DOI: 10.15849/IJASCA.240730.15

Int. J. Advance Soft Compu. Appl, Vol. 16, No. 2, July 2024
Print ISSN: 2710-1274, Online ISSN: 2074-8523

Utilizing Deep Learning Models in an Intelligent Facial Expression Classification System for Autism Disorder Diagnosis

Nesren Farhah

Department of Health Informatics, College of Health sciences, Saudi Electronic University, Riyadh 11673, Saudi Arabia

Abstract

Autism spectrum disorder (ASD) is a medical illness that markedly impacts the neurological function. An Individual with autism experiences challenges in acknowledging their names actively avoids establishing eye contact and demonstrates a limited capacity to express emotions. Biomedical image categorization is an important area of study that is increasingly popular among both researchers and clinicians for the purpose of detecting and diagnosing ASD disorders. The human face can function as a biological indicator since it has the ability to reflect the condition of the brain. Hence, it may be employed as a convenient and uncomplicated instrument for the early detection of ASDs. This research uses a variety of advanced techniques to detect individuals who have autism by analyzing their facial images. A range of deep learning (DL) models, including VGG16, InceptionV3, and EfficientnetB0, is used to identify autistic children by utilizing facial expressions detection. The research is carried out to empirically determine the optimal settings for the model optimizer and different hyperparameters in DL models with the goal of improving prediction accuracy. A conventional research dataset sourced from Kaggle comprising 2940 images of children analyzed as both having and not having autism is used to assess the DL models. The improved VGG16 model attained a 95% accuracy rate on the test dataset. The findings of the present study suggest that the improved VGG16 model outperforms the models created by prior studies. and can aid clinicians in validating the precision of their first assessment of ASD in pediatric patients.

Keywords: ASD, transferring learning, enhanced VGG16, diagnosis

1 Introduction

Autism spectrum disorder (ASD) is a complex neurological illness that disturbs speech, social communication, behavior, and cognitive development. ASD displays a wide range of indications, which is why it is referred to and described as a “spectrum” condition. People with ASD exhibit a range of impairments, which poses difficulties in accurately diagnosing and successfully treating the condition [1]. The precise cause of ASD is not completely understood; however, research proposes that a mixture of heritable and environmental aspects has a significant impact. Numerous genes have been recognized as possible contributions to the growth of ASD [2]. It is hypothesized

Received 20 April 2024; Accepted 8 July 2024

that genetic changes in aspects of these genes influence the expansion and functioning of the brain, resulting in the distinctive symptoms of ASD. Environmental variables have an influence on the likelihood that individuals will develop ASD both before and after birth. These variables encompass maternal illnesses during pregnancy, exposure to specific chemicals, and problems following childbirth [3]. Brain imaging studies have shown differences in brain structure and functioning in individuals with ASD. These disparities influence the way individuals handle and combine information, resulting in noticeable behavioral trends [4]. Presently, there is no definitive way to manage ASD. However, the early adoption of interventions and personalized therapy approaches significantly improve the general well-being of those affected by the condition. Practical behavior analysis (PBA) is a commonly used interactive behavioral intervention for persons with ASD.

ASD is a multifaceted condition that disrupts a person's everyday communication [5]. An individual with autism often encounters mild impairments and can occasionally need specialized assistance. Individuals with ASD commonly experience difficulties in communicating, resulting in their inability to effectively convey their thoughts, feelings, and intentions using spoken language, nonverbal gestures, or facial expressions during social interactions. While medical professionals frequently diagnose individuals with ASD by observing the neurophysiological indications associated with the disorder, at present, there is no ultimate physical signature or pathological method that can consistently recognize autism. Although an individual may not receive adequate therapy, an early diagnosis might offer some means to enhance their lifestyle [6]. The early detection of ASD symptoms might enhance the social lives of children by taking advantage of the malleability of brain development. Research also indicates that children who received interventions before the age of two obtained higher IQ scores than those who received medical assistance after the age of four [7]. According to a recent study, only a maximum of 30% of children with ASD are recognized after the age of three [8]. ASD is a neural illness that impacts many regions of the brain. It is a product of polymorphism, which refers to the genetic effect induced by the interplay of human genes [9]. According to research presented by the World Health Organization (WHO), around one in every 100 children has ASD. According to the Centers for Disease Control and Prevention (CDC), in 2021, the incidence of ASD was greatest in the USA, where around 1 in 44 children was affected. Furthermore, there are four times as many males as young male with ASD [10].

Numerous studies have examined the key attributes of autism from several perspectives, including the use of eye-tracking (ET) methods for facial-expression extractions [11], face recognition [12,13,14], biomedical image analysis [15], application development [16], and speech recognition [17]. Face recognition is a highly valuable approach for identifying a person's emotive condition and can provide accurate diagnoses of autism. Facial analysis is a widely used technique for examining human faces and identifying differences among abnormal and abnormal faces. It can be utilized to uncover important information that may help identify behavioral patterns [18,19]. Given the recent advancements in predictive analytics for recognizing face patterns, considerable attention is currently being paid within autism research to investigating and examining data on children with ASD to identify the condition at a younger age. To reorganize the process of identifying facial expressions based on features in various neural disorders, the authors in [20] implemented the convolutional neural network (CNN) approach to diagnose ASD. The initial development of CNN was trained to segment crucial facial features, while the second iteration was employed to identify and classify facial expressions. In 2018, Haque and Valles partitioned the system into four representations, which successfully recognized human facial expressions [21,22]. To do this, they made alterations to Kaggle's Facial Expression Recognition 2013 (FER2013) dataset, specifically aiming at young children with autism. These alterations involved manipulating the lighting effects by adjusting the contrast to create darker or brighter tones. Figure 1 provides a comprehensive summary of the many diagnostic techniques used for autism.

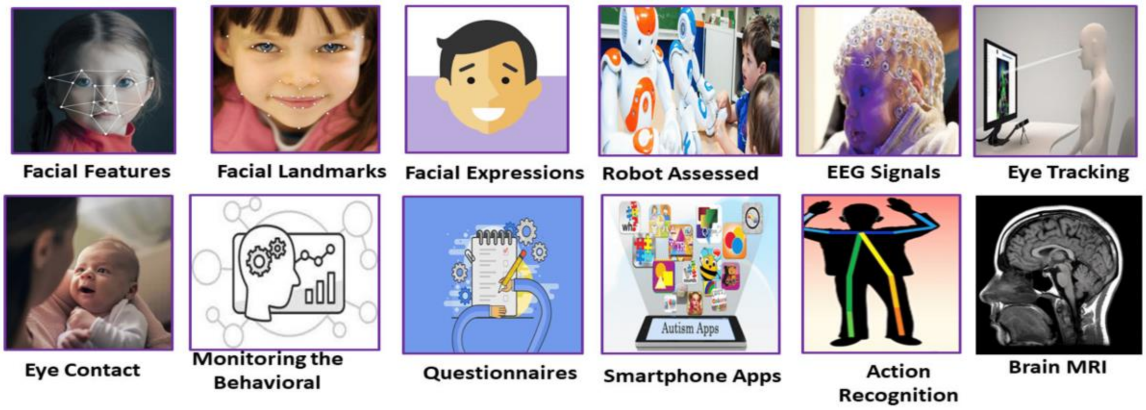


Figure 1. Various technologies for autism detection.

1.1 Contribution

The key objective of this project is to improve the diagnosis of autism by using DL algorithms to establish an autism prediction model with the maximum precision. The main aim is to develop a highly precise prediction model capable of determining if an individual, regardless of age (adolescent, child, or adult), has or does not have ASD. The objective is to utilize a standardized methodology to diagnose autism and transform it into DL algorithms capable of utilizing medical data to provide predictions and observations. This will ultimately result in improved methods for the early identification of ASD in the future.

2 Related Work

The field of study focused on diagnosing ASD based on facial traits is experiencing remarkable growth, mostly due to its significant socioeconomic implications for emerging nations. This approach has the potential to be a significant advancement in the early analysis of ASD, serving as a crucial instrument for the first screening of both children with ASD and typically developing (TD) children. Recent research has shown the capability of deep neural networks, namely, the use of CNN models, in diagnosing various diseases [23,24,25,26,27,28,29,30,31,32]. CNNs are extensively employed as feature extractors for research problems including object identification and image segmentation and categorization due to their exceptional capacity to learn by autonomously identifying concealed characteristics from a substantial number of pictures. While CNNs are very efficient and precise, the process of training these models demands a substantial investment of time and computer resources.

Akter et al. [33] utilized DL models to detect faces with ASD using a 2D-photo dataset obtained through the Kaggle platform. The researchers examined both narrow learning and DL techniques for identifying autism in children aged 2–14 and attained the maximum level of precision by using an enhanced version of the MobileNet-V1 model. In their study, Hosseini et al. [34] employed the MobileNet model to greatly enhance the accuracy of autism detection. The visual characteristics were collected from DL models, which utilized three completely connected layers followed by a custom dense layer to provide predictions for ASD. Nevertheless, to obtain more precision, the authors excluded images depicting young infants in the datasets. Consequently, they successfully decreased the occurrence of incorrect positive and negative results, resulting in an overall accuracy of around 95%. In a subsequent study, the authors in [35] utilized the same dataset and directed their investigation toward advanced DL models such as Xception and EfficientNetB, with a specific focus on the area under the curve (AUC) evaluation metric. Khosla et al. [36] employed the InceptionResNetV2, InceptionV3, and MobileNet methods for detecting ASD and observed that these models achieved good accuracy compared to previous studies. Alsaade and Alzahrani [37] utilized DL models to train on an ASD dataset. The Xception model attained the greatest accuracy level, namely, of 91%. CNN-based models utilized for feature mining from an autistic image dataset have undergone thorough training on the ImageNet dataset [38].

Khalaji et al. [39] projected a pre-processing method that is not influenced by external stimuli, resulting in high classification accuracy. Michelass et al. [40] introduced a new technology based on ET signs from films to evaluate combined considerations and categorize people as either having ASD or being TD. The main advancement was the implementation of "floating Regions of Interest," which monitor the eye's movement in connection to the meaning of an item. The model used an integration approach consisting of random forest classifiers to categorize people as having ASD or being TD, based on the trajectory characteristics retrieved from the ET signals. The authors in [41] developed a model for the identification of interactive autism by employing eye-movement analysis. The use of a sequential neural network in the ET visualization resulted in a noteworthy accuracy of 95.7% and an AUC of 84%. Cilia et al. [42] planned a method that combines ET with picturing and machine learning (ML). In this method, the ET examination paths were transmuted into pictorial representations in the form of images, after which a CNN model was trained to precisely classify these photographs. The findings suggest that employing visual representation facilitated the diagnostic process and achieved a notable degree of precision. The CNN attained an accuracy rate of 90%, a sensitivity rate of 83%, and a precision rate of 80%.

Gaspar et al. [43] presented a method for categorizing ASD using a kernel extreme learning machine (KELM). The training of the model was enhanced by employing data augmentation, and the accuracy of the KELM model was improved by utilizing the Giza pyramids construction (GPC) approach. The KELM achieved an accuracy rate of 95.8%. Zhong et al. [44] used four ML classifiers to distinguish between children with ASD and TD children by analyzing ET data. The authors employed forward feature selection to identify relevant attributes and fed them into ML classifiers. The support vector machine (SVM) attained the highest degree of accuracy, achieving 92.31%. The authors in [45] introduced a transfer learning technique for identifying ASD in people with strong cognitive abilities. The researchers utilized decision tree (TD), transfer learning, and logistic regression techniques on a dataset consisting of high-functioning ASD individuals and control patients. The findings exhibited a classification accuracy of 80.50%, indicating a high level of accuracy. Sun et al. [46] introduced a model that uses ET techniques to evaluate individually developing children and those with ASD at the same time. The program specifically targets stimuli related to limited interests. ASD was identified through network-based ML prediction (NBS-predict).

Omar et al. [47] used various TD approaches based on the AQ-10 and 250 real-world datasets for detecting ASD. The researcher in [48] proposed various ML approaches, such as KNN, SVM, and DT, for the diagnosis of ASD. Satu et al. [49] utilized several tree-based classifiers to analyze samples of individuals aged 16–30 years and determine the features distinguishing a “normal” from an “autistic” mature individual. Erkan et al. [50] applied KNN, SVM, and Random Forest (RF) to assess the comparative effectiveness of different approaches in identifying ASD. Thabtah et al. [51] employed the information gain (IG) method to generate reduced sets of features for both adults and adolescents. These smaller groups were subsequently utilized as inputs for logistic regression to detect ASD.

3 Materials and Methods

This subsection presents in depth the planned methodology to be used in developing an ASD classification system based on DL techniques than can diagnose autism in children from facial expression images. This section presents the dataset collection, data preprocessing, DL classification models, output evaluation metrics, and results analysis. The framework of this methodology is shown in Figure 2.

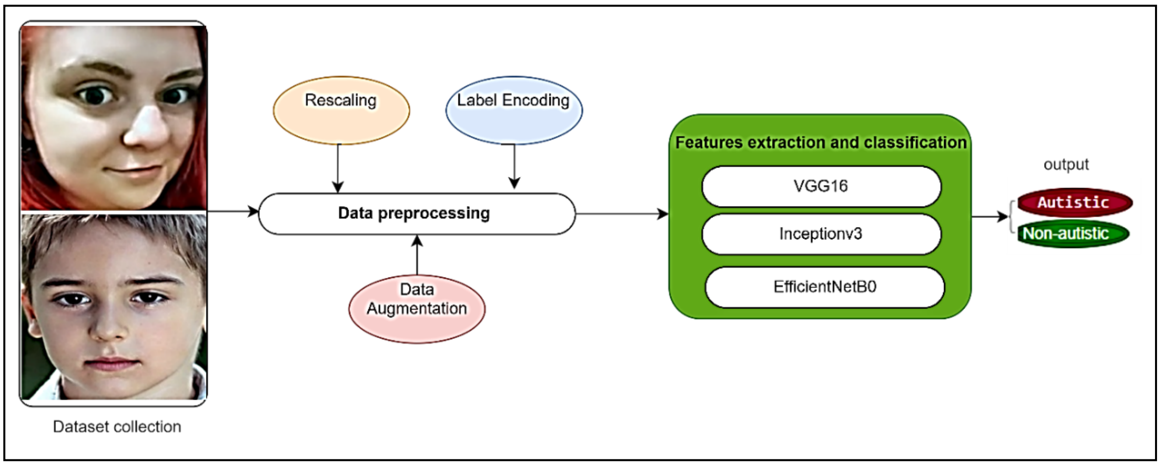


Figure 2: The framework of the proposed methodology

3.1 Dataset Collection

This research utilized a dataset of autistic children’s facial images from the Kaggle platform, an open-source resource for the research community [52]. The dataset included children aged 2–14 years, predominantly between 2 and 8 years old, with all images being 2D Red, Grean, Blue. The data were divided into two categories: the autistic category, which consisted of images of children diagnosed with autism, and the non-autistic category, which consisted of images of children without an autism diagnosis. Additionally, a test folder was used to evaluate the trained model, containing two subfolders labeled “autistic” and “non-autistic,” each with 100 images 224 × 224 × 3 in pixels. One subfolder comprised facial images of children with autism, and the other consisted of facial images of children collected arbitrarily from online explorations. In total, the dataset comprised 2,940 images, with 1,327 images of autistic children and 1,613 images of non-autistic children.

3.2 Data Preprocessing

For the autism recognition model utilizing facial expression features, we performed the following data preprocessing steps on the dataset.

Rescaling: All images were rescaled and normalized to a range of [0, 1] by applying a scaling factor of 1/255.

Data augmentation: We implemented different data resampling and augmentation methods to improve the dataset and avoid the overfitting problem: images were randomly rotated within a range of 40 degrees, parallel loosened to 20% of their width, and vertically transformed up to 20% of their height. Additionally, shear transformations up to 10% were applied, images were randomly flipped horizontally, and zooming was performed within a range of 20%. Pixels outside the boundaries of the input were filled using the "nearest" mode. Images were loaded from directories using ImageDataGenerator, targeting an image size of 224 x 224 pixels. The batch size for training was set to 128.

Label encoding: The binary class mode was used for classification, with classes labeled "non_autistic" and "autistic."

3.3 Feature Extraction and Classification Models

In this phase, we used advanced DL models, namely, VGG16, InceptionV3, and EfficientNetB0, for feature mining and classification. These models were refined on our particular dataset after being pre-trained on sizable image datasets. With its tiny but deep architecture, VGG16 can efficiently capture intricate patterns. To preserve a variety of properties, InceptionV3 uses simultaneous convolution procedures with varying kernel sizes. For the best accuracy and efficiency, EfficientNetB0 balances network depth, width, and resolution. Individually, each model was adapted by adding exclusive classification layers to allow the differentiation of autistic from non-autistic facial expressions.

3.3.1 The VGG16 Model

The VGG16 model [53] is a deep CNN architecture recognized for its straightforwardness and efficiency in image recognition investigation. The model has 16 different layers, namely, 13 convolutional layers and three completely linked layers, which were pre-trained on the ImageNet dataset for high-performance feature extraction. We applied this model for autism detection and improved it to include custom classifier layers tailored for this specific task.

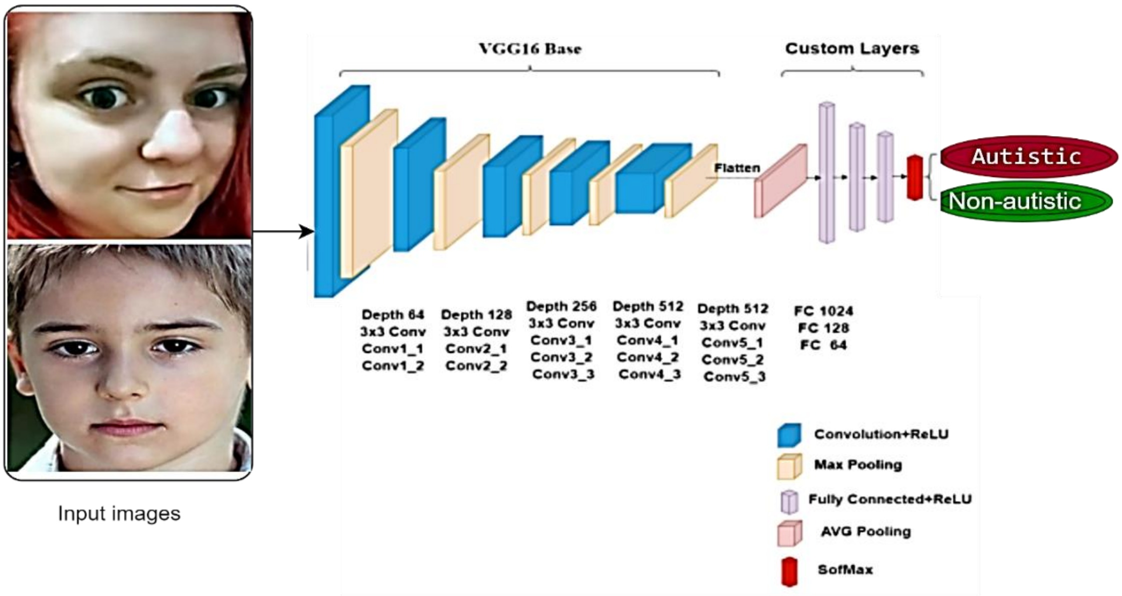


Figure 3: The structure of the VGG16 model

The base model is VGG16, pre-trained on ImageNet, with the top classification layers removed. To retain the valuable pre-trained weights, all layers except for the last few convolutional layers were frozen. The new classifier layers added to the model include a flatten layer, which converts the 2D feature maps to a 1D feature vector, followed by 256 hidden neurons presented in a dense layer as well as a ReLU, which is a nonlinear activation function. The nest layer is a fully connected layer including 95 neurons and a nonlinear softmax activation function for classification. The model was assembled with the help of the Adam optimizer, which was set a learning rate of 0.001, and the loss function was used to calculate the model errors and set to sparse categorical cross-entropy. The training process involved 100 epochs. Table 1 outlines the parameters used in the VGG16 architecture.

Table 1: Summarizing the VGG16 model parameters

Layer	**Parameters**
Base Model	VGG16
Input Shape	224 x 224 x 3
Frozen Layers	All except the last few convolutional layers
Flatten Layer	Yes
Dense Layer (ReLU)	256 units
Output Layer (Softmax)	95 units
Optimizer	Adam
Learning Rate	0.001
Loss Function	Sparse Categorical Cross-Entropy
Batch Size	128
Epochs	100

3.3.2 The Inceptionv3 Model

We also applied and modified the InceptionV3 architecture [54] for autism identification based on facial expression features to fit our particular needs. Due to its deep and efficient architecture, the InceptionV3 model is an excellent choice for capturing complex patterns in image data. To accept the custom layers specific to our research problem, the base model, InceptionV3, was pre-trained on the ImageNet dataset after the top classification layers were removed. To preserve already trained weights, all layers in the underlying InceptionV3 model were frozen. Figure 4 depicts the construction of the Inceptionv3 model.

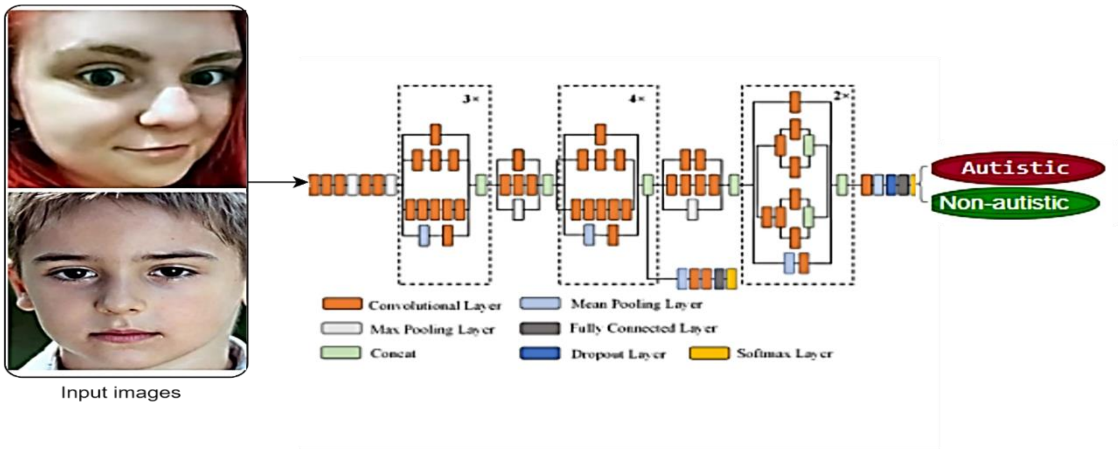


Figure 4: The architecture of the Inceptionv3 model

The newly added categorizer layers comprised a fully or completely connected dense layer with 128 units and a ReLU activation function, which was followed by a flatten layer that transformed the 2D feature maps into a 1D feature vector. A dropout layer, which can be used to avoid overfitting issues, was set to a rate value of 0.2, and the classification layer consisted of a fully connected dense layer with 64 units and a nonlinear softmax function for classification tasks. The model was trained by using the RMSprop optimizer, which was set a learning rate of 0.001, and the loss function was set to sparse categorical cross-entropy. The training process involved 100 epochs. Table 2 shows the various parameters utilized in the Inceptionv3 model structure.

Table 2: The hyperparameters applied in the Inceptionv3 model

Layer	**Parameters**
Base Model	VGG16
Input Shape	224 x 224 x 3
Flatten Layer	Yes
Dense Layer (ReLU)	256 units
Output Layer (Softmax)	95 units
Optimizer	Adam
Learning Rate	0.001
Loss Function	Sparse Categorical Cross-Entropy
Batch Size	128
Epochs	100

3.3.3 The EfficientNetB0 Model

The EfficientNetB0 model [55] is part of the EfficientNet family, which is designed to achieve high performance with fewer parameters by scaling depth, width, and resolution uniformly. EfficientNetB0 is particularly efficient for image classification tasks, making it well-suited for applications with limited computational resources. The EfficientNetB0 architecture was pre-trained on ImageNet, with its top layers removed. All layers in the base model were frozen. Figure 5 depicts the model structure.

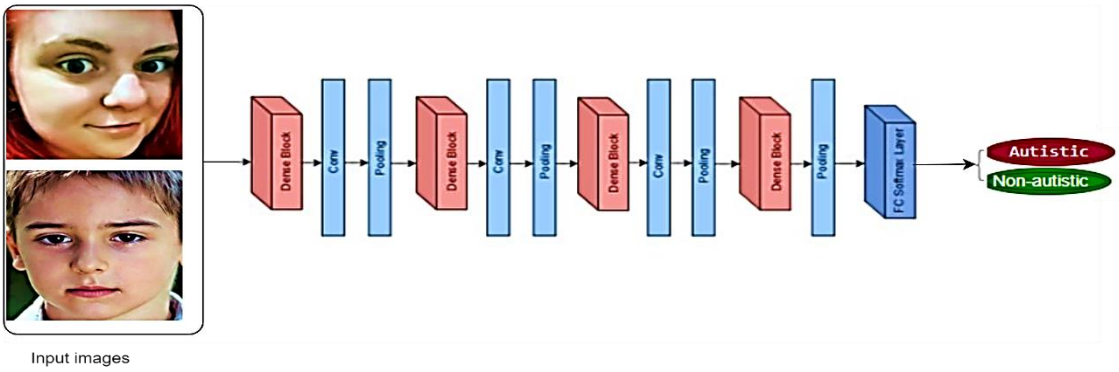


Figure 5: The structure of the EfficientNetB0 model

In this model architecture, we used a flatten layer, two dense layers with 256 neurons, and an activation function, which is ReLU. A rate of 0.5 as a value was set in the dropout layer, and in the last dense layer, there were two neurons representing the output dataset classes with a softmax activation function for the binary classification task. The dropout layer, which periodically deactivates neurons during training, aids in preventing overfitting, while the dense layers are designed to capture intricate patterns. Binary cross-entropy is utilized as the loss function when compiling the model using the RMSprop optimizer with a learning rate of 0.0001 and a decay rate of 1e-6. Additionally, effective feature extraction and dimensionality reduction are made possible by the dense block design in conjunction with global pooling layers. The parameters used in the model structure are shown in Table 3.

Table 3: Various hyperparameters applied in the EfficientNetB0 model.

Layer	**Parameters**
Input Shape	(224, 224, 3)
Optimizer	RMSprop
Learning Rate	0.0001
Loss Function	Binary Cross-Entropy
Activation Function	Softmax
Number of Units (Dense Layers)	256 (twice)
Dropout Rate	0.5
Number of Classes	2
Number of Epochs	100
Pre-trained Weights	ImageNet

3.4 Evaluation Metrics

Assessing the performance and testing results obtained by the proposed DL models, namely, VGG16, Inceptionv3, and EfficientNetB0, was crucial in gauging the effectiveness of the models. Several metrics are used to quantify performance, including accuracy, precision, recall, and F1-score, which are computed from the confusion matrix. The evaluation measures provide an alternative perspective on the models' advantages and disadvantages.

$$Accuracy = \frac{TP + TN}{FP + FN + TP + TN} \times 100 \quad (1)$$

$$F1 - score = 2 * \frac{precision \times Recall}{precision + Recall} \text{x}100\% \quad (2)$$

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives} x100\% \quad (3)$$

$$Recall = Sensitivity = \frac{True\ Positive}{True\ Positive + False\ Negatives} x100\% \quad (4)$$

4 Experimental Results

This section highlights the findings from our study on autism detection using DL methods-based models. We contrasted the performance and results obtained by three architectures, namely, VGG16, InceptionV3, and EfficientNetB0, which were trained and evaluated on a facial images dataset. The models were evaluated for their accuracy and capacity to classify images as autistic or non-autistic (Normal).

4.1 Testing Results of the VGG16 Model

The VGG16 model's classification results, shown in Table 4, illustrate its ability to distinguish amongst autistic and non-autistic images. For the autistic children, the model attained an F1-score of 97%, precision of 100%, and a recall of 95%, demonstrating an exceptional ability to correctly identify autistic cases with no false positives and few false negatives. There were 97 images in support of the autistic class. In the non-autistic class, the model had an F1-score of 62%, a recall of 100%, and precision of 44%, indicating a greater rate of false positives despite accurately recognizing all non-autistic instances. The support for the non-autistic class was four images.

Table 4: The testing classification results of the VGG16 model

	Precision	**Recall**	**F1-score**	**Support**	**Accuracy**
Autistic	100	95	97	97	
Non-autistic	44	100	62	4	**95**
Macro average	72	97	79	101	

The macro average calculation results for the F1-score, recall, and precision performance metrics were 72%, 97%, and 79%, respectively, across both classes. The overall accuracy of the model was 95%, indicating a high proportion of correct predictions. These results highlight the model's robust performance in detecting autism but also reveal a need for improvement in accurately classifying non-autistic cases. Figure 6 demonstrates the confusion matrix (CM) gained by the VGG16 model.

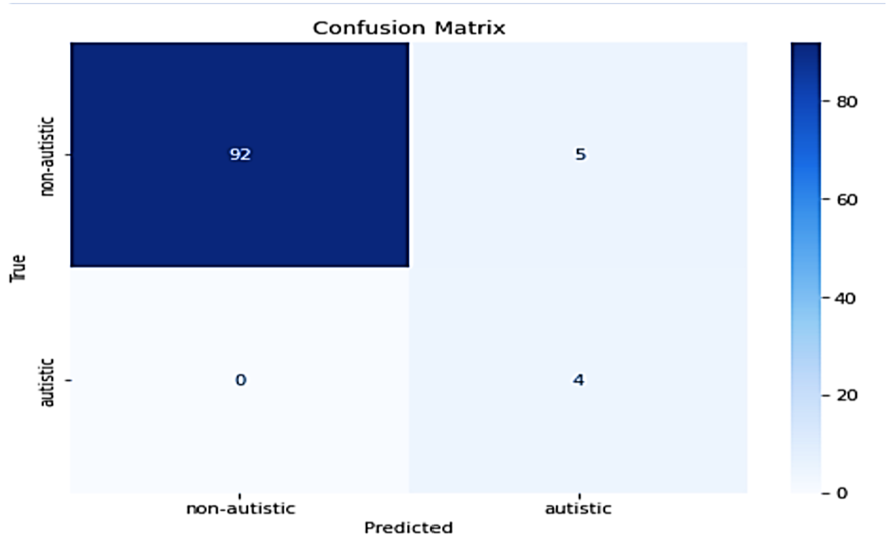


Figure 6: A CM of the VGG16 model

Figure 6 shows that the testing set contains 102 images, which are shown in a CM. Out of these images, 82 were accurately identified as non-autistic and four as autistic. However, five images were wrongly identified as autistic, resulting in false positives. Figure 7 visualizes the performance metrics for the VGG16 approach.

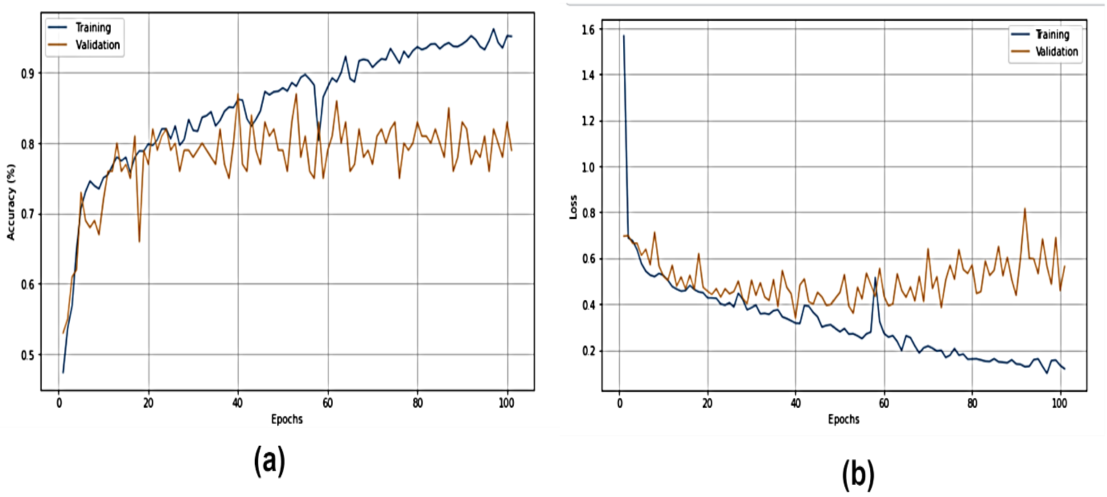


Figure 7: Performance metrics of the VGG16 model: a) training and validation accuracies, and b) model loss

As shown in Figure 7, the validation accuracy confirmed a notable improvement, increasing from 55% to 80%, while the training accuracy showed a substantial rise from 45% to 95%. Regarding the model's loss, the training loss significantly decreased from 16% to 1.5%, and the validation loss was mitigated, reducing from 75% to 55%.

4.2 Testing Results of the Inceptionv3 Model

This subsection presents the results of the Inceptionv3 model's autism detection tests. These findings are critical for assessing the model's test results and efficacy in differentiating between autistic and non-autistic children using facial expression traits. Table 5 displays the Inceptionv3 model's categorization testing performance, with a focus on its ability to distinguish between autistic and non-autistic children by utilizing facial expression traits.

Table 5: The testing classification results of the Inceptionv3 model

	Precision	Recall	F1-score	Support	Accuracy
Autistic	88	83	86	82	77
Non-autistic	42	53	47	19	
Macro average	65	58	66	101	

The model has a precision of 88% and 42% for the "Autistic" and "Non-autistic" categories, respectively, with recalls of 83% and 53%. Furthermore, the F1-scores were 86% and 47% for "Autistic" and "Non-autistic," respectively. These measurements provide vital insights into the model's capacity to reliably classify individuals based on facial expressions, which contributes to the progress of autism detection methods. Figure 8 gives the CM of the Inception model.

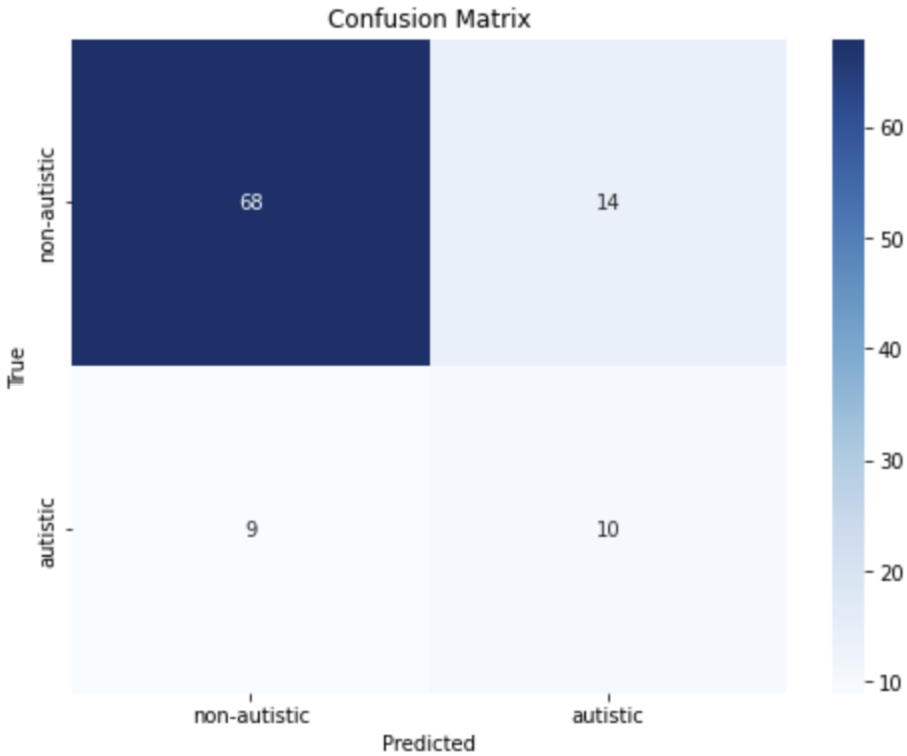


Figure 8: A CM of the Inceptionv3 model

As can be seen in Figure 8, the testing set contains 102 images, which are shown in a CM . Out of these images, 68 were accurately identified as non-autistic and 10 as autistic. However, 14 images were wrongly identified as autistic, resulting in false positives, and nine images were incorrectly predicted as non-autistic. Figure 9 displays the performance met rics of the Inception model.

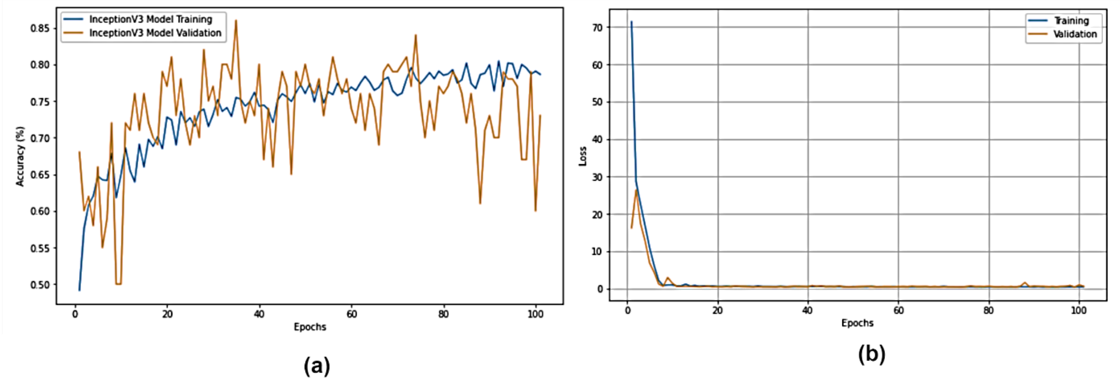


Figure 9: Evaluation metrics of the Inceptionv3 model: a) training and validation accuracies, and b) model loss

As shown in Figure 9, the validation accuracy saw a notable enhancement, increasing from 68% to 70%, while the training accuracy showed a significant rise from 45% to 79%. Regarding the model's loss, the training loss significantly decreased from 70% to 1%, and the validation loss was mitigated, reducing from 18% to 1%.

4.3 Testing Results of the EfficientnetB0 Model

This section presents the testing classification results of the EfficientNetB0 model for autism detection utilizing facial expression features, as shown in Table 6. The model had a precision of 100 for the "Autistic" class and 21 for the "Non-autistic" class. The recall scores were 89 for "Autistic" and 100 for "Non-autistic." The F1-scores were 94 and 35 for "Autistic" and "Non-autistic," respectively.

Table 6: Testing classification results of the EfficientNetB0 model

	Precision	**Recall**	**F1-score**	**Support**	**Accuracy**
Autistic	100	89	94	97	
Non-autistic	21	100	35	19	**89**
Macro average	61	94	65	100	

The overall accuracy of the model was 89. The macro average calculation and results for precision, recall, and F1-score were 61, 94, and 65, respectively. These metrics indicate the model's high precision and recall for the "Autistic" class, although the performance for the "Non-autistic" class suggests there are areas for improvement. The EfficientNetB0 model's accuracy and other metrics provide insights into its potential effectiveness and limitations in autism diagnosis based on facial expressions. Figure 10 shows the CM of the EfficientNetB0.

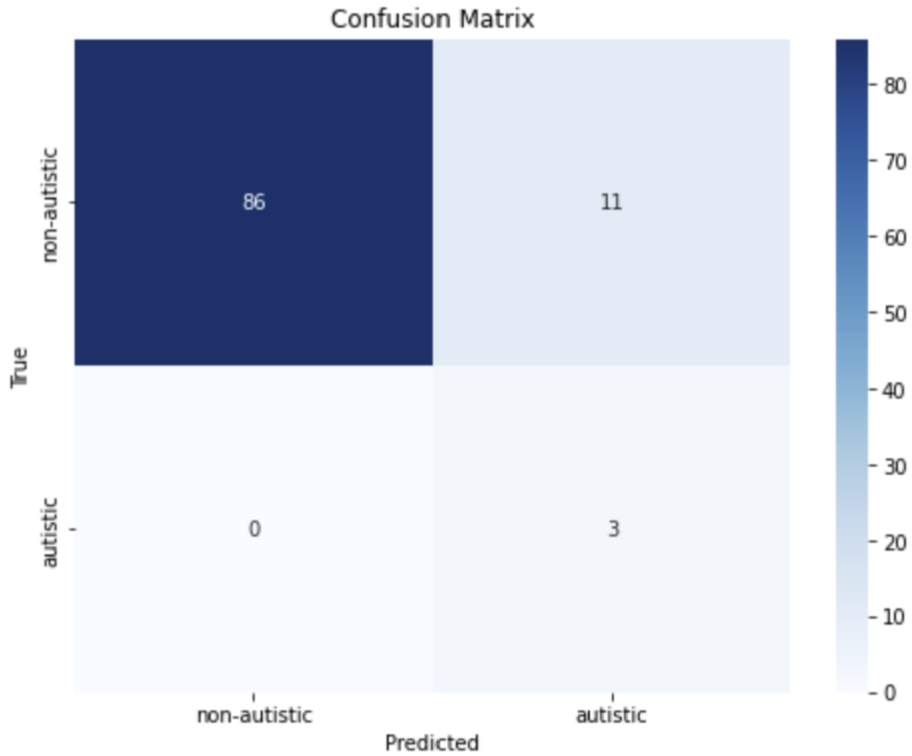
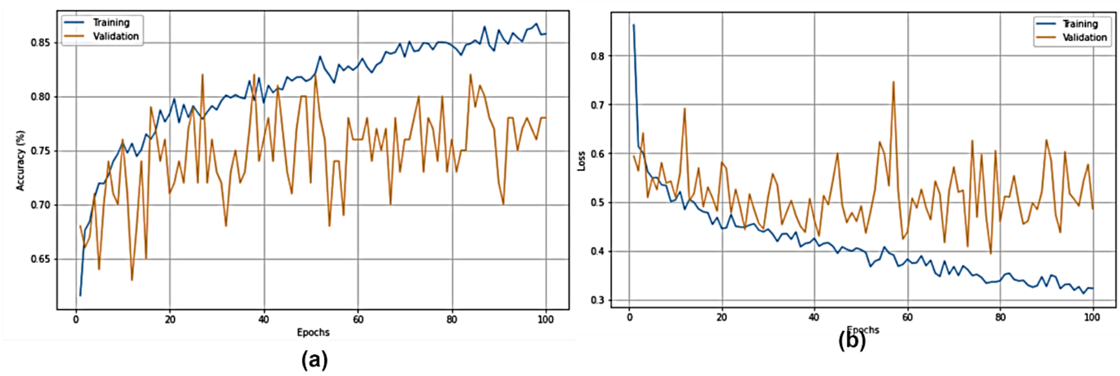


Figure 10: CM of the EfficientNetB0

As can be seen in Figure 10, 11 images were wrongly predicted as autistic, which resulted in false positives, and 86 images were correctly classified and predicted as non-autistic. Figure 11 depicts the learning curves of the EfficientNetB0 model.



As shown in Figure 11, the validation accuracy shows a notable enhancement, growing

Figure 11: The learning curves of the EfficientNetB0 model: a) training and validation accuracies, and b) model loss

from 68% to 79%, while the training accuracy showed a significant rise from 50% to 87%. Regarding the model's loss, the training loss significantly decreased from 88% to 3%, and the validation loss was mitigated, reducing from 60% to 50%.

5 Discussion

ASD is a complicated neurodevelopmental disorder marked by complications with social contact, communication, and recurrent behaviors, requiring early and precise diagnosis for prompt interventions. In this research, the performance and results attained by the used DL models, namely, VGG16, InceptionV3, and EfficientNetB0, were evaluated using facial expression characteristics to detect autism. The VGG16 model attained the greatest accuracy, of 95%, having perfect accuracy for autistic persons, but struggled with non-autistic individuals, indicating its high sensitivity but also a tendency to produce false positives. The InceptionV3 model performed moderately, with an accuracy of 77%, but there is tremendous opportunity for improvement, predominantly in minimizing false positives and enhancing non-autistic identification. The EfficientNetB0 model obtained 89% accuracy, outperforming InceptionV3 but not VGG16, with a strong capacity to identify autistic persons but low precision for non-autistic individuals, indicating many false positives. The VGG16 model, with 95% accuracy, is the most effective of the three, particularly in diagnosing autistic persons. However, the significant rate of false positives for non-autistic persons across all models suggests that sensitivity and specificity should be further refined and balanced. The InceptionV3 and EfficientNetB0 models, while demonstrating potential, show that there is substantial room for improvement in achieving reliable and balanced diagnostic performance.

Table 7 compares our study's findings to those of previous studies on ASD detection using the Kaggle dataset. Previous tests found accuracies of 87% with MobileNet [56], 90% and 91% with Xception [57][58], and 92% with MobileNet-V2 [59]. Our suggested model, VGG16, attained a 95% accuracy rate, demonstrating its superior capacity to detect ASD using facial expression characteristics. This demonstrates how the VGG16 model surpasses other existing models using diagnostic accuracy.

Table 7: Comparative analysis of our study results and those of previous research

Ref Id	**Model**	**Dataset**	**Accuracy %**
[56]	MobileNet	Kaggle dataset	87
[57]	Xception	Same	90
[58]	Xception	Same	91
[59]	MobileNet-V2	Same	92
Proposed	VGG16	Same	95

7 Conclusion

Studies have been performed to create computer systems that can help diagnose some neurodevelopmental abnormalities by analyzing photographs of faces, as facial morphology is known to be associated with brain developing issues. The main aim of the current study was to investigate several data-focused methods for diagnosing ASD utilizing advanced DL models, namely, VGG16, InceptionV3, and Efficientnet. The goal was to optimize the accuracy of predicting and diagnosing ASD. We employed the Kaggle ASD dataset, which includes 2940 images of different children and ages. Instead of prioritizing model and hyperparameter tweaking, our approach involved using augmentation and pre-processing strategies on the training set to identify the most efficient strategy for diagnosing ASD. After pre-processing and implementing a resampling strategy during training, the best performance was achieved. DL was utilized for each model examined in the study to attain superior outcomes to those of the original models. The VGG16 model, which is proposed for use, achieved the best accuracy, of 95%, and thus performed better than the other DL models. The suggested technique surpassed the accuracy and computing cost of current cutting-edge models. The outcomes of this study indicate that DL applications have the potential to enhance the diagnosis process of ASD. Additional research is advised to enhance the efficiency of these models and verify their efficacy on a wider scope.

References

1. Tisato, V.; Silva, J.A.; Longo, G.; Gallo, I.; Singh, A.V.; Milani, D.; Gemmati, D. Genetics and Epigenetics of One-Carbon Metabolism Pathway in ASD: A Sex-Specific Brain Epigenome? *Genes* **2021**, *12*, 782.
2. Ortega, M.A.; Álvarez-Mon, M.A.; García-Montero, C.; Fraile-Martínez, O.; Monserrat, J.; Martinez-Rozas, L.; Rodríguez-Jiménez, R.; Álvarez-Mon, M.; Lahera, G. Microbiota–gut–brain axis mechanisms in the complex network of bipolar disorders: Potential clinical implications and translational opportunities. *Mol. Psychiatry* **2023**, 1–29.
3. Chaste, P.; Leboyer, M. Autism risk factors: Genes, environment, and gene-environment interactions. *Dialog. Clin. Neurosci.* **2012**, *14*, 281–292.
4. Kohn, N.; Szopinska-Tokov, J.; Arenas, A.L.; Beckmann, C.; Arias-Vasquez, A.; Aarts, E. Multivariate associative patterns between the gut microbiota and large-scale brain network connectivity. *Gut Microbes* **2021**, *13*, 2006586.
5. Petrina, N.; Carter, M.; Stephenson, J. Recent developments in understanding friendship of children and adolescents with ASDs. In *Encyclopedia of ASDs*; Springer: Cham, Switzerland, 2021; pp. 3875–3880.
6. Kojovic, N.; Natraj, S.; Mohanty, S.P.; Maillart, T.; Schaer, M. Using 2D video-based pose estimation for automated prediction of ASDs in young children. *Sci. Rep.* **2021**, *11*, 1–10.
7. Georgoula, C.; Ferrin, M.; Pietraszczyk-Kedziora, B.; Hervas, A.; Marret, S.; Oliveira, G.; Rosier, A.; Crutel, V.; Besse, E.; Severo, C.A.; et al. A Phase III Study of Bumetanide Oral Liquid Formulation for the Treatment of Children and Adolescents Aged between 7 and 17 Years with ASD (SIGN 1 Trial): Participant Baseline Characteristics. *Child Psychiatry Hum. Dev.* **2022**, 1–13.
8. Zuckerman, K.E.; Broder-Fingert, S.; Sheldrick, R.C. To reduce the average age of autism diagnosis, screen preschoolers in primary care. *Autism* **2021**, *25*, 593–596.
9. Hettinger, J.A.; Liu, X.; Hudson, M.L.; Lee, A.; Cohen, I.L.; Michaelis, R.C.; Schwartz, C.E.; Lewis, S.M.; Holden, J.J. DRD2 and PPP1R1B (DARPP-32) polymorphisms independently confer increased risk for ASDs and additively predict affected status in male-only affected sib-pair families. *Behav. Brain Funct.* **2012**, *8*, 19.
10. Cardinal, D.N.; Griffiths, A.J.; Maupin, Z.D.; Fraumeni-McBride, J. An investigation of increased rates of autism in US public schools. *Psychol. Sch.* **2021**, *58*, 124–140.
11. Guillon, Q.; Hadjikhani, N.; Baduel, S.; Rogé, B. Visual Social Attention in ASD: Insights from Eye Tracking Studies. *Neurosci. Biobehav. Rev.* **2014**, *42*, 279–297.
12. Kanhirakadavath, M.R.; Chandran, M.S.M. Investigation of Eye-Tracking Scan Path as a Biomarker for Autism Screening Using Machine Learning Algorithms. *Diagnostics* **2022**, *12*, 518
13. Mujeeb Rahman, K.K.; Subashini, M.M. Identification of Autism in Children Using Static Facial Features and Deep Neural Networks. *Brain Sci.* **2022**, *12*, 94.
14. Aldhyani, T.H.H.; Verma, A.; Al-Adhaileh, M.H.; Koundal, D. Multi-Class Skin Lesion Classification Using a Lightweight Dynamic Kernel Deep-Learning-Based Convolutional Neural Network. *Diagnostics* **2022**, *12*, 2048.

15. Akter, T.; Satu, M.S.; Barua, L.; Sathi, F.F.; Ali, M.H. Statistical Analysis of the Activation Area of Fusiform Gyrus of Human Brain to Explore Autism. *Int. J. Comput. Sci. Inf. Secur. IJCSIS* **2017**, *15*, 331–337. [Google Scholar]
16. Aldhyani, T.H.H.; Nair, R.; Alzain, E.; Alkahtani, H.; Koundal, D. Deep Learning Model for the Detection of Real Time Breast Cancer Images Using Improved Dilation-Based Method. *Diagnostics* **2022**, *12*, 2505.
17. Schelinski, S.; Borowiak, K.; von Kriegstein, K. Temporal Voice Areas Exist in ASD but Are Dysfunctional for Voice Identity Recognition. *Soc. Cogn. Affect. Neurosci.* **2016**, *11*, 1812–1822.
18. Jiang, X.; Chen, Y.F. Facial Image Processing. In *Applied Pattern Recognition*; Bunke, H., Kandel, A., Last, M., Eds.; Studies in Computational Intelligence; Springer: Berlin/Heidelberg, Germany, 2008; pp. 29–48.
19. Garcia, C.; Ostermann, J.; Cootes, T. Facial Image Processing. *Eurasip J. Image Video Process.* **2008**, *2007*, 070872.
20. Yolcu, G.; Oztel, I.; Kazan, S.; Oz, C.; Palaniappan, K.; Lever, T.E.; Bunyak, F. Deep Learning-Based Facial Expression Recognition for Monitoring Neurological Disorders. In Proceedings of the 2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Kansas City, MO, USA, 13–16 November 2017; pp. 1652–1657.
21. Yolcu, G.; Oztel, I.; Kazan, S.; Oz, C.; Palaniappan, K.; Lever, T.E.; Bunyak, F. Facial Expression Recognition for Monitoring Neurological Disorders Based on Convolutional Neural Network. *Multimed. Tools Appl.* **2019**, *78*, 31581–31603.
22. Haque, M.I.U.; Valles, D. A Facial Expression Recognition Approach Using DCNN for Autistic Children to Identify Emotions. In Proceedings of the 2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), Vancouver, BC, Canada, 1–3 November 2018; pp. 546–551.
23. Yang, M.; Cao, M.; Chen, Y.; Chen, Y.; Fan, G.; Li, C.; Wang, J.; Liu, T. Large-scale brain functional network integration for discrimination of autism using a 3-D deep learning model. *Front. Hum. Neurosci.* **2021**, *15*, 277.
24. Ahsan, M.M.; E Alam, T.; Trafalis, T.; Huebner, P. Deep MLP-CNN model using mixed-data to distinguish between COVID-19 and Non-COVID-19 patients. *Symmetry* **2020**, *12*, 1526.
25. Ahsan, M.M.; Gupta, K.D.; Islam, M.M.; Sen, S.; Rahman, M.; Shakhawat Hossain, M. COVID-19 symptoms detection based on nasnetmobile with explainable ai using various imaging modalities. *Mach. Learn. Knowl. Extr.* **2020**, *2*, 490–504.
26. Ahsan, M.M.; Ahad, M.T.; Soma, F.A.; Paul, S.; Chowdhury, A.; Luna, S.A.; Yazdan, M.M.S.; Rahman, A.; Siddique, Z.; Huebner, P. Detecting SARS-CoV-2 from chest X-Ray using artificial intelligence. *IEEE Access* **2021**, *9*, 35501–35513.
27. Ahsan, M.M.; Mahmud, M.; Saha, P.K.; Gupta, K.D.; Siddique, Z. Effect of data scaling methods on machine learning algorithms and model performance. *Technologies* **2021**, *9*, 52.
28. Ahsan, M.M.; Nazim, R.; Siddique, Z.; Huebner, P. Detection of COVID-19 patients from CT scan and chest X-ray data using modified MobileNetV2 and LIME. *Healthcare* **2021**, *9*, 1099.
29. Ahsan, M.M.; Luna, S.A.; Siddique, Z. Machine-Learning-Based Disease Diagnosis: A Comprehensive Review. *Healthcare* **2022**, *10*, 541.
30. Ahsan, M.M.; Siddique, Z. Machine learning-based heart disease diagnosis: A systematic literature review. *Artif. Intell. Med.* **2022**, *128*, 102289.
31. Haque, R.; Islam, N.; Islam, M.; Ahsan, M.M. A comparative analysis on suicidal ideation detection using NLP, machine, and deep learning. *Technologies* **2022**, *10*, 57.

32. Pham, H.; Jones, M.A.; Gai, T.; Islam, W.; Danala, G.; Jo, J.; Zheng, B. Identifying an optimal machine learning generated image marker to predict survival of gastric cancer patients. In Proceedings of the Medical Imaging 2022: Computer-Aided Diagnosis, San Diego, CA, USA, 4 April 2022; Volume 12033, pp. 408–414.
33. Akter, T.; Ali, M.H.; Khan, M.; Satu, M.; Uddin, M.; Alyami, S.A.; Ali, S.; Azad, A.; Moni, M.A. Improved transfer-learning-based facial recognition framework to detect autistic children at an early stage. *Brain Sci.* **2021**, *11*, 734. [**Google Scholar**] [**CrossRef**]
34. Hosseini, M.P.; Beary, M.; Hadsell, A.; Messersmith, R.; Soltanian-Zadeh, H. Deep Learning for Autism Diagnosis and Facial Analysis in Children. *Front. Comput. Neurosci.* **2021**, *15*, 789998. [**Google Scholar**] [**CrossRef**]
35. Mujeeb Rahman, K.; Subashini, M.M. Identification of Autism in Children Using Static Facial Features and Deep Neural Networks. *Brain Sci.* **2022**, *12*, 94. [**Google Scholar**] [**CrossRef**]
36. Khosla, Y.; Ramachandra, P.; Chaitra, N. Detection of autistic individuals using facial images and deep learning. In Proceedings of the 2021 IEEE International Conference on Computation System and Information Technology for Sustainable Solutions (CSITSS), Bangalore, India, 16–18 December 2021; pp. 1–5.
37. Alsaade, F.W.; Alzahrani, M.S. Classification and Detection of ASD Based on Deep Learning Algorithms. *Comput. Intell. Neurosci.* **2022**, *2022*, 8709145.
38. Piosenka, G. Detect Autism from a Facial Image. Available online: **https://cutt.ly/ibIXt5a** (accessed on 20 September 2022).
39. Khalaji, E.; Eraslan, S.; Yesilada, Y.; Yaneva, V. Effects of data preprocessing on detecting autism in adults using web-based eye-tracking data. *Behav. Inf. Technol.* **2022**, 1–9.
40. Cardoso, T.V.; Michelassi, G.C.; Franco, F.O.; Sumiya, F.M.; Portolese, J.; Brentani, H.; Machado-Lima, A.; Nunes, F.L. ASD diagnosis based on trajectories of eye tracking data. In Proceedings of the 2021 IEEE 34th International Symposium on Computer-Based Medical Systems (CBMS), Aveiro, Portugal, 7–9 June 2021; pp. 50–55.
41. Radha, S.; Abinaya, M.N.; Darshini, K.D.; Kamalnath, K.N.; Jayanthi, P. Ensemble of Behavioral Features for Early Detection of ASD. In *International Conference on Soft Computing for Security Applications*; Springer Nature: Singapore, 2023; pp. 681–694.
42. Cilia, F.; Carette, R.; Elbattah, M.; Dequen, G.; Guérin, J.-L.; Bosche, J.; Vandromme, L.; Le Driant, B. Computer-Aided Screening of ASD: Eye-Tracking Study Using Data Visualization and Deep Learning. *JMIR Hum. Factors* **2021**, *8*, e27706.
43. Gaspar, A.; Oliva, D.; Hinojosa, S.; Aranguren, I.; Zaldivar, D. An optimized Kernel Extreme Learning Machine for the classification of the ASD by using gaze tracking images. *Appl. Soft Comput.* **2022**, *120*, 108654.
44. Zhao, Z.; Tang, H.; Zhang, X.; Qu, X.; Hu, X.; Lu, J. Classification of Children With Autism and Typical Development Using Eye-Tracking Data from Face-to-Face Conversations: Machine Learning Model Development and Performance Evaluation. *J. Med. Internet Res.* **2021**, *23*, e29328.
45. Kollias, K.-F.; Syriopoulou-Delli, C.K.; Sarigiannidis, P.; Fragulis, G.F. Autism detection in High-Functioning Adults with the application of Eye-Tracking technology and Machine Learning. In Proceedings of the 2022 11th International Conference on Modern Circuits and Systems Technologies, 2022 (MOCAST), Bremen, Germany, 8–10 June 2022; pp. 1–4.
46. Sun, B.; Wang, B.; Wei, Z.; Feng, Z.; Wu, Z.L.; Yassin, W.; Stone, W.S.; Lin, Y.; Kong, X. The Role of EEG Functional Connectivity Coupled with Eye Tracking in Early Diagnosis of ASD. 2023; *in press*.
47. Omar, K.S.; Mondal, P.; Khan, N.S.; Rizvi, M.R.K.; Islam, M.N. A Machine Learning Approach to Predict ASD. In Proceedings of the 2019 International Conference on

Electrical, Computer and Communication Engineering (ECCE), Bazar, Bangladesh, 7–9 February 2019; pp. 1–6.
48. Sharma, M. Improved Autistic Spectrum Disorder Estimation Using Cfs Subset with Greedy Stepwise Feature Selection Technique. *Int. J. Inf. Technol.* **2019**, *14*, 1251–1261.
49. Satu, M.S.; Sathi, F.F.; Arifen, M.S.; Ali, M.H.; Moni, M.A. Early Detection of Autism by Extracting Features: A Case Study in Bangladesh. In Proceedings of the 2019 International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST), Dhaka, Bangladesh, 10–12 January 2021; pp. 400–405.
50. Erkan, U.; Thanh, D.N. ASD Detection with Machine Learning Methods. *Curr. Psychiatry Res. Rev. Former. Curr. Psychiatry Rev.* **2019**, *15*, 297–308.
51. Thabtah, F.; Abdelhamid, N.; Peebles, D. A Machine Learning Autism Classification Based on Logistic Regression Analysis. *Health Inf. Sci. Syst.* **2019**, *7*, 12.
52. autistic children's facial images dataset, available at: https://www.kaggle.com/datasets/cihan063/autism-image-data Accessed: 5 June 2024.
53. Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4700-4708).
54. Simonyan, K., Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv preprint arXiv:1409.1556, 10 September 2014.
55. Efficient model. Tan, M., & Le, Q. (2019, May). Efficientnet: Rethinking model scaling for convolutional neural networks. In International conference on machine learning (pp. 6105-6114). PMLR.
56. Thabtah, F.; Kamalov, F.; Rajab, K. A New Computational Intelligence Approach to Detect Autistic Features for Autism Screening. Int. J. Med. Inform. 2018, 117, 112–124.
57. Alkahtani, H., Aldhyani, T. H., & Alzahrani, M. Y. (2023). Deep learning algorithms to identify ASD in children-based facial landmarks. Applied Sciences, 13(8), 4855.

Notes on contributors

Dr. Nesren Farhah is the Assistant Professor and Dean of Graduate Studies, Research, and Innovation at Saudi Electronic University. She began her academic journey with a Bachelor's degree in Biology from King Khalid University in Saudi Arabia. She then advanced her studies in the United States, earning a Master's degree followed by a PhD in Biomedical Informatics from Nova Southeastern University, Florida, and Rutgers University, New Jersey, respectively. Dr. Farhah's research interests encompass a broad spectrum of critical topics including health informatics, public health, e-health, intelligent systems, clinical decision support, and health data privacy and security. Her expertise and leadership are pivotal in fostering innovation and research in healthcare technology.