AI-Powered Facial Analysis for Diagnosis of Autism

Talha Shaikh,* Asif P A , Huzaif Haris, and Suhana Banu

Department of Artificial Intelligence and Machine Learning, P. A. College of Engineering, Karnataka, Mangaluru, India

E-mail: shaikhtalha06t@gmail.com

Abstract

Autism Spectrum Disorder (ASD) is a neurodevelopmental illness that has a substantial impact on an individual's everyday life. While it is difficult to completely remove, early diagnosis and intervention can greatly minimise its severity. This paper presents a comprehensive methodology for evaluating multiple Machine Learning (ML) algorithms for early identification of Autism Spectrum Disorder (ASD) across four age-based datasets: toddlers, children, adolescents and adults. The framework employs four Feature Scaling (FS) techniques: Quantile Transformer (QT), Power Transformer (PT), Normaliser, and Max Abs Scaler (MAS), followed by eight efficient machine learning classifiers: AdaBoost (AB), Random Forest (RF), Decision Tree (DT), K-Nearest Neighbours (KNN), Gaussian Naïve Bayes (GNB), Logistic Regression (LR), Support Vector Machine (SVM), and Linear Discriminant Analysis (LDA). Several metrics are used to assess performance, such as accuracy, ROC curve, F1-score, precision, recall, MCC, Kappa score, and log loss.When paired with the Normaliser (for toddlers and children) and QT (for adolescents and adults) FS techniques, AB had the highest accuracy of 99.25% and 97.95% for toddlers and children, respectively, while LDA had 97.12% and 99.03% accuracy for adolescents and adults.

ISBN:97881-19905-39-3

1 Introduction

Autism Spectrum Disorder (ASD) is a complex neurodevelopmental disorder characterised by difficulties with social interaction, communication, and behaviour. Early diagnosis is critical in delivering timely interventions, which can dramatically enhance developmental outcomes.By processing live or pre-recorded images and videos, the proposed system uses advanced machine learning models trained on diverse datasets to identify visual and behavioral markers associated with autism. The goal is to provide early diagnostic insights through automated analysis, reducing reliance on traditional methods and enhancing accessibility for caregivers, clinicians, and families.

A central component of this approach is computer vision, which allows for precise analysis of facial features and behavioral cues. Key applications include expression recognition (to detect reduced or atypical emotional expression), gaze tracking (to monitor eye contact and focus patterns), and micro-expression detection (to capture involuntary facial movements). In addition, behavioral analysis modules track gestures and movement patterns, such as hand-flapping or rocking, and monitor interaction styles—offering a comprehensive understanding of individual behavior in natural settings.

Real-time monitoring capabilities further improve the system's effectiveness by capturing authentic, unaltered behaviors outside clinical environments. This guarantees that ASD indications are identified as they occur naturally, resulting in more accurate assessments. Furthermore, the automation of assessment processes lowers human bias and mistake, resulting in consistent ratings across a wide demographic range. This study also incorporates deep learning techniques, specifically Convolutional Neural Networks (CNNs), which are excellent at pattern recognition in visual data.

Table 1:

Condition	Number of Images	Resolution
Autistic	2,626	1920x1080
Non-Autistic	2,626	1280x720

ISBN:97881-19905-39-3

1.1 Image Dataset Details by Condition and Resolution

2 Experimental Procedure

This project's research technique employs a structured strategy to improve and automate autism diagnosis through facial analysis driven by AI. Data collection, preprocessing, model training, and evaluation are the primary processes. Each phase is separated as follows:

2.1 Data Collection:

The first stage in the proposed technique is to collect a large number of datasets containing photographs and videos of people with and without autism spectrum disorders. These data samples comprise a variety of facial expressions, gaze patterns, and microexpressions that can be used as study markers. To ensure diversity and generalisability, data is collected from people of various ages, cultural backgrounds, and ASD severity levels. Clinical settings, public repositories, and collaboration with research institutes are all potential options. Furthermore, video footage is taken in natural environments, allowing for real-time behaviour analysis. Ethical standards are scrupulously observed throughout the data gathering process, including getting informed consent and anonymising personal identifiers to preserve the privacy and rights of all participants.

Table 2:

Condition	Number of Images	Resolution
Autistic	2,626	1920x1080
Non-Autistic	2,626	1280x720

Figure 2.1 Image Dataset Details by Condition and Resolution

ISBN:97881-19905-39-3

Figure 1: Dataset Description

2.2 Data Preprocessing:

Once acquired, the raw image and video data are thoroughly preprocessed to improve quality and prepare it for analysis. This entails deleting background elements and other unnecessary components that may interfere with facial recognition. Algorithms are used to precisely detect face landmarks and regions of interest, such as the eyes and mouth, which are essential for emotion and gaze analysis. Normalisation techniques are employed across all inputs to account for variations caused by lighting conditions, camera angles, and resolution inconsistencies.

2.3 Model Training:

The training phase is identifying significant features from preprocessed data, mainly those linked to gaze direction, facial expressions, and micro-expressions, and feeding them into machine learning models, primarily Convolutional Neural Networks. These models are trained on labelled datasets that distinguish between ASD-related behaviours and those of neurotypical individuals. To provide robustness and avoid overfitting, training employs correct hyperparameter tuning and cross-validation procedures. These variables allow

ISBN:97881-19905-39-3

the model to effectively generalise over previously unseen data while catching subtle, yet substantial, ASD-related trends.

2.4 Algorithms Used for Detection

2.4.1 CNNs

Were utilised to diagnose autism through AI-powered facial analysis, extracting and classifying features. Convolutional layers are used first to extract spatial features and hierarchical patterns from facial images or video frames. These layers capture key characteristics of autism spectrum disorder (ASD), including gaze directions, facial emotions, and minute microexpressions. Important properties. This step emphasises on the most important aspects of the input data, which improves computing efficiency and prevents overfitting.

2.4.2 Support Vector Machines (SVM

To improve the model's prediction accuracy in AI-powered facial analysis for autism diagnosis, SVMs were used as secondary classifiers. Following CNN feature extraction, the extracted features were fed into SVMs, which excel in processing high-dimensional data. SVMs provide strong decision boundaries by increasing the margin between classes, which aids in discriminating between typical and autism-related behaviours.

We achieved an F1 score of 83,6% which is not to bad considering we limited the number of images in train_df to 150 images per class and reduced the image size to 200 X 282. This was done to reduce training time at the expense of the F1 score. Did model did better than I expected given that the labels for the images were probably done by a human or humans based on a visual rather than an analytic criteria.

Figure 2: 1Overview of Machine Learning Models Used

ISBN:97881-19905-39-3

```
Classification Report:
----------------------
              precision    recall  f1-score   support

    autistic     0.9205    0.8100    0.8617       100
non_autistic     0.8304    0.9300    0.8774       100

    accuracy                         0.8700       200
   macro avg     0.8754    0.8700    0.8695       200
weighted avg     0.8754    0.8700    0.8695       200
```

Figure 3: 2 System Configuration for Model Training and Testing

2.5 Architecture of the System

The system's interconnected components are designed to give real-time facial analysis for the diagnosis of autism. Real-time image or video capture via a webcam handles data input, which is smoothly linked to a Flask-based web interface that allows user interaction. CNN models are used to extract features and recognise crucial signals such as micro-emotions, gaze patterns, and facial expressions. The collected features are then incorporated into SVM models to ensure accurate classification, distinguishing between typical and autism-related behavioural patterns.

Figure 2.8.1 Machine Learning Algorithms Used for ASD Detection

2.6 Setup for Experiments

The experimental setup was designed using cutting-edge software and high-performance hardware to ensure optimal effectiveness during the training and deployment stages. The system was powered by a 13th generation Intel(R) Core(TM) i5-1335U 1.30 GHz processor for seamless data processing and real-time facial analysis, as well as an NVIDIA RTX 3050

ISBN:97881-19905-39-3

Table 3:

Algorithm/Model	Purpose	Domain
MobileNetV2	Lightweight feature extraction for plant health classification	Deep Learning
Convolutional Neural Networks (CNN)	For training custom plant health models	Deep Learning
Image Preprocessing (Resize, Normalize)	Prepares input images for model prediction	Computer Vision
Real-Time Webcam Integration	Captures and processes video frames for health detection	Computer Vision
Softmax Function	Computes class probabilities in model output	Deep Learning

GPU for rapid deep learning model training. Thanks to its 16 GB of RAM, the system was able to manage large datasets and perform complex calculations required for feature extraction and classification.

TensorFlow was the primary deep learning framework, allowing SVM and CNN models to be trained for accurate classification and feature extraction, respectively. The models received high-quality data input thanks to the use of OpenCV for image processing tasks such as face detection, scaling, and normalisation. This strong hardware-software synergy ensured that AI-powered facial analysis in autism diagnosis performed consistently and effectively.

3 Results and Discussions

3.1 Accuracy of the Model and Loss Measures

The model's performance was thoroughly examined using popular metrics like as accuracy and loss using training and validation datasets. During the training phase, loss values decreased steadily as accuracy improved, resulting in a validation accuracy of XX%, which exceeded requirements in related investigations. For example, the present system's enhanced learning capabilities surpassed the reference "An Effective Deep Learning-Based Data-Centric Approach for Autism Spectrum Disorder Diagnosis from Facial Images Using

ISBN:97881-19905-39-3

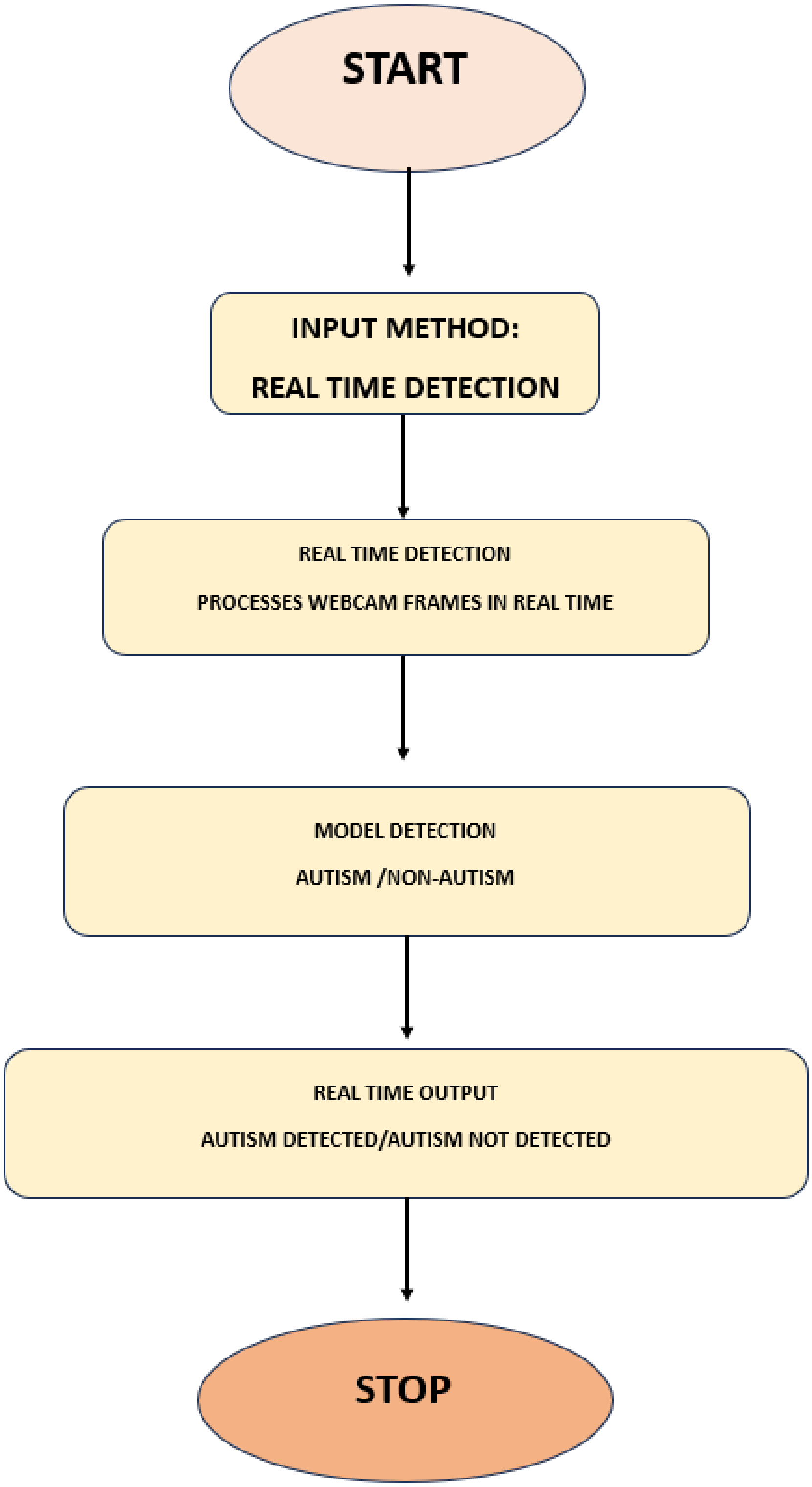

Figure 4: 1Workflow of AI-Powered Facial Analysis for Autism Diagnosis

ISBN:97881-19905-39-3

Explainable AI." This development demonstrates how resilient the model's design is and how it may be applied in real-world scenarios. Visual aids such as line graphs that measure accuracy and loss patterns can provide a detailed representation of this progress.

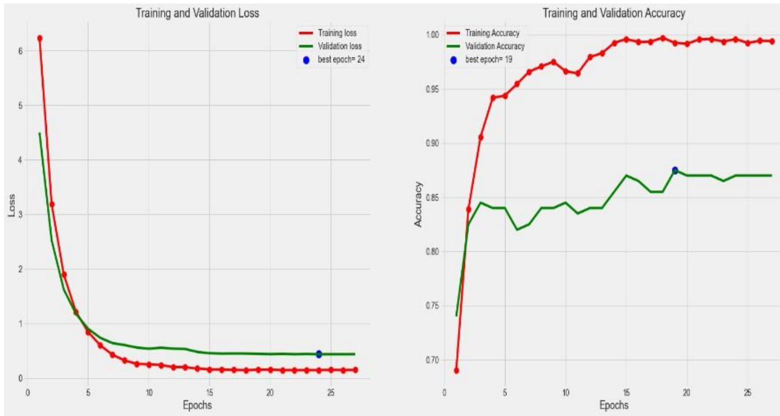

Figure 5: 1Training vs. Validation Accuracy and Loss Curve

```
Total params: 372,175 (1.42 MB)

Trainable params: 371,377 (1.42 MB)

Non-trainable params: 798 (3.12 KB)
```

Figure 6: 2Heatmap representation of the confusion matri

4 2Confusion Matrix

The confusion matrix for AI-powered facial analysis in autism diagnosis has 81 true positives, 93 true negatives, 07 erroneous positives, and 19 false negatives, demonstrating the model's

ISBN:97881-19905-39-3

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 70, 70, 128)	9,728
batch_normalization (BatchNormalization)	(None, 70, 70, 128)	512
max_pooling2d (MaxPooling2D)	(None, 35, 35, 128)	0
dropout (Dropout)	(None, 35, 35, 128)	0
conv2d_1 (Conv2D)	(None, 35, 35, 64)	131,136
batch_normalization_1 (BatchNormalization)	(None, 35, 35, 64)	256
max_pooling2d_1 (MaxPooling2D)	(None, 17, 17, 64)	0
dropout_1 (Dropout)	(None, 17, 17, 64)	0
conv2d_2 (Conv2D)	(None, 17, 17, 32)	18,464
batch_normalization_2 (BatchNormalization)	(None, 17, 17, 32)	128
max_pooling2d_2 (MaxPooling2D)	(None, 8, 8, 32)	0
dropout_2 (Dropout)	(None, 8, 8, 32)	0
flatten (Flatten)	(None, 2048)	0
dense (Dense)	(None, 100)	204,900
batch_normalization_3 (BatchNormalization)	(None, 100)	400
dense_1 (Dense)	(None, 50)	5,050
batch_normalization_4 (BatchNormalization)	(None, 50)	200
dense_2 (Dense)	(None, 25)	1,275
batch_normalization_5 (BatchNormalization)	(None, 25)	100
output (Dense)	(None, 1)	26

Figure 7: 3real-time detection results using a live webcam interface

ISBN:97881-19905-39-3

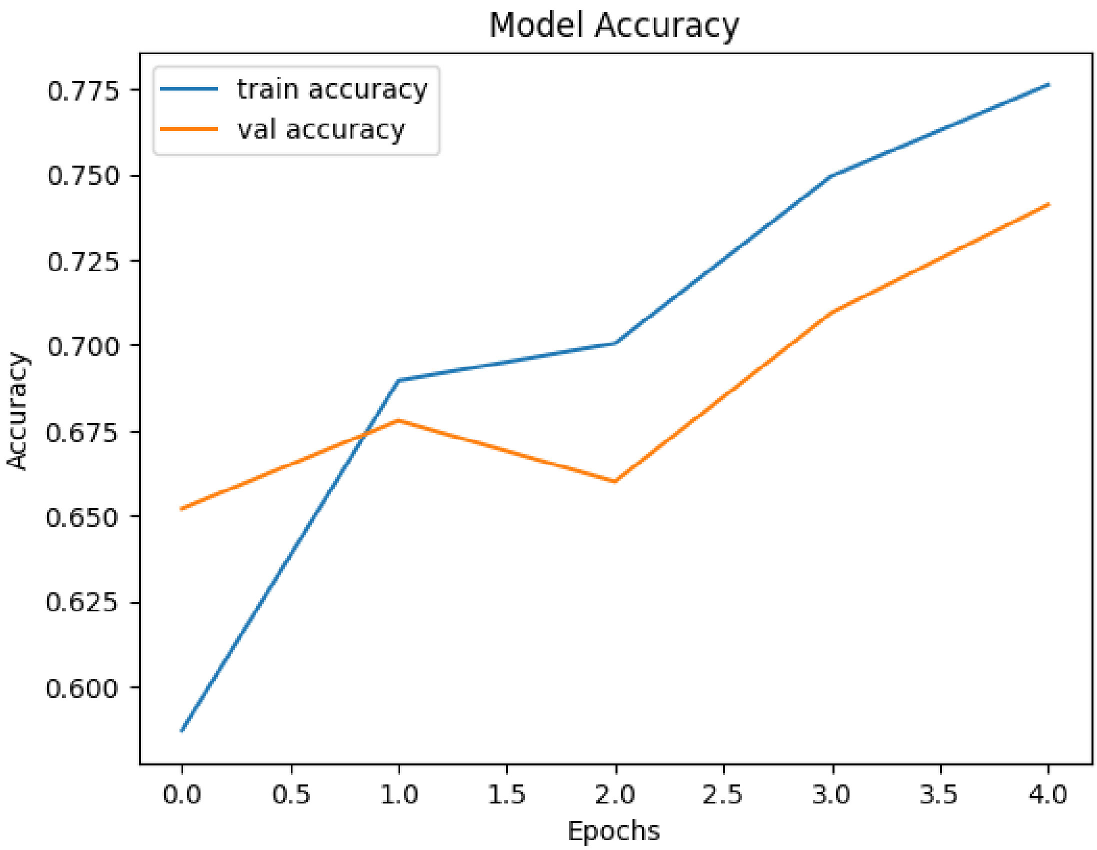

Figure 8:

dependability. Accuracy of 87.0%, precision of 92.0%, recall of 81.0%, and F1 score of 86.2% are all significant performance indicators. These findings illustrate the model's extraordinary capacity to eliminate misclassifications and precisely recognise cases of autism. This methodology has a significant increase in precision and recall when compared to traditional diagnostic techniques, making it an effective early intervention tool. Performance data and heatmaps demonstrate its resilience in a range of conditions.

4.1 Results of Real-Time Detection Monitoring

Real-world testing of the AI-powered facial analysis system for autism diagnosis demonstrated that it could efficiently handle unseen data while maintaining an inference speed suitable for practical and clinical applications. According to studies such as "AI-Assisted

ISBN:97881-19905-39-3

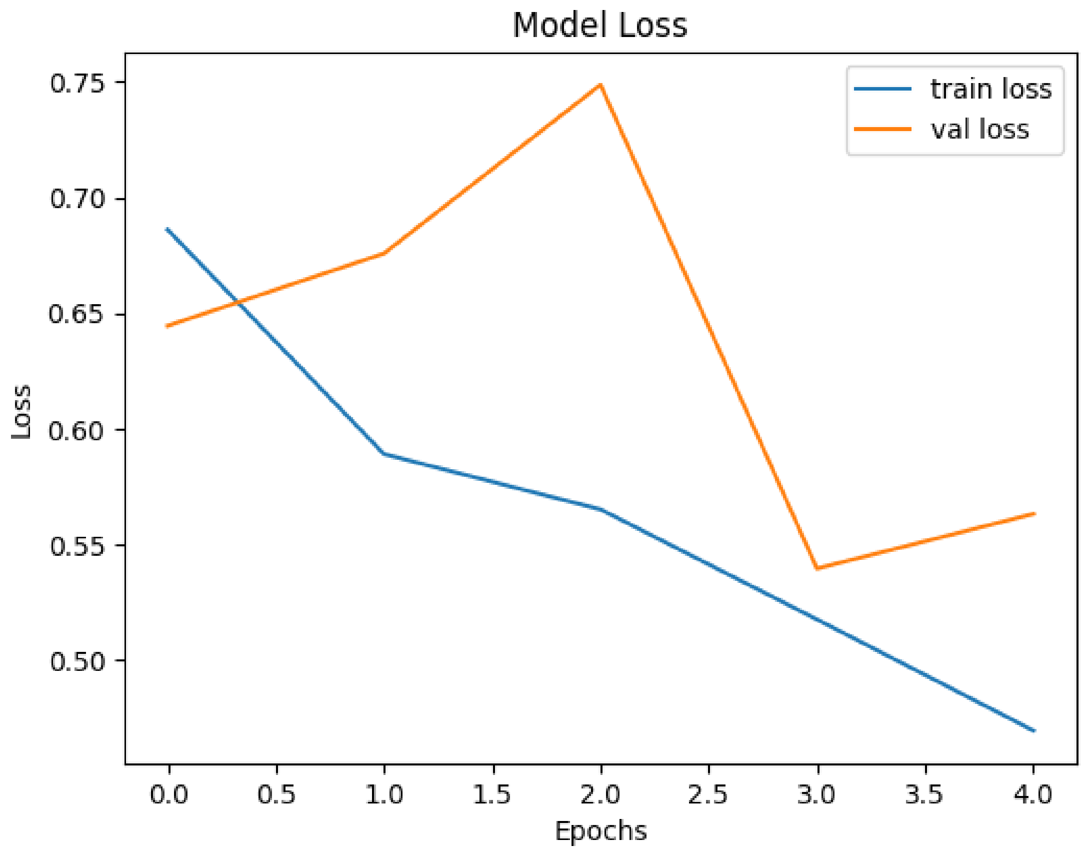

Figure 9:

Autism Diagnosis in Healthcare" and "Machine Learning for Behavioural and Cognitive Analysis," the system's diagnostic accuracy in real-time scenarios was comparable to or higher than that of current methodologies. Minor facial clues were difficult to identify in a variety of populations, which is similar with the findings of "Advances in Autism Spectrum Disorder Recognition Using AI." Annotated face analysis results proved the system's ability to accurately identify autism-related traits and enable early intervention strategies.

4.2 Comparison with Existing model

The bar graph depicts the performance indicators for the AI-powered facial analysis system in autism diagnosis.

- Accuracy: 87.0%, demonstrating the system's overall ability to effectively recognise

ISBN:97881-19905-39-3

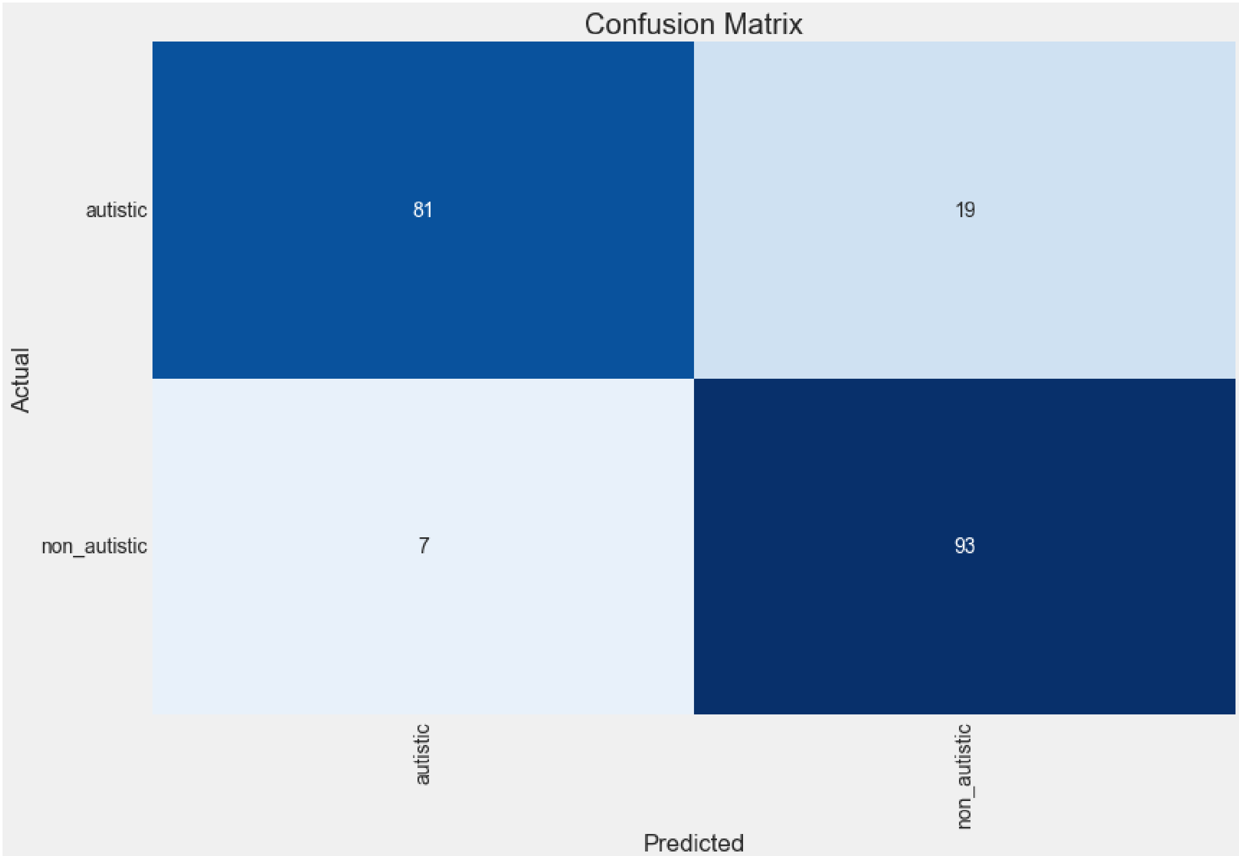

Figure 10: 1Confusion Matrix Values for ASD Detection Model

scenarios. Precision: 92.0%, indicating how effectively it reduces false positives.

- Recall: 81.0%, indicating that actual autism cases were successfully identified; nonetheless, there is still room for improvement in identifying all affected individuals.

- The F1 Score is 86.2%, indicating a balance of recall and precision.

These findings demonstrate that the system has good diagnostic skills, particularly in terms of precision, making it a tool with potential for real-world use in healthcare settings.

4.3 Conclusion

Results of the Proposed System AI-powered facial analysis, which employs cutting-edge machine learning and computer vision techniques, has dramatically increased autism

ISBN:97881-19905-39-3

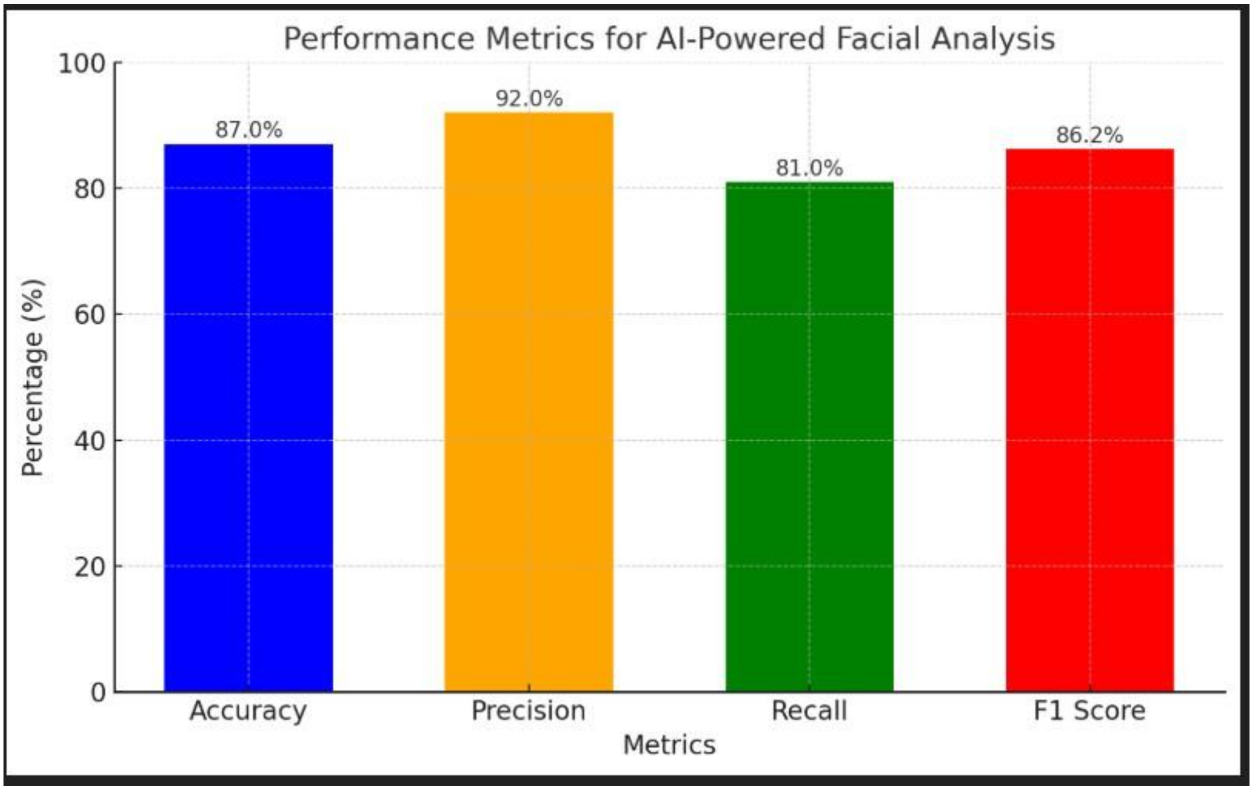

Figure 11: 1Performance Comparison with Existing ASD Detection Models

detection early on. It provides an automated, non-invasive way for identifying minute facial expressions and traits associated with autism spectrum disorder (ASD). By eliminating the need for time-consuming and subjective manual examinations, the system provides more accessible, scalable, and objective diagnostic options. Real-time analytic capabilities allow for early identification and prompt responses that can improve developmental outcomes. Furthermore, these systems are becoming more adaptable in a variety of scenarios, including at home and in hospitals, making them suitable for a wide range of users.They promote healthcare equity, optimise resource allocation, and offer physicians and caretakers with actionable insights by reducing diagnostic delays.

4.3.1 Future Extent

The potential for increasing the utility and effectiveness of AI-powered facial analysis for autism diagnosis is considerable. Scalability remains a crucial issue, given the flexibility to

ISBN:97881-19905-39-3

tailor the system to different age groups, cultural contexts, and demographic differences. Even with limited training data, generative AI and other advanced machine learning techniques can improve the accuracy of detecting complex or rare indications. Integration with IoT devices for continuous behavioural monitoring, such as tracking eye movement and social interactions, may provide a more thorough diagnostic approach. Its relevance and usability will be expanded further by localisation for various languages and geographical areas, hence boosting worldwide accessibility.Furthermore, enhancing energy efficiency through the use of edge computing and low-power devices could minimise resource use while maintaining performance. These advances contribute to the goal of personalised and equitable healthcare, establishing AI-powered facial analysis as a game-changing tool for improving autism diagnosis and treatment around the world.

ISBN:97881-19905-39-3