Gourd Vegetable Detection by Deep Learning Approach

BY

Kana Akter
ID: 203-15-3866

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

Supervised By
Md. Sazzadur Ahamed
Assistant Professor
Department of CSE
Daffodil International University

Co-Supervised By
Md. Zahid Hasan
Associate Professor
Department of CSE
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

14 JANUARY, 2024

APPROVAL

This project titled "**Gourd Vegetable Detection by Deep Learning Approach**, submitted by **Kana Akter**, and **Md. Sazzadur Ahmed Assistant Professor** 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 (BSc) and approved as to its style and contents. The presentation has been held in 27 January, 2024.

BOARD OF EXAMINERS

Chairman

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

27/1/24

Internal Examiner

Dr. Arif Mahmud
Associate Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

27/01/24

Internal Examiner

Mr. Shahadat Hossain
Senior Lecturer
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

27.01.24

External Examiner

Dr. Firoz Ahmed
Professor
Information & Communication Engineering
Rajshahi University

DECLARATION

I hereby declare that this thesis has been done by me under the supervision of of **Md. Sazzadur Ahmed Assistant Professor,** Department of CSE, and co-supervision of **Md. Zahid Hasan Associate Professor,** Department of CSE Daffodil International University. I also declare that neither this thesis nor any part of this thesis has been submitted elsewhere for the award of any degree or diploma.

Supervised by:

Md. Sazzadur Ahmed
Assistant Professor
Department of CSE
Daffodil International University

Co-Supervised by:

Md. Zahid Hasan
Associate Professor
Department of CSE
Daffodil International University

Submitted by:

Kana Akter

Kana Akter
ID: 203-15-3866
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

First I express my heartiest thanks and gratefulness to almighty God for His divine blessing makes me possible to complete the final year thesis based project successfully.

I really grateful and wish our profound our indebtedness to **Md. Sazzadur Ahmed,** Assistant Professor, Department of CSE Daffodil International University, Dhaka. Deep Knowledge & keen interest of our supervisor in the field of "Machine Learning" to carry out this project. His endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice ,reading many inferior draft and correcting them at all stage have made it possible to complete this project.

I would like to express our heartiest gratitude to **Dr. Sheak Rashed Haider Noori,** Professor and Head, Department of CSE, for his kind help to finish our project and also to other faculty member and the staff of CSE department of Daffodil International University.

Finally, I must acknowledge with due respect the constant support and patients of our parents.

ABSTRACT

The local market in Bangladesh offers a diverse range of around 100 varieties of vegetables for purchase and sale. The nomenclature of vegetables exhibits regional variations, leading to difficulties in identification for individuals. Amongst these, the gourd vegetables belonging to the Cucurbitaceous family, namely Sponge Gourd, Ridge Gourd, and Snake Gourd, Bottle Gourd, Bitter Gourd pose the greatest challenge due to their strikingly similar structural characteristics despite significant differences in quality. This study focuses on the application of image processing techniques for the identification of three specific vegetables. The utilization of image processing techniques in the field of agriculture is experiencing a steady growth in recent times. The utilization of Quality Test in image processing facilitates the diagnostic procedure and serves as a means to discern several categories of vegetables by analyzing their size, shape, and color. The suggested methodology starts with the acquisition of picture data. A total of 10000 photos of Sponge Gourd, Ridge Gourd, and Snake Gourd, Bottle Gourd, Bitter Gourd were obtained in the agricultural field. Upon the completion of image processing, I conducted training and testing utilizing the model 3 architectures. The obtained accuracy rate was 99% in both cases.

TABLE OF CONTENTS

LIST OF FIGURE

LIST OF TABLE

CHAPTER 1
INTRODUCTION

1.1 Introduction

Based on data provided by GSMA, there has been a notable rise in the number of mobile phone users relative to the global population. Specifically, the global population stands at 7.88 billion, while the number of mobile connections amounts to 10.37 billion. Among these connections, 3.8 billion are attributed to smartphone users. The use of smartphone cameras enables individuals to conveniently amass a collection of photographs or films. The incorporation of pictures in the domains of automated detection, classification, and single and multiple item identification has great importance. The use of image processing techniques in these areas has introduced a novel aspect to the field of study. The integration of photos has emerged as a significant component within the field of agricultural research. By employing images, it becomes feasible to digitally discern and classify various categories of vegetables and fruits within the agricultural sector [1]. The categorization of fruits and vegetables remains tough in image processing since fruits might have identical forms or textures, only exterior aspects can be determined by the picture [2]. The field of computer science has the potential to significantly impact the agricultural sector by using image processing techniques. The field of artificial neural networks has shown significant progress and expansion in recent times. The classification and operational precision exhibit remarkable performance. When confronted with challenges related to image processing, an alternate approach might be considered. Unsupervised learning and feature-rich extraction are two key attributes of neural networks that contribute to enhancing the network's overall performance. Within the broader context of imaging, the Convolutional Neural Network (CNN) has achieved notable success.

1.2 Motivation

The main catalysts that propel my work may be categorized into two separate domains. For example, advancements in the field of agriculture, along with significant technological innovations, have contributed to a shift in consumer perception, leading to the recognition of gourds as a vegetable. If this program achieves accurate identification of the diverse species of gourds, it has the potential to make a significant contribution to the progress of agriculture in Bangladesh. The implementation of automatic recognition is extremely helpful in the medical and nutraceutical fields. The fundamental concern is in the inability of some individuals to identify certain fruits, and our programmed aims to provide assistance to such persons in overcoming this challenge. This constitutes the principal rationale behind my desire to finalize this project.

1.3 Problem Definition

There is still a sizeable segment of Bangladesh's population that is not familiar with the gourd vegetable, particularly in the country's urban regions. since there are only subtle differences between the many gourd veggies. It is predicted that the aforementioned problem will be adequately addressed as a result of this endeavor. During the process of putting this initiative into action, I ran into difficulties that were associated with manual identification. The data labelling procedure needed manual involvement, which resulted in a significant increase in processing time. I have carried out each procedure methodically and with utmost attention to detail. There is also a problem that occurs when particular seasons prevent specific kinds of gourd vegetables from being readily available. The seasons bring about shifts in the availability of certain vegetables. When I first started working here, there were just five different kinds of fruit available to choose from among the little selection.

1.4 Research Questions

1. What will be the primary sources from which the data will be drawn, and what methods will be used to acquire it?

2. Inquiring into the topic of "Is it possible to accurately describe various gourd?"

3. Which of the following categories would best describe them?

4. How well can machine learning distinguish between the many varieties of gourd?

5. How deep learning architecture is built?

1.5 Expected Outcome

1. I anticipate that the model will be able to recognize gourd veggies with high accuracy as a result of this study.
2. It can assist individuals in the identification of Bangladeshi gourd vegetables.
3. The agriculture sector in Bangladesh is also supported by this initiative.

1.6 Report Layout

The data that we provide in our report will have been made public simultaneously with the release of our report:

Chapter 1 represents the examination of the study is dissected and evaluated. The preliminary analysis has significant importance within the context of this introductory portion. In this section, we will also examine the underlying motivations for our study. The problem elucidated in this chapter holds considerable significance. This component of the study provides an analysis and discourse on the topic.

Chapter 2 Illustrate the previous study. This chapter will undertake an analysis of many previously published studies that bear resemblance to my study, so facilitating a comparative evaluation with my own research.

Chapter 3 represents the methodology or approach that was used is analyzed in great detail, which provides a summary of the methodology or approach. In this part of the article, I will talk about the methodology that was used to carry out the study.

Chapter 4 is mainly experimental results. In this chapter, I will compare the outcomes obtained from three distinct models. I also explain how I can pick the best model out of the three available options.

Chapter 5 In this chapter, I will discuss the societal impact, as well as the plan for sustainability. I will explain how my project contributes to society as well as the requirements necessary to use this project.

Chapter 6 is the conclusion of the study paper is presented. Please examine this section to evaluate the efficacy of the model. The conclusion of the chapter entails a discussion on the constraints of the scope.

1.7 Research Objectives

1. To generate a thorough dataset on gourds.

2. To develop a customized Convolutional Neural Network (CNN) model.

3. In order to assist anyone encountering difficulties in identifying gourds within the context of Bangladesh.

4. The development of an Android application centered around gourd recognition is proposed.

1.8 Summary

This chapter provides a comprehensive overview of the introduction, goal, and rationale for the present study. Additionally, a detailed step-by-step explanation was provided on how to address the aforementioned issues. In the subsequent chapter, an exposition of pertinent literature will be provided.

CHAPTER 2
BACKGROUND STUDY

2.1 Introduction

Several attempts have been undertaken to utilize deep learning methodologies in order to identify and classify vegetables. Predictive modelling is a commonly employed application of machine learning and deep learning methodologies. Extensive scholarly inquiry has been devoted to the exploration of techniques for the recognition of vegetables in the context of identification. The courses focused considerable emphasis on the examination of deep learning techniques and their practical implementation in the context of problem-solving. The objective of this chapter is to furnish the reader with a comprehensive survey of the substantial corpus of study undertaken by several scholars in the aforementioned domain.

2.2 Related Works

The use of computer vision techniques in the agriculture industry has shown significant growth in recent years. Our analysis and comprehension are founded upon the works of several writers, a selection of which are referenced afterwards.

In their study, Dubey et al. (2021) classify various types of produce into several groups and draw attention to the concerns associated with fruit illnesses. The use of K-means clustering was implemented in the process of segmenting fruit for sickness detection. During disease classification, error rates ranging from 1% to 3% were recorded. Techniques derived from the discipline of computer vision are used. The researchers conducted an analysis of two credibility-based performances. The first performance focused on the identification of fruit and vegetable classification, which could potentially be utilized to assess the reliability of vegetables and fruits. The second performance involved the identification of fruit disease classification, which exhibited similarities to vegetable disease classification in terms of color and texture. There are a total of 2,615 photos, which may be categorized into 15 distinct kinds. The galleries include images of many fruits and vegetables, including Fuji apples, Granny Smith apples, plums, Asterix potatoes, Agata potatoes, cashews, oranges, honeydew melons, Tahitian limes, onions, nectarines, kiwis, watermelons, Spanish pears, and diamond peaches. The efficacy of this study is impeded by the limitation that just a single condition is addressed within these papers. The objective of this study is to provide the foundational basis for further investigations including unconventional species and diverse cultivars of food plants.

El-Salam (2) proposed a computer vision approach for the classification and categorization of many varieties of fruits and vegetables. During the course of this study, a comprehensive analysis was conducted on a total of 600 apple samples, resulting in the identification of a success rate of 94% in distinguishing between bad quality apples and those of excellent quality. The system has three cameras, each with a field of vision capable of encompassing 24 apples. Consequently, the system has the capacity to check 3000 apples every minute. The use of neural networks with fuzzy inference approaches has shown advantageous outcomes in the realm of picture feature categorization.

Zeng (2019) proposed a methodology designed for the detection and classification of fruits and vegetables. They have built a convolutional neural network model for the aim of performing image classification, which involves representing extracted aspects of pictures. The researchers used the VGG model for training purposes. The classification system encompasses a comprehensive range of 26 distinct categories. The proposed approach has a level of accuracy of 95.6 percent. The research incorporates a dataset of a total of 12,173 photos. The objective of this study is to investigate several methodologies for the recognition and categorization of commonly eaten fruits and vegetables in the foreseeable future.

Sakai et al. (2019) proposed a rudimentary but robust convolutional neural network (CNN) architecture aimed at achieving object category classification by including both feature extraction and item learning. The objective of this model was to investigate the process of object category identification by examining the acquisition of knowledge pertaining to the object. The study used image data from a diverse range of vegetables, including tomato, carrot, banana, cabbage, spinach, eggplant, and shiitake mushrooms. In contrast to the utilization of 20 pictures for each veggie, the training phase incorporates a cumulative total of 160 images. In contrast, a set of five shots capturing each vegetable will be subjected to the evaluation, resulting in a cumulative collection of forty images. In this study, a series of iterations ranging from 1 to 10 million are conducted. The performance has shown enhancement and is now deemed suitable for 3 million learning iterations; nevertheless, concurrently, the rate has been reduced for 10 million learning iterations. Based on the results obtained from the conducted investigation, it was observed that the recognition rate achieved a value of 97.58%, but the learning rate reached a value of 99.14. The focus of their research seems to be on object tracking with deep neural networks (DNN) in the near future.

The subject of discussion in the study conducted by Muhtaseb et al. [5] revolved on an application that utilizes photo histograms for the purpose of distinguishing between fruits and vegetables. The matching procedure considers both the hues and dimensions of the objects. The researchers used a set of four distinct attributes, namely lemon, apple, cucumber, eggplant, and pear, to describe the real qualities under investigation. At the time of this research, the findings indicate a 75% accuracy rate in fruit identification. The Chi-square Method was used to facilitate the comparison and alignment of photos in the preliminary inquiry. The objective of this piece is to examine the potential for expanding the assortment of fruits that are acknowledged in forthcoming times.

Dubey et al. (6) conducted a study that specifically addressed the challenge associated with accurately identifying fruits and vegetables. A computer vision processing system was used to analyze images, with the objective of determining the species and variety of the item under investigation. The subsequent stage entails an examination of the texturing characteristic. The findings of this investigation indicate a high degree of precision, namely 99%, as determined by the ISADH texture. Three techniques were implemented in this study. The first approach included the segmentation of pictures using a K-means clustering technique. The second methodology focused on the extraction of features.

Lastly, the third methodology utilized a multi-class support vector machine to classify various kinds of produce. A collection of 15 distinct images depicting various fruits and vegetables has been made available to us for use as reference models. There has been a deliberate endeavor to focus on the augmentation of a diverse range of attributes and the potential manifestations of these attributes in subsequent periods.

The study done by Xiaojun et al. [7] used deep neural networks and machine vision systems. The identification of vegetable and weed species using the CenterNet model is crucial for both vegetable identification and the delineation of bounding boundaries for these species. The segmentation based on color index was accomplished via the use of image processing techniques. The recommended strategy used the CenterNet model, which achieved an accuracy of 95.6%, a recall of 95.0%, and an F1-score of 0.953. The pictures of bok choy, commonly referred to as Chinese white cabbage, have been supplemented with additional image data. The training dataset consisted of 1150 photographs, with the input images being scaled to have dimensions of 512 pixels on each side, resulting in a total of 512 pixels overall. A novel approach was devised to address the issue of weed removal in the background, including the use of genetic algorithms for the assessment of the color index. The techniques used in this study were derived from Bayesian classification error. The focus of their research is on the utilization of in-situ videos as a prospective method for weed identification in the forthcoming period.

Ercole et al. [8] used an antibody-based biosensor in their investigation to examine a detection model for Escherichia coli (E. coli) organisms in vegetable meals. Bacteria were chosen as the subject of investigation due to their prevalence as a reliable indication of fecal contamination. The development of a recognition antigen-antibody system with the capability to detect E. coli bacteria in both food preparation water and environmental water is a plausible endeavor. At present, there is ongoing development of a monoclonal antibody, alongside efforts to enhance the immobilization procedure. The authors have used lettuce, sliced carrots, and rucola as instances of other vegetables that are at their disposal. The bacterial cells were subjected to a washing process using peptone water with a pH of 6.8. Subsequently, the cells were subjected to mechanical disruption using either a stomacher or a sonicator, with the aim of separating and retrieving the bacterial cells in the liquid medium. The PAB system has a greater level of sensitivity in comparison to traditional approaches. A cellular concentration of 10 cells per milliliter was seen after around 1.5 hours, suggesting a detection time that is 10 to 20 times faster compared to the standard colony-forming unit (CFU) approach.

Kaur et al. [9] proposed a methodology including an artificial neural network to assess the quality of vegetables. The traits that identify vegetables include their color, shape, and size. The Artificial Neural Network (ANN) methodology was used to conduct quality control inspections. The provided dataset has four images depicting various types of vegetables. In the foreseeable future, researchers want to prioritize the focus on quality evaluation via Artificial Neural Networks (ANN).

The study undertaken by Kagaya et al. [10] focused on the development of a model for the identification and detection of food items. In order to accurately classify the model of this fragmentation, a convolutional neural network was used. When using the formative model, it is said that there is an enhanced level of accuracy, specifically quantified as 93.8 percent. During the training phase of the study endeavor, a total of 1200 photos depicting food and 1000 images depicting nonfood items were shown. A total of 200 photographs were used throughout the testing phases. After doing 15 repetitions, the result was computed and afterwards decided. The researchers used image detection and identification methods to ascertain the nature of food items. Specifically, they utilized four distinct methodologies that were founded upon the principle of reliability.

Pouladzadeh and colleagues (2011) provide a comprehensive overview of a study investigating the domains of food detection, categorization, and analysis, which are intricately linked to the examination of eating patterns and the evaluation of dietary habits. The investigation included the determination of calorie counts via the use of several food items and their corresponding portion sizes. This article showcases a collection of three thousand distinct food images, each captured with diverse cameras and under varying lighting conditions. In order to facilitate the collection of data, it is divided into two components: a singular serving and a composite serving including several food items. The papers illustrate the use of techniques such as graph cut segmentation and deep learning for the purpose of food identification.

Probst et al. [12] provide a description of the food record system inquiry and provide a prototype of the Australian diet. For the purpose of this evaluation, a smartphone is being deployed in conjunction with picture processing and pattern recognition software. The study employs scale-invariant feature transformation, local binary patterns (LBP), and color as techniques to characterize food images. The bag-of-words approach was used for the task of image recognition. The dimensions of the provided image exceed 2000 pixels. Both the training sets, which were used throughout the codebook development process, and the test sets, which were extracted from the dataset, were drawn from a well-structured dataset. The researchers want to focus their future efforts on the detection and identification of food.

In their research, Rimi et al. [13] developed a machine learning-based model with the objective of discerning nine distinct plant species. In order to do this, the researchers used several machine learning approaches, including k Nearest Neighbors (k NN), Support Vector Machines (SVMs), and the Classification and Regression Tree (CART) algorithm. Furthermore, the use of deep learning techniques was implemented, specifically emphasizing transfer learning architectures such as VGG 16, Inception v3, and ResNet 50. The methods underwent training using a dataset of 4,320 training images that were scaled to a resolution of 224x224 pixels. In addition, a total of 1,080 test photos were used throughout the assessment process. The study of the performance indicated that the machine learning model had a maximum accuracy of 76.0 percent. By using the Inception V3 model with a total of 23 million trainable parameters, a remarkable accuracy rate of 98.0 percent was achieved. This stands in stark contrast to other approaches.

In the country of Bangladesh, there exist four unique varieties of potato. A research study conducted by Nuruzzaman et al. [14] has successfully developed an identification algorithm based on machine vision technology, which is capable of discerning between these different potato species. The researchers used several machine learning models, including random forests, linear discriminant analyses, logistic regression, support vector machines, naive Bayes, and the k-NN method, to conduct their analysis. Three unique analytic approaches, including color histogram, Hu moments, and Haralick texture, were used for image pre-processing. The Logistic Regression approach achieved a maximum attainable accuracy of 98% via the use of 200 distinct potato photographs.

In reviewing the previous discourse on pertinent literature, it is evident that there exists a diverse range of studies pertaining to fruits and vegetables. These studies encompass disease recognition, object recognition specifically pertaining to fruits and vegetables, recognition of different species of fruits and vegetables, as well as quality grading. However, it is noteworthy that no research has been conducted on visually similar fruits and vegetables or their species recognition. Furthermore, a comprehensive investigation into this specific sub problem of fruits and vegetables or their species recognition has yet to be undertaken. To clarify, none of these tasks has been accomplished.

2.3 Research Summary

Numerous research organisations appear to be pursuing a similar line of investigation, which has already been well explored within the domain of image categorization. The implementation of the project is distributed across multiple research groups. The findings of our investigation have yielded intriguing and innovative conclusions. Despite the existing scarcity of assets, it is anticipated that this business will continue to grow due to the inclusion of purchase data from newly introduced products.

CHAPTER 3
RESEARCH METHODOLOGY

3.1 Introduction

There are nine phases involved in my process for completing projects successfully. The schematic for each level is presented and shown in figure 3.1.

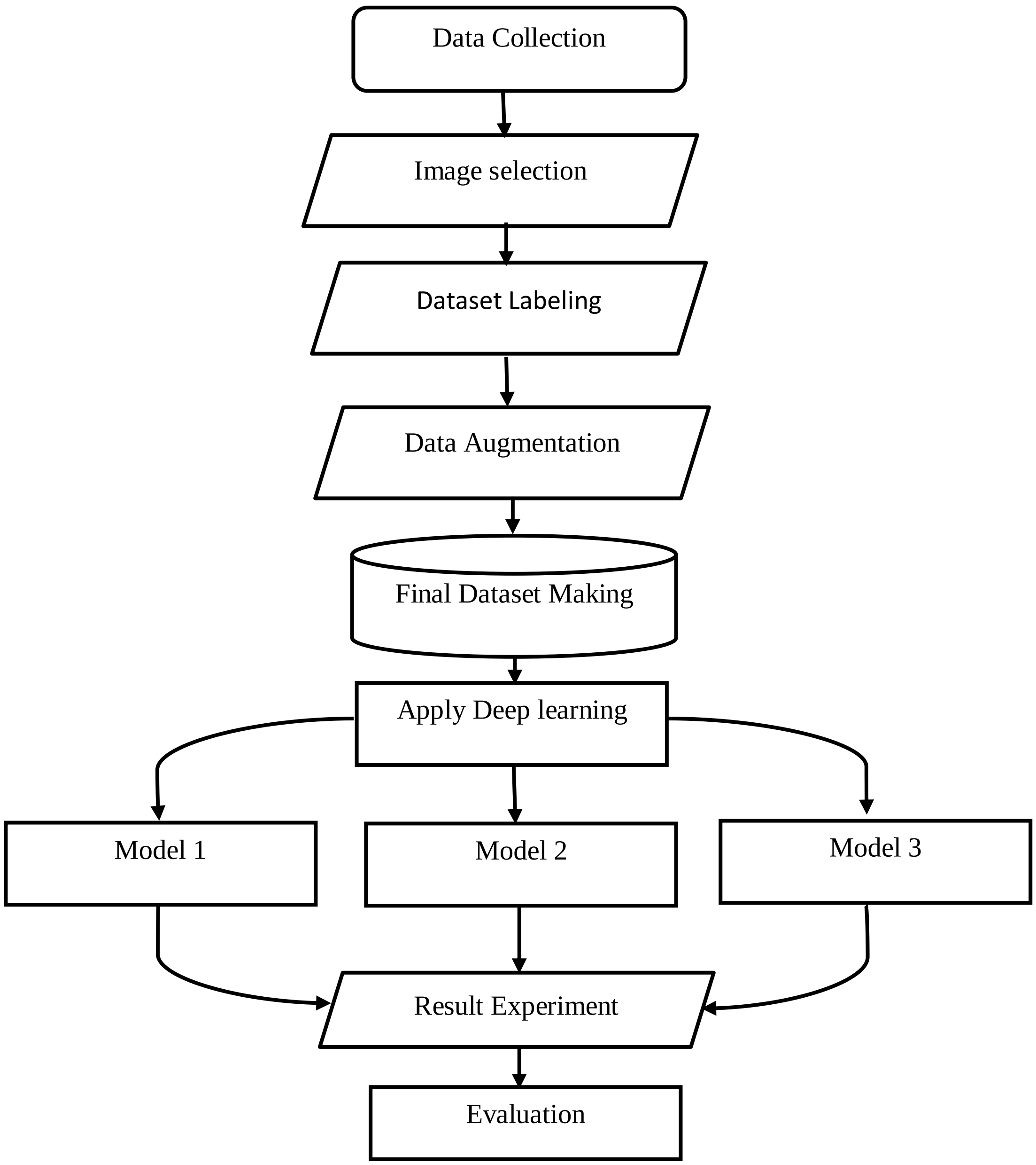


Figure 3.1: Methodology diagram

3.2 Data Collection

The area provided all of the data that was necessary for us to finish my job successfully. My information was gathered from a wide range of vegetable markets located all around the country of Bangladesh. There are around 2,000 distinct photo sets that are compiled as part of the data collection. Collect data from at least three different mobile devices, ideally from four or more of them, if at all feasible. The gathering of information from agricultural fields is my major purpose at this point. A total of around 1500 photographs were taken straight from the field, with the very last one being taken in a neighboring market.

3.3 Image selection

I went out to the field as well as the market, and I ended up taking almost two thousand photographs of the five distinct types of gourd vegetables. After I have collected all of the photographs, I then arrange each one in its own individual folder. Even though there weren't very many of them, I noticed that a couple of the images did not meet the stringent regulatory requirements and were fuzzy. On the other hand, I looked through my dataset and removed them manually from consideration. In addition, I used photographs that I selected because of their high quality in my study.

3.4 Dataset labeling

Nine distinct varieties of gourd vegetables from Bangladesh were utilised in the research that was carried out there. After the data gathering was finished, I proceeded to apply the dataset labelling techniques. The process needed a considerable amount of time to be completed. I did everything I could to ensure that the data were accurate and consistent. The distribution of the dataset is presented in Table 3.1 for reference. From this table we can see that bitter gourd contains 1865 images, bottle gourd contains 1871 images, ridge gourd contains 1914 images, snake gourd contains 1880 images, sponge gourd contains 1986 images. The total number of training size is 9516.

Table 3.1 DATASET REPRESENTATION

Fruits Name	Number of images
Bitter gourd	1865
Bottle gourd	1871
Ridge gourd	1914
Snake gourd	1880
Sponge gourd	1986

3.5 Data augmentation

The example of enhanced picture used for data augmentation is represented in Table 3.2. I employed a number of strategies, including sleep horizontal rotation vertical rotation vertical flip.

Table 3.2 DATASET REPRESENTATION.

Vegetable Name	Image
Bitter Gourd	
Bottle Gourd	
Ridge Gourd	
Snake Gourd	
Sponge Gourd	

3.6 Final Dataset Making

Data analysis is crucial to scientific research because it simplifies and improves data analysis. It simplifies objective data analysis for academics, preventing inspiration loss. Every real-world AI application needs a high-quality dataset. Any machine learning or deep learning model needs enough high-quality, relevant data sequences to succeed.

I create three datasets. One is training, validation, and test dataset. Descriptions of each dataset follow:

The table 3.3 shows that there are a total of 1865 pictures in a bitter gourd, 1871 in a bottle gourd, 1914 in a ridge gourd, 1880 in a snake gourd, and 1986 in a sponge gourd. There are a grand total of 9516 trainees.

Table 3.3 TRANING DATASET REPRESENTATION

Fruits Name	Number of images
Bitter gourd	1865
Bottle gourd	1871
Ridge gourd	1914
Snake gourd	1880
Sponge gourd	1986

Table 3.4 presents the testing dataset. Specifically, the bitter gourd exhibits a total of 156 photographs, the bottle gourd contains 192 pictures, the ridge gourd encompasses 537 pictures, the snake gourd comprises 648 pictures, and the sponge gourd encompasses 634 pictures. The total number of trainees amounts to 2167.

Table 3.4 TESTING DATASET REPRESENTATION

Fruits Name	Number of images
Bitter gourd	156
Bottle gourd	192
Ridge gourd	537
Snake gourd	648
Sponge gourd	634

The dataset used in the validation is listed in Table 3.5. In detail, there are a total of 65 images displayed on the bitter gourd, 80 photographs within the bottle gourd, 177 photographs on the ridge gourd, 216 photographs within the snake gourd, and 263 photographs within the sponge gourd. There are a grand total of 801 students enrolled.

Table 3.5 VALIDATION DATASET REPRESENTATION

Fruits Name	Number of images
Bitter gourd	65
Bottle gourd	80
Ridge gourd	177
Snake gourd	216
Sponge gourd	263

3.6 Custom CNN models

A convolutional neural network (CNN) model was specifically designed to address our specific research objectives, and further modifications were made to align the model with those criteria. In the subsequent section, a comprehensive evaluation of each of the several models is provided:

3.6.1 CNN model 1

The key elements behind our Convolutional Neural Network (CNN) model are concisely shown in Figure 3.2. The construction of this model involved the use of two convolutional layers, two maxpool2d layers, two dropout layers, and three dense layers. The color channel of this architecture is red green blue, so the input shape is (224, 224, 3) the padding of first convolution layer is same. The starting convolutional layer contains 32 numbers of filters and the filter size is 3 by 3. The size of first maxpooling layer is 2 by 2. The number of second convolution filter is 65 where activation is relu, padding is same and the size of filer is 3 by 3. After that I used maxpooling layer with size 2 by 2. The number of first dense filter is 128 with activation relu. After that I used a dropout layer with 50% dropout rate. And final dense layer size is 5 with activation softmax. The total number of trainable parameters of this model is 25718213.

```
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d (Conv2D)              (None, 224, 224, 32)      896
_________________________________________________________________
max_pooling2d (MaxPooling2D) (None, 112, 112, 32)      0
_________________________________________________________________
conv2d_1 (Conv2D)            (None, 112, 112, 64)      18496
_________________________________________________________________
max_pooling2d_1 (MaxPooling2 (None, 56, 56, 64)        0
_________________________________________________________________
flatten (Flatten)            (None, 200704)            0
_________________________________________________________________
dense (Dense)                (None, 128)               25690240
_________________________________________________________________
dropout (Dropout)            (None, 128)               0
_________________________________________________________________
dense_1 (Dense)              (None, 64)                8256
_________________________________________________________________
dropout_1 (Dropout)          (None, 64)                0
_________________________________________________________________
dense_2 (Dense)              (None, 5)                 325
=================================================================
Total params: 25,718,213
Trainable params: 25,718,213
Non-trainable params: 0
_________________________________________________________________
```

Figure 3.2 Custom CNN model 1

3.6.3 CNN model 2

Figure 3.3 represents the custom model 2. This is my most complex architecture among all three models. This model contains total three convolution layer three maxpooling layer three batch normalization layer and one globalaveragepooling2d layer. The first convolution layer contains 32 filter with size 3 by 3 with strides 2 by 2. The input shape of the model is 224 by 224 and I used same padding. The first maxpooling layer introduced with size 3 by 3. After this combination I used batchnormalization layer which helps model to faster calculation. The second convolution layer contains 64 filter with size 3 by 3 with strides 2 by 2. I used also same padding for this layer. The second maxpooling layer introduced with size 3 by 3. After this combination I used batchnormalization layer aslo. The third convolution layer contains 128 filters with size 3 by 3 with strides 2 by 2. I used also same padding for this layer. The second maxpooling layer introduced with size 3 by 3. After this combination I used batchnormalization layer. After all combination I used Global average pooling layer which helps to perform dimensionality reduction process.

```
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d (Conv2D)              (None, 112, 112, 32)      896
_________________________________________________________________
max_pooling2d (MaxPooling2D) (None, 38, 38, 32)        0
_________________________________________________________________
batch_normalization (BatchNo (None, 38, 38, 32)        128
_________________________________________________________________
conv2d_1 (Conv2D)            (None, 19, 19, 64)        18496
_________________________________________________________________
max_pooling2d_1 (MaxPooling2 (None, 7, 7, 64)          0
_________________________________________________________________
batch_normalization_1 (Batch (None, 7, 7, 64)          256
_________________________________________________________________
conv2d_2 (Conv2D)            (None, 4, 4, 128)         73856
_________________________________________________________________
max_pooling2d_2 (MaxPooling2 (None, 2, 2, 128)         0
_________________________________________________________________
batch_normalization_2 (Batch (None, 2, 2, 128)         512
_________________________________________________________________
global_average_pooling2d (Gl (None, 128)               0
_________________________________________________________________
dense (Dense)                (None, 256)               33024
_________________________________________________________________
dropout (Dropout)            (None, 256)               0
_________________________________________________________________
dense_1 (Dense)              (None, 128)               32896
_________________________________________________________________
dropout_1 (Dropout)          (None, 128)               0
_________________________________________________________________
dense_2 (Dense)              (None, 5)                 645
=================================================================
Total params: 160,709
Trainable params: 160,261
Non-trainable params: 448
```

Figure 3.3Custom CNN model 2

3.6.5 CNN model 3

Figure 3.4 represents the custom cnn model 3. This model is not too complex nor too simple. This model looks like model 2 but there are some differences here. I did not use any batchnormalization layer here, aslo there is no globalaveragepooling here. And the dense architecture is totally different than model to. In this modle first dense layer contains 128 filters with activation relu. After that I used 50% dropout then next layer contains 64 filters and after that I used 30% dropout. Finally dense contains 5 filters with softmax activation.

```
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d (Conv2D)              (None, 224, 224, 32)      896
_________________________________________________________________
max_pooling2d (MaxPooling2D) (None, 112, 112, 32)      0
_________________________________________________________________
conv2d_1 (Conv2D)            (None, 112, 112, 64)      18496
_________________________________________________________________
max_pooling2d_1 (MaxPooling2 (None, 56, 56, 64)        0
_________________________________________________________________
conv2d_2 (Conv2D)            (None, 56, 56, 128)       73856
_________________________________________________________________
max_pooling2d_2 (MaxPooling2 (None, 28, 28, 128)       0
_________________________________________________________________
flatten (Flatten)            (None, 100352)            0
_________________________________________________________________
dense (Dense)                (None, 128)               12845184
_________________________________________________________________
dropout (Dropout)            (None, 128)               0
_________________________________________________________________
dense_1 (Dense)              (None, 64)                8256
_________________________________________________________________
dropout_1 (Dropout)          (None, 64)                0
_________________________________________________________________
dense_2 (Dense)              (None, 5)                 325
=================================================================
Total params: 12,947,013
Trainable params: 12,947,013
Non-trainable params: 0
_________________________________________________________________
```

Figure 3.4CNN model 3

3.7 Summary

This chapter presents an overview of the methods employed in the research study. This chapter provides a detailed description of many steps. The architectural characteristics of three distinct models are also delineated. In the subsequent chapter, an account of the experimental findings pertaining to the three models will be provided, with the objective of identifying the optimal model.

CHAPTER 4
RESULT ANALYSIS

4.1 Introduction

The following section highlights studies that concentrate largely on description and make significant use of empirical data and real experiments. The first thing we need to do in this study is establish a working definition for the term "outcomes analysis." Because of this, the formation of a solid foundation will proceed more smoothly. It has been determined that the author should not have been given advance notice of the findings before the explanation of the implications of those findings. Anyone who is currently working on a research paper will find this material to be quite helpful. After the test is over, the scores are tallied, and a discussion of them follows.

4.2 Experimental Result

Table 4.1 ALGORITHM REPRESENTATION

CNN structure	Image size	Total Epochs	Best Epochs	Parameter usages	Accuracy	Lowest val loss
Model 1	(224,224,3)	25	4	activation = 'softmax' loss ='categorical_crossentropy' optimizer='adam', metrics=['accuracy']	97.19%	0.12
Model 2	(224,224,3)	30	9	activation = 'softmax' loss ='categorical_crossentropy' optimizer='adam', metrics=['accuracy']	99.17%	0.06
Model 3	(224,224,3)	25	7	activation = 'softmax' loss ='categorical_crossentropy' optimizer='adam', metrics=['accuracy']	98.71%	0.13

Table 4.1 represents the algorithm implementation part of my three models. From this table we can see that the input size of all model is (224,224,3). Total epocs of first model is 25 and I got best accuracy 97.19 from epochs number 4 where validation (val) loss is 0.12. The total epochs of model 2 is 30 and I got best accuracy from epochs 9 the accuracy rate is 99.17% where lowest validation loss is 0.06. and finally total epochs of model 3 is 25 I got best accuracy form 98.17 from epochs 7 where validation loss is 0.13.

Above discussion is about training time. from this analysis we can se that my most complex model 2 perform better than others while models are training period. From next analysis I will try to evaluate best model with real life data.

4.3 Model 1 CNN

The accuracy of model 1 during training and validation is depicted in Figure 4.1. The training performance of the model is commendable; nevertheless, the substantial disparity seen between the validation line and the training line suggests the presence of overfitting.

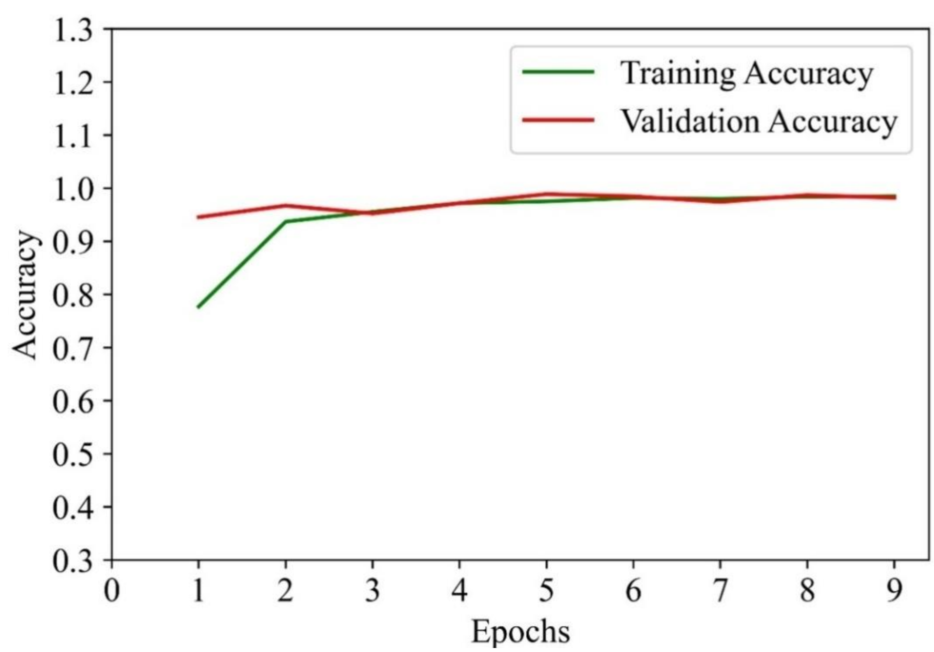


Figure 4.1: Training vs validation accuracy Custom CNN model 1

According to Figure 4.2, it can be observed that the model exhibits a notable level of overfitting. This is evidenced by the decreasing trend in training loss and the simultaneous increasing trend in validation loss as the number of epochs grows.

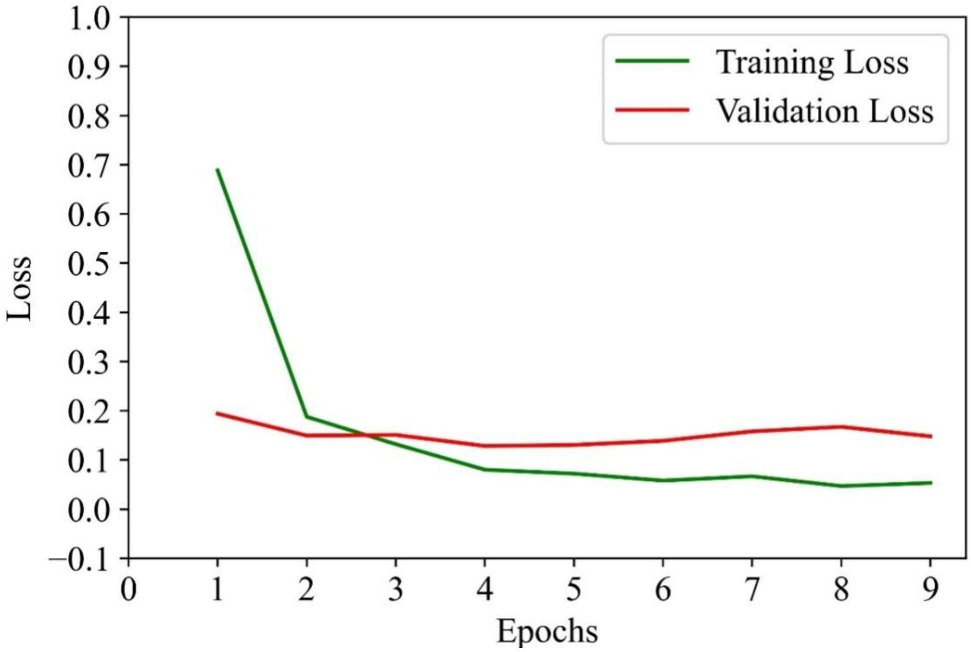


Figure 4.2: Training loss vs validation loss CustomCNN model 1

Table 4.2 Represents the classification report of model 2. From this table we can see that there are 3 factor I used to measure classification report performance.. Class column contains the names of gourd vegetable. For Bitter gourd we can see that the precision rate is 1 recall rate is 0.89. f1 score is 0.94. for bottle gourd the precision rate is 1.00. the recall rate is 0.89 the f1-sxore is 0.94. For ridge gourd the precision rate is 0.95 the rcallsocre is 0.98 for f1 score 0.96. for snake gourd precision rate is 1 recall also 1 and r1 score is also 1. For sponge gourd 0.98 for recall 0.96 for f1 score is 0.97. and the total accuracy of tis model is 0.97 among 801 images. For macro avg the lowest score is 0.96 and others score is 0.97.

Table 4.2: CLASSIFICATION REPORT OF MODEL 1

Class	Precision	Recall	F1-Score	Support
Bitter gourd	1.00	0.89	0.94	65
Bottle gourd	0.90	1.00	0.95	80
Ridge gourd	0.95	0.98	0.96	177
Snake gourd	1.0	1.00	1.00	216
Sponge gourd	0.98	0.96	0.97	263
Accuracy			0.97	801
Macro Avg	0.97	0.97	0.96	801
Weighted Avg	0.97	0.97	0.97	801

4.4 Model 2 CNN

Figure 4.3 presents a graphical depiction of the training accuracy and validation accuracy for our proprietary Convolutional Neural Network (CNN) model 2. The model had remarkable precision throughout the training phase; nonetheless, its highest validation accuracy after 9 iterations was roughly 99%. This is in spite of the fact that the model exhibited remarkable accuracy during the entire training process. The graph clearly indicates that the performance of the model under consideration is superior than that of model 1.

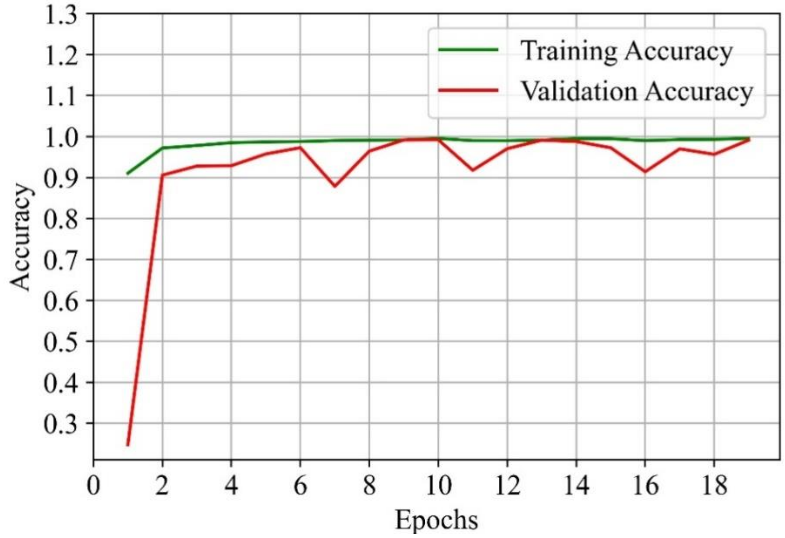


Figure 4.3: Trainingvs validation accuracyCustomCNN model 2.

The training and validation losses of CNN model 2 are compared in Figure 4.4. From this graph we can see that this model run over 18 epochs where training loss is very stable and there are very little gap between train and validation loss that also represents the lower overfitting.

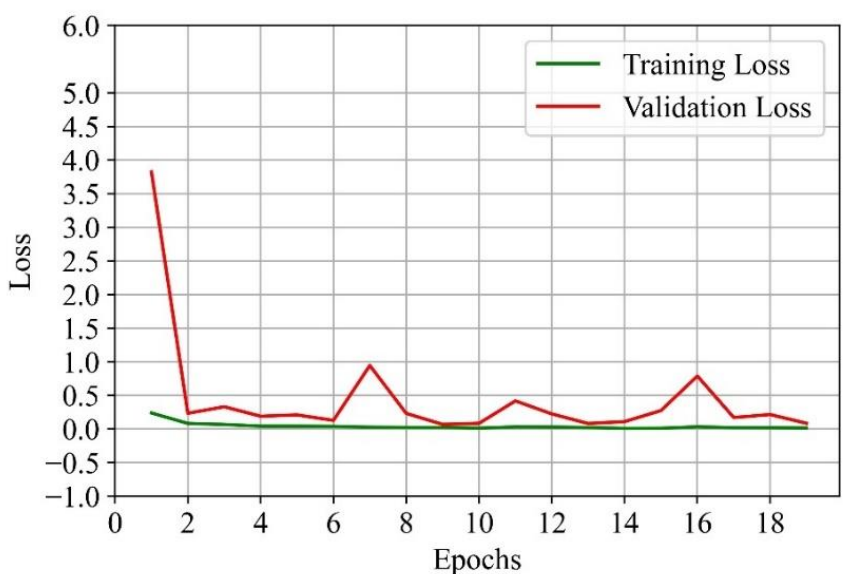


Figure 4.4: Training vs validation loss CustomCNN model 2

We can see the results of model 2's classification in Table 4.3. We can see from the chart that I utilized three metrics to evaluate the quality of the classification report. Vegetable names that are found in the class column are gourds. Our bitter gourd data shows a recall rate of 1.00 and a precision rate of 1. The rank of f1 is 1.00. There is an accuracy rate of 0.98 for bottle gourd. There is a recall rate of 0.99, 1 percent is the f1-sxore. The accuracy rate for ridge gourd is 0.98. A recall score of 1.00 is achieved with an F1 score of 0.99. Regarding snake gourd, the recall, precision rate, and r1 score are all 1. The sponge gourd has a recall of 1.00 and a f1 score of 0.99. And out of 801 photos, the model's overall accuracy is 0.99.

Table 4.3: CLASSIFICATION REPORT OF MODEL 2

Class	Precision	Recall	F1-Score	Support
Bitter gourd	1.00	1.00	1.00	65
Bottle gourd	0.98	1.00	0.99	80
Ridge gourd	0.98	1.00	0.99	177
Snake gourd	1.00	1.00	1.00	216
Sponge gourd	1.00	0.98	0.99	263
Accuracy			0.99	801
Macro Avg	0.99	1.00	0.99	801
Weighted Avg	0.99	0.99	0.99	801

4.5 Model 3

Accuracy in training compared to accuracy in validation is displayed in Fig. 4.5 for model 3. It is clear from looking at this graph that the model was run for a total of 12 epochs. where both the training precision and the number of epochs are continually being improved upon. However, the accuracy of the validation is superior to the accuracy of the training. which indicate that this model is not a more suitable choice for our dataset.

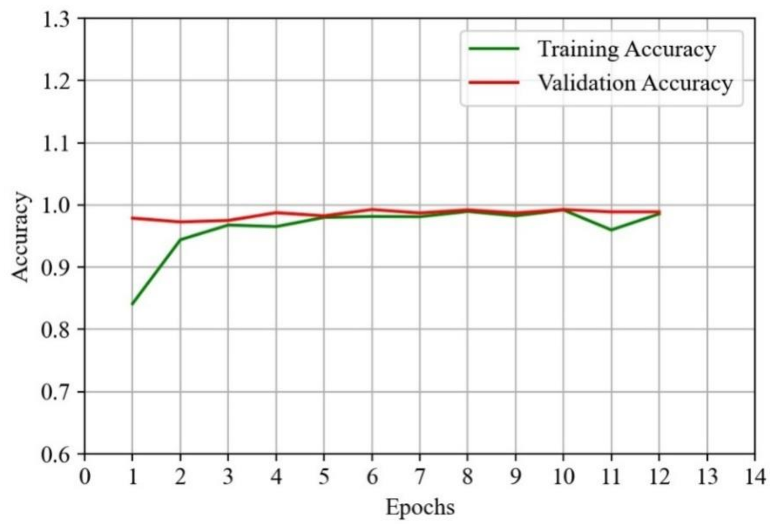


Figure 4.5: Training vs validation accuracyof vgg16.

Figure 4.6 illustrates the disparity that exists between the training loss and the validation loss for model 3. This graph demonstrates that the amount of training loss is progressively decreasing up until it reaches 10 epochs, but that when it hits 10 epochs, it begins to gradually grow. This trend continues until it reaches 20 epochs, at which point it begins to steadily decrease again. On the other hand, validation loss has been steadily climbing to ever-higher levels in tandem with the expanding number of epochs. This suggests that the model will never be able to provide a solution to my problem that is more efficient.

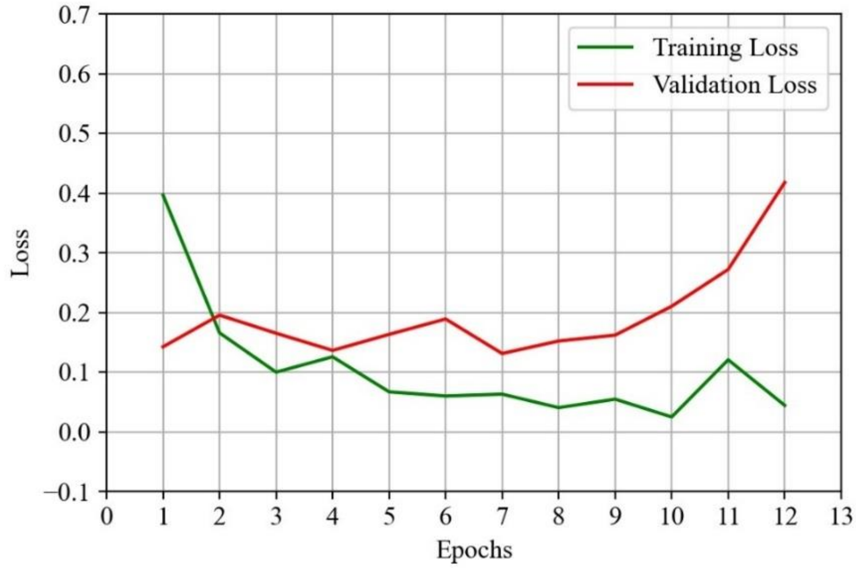


Figure 4.6: Training vs validation loss of model 3.

In Table 4.3, we are able to view the outcomes of the categorization performed by model 2. Based on the figure, it is clear that I applied three indicators in order to assess the quality of the classification report. Gourds are the names of varieties of vegetables that found in the class column. According to the data we have collected on bitter gourd, the recall rate is 1.00, and the precision rate is 1. F1 is ranked at a value of 1.00. In the case of bottle gourd, the accuracy rate is an impressive 0.98. A recall rate of 1,00 is achieved, and the f1-sxore is about 99 percent. Roughly 0.98 percent of the time, ridge gourd is accurate also the f1 score and recall socre 0.98. As far as the snake gourd is concerned, the recall, precision rate, and r1 score are all equal to 1. A recall of 98 percent and a f1 score of 0.98 and precision also 98 percent both associated with the sponge gourd. In addition, the model has an accuracy of 0.99 out of a total of 801 photographs.

From above analysis we can see that accuracy of both model 2 and model 3 is same. But there are some difference. There is very less score is achieved by model 2 about 98 percent. So I choose this model as final model.

Table 4.4 CLASSIFICATION REPORT OF MODEL 3

Class	Precision	Recall	F1-Score	Support
Bitter gourd	1.00	1.00	1.00	65
Bottle gourd	0.98	1.00	0.99	80
Ridge gourd	0.98	0.98	0.98	177
Snake gourd	1.00	1.00	1.00	216
Sponge gourd	0.98	0.98	0.98	263
Accuracy				801
Macro Avg	0.99	0.99	0.99	801
Weighted Avg	0.99	0.99	0.99	801

4.6 Evaluation

Differences between observed and expected values for sponge gourd, snake gourd, bottle gourd, bitter gourd, and ridge gourd are shown in Fig. 9. This evaluation is performed by real data which is never seen before by my model. Actual data is shown in blue bar, whereas the predicted data is shown in orange bar. For bitter gourd there are 65 real data available my model accurately detected all of them, for bottle gourd 80 image data was real and my model predict 80 out of 80 correctly. For ridge gourd 177 image is available and my model predict 177 correctly, for snake gourd there are 216 files available and my model also correctly detected these files. Finally for sponge gourd there are 263 image is available but my model correctly recognize about 210 images.

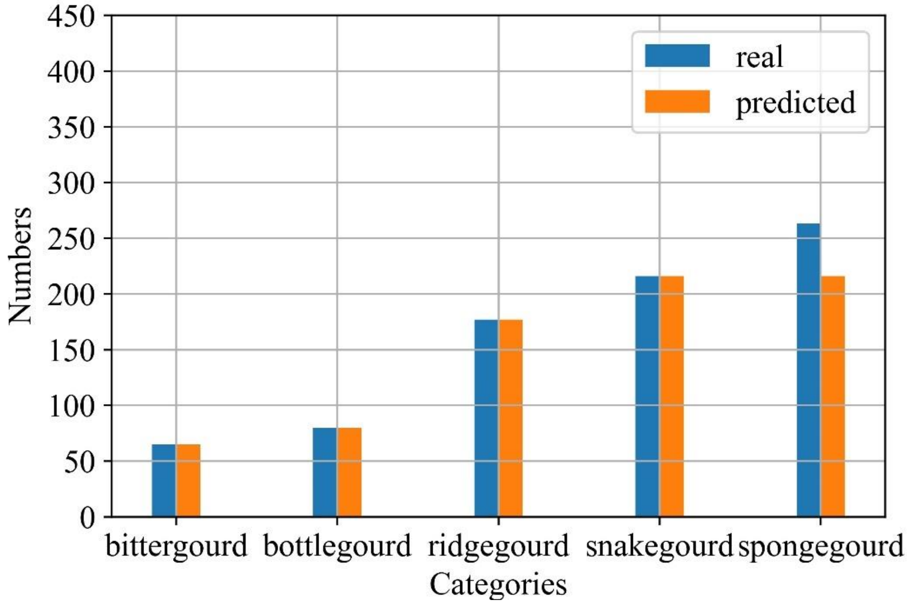
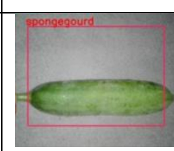


Fig. 9. Comparison of the test dataset.

The picture data utilized in this study consists of three types of vegetables, specifically ridge sponge gourd, snake gourd, bottle gourd, bitter gourd, and ridge gourd. The machine was trained using data from three different veggies in an attempt to recognize and classify them using computer vision techniques with the OpenCV library. The model used for testing was Model 2.

Table 4.5 opencv implementation

Real					
Predicted					

CHAPTER 5

IMPACT ON SOCIETY, ENVIRONMENT AND SUSTAINABILITY

5.1 Impact of Society

The primary effects of this attempt may be broken down into two distinct areas. The first is the influence on agriculture, and the second is the impact on technology. I presented a study that demonstrated how gourd vegetables look remarkably similar to one another, making it extremely challenging for non-farmers to differentiate between the varieties. In cases when correct identification of the gourd vegetable can result in improved market categorization.

5.2 Impact on Environment

This initiative has the potential to result in an augmentation of agricultural production, which may be seen as an environmental impact. Several aspects, including as energy usage, data center management, long-term sustainability, and the implementation of carbon offset strategies, have the potential to influence the environment. The energy consumption associated with the training of deep learning algorithms is evidently substantial because to the high computational demands, necessitating a significant reliance on non-renewable energy sources. The implementation of this project requires the utilization of costly cloud resources.

5.3 Ethical Aspects

The ethical problems that accompany this undertaking center on the protection of individuals' privacy and the safety of their data. To ensure the safety of all of the photographic data, we are taking exceedingly cautious precautions. We have additionally archived the model's weight and matrix data using the.h5 file format, which is known for its reliability. In conclusion, we think that the fact that our model was trained using a dataset that was evenly distributed provides it a significant absence of bias and makes it more fair in this regard.

5.4 Sustainability Plan

In order to improve this project, I intend to maintain both a certain degree of scope and long-term goals. One of my aims is to expand the dataset so that I may collect information on a larger range of issues; for example, the current dataset includes the classification of five different gourd vegetable species. All applicable data privacy standards will be followed, and data privacy and security will be maintained, as they are the primary objectives of this work.

CHAPTER 6
CONCLUSION AND FUTURE WORK

6.1 Summary of the Study

The utilization of automated vegetable recognition is seeing a growing trend. Despite the little familiarity that many individuals may have with gourds, there is a unanimous and enthusiastic endorsement of these botanical entities. The present thesis aims to provide a comprehensive analysis of a computer vision and deep learning framework designed to autonomously differentiate among five distinct types of perplexing gourds. The identification of vegetables is a well-studiedtopic, however, the automated classification of gourd vegetables is a significant challenge due to its reliance on factors such as item placements, forms, and colors. Nevertheless, a significant proportion of individuals exhibit a comparable physical manifestation. In order to tackle this matter, I want to construct three distinct convolutional neural network (CNN) models. Additionally, I presented the most optimal model that has been developed thus far, utilizing the outcomes obtained from our training process. Based on my own observations, it has been found that the most optimal outcomes are achieved by removing the background and making edits to the images. I, as well, attain favorable outcomes with enough practice and evaluation.

6.2 Conclusion

In this work I develop three different convolutional neural network to solve the gourd species recognition problem. I named my model as model 1 model 2 and model 3. My custom cnn model 2 architecture perform better among all the model. It generates about 99% accuracy on test data which is very permissible. I used this model for evaluation and prototyping with opencv library.

6.3 Recommendations

Some recommendation of my work is given bellow:

1. To improve data processing effectiveness and outcomes.
2. Improved data processing can recognize gourd vegetables.
3. To improve algorithm performance, consider raising complexity or switching to a specialized deep learning framework.

6.4 Future Work

I create three convolutional neural networks to recognize gourd species. Model 1, 2 and 3 are my models. My unique CNN model 2 architecture outperforms all others. It has 99% test data accuracy, which is acceptable. I evaluated and prototyped this model with opencv. In future I will increase the dataset by adding more species of gourd vegetables. I finally I will make an android app based on this research.

REFERENCE

1.Dubey SR, Jalal AS (2015) Application of image processing in fruit and vegetable analysis: a review. J Intell Syst 24:405–424

2. El-Salam A (2012) Application of computer vision technique on sorting and grading of fruits and vegetables

3. Zeng G (2017) Fruit and vegetables classification system using image saliency and convo-lutional neural network. In: IEEE 3rd Information technology and mechatronics engineer-ing conference (ITOEC), pp 613–617

4. Sakai Y, Oda T, Ikeda M, Barolli L (2016) A vegetable category recognition system using deep neural network. In: 10th International conference on innovative mobile and internet services in ubiquitous computing (IMIS), pp 189–192

5.Muhtaseb A, Sanaa S, Hashem T (2012) Fruit/Vegetable recognition. In: Proceedings of the 2012 student innovation conference, Hebron, Palestine

6. Dubey SR, Jalal AS (2013) Species and variety detection of fruits and vegetables from images. Int J Appl Pattern Recognit 1:108–126

7. Xiaojun J, Jun C, Yong C (2021) Weed identification using deep learning and image pro-

cessing in vegetable plantation. IEEE Access 9:10940–10950

8. Ercole C, Gallo MD, Mosiello L, Baccella S, Lepidi AA (2003) Escherichia coli detection in vegetable food by a potentiometric biosensor. Sens Actuators B-Chem 91:163–168

9. Kaur M, Reecha S (2015) ANN based technique for vegetable quality detection. IOSR J Electron Commun Eng (IOSR-JECE) 10(5):62–70

10. Kagaya H, Aizawa K, Ogawa M (2014) Food detection and recognition using convolutional neural network. In: Proceedings of the 22nd ACM international conference on multimedia

11. Pouladzadeh P, Abdulsalam Y, Shervin S (2015) Foodd: food detection dataset for calorie measurement using food images. In: International conference on image analysis and pro-cessing. Springer, Cham

12. Probst Y, Nguyen DT, Tran MK, Li W (2015) Dietary assessment on a mobile phone using image processing and pattern recognition techniques: algorithm design and system proto-typing. Nutrients 7(8):6128–6138

13. Rimi IF, Habib MT, Supriya S (2022) Traditional machine learning and deep learning mod-eling for legume species recognition. SN Comput Sci 3:430

14. Nuruzzaman M, Hossain MS, Rahman MM, Shoumik AS, Khan MA, Habib MT (2021) Machine vision based potato species recognition. In: 5th International conference on intelli-gent computing and control systems (ICICCS), pp 1–8

APPENDIX

One way to elucidate the challenging analytical procedures we underwent. Furthermore, there is a dearth of existing research in this particular domain. Our supervisor provided us with a significant amount of encouragement. Obtaining information was a significant obstacle for our team as well. A manual data gathering initiative was initiated. Conversely, the many employment categories present an extra hurdle. Through great exertion, we were able to achieve that outcome.

PLAGIARISM REPORT

Gourd Vegetable Detection by Deep Learning Approach

ORIGINALITY REPORT

24%	18%	13%	13%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Md. Mehedi Hasan, Khairul Alam, Sunzida Siddique, Tofayel Ahamed Topu, Md. Tarek Habib, Mohammad Shorif Uddin. "Chapter 2 Deep Learning Modeling for Gourd Species Recognition Using VGG-16", Springer Science and Business Media LLC, 2023 Publication	7%
2	dspace.daffodilvarsity.edu.bd:8080 Internet Source	6%
3	Submitted to CSU Northridge Student Paper	3%
4	Submitted to Daffodil International University Student Paper	1%
5	www.unitedgeneticsindia.com Internet Source	1%
6	123dok.com Internet Source	<1%