Autism Spectrum Disorder prediction in Toddlers using Deep Learning Techniques coupled with Questionnaires

Innocent Mochahari, Jimmy Brahma

Guided by Dr. Ranjan Maity (Assistant Professor)
Department of Computer Science and Engineering
Central Institute of Technology Kokrajhar, 783370, Kokrajhar

Abstract

In this paper, we describe a novel method for diagnosing Autism Spectrum Disorder (ASD) in Toddlers that integrates the Q-CHAT Questionnaire for Autism with the deep learning image classification models. The neurodevelopmental illness known as autism spectrum disorder (ASD) is characterized by issues with social interaction, communication, and repetitive behavior patterns. For those with ASD, early detection and intervention can greatly improve results. Our model makes use of the Q-CHAT, a popular screening tool for ASD in children to calculate a probabilistic autistic score, as well as a deep learning image classifier to learn the autistic features from an image. We improve the accuracy and robustness of ASD detection by fusing our proposed image classifier with the Q-CHAT replies. The efficiency of our method is demonstrated by our experimental findings on a sizable dataset, which indicate great possibilities for machine learning-based early ASD detection. The proposed techniques are evaluated on a publicly available ASD dataset which has 1463 images of Autistic children's faces and 1463 images for non-autistic children. Our VGG16 model demonstrated an impressive 88% accuracy on the testing data.

1. Introduction

A complex neurodevelopmental illness known as autism spectrum disorder (ASD) is characterized by deficiencies in social interaction, communication difficulties, and repetitive behavioral patterns. ASD is a significant public health concern because according to the world health organization it has been estimated that worldwide about 1 in 100 children has autism. Early detection and intervention make it possible for rapid therapeutic interventions and support services and they are essential to improve the long-term results for people with ASD.

An autistic person is generally not able to do social interaction and communication with other persons [1] or [7]. Autism Spectrum Disorder (ASD) impacts human brain development and the disorder’s origin is attributed to a blend of genetic and environmental factors. Signs can emerge around age three and persist throughout life.

Common challenges encompass various social and communication difficulties:

- Limited eye contact
- Reduced pain sensitivity
- Inadequate response to sounds
- Disinterest in cuddling
- Challenges in expressing gestures
- Minimal interaction with others
- Preference for solitude

To make a diagnosis of ASD, psychologists draw on a number of sources of information:

- Patient interviews.
- Observations of the patient's behavior.
- Tests of cognitive and language abilities.
- Medical tests to rule out other conditions.
- Interviews with parents, teachers or other adults who can answer questions about the patient's social, emotional and behavioral development.

With the rise of application of machine learning and deep learning-based models in the predictions of various human diseases, the early detection of autism by extracting the facial features from images now seems possible. In this paper, we describe a novel method for diagnosing ASD by transfer learning a pretrained CNN model and combining it with the Q-CHAT Questionnaire for Autism in Toddlers. We hope to increase the reliability and accuracy of early ASD detection by fusing our best CNN model with Q-CHAT responses.

The advancement of ASD identification and diagnosis can be achieved by combining machine learning models with clinical screening techniques. We may be able to shorten the time and resources needed for evaluation by automating the screening process, enabling early detection and intervention. Additionally, machine learning models can provide reliable and objective evaluations.

1.1 Objectives

The objectives of our project are listed below:

- Improve the accuracy and efficiency of diagnosing the Autism Spectrum Disorder by using deep learning techniques.
- Early identification and diagnosis which can lead to earlier interventions and treatments, that can improve outcomes for individuals with ASD.

1.2 Motivation

The urgent need for early diagnosis and treatment of autism spectrum disorder (ASD) is what inspired this effort. Children across the world are impacted by ASD, which has profound effects on their growth, social connections, and general well-being. The provision of appropriate support, treatments, and therapies that can greatly enhance children's outcomes and quality of life depends on the early detection of ASD in children.

Medical literature on the connection between facial morphology and autism has been thoroughly read by researchers. The morphological data was gathered using physical measurement techniques.
This takes a lot of time and is prone to mistakes. As a broad diagnosis of autism, these techniques are ineffective. The extraction of face traits from an image would seem to be a feasible way to create a low-cost, easily usable diagnostic tool given the advancements in neural networks. By transferring this technology to a publicly accessible website, parents would be able to submit one or more photographs of their child's face, complete the Q-chat questionnaire, and receive a likelihood as to whether or not their child has autistic features.

2. Related Work

Vaishali R, Sasikala R. et al. [1] proposed a method to identify Autism with optimum behavior sets. In this work they used Swarm intelligence based single-objective binary firefly feature selection wrapper. It was found that 10 features among 21 features of ASD dataset are sufficient to distinguish between ASD and non-ASD patients. The results obtained produced an average accuracy in the range of 92.12%-97.95% with optimum feature.

In 2015, Alessandro Crippa et al. [2] applied machine-learning procedure to the kinematic analysis of a simple reach, grasp, and drop task performed by preschool children with ASD in comparison to their mental-age-matched, typically developing peers which showed that the SVM algorithm reached a good mean individual classification in the comparisons between children with ASD and healthy controls with a maximum accuracy of 96.7 %. The classification accuracy that was achieved in this study was consistent with previous SVM applications to MRI data and to diffusion tensor imaging (DTI) data (Ingalhalikar et al. 2011; Deshpande et al. 2013) or with quadratic discriminant function application on diffusion tensor asymmetries (Lange et al. 2010).

Suman Raj et al [3] in 2019, attempted detection of Autism Spectrum Disorder using various machine learning and deep learning techniques. In their work it was found that after handling missing value, both the SVM and CNN based models show the same accuracy of prediction of about 98.30 % for ASD Child dataset. However for adolescent and adults, the CNN based model achieved highest accuracy result than all the other considered model building techniques. This study strongly suggested the possible implementation of a CNN based model for detection of Autism Spectrum Disorder instead of the other conventional machine learning classifier.

In 2019, K. S. Omar et al. [4] study suggests a brand-new approach to forecast Autism Spectrum Disorder (ASD) using cutting-edge machine learning methods to create a mobile application for predicting ASD in individuals of any age and to provide an efficient prediction model based on ML approach. As a result of this research, a mobile application was created based on the proposed prediction model and an autism prediction model was created by combining Random Forest-CART (Classification and Regression Trees) and Random Forest-Id3 (Iterative Dichotomiser 3). With the help of the AQ-10 dataset and 250 actual datasets gathered from individuals with and without autistic features, the suggested model was assessed. According to the evaluation findings, the suggested prediction model performs better for both types of datasets in terms of accuracy, specificity, sensitivity, precision, and false positive rate (FPR).

Angelina Lu and Marek Perkowski in [5] carried out experiments on computer vision, deep convolutional neural networks (CNNs) have gained popularity, especially in facial recognition systems. The VGG16 architecture has demonstrated great accuracy in image identification, it is the deep learning-based approach that their research is centered on using. To take advantage of the VGG16 model's capabilities, transfer learning is used with pre-trained models. They use Keras-VGGFace, a VGGFace implementation built on the Keras framework, combined with TensorFlow, an open-source machine learning platform. This strategy improved the facial recognition system's performance and accuracy by making use of the many resources and pre-trained models that were already accessible

3. Proposed Methodology

Fig. 3.1 shows the overview of the architecture of the workflow of our design.

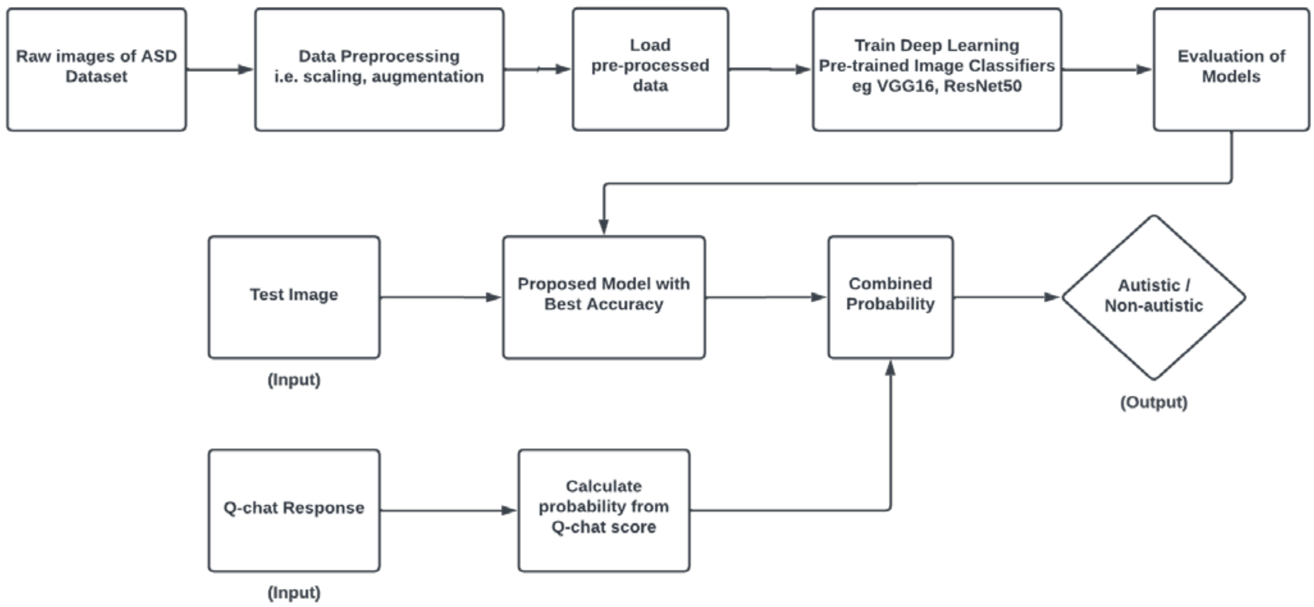


Fig. 3.1: Overall architecture for the proposed method

3.1 Experimental Setup

This work has been implemented in Jupyter Notebook using Python 3. For the implementation of our project, we utilized the following hardware components:

- Personal Laptop: We used our own laptops equipped with Nvidia GPU (1650) and an Intel i5 10th gen processor. This configuration provides the necessary computational power to run our current implementations.
- GPU Server: We also accessed a GPU server provided by our institute, which features an Ubuntu operating system equipped with Nvidia Tesla V100.

3.2 Data Collection and Preprocessing

We have collected a dataset of pictures that show toddlers with autistic and non-autistic classes from Kaggle (Autistic Children Facial Image Data Set) that pays particular attention to their faces. This is the only publicly available dataset and the creator has provided insights into their approach, asserting that their intention aligns with fair use principles outlined in Section 107 of the Copyright Act. The creator's goal was to develop a high-quality dataset for advancing deep learning AI classification models. They emphasize their non-profit motive and commitment to promoting teaching and research, which they believe falls within the scope of fair use. We use the dataset as a component of our research, mindful of the ongoing discussions related to copyright, fair use, and ethical data utilization. Fig 3.2 shows a few random images from the dataset.

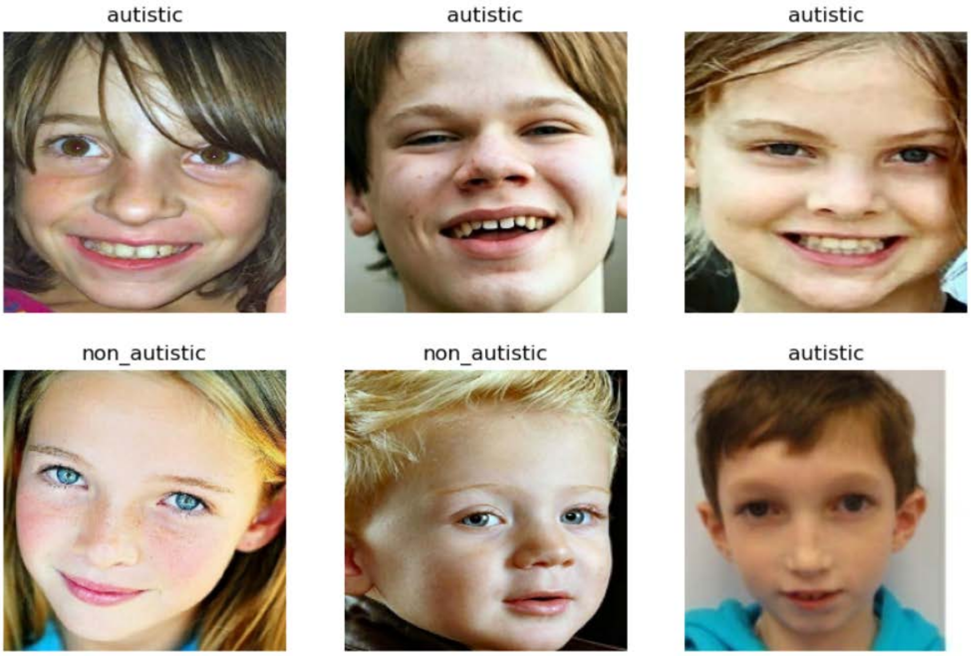


Fig 3.2 Images from the ASD Dataset

Data Pre-processing of images is the step taken to format images before they are used to train a model. This includes, but is not limited to, resizing, orienting, and color corrections. A good outcome is always produced by well-preprocessed data. We scaled the images to 200x200 pixels, and applied data augmentation techniques to enhance the quality of the dataset.

3.3 Training, Validation and Testing Split

The complete dataset contains three segments: training, validation, and testing. The training dataset comprises a total of 1263 images for both classes, while both the validation and testing datasets consist of 100 images each. This division enables effective model learning, tuning, and evaluation.

3.4 Pre-trained CNN Models

Convolutional Neural Networks (CNNs) are a special type of artificial neural network designed for tasks involving visual data. They are inspired by the human visual system and consist of layers that automatically learn and extract meaningful features from images. The key idea behind CNNs is the use of convolutional layers, which apply filters to input images to detect various patterns, textures, and structures.

Pretrained models like VGG16, ResNet50, and EfficientNetV2 are advanced architectures in the field of deep learning, specifically within computer vision. They belong to a class of Convolutional Neural Networks (CNNs).

3.4.1 VGG16 (Visual Geometry Group 16):

VGG16 is a CNN architecture developed by researchers at the University of Oxford [11]. It is characterized by its simplicity and uniformity in design, featuring layers with small 3x3 convolutional filters. Despite its straightforward structure, VGG16 is highly effective at image classification tasks. The VGG16 architecture is shown in Figure 3.3.

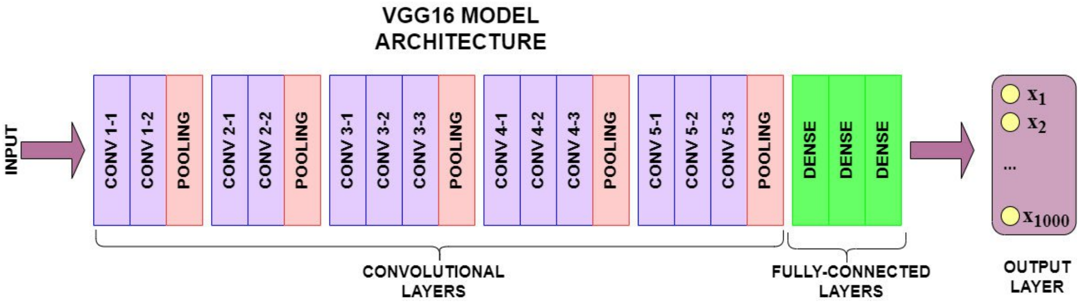


Fig 3.3 VGG16 model Architecture

3.4.2 ResNet50 (Residual Network 50):

ResNet50 is a member of the ResNet family, pioneered by Microsoft Research [12]. ResNet models tackle the challenge of training very deep networks by introducing residual blocks. These blocks allow the network to efficiently learn the difference between the current and target outputs, enabling the training of extremely deep architectures. ResNet50, with its 50 layers, is a popular choice as a starting point for various computer vision tasks. The ResNet50 architecture is shown in Figure 3.4.

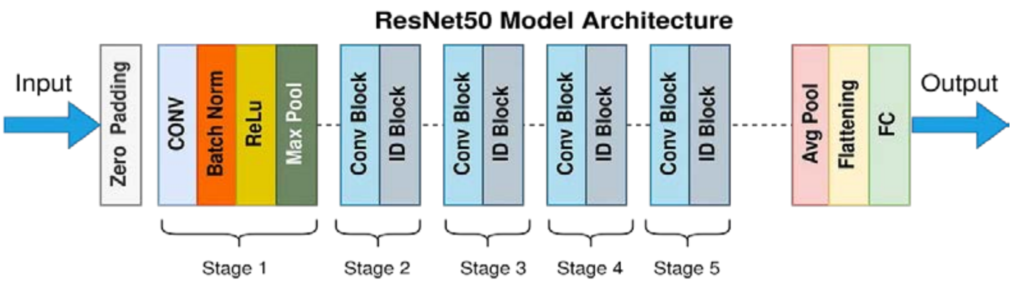


Fig 3.4 ResNet50 Model Architecture

4. Implementation and Result

4.1 Implementation

4.1.1 Implementation of Image Classifier

We developed a binary classification model using the VGG16 architecture pretrained on the ImageNet dataset. Our proposed VGG16 model comprises they key components:

- VGG16 Convolutional Base.
- Flattening Layer: Prepares features for processing.
- Dense Layers (256 neurons).
- Dropout Layer (0.5 rate).
- Output Layer (2 neurons).

To optimize the training process, we employed the Adam optimizer with a learning rate of 0.001. We implemented model checkpointing to save the best-performing model based on validation accuracy. This safeguard ensures that model parameters are stored only when an improvement in validation accuracy is observed. Additionally, we utilized a learning rate reducer function that dynamically adjusts the learning rate as the validation loss plateaus.

For the ResNet-50 model the components are:

- ResNet50 Convolutional Base.
- Flattening Layer.
- Dense Layers (32 neurons).
- Dropout Layer (0.5 rate).
- Output Layer (2 neurons).

During training, the ResNet50 base features are kept fixed, preserving their learned representations, while the added layers are optimized using the Adam optimizer and binary cross-entropy loss. To improve generalization, L2 regularization with a strength of 0.01 is applied exclusively to the dense layers. This regularization technique aids in preventing overfitting by encouraging simpler weight configurations, enhancing the model's capacity to perform well on new data.

4.1.2 Implementation of Q-Chat-10 score

Fig 4.1 shows a quick referral guide for parents to complete about their toddler (18 - 24 months) with concerns about autism. We take an example of a Q-chat response in the last column of the figure. The numbers in this column represent the options selected respectively.

Questions	1	2	3	4	5	example response
1- Does your child look at you when you call their name?	Always	Usually	Sometimes	Rarely	Never	3
2- How easy is it for you to make eye contact with your	Very easy	Quite easy	Quite difficult	Very difficult	Impossible	4
3- Does your child point to indicate that they want something? (e.g., a toy that is out of reach)	Many times a day	Several times a day	Several times a week	Less than once a week	Never	3
4- Does your child point to share an interest with you? (e.g., pointing to an interesting scene)	Many times a day	Several times a day	Several times a week	Less than once a week	Never	1
5- Does your child engage in pretend play? (e.g., taking care of dolls, talking on a pretend phone)	Many times a day	Several times a day	Several times a week	Less than once a week	Never	1
6- Does your child follow your gaze?	Many times a day	Several times a day	Several times a week	Less than once a week	Never	2
7- If you or someone in your family is visibly upset, does your child show signs of wanting to comfort you? (e.g., stroking your hair, hugging)	Always	Usually	Sometimes	Rarely	Never	5
8- Would you describe your child's first words as:	Very common	Quite common	Slightly uncommon	Very uncommon	My child does not speak	2
9- Does your child use simple gestures? (e.g., waving goodbye)	Many times a day	Several times a day	Several times a week	Less than once a week	Never	2
10- Does your child stare at nothing for no apparent reason?	Never	Less than once a week	Several times a week	Several times a day	Many times a day	4

Fig 4.1. Q-CHAT-10 and a response as an example

We assign the score according to Allison, Carrie, Bonnie Auyeung, and Simon Baron-Cohen. "Toward brief “red flags” for autism screening [6] in such a way that if the parent selects options from column 3, 4 or 5 then we assign 1 point each else we assign 0 point. After that we find the total sum and evaluate it further. If a child scores more than 3 out of 10, the health professional may consider referring the child for a multi-disciplinary assessment. The paper also concluded that the Q-Chat-10 score is directly correlated with having ASD or Not. That scoring is what determines the diagnosis. So if they get a 6 or higher, they automatically get a diagnosis of Yes. Therefore if our sum obtained is more than 6 points then we use a linear mapping function to assign the probability of having autism. This function is shown in Fig 4.2.

```
if (score > = 6)
        autistic_probability = (score - 6) / 4 * 0.5 + 0.5
else
        autistic_probability = score / 6 * 0.5
```

Fig 4.2. Function to calculate probability of autism from Q-Chat-10

For the total autism score we have combined the probability that we got from our best image classification model with the Q-Chat-10 probability score and divided it by 2. This gives us the final probabilistic score of having autism.

4.2 Result

4.2.1 Evaluation Metrics

In this section, we discuss the evaluation metrics used to assess the performance of our deep learning models. Properly evaluating the model's performance is crucial to gain insights into its capabilities and limitations. We focus on key metrics such as training and validation loss, training and validation accuracy, as well as testing accuracy. Additionally, we utilize confusion matrices to delve deeper into the model's classification performance on the testing dataset.

Training Loss and Accuracy: During the training process, the model's loss and accuracy on the training dataset are monitored. Training loss measures the dissimilarity between predicted values and ground truth labels. As training progresses, a reduction in training loss indicates improved convergence of the model. Training accuracy reflects the model's ability to correctly classify examples within the training dataset.

Validation Loss and Accuracy: To prevent overfitting, a validation dataset is employed to assess the model's performance on unseen data. Validation loss and accuracy provide insights into how well the model generalizes to new samples. Monitoring these metrics allows us to identify the point where the model starts overfitting, helping us make informed decisions regarding training duration and hyperparameter tuning.

Testing Accuracy: Once the model is trained and validated, its true performance is evaluated on a separate testing dataset that it has never encountered before. Testing accuracy quantifies the percentage of correctly classified instances and demonstrates the model's ability to predict real-world unseen data.

Confusion Matrix: A confusion matrix is a valuable tool to understand the classification performance in more detail. It provides a comprehensive breakdown of the model's predictions, including true positive, true negative, false positive, and false negative counts for each class. Fig 4.3 shows the confusion matrix for a binary classification model. From the confusion matrix, we can calculate metrics such as precision, recall, and accuracy.

<table>
<tr><td></td><td colspan="2">Predicted ASD</td></tr>
<tr><td rowspan="2">Actual ASD</td><td>True Positive (TP)</td><td>False Positive (FP)</td></tr>
<tr><td>False Negative (FN)</td><td>True Negative (TN)</td></tr>
</table>

Fig 4.3. Confusion Matrix

Precision = TP / (TP + FP)
Recall = TP / (TP + FN)
Accuracy = (TP + TN) / (TP + TN + FP + FN)

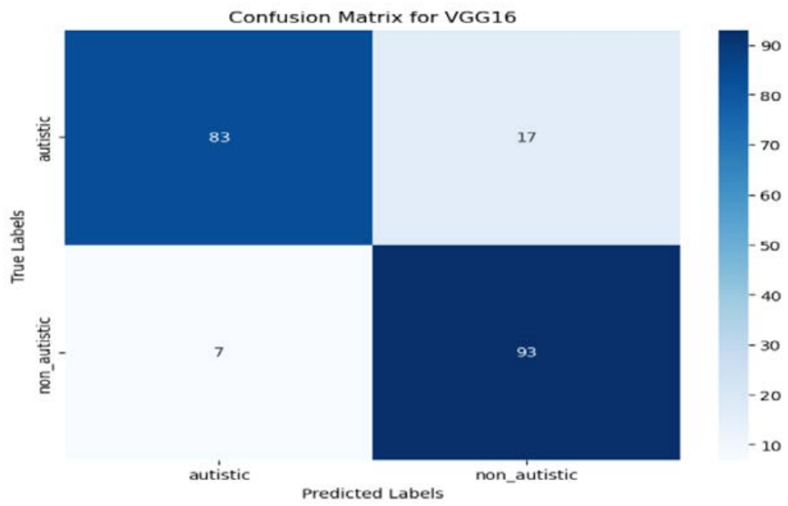
4.2.2 Experimental Findings

The experimental results of the two deep learning image classification models have been shown below.

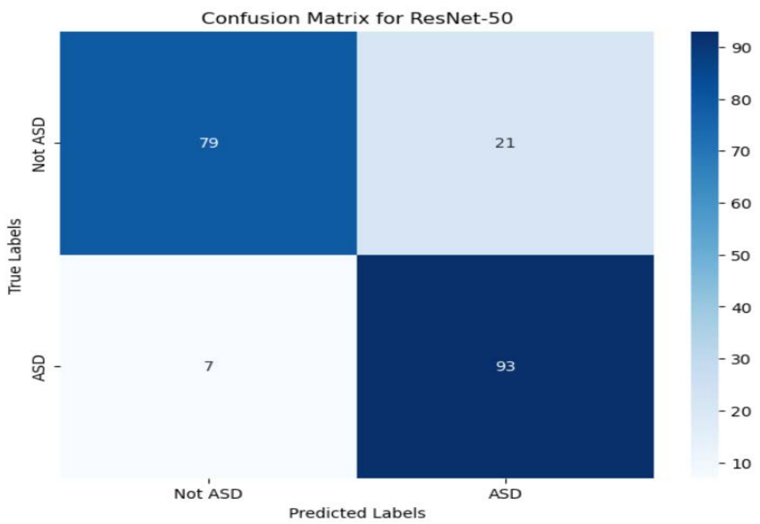
Table 4.1 Overall Results for Autism Spectrum Disorder Detection using image classifiers.

Model	Training Loss	Training Accuracy	Validation Loss	Validation Accuracy	Testing Accuracy
VGG16	0.043	0.9869	0.5624	0.865	0.88
ResNet50	0.211	0.9033	0.4618	0.8257	0.86

Confusion Matrix of the test dataset for both the models is shown in Fig 4.4.



(a)



(b)

Fig 4.4. Confusion matrix produced by the deep learning models (a) VGG16, (b) ResNet-50

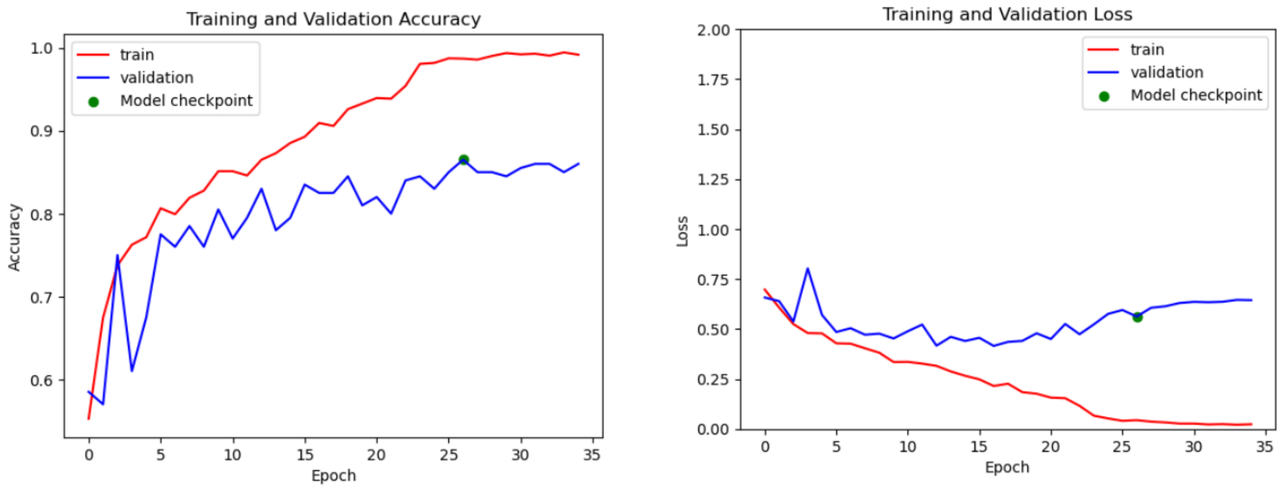
Table 4.2 Classification Report for VGG16.

	precision	recall	f1-score	support
ASD	0.92	0.83	0.87	100
Not ASD	0.85	0.93	0.89	100
accuracy			0.88	200
macro avg	0.88	0.88	0.88	200
weighted avg	0.88	0.88	0.88	200

Table 4.3 Classification Report for ResNet50.

	precision	recall	f1-score	support
ASD	0.92	0.79	0.85	100
Not ASD	0.82	0.93	0.87	100
accuracy			0.86	200
macro avg	0.87	0.86	0.86	200
weighted avg	0.87	0.86	0.86	200

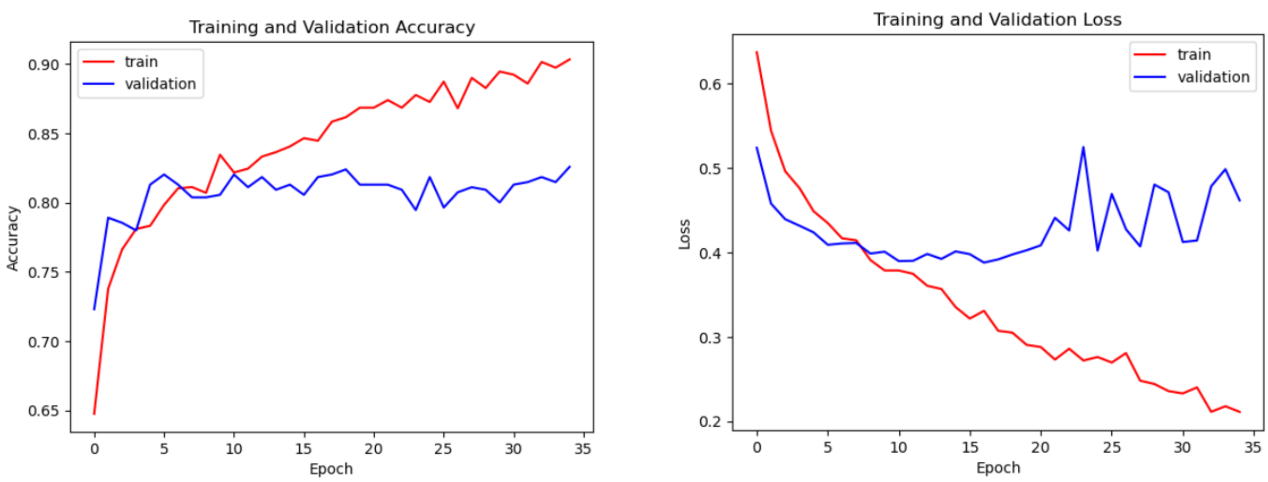
Graphs depicting training and validation loss, as well as training and validation accuracy:
Fig 4.5 shows the training and validation accuracy as well as loss for the VGG16 model.



(a) *(b)*

Fig 4.5 (a) Accuracy graph, (b) Loss graph of VGG16

Fig 4.5 shows the training and validation accuracy as well as loss for the ResNet-50 model.



(a) *(b)*

Fig 4.5 (a) Accuracy graph, (b) Loss graph of ResNet-50

5. Discussion

In our study, we explored the performance of two different classification techniques for our task. We experimented with pre-trained CNN models namely VGG16 and ResNet-50. After thorough evaluation, we found that VGG16 demonstrated the highest testing accuracy of 0.88, followed by ResNet-50 with an accuracy of 0.85.

The results indicate that VGG16 outperformed the other model in terms of classification accuracy. This finding suggests that this particular model is more effective in capturing the intricate features and patterns present in the dataset.

In conclusion, the Q-CHAT-10 questionnaire processing and image classification using the VGG16 model are both included in our system for diagnosing Autism Spectrum Disorder (ASD) by integrating the predictions from the VGG16 image classifier with the features acquired from the survey. This strategy makes use of both behavioral features suggested by the questionnaire responses and visual information from facial expressions. The foundation for data collection, preprocessing, feature extraction, and classification provided by our system architecture makes it easier to create an efficient ASD detection system. With more study and development, this method may help with early detection and intervention in ASD cases, improving outcomes for the affected individuals and their families.

References

[1] Vaishali, R., and R. Sasikala. "A machine learning based approach to classify autism with optimum behaviour sets." *Int. J. Eng. Technol* 7.4 (2018): 18.

[2] Crippa, Alessandro, et al. "Use of machine learning to identify children with autism and their motor abnormalities." *Journal of autism and developmental disorders 45 (2015): 2146-2156.*

[3] Raj, Suman, and Sarfaraz Masood. "Analysis and detection of autism spectrum disorder using machine learning techniques." *Procedia Computer Science 167 (2020): 994-1004.*

[4] K. S. Omar, P. Mondal, N. S. Khan, M. R. K. Rizvi and M. N. Islam, "A Machine Learning Approach to Predict Autism Spectrum Disorder," 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE), Cox'sBazar, Bangladesh, 2019, pp. 1-6, doi: 10.1109/ECACE.2019.8679454.

[5] Angelina Lu and Marek Perkowski, " Deep Learning Approach for Screening Autism Spectrum Disorder in Children with Facial Images and Analysis of Ethnoracial Factors in Model Development and Application", *Department of Electrical and Computer Engineering, Portland State University, Portland, OR 97207, USA;h8mp@pdx.edu*

[6] Allison, Carrie, Bonnie Auyeung, and Simon Baron-Cohen. "Toward brief "red flags" for autism screening: the short autism spectrum quotient and the short quantitative checklist in 1,000 cases and 3,000 controls." *Journal of the American Academy of Child & Adolescent Psychiatry* 51.2 (2012): 202-212.

[7] Thabtah, Fadi. "Machine learning in autistic spectrum disorder behavioral research: A review and ways forward. (2018) " Informatics for Health and Social Care : 1-20.

[8] Hyde, K., Novack, M.N., LaHaye, N. *et al.* Applications of Supervised Machine Learning in Autism Spectrum Disorder Research: a Review. *Rev J Autism Dev Disord* 6, 128–146 (2019). https://doi.org/10.1007/s40489-019-00158-x

[9] Dennis Paul Wall, J. Kosmicki, T.F. Deluca, E. Harstad, Vincent Alfred Fusaro. "Use of machine learning to shorten observation-based screening and diagnosis of autism." Translational psychiatry, 2 (4) (2012), p. e100

[10] Ridha Ilyas Bendjillali, Mohammed Beladgham, Khaled Merit, Abdelmalik Taleb-Ahmed. " Illumination-robust face recognition based on deep convolutional neural networks architectures " Laboratory of TIT, Department of Electrical Engineering, Tahri Mohammed University, Algeria Laboratory of IEMN DOAE. UMR CNRS 8520, University of Valenciennes, France

[11] Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." arXiv preprint arXiv:1409.1556 (2014).

[12] He, Kaiming, et al. "Deep residual learning for image recognition." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.