Bulletin of Business and Economics, 13(1), 557-572
https://bbejournal.com
https://doi.org/10.61506/01.00241

Autism Spectrum Disorder Detection in Children Via Deep Learning Models Based on Facial Images

Bawar Khan[1], Sohail Masood[2], Arslan Akram[3]

Abstract
Autism spectrum disorder (ASD) is a complicated psychiatric disease that causes difficulty in communicating with others, and restricted behavior, speech, as well as nonverbal interaction. Children with autism have unique facial characteristics that distinguish them from ordinarily developing children. Therefore, there is a requirement for a precise and automated system capable of early detection of autism in children, yielding accurate results. The objective of this research is to assist both families and psychiatrists in diagnosing autism through a straightforward approach. Specifically, the study employs a deep learning method that utilizes experimentally validated facial features. The technique involves a convolutional neural network along with transfer learning for the detection of autism. MobileNetv2, Xception, ResNet-50, VGG16 and DenseNet-121 were the pretrained models used for detection of autism. The evaluation of these models utilized a dataset sourced from Kaggle, comprising 2,940 facial images. We evaluated the five deep learning models using standard measures like recall, precision, accuracy, F1 score, and ROC curve. The proposed DenseNet-121 model outperformed existing transfer learning models, with 96% accuracy rate. With respect to performance evaluation, the proposed method exhibited superiority over the most recent models. Our model possesses the capability to support healthcare professionals in validating the precision of their initial screening for Autism Spectrum Disorders (ASDs) in pediatric patients.

Keywords: Autism spectrum disorder, Facial features, Convolutional neural network, Transfer learning, DenseNet-121

1. Introduction
A neuro-developmental illness known as autism spectrum disorder (ASD) causes a variety of challenges connected to communication, social interaction, and the presence of unpredictable behavioral patterns in both adults and children (Francés et al., 2023) . The early indications of ASD often occur during the very first year of childhood(Carette et al., 2018),and this could mean not making eye contact, not reacting to being called names, or not caring about the people who are caring for them. Only a tiny percentage of children initially exhibit no indicators of autism but subsequently begin to show symptoms between the ages of 18 and 24 months(Kanner, 1968). According to Alfalasi, In the US, one child in every 54 suffers with autism (Modak, n.d.). Autism affects about one in every 36 children in the United States, according to the Centres for Disease Control and Prevention in 2023 as shown in figure 1 surveillance year wise and Autistic spectrum disease affects more than 75,000,000 people, or approximately 1% of the global population (*69+ Autism Statistics*, n.d.).As reported by the World Health Organization (WHO), one out of every 100 children worldwide is identified with ASD features each year (Zhu et al., 2023). Recent study data(Goh et al., 2016) indicates that the majority of children with ASD do not receive a diagnosis until they have reached the age of three. Early ASD treatment is beneficial for toddlers who are still growing (Elbattah & Cilia, 2023).

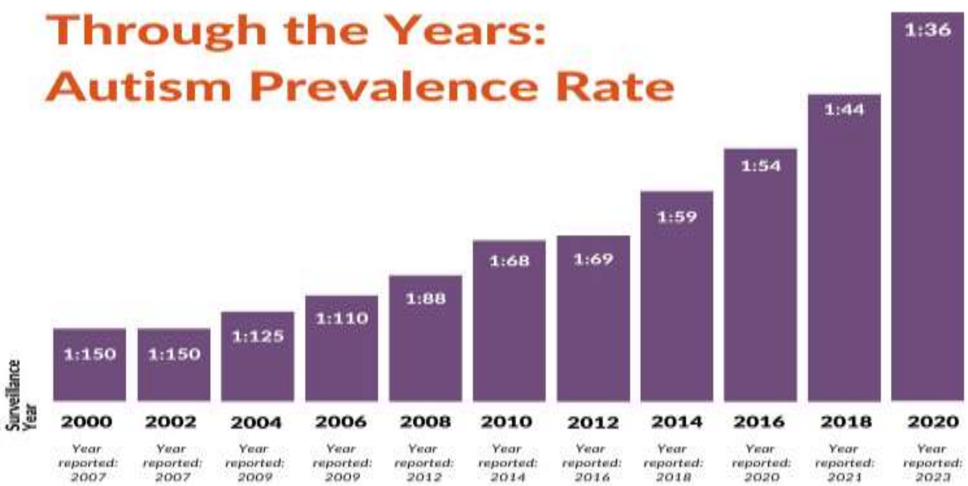

Figure 1: Prevalence of autism as reported by the CDC

However, because ASD cannot be accurately identified by displaying merely the actions of children in a clinic, a substantial amount of valuable time might be lost during the diagnosis process. A variety of clinical techniques can be used to identify autism as early as possible; however, they are laborious diagnostic procedure that are rarely utilized unless there is a significant predicted risk that the child will develop ASD and needs highly qualified medical professionals (Rahman et al., 2020). Many different approaches

[1] Corresponding Author, Faculty of Computing, Department of Computer Science and Information Technology, The Superior University, Lahore, 54000, Pakistan, bawerkhan@gmail.com
[2] Faculty of Computing, Department of Computer Science and Information Technology, The Superior University, Lahore, 54000, Pakistan
[3] Faculty of Computing, Department of Computer Science and Information Technology, The Superior University, Lahore, 54000, Pakistan, MLC Lab, Okara, 56310, Punjab, Pakistan

have been used to look over the key features of autism. These include facial-feature extractions(Rashid & Shaker, 2023)eye-tracking tactics (Kanhirakadavath & Chandran, 2022),(Thanarajan et al., 2023)face recognition (Mujeeb Rahman & Subashini, 2022)bio-medical image analysis (Aldhyani et al., 2022),application development(Schelinski et al., 2016),(Alsaade & Alzahrani, 2022), and recognition of speech (Mohanta & Mittal, 2022). Among these techniques, facial recognition stands out for its ability to reliably diagnose autism and determine an individual's mental and emotional state in real time. A widely utilized technique, it is employed to analyze human features and extract distinctive attributes that differentiate normal from abnormal expressions. Additionally, it is employed to mine substantial data in order to unveil patterns of behavior.

Machine learning, deep learning, and artificial intelligence can all be used to solve many social problems. AI has been used to help doctors treat autism. Researchers touted artificial neural networks as a means to extract traits of autism spectrum disorder patients that can distinguish them from controls. Researchers have utilized deep learning methods such as CNNs, RNNs (M. Li et al., 2019) ,(Pokorny et al., 2017),Xception,VGG16,VGG19,ResNet152V2,EfficientNetV2,Inception,ResNetV2,ResNet34,ResNet50,MobileNetv2,AlexNet, (Sridurga, P.D., et al., n.d.),(Kaur & Gupta, n.d.) (Ahmad et al., 2024), as well as the long short-term memory (LSTM) model (J. Li et al., 2020), (Ahmed et al., 2023), MR-CNN (Akram, Rashid, Jaffar, et al., 2023),for image classification. In recent years, more studies have used machine learning methods like multilayer perceptron (MLP), AdaBoostAB gradient boosting machine (GBM) random forest (RF),(Rabbi et al., 2021),DRLBP-SVM (Akram et al., 2022),XGBOOST (Akram, Rashid, Hajjej, et al., 2023),WLD-XGBOOST (Akram, Rashid, et al., 2024),and Curvelet-transformation with SVM (Iqbal & Khan, 2017; Akram, Khan, et al., 2024) to detect difference diseases from images. Deep learning models replace machine learning models because they can recognize complex patterns, extract hierarchical features, and handle massive volumes of data. Deep learning is vital in medicine for detecting and classifying wounded or disturbed bodily parts.

Deep learning models are effective in various domains like medical analysis, object detection, fashion recognition, and pattern identification. Training data samples and patterns is crucial for enhancing their performance and accuracy. Transfer learning is utilized to enhance the performance by applying models to new or limited datasets after training on extensive datasets. These models learn and extract features from the initial training, which can be used for new tasks or smaller datasets (Talo et al., 2019) . Transfer learning-based models are used to overcome resource limitations and save time. Pretrained models extract common features from new data samples, while additional layers are concatenated to capture hidden patterns. Transfer learning is famous for its ability to recognize intricate patterns and features, save time, and adapt to smaller sets of data by modifying features from a larger dataset.

Researchers are creating computer-aided decision support systems that employ machine learning to detect and treat autism in children. Surveys and interviews were utilized to collect data on children's behaviour and create ML-based prediction models for early ASD detection. These researchers did not use facial pictures to diagnose autism in children. Most people acquire text and numerical data via time-consuming and error-prone interviews and questionnaires. Children under eight may be apprehensive during interviews and questionnaires, giving incorrect answers due to insufficient understanding. Incorrect responses might skew the prediction model, diminishing accuracy. The constraint of machine learning systems is that the accuracy of classification is dependent on handcrafted features identified through feature extraction algorithms, and takes additional computational time to process large amounts of data. Previous studies using deep learning models yielded unsatisfactory results due to under fitting, overfitting, low-quality datasets, and complex architecture. We created a simplified prediction model using facial photographs from 2940 2-8-year-olds to address these concerns.

This study contributes to the field of autism spectrum disorder (ASD) by offering a robust strategy that uses pre-trained deep learning models. Furthermore, a revolutionary facial recognition technology is being developed to help medical professionals in the early detection of ASD, with a focus on lowering the condition's impacts and mortality risk. The following is the planned work's major scope:

- A robust approach has been designed to detect children with Autism Spectrum Disorder (ASD), employing five pre-trained deep learning algorithms, including Xception, ResNet50, MobileNetV2, VGG16 and DenseNet121.
- DenseNet121 demonstrated the most effective performance among the five deep learning algorithms that were pretrained.
- The dataset has also been evaluated using several pretrained models, including XceptionNet, ResNet50, MobileNetV2, VGG-16 and DenseNet-121 model.
- The proposed method's efficacy was evaluated in comparison to that of various state-of-the-art designs using metrics such as precision, accuracy, recall, and f1-score.
- To assist doctors in the early diagnosis of ASD, a facial recognition system has been developed.

The format of this article is as follows: Section 2 contains an overview of the literature. The techniques and materials are discussed in Section 3. Section 4 shows the results. Conclusion, limitations, and suggestions for future research are respectively addressed in Sections 5.

2. Literature Review

Many autistic spectrum disorder detection studies have been published. Most research on childhood autism spectrum disorder uses traditional machine learning as well as image processing algorithms. Before deep learning, machine learning algorithms were frequently utilized in image classification. However, feature extraction with machine learning technique, is a handmade process, and false extraction might result in missed identification and categorization of autism spectrum disorder (ASD). Therefore, DL algorithms including convolutional neural networks (CNN) are more popular in classification of images due to their automatic extraction as well as the important features of learning. Further research is necessary to validate the effectiveness of this approach, as its accuracy is still being investigated. Recently, a significant amount of work has already been made in using facial images to check for ASD.

Artificial intelligence has revolutionized the analysis of autism spectrum disorder (ASD) by means of biological pictures. (Rashid & Shaker, 2023) proposed convolutional neural network (CNN) along with transfer learning and facial impressions methodology to classify the condition. The models, Xception and VGG16, were trained using 2,940 face photos from the Kaggle platform. The Xception model obtained the maximum accuracy rate of 91%, followed by VGG16 with 75%. This breakthrough has saved lives and improved the understanding of ASD. The proposed model generated fewer results and did not calculate precision, recall, or f1-score. (Alsaade & Alzahrani, 2022) outlined the development of a web application using deep learning techniques to identify autism based on facial features. Three pre-trained models, Xception, VGG19, and NASNETMobile, were used for classification. The dataset used was 2,940 face images from Kaggle. With an accuracy of 91%, the Xception model outperformed VGG19 (80%) and NASNETMobile (78%). There are a few issues with this strategy, such as the overfitting problem and the requirement for increased accuracy.

The Deep learning models, specifically Xception and VGG19, to identify children's facial image landmarks for detecting Autism Spectrum Disorder (ASD) were also proposed in 2022 by (Sridurga, P.D., et al.,). The dataset consisting of total 2,940 facial images of healthy and autism patients were collected using the Kaggle platform. Xception achieved 86%, while VGG19 achieve 81% accuracy. The results indicate that the established approach is effective at detecting ASD. The suggested model exhibited lower accuracy, and additional assessment metrics like f1-score, recall, and precision weren't performed. (Rabbi et al., 2021) introduced an artificial intelligence algorithms to detect autism in children using images that may not be easily understood by the general public. Five algorithms, including Random Forest, Multilayer Perceptron, AdaBoost, Gradient Boosting Machine, and Convolutional Neural Network (CNN) achieved an accuracy of 92.31%. However, there is a lack of both the procedure and the evidence that supports it.

(Alkahtani et al., 2023) aimed to enhance the predictive performance of a convolutional neural network (CNN) model through the application of transfer learning techniques, specifically utilizing MobileNetV2 and a hybrid VGG19 model. Various machine learning methods, including logistic regression, linear support vector machine (linear SVC), random forest, gradient boosting, MLPClassifier, decision tree, and K-nearest neighbors, were employed to assess the model. The evaluation was conducted on a Kaggle dataset with 2940 images depicting both autistic as well as non-autistic children. The results showed a 92% accuracy rate for MobileNetV2, indicating superior transfer learning strategies compared to existing systems. The suggested model was not assessed using precision, recall, and f1-score evaluation metrics. The authors presented a sophisticated design that demanded excessive processing resources and effort, yet yielded unsatisfactory results.(Akter et al., 2021) used an enhanced six pre-trained CNNs with transfer learning approach and a number of machine learning classifiers as baseline classifiers after obtaining the facial picture dataset of autistic children from Kaggle. Based on experimental data, the improved MobileNet-V1 model exhibited exceptional performance, achieving the maximum accuracy of 90.67% and the least values of both fall-out as well as miss rate at 9.33%, outperforming other classifiers as well as models that have already been trained. The suggested framework has certain drawbacks; for example, the quality of the few face photos that have been used is not good. In the future, the authors of the study hope to enhance the overall accuracy of the model.

(Sadik, R., S. Anwar, and M. Reza, 2021) used a deep learning framework to identify autism based on facial expressions. They used a Convolutional Neural Network (CNN) with MobileNet and achieved 89% peak accuracy for validation and 87% test accuracy. The approach's robustness was also highlighted, with 87% F1-score and precision metrics. Although the recommended work has produced an unsatisfactory result, there is still room for improvement. (Gaddala et al., 2023) utilized deep CNNs to classify Autism Spectrum Disorder (ASD) using facial image analysis. The models, trained on facial image data, achieved an 84% accuracy rate. This research highlights the ability of deep learning methods to improve autism diagnosis methods, suggesting further research for fine tuning and optimization on larger scales as well as accuracy needs to be increased.

In a study by (Singh et al., 2023) , transfer learning was used to develop a model for detecting ASD from facial images. The kaggle 2940 dataset was used to train and test the suggested models. The models used in this study are VGG16, InceptionV3, MobileNet, Xception, EfficientNetB0, and EfficientNetB7. These models have accuracy levels of 86.3%, 86.1%, 88%, 87.7%, 85.6%, and 82.6%, respectively. The result of the models is not up to the mark to the suggested models yet has opportunity for growth.(Venkata Sai Krishna Narala et al., 2023) developed a deep-learning model using facial images to identify autism spectrum disorder. The model assessed a child's signs of autism and normal development, using specific facial measures to differentiate between those with ASD and those with typical development. The model was developed using Autism Image Datasets, with an Efficient Net convolutional neural network contributing to its 88% accuracy but this model needs to improve the accuracy.

(M. Ghazal et al., 2023) developed a deep transfer learning approach with AlexNet for early detection of ASD using facial characteristics. They used a Kaggle dataset of autistic and non-autistic children's facial images. Their system, ASDDTLA, achieved an accuracy of 87.7%, outperforming existing models, proving its efficacy in early ASD detection based on facial characteristics. Researchers hope to improve the model's performance. In the section of literature review, various scholarly articles are examined, presenting their respective proposed methodologies, accompanied by discussions on the datasets employed and the corresponding accuracy metrics. Table 1 summarizes the research study, including methodology, dataset, findings, and remarks. Reference (Rashid & Shaker, 2023)**,** (Alsaade & Alzahrani, 2022)**,** (Sridurga, P.D., et al., n.d.)**,** (Rabbi et al., 2021)**,** (Alkahtani et al., 2023)**,** (Akter et al., 2021)**,** (Gaddala et al., 2023)**,** (Singh et al., 2023) and (M. Ghazal et al., 2023) utilized a similar dataset comprising 2,940 images obtained from Kaggle. Several models employ intricate architectures, and they all require enhancements to yield more precise results. Nevertheless, there is still room to fine-tune their models and enhance overall performance. This study employed the transfer learning method to train pre-existing models in Autism Spectrum Disorder (ASD) detection using facial images for diagnostic purposes.

Table 1: Displays an overview of previously discussed research studies, including their suggested technique, dataset, results, and limitations

Ref	Techniques	Dataset	Result	Limitation /Future work
(Rashid & Shaker, 2023)	Xception VGG16	2,940	91% 75%	Accuracy need to be improved and to do data augmentation
(Alsaade & Alzahrani, 2022)	Xception VGG19 NASNETMobile	2940	91% 80% 78%	Overfitting issue as well as result should be improved
(Sridurga, P.D., et al.,)	Xception VGG19	2940	86% 81%	Need to improve accuracy
(Rabbi et al., 2021)	CNN	2940	92.31%	The procedure and supporting evidence are missing.
(Alkahtani et al., 2023)	MobileNetV2	2940	92%	A significant percentage of positive along negative outcomes were wrong.
(Akter et al., 2021)	MobileNet-V1	2936	90.67%	The suggested system has some drawbacks, such as using low-quality facial images. Improvements are needed for accuracy.
(Sadik, R., S. Anwar, and M. Reza, 2021)	MobileNet	3136	89%	Achieved satisfactory results but has room for growth, with fine tuning the model potentially improving model accuracy.
(Gaddala et al., 2023)	VGG16 & 19	2936	84%	Accuracy need to be improved
(Singh et al., 2023)	MobileNet Xception InceptionV3 EfficientnetB0 EfficientnetB7 VGG-16	2940	88 % 87.7 % 86.1 % 85.6 % 82.6 % 86.3 %	Accuracy need to be improved
(Venkata Sai Krishna Narala et al., 2023)	Efficientnet	2830	88%	Accuracy need to be improved
(M. Ghazal et al., 2023)	AlexNet	2940	87.7%	The models accuracy need to be improved

3. Materials and Methodology

This section outlines the comprehensive methodology proposed, encompassing the entire process from data acquisition to the evaluation of final results. The key stages in the proposed methodology include data preprocessing, which involves tasks such as data resizing and augmentation techniques. The preprocessed data is subsequently input into our chosen pre-trained models, like Xception, ResNet50, MobileNetV2, VGG16, and DensNet121, to identify the class of children as either autism or non-autism. These deep learning models were selected primarily due to their inherent advantages and benefits. These include their well-suited architectures, the advantages of transfer learning, design flexibility, and the substantial aid they have garnered from both the DL (deep learning) and ML (machine learning) communities. These models demonstrated enhanced proficiency in identifying and understanding complicated patterns and features within images, a critical factor for precise identification of autism.

Additional pre-trained models were assessed based on metrics including F-1 score, precision, accuracy, and recall. The basic structure of our methodology for classifying children as having autism or not is shown in Figure 2. This work primarily focuses on implementing pre-trained models by inserting updated final layers through layer freezing, fine-tuning, and training. The objective is to address the challenging task of classifying and detecting autism or non-autism in facial images. We processed the images using data resizing and augmentation techniques before feeding them to pre-trained models that automatically extracted the features. In terms of testing accuracy, from the above-used pretrained models DenseNet-121 performed the best among these models.

3.1. Dataset

We used images of the faces of children from a dataset accessible on Kaggle for this study (*Autism Dataset*,). This dataset is unique in that it is the only one of its kind that is freely available to the public. The ages of the children included in the sample varied from two to fourteen, with the bulk falling somewhere between two and eight years old. All of the images were standard 2D RGB colour images in JPEG format. This dataset had 2,940 children images totally, split into two groups: images of autistic children as well as images of non-autistic children. You can see a small sample of what the dataset looks like in figure 3. The entire dataset is split into an 80:20 ratio as shown in table 2, where 80% of the data is designated as a training set, while the rest of the 20% is reserved for testing. Moreover, 20% of the training part of data have been selected for verification.

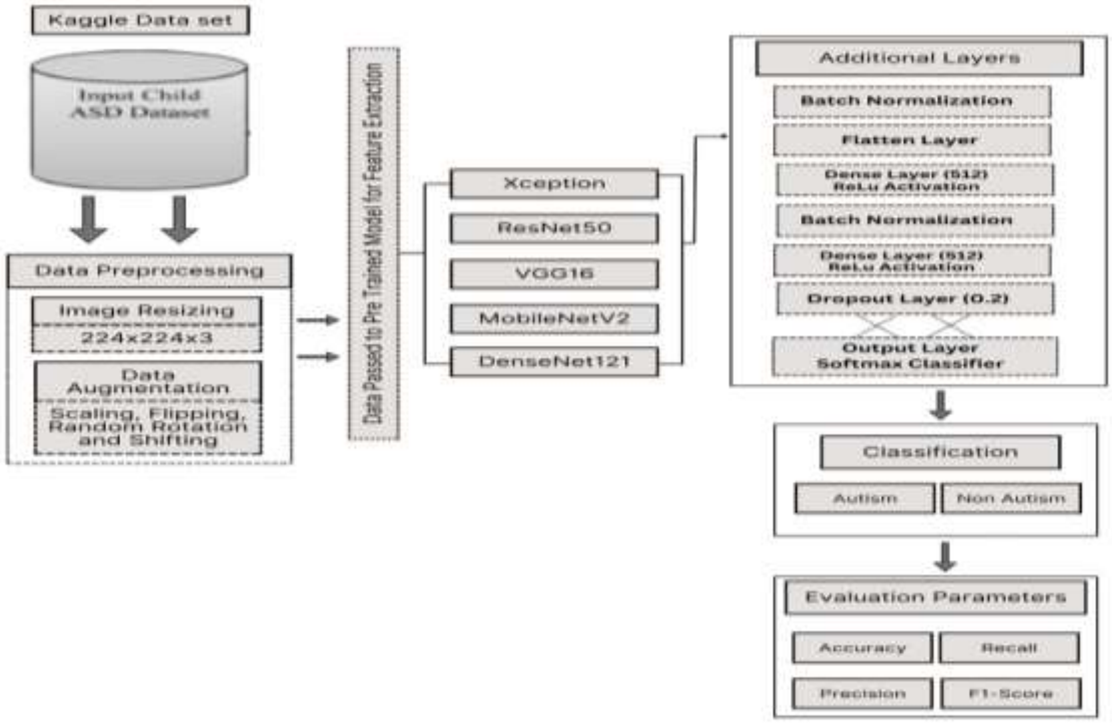

Figure 2: The pre-trained suggested models' workflow architecture

(a) (b)

Figure 3: Facial Images of (a) Autistic Children (b) Non-Autistic Children

Table 2: The dataset is divided into two phases: testing and training

Class	No of face Images	Training set	Testing set
Autistic	1470	1176	294
Non Autistic	1470	1176	294
Total Images	2940	2352	588

3.2. Pre-processing

Data preprocessing was done to clean up as well as crop the images because they were gathered from the internet by G. Piosenka. Before training the deep learning model, the images needed preprocessing because they varied in size, had different compression standards, and contained impurities. Therefore, preprocessing was necessary. The images had different sizes, even within each category. To make them consistent and improve results, all the images were changed to have a size of 224x224 pixels. Data augmentation techniques were employed on preprocessed data before feeding them into deep learning models. Utilizing this technique offers a primary benefit by enhancing model predictions and accuracy. Additionally, it proves advantageous in mitigating issues related to data overfitting. A horizontal flip entirely allows learning the features of both sides of the face. Normalization was employed as a scaling method, involving the rescaling of parameters for all images. This process adjusted the pixel values of the dataset from the original range of [0, 255] to a normalized range of [0, 1]. The data images undergo a rotation of 05 degrees, along with a height and width shift range spanning [-0.1, 0.1]. This range is employed to indicate the maximum allowable fractional adjustment in both width and height. Additionally, to avoid the model from focusing too much on specific details and overfitting, dropout layers were applied

3.3. Deep Learning Models

Convolutional neural networks (CNNs) have been utilized since the 1980s for the purpose of picture recognition and classification (Rawat & Wang, 2017). CNNs have demonstrated excellent performance in visual tasks in recent years, owing to advances in computing power, the availability of massive datasets for training, and the use of transfer learning for customizable classification (Duranta et al., 2023). Feature identification is the basic goal of any face recognition as well as object classification model. This makes it possible to categorize faces into two groups: those with autism and those without it. A method called "transfer learning" helps models to learn from a huge number of pictures. When applied to facial photos, convolutional neural network (CNN) models utilizing a transfer learning strategy can considerably improve early screening for autism spectrum disorder (Ghosh et al., 2021). To make binary classification easier, the layers of classification are modified. Through preprocessing the data before feeding it into the models, I have tried out a number of different deep learning models. This paper uses five models that have already been trained to find autism in face feature images. These models are Xception, ResNet50, MobileNetV2, VGG16, and DenseNet-121.

3.3.1. Xception Model

The architecture of this model is characterized by a simple modular design inspired by Google's Inception model. The system has three primary sections: entrance, center, and exit, each including unique convolutional layers and relu activation functions. The input image dimensions are 299 x 299 x 3. The entry flow processes the input and extracts features with dimensions of 19 x 19 x 728. Residual connections ensure that the output of each layer in a block is the maximum value, preserving the feature map as it undergoes nine passes through convolution layers in the central block. The last component of the processing pipeline generates an output with a standard-sized input image comprises 2048 features. Subsequently, these features are fed through a fully connected (FC) layer to the prediction layer, where adjustments are implemented in the final layers to enable binary classification. The Xception network uses the following mathematical equation to describe a pointwise along with depthwise separable convolution:

Point-wise-Conv $(W, y)_{(i,j)} = \sum_{m}^{M} W_m \cdot y_{(i,j,m)}$ (1)

Depth-wise-Conv $(W, y)_{(i,j)} = \sum_{k,l}^{K,L} W_{(k,l)} \odot y_{(i+k,j+l)}$ (2)

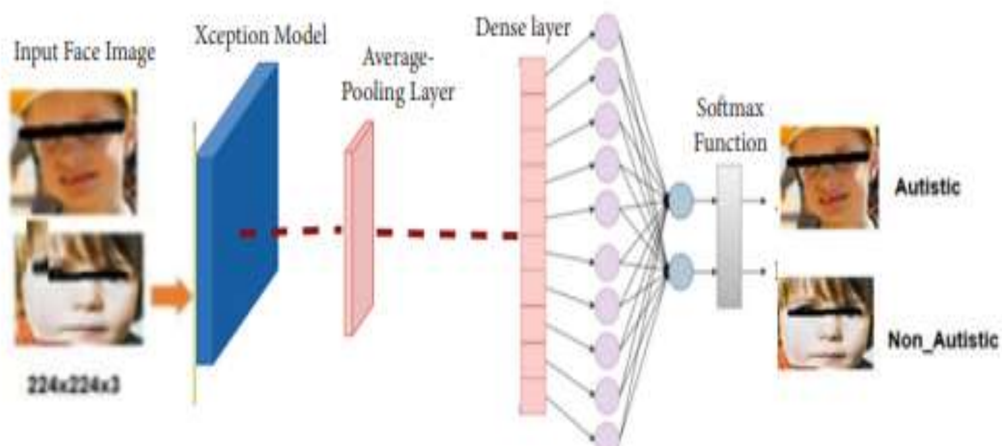

Figure 4: Xception model architecture

3.3.2. ReNet-50

Introduced in 2015, the ReNet-50 model is part of the family of residual network, distinguished by means of its utilization in residual blocks. These blocks play a crucial role in mitigating the vanishing gradient problem commonly associated with deep networks, facilitating the training of deeper models. ResNet50's input layer accepts an image with dimensions of 224x224. The architecture comprises four distinct stages, the quantity of layers and blocks varies according to each stage. The architecture of ResNet50 consists of sixteen residual blocks, with each block containing multiple convolutional layers connected by residual connections. The architecture also includes fully connected layers, pooling layers, and a softmax output layer specifically designed for classification purposes. ResNet focuses on learning residual representations of input data by fitting the residual mapping, which consists of multiple layers. The output of these layers is added to the input before being passed on to the next block. In residual networks, the following formula stabilizes and optimizes deep neural network training.

$X(out) = F(X(in) + X(in)$ (3)

where the residual block input is X_ (in). F (X_ (in) shows how the layers inside the block change the input. The residual block output, X_ (out), is the sum of the modified and original inputs. Adding input to the modified output helps ResNet efficiently learn the residual mapping, focusing on the challenging sections rather than starting from scratch. Residual learning simplifies training deep networks with enhanced performance.

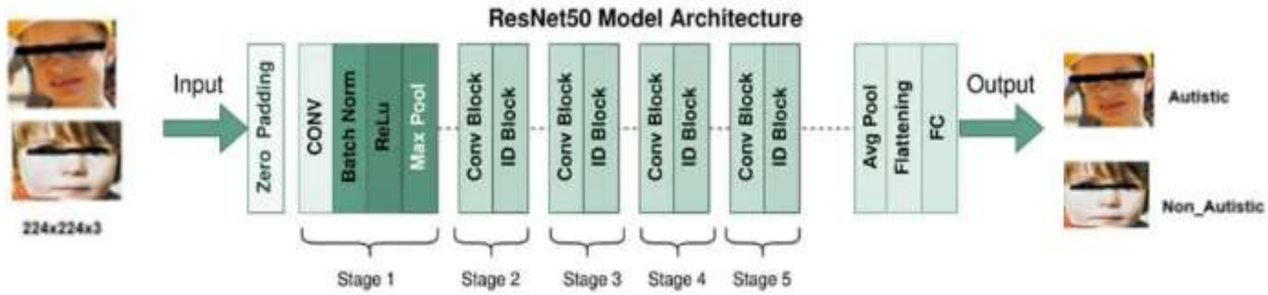

Figure 5: ResNet-50 model architecture

3.3.3. MobileNetv2

The MobileNet model, created by Andrew G in 2017, is specifically designed for inserted and mobile devices using constrained computational assets. It is a lightweight model tailored for implementing classification tasks on such devices. The goal of this MobileNetV2 architecture is to make a connection from one bottleneck layer towards another. MobileNetV2 expects images of 224x224 pixels with three color channels. There are 32 complete convolution layers that come before the 19 residual bottleneck layers in the inverted residual architecture. These convolution layers use nonlinear filter features to perform depth-based convolutions. The primary concept was to employ depth-wise separable convolutions to minimize computational demands and parameter count, all while preserving high accuracy. The skip connection as well as depthwise separable convolutions may be expressed mathematically as follows:

$$Y_{i,j,k} = \sum_{m,n} X_{i+m,j+n,k} \times K_{m,n,k} \tag{4}$$

The formula computes every element in the resulting feature map Y by calculating the weighted sum of the corresponding elements in the input feature map X, with the weights determined by the kernel K.

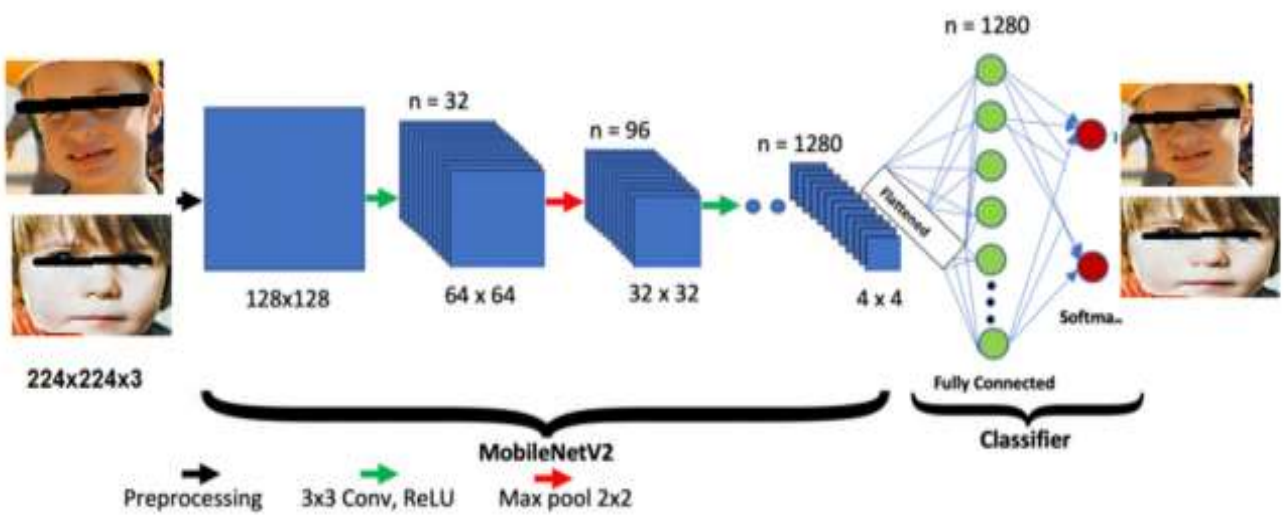

Figure 6: MobileNetV2 model architecture

3.3.4. Visual Geometry Group Network (VGG) Model

The VGG16 model, unveiled by the Visual Geometry Group at the University of Oxford in 2014, is a deep learning model. This model comprises 16 layers, featuring 13 convolutional and 3 fully connected layers. It accepts input images sized at 224 × 224 pixels with three color channels. There are 1000 neurons in the output layer corresponding to ImageNet's classes, using softmax activation. The VGG16 architecture contains mostly of convolutional layers, after which there are fully linked layers. The mathematical representation of a fully connected layer is:

$$FC(X) = \sigma\,(WX + b) \tag{5}$$

Where:

The input vector is represented by X. W stands for the matrix of weights. Here, b stands for the bias term. The activation function is denoted by σ.

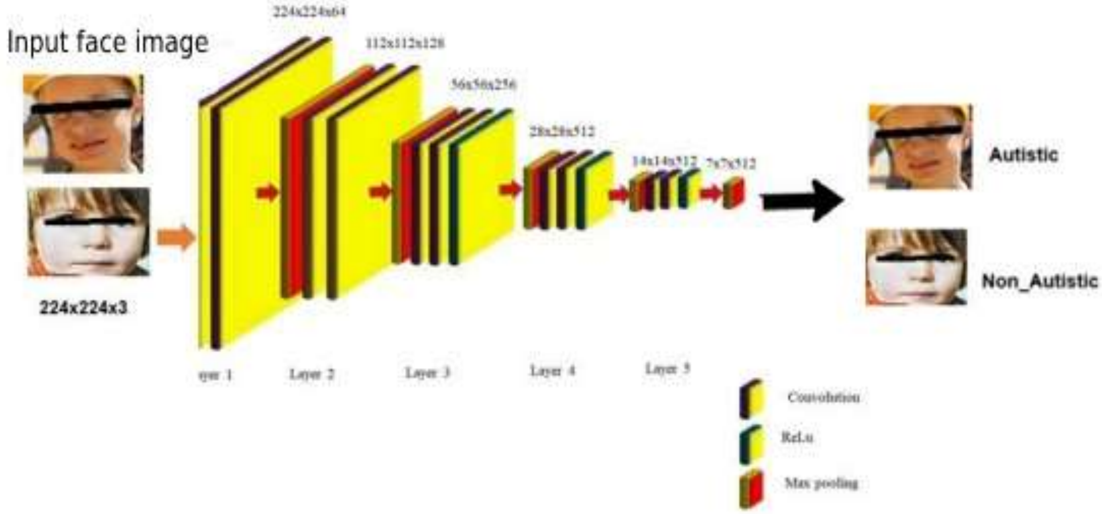

Figure 7: VGG16 model architecture

3.3.5. DenseNet121

Gao Huang introduced DenseNet121, a model for deep learning with 121 layers arranged into dense blocks, each comprising a defined amount of layers. The input to the network is an RGB image with a standard size, typically 224x224 pixels. Additionally, the model incorporates transition layers accountable for the feature map of down-sampling to reduce its size and Batch normalization along with ReLU activation functions are applied after each convolutional layers. Supporting layers, including global average pooling and softmax output layer, are also integrated into the structure, making DenseNet121 effective in image recognition with reduced vanishing gradient issues. The following is a representation of the dense block:

$X_l=H_l([X_{0,}X_{1,}\ldots X_1-_1])$ (6)

Here:

The output of the i-th layer in the dense block is denoted by X_l. H_l denotes the composition of operations performed at layer l. The notation $[X_{0,}X_{1,}\ldots X_1-_1]$ indicates the concatenation of the feature maps from each dense block layer that came before it.

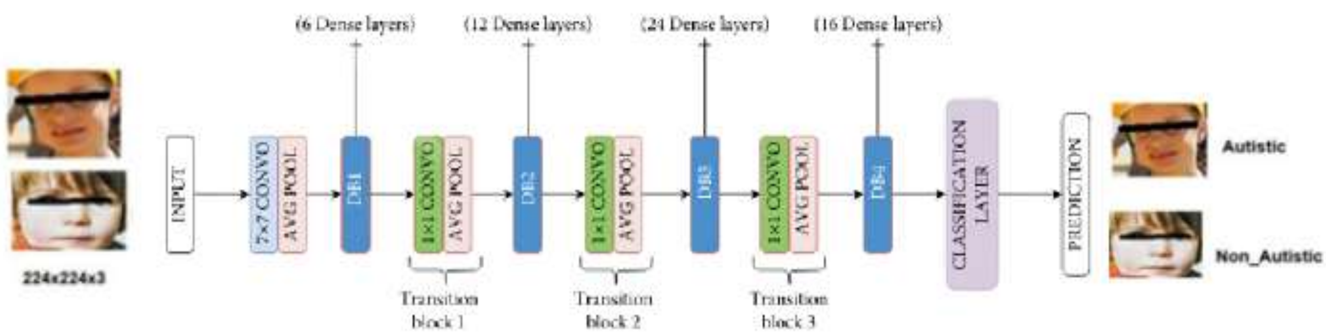

Figure 8: DenseNet-121 model architecture

Each network used in this study has a set number of layers along with parameters. Table 3 shows the whole number of layers and parameters for each model, which sheds light on how many filters the models employ. This information can help comprehend each model's relative complexity and capacity, as well as inform the choice of the best model for a specific task

Table 3: Pre-trained Deep Learning Model's no of layers with parameters summary

Attribute	Xception	ResNet50	MobileNetv2	VGG16	DenseNet121
Layers	71	50	53	16	121
Total Parameters	72516138	75242370	34641986	27828034	32997954
Trainable Parameters	51649538	51649538	32380418	13111298	25957378
Non Trainable Parameters	20866600	23592832	2261568	14716736	7040576

3.4. Proposed Additional Layers

The pretrained models in this study were adjusted using updated final layers that were added via layer freezing and transfer learning was applied for classifying images (facial images) to detect autism and this enables the underlying model to concentrate on learning new customs layers in order to train fresh data, rather than forgetting its prior knowledge. The additional layers are replacing the fully connected layer of the pre trained models. The model incorporates transfer learning by building upon a pretrained base. We used data preprocessing as well as augmentation techniques on 224 × 224 × 3 images before sending them to pre-trained models, which automatically pulled out the features. Figure 9 depicts the suggested final layers for pre-trained models, which include a batch normalisation layer, a flattening layer, and two dense layers with ReLU activation, a dropout layer, and a softmax classifier. The initial layer applies batch normalization to normalize the input data. Subsequently, a flattened layer is employed to convert the multi-dimensional feature maps into a one-dimensional array. The following dense layer comprises 512 units with a rectified linear unit (ReLU) activation function, promoting non-linearity in the model. Another batch normalization layer follows to stabilize and expedite the training process. The next dense layer introduces additional complexity to the model with 512 units and a ReLU activation function. A dropout layer is implemented with a dropout rate of 0.2 in order to mitigate the issue of overfitting, which deactivates a fraction of neurons at random during the training. Finally, a dense layer with 2 neurons, representing the output classes (autistic and non-autistic), is utilized, followed by a softmax activation for classification purposes.

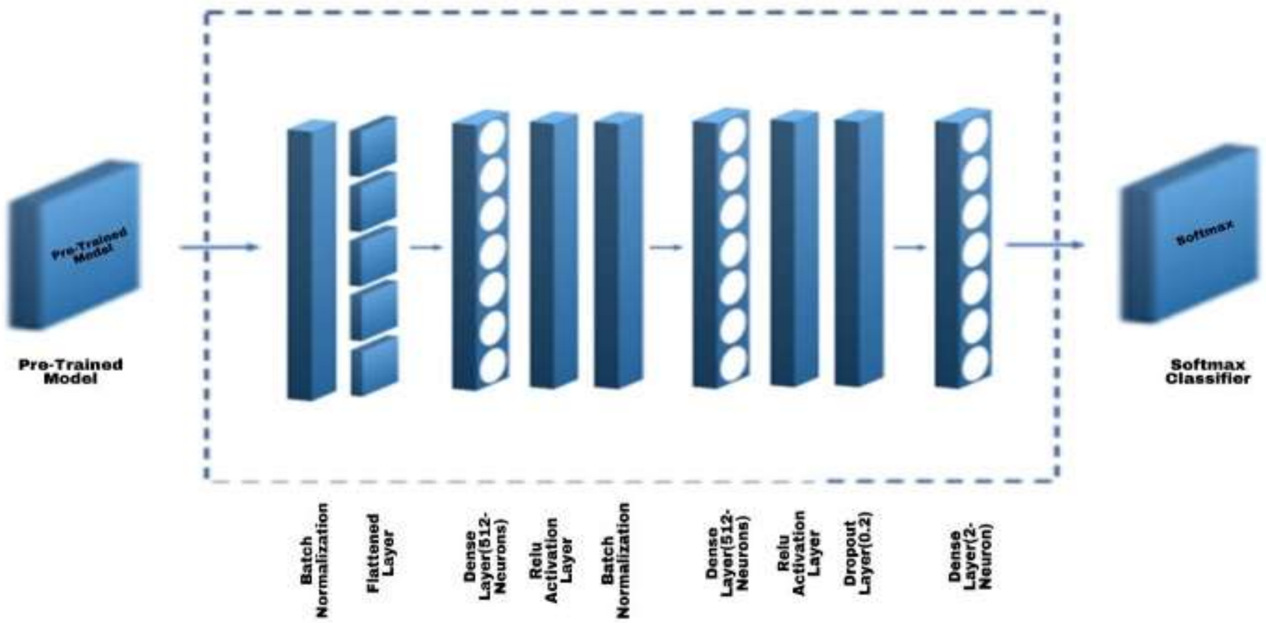

Figure 9: Schematic diagram of the proposed additional layers

3.5. Experimental Setup

This research comprises three main stages: (1) preprocessing and augmenting the dataset, (2) fine-tuning the deep convolutional neural network models with optimized hyper parameters, and (3) evaluating the models using relevant metrics for performance evaluation. Acquiring medical datasets is widely acknowledged as a challenging task, leading to difficulties in obtaining a sufficient number of datasets for training deep learning models. In this context, the Kaggle repository's autistic children dataset is utilized. The study involved implementing an intelligent autism detection system (ADS) by conducting experiments on various python libraries and hardware devices as shown in table 4.

Table 4: Displays the fundamental specifications for the ADS design

Sr.No	Hardware	Software & Libraries
1	Processor core i5	Drive on Google for dataset, Google Colab
2	8GB RAM	TensorFlow Keras , Panda, Seaborn,Matplot, Numpy, OpenCv
3		Python 3.10

3.6. Hyper Parameters

Table 5 provides a summary of the hyperparameters employed in the training of various deep learning models. Moreover, these hyperparameters were carefully chosen to enhance performance. A learning rate of $1e^{-4}$ was selected to facilitate a fine-tuning and gradual model convergence using our dataset. This value was established *via* practical testing and proven to be successful in our specific endeavour. Additionally, a batch size of 32 was chosen to optimize both memory utilization as well as training performance on our hardware configuration. The primary objective in determining this batch size was to attain precise and efficient results while optimizing the use of graphics processing unit (GPU) resources, all while avoiding memory overflow. We used softmax as our final selected classifier. The primary objective of the SoftMax classifier layer is to forecast the ultimate classification of the images and the class to which they correspond. The output is normalized into the distribution of probabilities across the class by the SoftMax function. The equations below demonstrate how the SoftMax function assesses the probability that an image is associated with a specific class.

$$S(Z_i)=\frac{\exp(Z_i)}{\sum_j \exp(Z_j)} \textit{ where } j = 1 \textit{ to } k \quad (7)$$

In the preceding mathematical equation, Z represents the length k's input vector, whereas exp (Z_i) symbolizes the exponential of the i-th element, and the sum of the exponential values for all elements in the input vector Z is given by $\sum_j \exp(Z_j)$. Finally, $S(Z_i)$ produces a distribution of probability for images over the k classes.

We employed categorical cross-entropy to calculate the average difference between expected as well as predicted values. Equation 8, which shows the loss measurement for the binary classification, uses the binary values 0 or 1 whereas y_{ij} to represent binary values and the probability p_{ij}.

$$\text{Categorical CE Loss} = -\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{NC} y_{ij} \cdot \log(p_{ij}) \quad (8)$$

As an optimizer, we selected Adam RMS Prop in order to attain the greatest loss mitigation feasible during the model training. By utilizing an adaptive gradient descent function, this technique of optimization accelerates the weights' approach to local minima. To enhance the robustness of the pretrained models and mitigate overfitting-underfitting issues, we conducted both training and testing operations over 50 epochs.

Table 5: Hyper parameters used to train different deep learning models

Hyper Parameters	Xception	ResNet50	MobileV2	VGG16	DenseNet121
Image Size			224x224		
Activation function			ReLU		
Output classification layer			Softmax		
Dense Layer			512		
Dropout value			0.2		
Learning rate			$1e^{-4}$		
Optimizer			RMSprop		
Loss-function			Categorical- Cross- Entropy		
Batch -size			32		
Epochs			50		

3.7. Performance Evaluation Metrics

This study employs distinct evaluation metrics for performance to assess the pre-trained models such as recall, f1 score, accuracy, precision, and ROC.The confusion matrix works as an indicator for evaluating classification performance by presenting a tabular representation of the accurate and erroneous outcomes in the testing results. The confusion matrix (CM) is a commonly used technique for evaluating the predictive performance of a trained model on a validation dataset. It features rows as well as columns that correspond to the true class labels, distinguishing between ASD and non-ASD. In the confusion matric, True Positives occur when children with ASD are correctly diagnosed, while False Positives happen when ASD are wrongly identified in normal children. True Negatives indicate accurate identification of normal children, whereas False Negatives occur when children with ASD are incorrectly labeled as normal. The mathematical expressions for accuracy, precision, recall F1-score, and ROC are as follows:

Accuracy: The accuracy of a classification model is a measure of performance that shows how well it works generally. The metric is determined by dividing the count of accurately predicted instances (True Positives along with True Negatives) by the overall count of instances in the dataset. The calculation for accuracy is:

$$\text{Accuracy} = \frac{TP + TN}{TP + FN + TN + FP} * 100 \quad (9)$$

Precision: Precision quantifies classification model accuracy in positive predictions. Precision is calculated by dividing the count of true positives by the sum of false positives and true positives. Here's the precision formula:

$$\text{Precision} = \frac{TP}{TP + FP} \quad (10)$$

Recall: Recall is a performance statistic that assesses how well a classification model captures all pertinent instances of a positive class. Some other terms for recall include sensitivity and true positive rate. It is calculated as True Positives divided by the sum of True Positives along with False Negatives. Recall formula is as follows:

$$\text{Recall} = \frac{TP}{TP + FN} \tag{11}$$

F1-Score: The F1 score, sometimes referred to as the F1 measure, combines precision and recall into a single result in a confusion matrix, making it a valuable performance metric. It is especially beneficial when deciding whether precision or recall is more important. The F1 score is computed employing the following equation:

$$\text{F1} - \text{Score} = \frac{2TP}{2TP + FP + FN} * 100 \tag{12}$$

ROC: The Receiver Operating Characteristic curve, or ROC, is a graphical plot that shows how well a classification model works at different levels for classification problems. The ROC curve is made by plotting the rate of true positives against the rate of false positives at each baseline level.

The formula for calculating the true positive rate (TPR) is:

$$\text{True Positive Rate (TPR)} = \frac{TP}{TP + FN} \tag{13}$$

The false positive rate (FPR) is calculated using the following formula:

$$\text{False Positive Rate (FPR)} = \frac{FP}{FP + TN} \tag{14}$$

4. Experimental Results and Discussion

Early detection of autism is crucial for safeguarding the well-being of numerous children. Creating an intelligent AI-based system can play a significant role in detecting autism at an earlier stage. In this study, we applied five algorithms classifications to differentiate between autistic as well as non-autistic children using the dataset focused on autistic children. Several studies were conducted to evaluate the proposed approach for identifying autism spectrum disorder (ASD). The study used 80% of the data for training as well as validation, with 20% left over for testing. The pretrained models were trained using a standard set of hyperparameters. During training, the learning rate was set to 1e-4with batch sizes of 32 and 50 epochs respectively. Every model was trained and then evaluated in order to evaluate how well it performed in the detection of Autism Spectrum Disorder. The experimental findings were focused on the following areas:

- Detailed discussion of the results of the proposed DenseNet 121 Model.
- Comparison of results with other deep learning models.
- Comparison of the proposed method's accuracy performance with SOTA Methods.

4.1. Performance Evaluation of Proposed DenseNet121 Model

The intention of this research is to present the DenseNet 121 model to facilitate early detection of autism spectrum disorder (ASD). As previously mentioned, all of the pre-trained models in the study as well as our suggested DenseNet-121 model were used in the training process. Following training, testing and validation data were utilized to assess the proposed DenseNet 121 model's performance. Figure 10 presents the training along with validation results. The loss curves with training epochs are shown in Figure 11. The graph for the suggested model shows that, with the specified hyperparameters, the accuracy of the training and validation sets increased progressively over a shorter time. By analyzing the graph, we can see that after the second epoch, the training accuracy grew steadily to reach 78.16%, having started at just over 73.27%. Nevertheless, there was a minor drop in training accuracy in the sixth period. After the third epoch, the validation accuracy increased to 82.55% from a starting point of 77.23%. By the end of the training period, the model's training accuracy has increased steadily to 99%. However, the validation accuracy likewise grows steadily, beginning at 77.23% and ending at 94.47% in the last era. This shows that the model generalizes well to previously unseen validation data in addition to learning from the training set of data well.

Figure 10: Proposed Model Training as well as Validation Accuracy curve graph

Figure 11 displays the training as well as validation losses. The graph analysis reveals that the training error significantly decreased throughout the course of epochs, with the training loss starting at 63% and reaching 0.02% at the 50th epoch. As an additional indication that the model is successfully minimizing error on both training as well as validation sets, the validation loss drops from 55% to 18%. During training, the custom parameters introduced to the top of the model were adjusted using the RMSprop optimizer to reduce the inaccuracy in the model's output. Several tactics were employed to prevent the deep learning models from becoming overfit. The convergence of training and validation metrics implies a well-trained and well-generalizing model, demonstrating that the model converged sufficiently to identify ASD using early childhood facial image data.

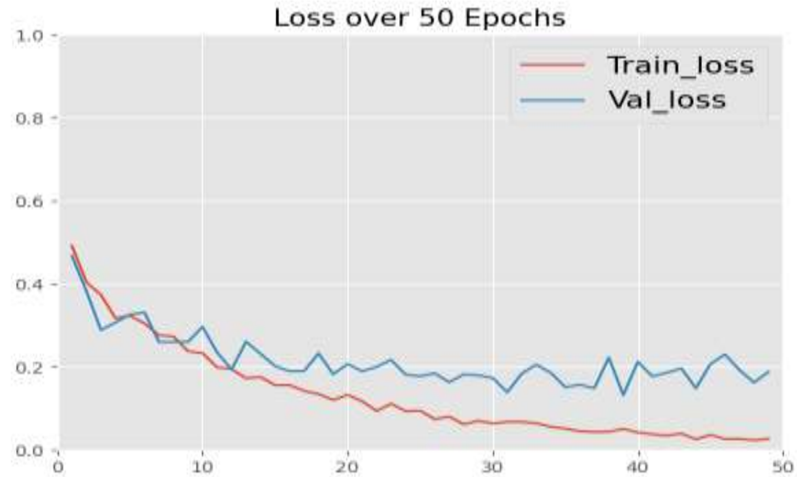

Figure 11: Proposed Model Training as well as Validation Loss curve graph

Table 6 also shows an analysis of how well the suggested DenseNet-121 model did on a publicly available dataset of autism that is split into groups for people with and without autism. The model was assessed pertaining to recall, accuracy, precision, and f1 score. The overall accuracy of our suggested model was 96% for each of the two classes. The accuracy, recall, and f1 score of our autistic class were 99%, 93%, and 96%, respectively. In contrast, the non-autistic class achieved 93%, 99%, and 96% in precision, recall, and f1 score, accordingly.

Table 6: The suggested mode's performance evaluation

Class	Precision(%)	Recall(%)	F1 –Score (%)	Accuracy (%)
Autistic	99	93	96	96%
Non Autistic	93	99	96	

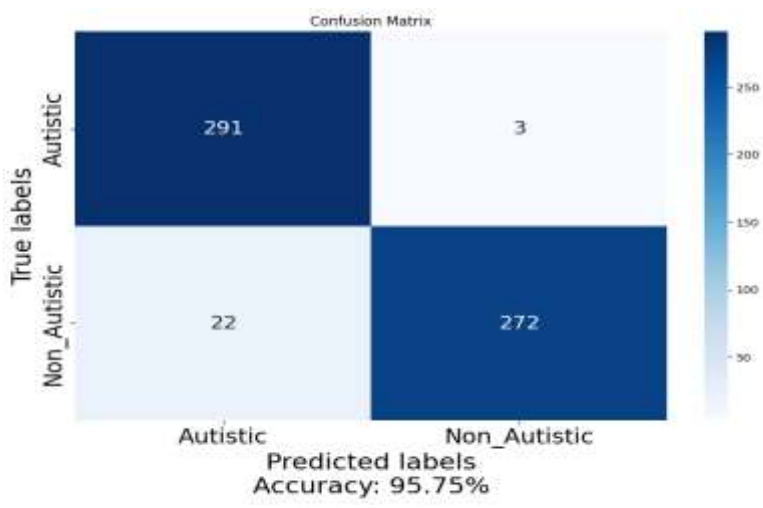

Figure 12: Confusion matrix of the proposed DenseNet121 Model

A confusion matrix is an effective way to assess how well deep learning models perform, especially when it comes to classification problems. The model's predictions can be analyzed to find trends, including which classes and to what extent they are misclassified. This can be helpful in determining areas in which the model requires improvement and in evaluating how well various models perform in comparison. Similarly, as shown in Figure 13, a confusion matrix was used to visually measure a model's performance on testing. The confusion matrix that is displayed shows correct guesses off-diagonally in grey and right predictions diagonally in blue. The X-axis displays the real labels in blue, while the Y-axis displays the anticipated labels in the same blue colour. The DenseNet 121 model correctly predicted 291 images, which are coloured blue in the confusion matrix, to be in the autistic class.

The data suggests that on 292 times, the model's predictions matched the actual data for the autistic class. On the other hand, three pictures were mistakenly labelled as not belonging to the autism class by the DenseNet121 model. This is categorized as a genuine negative and is known as misclassification. The model correctly predicted 272 samples in the non-autistic class, which is categorized as a false negative. Furthermore, just 22 pictures—also known as false positives—were misclassified (Figure 11). The DenseNet121 model produced good accuracy results, based on the evaluation measures.

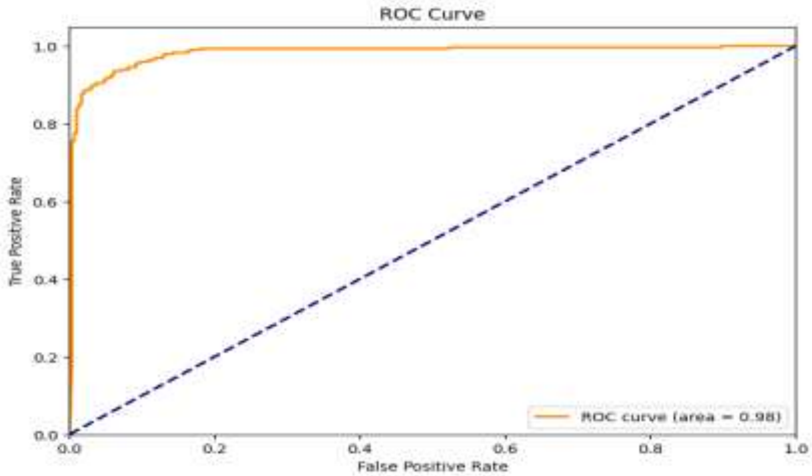

Figure 13: ROC plot for the proposed DenseNet 121 model with the value under curve value

Figure 12 above displays the suggested model's receiver operating characteristics, or ROC. This curve graph's key advantage is that it visualizes model performance, demonstrating the tradeoff between true positive & false positive rates. Because our suggested model was subjected to binary classification, the AUC (area under the curve) values of two distinct classes are displayed in the ROC curve. The AUC for the autistic as well as non-autistic groups was 0.98, respectively.

4.2. The result comparison with other deep learning models.

The study compared the performances and efficiencies of five deep learning pre trained models: ResNet-50, Xception, MobileNetv2, VGG16 and DenseNet-121.Each deep learning model in each study employed the same set of parameters. Further, tables 7, assessing the efficiency of various pre-trained DL (Deep Learning) models with respect of accuracy, recall, precision, as well as f1-score respectively. The accuracy of a classification model is a measure of performance that shows how well it works generally. Precision evaluates the correctness of positive expectations provided by models, with high values indicating fewer false positive inaccuracies. Meanwhile, recall assesses the capacity of models to cover all the favorable illustrations. The f1-score serves as a consolidated metric that makes up balance between precision and recall, offering a comprehensive evaluation of a model's performance.

Table 7: Comparison of the proposed model's performance with other deep learning models

Model	Precision(%)	Recall (%)	F1 Score (%)	Accuracy (%)
ResNet50	83	82	82	82
Xception	92	92	92	92
MobileNetV2	85	85	85	85
VGG16	94	94	94	94
DenseNet-121	96	96	96	96

Table 7 provides a clear presentation of the statistical analysis for various DL models regrading to precision, recall, and f1-score and accuracy. Though looking at the precision values, the precision of the DenseNet-121 model is 96% which is 13%, 4%, 11% and 2% higher, respectively than that of the ResNet50, Xception, MobileNetv2 and VGG16. The recall as well as f1-score of the DenseNet 121 is 96% which is 14%, 4%, 11% and 2% higher, respectively than that of the ResNet50, Xception, MobileNetv2 and VGG16. The proposed DenseNet121 model's prediction accuracy is 96% which is 14%, 4%, 11% and 2% higher, respectively, than that of the ResNet50, Xception, MobileNetv2 and VGG16. According to the DenseNet-121 model, 96% of children with positive autism traits are correctly identified as autistic. Through the above experimental investigation, we can approximate that DenseNet-121 has out-performed the various other deep learning models in all valuation parameters in classifying autistic children based on facial images.

4.3. Evaluation of the suggested method's accuracy performance in comparison to SOTA Methods

Despite the fact that a variety of deep learning algorithms have been developed for ASD classification, as far as our knowledge extends, the prior studies in the literature review used a transfer learning strategy to detect ASD using facial photos and the same dataset. Consequently, as our study used the same dataset and the same transfer learning model with the same hyperparameters, our comparison is restricted to it. However, we have also included other deep learning experiments with excellent performances to critically assess the performance of our suggested model. Findings of the proposed system in comparison to the current system are summarized in Table 8. We conducted a literature review to determine the optimal level of accuracy for this dataset. (*Autism*

Dataset,) , discovered ten research investigations that employed it in their experiments. The initial three studies were carried out by (Rashid & Shaker, 2023), (Alsaade & Alzahrani, 2022) ,and (Sridurga, P.D., et al.,) , whom tried to classify autistic and non-autistic using the Xception model, achieved 91 %, 91%, and 86% accuracy on the validation data respectively. The fourth investigation, carried out by (Rabbi et al., 2021) in 2021**,** used the CNN model, achieved 92.31% accuracy on the validation data. The fifth to seventh study, which was performed by (Alkahtani et al., 2023)**,** (Akter et al., 2021) ,and (Singh et al., 2023) , used the MobileNet model to detect autism facial images and achieved 92%,90.67% and 88% validation accuracy respectively. An eighth study that was done by (Gaddala et al., 2023) in 2023, which implemented the VGG16/19 model on dataset consisting of faces of children, achieved 84% accuracy. The last study that was done by (M. Ghazal et al., 2023) in 2023, who used the AlexNet model on facial images of children, achieved 87.7% accuracy. They all used the same dataset for experiments. Our proposed model outperformed other reviewed methodologies, achieving an accuracy of 96 % with a minimal 4% Misclassification Rate (MCR). This performance surpasses all other existing studies, showcasing the superiority of our expert system. Notably, our approach demonstrated a 7.69% improvement in accuracy compared to recent studies, attributed to its ability to handle the complexity of the dataset.
Detecting ASD at a younger age through facial image analysis could have profound effects, not only on the child, but on the parents and the doctor engaged. An early diagnosis offers the advantage of a quick assessment by the physician, who can rapidly determine whether the child is on the autism spectrum or developing typically based on facial images. Early ASD diagnoses offer a distinct advantage, yet the process turns into more demanding when a professional expert manually assesses a child's facial traits to diagnose autism or typical development through visual interpretation. Enhancing the model's training with a larger dataset improves its accuracy. If integrated into a mobile application, parents could independently screen their children, streamlining the preparation for referrals or diagnostic testing.

Table 8: A Comparison of the newly proposed model with the current advanced approaches

Ref.	Year	Techniques	Dataset	Accuracy
(Rashid & Shaker, 2023)	2023	Xception	2,940	91%
(Alsaade & Alzahrani, 2022)	2022	Xception	2940	91%
(Sridurga, P.D., et al.,2022)	2022	Xception	2940	86%
(Rabbi et al., 2021)	2021	CNN	2940	92.31%
(Alkahtani et al., 2023)	2023	MobileNetV2	2940	92%
(Akter et al., 2021)	2021	MobileNet-V1	2936	90.67%
(Gaddala et al., 2023)	2023	VGG16 & 19	2936	84%
(Singh et al., 2023)	2023	MobileNet	2940	88 %
(M. Ghazal et al., 2023)	2023	AlexNet	2940	87.7%
Proposed Model	2024	DenseNet121	2940	96%

5. Conclusion and Future work

This paper presents a novel methodology employing Convolutional Neural Networks (CNNs) and transfer learning for the binary classification of Autism Spectrum Disorder (ASD) patients based on facial images. The study utilizes five pre-trained CNN models, implementing transfer learning techniques on a publicly available Kaggle dataset. The primary objective was to enhance the precision and efficiency of detecting autism in children at an early stage. Various CNN models, including Xception, ResNet-50, MobileNetv2, VGG-16, and DenseNet121, were employed for image detection and classification. The training of these CNN models utilized the RMSprop optimizer, and the training process extended over 50 epochs. The batch size is 32. The results show that our model performed remarkably well on the test data, with an accuracy of 96%, surpassing the performance of the DenseNet-121 model, which previously held the record for the highest accuracy on this dataset. These results highlight the potential of leveraging deep learning-based models as efficient and precise tools for both specialists and families in the swift and accurate diagnosis of autism. In the future, our next efforts will focus on improving autism identification through the use of pre-trained models in the Internet of Medical Things (IoMT). By leveraging cutting-edge technology, we hope to contribute to ongoing research in this crucial area, ultimately improving early detection as well as intervention options for persons with autism spectrum disorders.

References

69+ Autism Statistics: How Many People Have Autism? (n.d.). Retrieved February 6, 2024, from https://www.crossrivertherapy.com/autism-statistics

Ahmad, I., Rashid, J., Faheem, M., Akram, A., Khan, N. A., & Amin, R. U. (2024). Autism spectrum disorder detection using facial images: A performance comparison of pretrained convolutional neural networks. *Healthcare Technology Letters*, htl2.12073.

Ahmed, Z. A. T., Albalawi, E., Aldhyani, T. H. H., Jadhav, M. E., Janrao, P., & Obeidat, M. R. M. (2023). Applying Eye Tracking with Deep Learning Techniques for Early-Stage Detection of Autism Spectrum Disorders. *Data*, *8*(11), 168.

Akram, A., Khan, I., Rashid, J., Saddique, M., Idrees, M., Ghadi, Y. Y., & Algarni, A. (2024). Enhanced Steganalysis for Color Images Using Curvelet Features and Support Vector Machine. *Computers, Materials & Continua*, *78*(1), 1311–1328.

Akram, A., Ramzan, S., Rasool, A., Jaffar, A., Furqan, U., & Javed, W. (2022). Image splicing detection using discriminative robust local binary pattern and support vector machine. *World Journal of Engineering*, *19*(4), 459–466.

Akram, A., Rashid, J., Hajjej, F., Yaqoob, S., Hamid, M., Arshad, A., & Sarwar, N. (2023). Recognizing Breast Cancer Using Edge-Weighted Texture Features of Histopathology Images. *Computers, Materials & Continua*, *77*(1), 1081–1101.

Akram, A., Rashid, J., Jaffar, A., Hajjej, F., Iqbal, W., & Sarwar, N. (2024). Weber Law Based Approach for Multi-Class Image Forgery Detection. *Computers, Materials & Continua*, *78*(1), 145–166.

Akram, A., Rashid, J., Jaffar, M. A., Faheem, M., & Amin, R. U. (2023). Segmentation and classification of skin lesions using hybrid deep learning method in the Internet of Medical Things. *Skin Research and Technology*, *29*(11), e13524.

Akter, T., Ali, M. H., Khan, Md. I., Satu, Md. S., Uddin, Md. J., Alyami, S. A., Ali, S., Azad, A., & Moni, M. A. (2021). Improved Transfer-Learning-Based Facial Recognition Framework to Detect Autistic Children at an Early Stage. *Brain Sciences*, *11*(6), 734.

Aldhyani, T. H. H., Nair, R., Alzain, E., Alkahtani, H., & Koundal, D. (2022). Deep Learning Model for the Detection of Real Time Breast Cancer Images Using Improved Dilation-Based Method. *Diagnostics*, *12*(10), 2505.

Alkahtani, H., Aldhyani, T. H. H., & Alzahrani, M. Y. (2023). Deep Learning Algorithms to Identify Autism Spectrum Disorder in Children-Based Facial Landmarks. *Applied Sciences*, *13*(8), 4855.

Alsaade, F. W., & Alzahrani, M. S. (2022). Classification and Detection of Autism Spectrum Disorder Based on Deep Learning Algorithms. *Computational Intelligence and Neuroscience*, *2022*, 1–10.

Autism dataset. *Accessed 8th December*, 2023, from https://www.kaggle.com/datasets/cihan063/autism-image-data

Carette, R., Cilia, F., Dequen, G., Bosche, J., Guerin, J.-L., & Vandromme, L. (2018). Automatic Autism Spectrum Disorder Detection Thanks to Eye-Tracking and Neural Network-Based Approach. In M. U. Ahmed, S. Begum, & J.-B. Fasquel (Eds.), *Internet of Things (IoT) Technologies for HealthCare* (Vol. 225, pp. 75–81). Springer International Publishing.

Duranta, D. U. S., Ali, M. S., Das, A. A., Rahman, M. M., Ahsan, M. M., Miah, M. S., & Islam, M. K. (2023). Enhancing Atrial Fibrillation detection accuracy: A wavelet transform filtered single lead ECG signal analysis with artificial neural networks and novel feature extraction. *Machine Learning with Applications*, *12*, 100472.

Elbattah, M., & Cilia, F. (2023). *Applications of Machine Learning Methods to Assist the Diagnosis of Autism Spectrum Disorder* [Preprint]. Open Science Framework.

Francés, L., Caules, J., Ruiz, A., Soler, C. V., Hervás, A., Fernández, A., Rodríguez-Quiroga, A., & Quintero, J. (2023). An approach for prevention planning based on the prevalence and comorbidity of neurodevelopmental disorders in 6-year-old children receiving primary care consultations on the island of Menorca. *BMC Pediatrics*, *23*(1), 32.

Gaddala, L. K., Kodepogu, K. R., Surekha, Y., Tejaswi, M., Ameesha, K., Kollapalli, L. S., Kotha, S. K., & Manjeti, V. B. (2023). Autism Spectrum Disorder Detection Using Facial Images and Deep Convolutional Neural Networks. *Revue d'Intelligence Artificielle*, *37*(03), 801–806.

Ghosh, T., Banna, Md. H. A., Rahman, Md. S., Kaiser, M. S., Mahmud, M., Hosen, A. S. M. S., & Cho, G. H. (2021). Artificial intelligence and internet of things in screening and management of autism spectrum disorder. *Sustainable Cities and Society*, *74*, 103189.

Goh, K. L., Morris, S., Rosalie, S., Foster, C., Falkmer, T., & Tan, T. (2016). Typically developed adults and adults with autism spectrum disorder classification using centre of pressure measurements. *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 844–848.

Iqbal, S., & Khan, W. (2017). Exploring the Nexus of Service Quality, University Image, and Word-of-Mouth Communication: A Comparative Study of Public and Private Universities. *Journal of Policy Options*, *4*(3), 60-69.

Kanhirakadavath, M. R., & Chandran, M. S. M. (2022). Investigation of Eye-Tracking Scan Path as a Biomarker for Autism Screening Using Machine Learning Algorithms. *Diagnostics*, *12*(2), 518.

Kanner, L. (1968). Autistic disturbances of affective contact. *Acta Paedopsychiatrica*, *35*(4), 100–136.

Kaur, N., & Gupta, G. (n.d.). *Autism Disorder Detection Based on Deep Learning: Deep Analysis*.

Li, J., Zhong, Y., Han, J., Ouyang, G., Li, X., & Liu, H. (2020). Classifying ASD children with LSTM based on raw videos. *Neurocomputing*, *390*, 226–238.

Li, M., Tang, D., Zeng, J., Zhou, T., Zhu, H., Chen, B., & Zou, X. (2019). An automated assessment framework for atypical prosody and stereotyped idiosyncratic phrases related to autism spectrum disorder. *Computer Speech & Language*, *56*, 80–94.

M. Ghazal, T., Munir, S., Abbas, S., Athar, A., Alrababah, H., & Adnan Khan, M. (2023). Early Detection of Autism in Children Using Transfer Learning. *Intelligent Automation & Soft Computing*, *36*(1), 11–22.

Modak, D. S. (n.d.). *Early detection of autism spectrum disorder (ASD) using Machine Learning techniques*.

Mohanta, A., & Mittal, V. K. (2022). Analysis and classification of speech sounds of children with autism spectrum disorder using acoustic features. *Computer Speech & Language*, *72*, 101287.

Mujeeb Rahman, K. K., & Subashini, M. M. (2022). Identification of Autism in Children Using Static Facial Features and Deep Neural Networks. *Brain Sciences*, *12*(1), 94.

Pokorny, F. B., Schuller, B., Marschik, P. B., Brueckner, R., Nyström, P., Cummins, N., Bölte, S., Einspieler, C., & Falck-Ytter, T. (2017). Earlier Identification of Children with Autism Spectrum Disorder: An Automatic Vocalisation-Based Approach. *Interspeech 2017*, 309–313.

Rabbi, Md. F., Hasan, S. M. M., Champa, A. I., & Zaman, Md. A. (2021). A Convolutional Neural Network Model for Early-Stage Detection of Autism Spectrum Disorder. *2021 International Conference on Information and Communication Technology for Sustainable Development (ICICT4SD)*, 110–114.

Rahman, Md. M., Usman, O. L., Muniyandi, R. C., Sahran, S., Mohamed, S., & Razak, R. A. (2020). A Review of Machine Learning Methods of Feature Selection and Classification for Autism Spectrum Disorder. *Brain Sciences*, *10*(12), 949.

Rashid, A., & Shaker, S. (2023). Autism spectrum Disorder detection Using Face Features based on Deep Neural network. *Wasit Journal of Computer and Mathematics Science*, *2*(1), 74–83.

Rawat, W., & Wang, Z. (2017). Deep Convolutional Neural Networks for Image Classification: A Comprehensive Review. *Neural Computation*, *29*(9), 2352–2449.

Sadik, R., S. Anwar, and M. Reza,. (2021). AutismNet: Recognition of Autism Spectrum Disorder from Facial Expressions using MobileNet Architecture. *International Journal of Advanced Trends in Computer Science and Engineering*, *10*(1), 327–334.

Schelinski, S., Borowiak, K., & Von Kriegstein, K. (2016). Temporal voice areas exist in autism spectrum disorder but are dysfunctional for voice identity recognition. *Social Cognitive and Affective Neuroscience*, *11*(11), 1812–1822.

Singh, A., Laroia, M., Rawat, A., & Seeja, K. R. (2023). Facial Feature Analysis for Autism Detection Using Deep Learning. In A. E. Hassanien, O. Castillo, S. Anand, & A. Jaiswal (Eds.), *International Conference on Innovative Computing and Communications* (Vol. 703, pp. 539–551). Springer Nature Singapore.

Sridurga, P.D., et al.,. (n.d.). Detecting Autism Spectrum Syndrome using VGG19 and Xception Networks. *International Journal of Research in Engineering, Science and Management*, *5*(12, December 2022).

Talo, M., Baloglu, U. B., Yıldırım, Ö., & Rajendra Acharya, U. (2019). Application of deep transfer learning for automated brain abnormality classification using MR images. *Cognitive Systems Research*, *54*, 176–188.

Thanarajan, T., Alotaibi, Y., Rajendran, S., & Nagappan, K. (2023). Eye-Tracking Based Autism Spectrum Disorder Diagnosis Using Chaotic Butterfly Optimization with Deep Learning Model. *Computers, Materials & Continua*, *76*(2), 1995–2013.

Venkata Sai Krishna Narala, M. S., Vemuri, S., & Kattula, C. (2023). Prediction of Autism Spectrum Disorder Using Efficient Net. *2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS)*, 1139–1143.

Zhu, F., Wang, S., Liu, W., Zhu, H., Li, M., & Zou, X. (2023). A multimodal machine learning system in early screening for toddlers with autism spectrum disorders based on the response to name. *Frontiers in Psychiatry*, *14*, 1039293.