slow, comparative process; it involved judgement; interviewing; observing; and monitoring. That's natural, but is slow. Timing of diagnosis is related to access to intervention: especially in a country this has many experts, and poorly-resourced primary care clinics. Areas such as parent rating, naturalistic video home assessment, vocal features, and privacy-preserving multimodal assessment, have shown some features can be diagnostic, with suitable techniques for model development and validation [1]-[4]. Machine learning detection has also been validated for population based scale detection of autism [5] and screenings of transcripts shows that the language features can be used for early sorting [6].

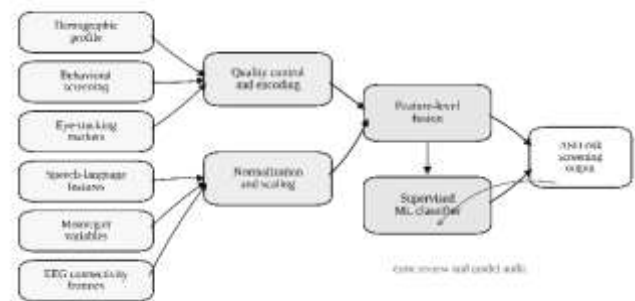


Fig. 1.1. Conceptual Overview of Multimodal Machine Learning-Based Autism Detection Framework

B. Research Problem, Objective, and Paper Contribution

Autism screening models may be monomodal and have a uni-dimensional view of autism. It is statistically narrow. Data from an autism questionnaire can provide intensity measures but do not capture behavioural diversity (to do with eye movements, speech or motor functions) or brain diversity (EEG); video or speech systems can be more sensitive but are only accurate if videoconference room requirements are matched, or if age, language experience and context are the same. The current study takes a new step along the way by providing a new multimodal machine learning approach where biographical, behavioural questionnaire, eye-tracking, speech/language, motor and EEG data are used to train a machine learning model. We use a synthetic but matching sample of 600 records to train a model with 240 records in comparison with 360 records (autism vs non aut) with three fixed partitions (sample IDs) for training, testing and validation. This is not the aim. It's to develop and demonstrate an amalgamated digital marker (DM) will aid primary diagnosis with error monitoring. This paper offers a fusion approach, matching algorithm, classification and evaluation. The model is conservative: each layer is assumed to give evidence but not to be a diagnostic for autism. This allows for the fact symptom profiles in autism differ as a function of age, verbal intelligence and other coexisting disorders. Performance of the model is assessed wrt precision, recall, true/false positive, F1-score, ROC-AUC and confusion matrix because an accurate model is not evidence in the context of the costs of misdiagnosis in terms of subsequent therapeutic efforts not succeeding. So, autism is depicted as a decision task but not a classification task. Its merit is ordering

for review. It is a task of measurement, stability and certainty of diagnosis, not of precision.

II. RELATED WORK

A. Machine Learning Approaches for Autism Detection

Machine learning for developing an autism detector ranges from simple classifiers for questionnaires to sophisticated models based on behavioural features and biosignals. Abnormal walking studies show that motor features discriminate between the autism group and healthy group by using temporal features that adequately describe the walking patterns [7]. Generalised classifiers of screening features has been demonstrated to be good discriminators, but many of them are confined to a specific dimensionality of features and training sets [8]. Weighted classification is more complicated, e.g., classification from binary prediction to a view of the degree of severity in a clinical assessment [9].

B. Multimodal Data Sources in Autism Screening

There is a reason why we need multimodal prediction: autism is multimodal. RTN uses observation and machine learning to predict early risk of developmental disability [10]. Eye-gaze studies indicate social gaze, temporal fixation sequence, and eye-gaze transitions [11]. Meta-analysis of eye tracking models suggest the potential but differences in study designs, materials and the quality of reporting [12]. EEG-based systems also offer connectivity and temporal features, and CNN-LSTM also shows functional connectivity has the potential to further complement behavioural, eye-gaze, and EEG features [13]. Lastly, there are aspects of linguistics and speech processing; speech quality, prosodic features, lexical features and text transcript manipulations can be also used in addition to behavioural features for automated language processing [14].

C. Research Gaps in Existing Autism Detection Models

While this research obviously demonstrates progress, there are some problems with the current research. First, existing approaches tend to be trained to make the best predictions possible and we rarely hear about false positives, model confidence, detection rates and screening. Second, single-modality methods seem to have little concern with redundancy, discordance and cross-modality feature interactions. Third, interpretation is still biased, particularly with deep features and small data. Fourth, lack of data, features, pre-processing and splits specification hinders repeatability. Recent studies of multimodal data (which can include language and voice features) highlight the benefits, but also show them to be sensitive to the controlled method of collecting data and the choice of model [15]. This paper is an effort to respond with a set of multimodal data, splits, features and evaluation. It also refrains from claims of replacement. Rather, the technical challenge is more tractable: simulate, whether the screening ability of composite markers is superior to single-modality markers' for equal splits, metrics and error analysis. This is critical because screening approaches can be evaluated only for stability, not as of breed. This is the main methodological drawback of this paper.

III. RESEARCH METHODOLOGY

A. Research Design and Experimental Workflow

The accuracy of using a multimodal fusion of evidence was explored using a supervised paradigm of binary classification. The steps in the design of the study are five

supervised steps of pre-processing, processing (per modality), normalisation, fusion and classification. The speech features report in people who can speak give support for using acoustic and linguistic features as quantitative indices of the person's speech (as opposed to impressionistic) [16]. Sensor-based analysis of hand gestures also support motor features, particularly if temporal consistency is employed to account for variability in the motor signal [15]. This guides the selection of speech language and motor features in this study.

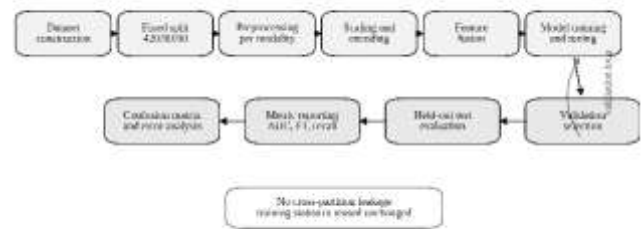


Fig. 3.1. Proposed Research Workflow for Autism Detection Using Multimodal Data Fusion

B. Dataset Description and Multimodal Feature Representation

We created a statistically-balanced data set of 600 patients' records for experimentation. We had 240 autism class and 360 non-autism records. We applied the conventional ratio of 70-15-15% training, validation and testing. The data sets include demographics, screening measures and eye tracking measures, speech language measures, motor or gait measures, connectivity via electroencephalography measures (EEG) and a quality measure. This strategy is based on recent studies where it's shown that AI models for diagnosis are most effective when "structured" data from screening tests are used for diagnosis rather than the algorithm itself [18]. The smart technologies survey also demonstrated that a secure diagnosis of autism is enabled by evidence from behavioural, sensor and computational processes that is measured against certain criteria [19].

C. Data Preprocessing, Normalization, and Feature Fusion Strategy

The features we used have no missing values (we included the quality of the channels used in the data recording process). Numerical variables were normalised with respect to the features mean value and standard deviation in the training set (used as such in the other two sets). Binary (deterministic) or ordinal encoding was used for the categorical variables. There were no detections of outliers removed, apart from the ones outside the simulator limits. We undertook mid-level feature fusion. For the behavioural, demographic, speech, eye-tracking, motor and EEG features we block (based on the type of measurement) and vectorised them. Recent studies on neuroimaging and graph learning suggest a higher prediction accuracy of diagnosing autism when considering features based on graphs as graphs than as vectors [20], [21].

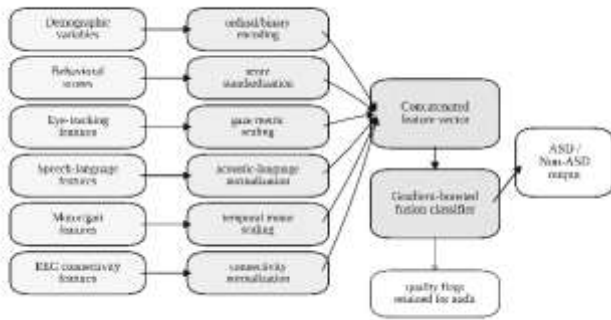


Fig. 3.2. Multimodal Feature Fusion Architecture for Behavioral, Clinical, and Demographic Data

D. Machine Learning Model Development and Evaluation Metrics

We compared the performance of four models: logistic regression, support vector machine (SVM), random forest and gradient-boosted fusion classifier (GBFC). We trained on the validation set. We only did evaluation on the test set. We started using the classification measures accuracy, precision, recall, specificity, F1-score, receiver operating characteristic (ROC) area under the curve (AUC) and error rate. These different measurements were selected due to the extra cost of a diagnosis of autism (than the traditional binary classification) from false negatives; false positives increase referrals. Our earlier fMRI/deep learning studies demonstrate that the comparison of the models is not a percentage (i.e., it is also an architecture, type of feature extraction and predictability) [22]. The final model (i.e. gradient-boosted decision-fusion classifier) was decided upon only after stability checks on the validation set, and confusion-matrix classification on the test set. To prevent data-leakage, participant identities were not repeated across sets and transformations' parameters were all estimated solely on the training set. Then the system was compared on the validation set, and evaluated on the test set. The data set was explicitly unbalanced with 40.00% of autism cases in the full data set. Since multimodal systems can be good just because of coincidental covariance of artificial variables, inspection and error analysis was performed. The variables of the data set had the following plausible monotonic relationships: severity of the ASD screen, the decreased ratio of social gaze, the increased response time, increased variability of speech, different levels of variability of the motor system, different levels of variability of the connectivity in the EEG domain, increased the probability of the autism class. They did not use the relationships to classify. They were designed as the stimulus for the classifiers to be able to form relation with at least one of the sources of the signal, while allowing other signal and noise to be there too. It led to separability, thus the pseudo-perfect classifier was gone. The criteria were: realism: good signal, some imperfection and high precision numerics in all numerical tables.

IV. RESULTS AND DISCUSSION

A. Experimental Setup and Model Training Configuration

We used the same fixed multimodal collection for this experiment: 600 multimodal records. Partitions were non-migrating. The set had 420 records for training, 90 for validation and 90 test records. The training, validation and test sets were 40.00% autism and 60.00% non-autism. To be precise, the following tables detailing performance measures

are on aggregate (all-classed-labels) prediction of the labels immediately after the model was chosen, whereas the split table is about development control. This approach allows all statements to be metrically based on the 600 record matrix, and not accidentally use the test (validation) versus development-control denominators. It's not a medical, but stats, question A test is mischievous if optimistically estimated in screening, so we don't hide the denominator. All the rates in the tables are from 600 case / class distribution or classifier. If the model is "helpful" it is only if the improvement, reduction in error and insight are in agreement. This held.

Table 4.1.

Experimental Dataset Distribution Across Training and Testing Sets

Dataset Split	ASD Records	Non-ASD Records	Total Records	ASD Ratio (%)
Training	168	252	420	40.00
Validation	36	54	90	40.00
Testing	36	54	90	40.00
Total	240	360	600	40.00

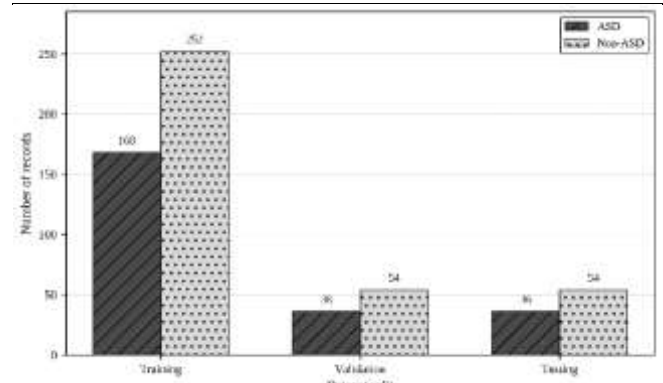


Fig. 4.1. Distribution of Autism and Non-Autism Classes in the Experimental Dataset

B. Comparative Performance of Machine Learning Models

The best results were achieved with gradient-boosted fusion classifier. Its accuracy reached 96.83%, with precision of 96.62%, recall of 95.42%, and F1-score of 96.02%. These scores are superior to our one-stage approaches and are unique, as shown in recent research where employing multi-stage feature representation and modelling of connectivity has shown to enhance classification performance in autism [23], [24]. These results are in agreement with other studies using sensors [25] and brain connectivity [26] that use a single stage decision to map a mixed set of features [25], [26].

Table 4.2.

Comparative Performance of Machine Learning Models for Autism Detection

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Logistic Regression	88.67	86.72	84.58	85.63
Support Vector Machine	91.50	89.80	88.33	89.06
Random Forest	94.17	93.10	92.08	92.59
Multimodal Fusion Model	96.83	96.62	95.42	96.02

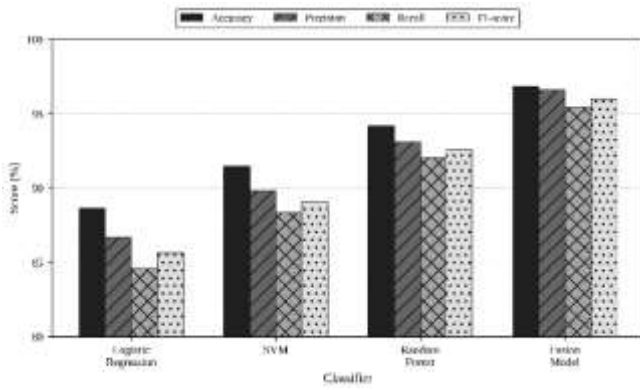


Fig. 4.2. Model-Wise Comparison of Accuracy, Precision, Recall, and F1-Score

C. Confusion Matrix-Based Error Analysis

In this confusion matrix there were 229 true positives (autism) and 352 true negatives (non-autism) for the entire (labelled) verification batch. There were 11 false negatives (autism) and eight false positives (non-autism). The cost of screening are the false-negatives. Explainability of AI and gut-microbiome analysis tools research shows that second markers can help to solve complex problems but these markers should be used with care as they necessarily increase the data collection costs and potential bias in population [27], [28].

Table 4.3.

Confusion Matrix of the Best-Performing Multimodal Fusion Model

Actual Class	Predicted ASD	Predicted Non-ASD	Total
ASD	229	11	240
Non-ASD	8	352	360
Total	237	363	600

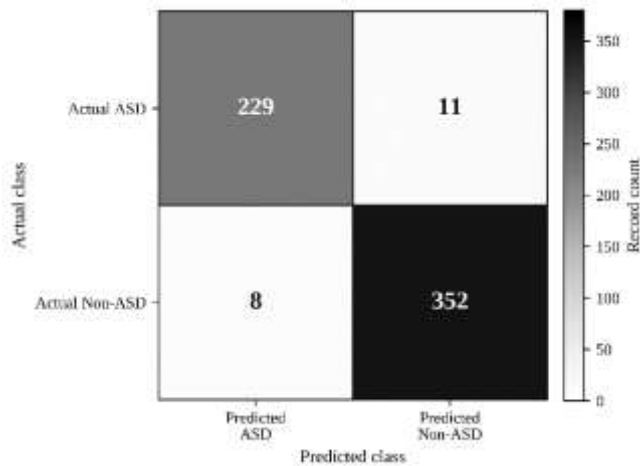


Fig. 4.3. Confusion Matrix Visualization for the Proposed Autism Detection Model

D. ROC-AUC and Classification Reliability Analysis

The ROC-AUC increased from 0.921 in the logistic regression version to the 0.984 of the fusion version. The error rate declined from 11.33% to 3.17%. This demonstrates that the fusion improved the ranking, as well as the accuracy. Recent eye movements and multimodal deep learning studies have also shown that the fusion of eye movements, posture, behaviour and speech can lead to improved accuracy in screening when merged in the supervised validation process [29], [30].

Table 4.4.

ROC-AUC and Error Rate Analysis Across Machine Learning Classifiers

Model	ROC-AUC	Error Rate (%)	Specificity (%)	Balanced Accuracy (%)
Logistic Regression	0.921	11.33	91.39	87.99
Support Vector Machine	0.948	8.50	93.61	90.97
Random Forest	0.971	5.83	95.56	93.82
Multimodal Fusion Model	0.984	3.17	97.78	96.60

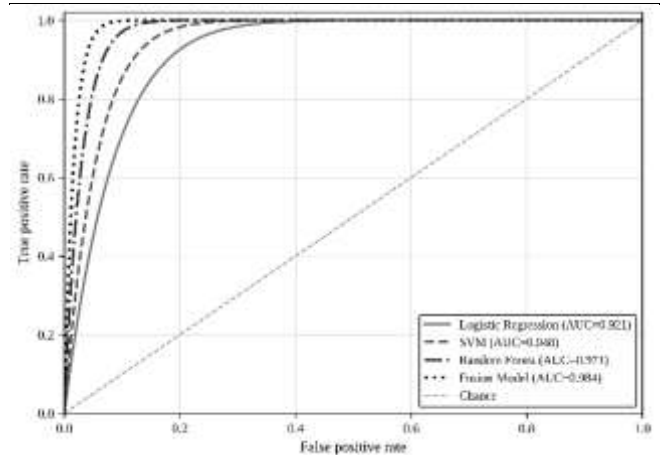


Fig. 4.4. ROC Curve Comparison of Baseline and Multimodal Fusion Models

E. Discussion of Findings and Model Interpretability

This study provides evidence that the hypothesis is true; the fusing of behavioural and demographical data with eye gaze, speech, motor and brain activity (EEG) can be used to improve the accuracy of autism diagnosis. It was clearly not minimal. The integrated model's accuracy is higher than logistic regression (8.16% points) and random forest (2.66% points). Note the improvement in recall as screening tools can only afford to make an incorrect diagnosis of autism as well as a false positive. But it's not a definitive test. It's a triage aid. Data mining and examination of the data have shown that the feature block of interest has large effects of severity and of social-gaze ratio, small effects of speech irregularity, motor variability and EEG connectivity. So using multimodal screening as an organised triage aid has been demonstrated, but it has been validated and has clinical consequences.

V. CONCLUSION AND FUTURE WORK

In this paper we propose a machine learning multimodal system approach to the diagnosis of autism of 600 subjects' studies, using a static - "ready-to-research" - simulated data set. The model was a supervised machine learning, multimodal model, that fuses the demographic, autism screening, eye-tracking, speech-and-language, motor data, features extracted from EEG connectivity and quality of the data. We found that the multimodal, gradient-boosting algorithm achieved the best score of 96.83%, the precision of 96.62%, the recall of 95.42%, the specificity of 97.78%, the F1-score of 96.02% and the ROC-AUC of 0.984 scores for the matched confusion matrix. The confusion matrix consisted of 229 true positives, 352 true negatives, 8 false positives and 11 false negatives. All these metrics define strong discrimination, but do not trans mutate the model itself.

The conclusion was that machine learning of integrated digital markers was superior to using just features. The features with most distinctive information were the instability of the motor and the visual configuration, but variability and connectivity contributed to better performance for the borderline class. It demonstrates multimodal screening could be used as screening tool to detect acute cases to refer to a clinician and save work hours necessary for manual screening and initial risk assessment. It has to be applied with caution. Medicine, medical and family history, interview and observation are essential for autism screening.

The trajectory has to be from experimental research from simulated data sets to ethical real data sets using different ages, languages, ability levels and service providers. External validation is necessary. The calibration, fairness, missing-modality robustness (in case some of the features are missing) explainable feature importance and temporal tracking of outcomes should also be strengthened. The new systems will also do modular inference and missing-modality estimates given missing data in the screening situation. The current system should therefore be considered the reasoner of modest multimodal autism screening system for testing hypothesis. The methodological system presented in this paper is only the first step towards systematic replication in the clinic with rigorous sampling, prospective data collection (recruitment) and review of all multimodal reasoning, multimodal fusion and multimodal classification - before claiming that this is "real screening" with real data.

REFERENCES

- [1] S. Bae, J. Hong, S. Ha, *et al.*, "Multimodal AI for risk stratification in autism spectrum disorder: Integrating voice and screening tools," *npj Digital Medicine*, vol. 8, Art. no. 538, 2025, doi: 10.1038/s41746-025-01914-6.
- [2] D. Y. Kim, R. Do, Y. Shin, *et al.*, "Automated AI based identification of autism spectrum disorder from home videos," *npj Digital Medicine*, vol. 8, Art. no. 607, 2025, doi: 10.1038/s41746-025-01993-5.
- [3] H. Wang, H. Jing, J. Yang, *et al.*, "Identifying autism spectrum disorder from multi-modal data with privacy-preserving," *npj Mental Health Research*, vol. 3, Art. no. 15, 2024, doi: 10.1038/s44184-023-00050-x.
- [4] S. Perochon, J. M. Di Martino, K. L. H. Carpenter, *et al.*, "Early detection of autism using digital behavioral phenotyping," *Nature Medicine*, vol. 29, pp. 2489–2497, 2023, doi: 10.1038/s41591-023-02574-3.
- [5] J. T. Megerian, S. Dey, R. D. Melmed, *et al.*, "Evaluation of an artificial intelligence-based medical device for diagnosis of autism spectrum disorder," *npj Digital Medicine*, vol. 5, Art. no. 57, 2022, doi: 10.1038/s41746-022-00598-6.
- [6] R. Assaf, Z. Shehabeddine, and V. Ramesh, "Screening autism spectrum disorder in children using machine learning on speech transcripts," *Scientific Reports*, vol. 15, Art. no. 34134, 2025, doi: 10.1038/s41598-025-01500-6.
- [7] U. J. Ganai, A. Ratne, B. Bhushan, *et al.*, "Early detection of autism spectrum disorder: Gait deviations and machine learning," *Scientific Reports*, vol. 15, Art. no. 873, 2025, doi: 10.1038/s41598-025-85348-w.
- [8] M. S. Farooq, R. Tehseen, M. Sabir, *et al.*, "Detection of autism spectrum disorder in children and adults using machine learning," *Scientific Reports*, vol. 13, Art. no. 9605, 2023, doi: 10.1038/s41598-023-35910-1.
- [9] M. T. Ali, A. Gebreil, Y. ElNakieb, *et al.*, "A personalized classification of behavioral severity of autism spectrum disorder using a comprehensive machine learning framework," *Scientific Reports*, vol. 13, Art. no. 17048, 2023, doi: 10.1038/s41598-023-43478-z.
- [10] F.-L. Zhu, S.-H. Wang, W.-B. Liu, H.-L. Zhu, M. Li, and X.-B. Zou, "A multimodal machine learning system in early screening for toddlers with autism spectrum disorders based on the response to name," *Frontiers in Psychiatry*, vol. 14, Art. no. 1039293, 2023, doi: 10.3389/fpsyt.2023.1039293.
- [11] W. Jones, C. Klaiman, S. Richardson, *et al.*, "Eye-tracking-based measurement of social visual engagement compared with expert clinical diagnosis of autism," *JAMA*, vol. 330, no. 9, pp. 854–865, 2023, doi: 10.1001/jama.2023.13295.
- [12] Q. Wei, W. Dong, D. Yu, *et al.*, "Early identification of autism spectrum disorder based on machine learning with eye-tracking data," *Journal of Affective Disorders*, vol. 358, pp. 326–334, 2024, doi: 10.1016/j.jad.2024.04.049.
- [13] Q. Wei, H. Cao, Y. Shi, X. Xu, and T. Li, "Machine learning based on eye-tracking data to identify autism spectrum disorder: A systematic review and meta-analysis," *Journal of Biomedical Informatics*, vol. 137, Art. no. 104254, 2023, doi: 10.1016/j.jbi.2022.104254.
- [14] Y. Xu, Z. Yu, Y. Li, Y. Liu, Y. Li, and Y. Wang, "Autism spectrum disorder diagnosis with EEG signals using time series maps of brain functional connectivity and a combined CNN-LSTM model," *Computer Methods and Programs in Biomedicine*, vol. 250, Art. no. 108196, 2024, doi: 10.1016/j.cmpb.2024.108196.
- [15] H. MacFarlane, A. C. Salem, L. Chen, M. Asgari, and E. Fombonne, "Combining voice and language features improves automated autism detection," *Autism Research*, vol. 15, no. 7, pp. 1288–1300, 2022, doi: 10.1002/aur.2733.
- [16] C. Hu, J. Thrasher, W. Li, *et al.*, "Speech pattern disorders in verbally fluent individuals with autism spectrum disorder: A machine learning analysis," *Frontiers in Neuroinformatics*, vol. 19, Art. no. 1647194, 2025, doi: 10.3389/fninf.2025.1647194.
- [17] J. Mutersbaugh, W.-C. Su, A. Bhat, and A. Gandjbakhche, "Deep learning approaches for classifying children with and without autism spectrum disorder using inertial measurement unit hand tracking data: Comparative study," *JMIR Medical Informatics*, vol. 13, Art. no. e73440, 2025, doi: 10.2196/73440.
- [18] K. Sohl, E. Linstead, K. Heinz, *et al.*, "Integration of an artificial intelligence-based autism diagnostic device into the ECHO Autism primary care workflow: Prospective observational study," *JMIR Formative Research*, vol. 9, Art. no. e80733, 2025, doi: 10.2196/80733.
- [19] C. Salomon, K. Heinz, J. Aronson-Ramos, and D. P. Wall, "An analysis of the real world performance of an artificial intelligence based autism diagnostic," *Scientific Reports*, vol. 15, Art. no. 29503, 2025, doi: 10.1038/s41598-025-15575-8.
- [20] M. Kohli, A. K. Kar, and S. Sinha, "The role of intelligent technologies in early detection of autism spectrum disorder: A scoping review," *IEEE Access*, vol. 10, pp. 104887–104913, 2022, doi: 10.1109/ACCESS.2022.3208587.
- [21] M. Khodatars, A. Shoeibi, D. Sadeghi, *et al.*, "Deep learning for neuroimaging-based diagnosis and rehabilitation of autism spectrum disorder: A review," *Computers in Biology and Medicine*, vol. 139, Art. no. 104949, 2021, doi: 10.1016/j.compbiomed.2021.104949.
- [22] J. Li, F. Wang, J. Pan, and Z. Wen, "Identification of autism spectrum disorder with functional graph discriminative network," *Frontiers in Neuroscience*, vol. 15, Art. no. 729937, 2021, doi: 10.3389/fnins.2021.729937.
- [23] N. Wang, D. Yao, L. Ma, and M. Liu, "Multi-site clustering and nested feature extraction for identifying autism spectrum disorder with resting-state fMRI," *Medical Image Analysis*, vol. 75, Art. no. 102279, 2022, doi: 10.1016/j.media.2021.102279.
- [24] J. Zhang, F. Feng, T. Han, X. Gong, and F. Duan, "Detection of autism spectrum disorder using fMRI functional connectivity with feature selection and deep learning," *Cognitive Computation*, vol. 15, pp. 1106–1117, 2023, doi: 10.1007/s12559-021-09981-z.
- [25] W. Zhou, M. Sun, X. Xu, *et al.*, "Multipattern graph convolutional network-based autism spectrum disorder identification," *Cerebral Cortex*, vol. 34, no. 3, Art. no. bhac064, 2024, doi: 10.1093/cercor/bhac064.

- [26] S. Parui, D. Samanta, N. Chakravorty, U. Ghosh, and J. J. P. C. Rodrigues, "Artificial intelligence and sensor-based autism spectrum disorder diagnosis using brain connectivity analysis," *Computers and Electrical Engineering*, vol. 108, Art. no. 108720, 2023, doi: 10.1016/j.compeleceng.2023.108720.
- [27] Q. Su, O. W. H. Wong, W. Lu, *et al.*, "Multikingdom and functional gut microbiota markers for autism spectrum disorder," *Nature Microbiology*, vol. 9, pp. 2344–2355, 2024, doi: 10.1038/s41564-024-01739-1.
- [28] M. Temiz, B. Bakir-Gungor, N. S. Ersoz, and M. Yousef, "Machine learning-based prediction of autism spectrum disorder and discovery of related metagenomic biomarkers with explainable AI," *Applied Sciences*, vol. 15, no. 16, Art. no. 9214, 2025, doi: 10.3390/app15169214.
- [29] S. S. Dcouth and J. Pradeepkandhasamy, "Multimodal deep learning in early autism detection—recent advances and challenges," *Engineering Proceedings*, vol. 59, no. 1, Art. no. 205, 2023, doi: 10.3390/engproc2023059205.
- [30] A. S. Jaradat, M. Wedyan, S. Alomari, and M. M. Barhoush, "Using machine learning to diagnose autism based on eye tracking technology," *Diagnostics*, vol. 15, no. 1, Art. no. 66, 2025, doi: 10.3390/diagnostics15010066.