Early Detection of Autism Spectrum Disorder in Children Using Facial Images and AI-Based Deep Learning Models

D.N. Vasundhara[1], P. Lakshmi Prasanna[2], Ch.V.K.N.S.N. Moorthy[3*] and N. Venkata Sailaja[4]

[1,2,4]VNR Vignana Jyothi Institute of Engineering and Technology, Hyderabad, 500090, India
[3]Vasavi College of Engineering, Hyderabad, 500031, India

Author Designation: [1,4]Senior Assistant Professor, [2]Assitant Professor, [3]Professor

*Corresponding author: Ch.V.K.N.S.N. Moorthy (e-mail: krishna.turbo@gmail.com).

Abstract The condition in which an individual's communication abilities and social interactions are affected during early development is called autism spectrum disorder (ASD). Such people feel more anxious, react to emotions in different ways and experience sounds, lights or touch differently. Early diagnosis and timely interventions can significantly improve behavioural and language development. Traditional methods for ASD detection, such as subjective questionnaires and behavioural observations, are often time-consuming, costly and reliant on specialized expertise, leading to a delay in intervention. To address these challenges, this work introduces a novel approach for the early identification and detection of ASD in children using facial images. The research focuses on utilizing YOLOv8n for facial features identification and ASD classification, harnessing its advanced object detection capabilities to enhance both accuracy and efficiency. This paper also presents a comparative analysis of YOLOv8n and YOLOv9c to conclude with the most effective solution. This study facilitates real-time ASD screening, providing healthcare professionals with timely and effective interventions in clinical environments.

Key Words Autism Spectrum Disorder (ASD), Facial Image Analysis, YOLO, Real-time screening, AI-based Diagnostics

INTRODUCTION

The condition in which an individual's communication abilities and social interactions are affected during early development is called autism spectrum disorder (ASD). Such people feel more anxious, react to emotions in different ways and experience sounds, lights or touch differently [1]. Children with autism exhibit a range of symptoms, which are more common, but not everyone with autism has the same ones. While certain individuals with ASD may also experience cognitive impairment or disabilities, the majority possess average to above-average intelligence. This condition, which typically emerges within the first three years of life, is complex and often characterized by difficulties in behaviour and communication. The severity varies from person to person; some of them face mild difficulties, while others require full-time care due to severe impairments [1]. The child with Autism requires more family support. ASD is typically identified in children between the ages of two and three [2].

Growing Prevalence of ASD Around the World
The prevalence of AD has been rising on a global scale. The term "prevalence" refers to the proportion of individuals within a population who are affected by a particular condition at a specific time. It is typically represented either as a percentage (e.g., 1%) or a ratio (e.g., 1 in 100). As of 2022, approximately 1 in 31 children in the US were diagnosed with Autism, according to the CDC (Centers for Disease Control and Prevention), most recent surveillance data gathered through the ADDM (Autism and Developmental Disabilities Monitoring Network). According to a 2023 ETHealthWorld research, over 18 million Indians were diagnosed with autism disorder. Moreover, projections by the Indian Academy of Pediatrics (IAP) in 2024 estimate that approximately 3% of India's population [3] has been diagnosed with ASD. This marks a continued increase in autism prevalence compared to previous years. The prevalence of ASD is not limited by demographic boundaries; it has been observed across various cultural, racial and economic backgrounds. However, this condition is notably three times more likely among boys than girls. The child with ASD may include other developmental conditions such as ADHD (Attention Deficit and Hyperactivity Disorder) and cerebral palsy, visual impairments.

Vasundhara *et al.*: Early Detection of Autism Spectrum Disorder in Children Using Facial Images and AI-Based Deep Learning Models

jpms

Table 1: Historical Trends in Autism Prevalence

Year of Surveillance	Year of Birth	Number of Data Collection Sites	Prevalence per 1000 (Range of Estimates)	Estimated Ratio (1 in X Children)
2022	2014	16	32.2 (9.7-53.1)	1 in 31
2020	2012	11	27.6 (23.1-44.9)	1 in 36
2018	2010	11	23.0 (16.5-38.9)	1 in 44
2016	2008	11	18.5 (18.0-19.1)	1 in 54
2014	2006	11	16.8 (13.1-29.3)	1 in 59
2012	2004	11	14.5 (8.2-24.6)	1 in 69
2010	2002	11	14.7 (5.7-21.9)	1 in 68
2008	2000	14	11.3 (4.8-21.2)	1 in 88
2006	1998	11	9.0 (4.2-12.1)	1 in 110
2004	1996	8	8.0 (4.6-9.8)	1 in 125
2002	1994	14	6.6 (3.3-10.6)	1 in 150
2000	1992	6	6.7 (4.5-9.9)	1 in 150

Table 1 outlines the historical progression of autism prevalence among eight-year-old children, based on ADDM Network surveillance across various years [4].

Symptoms and Signs of Autism in Children

Identifying ASD based on symptoms is a difficult task. As the symptoms are not same in all the individuals. The signs of autism include difficulties with Social Communication and Interaction, repetitive/restricted patterns of behaviour and interest, delayed speech and language development. Further, these symptoms also appear sometimes in children without ASD too. The variation is the frequency of the appearance of such symptoms. In addition to that, the challenges and their severity is also not same in the effected children. Some may be affected very little where as others may need continuous support to overcome it. Consequently, the complexity of detecting and diagnosing is very high especially during early childhood. The following are a few challenges faced by children with ASD.

Diagnosis Process of ASD

Autism diagnosis for children involves a step-by-step detection that will be conducted by psychologists for children aged 2-3 years. In general, parents and caretakers begin to identify the signs of autism, like communication problems, social interaction and behaviours, which are usually identified during checkups, at school or home. For all these kinds of issues, paediatricians or psychologists perform developmental screening, typically around 18-24 months, to isolate delays in speech, motor skills or social behaviour. A detailed further assessment is to be performed by the specialists, observing the child's detailed development history and standardized diagnostic tools.

Existing ASD diagnostic methods include several well-known tools and procedures. One important and widely used tool for initial autism assessment is the M-CHAT (Modified Checklist for Autism in Toddlers). It is a list of questionnaires filled out by parents to identify early indications of autism in children. There is another assessment called ADOS (Autism Diagnostic Observation Schedule) that involves a play-based and interactive evaluation process for analysing communication skills and repetitive behaviours. The ADI-R (Autism Diagnostic Interview-Revised) involves a structured conversation with parents and caregivers to document comprehensive information about the child's developmental background and behavioural patterns of autism. CARS (Childhood Autism Rating Scale) is used to identify the severity of symptoms. Using these tools, paediatricians or psychologists diagnose and plan for the treatment [5].

The conventional process of autism, above methods have significant challenges. A major limitation is a huge dependency on assessments and parents/caregivers' inputs. The observations and the inputs are captured to diagnose the symptoms based on the individual interpretation may take time, lead to inconsistent outcomes and incur more cost for the families. Along with these problems, the availability of experts who can perform diagnostic services is very limited in many regions. Lack of access makes it difficult to assess the children. As a result, most of the children are not able to be diagnosed at an early stage of autism. These challenges highlighted the requirement for faster, more objective and accessible approaches to ASD diagnosis.

Automation of the ASD screening process offers a promising approach. AI-based systems leverage computer vision. Deep learning and machine learning techniques are used to examine facial features or behavioural cues. Reducing human intervention and subjectivity, these systems can help produce consistent, repeatable results and operate much faster than the traditional evaluation process. The major advantage is that it can be deployed as either mobile or web-based applications to improve accessibility even in remote areas. Automation enables early and accurate detection of the symptoms, enabling timely interventions, which are very important for enhanced outcomes in children with ASD. By reducing the need for prolonged expert involvement and enabling a mass-scale screening process, AI-based automated solution fills the gap in early diagnosis and support more inclusive healthcare solutions.

Literature Summary

Autism Spectrum Disorder aims for early diagnosis through an automated process subjected to various computational studies. Multiple computational techniques, including machine learning, deep learning and object detection frameworks, are being used for Facial image-based detection (Table 2).

Logistic regression model has demonstrated potential in identifying ASD from behavioural datasets, emphasizing the

Vasundhara *et al.*: Early Detection of Autism Spectrum Disorder in Children Using Facial Images and AI-Based Deep Learning Models

joms

Table 2: Comparison of Relevant Studies on ASD Detection Using Facial Images

Ref.	Authors	Method Used	Dataset Used and Source	Limitation of the study	Accuracy (as per paper)
[1]	Reham Hosney, Fatma M. Talaat, Eman M. El-Gendy, Mahmoud M. Saafan	Attention-based YOLOv8 (AutYOLO-ATT) for facial expression recognition	Custom facial expression dataset of autistic and typical children (6 emotions)	Limited dataset diversity, potential challenges in generalizing to varied real-world conditions	97.2%
[2]	Akhil Kumar, Ambrish Kumar, Dushantha Nalin K. Jayakody	Enhanced YOLOv7-tiny model	Self-annotated autism face dataset (Publicly available)	Model Complexity higher, limited generalization to diverse expressions	Mean Average Precision (mAP): 79.56%; Intersection over Union (IoU): 51.99%
[18]	Diwan, T., Anirudh, G., Tembhurne, J.V.	YOLO-based comparative review	Multiple datasets	General survey; no implementation or real-time deployment	–
[19]	Hussain, M.	YOLOv1-YOLOv8 comparative review	Multiple datasets	Earlier YOLO versions have poor object detection, low recall	-
[17]	Chakradhar, K., Tharun. K, Reddy, Somasundaram	CNN based Deep Learning Models (YOLOv8n, Detectron2, VGG16, ResNet50, Xception, Inceptionv3 and MobileNetV2)	Kaggle ASD facial image dataset	Overfitting risk with limited data	94%
[10]	Khosla Y., Ramachandra P., Chaitra N.	MobileNet Transfer Learning	Kaggle ASD Face dataset	Overfitting in small datasets; poor generalization	87%
[11]	Pranavi Reddy and Andrew J.	VGG16, VGG19, EfficientNetB0	Kaggle	VGG models are computationally expensive; less efficient in real-time	87.9%
[12]	Awaji, B., Senan, E.M., Olayah, F.*et al.*	Hybrid CNN feature extraction (VGG16, ResNet101, MobileNet) combined with M.L classifiers (XGBoost, Random Forest)	Kaggle	Model interpretability is low; high training time	98.8% (best hybrid RF model with VGG16-MobileNet features)
[16]	Alkahtani, H., Aldhyani, T.H.H. and Alzahrani, M.Y.	Deep Learning using MobileNetV2 on Facial landmarks	Kaggle	Small and Limited dataset, Lack of clinical validation, No multimodal analysis	92%
[20]	Uddin, M.Z., Shahriar, M.A., Mahamood, M.N.*et al.*	Systematic review of deep learning approaches for image/video-based ASD detections.	Utilizes various public and private datasets (2017-June 2023)	Data Limitation, Lack of standardization	-

significance of complete diagnostic information for improved outcomes [6]. Feature selection techniques like Chi-Square and information gain have been utilized to recognize the minimal subsets of diagnostic traits [7]. To reduce the diagnostic burdens [8], classifier-based recommender systems with integrate decision trees and random forests have enhanced the screening efficiency. Especially to highlight the importance of multiple data fusion for robust diagnosis [9], comparative evaluation using SVM, CNN on ABIDE and NDAR datasets has been used. Privacy-preserving methods like federated learning were introduced that support the classification while maintaining data confidentiality across distributed sources [10]. Machine learning methods have helped in autism detection using behavioural data, but they often struggle to work with complex image data. These models usually need manual feature selection, which limits their ability to detect subtle facial cues. To overcome this, recent studies have moved towards deep learning models like CNNs, which can automatically learn important features from facial images and offer better accuracy.

Deep learning approaches use pre-trained convolutional neural networks to classify ASD. Transfer learning using architectures like MobileNet and InceptionV3 has demonstrated strong performance. MobileNet offers advantages in latency and real-time applications [11]. In a

study [12], a comparison of VGG16, VGG19 and EfficientNetB0 is performed and it is concluded that EfficientNetB0 is effective among them. Feature fusion strategies combining CNN embedding with ensemble classifiers, such as XGBoost and Random Forest, along with t-SNE-based dimensionality reduction, have refined classification outcomes [13]. Lightweight CNN such as DenseNet121 and EfficientNetB0 have also been assessed, reinforcing the suitability of convolutional approaches for facial analysis [14]. Additionally, investigations into eye-tracking data have revealed useful behavioural indicators, though challenges like inconsistency and lack of standardization persist [15]. A thorough analysis of neuroimaging-based deep learning models has highlighted the importance of model interpretability and the difficulties arising from data heterogeneity in medical imaging [16]. A mobile app that incorporates MobileNet and facial landmark analysis has shown a very good potential for practical use, though data quality and generalization continue to be crucial factors [17]. While CNN-based models have improved image classification, they mainly give a single prediction for the whole image and don't point out where the features are. In real-time applications or detailed analysis, it's not suitable. So, researchers used object detection methods like YOLO, which can detect and classify features at the same time and are also faster for real-time use.

Vasundhara *et al.*: Early Detection of Autism Spectrum Disorder in Children Using Facial Images and AI-Based Deep Learning Models

jpms

Object detection techniques, utilizing YOLO architectures, have been extensively employed. Enhancements using attention mechanisms in YOLOv8 have facilitated the classification of facial expressions linked to ASD, effectively detecting subtle emotional cues [1]. Enhancements to YOLOv7-tiny, such as the inclusion of dilated convolutions and extra detection heads, have facilitated the efficient identification of visual features specific to ASD [2]. The combination of YOLOv8n and Detectron2 for facial pattern recognition has aided in distinguishing between autistic and neurotypical traits [18]. YOLO is considered the best architecture for real-time processing and an effective balance between speed and efficiency [19]. It has integrated innovations such as anchor-free detection, spatial pyramid pooling and transformer modules, greatly enhancing its capability in image analysis tasks [20].

In addition to individual models, integrated systems are suggested for ASD detection and a comprehensive review of recent approaches entitled image and video-based deep learning systems as a crucial tool for early diagnosis. It also addresses challenges like limited data availability and the absence of standardized benchmarks [21]. Review of comparing technological tools to traditional diagnostic methods entitled enhanced accessibility and operational efficiency, mainly in environments with limited resource availability [22]. A comprehensive system integrating mobile applications, web-based inference engines and self-learning features has shown promise for scalable implementation and user interaction [5].

There are Many existing approaches depend on behavioural data or facial expression images that often suffer from inconsistency and lack of clarity. Although considerable advances have been made in ASD detection using computational methods, CNN and related deep learning models shows promising result, but challenges such as limited dataset diversity, variable image quality and difficulties in real-time implementation remains the same [22]. These issues highlight the need for more reliable and versatile systems capable of capturing subtle facial characteristics across diverse conditions. This motivates the development of improved methods that enhance early and accurate screening of ASD in children.

Dataset Description
This study utilizes a publicly available dataset from the Roboflow platform, which includes 14,352 annotated facial images of children that have been labelled as autistic or non-autistic. Each image contains bounding box annotations to assist object detection and classification tasks. The dataset is divided into three groups for testing, validation and training. Such splitting reduces the bias and overfitting possibility since the models are trained and evaluated on separate balanced sets [23].

The dataset utilized in this research is publicly available. It comprises only de-identified facial images with no personally identifiable information (PII). According to the Roboflow dataset sharing policies, it is understood that before making this dataset public, the curators obtained adequate consent and ethical clearance. Given the child's facial data aspect, there is ethical sensitivity. There was no personal or clinical collection for this research.

Data Preprocessing and Augmentation
The facial image dataset for this research was obtained from Roboflow, which provides built-in preprocessing and augmentation features at the time of dataset export. Every image was automatically adjusted to a fixed resolution of 640×640 pixels, converted into RGB colour format and normalized so that input shape and colour are consistent across the dataset. To enhance the model's ability to generalization and reduce impact of variations in illumination, facial expressions and head pose, the Roboflow augmentation pipeline was applied. These include horizontal flipping, random rotations, brightness and contrast adjustment, cropping and zooming. These augmentations help the model focus on learning facial features relevant for classification. There was no need for additional manual preprocessing because these augmentations take place during dataset export within Roboflow. every image was automatically adjusted to a fixed resolution of 640×640 pixels.

Proposed Framework for Autism Detection by AI
In this proposed study, facial images of children are analysed using deep learning techniques to support early autism screening. Two object detection models, YOLOv8n and YOLOv9c, were applied to classify the images as autistic or non-autistic. These models were trained on publicly available annotated datasets that include bounding box labels on the facial images. YOLO was preferred over other models because it offers real-time detection along with high classification accuracy. Both models effectively capture subtle facial cues associated with ASD and perform significantly better than conventional CNN models.

System Architecture
The Yolo Model: You Only Look Once (YOLO) is a real-time object detection model known for its accuracy, efficiency and speed. As a one-stage architecture, YOLO performs both classification and localization in a single forward pass of the network. This streamlined design minimizes inference time, making it highly appropriate for applications in real time. The model divides the input image into multiple regions and predicts bounding boxes along with corresponding class scores, enabling quick object detection with minimal computational load. In the diagnosis of autism in children, Yolov8 efficiency enables timely and early detection.

YOLOv8 Architecture
YOLOv8, the latest iteration in the YOLO family, introduces a more refined and efficient architecture designed for object identification, image classification and segmentation. It builds

Vasundhara *et al.*: Early Detection of Autism Spectrum Disorder in Children Using Facial Images and AI-Based Deep Learning Models
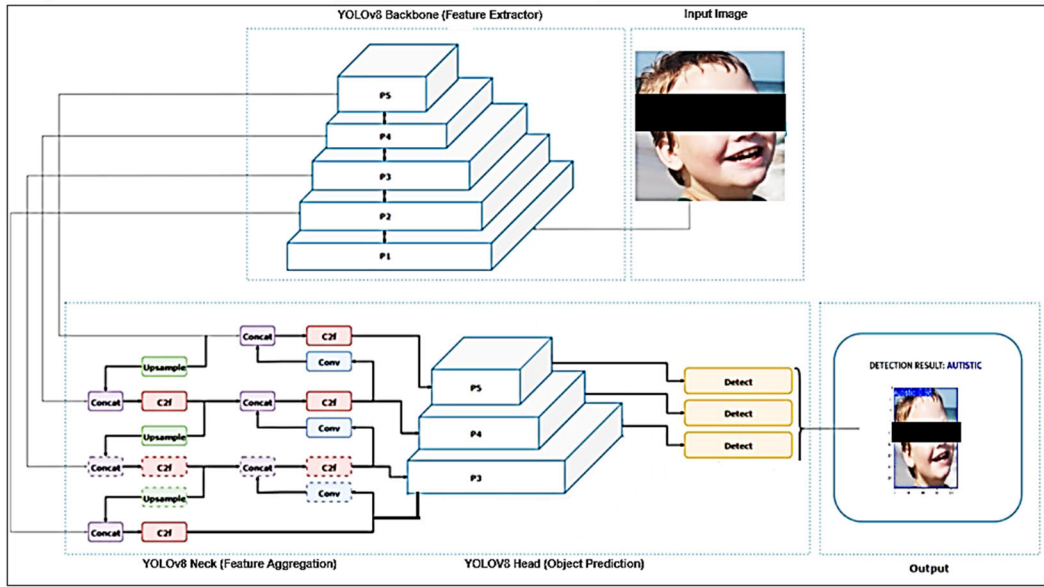
jpms

Figure 1: Modified YOLOv8 architecture for autism detection. The backbone extracts hierarchical features (P1–P5) to identify facial symmetry and textures

upon previous YOLO versions by incorporating modern deep learning techniques, including an anchor-free design, decoupled head and lightweight modules, making it faster and more accurate than its predecessors. YOLOv8 is particularly suitable for edge deployments and medical uses because of its compact size and high functionality. As shown in Figure 1, the structure of YOLOv8 contains three main components: the feature-extraction Backbone, the transitional Neck module and the output-generating Head [24].

Backbone

This component handles the initial stage of processing by extracting relevant features from the input image. YOLOv8 utilizes a CSPDarknet-based structure enhanced with convolutional blocks and activation functions. It captures low-level and high-level features from the input image and forwards them to the next stage. Each convolutional layer in the Backbone processes the input feature map using the standard formula:

$$\text{Output Size} = \left\lceil \frac{(w - F + 2 \times P)}{s} \right\rceil + 1$$

(1)

Where:
W : Width/height of the input
F : Size of the kernel (filter)
P : Padding applied around the input
S : Stride length

The Backbone extracts deep spatial features such as facial symmetry, eye spacing, philtrum width and face shape, for generating hierarchical feature maps from the input image. These feature maps range from low-level characteristics like textures and edges at P1 to high-level semantic features such as overall facial structure and symmetry at P4-P5.

Neck

The neck module is used for feature aggregation across multiple scales where different scales of features P3–P5 are refined and fused. YOLOv8 employs a PAN (Path Aggregation Network) or FPN (Feature Pyramid Network), which helps to detect objects of various sizes and enhance feature fusion between the backbone and head. C2f (Cross-Stage Partial Fusion) uses residual-like connections and partial convolutions and reduces redundancy while enhancing the quality of the feature maps, no unique formula, but spatial dimensions typically preserved or slightly altered depending on kernel and stride.

- **Upsample** doubles the feature map dimensions:

$$\text{Output size} = \text{Inputsize} \times 2$$

(2)

- **Concat** merges feature maps along the channel axis to increase the spatial resolution:

$$\text{Output Channel} = C1 + C2$$

(3)

- **Conv** uses the same standard convolution formula for further processing

This feature aggregation allows for effective fusion of detailed and abstract features, preserving both spatial resolution and high-level meaning.

Head

The head of YOLOv8 is decoupled, meaning it independently predicts object scores, class probabilities and bounding box coordinates. Unlike older YOLO versions, YOLOv8 introduces anchor-free detection, which simplifies training and improves

Vasundhara *et al.*: Early Detection of Autism Spectrum Disorder in Children Using Facial Images and AI-Based Deep Learning Models

jOMS

flexibility for detecting irregular or uncommon patterns beneficial for medical imaging tasks such as ASD detection. For each scale.

The Detect layer estimates total predictions per scale as follows:

$$Total\ Prediction = S^2 \times A \times (B+C)$$

$$(4)$$

Where:
S   : Spatial size (e.g., 20 for P5 → 20×20)
A   : Number of anchors (YOLOv8 is anchor-free, so usually 1)
B   : Bounding box coordinates (typically 4: x, y, w, h)
C   : Number of classes (in your case, 2: Autistic, Non-Autistic)

The model outputs a set of bounding boxes around detected regions along with confidence scores and class labels. The anchor-free mechanism also improves generalization on varied datasets without the need for extensive tuning of anchor box parameters.

- **Bounding Box** prediction (x,y,w,h) is generally normalized between 0–1
- **Confidence Score** is the sigmoid of objectness:

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

$$(5)$$

- **Class Probability** also uses sigmoid (for binary/multi-label) or softmax (multi-class) = $\sigma(c)$ or Softmax (c_i)

This offers real-time autism screening through facial feature detection, enabling model classification alongside face detection for precise diagnosis.

RESULTS AND DISCUSSION

The YOLOv8n and YOLOv9c models were trained with Ultralytics YOLO framework. Training was conducted for 50 epochs at a 640×640 image size. Model performance was assessed using evaluation metrics such as mAP (mean average precision), accuracy, precision, recall and F1-score. Stable mAP and classification metrics imply the model is very reliable. Training and evaluation were carried out multiple times and results were highly consistent across all the runs, indicating that stable behaviour of the model.

Performance Metrics

The model's performance is assessed by using metrics such as precision, recall and F1-score. Once the model is implemented, these metrics are determined with the following mathematical equations based on the true positives (TP), false positives (FP) and false negatives (FN):

$$Precision = \frac{TP}{TP + FP}$$

$$(6)$$

$$Recall = \frac{TP}{TP + FN}$$

$$(7)$$

Mean Average Precision (mAP)

- **mAP@0.5**: Represents the average precision when the Intersection over Union (IoU) is set at a fixed threshold of 0.5
- **mAP@0.5:0.95**: Calculate the average precision over multiple IoU thresholds, ranging from 0.5 to 0.95 in increments of 0.05.

This measure evaluates both localization and classification accuracy over varying thresholds:

$$F1 - Score = 2 \times \frac{(Precision \times Recall)}{(Precision \pm Recall)}$$

$$(8)$$

Model Evaluation Process
Both YOLOv8n and YOLOv9c models were trained and evaluated using the same annotated dataset, consisting of child facial images labelled as autistic and non-autistic. The evaluation focused on the following performance aspects:

- **Accuracy of Predictions:** Correct classification of facial images as Autistic or Non-Autistic
- **Confidence Scores:** The certainty associated with each prediction, reflecting model confidence
- **Detection of Relevant Facial Features:** The ability to extract patterns or indicators linked to ASD
- **Real-Time Inference Capability:** How efficiently the model produces results for practical deployment scenarios

The evaluation outcomes for both models are illustrated in Table 3, Figures 2, 3 and 4.

The normalized confusion matrices for YOLOv8n(a) and YOLOv9c(b) are shown in Figure 2. It can be observed that YOLOv8n achieves higher true positive rates for both the Non- Autistic and Autistic classes, with fewer misclassifications as Background. In contrast, YOLOv9c shows slightly more confusion, particularly between the Autistic and Background classes, indicating less clear classification. Based on these findings, YOLOv8n is superior at differentiating facial characteristics associated with ASD.

Precision vs Recall curves for YOLOv8n and YOLOv9c are in Figure 3. YOLOv8n achieves more Precision and Recall for both classes, along with higher mAP@0.5, which

Table 3: Performance Comparison Between YOLOv8n vs YOLOv9c Models

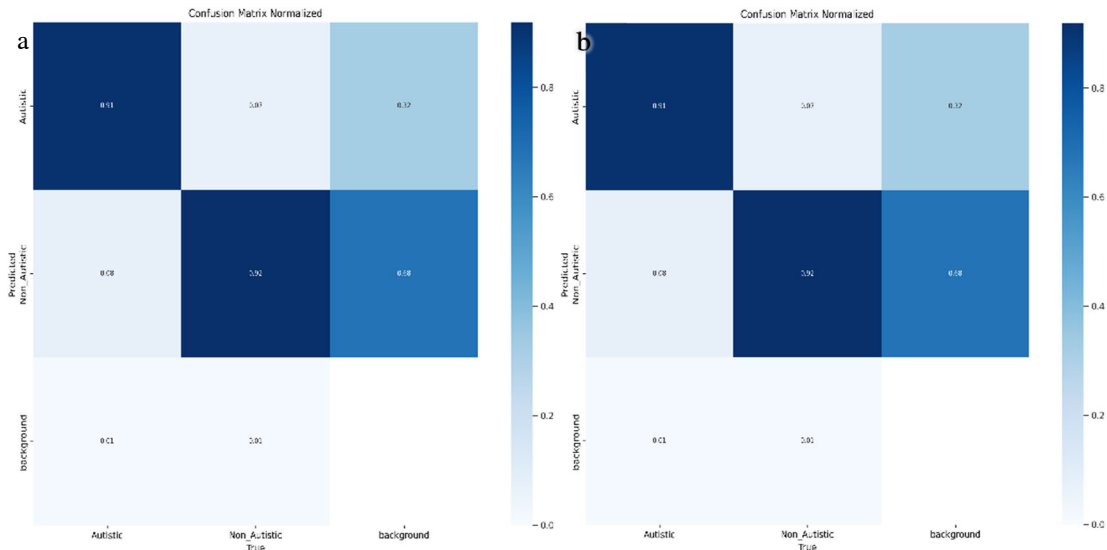| Metric	YOLOv8	YOLOv9
Precision	91.95%	91.29%
Recall	92.77%	90.29%
F1-Score	92.36%	90.79%
mAP@0.5	94.76%	93.32%
mAP@0.5:0.95	94.73%	93.21%

Figure 2a,b: Normalized Confusion Matrix (a) YOLOv8n, (b) YOLOv9c
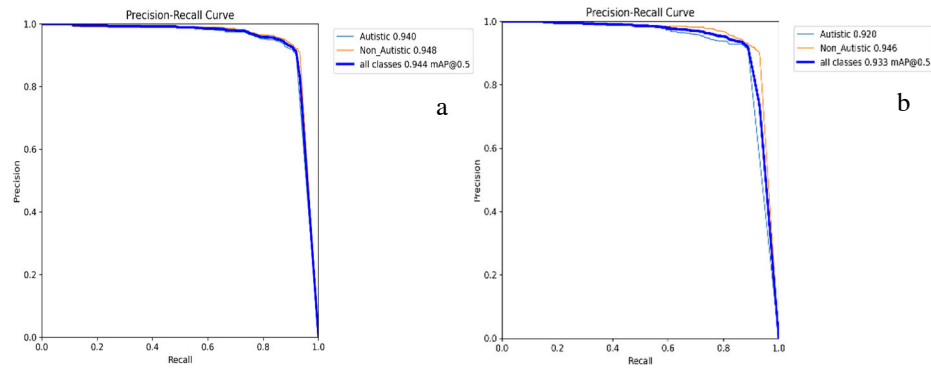

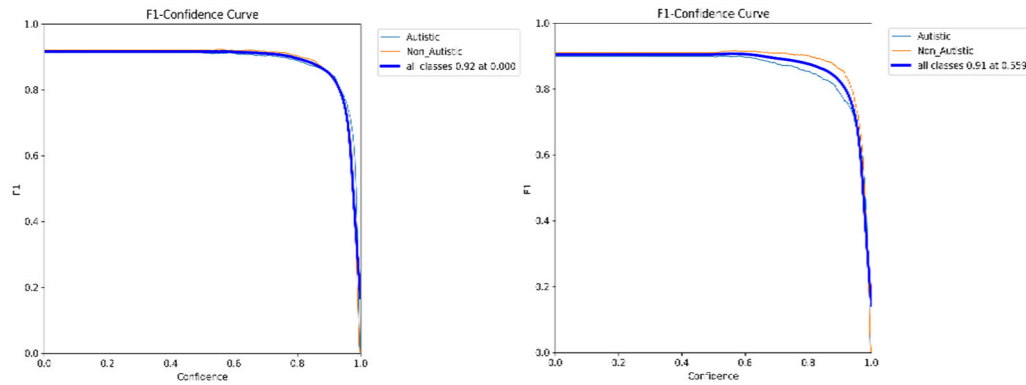Figure 3a,b: Precision vs Recall (a) YOLOv8n, (b) YOLOv9c

Figure 4a,b: F1-Confidence Curve (a) YOLOv8n, (b) YOLOv9c

Table 4: Comparative Analysis of Existing Methods and the Proposed YOLO-based ASD Detection Model

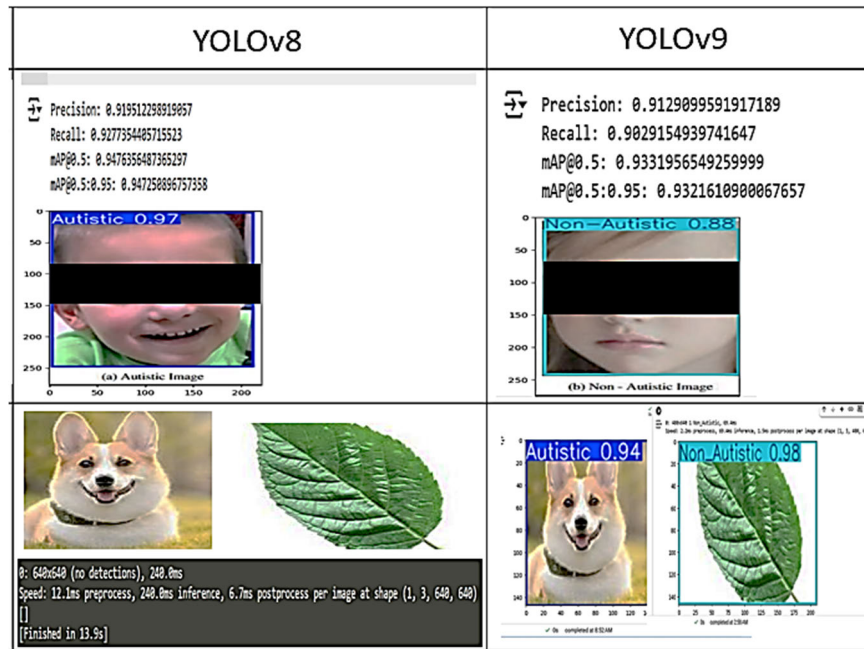| Ref No.	Method Used	Dataset Used	Limitation	Accuracy/ mAP
[1]	Attention-based YOLOv8 (AutYOLO-ATT)	Custom facial expression dataset (6 emotions) of autistic and typical children	Limited dataset diversity, challenges in generalizing to real-world settings	97.2%
[2]	Enhanced YOLOv7-tiny	Self-annotated autism face dataset (Public)	Model complexity, lower generalization to varied expressions	79.56%
Proposed	YOLOv8n / YOLOv9c	Roboflow (14,352 facial images with bounding boxes)	-	YOLOv8n: 94.76%, YOLOv9c: 93.32%

Figure 5: Visual Comparison of YOLOv8n and YOLOv9c Outputs on Test Images

means it can differentiate better between autistic and non-autistic children. These curves confirm that YOLOv8n is more reliable in identifying features associated with autism, while generating fewer false positives.

Figure 4 presents the F1-Confidence curves for both YOLOv8n and YOLOv9c models. It shows how well each balances precision and recall at different confidence levels. The YOLOv8n keeps a higher and more stable F1-score (~0.92) throughout the confidence thresholds, with a sharp drop only after 0.85, since it was tested on consistent performance. YOLOv9c achieves a slightly lower F1-score value (~0.91) and drops after 0.75; its predictions are more sensitive in confidence than earlier versions. Results confirm that YOLOv8n is robust and reliable, so it can be used for real-time ASD screening.

In this work, we analysed the YOLO-based object detection models (YOLOv8n, YOLOv9c) for detecting autism in children using facial images. The models use annotated facial images, which enhance the real-time processing performance. Although attention-based methods like Hosney [1] achieved high accuracy, they used more complex architectures and emotion-specific datasets, which limit their applicability in practical scenarios and deployment on edge devices." Another study [2] uses the YOLOv7-tiny model, which produces competitive results but has slightly lower accuracy and the computational load is higher in real-time processing. This work aims to achieve a trade-off between detection accuracy and deployment efficiency.

Comparison of YOLOv8n and YOLOv9c is presented in Figure 5, as:

- Both models accurately detected and classified the facial images of children, demonstrating reliable performance on relevant data

- YOLOv8 effectively identified non-human inputs such as a dog and a leaf, showcasing superior robustness and filtering capability
- In contrast, YOLOv9 misclassified these irrelevant objects, predicting a dog as "Autistic" and a leaf as "Nonautistic," highlighting its vulnerability to false detections on irrelevant inputs

CONCLUSIONS

The proposed YOLO models perform significantly better than the enhanced YOLOv7-tiny model across all major evaluation metrics. The effectiveness of autism screening is significantly influenced by the low number of false positives and false negatives provided by the model. YOLOv8n and YOLOv9c models are implemented for facial image-based autism spectrum disorder and the accuracy measures like precision, recall and F1 score are calculated. Though both models are promising, YOLOv8n outperforms YOLOv9c in overall ASD detection capabilities. This is attributed to its filtering mechanism, which suppresses irrelevant ASD feature detections and highlights pertinent facial features. On the contrary, YOLOv9c fails to filter correctly, which leads to over-detection and misclassification. With better handling of anchors, optimized architecture of detection heads in YOLOv8n enables reliable and precise classification of faces, a requirement in medical imaging where small details matter.

The research provided the effectiveness of deep learning models, particularly those based on YOLO architecture, in the real-time diagnosis of autism spectrum disorder. Making a wrong prediction could lead to major consequences; a false negative may delay early intervention, while a false positive can cause unnecessary stress for the parents. More

Vasundhara *et al.*: Early Detection of Autism Spectrum Disorder in Children Using Facial Images and AI-Based Deep Learning Models

jpms

importantly, YOLOv8 gives the lowest number of false positives and false negatives, which directly impacts the reliability of autism screening. However, identifying the autism spectrum disorder is a complex process that requires multiple diagnostic methods. A single test or approach is insufficient to diagnose autism properly, which is a complicated illness. A thorough assessment that involves behavioural observations, gathering child development history from parents through questionnaires, eye-tracking, facial features analysis and other clinical evaluations is necessary.

Hence, the proposed model works well as a decision-support tool, assisting healthcare professionals in speeding up the initial screening, especially in under-resourced settings.

Ethical Statement

The dataset utilized in this research is publicly available. It comprises only de-identified facial images with no personally identifiable information (PII). According to the Roboflow dataset sharing policies, it is understood that before making this dataset public, the curators obtained adequate consent and ethical clearance. Given the child's facial data aspect, there is ethical sensitivity. There was no personal or clinical collection for this research.

REFERENCES

[1] Hosney, R.*et al.* "AutYOLO-ATT: An attention-based YOLOv8 algorithm for early autism diagnosis through facial expression recognition." *Neural Computing and Applications*, vol. 36, 2024, pp. 17199–17219. https://doi.org/10.1007/s00521-024-09966-7.

[2] Kumar, A.*et al.* "Ambiguous facial expression detection for autism screening using enhanced YOLOv7-Tiny model." *Scientific Reports*, vol. 14, 2024, article 28501. https://doi.org/10.1038/s41598-024-77549-6.

[3] Maenner, M.J.*et al.* "Prevalence and characteristics of autism spectrum disorder among children aged 8 years — Autism and Developmental Disabilities Monitoring Network, 11 sites, United States, 2020." *MMWR Surveillance Summaries*, vol. 72, no. 2, 2023, pp. 1–14. https://doi.org/10.15585/mmwr.ss7202a1.

[4] Centers for Disease Control and Prevention. Data and Statistics on Autism Spectrum Disorder. CDC, 27 May 2025, https://www.cdc.gov/autism/data-research/index.html

[5] Shahamiri, S.R. and F. Thabtah. "Autism AI: A new autism screening system based on artificial intelligence." *Cognitive Computation*, 2020. https://doi.org/10.1007/s12559-020-09743-3.

[6] Thabtah, F.*et al.* "A machine learning autism classification based on logistic regression analysis." *Health Information Science and Systems*, vol. 7, no. 1, 2019, article 12. https://doi.org/10.1007/s13755-019-0073-5.

[7] Priyadarshini, I. "Autism screening in toddlers and adults using deep learning and fair AI techniques." *Future Internet*, vol. 15, no. 9, 2023, article 292. https://doi.org/10.3390/fi15090292.

[8] Shinde, A. and D. Patil. "A multi-classifier-based recommender system for early autism spectrum disorder detection using machine learning." *Healthcare Analytics*, 2023, article 100211. https://doi.org/10.1016/j.health.2023.100211.

[9] Talukdar, J.*et al.* "A comparative assessment of most widely used machine learning classifiers for analysing and classifying autism spectrum disorder in toddlers and adolescents." *Healthcare Analytics*, vol. 3, 2023. https://doi.org/10.1016/j.health.2023.100178.

[10] Farooq, M.S.*et al.* "Detection of autism spectrum disorder (ASD) in children and adults using machine learning." *Scientific Reports*, vol. 13, no. 1, 2023, article 9605. https://doi.org/10.1038/s41598-023-35910-1.

[11] Khosla, Y.*et al.* "Detection of autistic individuals using facial images and deep learning." *Proceedings of CSITSS 2022*, 2022. https://doi.org/10.1109/CSITSS54238.2021.9683205.

[12] Reddy, P. and A. J. "Diagnosis of autism in children using deep learning techniques by analyzing facial features." *Engineering Proceedings*, vol. 59, 2023, article 198. https://doi.org/10.3390/engproc2023059198.

[13] Awaji, B.*et al.* "Hybrid techniques of facial feature image analysis for early detection of autism spectrum disorder based on combined CNN features." *Diagnostics*, vol. 13, no. 18, 2023, article 2948. https://doi.org/10.3390/diagnostics13182948.

[14] Thamilselvan, R.*et al.* "Autism spectrum disorder diagnosis using deep learning techniques." *Proceedings of ICC-ROBINS 2024*, 2024, pp. 402–407. https://doi.org/10.1109/ICC-ROBINS60238.2024.10533978.

[15] Wei, Q.*et al.* "Machine learning based on eye-tracking data to identify autism spectrum disorder: A systematic review and meta-analysis." *Journal of Biomedical Informatics*, 2023. https://doi.org/10.1016/j.jbi.2022.104254.

[16] Khodatars, M. *et al.* "Deep learning for neuroimaging-based diagnosis and rehabilitation of autism spectrum disorder: A review." *Computers in Biology and Medicine*, vol. 139, 2021, article 104949. https://doi.org/10.1016/j.compbiomed.2021.104949.

[17] Alkahtani, H.*et al.* "Deep learning algorithms to identify autism spectrum disorder in children-based facial landmarks." *Applied Sciences*, vol. 13, no. 8, 2023, article 4855. https://doi.org/10.3390/app13084855.

[18] Chakradhar, K.*et al.* "Detection of autism spectrum disorder using deep learning models." *Proceedings of ICICI 2024*, 2024, pp. 257–262. https://doi.org/10.1109/ICICI62254.2024.00050.

[19] Diwan, T.*et al.* "Object detection using YOLO: Challenges, architectural successors, datasets and applications." *Multimedia Tools and Applications*, vol. 82, 2023, pp. 9243–9275. https://doi.org/10.1007/s11042-022-13644-y.

[20] Hussain, M. "YOLOv1 to v8: Unveiling each variant–A comprehensive review of YOLO." *IEEE Access*, vol. 12, 2024, pp. 42816–42833. https://doi.org/10.1109/ACCESS.2024.3378568.

[21] Uddin, Md.Z.*et al.* "Deep learning with image-based autism spectrum disorder analysis: A systematic review." *Engineering Applications of Artificial Intelligence*, vol. 127, part A, 2024. https://doi.org/10.1016/j.engappai.2023.107185.

[22] Taneera, S. and R. Alhajj. "Diagnosis of autism spectrum disorder: A systematic review of clinical and artificial intelligence methods." *Network Modeling Analysis in Health Informatics and Bioinformatics*, 2025. https://doi.org/10.1007/s13721-024-00499-6.

[23] "Autism Dataset." Roboflow Universe, Roboflow, https://universe.roboflow.com/detection-ifxlr/autism-7ywpu/dataset/3/images?split=train

[24] Karna, N., M. A. P. Putra, S. M. Rachmawati, M. Abisado, and G. A. Sampedro. "Towards Accurate Fused Deposition Modeling 3D Printer Fault Detection Using Improved YOLOv8 with Hyperparameter Optimization." IEEE Access, vol. 11, 2023, pp. 69395–69407, https://doi.org/10.1109/ACCESS.2023.3293056.