Hand gesture recognition or segmentation using a convolutional neural network with synthetic data augmentation for classification

BY

Billal Hossain
ID: 201-15-14344

This Report Presented in Partial Fulfillment of the Requirements for the Degree of Bachelor of Science in Computer Science and Engineering

Supervised By

Dr. Sheak Rashed Haider Noori
Professor & Associate Head
Department of Computer Science and Engineering
Daffodil International University

Co-Supervisor By

Dr. Md Zahid Hasan
Associate Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

JANUARY 2023

APPROVAL

This project, titled "**Hand gesture recognition or segmentation using a convolutional neural network with synthetic data augmentation for classification,**" submitted by **Billal Hossain** to the Department of Computer Science and Engineering, Daffodil International University, has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on January 19, 2023.

BOARD OF EXAMINERS

Dr. Touhid Bhuiyan — **Chairman**
Professor and Head
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

19.01.23

Dr. Md Zahid Hasan — **Internal Examiner**
Associate Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

19.01.23

Saiful Islam — **Internal Examiner**
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

19.01.23

Dr. Md. Sazzadur Rahman — **External Examiner**
Associate Professor
Institute of Information Technology
Jahangirnagar University

DECLARATION

We hereby declare that, this project has been done by us under the supervision of **Dr. Sheak Rashed Haider Noori**, **Department of CSE** Daffodil International University. We also declare that neither this project nor any part of this project has been submitted elsewhere for award of any degree or diploma.

Supervised by:

Dr. Sheak Rashed Haider Noori
Professor & Associate Head
Department of Computer Science and Engineering
Daffodil International University

Co-Supervisor by:

Dr. Md Zahid Hasan
Associate Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Billal Hossain

Submitted by:

Billal Hossain
ID: 201-15-14344
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

First, we express our heartfelt thanks and gratitude to the almighty God for His divine blessing that makes it possible for us to complete the final year project or internship successfully.

We are really grateful and wish our profound indebtedness to Dr. Sheak Rashed Haider Noori, Professor, Department of CSE, Daffodil International University, Dhaka. His guidance and support increase our level of confidence to complete this research project accurately and correctly. Deep Knowledge and keen interest of our supervisor in the field of "CNN" to carry out this project. His endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior drafts and correcting them at all stages have made it possible to complete this project.

We would like to express our heartfelt gratitude to Professor Dr. Touhid Bhuiyan, Professor and Head**,** Department of CSE, for his kind help in finishing our project, as well as to other faculty members and the staff of the CSE department at Daffodil International University.

We would like to thank our entire course mate at Daffodil International University, who took part in this discussion while completing the course work.

Finally, we must acknowledge, with all due respect, the constant support and patience of our parents.

ABSTRACT

Recognition of hand gestures is a crucial component of human connection and communication. Hand gesture recognition can be used to overcome linguistic barriers as well as facilitate user engagement in the human-computer interface (HCI). The scientific community has recently given hand gesture recognition a lot of attention for a variety of applications, such as enhanced driver assistance systems, prosthetics, and robot controls. For example, hand gesture recognition may be used to understand sign language. In this paper, we proposed two models of convolutional neural network such as InceptionV3 and VGG16 for hand gesture recognition with synthetic data augmentation for classification using own build dataset. We will have to put in a lot of work to achieve the goal, including data preparation and cleaning. After pre-processing the dataset, we augment all of the data classification. After augmenting the data for the feed model, we employ a number of classification techniques to identify the pictures and ensure excellent accuracy in the outcome decision. Among them, Inception V3 and VGG-16 are the best outcomes. The presented architectures' classification results exhibit a high degree of accuracy, approaching 99.59%, and VGG16 has an accuracy of 97.21%.

TABLE OF CONTENTS

CONTENTS	PAGE

LIST OF FIGURE

CONTENTS **PAGE**

LIST OF TABLE

CONTENTS **PAGE**

Chapter 1
Introduction

1.1 Introduction

Hand gestures are essential to how we connect with the world around us and play a significant role in nonverbal communication [13]. Also using sign language to express their feelings are young people who are deaf or hard of hearing. People who have hearing loss or are hard of hearing regularly employ hand gestures in their daily lives. In linguistics, bodily language also has this crucial element. The study of HGR has attracted increasing interest from scientists all around the world. Everything is closer to our hands due to technology. Modern technology is improving the methods by which humans and robots communicate. For vision-based static gesture identification combining two datasets, such as the NUS hand gesture and American Sign Language (ASL) datasets, Yong Soon Tan et al. (2021) reported a CNN model that was called EDenseNet. The average accuracy of the EDenseNet grew from 98.50% without augmentation to 99.64% after augmentation [1]. Naoto Ageishi et al. [2] proposed a model DNN for just a camera-based hand gesture recognition system using the American Hand Sign (AHS) dataset and obtained an accuracy of 97.31%, but wide pictures were not acceptable. In the long run, they would like to add sensor fusion.

The most well-known application that has the potential to have an immediate and lasting impact is sign language recognition. This hand gesture detection application's primary goal is to categorize and identify the movements. For this research, I collect a number of data points from several people in different places and prepare my dataset. After I have created my own dataset, it is time to preprocess and augment it. In order to recognize hands, I utilize a number of techniques and principles from other approaches, such as image preprocessing and convolutional neural networks. Applications for hand gesture recognition are many. For instance, we can use sign language to converse with deaf people who are unable to hear.

The number of convolutional neural network algorithms that aid in the detection and identification of the gesture that every method is intended to target This paper presents a number of algorithms and identifies the method that performs the best, offering more accuracy and more prompt and flexible outcomes. To make this discovery, the structure and process of hand movement recognition were evaluated using the CNN transfer InceptionV3 and VGG16 algorithms. A critical turning point in the invention of CNN

detectors were the Inception network. [14] CNN hoped for a great outcome from the start. The network in Inception was complicated. It employed several methods to improve results. As a result of its ongoing development, the network has evolved into several variants, including Inception v1, v2, v3, and ResNet. A convolutional neural network with 48 layers thick is called Inception-v3 [15]. Besides, I also used the VGG16 algorithm. A most important innovation that led to further improvements was VGG16. Convolutional neural network modeling is used. The top-performing classifier in the Imagenet database was found to be VGG16 across all settings.

Additionally, deep learning faces the issue of data scarcity, where a lack of training data results in unsatisfactory generalizations. However, the quality of generalizations based on unobserved data is also influenced by the diversity and volume of training data. This study has a few limitations, such as the fact that it only works with 0 to 9 numerical digits. I would like to learn more about the basic signs of hand gestures.

1.2 Motivation

In general, children with autism often utilize sign language as a technique of communication. To communicate with ordinary people, hand gestures are necessary. I have a cousin who is deaf and hard of hearing. Not only my cousin, but approximately 9.6% of the population in Bangladesh is deaf or hard of hearing. Bangladesh has a high rate of discrimination against deaf people and limited educational opportunities for them. I have felt bad about seeing such children since I was a child, and I have always wanted to do something to help them and get motivated by it. Deaf and mute people can communicate, participate in conversations, study, and generally live their lives like a typical human would by using sign language.

1.3 Rationale of the Study

People are becoming more and more dependent on technology in the present world. We would like to use this cutting-edge technology to help people with hearing and speech impairments live modern lives. For example, a few days ago, when I went to my house to meet my cousin, I saw he had made a video call with his friend who lives outside of Bangladesh. After that, I heard he has a number of friends from abroad, and he always communicates with them through his mobile phone using HGR. Our key objective is to enable them to interact with all regular people and participate in all of our activities. In our society, we aim to make these individuals assets rather than burdens.

1.4 Expected Output

There are several systems for recognizing hand gestures. I would like to improve the suggested system's height accuracy by using data augmentation. I want to train my dataset using data segmentation, then use the convolutional neural network model to get the maximum performance.

1.5 Report Layout

First of all, this article begins with some primary or vital facts. Besides, six chapters altogether were included in our report. Every chapter is read from a different angle, and each portion is broken down into multiple components that are all thoroughly explained. The topics of this research paper are as follows:

The introduction, inspiration, purpose of the study, and expected outcome are all presented in Chapter 1 about hand gesture recognition or segmentation using a convolutional neural network with synthetic data augmentation for classification.

Then in Chapter 2, the principal overview of the literature is discussed, along with related work in the field of hand gesture recognition. After that, the third chapter, "Research Methodology," contains a description of the research subject, instruments, data collection process, and statistical analysis, as well as the requirements for implementation. The experimental setup, the experimental results, and the analysis will all be in Chapter 4. Discussion and test result. Last but not least, chapter five has a summary of the study, a conclusion, and implications for future research.

Chapter 2
Background

2.1 Preliminaries

Reading pertinent literature reviews serves as an introduction to any research topic. We started our investigation by reading relevant literature reviews. Our study focuses on classifying hand gestures using a convolutional neural network and synthetic data augmentation. The most natural way for people to communicate is through gestures. Today, gesture recognition is employed for HGR. There is a plethora of work on hand gesture recognition nowadays. Before categorization, a hand gesture recognition technique often comprises numerous pre-processing phases. In order to extract significant characteristics, the input picture is often pre-processed using image processing techniques like filtering or segmentation. The features are then retrieved using manually created feature extraction algorithms from the preprocessed picture.

2.2 Related Works

Several studies of hand gesture identification have been undertaken over the past few decades. Gibran Benitez-Garcia et al. [3] wanted to avoid optical flow computation and RGB frames combined with hand segmentation masks. They used a lightweight semantic segmentation method (FASSD-Net) to get high accuracy from Temporal Segment Networks (TSN) and Temporal Shift Modules (TSM), useful HGR methods, by using the IPN Hand dataset. They would like to plan to merge TSN and TSM into one architecture. Priyanka Parvathy et al. [4] proposed a vision-based HGR using machine learning that included segmentation, feature extraction, and classification. They used a database of Sebastian Marcel's static hand gestures, with 96.5% accuracy within 0.024 s. Muni Oudah et al. [5] worked on several reviews of HGR papers, focusing on computer vision techniques, classification algorithms and drawbacks, the number and types of gestures, datasets, detection areas, camera types, segmentation, and so on. The paper by MUNEER AL-HAMMADI et al. [6] uses several deep learning architectures such as hand segmentation, local and global feature representations, and sequence feature globalization and recognition for dynamic hand gesture recognition. They deployed it to assess a tough and actual sign language dataset. To optimize the duration of the input clip and test the system for real-time hand gesture detection, they will eventually employ different temporal aspect modeling methodologies and conduct a large number of trials. In terms of recognition rate, the outcome exceeded cutting-edge techniques, proving its usefulness.

A framework to categorize hand-gesture signatures produced by an ultra-wideband (UWB) impulse radar was suggested by SRUTHY SKARIA et al. [7] using deep-learning algorithms and comparing four different categorization architectures. Above 96% accuracy was achieved, and even when comparable movements were used, there was very little chance of error. MUNEER AL-HAMMADI et al. [8] presented a 3DCNN technique that used transfer learning to overcome the lack of three labeled datasets for hand gesture identification. They obtained two accuracy modes: the signer-dependent mode yielding 98.12%, 100%, and 76.67% accuracy, and the signer-independent mode yielding 84.38%, 34.9%, and 70% accuracy.

Yuanyuan SHI et al. [9] focused on the feature extraction technique used to extract the spatiotemporal structural information from a video sequence and examined the benefits and drawbacks of current technology while also addressing likely new study areas.

Abdullah Mujahid et al. [10] introduced a lightweight model for gesture detection that does not need preprocessing, picture filtering, or image augmentation. It is built on YOLOv3 and DarkNet-53 convolutional neural networks. The suggested model recognized motions accurately except in a challenging context and minimal picture mode. They used a data model in the YOLO and Pascal VOC formats. a YOLOv3-based model with a corresponding F-1 score of 97.68, 94.88, 98.66, and 96.70% for accuracy, precision, recall, and score. By comparing their model to the Single Shot Detector and Visual Geometry Group, they found that their accuracy ranged from 82 to 85%. In the long run, they will concentrate on hybrid approaches with intelligent mobile applications. A team led by Bin Hu and Jiacun Wang [11] developed a method for controlling UAVs in flight where a deep-learning neural network serves as the system's central component. They each created their own basic dataset. They obtained an average accuracy for a 2-layer completely connected neural network in static analysis of 96.7% on scaled samples and 12.3% on non-scaled datasets. The 5-layer fully connected neural network succeeds after that, achieving an accuracy of 98.0% on scaled datasets and 89.1% on non-scaled datasets. The 8-layer neural network convolutional produces an accuracy of 89.6% on scaled datasets and 96.9% on non-scaled samples.Yong Soon Tan et al. [12] developed EDenseNet for vision-based HGR by combining three datasets: two ASL datasets and one NUS dataset. They were acquired with an average accuracy of 98.50% without enhanced data and 99.64% with augmented data.

2.3 Summary of Research

The primary goal of this study is to advance the convolutional neural network area. In this work, we suggest two methodologies: Inception V3 and VGG16. Sign language's nine digits are 0 to 9. We collected 23,153 hand gestures. Our dataset is unique since it was manually built while keeping both genders. For the dataset, we created 10 classes, namely 0, 1, 2, 3, 4, 5, 6, 7, 8, and 9. After that, we cleaned up our dataset by removing noise and other elements. During the model development, the dataset is divided into three files: train, test, and valid. After completing each stage, it's ready to start the model training process. Lastly, the outcomes of our experiment were consistent with the models we suggested.

2.4 Scope of the Problem

The study being done on this topic is quite comparable to prior studies on hand gesture recognition. We are performing hand gestures using the digits. We are facing a number of issues, including the fact that people are unwilling to give their hands to click for data collection.

2.5 Challenges

Without taking on challenges, we cannot succeed. We always welcome every challenge. The majority of individuals are uncomfortable using hand gestures; it was difficult to compile a collection of them. We spent a lot of time collecting the right dataset, as this was one of our main issues. We had considerable difficulty deciding which model to use for the image processing system.

Chapter 3

Research Methodology

3.1 Introduction

Every research study, including this one, uses a certain approach. In this section, I would like to describe my work procedure part by part in depth. The goal of the study project is to learn something new by using creative problem-solving techniques. To begin, I present how to collect raw data from the environment, filter out noise, and segment it. What technique I used to manipulate the data, how the input worked, how the outputs were statistically analyzed, and what was needed for implementation. The dataset must undergo some pre-processing before being fed into any model. We will go through every step of the process in this part, from gathering data to receiving the desired outcomes. We will really need the right dataset in order to get an accurate result. For the purpose of achieving the best outcomes, we have constructed distinct datasets depending on nine hand gestures. Figure 3.1.1, which summarizes the overall study effort, contains the entire work method.

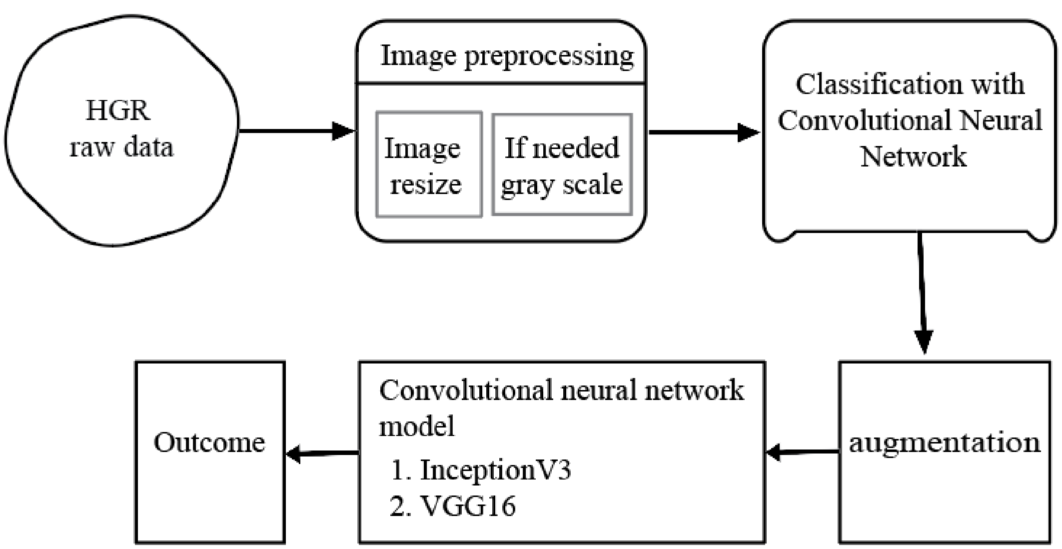

Fig: 3.1.1 Architecture of hand gesture recognition

3.2 Research Subject and Instrumentation

The proposed study title is “**Hand gesture recognition using a convolutional neural network with synthetic data augmentation for classification."** First of all, we are working with a 0–9 digit. To take data from the environment and train the machine, At the moment, a lot of research is being done in the field of convolutional neural networks. Because our suggested model contains various mathematical functions, we implemented our study procedure using a number of CNN algorithms. It will be quite challenging for us to work with a convolutional neural network model. Because a powerful PC, GPU, and other equipment are required. we needed a number of instruments, such as

1. Google Colab
2. 12GB RAM with Intel Core i3
3. 120 GB SSD
4. 1TB hard disk
5. Windows 10 Enterprise
6. NumPy
7. Python 3.8,
8. Tensor flow
9. keras
10. Jupiter notebook and so on.

3.3 Data Collection Procedure

Currently, there is a lot of research being done on hand gesture recognition. The most important part of research is data collection; without data or a dataset, we cannot do any kind of research or work. As a result, for our proposed research, we collected several data. But we did not employ any datasets or resources. In this section, we gathered information on both genders with the assistance of our society, such as men and women. We use our phone's camera to take pictures in sign language and with naked hands against various backgrounds. Then take pictures of as many individuals as possible holding their hands in 10 various signs. We collected 23,153 hand gestures. Our dataset is unique since it was manually built while keeping both genders. For the dataset, we created 10 classes, namely 0, 1, 2, 3, 4, 5, 6, 7, 8, and 9.

Table 3.3.1 Data Size of each class

Name of sign Image	Amount of Images
Zero	2321
One	2325
Two	2307
Three	2315
Four	2311
Five	2322
Six	2318
Seven	2326
Eight	2317
Nine	2324

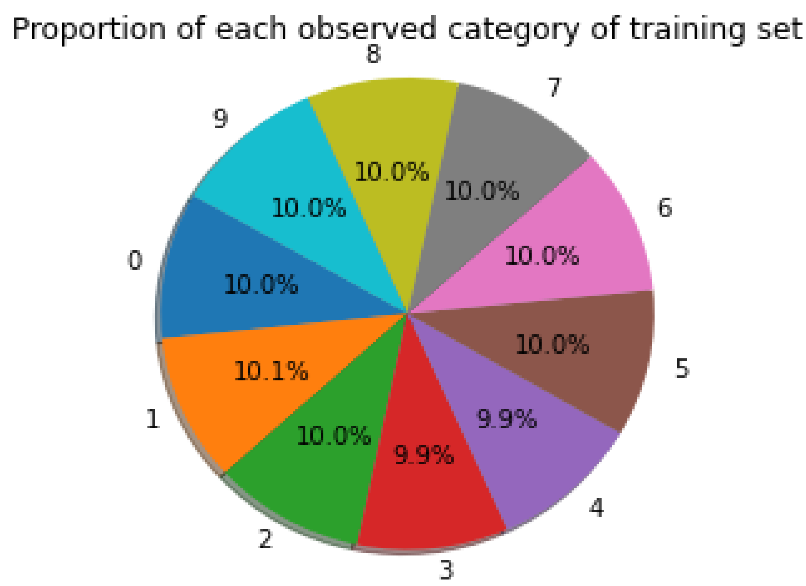

Fig. 3.3.2 Pie Chart: Proportion of Train Data

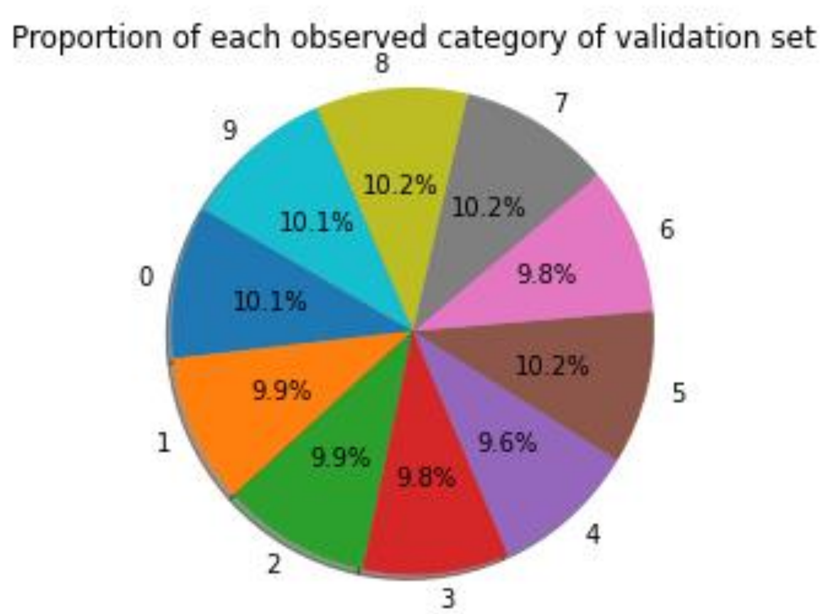

Fig. 3.3.3 Pie Chart: Proportion of Validation Data

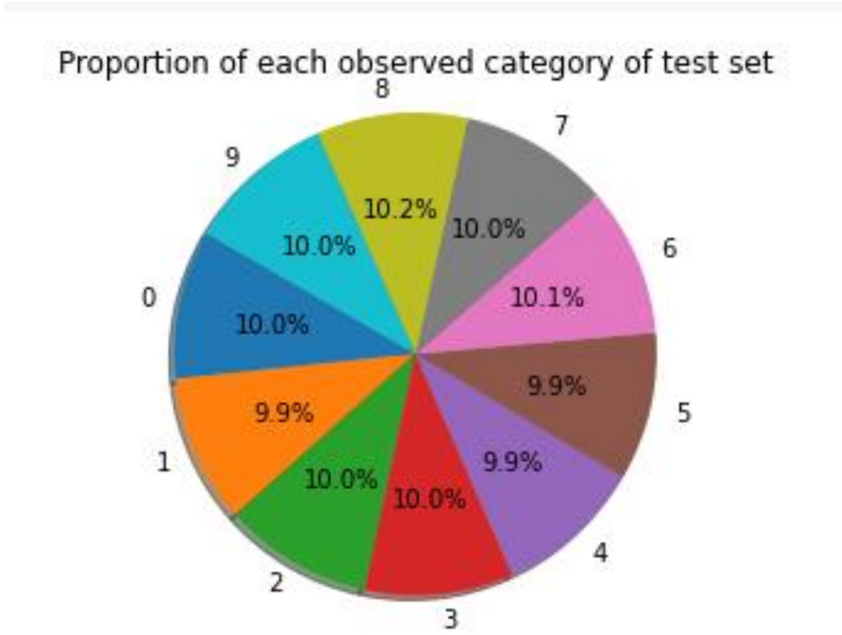

Fig. 3.3.4 Pie Chart: Proportion of Test Data

3.4 Data Preprocessing

Preparing original data to be used by a machine learning algorithm is a process known as data preparation. The first and most important stage in developing a machine learning algorithm is this one. The resolution of the picture frames is increased using image processing techniques. The best data from the raw dataset was chosen. then employ ImageDataGenerator from Keras for image preprocessing. There are several properties, like rescale, rotation, width-shift-range, shear, zoom, horizontal, and so on. For image resizing, we use a width and height of 250 pixels.

Table 3.4.1 Amount of all types of data

Name of Sign Image	Training Data	Validation Data	Testing Data
Zero	1615	238	468
One	1626	233	466
Two	1605	233	469
Three	1600	230	468
Four	1597	226	463
Five	1617	240	465
Six	1618	230	472
Seven	1619	240	470
Eight	1604	240	476
Nine	1619	237	469

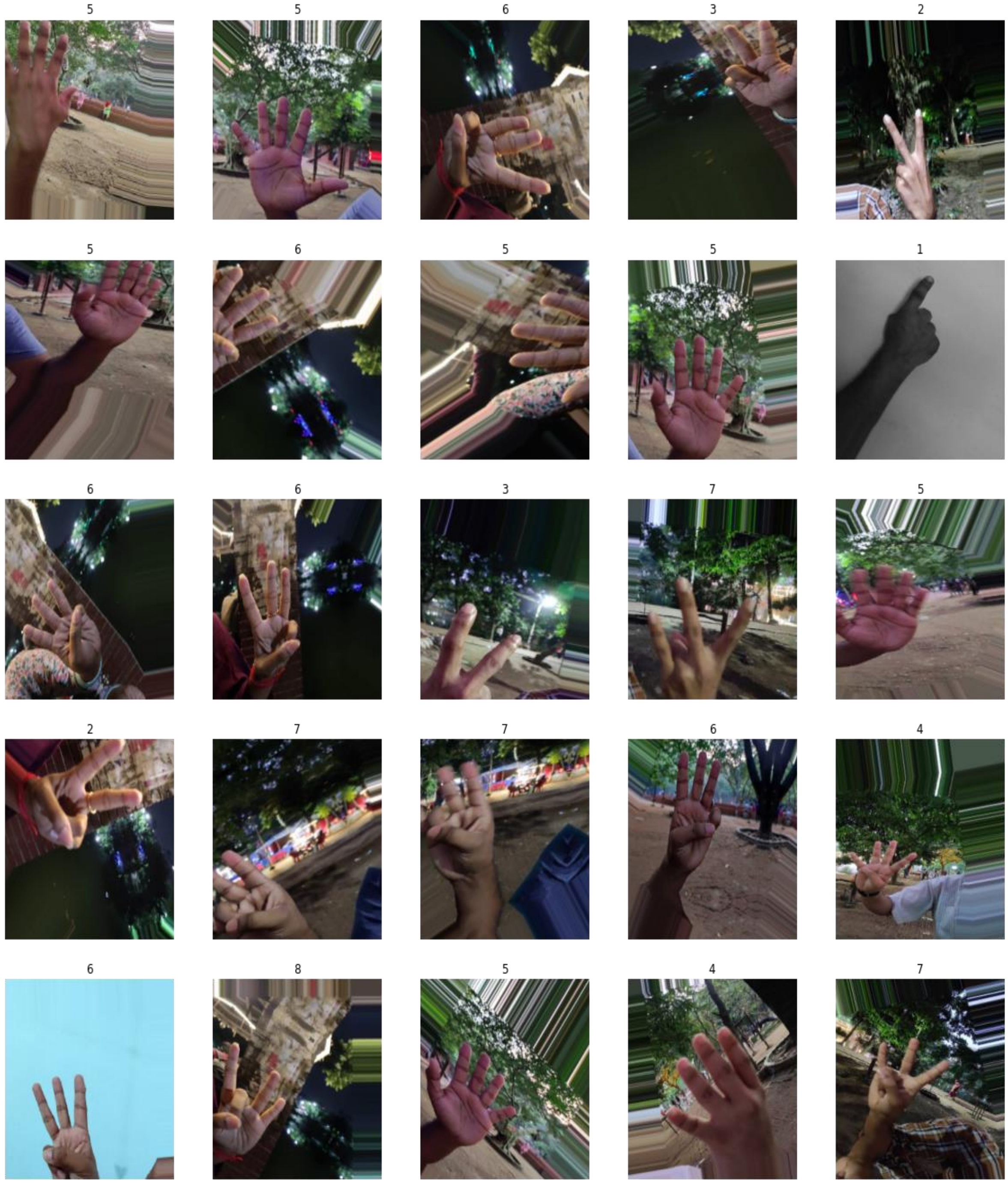

Fig. 3.4.1. Preprocessed Data

3.5 Statistical Analysis

Around 23,153 hand gestures for different phrases were acquired for this investigation. In this experiment, we utilized 16120 samples for training, 2347 samples for validation, and 4686 samples for testing.

3.6 Proposed Methodology

There are several deep learning techniques that may be used to create convolutional neural network models.

Chapter 4

Experimental Results and Discussion

4.1 Introduction

We will go through the two algorithms used in this project in detail before moving on to this chapter. Several convolutional neural network techniques were used in our research to put our findings into practice. In this area, you may see how various models are put to use, how they operate, and how their results are presented graphically.

4.2 Experimental Setup

After putting all of the suggested techniques into practice, we evaluate which model provides the greatest accuracy and is most appropriate for the recommended convolutional neural network techniques. Not only we are seeking the highest level of accuracy, but also the best solution in terms of numerical output. We evaluate the accuracy score, recall, and precision value to identify the optimal model to reflect our offered methods.

Accuracy: The segmentation model's accuracy score, which is related to the model's percentage of accurate predictions, was used to evaluate the model's value.

Accuracy = Accurate Predictions / Total Predictions.

Precision: A genuine, accurate prediction based on observation's accuracy rate is what is meant by precision. It also goes by the name "PPV" (Positive Predicted Value).

Precision = True Positive / (True Positive + False Positive)

Recall: Recall is calculated by dividing the number of true positives by the sum of all true positives and false negatives in the dataset. It serves as a gauge for how comprehensive the model is.

True Positive / (True Positive + False Positive) = Recall

4.3 Experimental Results & Analysis

Convolutional neural networks are designed to evaluate how much better than average the performance of the model in use is. The classification techniques can absolutely produce exact results based on class and attain exceptional accuracy. After testing, we discovered a variety of results with acceptable accuracy. We are all aware that no computer can ever deliver 100% accurate results. The model we used for our project provided accuracy, which was good; however, it wasn't always correct. We will examine and graphically display our experimental findings in this area.

4.3.1 Inception V3 model:

To obtain good results using our custom-built dataset. We analyzed two convolutional neural network techniques, among them Inception v3, which we now show here.

Table 4.3.1.1 Training Report

Algorithm Name	Epoch	Precision	Recall	Accuracy
Inception V3	1	0.9692	0.9048	0.9342
	2	0.9888	0.9847	0.9866
	3	0.9916	0.9901	0.9908
	4	0.9930	0.9925	0.9926
	5	0.9936	0.9930	0.9933

Table 4.3.1.2 Validation Report

Algorithm Name	Epoch	Precision	Recall	Accuracy
Inception V3	1	0.9940	0.9919	0.9923
	2	0.9915	0.9915	0.9915
	3	0.9979	0.9974	0.9974
	4	0.9966	0.9966	0.9966
	5	0.9957	0.9957	0.9957

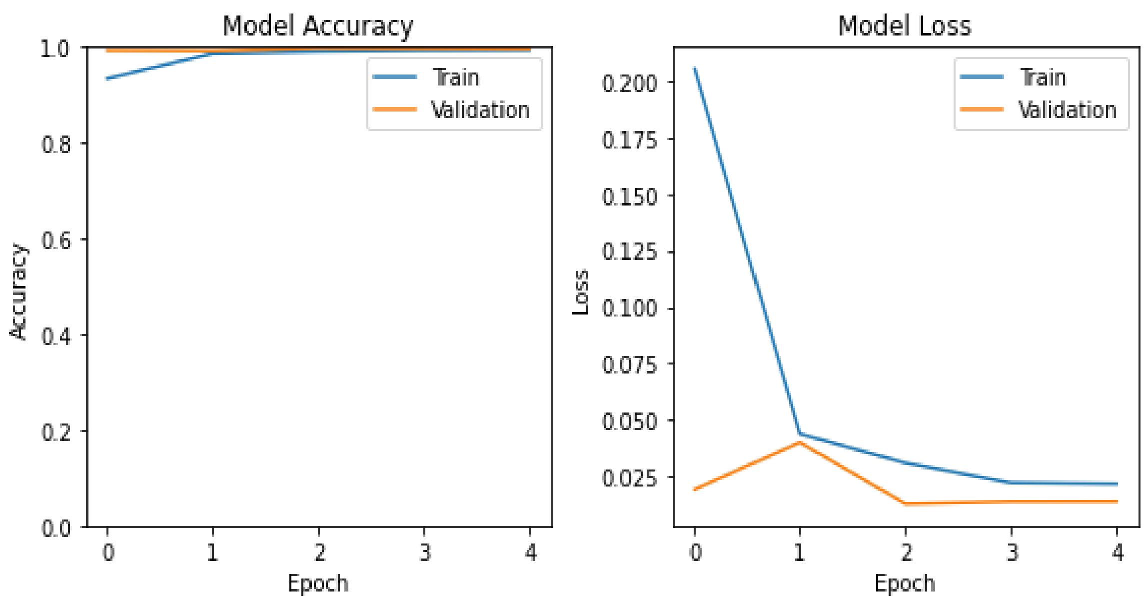

Figure: 4.3.2.1 plots the training and validation accuracy and loss values of Inception V3

Finally, Test Loss: 0.01, Test Precision: 1.00, Test Recall: 1.00, Test Accuracy: 99.59%.

4.3.2 VGG16

Table 4.3.2.1 VGG16 Report

Algorithm Name	Epoch	Training Accuracy	Testing Accuracy
VGG 16	1	0.8800	0.9497
	2	0.9783	0.9501
	3	0.9873	0.9783
	4	0.9898	0.9744
	5	0.9893	0.9732

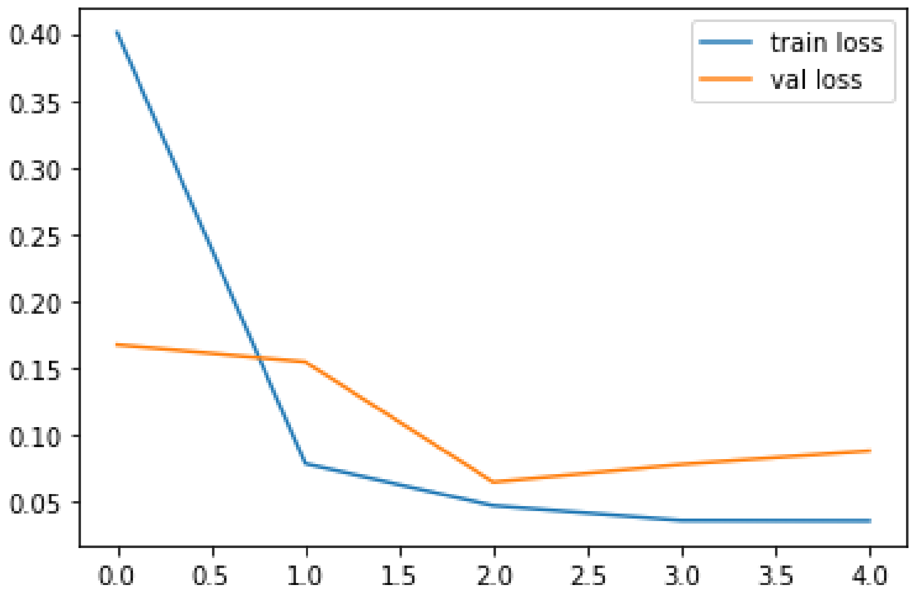

Figure: 4.3.2.2 plots the training and validation loss values of VGG 16

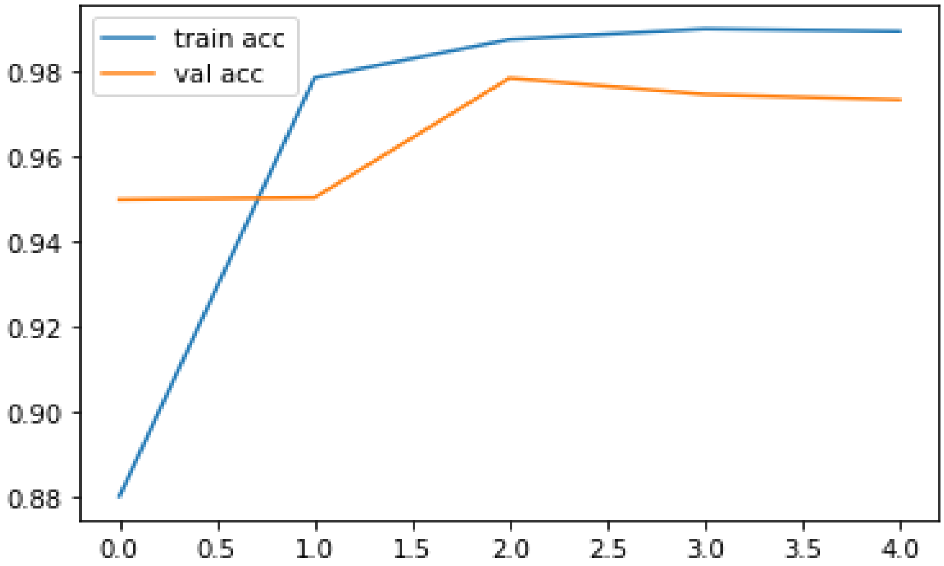

Figure: 4.3.2.3 plots the training and validation accuracy value of VGG16

Table 4.3.2.2: Table of Comparative Results between Inception v3 and VGG16

Algorithm	**Accuracy**
Inception V3	99.59%.
VGG16	96.51%

4.4 Discussion

In this study, we presented two models, Inception v3 and VGG16. On the other hand, we have applied and obtained outcomes for the Inception V3 and VGG16 techniques. The 23153 hand gesture or sign language images included in our experiment were examined using ten different classification approaches and two convolutional neural network algorithms. These ten numbers are 0–9: 0, 1, 2, 3, 4, 5, 6, 7, 8, and 9 for easy recognition by a machine.

Chapter 5

Impact on Society, Environment and Sustainability

5.1 Impact on Society

Hand gesture recognition is very crucial to our society. By creating more interaction with hand gesture recognition technology and robots, this study will have a profound societal impact. It is especially important for those who are unable to speak or hear. They can communicate with normal people using this technology. I believe it has the potential to have a significant impact on society. In general, autistic people are unable to speak or hear and are constantly isolated from society. This could be the education sector, sports sector, job sector, and so on. That is why they can easily communicate with all of Sector using hand gestures and signs. Using hand gesture technology, they can learn more from school and will act as a valuable asset to society.

5.2 Impact on Environment

Hand gesture recognition is used not only by autistic people who are blind or deaf but also by normal people who are speaking to non-native speakers. Different environments are crucial for hand gesture recognition.

5.3 Ethical Aspects

According to the ethical consideration, the study models and dataset violate neither confidentiality nor any fundamental human rights. Although we used a variety of sources to assemble the data for this study, the majority of it was created by ourselves. While carrying out our work, we made no claims to the work produced by any other companies or people. We worked on our personal computers. Honesty, respect for the law, integrity, validity, and openness were all factors considered during the study process.

5.4 Sustainability Plan

This study is obviously sustainable. The primary objective of this project is to categorize hand gestures using convolutional neural network techniques and data augmentation. Many long-term upgrades to the corporation can be made using our method. Before we begin, there are approximately 9.6% of the people who are blind or deaf in Bangladesh. It is gradually increasing day by day, which is why we need to conduct more and more research on hand gesture recognition to make it more valuable.

Chapter 6

Conclusion, Future Research

6.1 Conclusions

In the modern world, during a period of modern technology, anything is possible with the aid of smart technology. We proposed two convolutional neural network models, InceptionV3 and VGG16, for hand gesture recognition with synthetic data augmentation for classification using our own built dataset in this study. In this experiment, two convolutional neural network classifier methods are used to more accurately identify hand gestures on a custom dataset. We have utilized two algorithms in our study, which has produced a range of outcomes. The classification of the provided results in architectures shows a high degree of accuracy, close to 99.59%, while VGG16 has an accuracy of 97.21%. To be tested in the future, we want to create a dataset that combines sign recognition with digit and alphabet recognition.

6.2 Implication for Further Study

We discovered various restrictions when doing so, such as the fact that we had no prior experience with limited domains. In our upcoming study, we can explore the prospect of employing a deep learning technique as a classification algorithm and collecting data on more advanced and dynamic basic gestures. Furthermore, our goal is to work from A to Z and do work the fundamental signs of deafness and hard of hearing.

References:

[1] Asif, Ali Raza, et al. "Performance evaluation of convolutional neural network for hand gesture recognition using EMG." *Sensors* 20.6 (2020): 1642.

[2] Ageishi, Naoto, Fukuchi Tomohide, and Abderazek Ben Abdallah. "Real-time Hand-Gesture Recognition based on Deep Neural Network." SHS Web of Conferences. Vol. 102. EDP Sciences, 2021.

[3] Benitez-Garcia, Gibran, Lidia Prudente-Tixteco, Luis Carlos Castro-Madrid, Rocio Toscano-Medina, Jesus Olivares-Mercado, Gabriel Sanchez-Perez, and Luis Javier Garcia Villalba. "Improving real-time hand gesture recognition with semantic segmentation." Sensors 21, no. 2 (2021): 356.

[4] Parvathy, Priyanka, Kamalraj Subramaniam, G. K. D. Prasanna Venkatesan, P. Karthikaikumar, Justin Varghese, and T. Jayasankar. "Development of hand gesture recognition system using machine learning." Journal of Ambient Intelligence and Humanized Computing 12, no. 6 (2021): 6793-6800.

[5] Oudah, Munir, Ali Al-Naji, and Javaan Chahl. "Hand gesture recognition based on computer vision: a review of techniques." journal of Imaging 6.8 (2020): 73.

[6] Al-Hammadi, Muneer, Ghulam Muhammad, Wadood Abdul, Mansour Alsulaiman, Mohammed A. Bencherif, Tareq S. Alrayes, Hassan Mathkour, and Mohamed Amine Mekhtiche. "Deep learning-based approach for sign language gesture recognition with efficient hand gesture representation." IEEE Access 8 (2020): 192527-192542.

[7] Skaria, Sruthy, Akram Al-Hourani, and Robin J. Evans. "Deep-learning methods for hand-gesture recognition using ultra-wideband radar." IEEE Access 8 (2020): 203580-203590.

[8] Al-Hammadi, Muneer, Ghulam Muhammad, Wadood Abdul, Mansour Alsulaiman, Mohamed A. Bencherif, and Mohamed Amine Mekhtiche. "Hand gesture recognition for sign language using 3DCNN." IEEE Access 8 (2020): 79491-79509.

[9] Yuanyuan, S. H. I., L. I. Yunan, F. U. Xiaolong, M. I. A. O. Kaibin, and M. I. A. O. Qiguang. "Review of dynamic gesture recognition." Virtual Reality & Intelligent Hardware 3, no. 3 (2021): 183-206.

[10] Mujahid, Abdullah, Mazhar Javed Awan, Awais Yasin, Mazin Abed Mohammed, Robertas Damaševičius, Rytis Maskeliūnas, and Karrar Hameed Abdulkareem. "Real-time hand gesture recognition based on deep learning YOLOv3 model." Applied Sciences 11, no. 9 (2021): 4164.

[11] Hu, Bin, and Jiacun Wang. "Deep learning based hand gesture recognition and UAV flight controls." International Journal of Automation and Computing 17.1 (2020): 17-29.

[12] Tan, Yong Soon, Kian Ming Lim, and Chin Poo Lee. "Hand gesture recognition via enhanced densely connected convolutional neural network." Expert Systems with Applications 175 (2021): 114797.

[13] “Nonverbal communication and body language,” HelpGuide.org. [Online]. Available: https://www.helpguide.org/articles/relationships-communication/nonverbal-communication.htm. [Accessed: 18-Jan-2023].

[14] B. Raj, “A simple guide to the versions of the inception network,” Medium, 31-Jul-2020. [Online]. Available: https://towardsdatascience.com/a-simple-guide-to-the-versions-of-the-inception-network-7fc52b863202. [Accessed: 18-Jan-2023].

[15] “Deep Network designer,” Inception-v3 convolutional neural network - MATLAB. [Online]. Available: https://www.mathworks.com/help/deeplearning/ref/inceptionv3.html;jsessionid=49b5b004065179659f3875ca09dd. [Accessed: 18-Jan-2023].

[16] G. L. Team, “Introduction to VGG16: What is VGG16?,” Great Learning Blog: Free Resources what Matters to shape your Career!, 18-Nov-2022. [Online]. Available: https://www.mygreatlearning.com/blog/introduction-to-vgg16/#VGG%2016%20Architecture. [Accessed: 18-Jan-2023].

Hand gesture

ORIGINALITY REPORT

16%	16%	3%	13%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	dspace.daffodilvarsity.edu.bd:8080 Internet Source	13%
2	www.researchgate.net Internet Source	1%
3	www.mdpi.com Internet Source	1%
4	Submitted to Daffodil International University Student Paper	1%

Exclude quotes	Off	Exclude matches	< 1%
Exclude bibliography	On		