International Journal of Multidisciplinary Engineering in Current Research - IJMEC
Volume 10, Issue 6, June-2025, http://ijmec.com/, ISSN: 2456-4265

# Autism Spectrum Disorder Detection Using Yolov9

[1]k. Sathwik, [2]Nachiketh R, [3]Rohith K, [4]Gokul M & [5]MShona
[1,2,3,4] B-Tech, Department of Computer Science and Engineering, M S Ramaiah University of Applied Sciences. Bengaluru, Karnataka, India
[5]Assistant Professor, Department of Computer Science and Engineering, M S Ramaiah University of Applied Sciences. Bengaluru, Karnataka, India Email: mshona.cs.et@msruas.ac.in

***ABSTRACT:***

***Autism spectrum disorder (ASD) is a neurodevelopmental illness that is marked by difficulties in social interaction, communication, and repetitive activities. It is a complex condition. For the purpose of image-based detection of autism spectrum disorder (ASD) through facial analysis, traditional deep learning techniques, particularly Convolutional Neural Networks (CNNs), have experienced widespread application. The performance of CNN-based classifiers is satisfactory; however, they frequently have a high computational cost, delayed inference, and restricted accuracy when it comes to identifying subtle facial traits that are related to autism spectrum disorder (ASD). This study presents an improved detection strategy that makes use of YOLOv9, a cutting-edge real-time object recognition model that is already well-known for its better speed and accuracy. The goal of this work is to overcome these limitations. The YOLOv9 algorithm is able to effectively recognize patterns in facial expressions that are symptomatic of autism with a much-reduced latency, which makes it acceptable for applications that include early screening. The performance of the proposed model is superior to that of standard CNN approaches because it makes use of the sophisticated feature extraction capabilities and attention mechanisms implemented in YOLOv9. The results of the experiments show that the detection precision, real-time performance, and robustness against variations in image quality and lighting have all been significantly enhanced. The findings of this study represent a significant step toward the development of autism detection techniques that are scalable, quick, and reliable in clinical and non-clinical settings.***

***Keywords: - ASD, YOLOv9, CNN, ML & DL***

## 1-INTRODUCTION

A complicated neurological condition that impacts social interaction, speech, and behavior, autism spectrum disorder (ASD) is a condition that affects individuals with autism. It is possible for people who have autism spectrum disorder to display behaviors that are repetitive, have difficulties interpreting emotions, and have delayed speech or non-verbal signs. Early diagnosis is sometimes delayed due to subjective evaluations and a lack of professionals, despite the fact that it is extremely important. In addition to being time-consuming and costly, traditional diagnostic approaches are also prone to human error (human error). Automatically automating the detection process through the use of facial analysis, behavior tracking, and pattern recognition is a promising option that can be offered by AI-based systems. From facial photos or videos, deep learning models, particularly those used in computer vision, are able to recognize subtle signs of autism spectrum disorder (ASD). These methods improve the speed of diagnosis while also improving its accuracy and accessibility. Utilizing AI in early autism spectrum disorder screening can result in more rapid interventions and improved outcomes over the long term.

Autism spectrum disorder (ASD) is a condition that is currently affecting an increasing number of people all over the world. This condition is characterized by difficulties in social communication and restricted behavioral patterns. When it comes to enhancing the quality of life through timely intervention, early detection is an extremely important factor. In order to examine face characteristics that are suggestive of autism spectrum disorder (ASD), existing methods mostly rely on convolutional neural networks (CNNs). The CNN-based models, on the other hand, frequently have a limited real-time performance, a reduced sensitivity to tiny facial alterations, and a large computational cost. These limitations make it difficult to scale up automated autism spectrum disorder screening systems and reduce their overall usefulness, particularly in settings with limited resources. For this reason, there is a requirement for a detection method that is more reliable, quicker, and more accurate in order to enhance the early diagnosis of autism through the use of image analysis.

This research is being conducted with the primary purpose of developing a system that is capable of diagnosing Autism Spectrum Disorder (ASD) through the use of facial image analysis in a manner that is both accurate and efficient. Through the utilization of the more advanced capabilities of YOLOv9, it intends to overcome the constraints that are associated with typical CNN models. Automatic extraction of facial traits associated with autism is the primary focus of this study, which does not involve any human interaction. Through the provision of a screening instrument that is both quick and non-invasive, it aims to cut down on diagnostic delays. Additionally, one of the objectives is to guarantee that the model functions effectively across a wide range of facial datasets and lighting circumstances. For the purposes of clinical and assistive applications, the system is designed to be both cost-effective and scalable respectively. The end goal of the project is to provide support for early intervention and to improve the quality of life for people who have autism spectrum disorder (ASD).

## 2. LITERATURE REVIEW

An innovative computer-aided grading system for infants and toddlers (between the ages of one and four years) was proposed by Haweel et al. [1] on the basis of an investigation into the activity of the brain in response to a speech experiment. Subah et al. [2] proposed an ASD detection model by making use of the functional connectivity properties of resting-state functional magnetic resonance imaging (fMRI) data. Two of the most used brain atlases, namely Craddock 200 (CC200) and automated anatomical labeling (AAL), as well as two atlases that are not commonly used, namely bootstrap analysis of stable clusters (BASC) and Power, are utilized in the model that they have proposed.

In order to construct a sign language recognition model, Masood et al. [3] utilized the deep convolutional neural network (CNN) inception model to train the model on spatial data. Additionally, they utilized the recurrent neural network (RNN) to train the model on temporal features that are present in video sequences. Eslami et al. [4] presented a model that was based on the deep learning and machine learning technique, using CNN and SVN classifier respectively, for the purpose of identifying attention-deficit/hyperactivity disorder (ADHD) and autism spectrum disorder (ASD). The model is utilized to identify the characteristics of autism by utilizing functional and structural state magnetic resonance imaging (MRI) of the subject. It was demonstrated that the deep learning (DL) model performs in a manner that is superior to the machine learning (ML) model that was proposed.

Using a variational autoencoder, Choi [5] was able to translate high-dimensional and multidimensional data into two-dimensional characteristics. phenomenon of stereotyped motor movements (SMM) in autistic individuals. These movements have an impact on learning and social abilities, and they include body swaying and elaborate hand motions. CNN is responsible for processing the multi-sensor accelerometer measurements that are received from SMM in order to extract a variety of properties. CNN was used as the foundation for the entirely automated brain tumor segmentation approach that was proposed by Havaei et al. [7].

Devika et al. [8] proposed an autism spectrum disorder (ASD) detection model that is based on structural magnetic resonance imaging (MRI). MRI data from only healthy patients was employed by the authors for this aim, and a generative adversarial network (GAN) was used to train the system to recognize spatio-temporal patterns in skeletal brain connections. Two alternative baselines, a more complex self-attention GAN and a more fundamental UNet, were compared to the model in order to evaluate its performance. In addition, when compared to cross-sectional data, longitudinal data produced results that were 17–28% more accurate (one scan for each patient).

Khadem-Reza et al. [9] proposed a unique method for distinguishing individuals with autism spectrum disorder (ASD) from controls by utilizing structural magnetic resonance imaging (MRI) data. By simultaneously employing the volume and surface features of the structural photographs, this approach is able to achieve further improvements in its accuracy. In their analysis, the authors utilized a

Additionally, he demonstrated a functional connectivity pattern that is associated with autistic children. Researchers Rad et al. [6] investigated the variety of machine learning and deep learning methodologies, which resulted in diagnosis accuracy rates of 86.29 percent, 71.1 percent, 86.5 percent, and 88.46 percent, respectively. In terms of diagnostic accuracy, the artificial neural network (ANN) provided the highest conceivable outcome.

## 3. IMPLEMENTATIONS OF AUTISM DETECTION MODEL

### 3.1 System Model

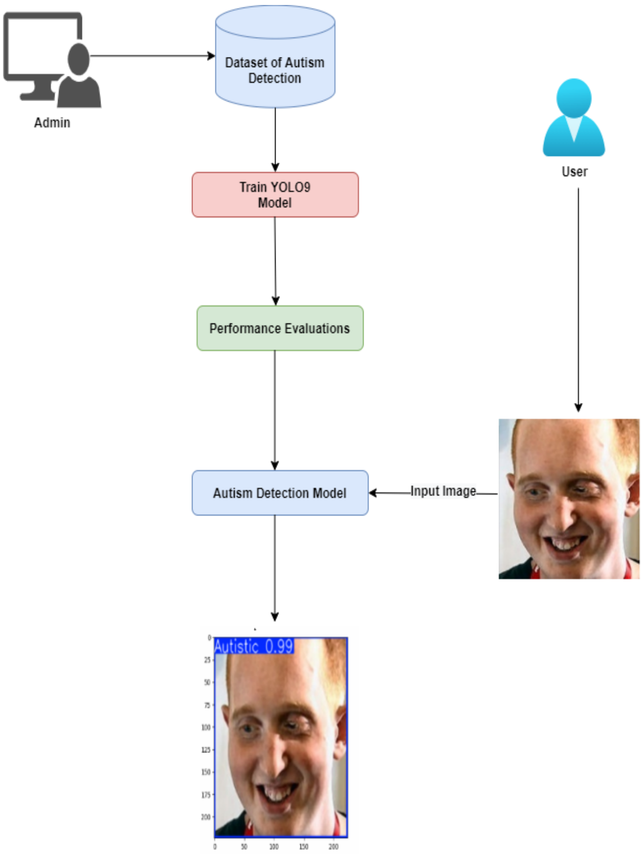


Figure.3.1 Autism Detection Model

As shown in figure.3.1, the autism detection model is constructed with a number of different modules, including a collecting dataset, a training model, model evaluations, and an autism detection model.

### 3.2 System Modules

#### 3.2.1 Collection of Autism Detection Dataset:

In this module, we are acquiring the Autism Detection image dataset from Roboflow Universe a platform commonly used for sharing computer vision datasets. The dataset is designed for detecting autistic and non-autistic, using computer vision techniques. This image dataset contains a total of 5931 images and images in the dataset are annotated with bounding boxes marking the location of autistic and non-autistic.

#### 3.2.2 Model (YOLOv9) Training:

The initial stage in the process of training a YOLOv9 model for weapons detection is to load the pre-trained YOLOv9 model; this serves as the starting point for the training process. After it has been loaded, the model can then be trained using the dataset that you have chosen. During the training phase, the configuration of the dataset is specified by means of a configuration file. This file specifies the paths of the training, validation, and test datasets, as well as the names of the classes corresponding to each dataset. It is during this step that important parameters are configured. These parameters include the number of epochs, the batch size, the picture size, and the number of workers who load data. An example of this would be the training process being set to run for fifty epochs, with a batch size of sixteen and an image resolution of 640 by 640 pixels. The model is able to adapt to your specific detection needs by fine-tuning the pre-trained weights on your annotated Autism detection dataset. This allows the model to achieve improved accuracy in detecting autistic and non-autistic individuals in a variety of scenarios.

#### 3.2.3 Model Evaluations:

The performance of the YOLOv9 model on the test dataset is evaluated in this module by the Model evaluation section. Metrics such as precision, recall, and the F1 score are utilized in this evaluation. The accuracy of detection is measured by these metrics, which also provide insights across a variety of threshold positions. The process of visualizing predictions on test images brings to light both strengths and shortcomings, such as the ability to handle images that are small or obscured. Through the identification of false positives and false negatives, error analysis helps to direct improvements. This procedure guarantees the dependability of the model and makes it ready for implementation in situations that are based in the actual world.

#### 3.2.4 Autism Detection Model:

At the beginning of the process, this module loads a collection of input photographs from a certain folder and then gives the model the opportunity to make predictions about the objects (such as autistic and non-autistic) that are contained inside these images. In order to provide insights into the detection accuracy of the model, the predictions are shown using bounding boxes and class labels.

## 4. EXPERIMENTAL RESULTS

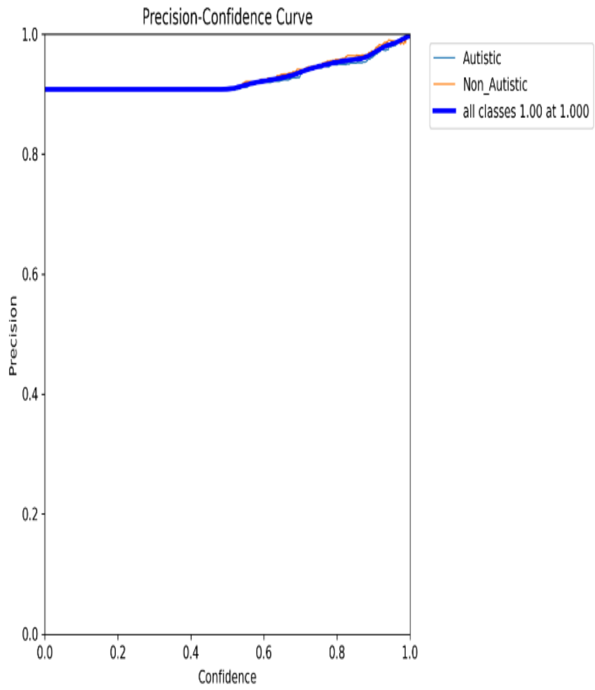


Figure.2 Precision curve

From figure.2 across the entire confidence range, the curve demonstrates a nearly flat and high level of precision for both the autistic and non-autistic classes.

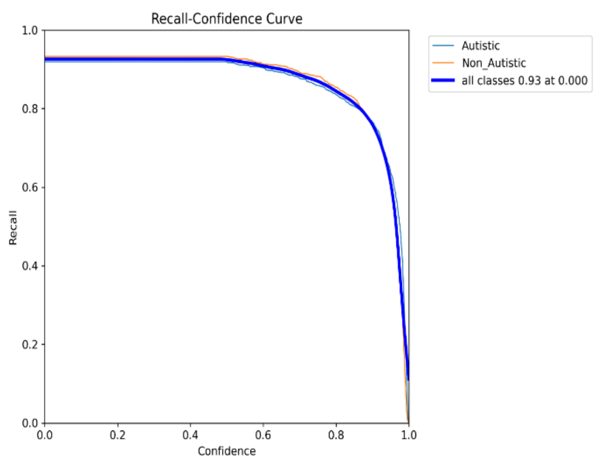


Figure.3 Recall Curve

From figure.3 as the confidence level grows, the recall for both classes, namely Autistic and Non-Autistic, begins at a high level (about 0.93), and then steadily decreases.

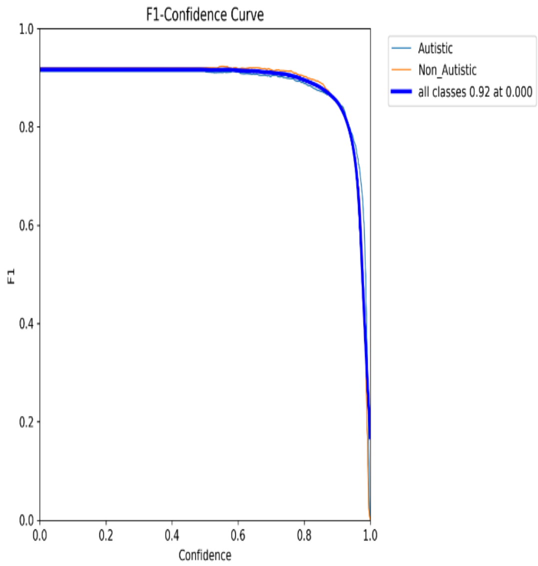


Figure.4 F1-score Curve

From figure.4, at confidence levels that are low, the F1-score for both the Autistic and Non-Autistic classes is extremely high, measuring approximately 0.92. As confidence improves, the curve experiences a modest decrease, particularly beyond a value of approximately 0.85, where fewer predictions are produced but they are more certain.

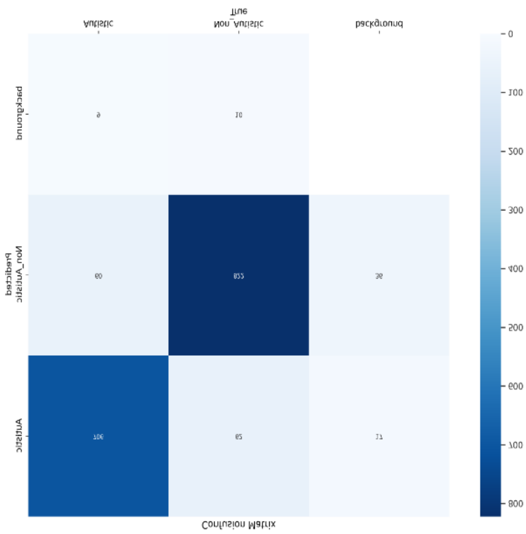

Figure.5 Confusion Matrix

From figure.5, with F1-scores that are greater than 90% for both the Autistic and Non-Autistic classes, the confusion matrix demonstrates that the model is capable of achieving a high level of classification accuracy. The robustness of the YOLOv9-based detection system is highlighted by the fact that it has a minimal amount of misclassification and background interference.

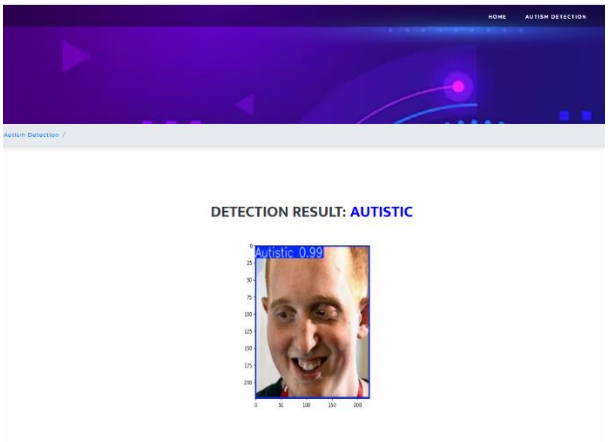


Figure.6 Detection of AUTISTIC

From figure.6, the YOLOv9 trained model will return the detected result as AUTISTIC for the given input image.

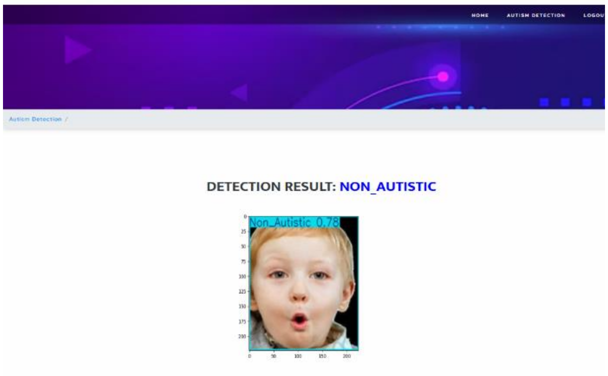


Figure.7 Detection of NON_AUTISTIC

From figure.7, the YOLOv9 trained model will return the detected result as NON_AUTISTIC for the given input image.

## 5. CONCLUSION

The results of this study reveal that YOLOv9 is effective in identifying individuals with autism spectrum disorder through the use of facial image analysis. When compared to conventional CNN-based classifiers, the YOLOv9-based model that has been developed offers higher detection accuracy, shorter inference times, and improved robustness across a wide range of visual circumstances. A strong candidate for scalable, non-invasive autism spectrum disorder screening tools, it is characterized by its lightweight architecture and its ability to do processing in real time. The focus of future research will be on expanding this approach to video-based behavioral analysis and integrating multi-modal data in order to further improve diagnostic precision.

## 6.REFERENCES

[1] R. Haweel et al., "A Novel Grading System for Autism Severity Level Using Task-Based Functional MRI: A Response to Speech Study," IEEE Access, vol. 9, pp. 100570–100582, 2021, doi: 10.1109/ACCESS.2021.3097606.

[2] F. Z. Subah, K. Deb, P. K. Dhar, and T. Koshiba, "A deep learning approach to predict autism spectrum disorder using multisite resting-state fmri," Applied Sciences (Switzerland), vol. 11, no. 8, 2021, doi: 10.3390/app11083636.

[3] S. Masood, A. Srivastava, H. C. Thuwal, and M. Ahmad, "Real-time sign language gesture (word) recognition from video sequences using CNN and RNN," Advances in Intelligent Systems and Computing, vol. 695, pp. 623–632, 2018, doi: 10.1007/978-981-107566-7_63.

[4] T. Eslami, F. Almuqhim, J. S. Raiker, and F. Saeed, "Machine learning methods for diagnosing autism spectrum disorder and attention-deficit/hyperactivity disorder using functional and structural MRI: A survey," Frontiers in Neuroinformatics, vol. 14, 2021, doi: 10.3389/fninf.2020.575999.

[5] H. Choi, "Functional connectivity patterns of autism spectrum disorder identified by deep feature

learning," 2017, [Online]. Available: http://arxiv.org/abs/1707.07932.

[6] N. M. Rad, S. M. Kia, C. Zarbo, G. Jurman, P. Venuti, and C. Furlanello, "Stereotypical motor movement detection in dynamic feature space," IEEE International Conference on Data Mining Workshops, ICDMW, vol. 0, pp. 487–494, 2016, doi: 10.1109/ICDMW.2016.0076.

[7] M. Havaei et al., "Brain tumor segmentation with Deep Neural Networks," Medical Image Analysis, vol. 35, pp. 18–31, 2017, doi: 10.1016/j.media.2016.05.004.

[8] K. Devika, D. Mahapatra, R. Subramanian, and V. R. M. Oruganti, "Outlier-based autism detection using longitudinal structural MRI," IEEE Access, vol. 10, pp. 27794–27808, 2022, doi: 10.1109/ACCESS.2022.3157613.

[9] Z. K. Khadem-Reza and H. Zare, "Automatic detection of autism spectrum disorder (ASD) in children using structural magnetic resonance imaging with machine vision system," Middle East Current Psychiatry, vol. 29, no. 1, 2022, doi: 10.1186/s43045-02200220-1.