GREEN UNIVERSITY OF BANGLADESH (GUB)

Computer Vision Based Autism Classification System using Deep Learning

Submitted by

Badhon Parvej (183002111)

Afrin Sufian (183002048)

Sikder Md Mahbub Alam (183002005)

A thesis/project submitted to the Department of Computer Science & Engineering
for the partial fulfillment of the degree of
Bachelor of Science in Computer Science & Engineering

Supervised by

Dr. Muhammad Aminur Rahaman

Associate Professor, Department of CSE

Department of Computer Science & Engineering

Green University of Bangladesh

220/D Begum Rokeya Sarani, Dhaka-1207

November, 2022

Declaration

We hereby declare that this thesis is based on the results found by ourself. Materials of work found by other researcher are mentioned by reference. This thesis, neither in whole nor in part, has been previously submitted for any degree.

Badhon Parvej	Afrin Sufian	Sikder Md Mahbub Alam
ID: 183002111	ID: 183002048	ID: 183002005

Certificate

This is to certify that the thesis entitled **Computer Vision Based Autism Classification System using Deep Learning** has been prepared and submitted by **Badhon Parvej, Afrin Sufian and Sikder Md Mahbub Alam** in partial fulfillment of the requirement for the degree of Bachelor of Science in Computer Science and Engineering on November 03, 2022.

Dr. Muhammad Aminur Rahaman
Supervisor

Accepted and approved in partial fulfillment of the requirement for the degree Bachelor of Science in Computer Science and Engineering.

Prof. Dr. Md. Saiful Azad
Chairman

Mr. Syed Ahsanul Kabir
Member

Mr. Saurav Chandra Das
Member

Md. Riad Hassan
Member

Acknowledgments

We would like to express our greatest appreciation to all individuals who have helped and supported us throughout the project. We are thankful to our supervisor Dr. Muhammad Aminur Rahaman, Associate Professor, Department of CSE, Green University of Bangladesh for his ongoing support during the project, from initial advice, and encouragement, which led to the final report of this project. We would also like to thank Mr. Palash Roy, Lecturer Department of CSE, Green University of Bangladesh who was always there in our computer lab for assistance.

A special acknowledgement goes to our classmates who helped us in completing the project by exchanging interesting ideas and sharing their experience.

We wish to thank our parents as well for their undivided support and interest who inspired us and encouraged us to go our own way, without whom we would be unable to complete our project.

At the end, we want to thank our friends who displayed appreciation to our work and motivated us to continue our work.

Abstract

A complicated neuro-developmental illness called autism spectrum disorder (ASD) impacts communication, language, and social skills.In social interactions,it influences both spoken and nonverbal communication. Early ASD detection and diagnosis are crucial for early educational planning and treatment, family support, and timely delivery of appropriate medical care for the child. It may be possible to develop the best therapeutic strategy at the correct moment with the aid of early ASD screening. By examining facial features, it is possible to recognize ASD by studying the significant markers encoded in human faces. In order to categorize them, we have collected facial images of children with ASD from the public domain and employed deep learning classifiers. it is becoming increasingly important to develop automated diagnosis techniques for ASD. This was accomplished using a dataset of autistic children that contains 2540 facial images of children with autism and typical children. We examined the outcomes of the existing methods and updated the VGG 16 for training and fitting the model. Subsequently, by making major adjustments to the epoch number batch size and other variables, we increased the accuracy from 77% to 86.3% and also calculate precision, specificity, and sensitivity.Additionally, we have enhanced training and testing in comparison to base code. In addition, we have calculated a number of factors to help us evaluate the outcomes more effectively, including true positive, true negative, false positive, and false negative results, as well as accuracy, precision, sensitivity, and specificity of the prediction and result in the submissions.csv file. We have also implemented the vgg19 model of cnn and compared the results of both the models.The highest accuracy result of 88% of VGG 19, followed by VGG16 86.3%.

TABLE OF CONTENTS

List of Figures

List of Tables

Chapter 1

Introduction

ASD is a neurodevelopmental disease that impacts social relationships by affecting both verbal and nonverbal communication. Early ASD detection and diagnosis are crucial for early educational planning and treatment, family support, and timely delivery of appropriate medical care for the child. Social communication is the main emphasis of the revised ASD diagnosis criteria[1]. Mostly because they have communication issues. He or she might not talk as early as other kids. He or she might not want to look people in the eye. Their conduct and emotions are reflected in their body language. They have the same facial pattern as well. Children with autism were compared to typically developing kids who have autistic relatives in the study. Children with autism were shown to have more facial asymmetries than typically developing kids. Facial asymmetry is another indicator of how severe autism symptoms are. ASD often affects people of all racial, ethnic, and socioeconomic backgrounds. Boys experience the disease far more frequently than girls do[2]. Boys are four to five times as likely than females to have ASD. Over the past 20 years, there has been a steady rise in the frequency of ASD; current estimates put the number of affected children at 1 in 36. An increased incidence of ASD has been associated with hereditary variables, parental history of mental problems, preterm births, and prenatal exposure to psychotropic medicines or pesticides. [3][4] The correct therapy strategy might be developed and implemented at the right time with the aid of early ASD screening. DL approaches have been extensively employed throughout the past century to evaluate and categorize medical pictures for the diagnosis of illnesses. These methods aid medical professionals in the detection and identification of illnesses. By examining facial characteristics, eye contact, and head movement, it is possible to recognize ASD in individuals whose faces encode significant signals. In order to better accurately identify children with ASD in the early stages, an enhanced transfer-learning-based autism face recognition framework is presented in this study.

1.1 Problem Domain

Children with ASD are denied the opportunity to learn social skills. They are unable to recognize others' emotions or facial expressions. They don't want to interact with anything on their own. They like to play by themselves and stick to their established schedule. ASD is a chronic disorder that may be quite stressful for both the affected individual and their family. Every ASD child is different and has different experiences. The problem is we can't identify the child is autism affected or not in the early stage. Even if we know that the children are autism affected but still we don't know which level or what type of autism they have. That is why they don't get proper treatment. So by the time when we finally can identify autism, it becomes late to fast improve behavior. Sometimes at first, the doctor cannot suggest Proper treatments. Also, some families cannot understand that their children are autism affected and sometimes it becomes impossible to take their children to hospital because of circumstances.

1.2 Motivation

Autism cannot be detected automatically and sometimes doctors cannot identify correctly. We are developing a system to automatically identify and classify autism child. Doctors don't know how to prevent ASD in children. They do know that it is not caused by whether what a parent does to raise a child, or is it linked to childhood vaccines. But by spotting and treating ASD can reduce symptoms and enhance autism child's normal development which can improve child's quality of life. We are trying to make a Computer Vision Based Autism Classification System and adding deep learning from Eye-Tracking Data of the autistic children for their Behavioral Improvement. For this we are going to use CNN (convolutional neural network) classifier to train our model to detect autism from a facial image.

1.3 Aims and Objectives

Autism spectrum disorder (ASD) is a problem that affects a child's nervous system and growth and development. A child with ASD often has problems communicating. They may have trouble developing social skills. To develop an Image Based Autism Detection System for Autistic children by using Computer Vision and Deep Learning methods.

- Facial images of autistic and non-autistic children are given and train the model using some of the dataset.
- After training testing our model to detect autism from facial image to develop an eye-tracking system for Autistic child to improve their behavior by using Computer Vision and Deep Learning.

1.4 Research Questions

- How can we use make the autism classification system?
- What are the test cases of the system?
- How much accuracy this system can provide?
- How this classification system works for autism children?
- How to improve the accuracy and minimize computational cost?
- What are the better ways to improve the accuracy and minimize computational cost?

1.5 Proposed Idea

In contrast to the conventional method of diagnosis, our idea makes use of a convolution neural network classifier to predict early autism symptoms in children using facial features in photos with a lower cost, short time, and a higher degree of accuracy. Figure 3.1 shows the block diagram of the system architecture of our deep learning model. We have compared our results to the existing result of the authors

1.6 Research Contribution

This thesis tells us about how can we identify autism using facial features in photos at the early age of children and also the power of computer vision.

How people can be benefitted from this system:

- Doctors can identify autism in children.
- It is more accurate than other model.

1.7 Applications

We can use our system to identify autism in many cases :

- Doctors can identify autism in children. so it would be a great help to doctors.
- Parents can monitor their children. .
- It can be implemented in school and children hospital so that autism can be identified automatically.

1.8 Thesis Outline

- Chapter 1: In this chapter discuss introduction of the research, Problem domain, motivation, Aim & objective about the work, research question, Proposed Idea, research contribution and Application of the thesis work.

- Chapter 2: Literature Review as discuss about existing system which were studied. For our research. There we describe the Standard system. comparison with existing work with that system main concept.

- Chapter 3: This chapter is all about system design and working procedure related. From the system design to all steps and process discuss.

- Chapter 4: Experimental result and performance analysis with the proposed systems. There we have taken some data from existing system as training and creating test case to testing the system finally analyze data and shown the result.

- Chapter 5: General discussion and limitation of this work, the practical implementation, limitation and future of our proposed system.

Chapter 2

Literature Review

2.1 Introduction

In this research, we were read some related journals and conference papers, some papers about visualizing the eye-tracking patterns of ASD-diagnosed individuals with a particular focus on children at early stages of development.[5], unsupervised machine learning to discover clusters in Autism Spectrum Disorder[6], analysis and comparison between EEG, Eye and combined data.[7], etc., Previously, several people created these models to visually diagnose autism. They had all adopted various strategies. According to Rad and Furnanello[8], ”the stereotyped motor movement of the patients stood out enough to be observed, although most of the investigations were largely oriented around the social and communicational challenges of the young children with ASD.

2.1.1 Visualizing eye tracking and facial patterns

The aim of this kind of study is to investigate if eye tracking data obtained from persons with and without high functioning autism can be utilized to identify autism[9]. They observe the eye movements of adults with and without autism while they browse web pages in search of information. then they train machine learning classifiers to detect the condition using the captured eye-tracking data[10]. Numerous conditions and illnesses that lack a clinical biomarker run the risk of being misdiagnosed or identified in an advanced stage. Autism Spectrum Disorder (ASD) is one such neurodevelopmental condition that interferes with social interaction and communication. The term ”spectrum” is used to describe the many kinds and amounts of care that various people may need because autism is such a widely varied illness, whereas ”high-functioning autism” denotes a high degree of independence and aptitude.[11][12] Despite having an IQ in the normal range, people with high functioning autism may process information differently, especially in circumstances requiring social engagement, comprehension of semantics and pragmatics, or the transfer of knowledge from one field to another.

2.1.2 Discover cluster in autism spectrum disorder

A lifelong condition marked by social and communication deficits is autism spectrum disorder (ASD). This study seeks to identify clusters in ASD using unsupervised machine learning. The essential concept is to learn clusters based on the eye-tracking scanpaths' visual representation. Utilizing compressed representations developed by a deep autoencoder, the clustering model was trained. The outcomes of the experiments show a positive trend of clustering structure.[13] The clusters are further investigated to offer intriguing insights into the features of the gaze behavior involved in autism.

2.1.3 The contribution of machine learning in ASD

According to the Diagnostic and Statistical Manual of Mental Disorders, autism spectrum disorder (ASD) is a highly complex neurodevelopmental condition with complex etiological causes that affects 1% of the world's population and is characterized by difficulties with social interaction and communication as well as repetitive behaviors and interests[14]. Kanner, who initially introduced it, said it involved ”resistant to change” and ”desire for sameness.” Autism (self) and psychopathy were combined to form Asperger's definition of ASD as ”autistic psychopathy” in. (personality). ASD has a high male-to-female ratio of 4:1 on average, rising sharply to 10:1 in cases of ”high functioning autism” or Asperger syndrome, and falling to 2:1 in cases of co-occurring moderate-to-severe intellectual handicap[15] . Eye gaze loss is a common symptom of ASD, and while it cannot cause autism, it is a crucial part of many diagnostic testing. Along with limited social interaction and communication, restricted, monotonous, and stereotyped conduct is also present. Eye gaze anomalies are linked to both social and non-social cues in people with ASD[16][17]. When it comes to social and facial cues, people with ASD frequently suffer with selectively attending to biological movements, such as body motions and facial expressions, as well as other people's gaze. Individuals with ASD frequently display visual inequalities in their attention to faces, in contrast to individuals who are typically developing[18][19].

2.2 Conclusion

In those different literature review papers with the focus of what features will classified autistic child with typical child and deep learning model. we choose facial features and CNNs to identify autism that will classify autistic child from typical child.

Chapter 3

The Design Methods and Procedures

3.1 Introduction

This study presents a deep learning model based on VGG16 and VGG19 to classify autism using facial features of autistic and typical development children. Facial features can determine if a child has autism or is normal.VGG16 and VGG19 models extracted important facial features from the images. One advantage of deep learning algorithms is the capability to extract tiny details from an image, that an individual can not detect with the bare eye.

3.2 System Design

This figure 3.1 shows the architecture of the proposed system for our model. Our proposed system working on autism classification Autism spectrum disorder (ASD) vs Typically developing (TD). We are finding out the autistic child using the best accuracy result. So we used CNN model for getting better performance and used the Confusion matrix.

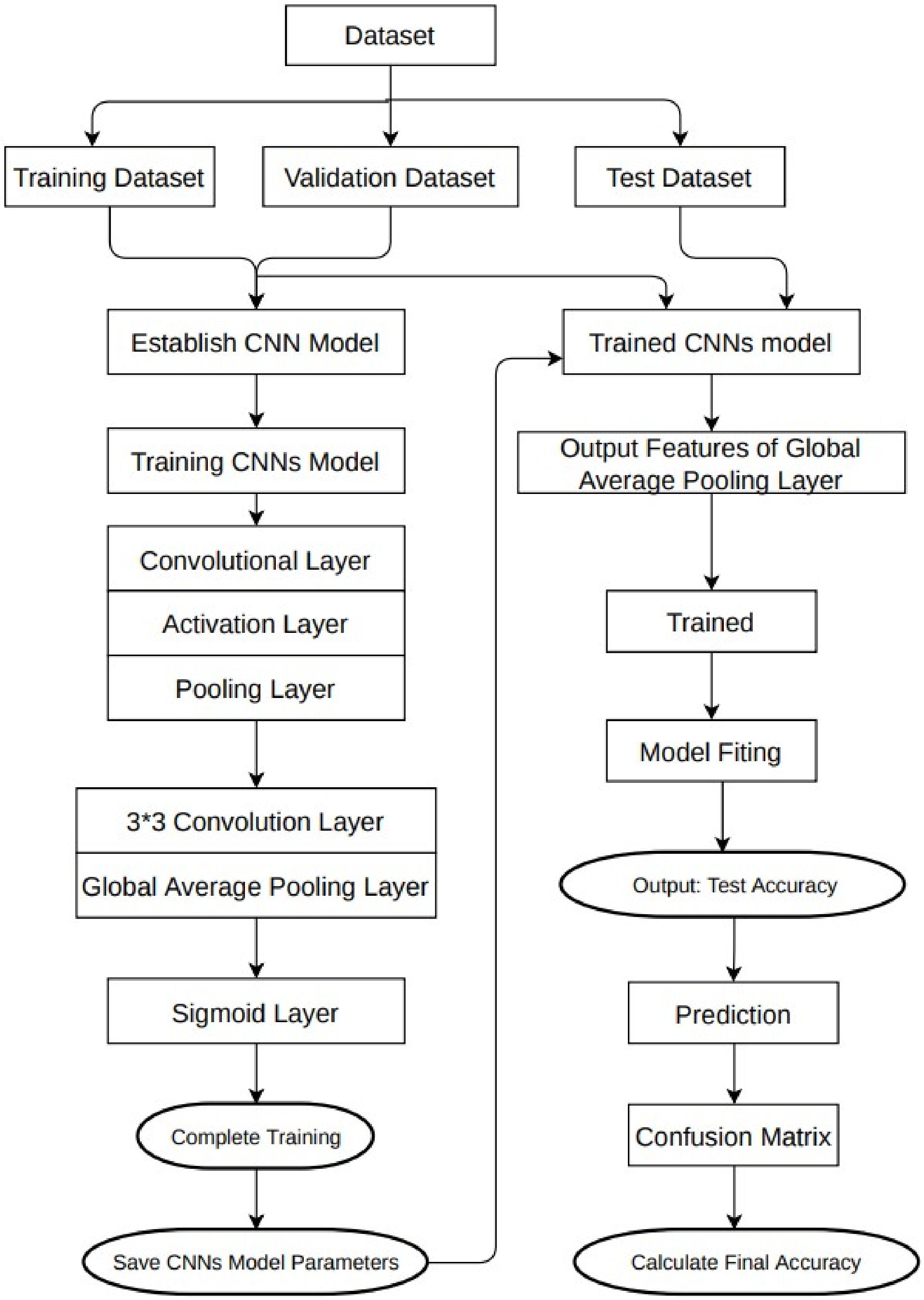

Figure 3.1: System architecture for our proposed model.

3.3 Working Procedure

- **Training**: Dataset for training will start by putting our dataset through training two files in our dataset have the names autistic and non-autistic with the words being separated by a period and the first word indicating the kind whether the kid is autistic or not and the second word denotes the photo number.

- **Testing**: Dataset for load testing has the following file names: ”non-utistic” and ”Autistic.” Multilayer neural networks will be used in place of deep neural networks. Convolutional neural networks will be utilized because the dataset only contains images.

- **Prepairing model**: Then, we’ll set up the model by taking the actions outlined below: Next, we’ll discuss hyperparameters. The output layer is then flattened to one dimension. A completely connected layer with 512 hidden units and ReLU activation will also be added. The dropout rate will then be increased by 0.5. Dropout regularization disables a fraction of the neurons in a hidden layer at random. The percentage of neurons that are deactivated in the preceding layer is determined by the dropout rate of 0.5 that we can apply to the Keras library after any hidden layer. A final sigmoid layer will then be added for classification.

- **Data preprocessing** :Data will be preprocessed before being extracted from photos as a matrix.

- **Training dataset prepairing** : Training Generator: This is where the train data will be created and prepared. Validation Generator: We may ensure the caliber of the performance by using validated data to filter the data. Validation data are those used to evaluate the loss and any model metrics at the end of each epoch. The model won’t be trained with this data.

- **Model fitting**: We will now train the data for a predetermined batch size and number of epochs.

- **Prediction**:We will now use the model that was previously assessed to predict the outcome for the test dataset. The following will be produced: Image name: actual name(image name, 0 or 1), for example, autistic.127.jpg(1); 0/1 is a forecast.

3.3.1 Dataset

Our study compared the facial images of autistic and typically developing kids that were taken from the open source platform, which is accessible to everyone online[20]. In this table 3.1 show that the number of photos 2940 face made up the dataset, half of which were of children with autism and the other half of which were of typically developing kids. This information was compiled using online resources, including Facebook pages and websites dedicated to autism.

Table 3.1: Splitting of the dataset for training, testing, and validation.

Total face images	Training set	Validation set	Testing set
2940	2540	100	300

Children's activities that were captured on camera:

- Recite lines from movies or television shows.
- Chores that are repeated and include tools like pencils or action figures.
- Repetitive motions involving the arms, hands, fingers, and body.
- When talking about mutual interests, successes, or emotions.
- Whilst talking and maintaining eye contact.
- Unusual reactions to scents, tastes, and sounds.
- While playing with others.

3.3.2 Preprocessing

The photos were cleaned up and cropped as part of the data preprocessing. Piosenka's collection of data from online sources necessitated preprocessing before the deep learning model could be trained on them. The face in the original image was automatically cropped by the dataset builder[21]. The dataset was divided into 2,540 training photos, 100 validation images, and 300 testing images. The normalization approach was used to scale; the dataset was rescaling all of the picture parameters from [0, 255] to [0, 1].

3.3.3 Convolutional Neural Network Models

A branch of artificial intelligence known as ”computer vision” has been extensively developed to help humans in their daily lives, such as through medical applications[22]. As a result, the CNN algorithm has contributed to study on behavior and psychology as well as the diagnosis of diseases.

3.3.4 Basic Components of the CNN Model

Undoubtedly one of the most well-known deep learning techniques is the convolutional neural network (CNN). It accepts the input image and assigns weights and biases that can be learned importance in order to determine the image's class. By connecting and communicating with other cells, a neuron can be thought of as a mimic of the contact pattern of the neurons in the human brain. We shall discuss the input layer, convolutional layer, activation function, pooling layer, fully connected layer, and output prediction in this section of the article.

3.3.5 Convolutional Layer with a Pooling Layer

The convolutional layer receives a picture as a matrix of pixel values as input. The convolutional layer's goal is to simplify the images without sacrificing any of the key characteristics that will aid in the detection of autism. Low level features like edges and color are extracted by the CNN model's first layer. We were able to add more layers to the CNN model because of how it was built, allowing it to extract high-level features that will aid with visual comprehension. The number of weights was decreased by utilizing either the max pooling or average pooling approaches because the convolution layer outputs a large number of parameters that could significantly slow down the matrices' arithmetic operations. The maximum values in each window of the stride are the basis for max pooling, whereas the mean value of each window of the stride is the basis for average pooling. The model used in this investigation was based on maximum pooling.

3.3.6 Fully Connected Layer and Activation Function.

The fully connected (FC) layer is a nonlinear collection of high- level features that took input from the hidden layers and were shown as outputs. In the FC layer, the input image is shown as a column vector. The model is composed of the back- propagation pathways and the forward neural network. The forward neural network's output layer is flattened. The neural network in backpropagation undervalues the loss mistakes by using the number of training iterations, and it discovers more features. Most deep learning models perform very well as the number of hidden layers and training iterations are increased, enabling the neural network to effectively extract low-level input . After obtaining the parameters from the FC layer, the softmax classifier calculates the characteristics to anticipate the output, as illustrated. A photograph falls under class 0 if its Softmax output is zero, and class 1 if it has a Softmax output of one. In this study, class 0 denotes an autistic person, whereas class 1 denotes a healthy person.

3.3.7 Deep Learning Models

Our paper is based on two pre-trained models for autism detection using facial feature images: VGG16 and VGG19.

3.3.7.1 VGG16

A convolutional neural network is also comprehended as a ConvNet, which is a type of artificial neural network. A convolutional neural network has an input layer, an output layer, and various hidden layers. VGG16 is a type of CNN (Convolutional Neural Network) show the figure 3.2 that is considered to be one of the best computer vision models to date. The creators of this model evaluated the networks and increased the depth using an architecture with very small (3×3) convolution filters, which showed a significant improvement on the prior-art configurations.

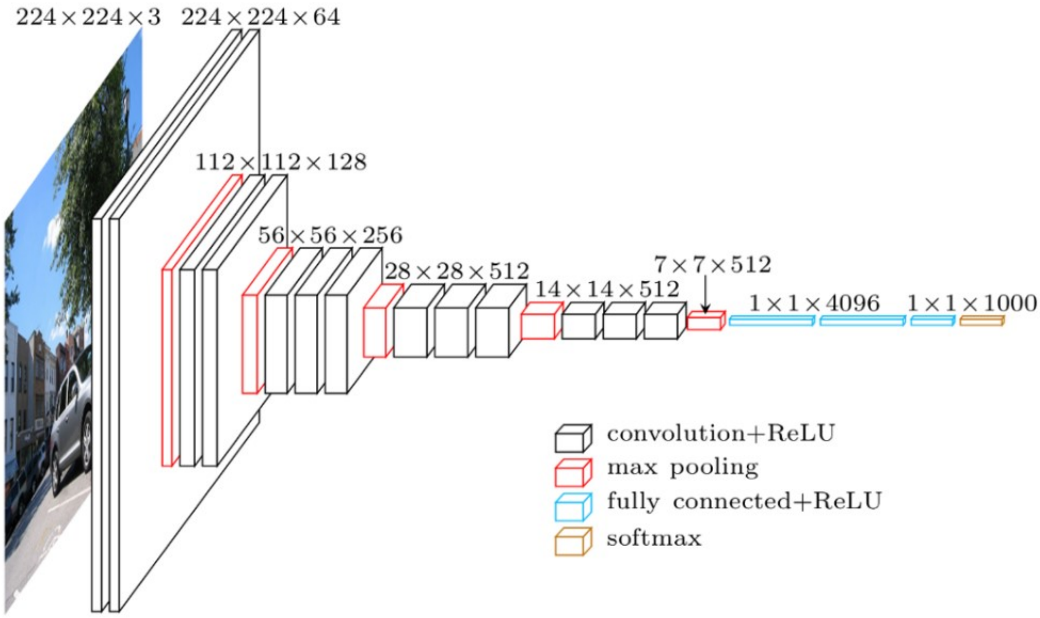

Figure 3.2: VGG16 model architecture.

3.3.7.2 VGG19

VGG19 stands for the visual geometry group network (VGG19), in this figure 3.3 is a in-depth artificial neural network model with a 19 artificial multilayers process. VGG19 is based on the CNN technique, is normally implemented on the ImageNet dataset, and is valuable because of its easy implemenatation, as 3×3 convolutional layers are connected to its upper side to increase with the deep level.

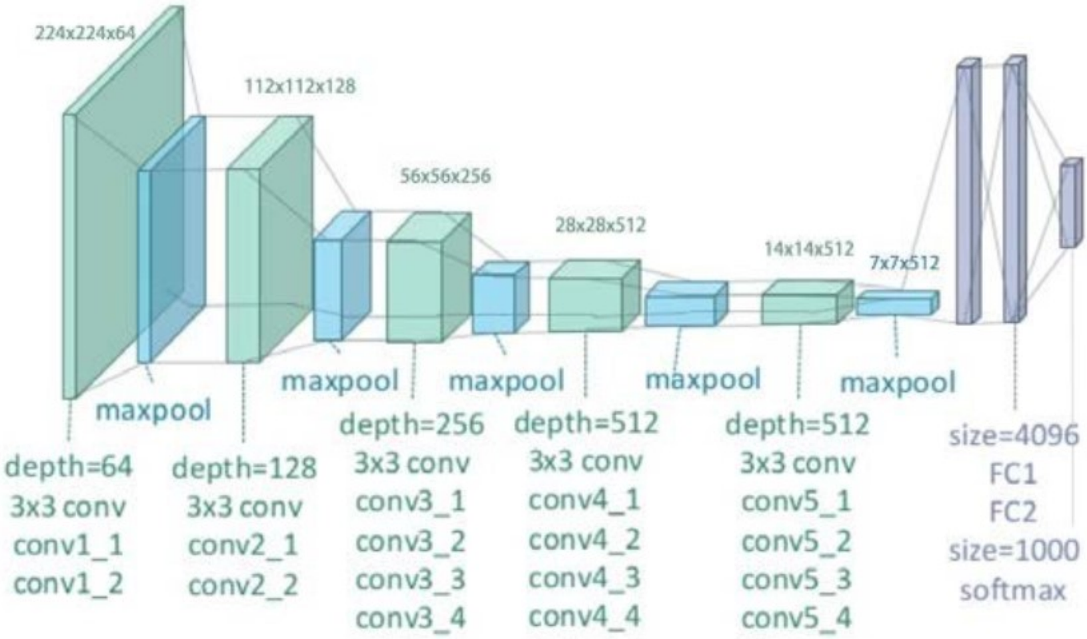

Figure 3.3: VGG19 model architecture.

- Input layer: This layer's purpose is to receive an image input with the dimensions 224*224*3. To keep the input size of the image consistent for the ImageNet race, the model's creators chose the central 224*224*3 patch in each image.

- Convolutional layers: After the convolution stride is held at 1 pixels, the convolutional layers of the VGG use a minimal receptive field of 3*3, which is the smallest practical size that still releases left and right as well as up and down to preserve the spatial resolution. The number of pixel changes across the input matrix is known as the stride.

- Hidden layers: All the hidden layers of the VGG network use ReLU. VGG rarely affects Local Response Normalization due to its longer training time and memory requirements (LRN). It also has no effect on the model's overall accuracy.

- Fully connected layers: There were 4,096 channels total in the first two FC layers and 1,000 channels in the third tier.

3.4 Implement

3.4.1 Train dataset

Figure 3.4 shows the dataset has been divided into two categories: autistic and non-autistic, respectively(0,1).

	filename	category
0	Non_Autistic.1059.jpg	0
1	Non_Autistic.204.jpg	0
2	Non_Autistic.138.jpg	0
3	Non_Autistic.1268.jpg	0
4	Autistic.961.jpg	1

Figure 3.4: Result of train data.

3.4.2 Test dataset

We will now load the test dataset, which consists solely of filenames show in figure 3.5:

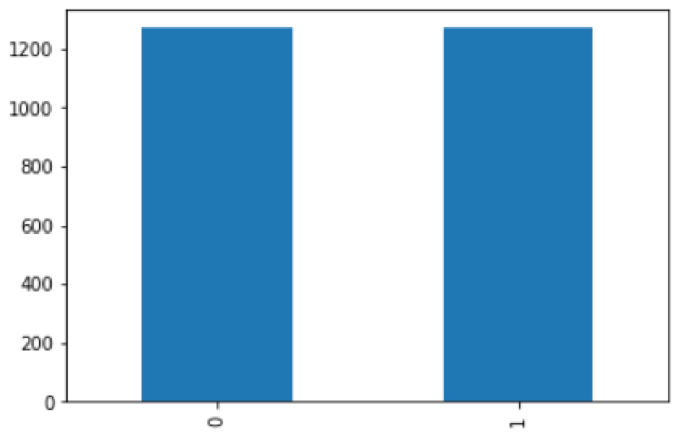

Figure 3.5: Result of test dataset

3.4.3 Sample image

We'll now load the example image as described below in figure 3.6 and 3.7:

	filename
0	Autistic.139.jpg
1	Autistic.4.jpg
2	Autistic.43.jpg
3	Autistic.82.jpg
4	Autistic.89.jpg

Figure 3.6: Sample image list.

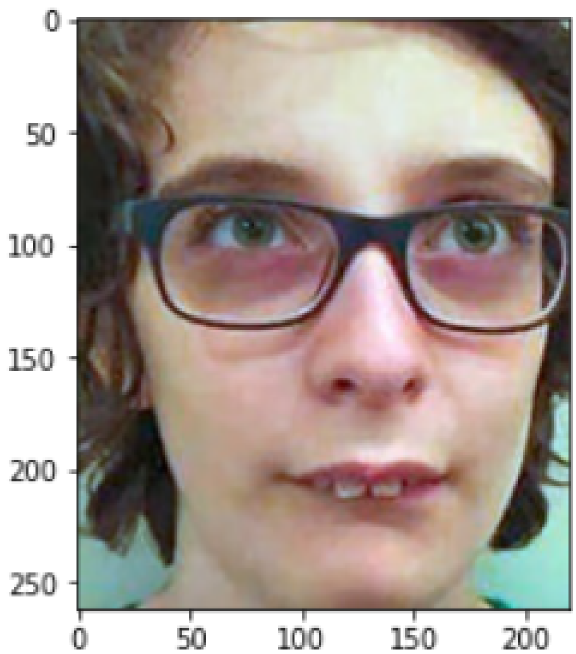

Figure 3.7: Sample image

3.4.4 Model Train

Figure 3.8 shows that our data was trained using block layer. There are total parameters 14,977,857 and trainable parameters is 14,977,857. We see that non-trainable parameters result 0. Figure 3.9 shows that the epoch was run to get the validation accuracy and loss. It is very impact for final accuracy of results.

```
Model: "model"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 input_1 (InputLayer)        [(None, 224, 224, 3)]     0

 block1_conv1 (Conv2D)       (None, 224, 224, 64)      1792

 block1_conv2 (Conv2D)       (None, 224, 224, 64)      36928

 block1_pool (MaxPooling2D)  (None, 112, 112, 64)      0

 block2_conv1 (Conv2D)       (None, 112, 112, 128)     73856

 block2_conv2 (Conv2D)       (None, 112, 112, 128)     147584

 block2_pool (MaxPooling2D)  (None, 56, 56, 128)       0

 block3_conv1 (Conv2D)       (None, 56, 56, 256)       295168

 block3_conv2 (Conv2D)       (None, 56, 56, 256)       590080

 block3_conv3 (Conv2D)       (None, 56, 56, 256)       590080

 block3_pool (MaxPooling2D)  (None, 28, 28, 256)       0

 block4_conv1 (Conv2D)       (None, 28, 28, 512)       1180160

 block4_conv2 (Conv2D)       (None, 28, 28, 512)       2359808

 block4_conv3 (Conv2D)       (None, 28, 28, 512)       2359808

 block4_pool (MaxPooling2D)  (None, 14, 14, 512)       0

 block5_conv1 (Conv2D)       (None, 14, 14, 512)       2359808

 block5_conv2 (Conv2D)       (None, 14, 14, 512)       2359808

 block5_conv3 (Conv2D)       (None, 14, 14, 512)       2359808

 block5_pool (MaxPooling2D)  (None, 7, 7, 512)         0

 global_max_pooling2d (Globa  (None, 512)              0
 lMaxPooling2D)

 dense (Dense)               (None, 512)               262656

 dropout (Dropout)           (None, 512)               0

 dense_1 (Dense)             (None, 1)                 513

=================================================================
Total params: 14,977,857
Trainable params: 14,977,857
Non-trainable params: 0
```

Figure 3.8: Model was prepared with output.

```
Epoch 1/100
190/190 [==============================] - 91s 295ms/step - loss: 0.6936 - accuracy: 0.5686 - val_loss: 0.6888 - val_accuracy: 0.5119
Epoch 2/100
190/190 [==============================] - 50s 264ms/step - loss: 0.6781 - accuracy: 0.5871 - val_loss: 0.7158 - val_accuracy: 0.6071
Epoch 3/100
190/190 [==============================] - 51s 269ms/step - loss: 0.5977 - accuracy: 0.6979 - val_loss: 0.5530 - val_accuracy: 0.7579
Epoch 4/100
190/190 [==============================] - 51s 265ms/step - loss: 0.5402 - accuracy: 0.7318 - val_loss: 0.4618 - val_accuracy: 0.7659
Epoch 5/100
190/190 [==============================] - 53s 275ms/step - loss: 0.4999 - accuracy: 0.7564 - val_loss: 0.4890 - val_accuracy: 0.7778
Epoch 6/100
190/190 [==============================] - 51s 268ms/step - loss: 0.4808 - accuracy: 0.7652 - val_loss: 0.4333 - val_accuracy: 0.8056
Epoch 7/100
190/190 [==============================] - 52s 273ms/step - loss: 0.4776 - accuracy: 0.7863 - val_loss: 0.4207 - val_accuracy: 0.7857
Epoch 8/100
190/190 [==============================] - 51s 270ms/step - loss: 0.4408 - accuracy: 0.7999 - val_loss: 0.4690 - val_accuracy: 0.7738
Epoch 9/100
190/190 [==============================] - 52s 274ms/step - loss: 0.4053 - accuracy: 0.8171 - val_loss: 0.4191 - val_accuracy: 0.7976
Epoch 10/100
190/190 [==============================] - 53s 276ms/step - loss: 0.3996 - accuracy: 0.8245 - val_loss: 0.4805 - val_accuracy: 0.7738
Epoch 11/100
190/190 [==============================] - 51s 266ms/step - loss: 0.4001 - accuracy: 0.8193 - val_loss: 0.3993 - val_accuracy: 0.8056
Epoch 12/100
190/190 [==============================] - 51s 268ms/step - loss: 0.3588 - accuracy: 0.8474 - val_loss: 0.3533 - val_accuracy: 0.8214
Epoch 13/100
190/190 [==============================] - 51s 267ms/step - loss: 0.3519 - accuracy: 0.8505 - val_loss: 0.4168 - val_accuracy: 0.8135
Epoch 14/100
190/190 [==============================] - 52s 271ms/step - loss: 0.3527 - accuracy: 0.8452 - val_loss: 0.4415 - val_accuracy: 0.7897
Epoch 15/100
190/190 [==============================] - 51s 267ms/step - loss: 0.3417 - accuracy: 0.8452 - val_loss: 0.3841 - val_accuracy: 0.8254
Epoch 16/100
190/190 [==============================] - 51s 266ms/step - loss: 0.3252 - accuracy: 0.8602 - val_loss: 0.3865 - val_accuracy: 0.8413
Epoch 17/100
190/190 [==============================] - 51s 266ms/step - loss: 0.2962 - accuracy: 0.8716 - val_loss: 0.5593 - val_accuracy: 0.7659
Epoch 18/100
190/190 [==============================] - 51s 267ms/step - loss: 0.3096 - accuracy: 0.8646 - val_loss: 0.5506 - val_accuracy: 0.7341
Epoch 19/100
190/190 [==============================] - 51s 268ms/step - loss: 0.2873 - accuracy: 0.8755 - val_loss: 0.4620 - val_accuracy: 0.8095
Epoch 20/100
190/190 [==============================] - 52s 272ms/step - loss: 0.2706 - accuracy: 0.8830 - val_loss: 0.4651 - val_accuracy: 0.7698
```

Figure 3.9: Run of 100 epoch

3.5 Conclusion

The heart of our proposed technique relies on a deep learning model. The first is the training part. We use multi-layer neural networks as deep neural networks. We use CNN because we are working with images. Model fitting was carried out. We train the data for a fixed number of batch size and epoch.

Chapter 4

Experimental Result and Performance Analysis

4.1 Introduction

We have proposed VGG16 AND VGG19 as the sole computer vision algorithm for our thesis. The proposed system takes pre-processed data and goes through a number of things.

4.2 Experimental Setup

A test was conducted on various Python libraries and hardware tools are being used to create sophisticated autism diagnosis systems (ASD). Table 4.1 shows the requirements of our system.

Table 4.1: Experimental Setup for our system.

Hardware	Software/libraries
Processor core I5	Keras library
	TensorFlow library
8 GB RAM	Matplotlib
	Numpy

4.2.1 Evaluation Matrix.

For the two pretrained models, this study employs a variety of performance assessment criteria, including a confusion matrix, accuracy, sensitivity, precision, and specificity. An example of a measure of classification performance is a confusion matrix, which is a table containing the true and false values of the test results. In the confusion matrix of the VGG16 model, the True Positives represented 134 of the 150 autistic children, the False Negatives represented 25, the True Negatives represented 125 of the 150 normally developing children, and the False Positives represented 16.In the confusion matrix of the VGG19 model, the True Positives represented 130 of the 150 autistic children, the False Negatives represented 16, the True Negatives represented 134 of the 150 normally developing children, and the False Positives represented 20. The following are the formulae for these measures:

Accuracy= (TP + TN)/(FP+FN+TP+FN)*100%

Specificity= TN/(TN+FN)*100%

Sensitivity= TP/(TP+FP)*100%

Where TN stands for True Negative, FP for False Positive, TP for True Positive, and FP for False Positive. Specificity is the model's ability to recognize normal children, whereas sensitivity is the model's ability to recognize autistic children. In this table 4.2 shows us how many layers we used and the batch size and number of epoch , we also used one optimizer and a activation function.

Table 4.2: Parameters used in the pretrained deep learning models for our system.

Parameter name	Value
Global max pooling layer size	3 * 3
Batch size	12
Number of epochs	100
Dense layer	128, 64
Output classification layer	Softmax
Optimizer	ADAM
Activation function	Rule

4.3 Results

4.3.1 Predicted data for VGG16

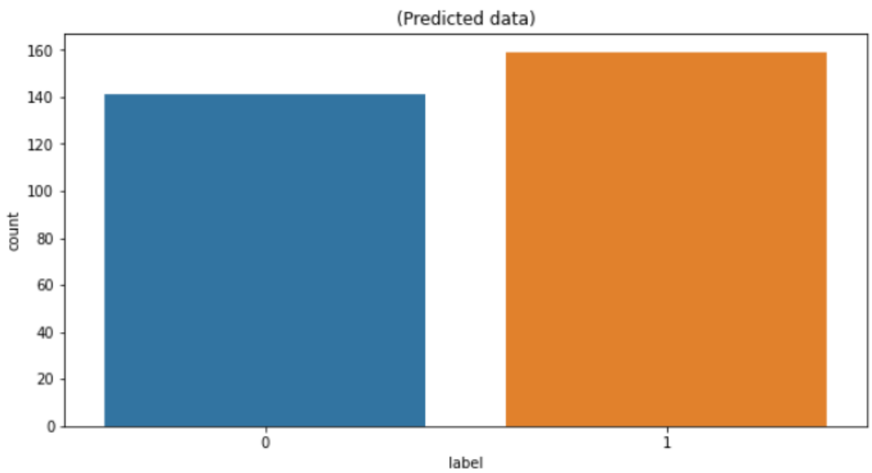

Figure 4.1: VGG16 Predicted Data

```
predicted Autistic :  159
predicted Non Autistic :  141
Actual Autistic :  150
Actual Non Autistic :  150
```

Figure 4.2: Predicted data

In this figure 4.1 and 4.2 show VGG16 out of 300 test data set(actual autistic photo : 150 and actual Non-autistic photo :150) predicted autistic 159 photos and Non-autistic 141 photos.

```
Actual Non Autistic percentage in total test data:  50.0 %
Predicted Non Autistic percentage in total test data:  47.0 %
Actual Autistic percentage in total test data:  50.0 %
Predicted Autistic percentage in total test data:  53.0 %
```

Figure 4.3: Percentage of image classified

Figure 4.3 shows us the percentage output of test data .this gives us an overview of how this algorithm will work on any releated test data. Predicted Non-Autistic percentage :51.3%(actual non-autistic 50%) predicted Autistic percentage : 48.6%(actual autistic 50%)

```
True positive :  134
True Negative :  125
false Positive :  16
false Negative :  25
```

Figure 4.4: Confusion matrix outcomes

Figure 4.4 shows the four confusion matrix value for calculate final accuracy and confusion matrix was employed to forecast accuracy and other outcomes.

- **Accuracy** : 86.33 %
- **Precision** : 89.33 %
- **Sensitivity** : 84.27 %
- **Specificity** : 88.65 %

4.3.2 Predicted data for VGG19

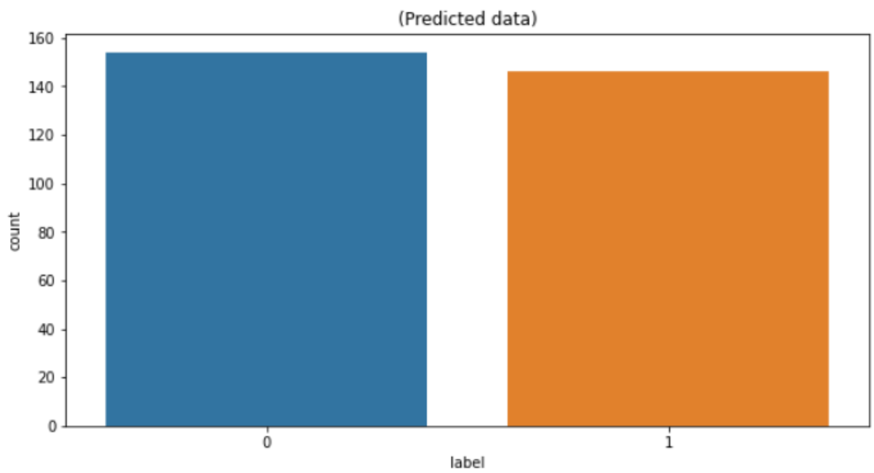

Figure 4.5: VGG19 Predicted data

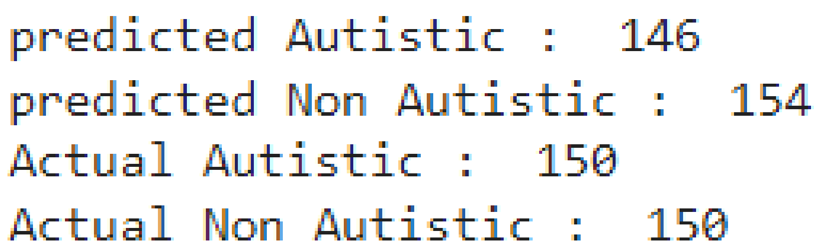
predicted Autistic : 146
predicted Non Autistic : 154
Actual Autistic : 150
Actual Non Autistic : 150

Figure 4.6: Number of photos that were categorized

Figure 4.6 shows VGG19 out of 300 test data set(actual autistic photo : 150 and actual Non-autistic photo :150) predicted autistic 146 photos and Non-autistic 154 photos.

```
Actual Non Autistic percentage in total test data:  50.0 %
Predicted Non Autistic percentage in total test data:  51.33333333333333 %
Actual Autistic percentage in total test data:  50.0 %
Predicted Autistic percentage in total test data:  48.66666666666667 %
```

Figure 4.7: Percentage of image classified

Figure 4.7 shows us the percentage output of test data. This gives us an overview of how this algorithm will work on any related test data. Predicted Non-Autistic percentage: 51.3%(actual non-autistic 50%) predicted Autistic percentage: 48.6%(actual autistic 50%).

```
True positive :  130
True Negative :  134
false Positive :  20
false Negative :  16
```

Figure 4.8: Confusion matrix outcomes

Figure 4.8 shows the four confusion matrix value for calculate final accuracy and confusion matrix was employed to forecast accuracy and other outcomes.

- **Accuracy** : 88.0%
- **Precision** : 86.66%
- **Sensitivity** : 89.04%
- **Specificity** : 87.01%

In this table 4.3 we are showing the comparison Precision Specificity Sensitivity and final Accuracy. In this System we are train different type autistic child and non-autistic child on their facial features. VGG19 achieved highest Accuracy 88.0%.

Table 4.3: Testing results of the pretrained deep learning models for our system.

Model name	Precision	Specificity	Sensitivity	Accuracy
VGG16	89.33%	88.65%	84.27%	86.33%
VGG19	87.01%	86.66%	89.09%	88%

4.4 Result of Proposed System

The test results from the investigations done to look for ASD are shown in this section. The testing outcomes for the used deep learning models are presented in table 4.3. Two separate pretrained deep learning models, namely VGG16 and VGG19, were used in these tests to identify ASD. Each model underwent training and testing to identify the characteristics that, based on face features, classify youngsters as autism or normal. The confusion metrics for the two deep learning models are shown in Figures 4.9 and 4.10.

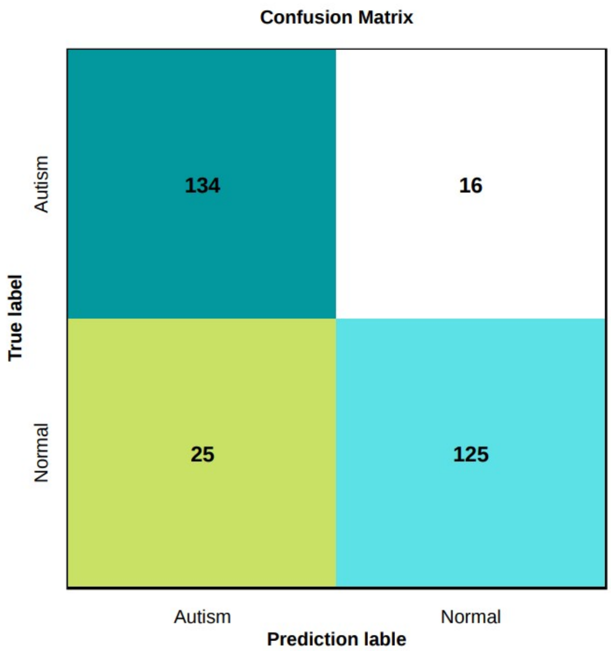

Figure 4.9: Confusion Matrix of VGG16.

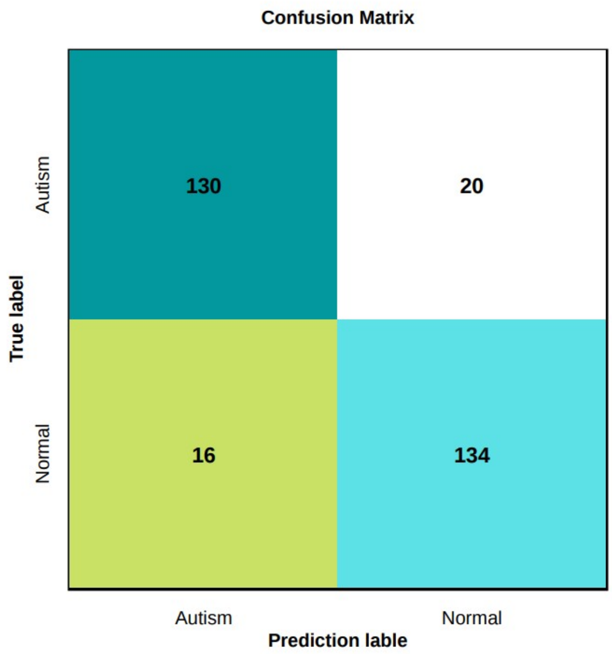

Figure 4.10: Confusion Matrix of VGG19

They demonstrate that among the two models, the VGG19 model had the greatest testing accuracy (88%) while the VGG16 model had the lowest performance level (86.3%). The VGG19 model demonstrated the superior accuracy, with just a tiny proportion of mistakes, despite the dataset being compiled from Internet sources by the data generator, which clearly indicated the difference in ages and the caliber of the photos. Figures 4.12 and 4.14, where the y-axis represents the score percentage and the x-axis the number of epochs, show the performance of the VGG19 model for the training and validation of the data for ASD detection. The VGG19 model's accuracy increased from 56 to 98% during the training phase after 100 iterations, then it decreased to 82% during the validation phase. Figure 4.14 shows its training and validation losses.Figures 4.11 and 4.13 depict how well the VGG16 model performed in identifying ASD. The visual plot demonstrates that the model's performance was subpar. Figure 4.11, with the score% on the y-axis and the number of epochs on the x-axis, displays the accuracy percentages for both training and validation. It performed well, however in comparison to other deep learning models, it performed worse. Figure 4.13 displays its entropy training and validation losses. Figure 4.11 displays how it performs.

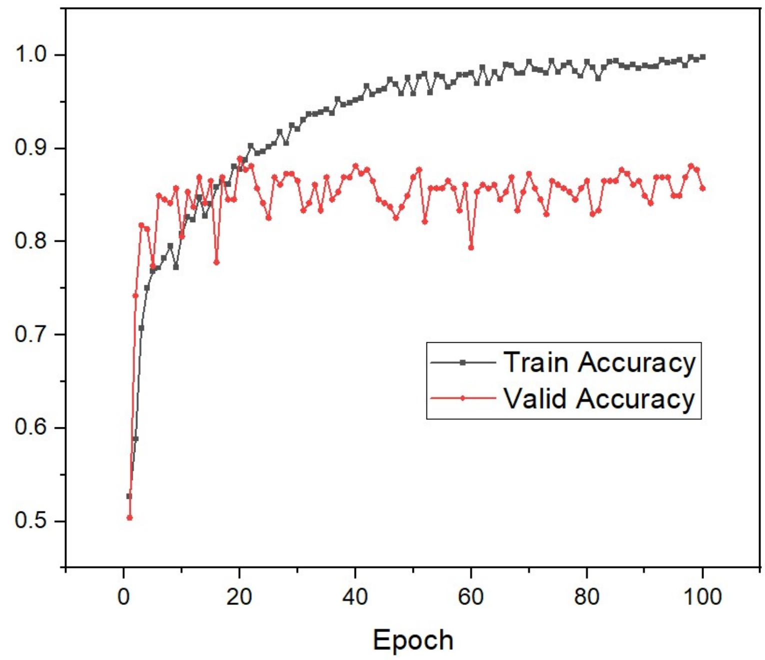

Figure 4.11: VGG16 Train and validation accuracy

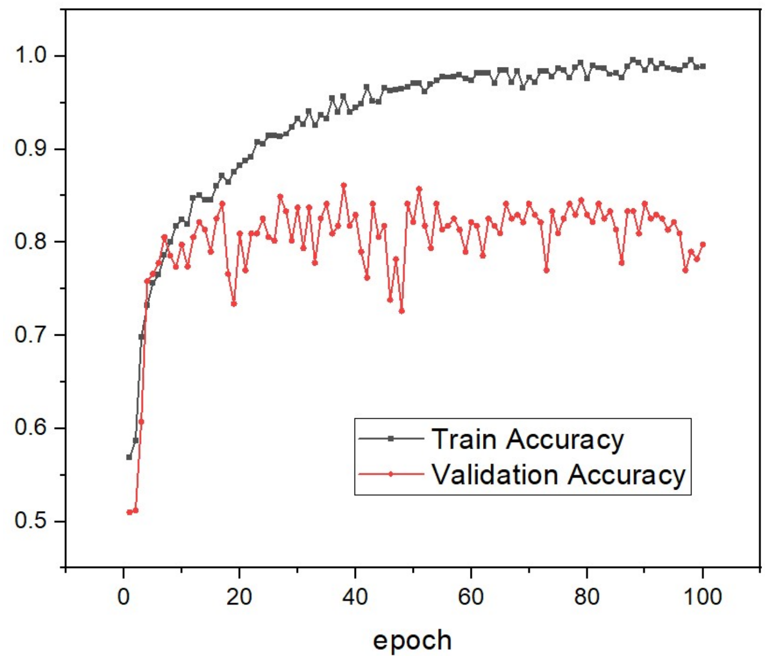

Figure 4.12: VGG19 Train and validation accuracy

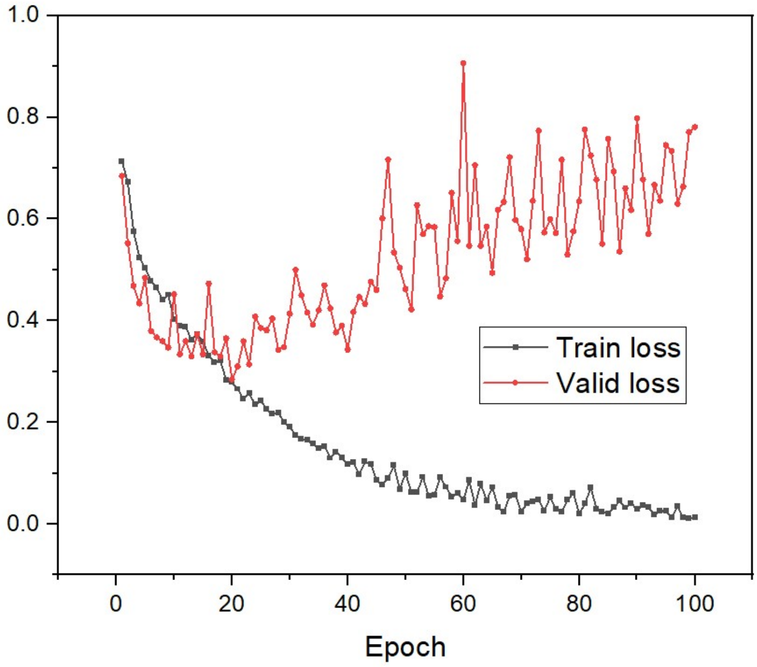

Figure 4.13: VGG16 Train and validation loss

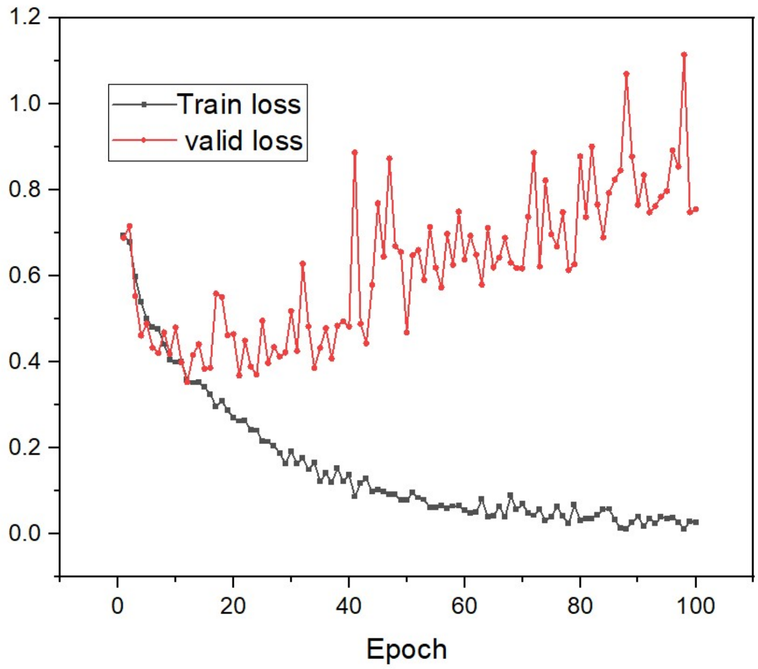

Figure 4.14: Train and validation loss

The status of the comparison between VGG16 and VGG19 is finally shown in figure

4.15. The VGG19 model was shown to be the best deep learning model for diagnosing ASD. The VGG19 achieved heights accuracy that is 88%.

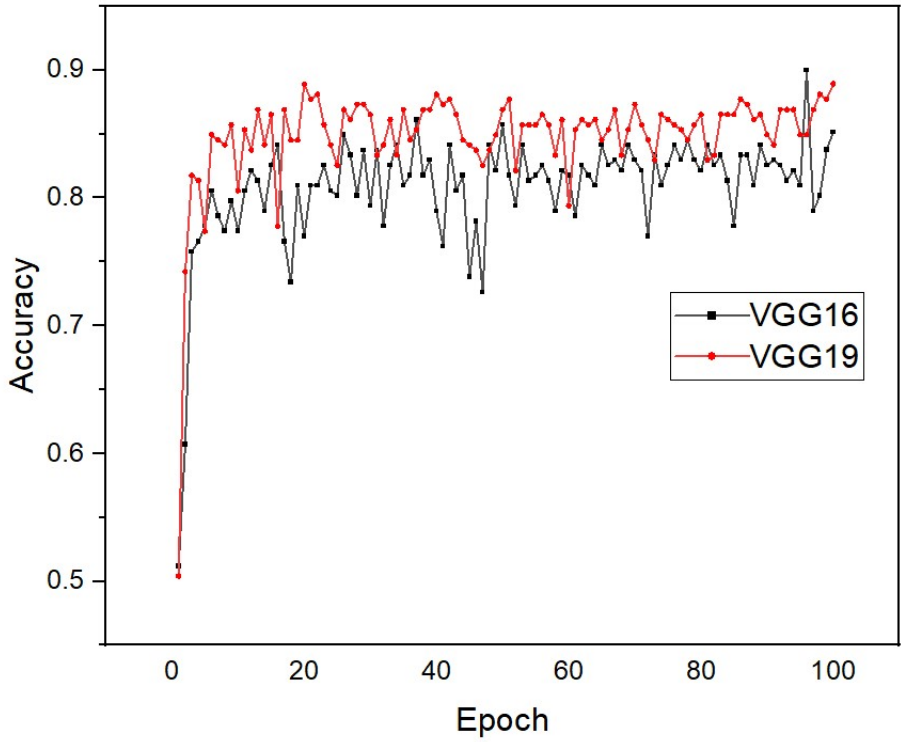

Figure 4.15: Comparison between VGG16 and VGG19

4.5 Conclusion

For the purpose of identifying and categorizing ASD patients, we suggested a CNN (Convolutional Neural Network) architecture in this study. The training period will be shortened since our suggested CNN architecture can achieve better classification performance with less parameters. As a result, when compared to previous models of a similar nature, our suggested model is quicker and less complicated.

Chapter 5

Conclusion

5.1 General Discussion

The medical expenditures associated with autistic spectrum disorder (ASD), a neurodevelopmental condition, are considerable. Medical testing cannot identify autism. A behavioral prediction is therefore necessary. This challenge can be solved with the use of machine learning. In this study, we employ CNN, the most successful machine learning algorithm, to identify individuals with autism. This method provided the best forecast outcome. The most accurate prediction was made using CNN-based models for the ASD dataset. 2940 children's photos from a collection that is divided into children with and without autism were used in this experiment. CNN. The ability to detect a new child with ASD right away has the potential to significantly lessen the nervous system issues in our society.

5.2 Limitation of the research

- We used only image dataset.
- This model can not work head movement and eye tracking features.

.

5.3 Practical Implications

- Doctors can identify autism in children. so it would be a great help to doctors
- Parents can monitor their children.
- It can be implemented in school and children hospitals so that autism can be identified automatically

5.4 Future Works

- We'll try to increase the accuracy and also try to find other deep learning algorithm
- Also try to implement in GUI, so that doctors can easily test.
- There is no research about Bangladeshi autism child , so we will try to gather data and implement deep learning algorithms .

References

[1] F. Tamilarasi and J. Shanmugam. Convolutional neural network based autism classification. pages 1208–1212, 06 2020.

[2] Md Tahmid Zoayed, Sayma Arshe, and Plabon Banik. Autism detecting model using image. 06 2022.

[3] Luciano Nunes, Plácido Pinheiro, Mirian Pinheiro, Monica Pompeu, Marum Simão Filho, Rafael Comin-Nunes, and Pedro Pinheiro. *A Hybrid Model to Guide the Consultation of Children with Autism Spectrum Disorder*, pages 419–431. 10 2019.

[4] Yi L. Liu W, Li M. Identifying children with autism spectrum disorder based on their face processing abnormality: A machine learning framework. pages 888–98, 4 2016.

[5] Peebles D. Thabtah F. A new machine learning model based on induction of rules for autism detection. pages 264–286., 3 2016.

[6] Seyed Reza Shahamiri and Fadi Thabtah. Autism ai: a new autism screening system based on artificial intelligence. *Cognitive Computation*, 12, 07 2020.

[7] Jordan Hashemi, Geraldine Dawson, Kimberly L. H. Carpenter, Kathleen Campbell, Qiang Qiu, Steven Espinosa, Samuel Marsan, Jeffrey P. Baker, Helen L. Egger, and Guillermo Sapiro. Computer vision analysis for quantification of autism risk behaviors. *IEEE Transactions on Affective Computing*, 12(1):215–226, 2021.

[8] James M. Rehg, Gregory D. Abowd, Agata Rozga, Mario Romero, Mark A. Clements, Stan Sclaroff, Irfan Essa, Opal Y. Ousley, Yin Li, Chanho Kim, Hrishikesh Rao, Jonathan C. Kim, Liliana Lo Presti, Jianming Zhang, Denis Lantsman, Jonathan Bidwell, and Zhefan Ye. Decoding children's social behavior. In *2013 IEEE Conference on Computer Vision and Pattern Recognition*, pages 3414–3421, 2013.

[9] Marco Leo, Marco Del Coco, Pierluigi Carcagni, Cosimo Distante, Massimo Bernava, Giovanni Pioggia, and Giuseppe Palestra. Automatic emotion recognition in robot-children interaction for asd treatment. In *2015 IEEE International Conference on Computer Vision Workshop (ICCVW)*, pages 537–545, 2015.

[10] Fernando De la Torre, Wen-Sheng Chu, Xuehan Xiong, Francisco Vicente, Xiaoyu Ding, and Jeffrey Cohn. Intraface. In *2015 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)*, volume 1, pages 1–8, 2015.

[11] Yin Li, Alireza Fathi, and James M. Rehg. Learning to predict gaze in egocentric video. In *2013 IEEE International Conference on Computer Vision*, pages 3216–3223, 2013.

[12] Jordan Hashemi, Qiang Qiu, and Guillermo Sapiro. Cross-modality pose-invariant facial expression. In *2015 IEEE International Conference on Image Processing (ICIP)*, pages 4007–4011, 2015.

[13] T. Ojala, M. Pietikainen, and T. Maenpaa. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(7):971–987, 2002.

[14] Patrick Lucey, Jeffrey F. Cohn, Takeo Kanade, Jason Saragih, Zara Ambadar, and Iain Matthews. The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops*, pages 94–101, 2010.

[15] Tanaya Guha, Zhaojun Yang, Ruth B. Grossman, and Shrikanth S. Narayanan. A computational study of expressive facial dynamics in children with autism. *IEEE Transactions on Affective Computing*, 9(1):14–20, 2018.

[16] Tomas Simon, Hanbyul Joo, Iain Matthews, and Yaser Sheikh. Hand keypoint detection in single images using multiview bootstrapping. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 4645–4653, 2017.

[17] Kyle Krafka, Aditya Khosla, Petr Kellnhofer, Harini Kannan, Suchendra Bhandarkar, Wojciech Matusik, and Antonio Torralba. Eye tracking for everyone. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2176–2184, 2016.

[18] Ming Jiang and Qi Zhao. Learning visual attention to identify people with autism spectrum disorder. In *2017 IEEE International Conference on Computer Vision (ICCV)*, pages 3287–3296, 2017.

[19] Romuald Carette, Mahmoud Elbattah, Gilles Dequen, Jean-Luc Guérin, and Federica Cilia. Visualization of eye-tracking patterns in autism spectrum disorder: Method and dataset. In *2018 Thirteenth International Conference on Digital Information Management (ICDIM)*, pages 248–253, 2018.

[20] Mohammed Saeed Alzahrani Fawaz Waselallah Alsaade. ”classification and detection of autism spectrum disorder based on deep learning algorithms”, computational intelligence and neuroscience. In *2017 IEEE International Conference on Computer Vision (ICCV)*, 2022.

[21] Konstantinos-Filippos Kollias, Christine K. Syriopoulou-Delli, Panagiotis Sarigiannidis, and George F. Fragulis. The contribution of machine learning and eye-tracking technology in autism spectrum disorder research: A systematic review. *Electronics*, 10(23), 2021.

[22] Mahmoud Elbattah, Romuald Carette, Gilles Dequen, Jean-Luc Guérin, and Federica Cilia. Learning clusters in autism spectrum disorder: Image-based cluster-

ing of eye-tracking scanpaths with deep autoencoder. In *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 1417–1420, 2019.

List of Acronyms

ASD	Autism spectrum disorder
TD	typically developing
CNNs	Convolutional neural network
VGG	Visual Geometry Group

List of Symbols

®	Trademark Symbols