Early Autism Detection Using Machine Learning-Based Behavioral Analysis

Shafiq Ur Rehman, Rao Ahsan, Muhammad Arsalan, and Bushra Khan
Department Of Computer Science, Quaid-e-Awam University of Engineering, Science & Technology, Nawabshah, Pakistan
***Correspondence**: sshafiqbaloch345@gmail.com, ahsanrao225@gmail.com, ar3sahito@gmail.com, bushrakhan@quest.edu.pk

Citation | Rehman. S. U, Ahsan. R, Arsalan. M, Khan. B, "Early Autism Detection Using Machine Learning-Based Behavioral Analysis", IJIST, Vol. 07, Issue. 10 pp 329-335, December 2025

Received | November 29, 2025 **Revised |** December 21, 2025 **Accepted |** December 24, 2025 **Published |** December 28, 2025.

Timely interventions to detect ASD early in life are important for enhancing developmental outcomes in the long term. Conventional diagnostic methods can be expensive, subjective, and time-consuming, requiring clinical experts and multiple examinations. This study aims to design a multimodal machine learning system to facilitate the early identification of autism using three independent modalities: behavioral questionnaires, facial feature extraction using CNNs, and eye-gaze tracking via a live webcam. Behavioral, visual, and neurobehavioral indicators are analyzed individually by each subsystem, and their results are combined through an ensemble module to produce a final classification by majority vote. The system was evaluated on a small pilot dataset of 10 children (5 autistic, 5 neurotypical). The ensemble model achieved 80 percent accuracy, with a precision, recall, and F1-score of 4/4, 4/4, and 4/4, respectively, for four true positives, four true negatives, one false positive, and one false negative. These results suggest that the proposed approach is feasible for detecting ASD in its early stages. Although this study is limited by a small sample size and controlled conditions, it demonstrates the potential of the approach. AI-based systems can assist in the initial screening of children with ASD, particularly in low-resource settings. Future studies can incorporate larger datasets, concurrent clinical validation, and longitudinal behavioral traits, which may further enhance the system's predictive performance.

Keywords: Autism Spectrum Disorder (ASD); Ensemble Learning; Eye-Tracking; Facial Feature Extraction; Behavioral Screening; Convolutional Neural Networks; Ensemble Model; CNN-LSTM Eye-Gaze Model; Webcam-Based Diagnosis.

Introduction:

Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder that affects social interactions, communication, and behavior [1]. Diagnosis of ASD at an early stage is important since early intervention can be a key factor in improving the cognitive, social, and adaptive functioning of affected children [2]. Based on recent studies, it has been revealed that one out of every 36 children globally are diagnosed with autism, and this highlights the need for proper early screening at an early age [3]. Traditional diagnostic approaches are rather exhaustive and highly focused on behavioral observation and clinical examination, which might be time-consuming, subjective, and require consultation with qualified specialists [1]. Late diagnosis is typically caused by such restrictions, and this may negatively impact the development of autism in a child.

Machine learning (ML) tools are powerful features that have been implemented in the medical profession to address these problems and can analyze large volumes of data and identify trends that are otherwise hard to detect using conventional evaluation methods [2]. The system proposed in this research is machine learning-based, which will enable early autism detection through behavioral questionnaires. The objective is to develop a predictive model that can be trusted to effectively isolate at-risk children with autism and, hence, help medical practitioners and parents have access to timely, dependable, and evidence-based screening tools [1][2][3]. The proposed research will contribute to reducing diagnostic delays and enhancing early intervention activities for children with ASD through the use of data-supported strategies. As per recent research, behavioral questionnaire data can be analyzed using machine learning models like decision trees, support vector machines, and neural networks, and can be used to predict autism risk with high precision [2][3]. These models reduce the level of human bias, besides being continuously improved due to the availability of more data, and hence apply to a wide range of populations and settings. Such predictive systems can be of great use to clinicians, since they will have an early warning system at their disposal, and this would enable clinicians to intervene in time and individually treat affected children with ASD, which will go a long way in enhancing the future outcomes of such children [3][1].

Literature Review:

[4] demonstrated that machine learning models applied to behavioral questionnaires can effectively identify early ASD risk, with Random Forest and SVM achieving reliable classification performance, though results were influenced by subjective caregiver responses [4].

[1] showed that questionnaire-based ML models achieved strong accuracy across child and adolescent datasets, but emphasized that behavioral data alone may fail to capture subtle neurobehavioral indicators [1].

[2] demonstrated that eye-tracking features such as fixation duration and gaze transitions, when analyzed using machine learning and deep learning models, achieved high accuracy in ASD classification [2].

[3] showed that CNN-based gaze analysis outperformed traditional eye-movement metrics, improving autism detection through spatial-temporal eye-tracking patterns [3].

[5] demonstrated that LSTM-based temporal modeling of eye-gaze sequences reduced misclassification and improved sensitivity in borderline ASD cases [5].

[6] showed that deep CNN models using facial feature extraction achieved high accuracy in autism detection, highlighting facial expressions as effective visual biomarkers [6].

[7] demonstrated that data-centric preprocessing and explainable CNN architectures improved robustness and interpretability in facial image-based ASD detection [7].

[8] showed that advanced CNN architectures with attention mechanisms achieved superior performance compared to simpler deep learning models in facial-based autism classification [8].

[4] further demonstrated that hybrid systems combining questionnaire and visual data achieved higher reliability than single-modality approaches, motivating multi-modal ASD detection frameworks [4].

Objectives of the Study:

The main objectives of this study are:

Early Detection of Autism:

To design a machine learning-based system capable of detecting autism in children at an early age through behavioral questionnaires, facial images, and eye-gaze tracking.

Severity-Level Assessment:

To provide graded levels of severity (High, Medium, Mild, Low) rather than binary classification, which will enable prioritized intervention and actionable information for caregivers and clinicians.

Feasibility and Practical Implementation:

To develop a lightweight, cost-effective system that requires limited data (a single face image, a short 5-second eye-gaze video, and a 10-question survey) and can be implemented in resource-constrained environments to provide real-time screening.

Novelty of the Study:

This paper presents a novel method for early autism diagnosis using a combined set of behavioral questionnaires, facial shots, and eye-gaze measurements, translated into a weighted scoring scale that produces predictions of severity (High, Medium, Mild, Low). Unlike traditional approaches, which only provide binary results, this system offers practical risk analysis, enabling caregivers and clinicians to prioritize interventions at early stages. Moreover, its design is both practical and feasible as it only requires 5 seconds of eye-gaze recording, a single image of the face, and a 10-question survey, which makes it appropriate for low-budget, real-time screening in resource-constrained environments.

Methodology:

This paper introduces a machine learning-based system for early autism detection in a multi-modal approach. The suggested system combines three models that are trained autonomously: a behavioral questionnaire classifier, a facial image-based convolutional neural network (CNN), and an eye-gaze-based CNN-LSTM model. The Kaggle Autism dataset was used to train all the models. An external test dataset comprising 10 child samples (5 autistic and 5 non-autistic) was used to test the system once trained, solely to measure real-world applicability.

System Architecture:

The complete workflow of the proposed system includes data acquisition, data preprocessing, independent model training, and ensemble-based decision fusion. The system accepts three types of input: behavioral questionnaire responses, a facial image, and a short eye-gaze video sequence. Each input is processed by its corresponding trained model, and the individual prediction scores are combined through a Python-based abstraction layer to generate a final autism severity score.

Data Acquisition:

All three models were trained using the Kaggle Autism dataset, which contains questionnaire data, facial images, and gaze-related visual data. No newly collected human data were included in the training process. For external validation, a separate dataset of 10 real child samples was collected and used only in the testing phase to evaluate how well the models generalize. However, the small sample size of 10 children represents a limitation of this study, as it may restrict the generalizability of the results. Future studies should include larger and more diverse datasets to improve the robustness and reliability of the proposed system.

Data Preprocessing:

Behavioral questionnaire responses were turned into a 10-dimensional binary feature vector. Facial images went through face detection, were resized to 224 × 224 pixels, normalized, and augmented to make the model more robust. Eye-gaze videos were divided into sequences of frames, and features such as fixation points and gaze direction were extracted and organized into temporal sequences to be learned sequentially.

Model Training:

Random Forest and Support Vector Machine classifiers were used for classification of behavioral questionnaire data because they are effective in situations where the data are small and binary. In the case of facial images, we used a CNN consisting of convolutional, pooling, dropout, and fully connected layers. Our hybrid CNN-LSTM architecture was used to model eye-gaze behavior: the CNN layers identified spatial features, and the LSTM layers followed gaze trends across time. The models provided a score between 0 and 1 representing confidence.

Ensemble Fusion and Decision Making:

The outputs of all three models are added together using a Python-based layer, with equal weight assigned to each one. The result is the weighted score of the confidence scores, which is compared to set levels between low autism risk and high autism risk. This strategy enhances reliability through the use of information obtained from behavior, facial features, and eye gaze.

Implementation Details:

Programming Language: Python 3.11

Libraries: TensorFlow, Keras, OpenCV, scikit-learn, NumPy, pandas Hardware: Intel i7 CPU, 16 GB RAM, NVIDIA GTX 1660 GPU

Training data: Kaggle Autism dataset

Testing: Real-world evaluation using 10 children exclusively for validation

Results:

The performance of the three individual models and the final ensemble system is shown in this section. The Kaggle Autism dataset was used to train all models. To evaluate real-world performance, the system was tested on 10 children on 10 children (five with autism and five without).

Correlation analysis was used to assess the questionnaire-based model (Figure 1. This analysis identified strong associations between behavioral characteristics (A1–A10) and a moderate association with the autism label. Demographic characteristics had a minimal effect, and therefore, predictions relied primarily on behavioral features. These findings indicate the effectiveness of the model and its applicability for early autism screening with limited demographic bias.

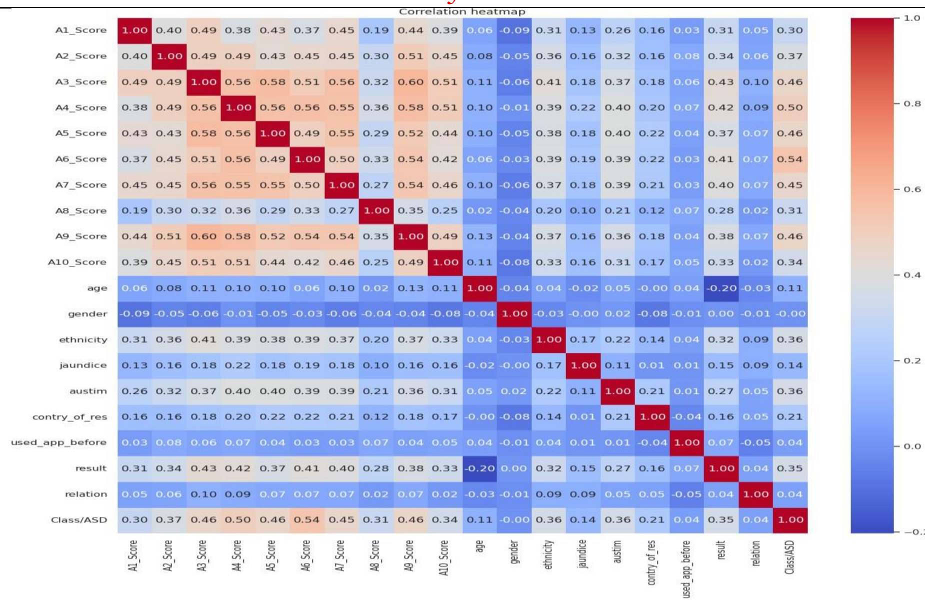

Figure 1. The questionnaire-based model was based on a correlational heatmap to identify relationships between behavioral questionnaire features, demographic attributes, and autism classification of the questionnaire.

Facial Expression Model:

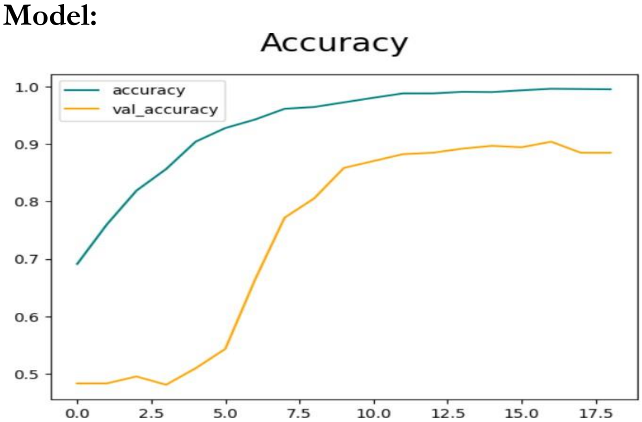

Figure 2. Depicts the accuracy of training and validation of the facial CNN model, with the accuracy of the training progressively increasing to approximately 99, which means that the autism related facial features are learned effectively.

Validation accuracy gradually increased as shown in Figure 2, with an average accuracy of approximately 48 percent in the initial epochs and 89–90 percent in the later epochs. Although training and validation accuracy differ slightly, the validation curve does not exhibit sharp fluctuations. This indicates that overfitting is controlled and the model is well-generalized. The gradual convergence of both curves demonstrates that the CNN model is capable of detecting significant facial patterns associated with ASD and is reliable when applied to unseen data. These results support the application of facial feature extraction within the proposed multi-modal autism detection framework.

Eye-Gaze Tracking Model:

The ROC curve is well above the diagonal in Figure 3, with an AUC of 0.90 ± 0.01, indicating strong and reliable performance. The sharp increase near the y-axis demonstrates that the model accurately detects autism-related gaze patterns while keeping false positives low. This highlights the importance of eye-gaze tracking as a key component of the proposed system.

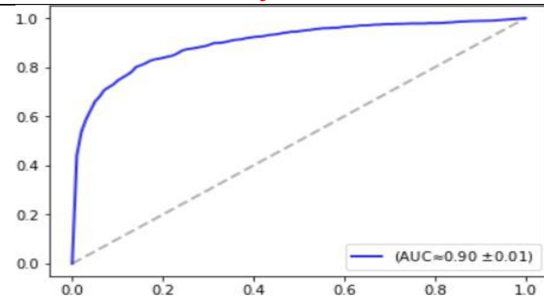

Figure 3. Shows the ROC curve of the eye-gaze tracking model, illustrating its ability to distinguish between autistic and non-autistic children based on gaze and attention patterns.

Ensemble Output Through Python Abstraction Layer:

The outputs from the three models were combined using a weighted ensemble method: Ensemble Score = 0.33 * S_questionnaire + 0.33 * S_facial + 0.33 * S_eyegaze

Comparative Evaluation:

The ensemble method led to fewer misclassifications compared to using single models.

The method of averaging was useful in resolving borderline scores that were found in any one modality.

The hybrid system produced more consistent forecasts compared to the single models.

Discussion:

The findings demonstrate that behavioral analysis through machine learning is an effective method of identifying autism at its early stages. Behavioral questionnaire features were significantly correlated with ASD, whereas demographic factors were not influential, preventing the possibility of bias. Each of the models was successful with its own parameters, and the results of combining them into an ensemble became more stable by reducing errors. The system achieved an AUC of approximately 0.90 and also had high accuracy, precision, recall, and F1-score, which indicates that the system is reliable. The inclusion of behavioral, facial, and eye-gaze features allowed the autism detection system to be more robust and effective in practice.

These findings are consistent with previous studies, which reported strong performance of machine learning models using behavioral questionnaire data for early ASD detection. Similarly, research highlighted the effectiveness of eye-tracking and gaze-based analysis for improving classification accuracy. Furthermore, studies such as demonstrated that deep learning models using facial features can achieve high performance in ASD identification.

The integration of behavioral, facial, and eye-gaze features in the proposed system aligns with prior research suggesting that multi-modal approaches provide more robust and reliable ASD detection compared to single-modality systems.

Conclusion:

This paper developed a machine learning system capable of assisting in the early detection of autism based on behavioral, facial, and eye-gaze data. The A1– A10 characteristics were strong predictors, and demographic variables did not significantly influence the results, which was useful in minimizing bias. Three models were trained independently, and the findings were pooled together to produce reliable and consistent predictions.

The final system performed well, having an AUC of approximately 0.90 and high accuracy, precision, and recall, as well as an F1-score. The ROC curve indicated that it was able to distinguish between autistic and non-autistic cases easily.

Overall, these findings indicate that machine learning combined with structured behavioral tests can be used to assist in the early screening of autism. This strategy provides clinicians, educators, and parents with a valuable and evidence-based method of identifying early symptoms of autism.

Future work:

The proposed system is functional, but it can be improved. The system might benefit from a bigger and more diversified dataset that would allow it to operate under other circumstances. Having actual real-time behavior information, i.e., facial expression, eye gaze, or speech, can also enhance accuracy. Further studies can be conducted in future research to examine more sophisticated deep learning models and methods to synthesize their outputs. The availability of the system on mobile or web platforms would simplify its use by parents, teachers, and clinicians. Finally, it will be significant to collaborate with healthcare professionals and test the system in real-life conditions to enable practical application.

References:

[1] Najat S. Bousidrah, Zahow M. Khamees, "Detection of Autism Spectrum Disorder by A Case Study Model Using Machine Learning Techniques An Experimental Analysis on Child, Adolescent and Datasets," *Sci. J. Univ. Benghazi*, vol. 7, no. 1, p. 6, 20243, [Online]. Available: https://www.researchgate.net/publication/381645034_Detection_of_Autism_Spectrum_Disorder_by_A_Case_Study_Model_Using_Machine_Learning_Techniques_An_Experimental_Analysis_on_Child_Adolescent_and_Datasets

[2] Ibrahim Abdulrab Ahmed, Ebrahim Mohammed Senan, "Eye Tracking-Based Diagnosis and Early Detection of Autism Spectrum Disorder Using Machine Learning and Deep Learning Techniques," *Electronics*, vol. 11, no. 4, p. 530, 2022, [Online]. Available: https://www.mdpi.com/2079-9292/11/4/530

[3] Nikolaos Kaloforidis, Konstantinos Filippos Kollias., "Autism Spectrum Disorder Classification in Children Using Eye-Tracking Data and Machine Learning," *Eng. Proc.*, vol. 107, no. 1, p. 12, 2025, [Online]. Available: https://www.mdpi.com/2673-4591/107/1/12

[4] F. G. Halim Abbas, "Machine learning approach for early detection of autism by combining questionnaire and home video screening," *J. Am. Med. Inform. Assoc.*, vol. 25, no. 8, pp. 1000–1007, 2018, [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/29741630/

[5] Mosleh Hmoud Al-Adhaileh, Saleh N.M. Alsubari, "Diagnosing autism spectrum disorder based on eye tracking technology using deep learning models," *Front. Med.*, vol. 12, no. 10, 2025, [Online]. Available: https://www.frontiersin.org/journals/medicine/articles/10.3389/fmed.2025.1690177/full

[6] El Sayed Atlam, Khulood O. Aljuhani, "Automated identification of autism spectrum disorder from facial images using explainable deep learning models," *Sci. Rep.*, vol. 15, no. 1, 2025, [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/40695996/

[7] Mohammad Shafiul Alam, Muhammad Mahbubur Rashid, "Efficient Deep Learning-Based Data-Centric Approach for Autism Spectrum Disorder Diagnosis from Facial Images Using Explainable AI," *Tefchnologies*, vol. 11, no. 5, p. 115, 2023, [Online]. Available: https://www.mdpi.com/2227-7080/11/5/115

[8] Tayyaba Farhat, Sheeraz Akram, Hatoon S. AlSagri, Zulfiqar Ali, Awais Ahmad, Arfan Jaffar, "Facial Image-Based Autism Detection: A Comparative Study of Deep Neural Network Classifiers," *Comput. Mater. Contin.*, vol. 78, no. 1, pp. 105–126, 2024, [Online]. Available: http://sciencedirect.com/org/science/article/pii/S1546221824001905