FRESH AND ROTTEN FRUITS CLASSIFICATION USING DEEP LEARNING ALGORITHM.

BY

SUMAEYA AKTER

ID: 221-25-128

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

Supervised By

MD. SAZZADUR AHAMED
Assistant Professor
Department of CSE
Daffodil International University

Co-Supervised By

MR. ABDUS SATTAR
Assistant Professor
Department of CSE
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

17 January 2023

APPROVAL

This Thesis titled **"FRESH AND ROTTEN FRUITS CLASSIFICATION USING DEEP LEARNING ALGORITHM"**, submitted by Sumaeya Akter, ID No: 221-25-128, 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 M.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on 17-01-2023.

BOARD OF EXAMINERS

Chairman

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

Internal Examiner

Ms. Naznin Sultana
Associate Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Internal Examiner

Mr. Md. Sadekur Rahman
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

External Examiner

Dr. Mohammad Shorif Uddin, PhD
Professor
Department of Computer Science and Engineering
Jahangirnagar University

DECLARATION

We hereby declare that this project has been done by us under the supervision of **Md.Sazzadur Ahamed**, Assistant Professor, 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:

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

Co-Supervised by:

Mr.Abdus Sattar
Assistant Professor
Department of CSE
Daffodil International University

Submitted by:

Sumaeya Akter

Sumaeya Akter
ID: 221-25-128
Department of CSE
Daffodil International University

©Daffodil International University ii

ACKNOWLEDGEMENT

First, I express our heartiest thanks and gratefulness to almighty God for His divine blessing making us possible to complete the final year project successfully.

I really grateful and wish our profound our indebtedness to Supervisor **Md.Sazzadur Ahamed**, Assistant Professor, Department of CSE Daffodil International University, Dhaka. Deep Knowledge & keen interest of our supervisor in the field of "Machine Learning and Artificial Intelligence" 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 Thesis.

I would also like to thank my co-supervisor Mr. **Abdus Sattar**, Assistant Professor, Department of CSE Daffodil International University, Dhaka. When i face any problem, he helped us with valuable ideas and suggestions. He motivated us and help us to complete this work.

I would like to express our heartiest gratitude to **Prof. Dr. Touhid Bhuiyan**, Head, of the Department of CSE, for his kind help to finish our project and also to other faculty members and the staff of CSE department of Daffodil International University.

I would like to thank our entire coursemates at Daffodil International University, who took part in this discussion while completing the coursework.

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

ABSTRACT

Producing fruit was a common and important part of agricultural life in Bangladesh. The growing of fruit has gone through a period of profound change. We obtain nutrients from fruits. It is a provider of vitamin and mineral supplements. Additionally, rotten fruit can be seen toward the end of the container in the consumer. As a result, I devised this research plan in an attempt to stop the fruit from going bad. The most recent breakthroughs in computer vision and deep learning are being utilized in our process to determine whether or not the fruit has gone bad. At this point, everything is sophisticated and dependent on new technology. Everywhere we look, we see people using their phones and other electronic devices. Following that, a device was constructed right here. A clever and speedy piece of equipment that can distinguish between fresh and rotting fruit. We make use of the CNN architecture of the red fruit and the clean fruit in our work. Because it makes use of photographic data, this device is able to determine the percentage of red and fresh fruit. If we had a vast farm and a savory factory where we could detect rotting fruit, then this device might make our work simpler. We want our artwork to be even more precise, but we need to do it in a lot less time. With an accuracy of nearly 98%, Model 3 delivered the best results of 10000 natural image dataset. Therefore, for the implementation of our job, I used model number 3.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Introduction

A lot of seasonal fruits are available in Bangladesh. These fruits are produced in large quantities at a single time, i.e., fruit is available in the market for a specific period. This is due to the climate of Bangladesh. For this, to meet the needs of the people of our country throughout the year, we either have to store many fruits or else we have to rely on fruit juice. Many problems arise when many fruits are bought or stored together. Seasonal fruits begin to rot very quickly, so it becomes complicated to choose to keep many of them at once. This becomes very difficult, especially for industries, as they supply fruit juice to the market throughout the year. So, they must separate the good fruit from the rotten fruit. A fungus forms on the fruit's skin, spreading widely to other fruits. It causes a lot of damage. It is challenging for humans to sort out good fruit and rotten fruit without any intelligence system. An excellent way to quickly select rotten and good fruits is with pictures of fruits. Computer vision and deep learning techniques are perfect for detecting rotten fruit and sound fruit by image processing. Picture A critical test for investigating human reasoning and deep learning is the classifier. This one is probably the most popular progress and exam topic to address our daily issues and realities. We all encounter various problems in our daily lives that can be effectively handled and solved using the human brain's classification and processing power By observing scenarios like comprehending which organic product is acceptable or the degeneration of which food is excellent for your health or even worse, I can tell which illness is affected or which organic product or plant we should eat. I've struggled to find a problematic solution to the problem of good and ruined organic food. I will utilize calculations and human reasoning to decide which biological product is unique and which is destroyed by CNN architecture. We now purchase natural products even though the superstore provides them due to the illumination. Everyone may see how sifting soil products causes them to want to spend money and buy the newest organic product in malls using my framework.

1.2 Motivation

Nowadays, many people purchase fruits from supermarkets or wholesalers. Lack of time prevents them from determining whether the fruit is fresh or rotting. They never inspect fruits before buying them because doing so takes time. Because of this, a lot of edible fruit goes to waste every year. The economic and health costs of eating these rotten, toxic fruits are enormous. We plan to create an automated method to assess if the fruit is fresh or rotten to help with this cost load and solve related health issues. With the advancements in AI and image processing technologies currently being made, we can create these solutions from scratch. As a result, this research work lessens this challenge and develops a system that allows for quick scanning and results in delivery.

1.3 Rational of The Study

Recent advances in artificial intelligence have prompted revolutions across several industries. Artificial intelligence helps us find answers to numerous issues. CNN, an example of an image classifier, is used to detect cancer cells and diagnose leaf illnesses, among many other applications. Therefore, we believe it holds the key to our problem's resolution. As an example, the classification of fresh food is now a common use of AI, and it exemplifies the original challenge. In light of this, we have settled on employing AI to determine whether or not a piece of fruit is ripe. Using state-of-the-art methods in artificial intelligence and By processing photos, we could build these tools from scratch. Thankfully, this problem has been mitigated, and a system is in place that lets scanner users access the data they've gathered. In this research, we aimed to create a self-correcting system using deep learning to circumvent this problem. In terms of fruit condition, our technology can tell the difference between fresh and spoiled. And when it comes to identifying fruits, it's 98% accurate.

1.4 Research Questions

- The data set will be produced when and how?
- Is it possible to distinguish between the behavior of fresh and rotten fruits?
- How are rotten and fresh fruits classified?
- How will this work benefit people?

- Do people use it easily?
- what will arise the new problems for its usability?

1.5 Expected Outcome

The state-of-the-art in artificial intelligence and image processing has allowed us to create these solutions from scratch. I simplify this process and develop a means for users to scan and obtain the resulting data. We have created a deep learning-based technological intelligence system to avoid this issue. According to the condition of the fruit, our method may be able to distinguish between fresh and damaged fruit. The illness can be correctly diagnosed in 98% of cases.

1.6 Project Management and Finance

I used Google Collaborate's premium version to enhance the 12-hour limit to 24 hours for algorithm tuning. And the price is just $10.

1.7 Report Layout

Chapter 1: This introductory chapter will describe my program, its goals, the problem I was attempting to tackle, our proposed solution, and an estimated completion date. This section is also where I detail the rationale for our study.
Chapter 2: Chapter 2 discusses the background of the study, relevant research, and the current situation in Bangladesh. In addition to a detailed examination of the work's historical context, a quick summary of that context's history is also offered.
Chapter 3: The research strategy will be outlined in Chapter 3. This section elaborates on the process or methodology. Procedures for collecting data will be outlined here.
Chapter 4: In Chapter 4, we'll take a look at the proposed model's efficacy with the use of a classification report and an accuracy table.
Chapter 5: As the final section of the study, Chapter 5 wraps up everything discussed throughout. Here, we will briefly review the model's output. In this part, we also give a comparison of the two methods' accuracy. The model's online output and implementation are discussed here as well. The

chapter concludes with a critical analysis of the book's weaknesses. It also subtly includes details about upcoming initiatives.

1.8 Summary

The fundamental structure of our organization is laid forth in this chapter. This is the first and most important chapter in our book. This chapter presents an outline of our general framework, along with a few related frameworks, our inspirations, our objectives, and our commitments to this framework. Finally, in this chapter, I look at my overall framework strategy and the several approaches that could be taken to address our particular predicament.

CHAPTER 2

BACKGROUND STUDY

2.1 Terminologies

There have been a lot of studies done on the diagnostic and therapeutic potential of herbal remedies for a variety of diseases. In the realm of prediction, one of the most common applications of deep learning and master learning may be found. There have been a variety of investigations on the possibility of discovering methods to either predict or identify plant diseases. In this work, the problem is solved by investigating and applying a number of different machine-learning algorithms. This article provides a concise summary of the labor put in by a great number of specialists in the field.

2.2 Related Works

In this section, we'll talk about what we've accomplished. The results of overviews of our studies on the related work are presented in Section II, where we also examine this endeavor's potential benefits and drawbacks.

Azizah et. al. [1]. developed a machine vision system for detecting defects in fruit skin in the study. SVM is a machine-learning technique. The vector machine can handle a large enough dataset (SVM). The transfer to a machine learning approach relies heavily on the accuracy of the keys and characteristics used, so their accuracy is paramount to the classification process as a whole. Gains in efficiency are possible with deep learning models. These models aid in the classification of massive image datasets.

VenkataRamReddy et. al. [2] Reddy has a leg up on the competition when it comes to defect and non-standard fruit quality thanks to his ground-breaking expertise in image processing. Use it to find imperfections on the mango's skin. First, the researcher hand-selects imperfect fruits and divides them into quality categories. With an accuracy rate of 97.5%, this model was very close to flawless. This cutting-edge system analyzes fruit quality in real time over the internet using Laser Backscattering Imagery and the CNN Hypothesis.

For whatever reason, that may be. At work, all three authors [3] were forced to employ machinery that could identify imperfect fruit even when the skin was still on. SVM classifies color characteristics and structure. Support vector machines only give accurate findings on a select few data sets (SVMs). When using machine learning to sort data, the features and qualities utilized to analyze and classify the data are crucial—using models to increase comprehension and performance. These techniques can be used to model images for large-scale databases.

By recognizing and categorizing convex surfaces in RGBD images, Nyarko et al.[4] have created a brand-new approach for classifying fruits. The RGB-D image is first divided into thirds by a ridge on the convex side of the picture. Each convex surface is then given a suitable descriptor description, which is used to classify the surface. As a fresh method of describing surfaces that are roughly convex, the CTI (Convex Temple Instance) descriptor is suggested. The approach approximates surfaces by using the descriptors that each side of the convex polyhedron represents. Because of this, computing descriptors is a quick and effective procedure. The descriptor used to describe 3D point clouds and the SHOT descriptor is quite comparable. Two iterations of the CTI and SHOT descriptors—one that uses colour and the other that doesn't—are compared and contrasted in this evaluation. Using a k-neighbour classifier, measured surfaces are split into two categories: fruit and other. The accuracy of its numerical calculations is the critical advantage of the proposed expert system over earlier fruit recognition methods. Precision is crucial because the system is intended to work as a stand-alone fruit picker.

The apples' freshness or decay was determined by analyzing the defects Roy et al. [5] caused in the peel. Semantic deep learning methods isolate the spoiled region from the rest of apple's RGB image. UNet and its improved successor, Enhanced Units, are useful tools for achieving accurate segmentation (En-UNet). If you're looking for improved accuracy above UNet's 95.36 percent, the proposed En-UNet model is the way to go. Training and validation accuracy were both 97.46%. When compared to En-IoU UNet's average of 0.95%, UNet's score of just 0.6% seems almost laughable.

To help blueberry breeders, Ni et al. [6] did a thorough study segmenting blueberry fruit qualities such as compactness of the cluster, fruit ripeness, and berry abundance. If you want to know when your blueberries are ready to pick, you can use a disguise. A thorough evaluation of the completed r-CNN model has been conducted. Validation data accuracy is 78.3%, and test data accuracy is 71.61% when using a union threshold of 0.5 or above.

Brahimi et al. [7] used a sizable dataset to predict various contemporary technical developments. This dataset has 14828 sick tomato leaves in total. Convolutional neural networks (CNNs) are a type of machine learning that trains categorization without the usage of custom functions using picture files as input. It was initially intended to be a teaching strategy. We used visual representation techniques for symptom localization and recognition in order to evaluate the quality of the deep model that was given to us. The research's findings are exciting since they provide models that are incredibly shallow, reach an accuracy of 99.18%, and could be useful tools for safeguarding farmers against risks associated with tomatoes.

2.3 Comparative Analysis and Summary

I chose CNN because,

- After reviewing several scholarly articles and projects, we can confidently say that it functions admirably with agricultural, natural, and other categories. What we need for our purposes is best served by this.
- Comparatively, CNN has the highest accuracy of all trained image classification algorithms, at 100% percent or more.
- It's user-friendly, and lots of tools can be put to use for expansion.
- Picture comparisons, such as between new and ripe red fruits, are also very effective. Not to your face, in your face, etc.
- It's possible that with proper use of CNN layers and sufficient training, I can get valuable results with high accuracy.

2.4 Scope of the Problem

I've been working on a system that can easily distinguish between rotten and edible produce. In conclusion, the convolutional neural network is quite useful. Not too far in the future, we'll be able to put these kinds of technologies to use in a number of other settings, including huge industrial malls. All of our work will be freely available to anyone who wants it, and we'll do everything in our power to make that process as easy as possible. Many wholesale markets across the country sell red fruit, which makes them ideal places for magistrates to do their jobs.

2.5 Challenges

The most challenging aspect of our work is organizing data sets for use in the future. Our data set or future changes have been accurately determined using powerful ML and image processing technologies. We also face the problem of a lack of available resources and employment opportunities in Bangladesh.

Data Collection:

Accurate data collection is essential for deep learning. We couldn't gather as much data because deep understanding requires a lot of data from different wholesale locations and marketplaces. We only captured 10000 raw pictures of both fresh and rotten food. as a result of the fruit vendors' permission for us to purchase their fruit from them.

Model Selection:

For deep learning, a wide variety of models are available. It is crucially essential for them to pick a suitable model. Choosing the right model and reliable data is, in fact, a breeze. To classify images, a wide variety of models exist. Using the vgg16 and resnet50 architectures, we analyzed how well our model performed. To implement our model, we turn to the Google TensorFlow library.

CHAPTER 3

RESEARCH METHODOLOGY

3.1 Introduction

The methods and practical considerations of my research are discussed here. The tools, data, subjects, preprocessing, stats, and deployment are all covered. What you're looking at is a schematic used in academic study. Our research process has been laid out in a logical order for you to follow along. The method is depicted in Figure 3.1.

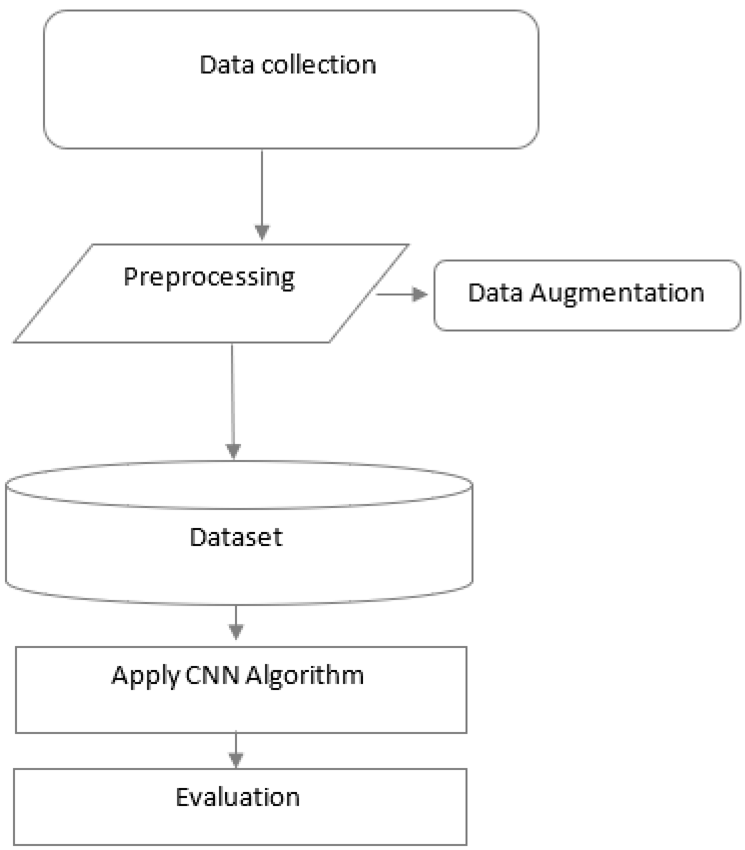

Figure 3.1 Methodology Diagram

3.2 Data Collection

The collecting of exhaustive amounts of data is essential for having an in-depth understanding, just as it is with machine learning. Due to the fact that Deep Learning requires a substantial amount of data, we collected it from a wide variety of retailers and other retail locations. We were on our own when we took shots of the fruit. We took almost eleven thousand images in total. We chose three different fruits and photographed them from a variety of perspectives before selecting our favorites. I took three of my fruits and photographed them while they were still fresh, and I also photographed three of my fruits after they had spoiled. The outcome was one, but we now have a level six dataset with three good results and many bad photos of these three results. The outcome was one. In order to accomplish this, we were required to visit a number of the neighborhood markets and obtain the proprietors' approval before photographing the fruit stalls within them. In spite of the fact that it required a significant amount of time, the quality of our data improved significantly, and we obtained a large number of photos for categorization. Our research utilized this information well as a data set.

3.2.1 Data Pre-processing

Following the completion of the data collection phase, we put it through two distinct processing steps before utilizing the method. Enhanced data. More data. According to the authors of the study, the risk of making a mistake in classification can, in the vast majority of cases, be mitigated by the use of preprocessing. [10]

3.2.2 Data augmentation:

The augmentation strategy that was implemented during the data increase operation is shown in figure 3.2. A number of different methods, including scale, Tilt, rotational, diagonal flip, and perpendicular flip, were utilized in the process of increasing the data. The better data can be augmented, the better the program will work.

Data augmentation:

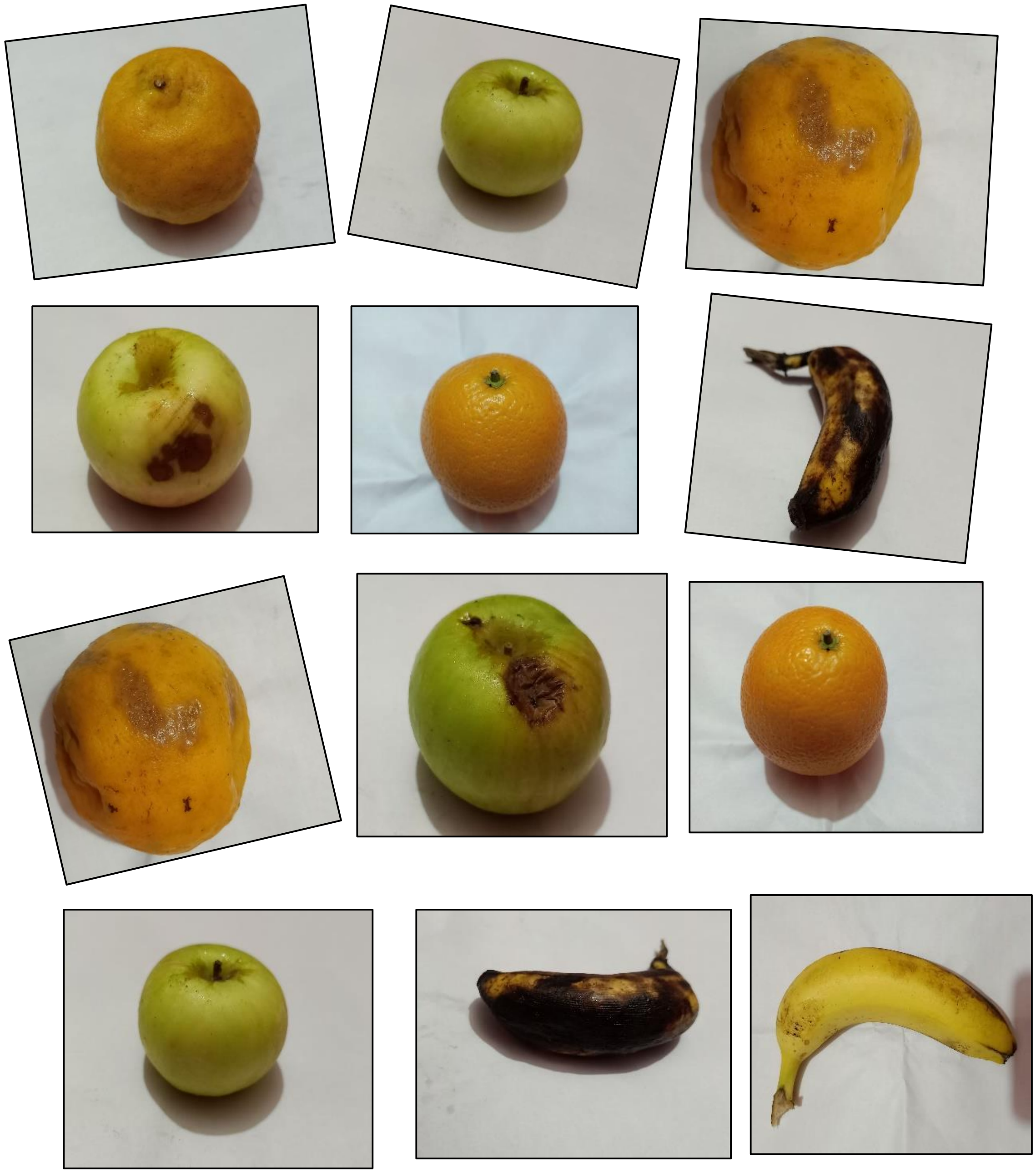

Figure 3.2: Data augmentation

There was a lot of noise in our collection that we made transparent in this step. I have removed the background of the data in the above image. It has facilitated machine understanding. Then increased the quality of the photos to work with them at any scale. Another task in this step is to rotate to a specific angle so that our algorithm in Apple can understand the data. I have tried to do the best possible way of augmentation so that the data is elementary and transparent for the machine to understand

3.2.3 Dataset

After completing the process of improving the pictures, the total number of photographs in my collection was 10000. My picture dataset includes photos that have had their backgrounds removed as well as been preprocessed and supplemented. In addition, some of the data that was collected in its original form is included in this collection. The information I have been working with can be seen in graphical form in the following picture.

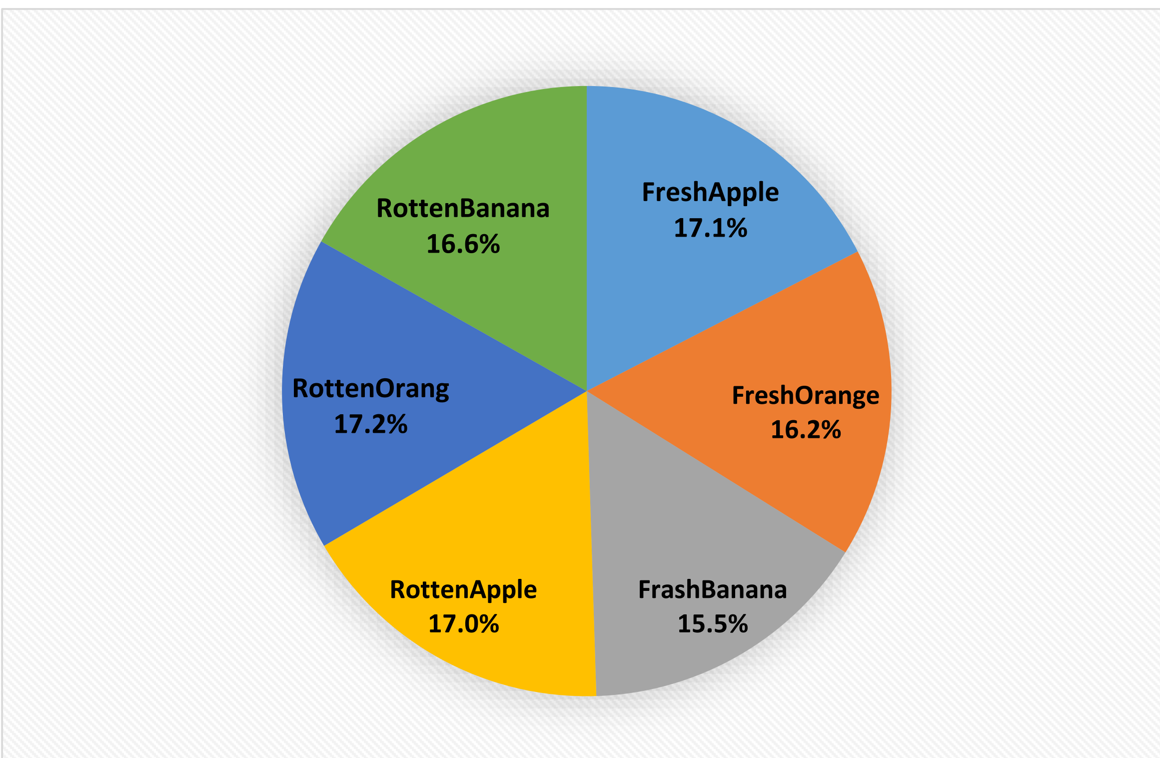

Figure 3.3 Training dataset representation

As part of the process of illustrating my training data set, I sorted all 10,000 of our photographs into one of six groups. To begin, fresh apples contain 17.1%, fresh bananas include 15.5%, and fresh oranges contain 16.2%. These three levels of fresh fruits are then compared to one another. After that, I have three levels of rotten fruits to make the dataset more versatile. The rotten orange level is 17.2%, the rotten banana level includes 16.6%, and the rotten apple level has 17.0%. Finally, I gave it my best shot at dividing the overall number of images into a comparable amount for each category. The dataset I would be using for my research seems to be in good shape, but when I trained on it, I discovered it was a very good dataset.

Figure 3.3 where you can find the representation graph that pertains to the training dataset.

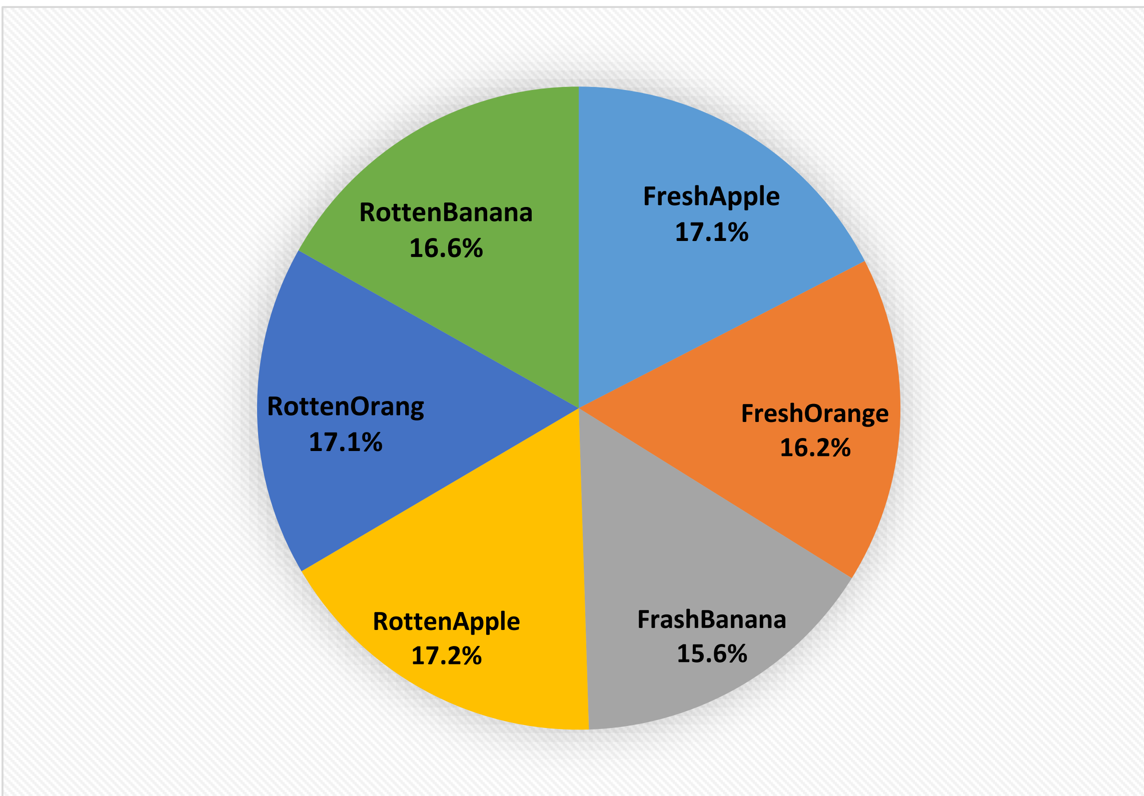

Figure 3.4 Validation dataset representation

As a step toward illustrating the validity of my data collection, I organized all 10,000 of our images into one of the six categories that were available to me. To begin, freshly picked apples have 17.1%, freshly picked bananas have 15.6%, and newly picked oranges have 16.2%. The quality of these three tiers of fresh fruits is then evaluated in relation to one another. After that, I have three different levels of spoiled fruits so that the dataset can be used in a variety of contexts. The level of rot in the rotten orange is 17.1%, the level of rot in the rotten banana is 16.6%, and the level of rot in the rotten apple is 17.1%. In the end, I gave it the best try I could by attempting to divide the total number of photographs into an amount that was equivalent for each category. When I trained on the dataset I would be utilizing for my research, I realized that it was a decent dataset. This is despite the fact that the dataset appeared to be in good shape. From this observation, it can be said that my dataset is excellent. There is no significant difference between the training data and the validation dataset. Every Value came almost the same. A slight difference was observed between the two levels, which was very little. This implies that our dataset is transparent and workable. The representation graph for the validation dataset can be found in Figure 3.4. This is the location where you can find it.

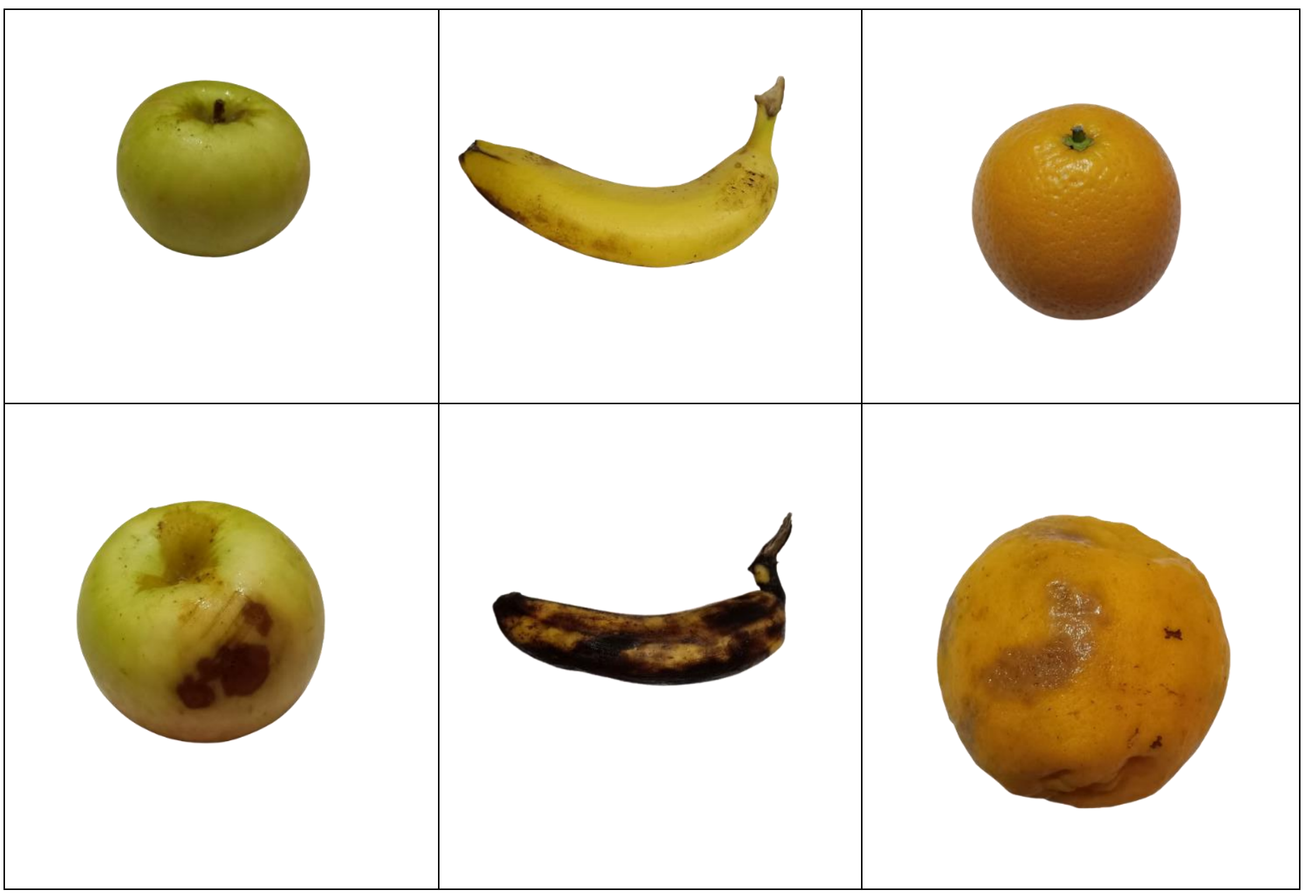

Figure 3.5: Dataset representations

Figure 3.5 displays the final classification we came up with for the dataset. Without any noise, all of our fruit data is presented here for our perusal. It does not appear that there is a problem with the rotation. It demonstrates a very pleasant organization of the dataset. It has been demonstrated that the project classification dataset is validated to what extent.

3.3 Statistical Analysis

This part shows the statistical analysis of our dataset. Here I show the result of the training and testing dataset evaluation. It is a very well prepossessing dataset and it shows a very well statistical analysis report.

Table 3.1: Training data Size

Fruits Name	**Amount**
Rotten Apples	1817
Rotten Bananas	1816
Rotten Oranges	1700
Fresh Apples	1900
Fresh Banana	1850
Fresh Oranges	1830

Table 3.2: Testing data size

Fruits Name	Amount
Rotten Apples	395
Rotten Bananas	381
Rotten Oranges	388
Fresh Apples	601
Fresh Banana	503
Fresh Oranges	403

I have analyzed 10000 pieces of data for training. My dataset level was six. At the end of the analysis, the report gave very good results. All the classes are in very close numbers. No means that our training data set is balanced.

3.4 Applied Mechanism

Convolutional networks view images not as two-dimensional patterns but rather as three-dimensional volumes, which is counter to human perception. Convolutional networks, in contrast to the two-dimensional field of vision we possess, are able to discern the depth of information contained within an image.

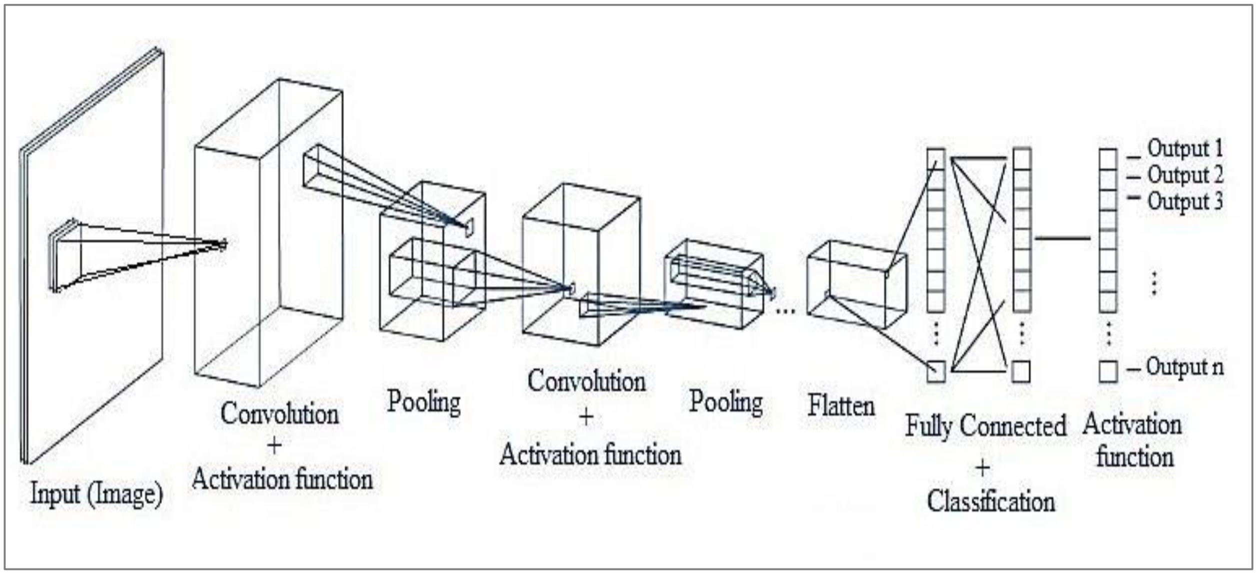

Figure 3.6 CNN architecture

As a result of the RGB encoding of an image, a convolutional network needs to first consume the image in three distinct depth layers or channels before it can process the image. Before being utilized by the convolutional network, pixel squares are subjected to filtering. The job of a filter, which is also known as a kernel, is to search for patterns within a collection of pixels. The number of steps that a filter's path through an input image takes is related to the size of the activation map that represents that path. In order to make graphics, patterned pixels are utilized. The creation of new activation maps has resulted in the production of this brand-new volume. Because of the larger dimensionality of these images, their processing of them requires a significant investment of both time and the computational resources available. Convolutional networks have developed a solution to this problem by utilizing techniques for dimensionality reduction, such as filter stride and downsampling.

3.5 Prerequisites for Implementation

In order to implement the project, we made use of a variety of machine-learning libraries, which are all displayed below in their respective versions.
python:

The most recent release of Python is version 3.8. It is a programming language of the highest possible standard. It is utilized by the majority of studies to carry out their research. As a result of how easy it is to learn, it is highly recommended for programming languages that are used in occupations that are dependent on AI, and it is especially well-liked by the future generation of programmers.

Google CoLab:

In Google CoLab, the distribution of the open-source programming language Python is available for usage at no additional cost. Even though we might do our work online using Jupiter Notebook, we still have unrestricted access to a virtual GPU. The primary benefits of participating in this Google CoLab are as follows.

Specifications for both hardware and software include: an operating system (Windows 10 is recommended), a hard drive, a web browser with more than 4 gigabytes of storage space, and a certain amount of random access memory (RAM) (more than 4 GB)

CHAPTER 4

EXPERIMENTAL RESULT AND EVALUATION

4.1 Introduction

The scientific evidence and experimental procedure are the primary themes of this chapter. When we assess it, the first thing we do is think about the findings. The Implications section ought to be constructed in such a way that the findings are presented without the necessity for awareness or examination on the part of the reader. The section on research papers has some suggestions for further reading. Documentation of the results of the test and the findings will be included in this chapter. The findings ought to be included in the section titled "Implications," but the reader should not be aware of them or examine them. You might find some helpful hints in the area that is dedicated to research papers. This is where the results of the test, as well as an analysis of it, will be displayed. All of the algorithm results and models that contribute to the success of my project are explained in detail here. In this section, I provide details of an explanation of all of the used models and algorithms.

4.2 Experimental Result

When conducting research, it is common practice to experiment with changing the values of one or more variables that are not reliant on the outcome of the investigation [11]. As a consequence of this, it needs more than just a single training course in a variety of settings to carry out an experiment involving machine learning.

At this stage, I perform an analysis of the findings from my study, and I make use of three models in order to make the findings from the applied algorithm more evident. I do this so that I can better understand the implications of the results. The three different types each display their unique take on three various classification reports. In addition, I used a variety of parameters in order to evaluate the reliability of the results that were provided by each model. I did this so that I could decide whether or not the findings were accurate. Due to the precision, recall, and f1 parameters of the three models, the conclusions that they arrive at are understandable and satisfying. In the decision, in order to determine the reliability of my overall findings, we made use of three distinct parameters. The degree of precision was relatively uniformly distributed across all three criteria.

Table 4.1 Accuracy Table for Model 1

Label Name	Precision	Recall	F1-Score	Support,
Fresh Apple	0.91	1.00	0.95	345
Fresh Banana	0.99	0.99	0.99	330
Fresh Orange	0.97	0.95	0.96	330
Rotten Apple	0.96	0.95	0.95	555
Rotten Banana	1.00	1.00	1.00	465
Rotten Orange	0.98	0.93	0.96	313
Accuracy			0.97	2340
Macro avg	0.97	0.97	0.97	2340
Weighted avg	0.97	0.97	0.97	2340

The report on Model 1's classification can be seen in Table 4.1 below. The apple, the orange, and the banana were the three common fruits that I used for my research. I compared the fresh versions of these fruits to their counterparts that had already begun to rot. The results for Precision, Recall, and F1-Score for the level Rotten Banana were all 1, which is the same as the height number for all of the other levels. In the scenario involving Fresh apple, three metrics are carried out admirably: the Precision value is 0.91%, the Recall value is 1.00, and the F1-Score value is 0.95%. The following maximum value of Precision, Recall, and F1-Score for fresh bananas is 0.99%. This value reflects the closest highest value. The Precision, which is displayed as 0.97%, the Recall, which is expressed as 0.95%, and the F1-Score are expressed as 0.96%. It is not a value that is too far apart from our primary accuracy. Apple that has gone bad has a Precision score of 0.96%, a Recall score of 0.95%, and an F1-Score of 0.95%. The final level demonstrates a solid value with a Precision score of 0.98%, a Recall score of 0.93%, and an F1-Score of 0.96% for the rotten apple.

When it came to good pages, there was not much of a difference in Precision, Recall, or F1-Score. The results for the Macro average and the Weighted average are the same for all three parameters. Both the Macro avg and the Weighted avg provided findings that were equivalent in terms of accuracy. This indicates that our algorithms have successfully produced accurate and reliable results. Based on the findings of our research, Table 1 has an accuracy of 97% out of a total of 2340 data sets.

Table 4.2 Accuracy Table for Model 2

Label Name	**Precision**	**Recall**	**F1-Score**	**Support,**
Fresh Apple	0.96	1.00	0.98	345
Fresh Banana	0.99	1.00	0.99	330
Fresh Orange	0.99	0.95	0.97	330
Rotten Apple	0.96	0.96	0.96	555
Rotten Banana	0.99	1.00	1.00	465
Rotten Orange	0.97	0.96	0.96	313
Accuracy			0.98	2340
Macro avg	0.98	0.98	0.98	2340
Weighted avg	0.98	0.98	0.98	2340

The analysis of Model 2's classification is presented in Table 4.2, which can be found below. Apples, oranges, and bananas are the three most prevalent fruits in the world, and I utilized them as examples in my research. When I compared the fresh versions of these fruits to their equivalents that had already started to decompose, I found that the fresh versions were superior. The scores for the level Rotten Banana's Precision are 0.99%, and its Recall and F1-Score were 1, which is the same as the height number for all of the previous levels. When it comes to the situation involving fresh apples, three criteria are performed exceptionally well: The value of the Precision component is 0.96%, the value of the Recall component is 1.00, and the value of the F1-Score component is

0.98%. For rotten apples, the following maximum value for Precision, Recall, and F1-Score is 0.96%. This number is the highest value that is closest to this one. The Recall value, which is shown to be 1%, the Precision value, and the F1-Score value, all of which are shown as having the same value for Fresh Banana as 0.99%. It is not a value that is significantly off from the primary precision that we strive for. The F1-Score for Fresh Orange is 0.97%, its Precision score is 0.99%, while its Recall score is 0.95%, and its F1-Score is also 0.99%. The final level exhibits a strong value with a Precision score of 0.97%, a Recall score of the same value, and an F1-Score with the same value as the Rotten Orange. There was not a significant gap in Precision, Recall, or F1-Score when it came to good pages. All three parameters produce identical results for the Macro average as well as the Weighted average. Findings derived from the Macro average and the Weighted average were virtually identical in terms of how accurately they represented the data. This suggests that our algorithms were successful in producing findings that are accurate and reliable. Table2, out of a total of 2340 data sets, has an accuracy of 98%, which we determined based on the outcomes of our research. Its accuracy is improved by one point compared to model 1. Hence it will indicate that it is superior to model 1.

Table 4.3 Accuracy Table for Model 3

Label Name	**Precision**	**Recall**	**F1-Score**	**Support,**
Fresh Apple	0.95	1.00	0.97	345
Fresh Banana	0.99	1.00	1.00	330
Fresh Orange	0.99	0.93	0.96	330
Rotten Apple	0.99	0.95	0.97	555
Rotten Banana	1.00	1.00	1.00	465
Rotten Orange	0.91	0.98	0.95	313
Accuracy			0.98	2340
Macro avg	0.97	0.98	0.97	2340
Weighted avg	0.98	0.98	0.98	2340

The examination of the classification of Model 3 is provided in Table 4.3, which can be located further down this page. I did some of my study using apples, oranges, and bananas as examples because these are the three fruits that are consumed the most frequently all around the world. When I contrasted the fresh versions of these fruits with their counterparts that had already begun to rot, I discovered that the fresh versions of the fruits were superior to their counterparts in every way. The scores for Precision, Recall, and F1-Score on the level Rotten Banana were all 1, which is the same as the height number on each of the earlier levels. When it comes to the circumstances surrounding fresh apples, there are three criteria that have been met particularly well: It has been determined that the Precision component has a value of 0.95%, the Recall component has a value of 1.00, and the F1-Score component has a value of 0.97%. The greatest possible value for Precision is 0.99%, the maximum possible value for Recall is 0.95%, and the maximum possible value for F1-Score is 0.97% for Rotten apples. This particular number represents the maximum possible value that is comparable to this one. The Recall and F1-Score values have both been displayed quite clearly, and it would appear that they have reached an identical point. This point is indicated to be 1%, and the Precision value for Fresh Banana is 0.99%. This value does not deviate from the primary Precision that we aim to achieve by a substantial margin. Fresh Orange has an F1-Score of 0.96%, a Precision score of 0.99%, and a Recall score of 0.93%; additionally, its F1-Score is also 0.99%. The last level has a strong value, with a Precision score of 0.91%, a Recall score of 0.98%, and an F1-Score value of 0.95% for the Rotten Orange. These three scores combine to make the overall score. When it came to good sites, there was not a discernible difference in Precision, Recall, or F1-Score between the two groups. The Macro average and the Weighted average both yield the same results regardless of which of the three factors are changed. The results obtained from the Macro average and the Weighted average were extremely similar in terms of how precisely they portrayed the facts. This provides support for the notion that our algorithms were successful in providing findings that are accurate and reliable. According to the findings of our investigation, Table3 possesses an accuracy of 98% out of a total of 2340 data sets. We came to this conclusion based on the findings. When compared to model 1 and model 2, its Precision is improved by one point because of this improvement. As a result, it will demonstrate that it is more effective than model 3. Therefore, I decided to go with model 3 for the further implementation of my following step.

Model 1 Analysis

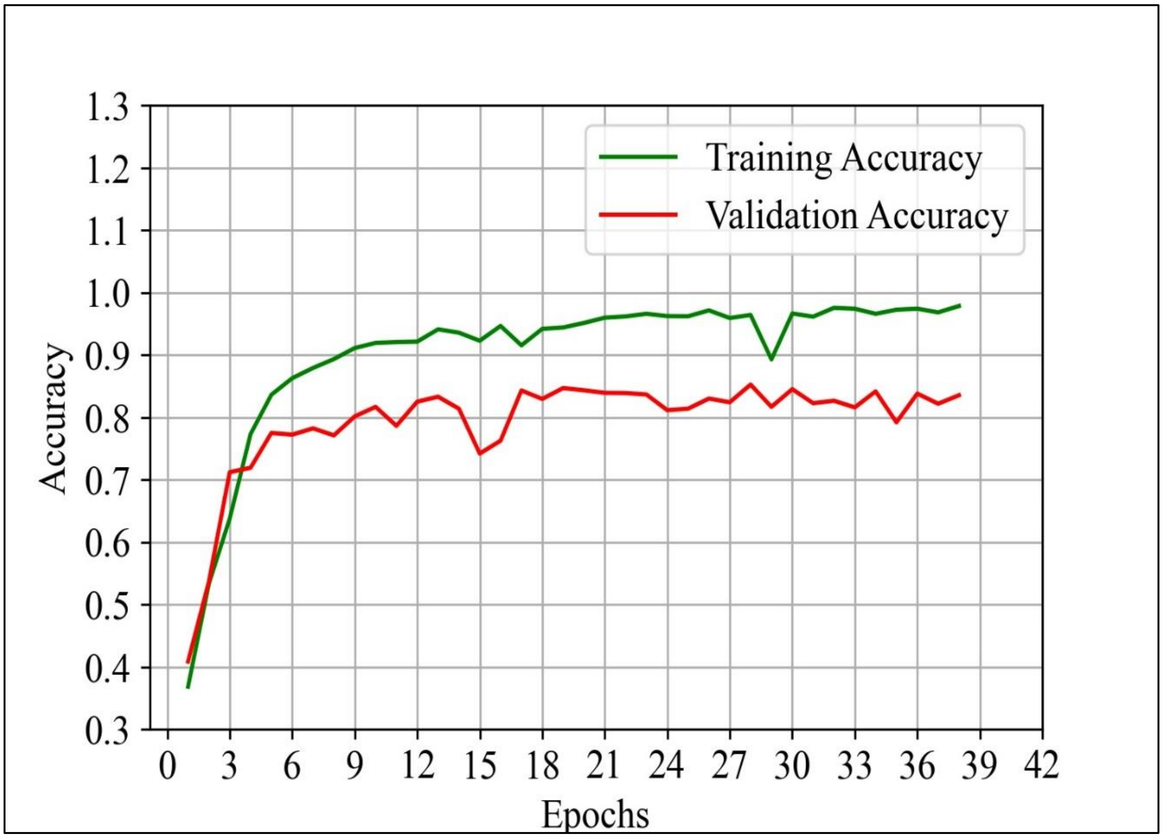

Figure 4.1 Training vs validation accuracy for Model 1

The degree of accuracy with which the first model was trained is illustrated in Figure 4.1. The accuracy of the validation is shown by the red line, while the accuracy of the training is shown by the green line. In this particular model, the accuracy of the training and validation phases steadily improves over time. A zigzag line is produced as a result of training as well as validation. In addition, the two lines of shows each have their own unique characteristics. This event suggests that our dataset has a somewhat slow pace of learning. The accuracy of validation for our model 1 is 97.00 percent, which is the highest percentage that can be equaled by the accuracy of another model. However, it reveals that the dataset is not balanced. This model is not stable for the dataset.

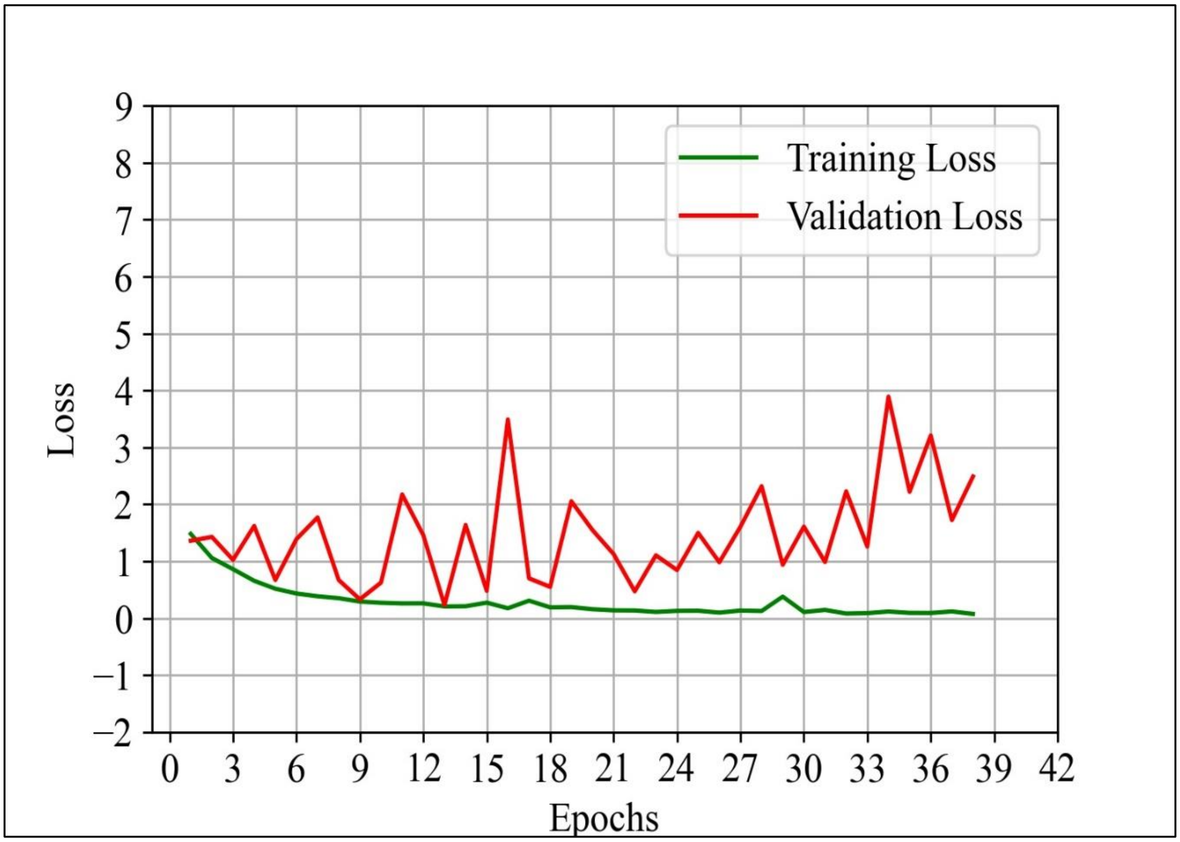

Figure 4.2 Training vs validation Loss for Model 1

A training loss contrasted with a validation loss over time is one of the measurement combinations that is used the most frequently. In contrast to the training loss, which displays how accurately model 1 reproduces the initial data, the validation loss demonstrates how accurately model 1 generates new data. In other words, it evaluates how well model 1 produces new data. One example of the training versus validation of loss model 1 is displayed in figure 4.2. Model 1 produced relatively zigzagging curves for both the guarantee losses and the activity losses. In a similar vein, the validation loss and the training loss are both reduced when the duration of the epochs is stretched out over longer periods of time. Rapid learning takes place, and there is no indication that model 1 has been overfitting to this data set in any way. However, the results demonstrate that a significant amount of data was lost throughout the validation process. The line of training loss represents the state, but validation resulted in a significant loss of data; hence, it is not an appropriate model for the dataset I have chosen. As a result, I decided against using it for my subsequent implementation.

Model 2 Analysis

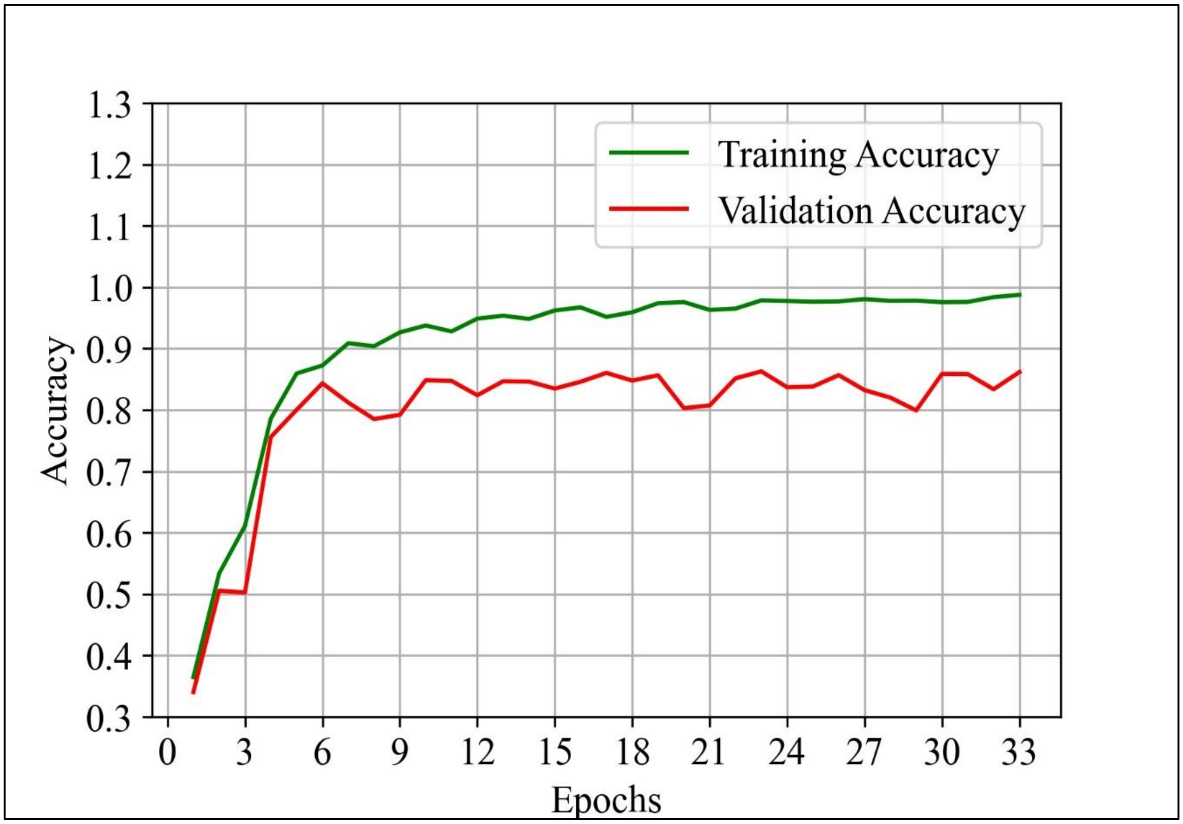

Figure 4.3 Training vs validation accuracy for Model 2

Figure 4.2 provides an illustration of the degree of precision that was achieved through the training of the first model. The red line illustrates how accurate the validation was, while the green line illustrates how accurate the training was. Within the context of this specific model, the precision of both the training and the validation phases continuously increases over the course of time. The combination of training and validation leads to the production of a line from model1 that has fewer zigzags in it. In addition, each of the two distinct lines of shows has its own set of individual qualities. Because of this incident, it appears that our dataset has a somewhat sluggish rate of learning. The accuracy of validation for our model 2 is 98 percent, which is the highest percentage that can be matched by the accuracy of another model. There is no other model whose accuracy comes close to matching ours. Nevertheless, it demonstrates that the dataset does not have a balanced distribution of values. This model2 does not exhibit very high levels of stability for the dataset. On the other hand, it demonstrates that the output is generated far more effectively than the model1.

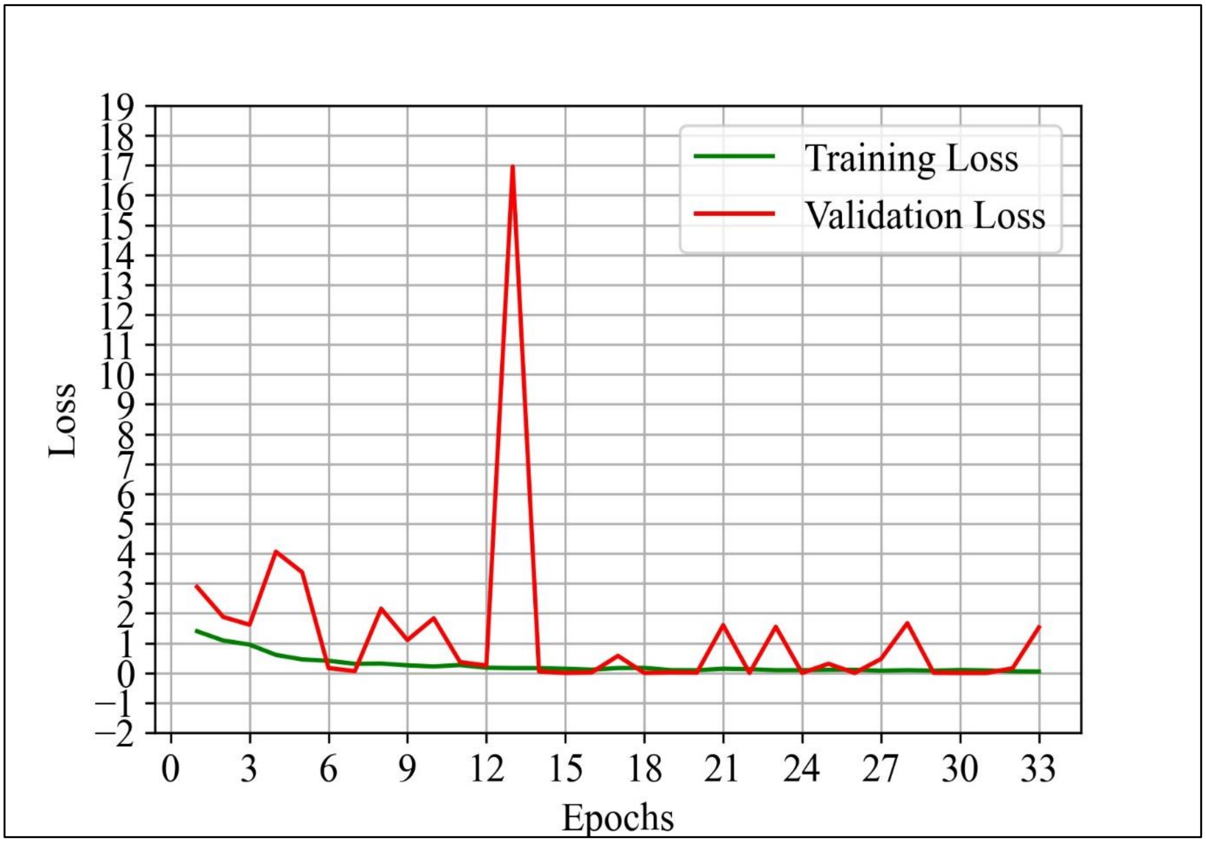

Figure 4.4 Training vs validation Loss for Model 2

Standard methods of measurement often involve a combination of training loss and validation loss over time. In contrast to the training loss, the validation loss shows how well model 2 can make up its own data. Specifically, it assesses how well the second model can produce novel results. Figure 4.4 depicts a hypothetical instance of either training or validating the loss model 2. There was a large discrepancy between the validation and training losses in Model 2. In a single hop, it jumps from a fairly lofty height to a sky-scraping one. Reducing training and validation losses by using longer epochs may show a similar pattern. There is no evidence that model 2 overfits this data set, and the model is able to learn quickly. However, the findings show that a substantial quantity of information was lost throughout the validation procedure. This model is inappropriate for the dataset I've picked because validation caused a large loss of data while the line of training loss indicates the state. So, I decided against including it in my next deployment.

Model 3 Analysis

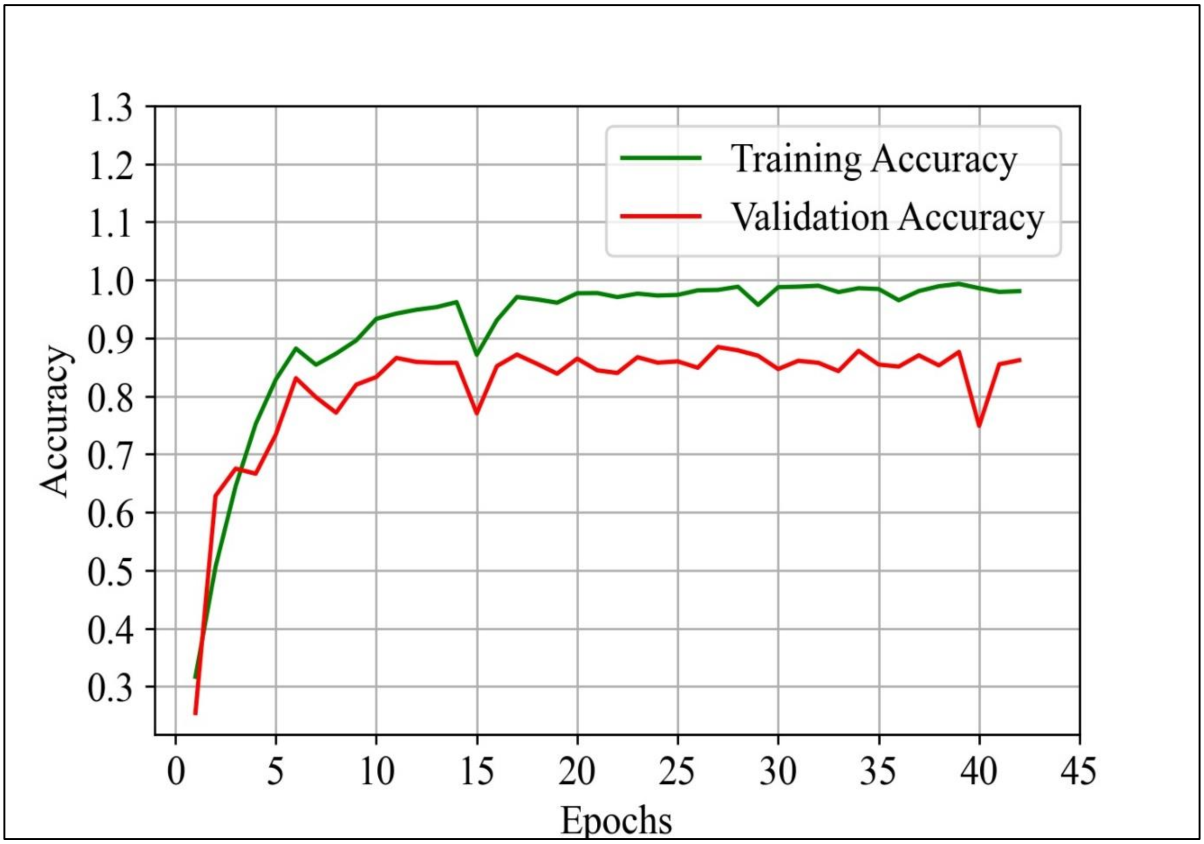

Figure 4.5 Training vs validation accuracy for Model 3

Figure 4.5 provides a visual representation of the level of precision with which the initial model was trained. The red line represents the accuracy of the validation, whereas the green line represents the accuracy of the training. In this particular model, the accuracy of both the training and the validation phases steadily improves throughout the course of their respective time periods. The combination of training and validation will result in the production of a clean zigzag line. In addition, each of the two lines of programs has its own distinct qualities that set it apart from the other. Based on the occurrence of this event, it would appear that our dataset performs better than the other two models. The validation accuracy of our model 3 is 98 percent, which is the greatest proportion possible. On the other hand, it demonstrates that the dataset has a more even distribution than any of the other two models. This model is reliable when applied to the dataset. As a result, I chose the model to be used for the upcoming assignment, and I will use it in the application for my project. It performed exceptionally well in detecting fruits, whether they were fresh or rotten.

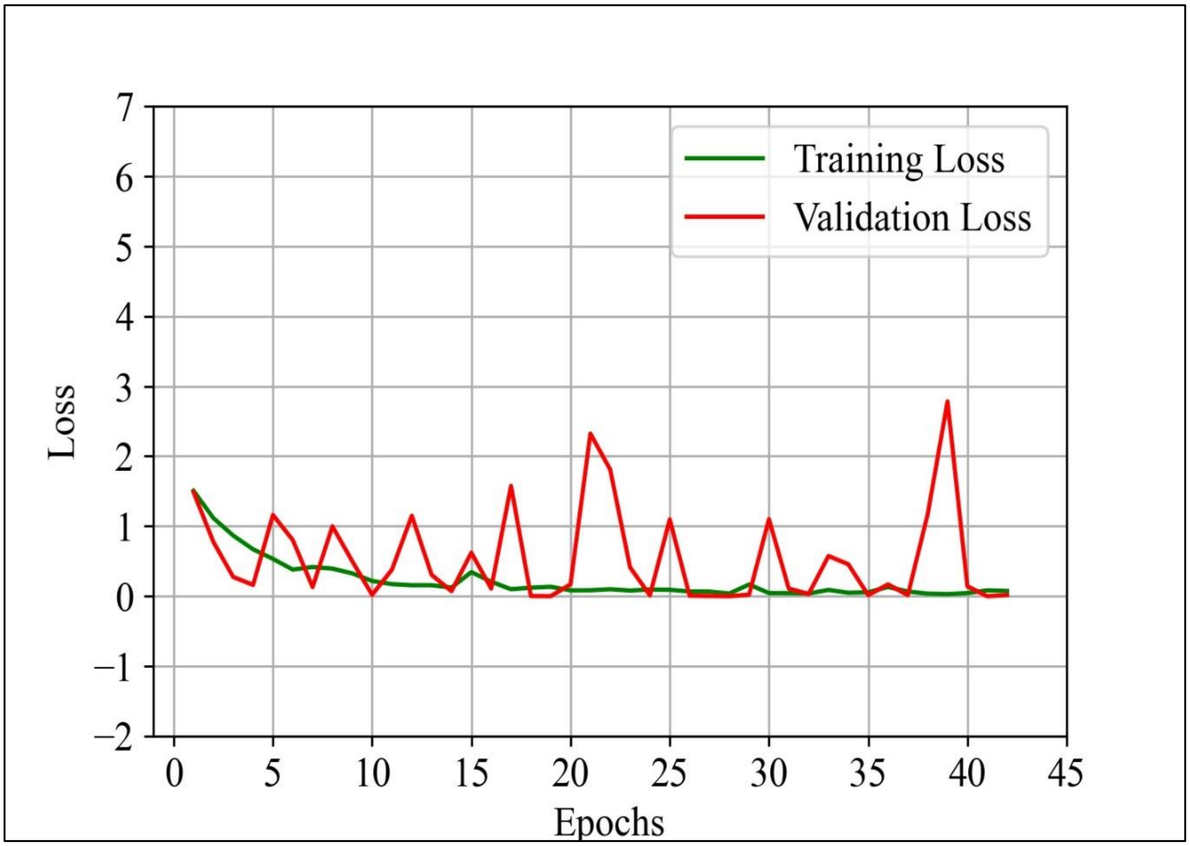

Figure 4.6 Training vs validation Loss for Model 3

The comparison of a training loss to a validation loss over time is one of the most popular measurement combinations. There is a time component attached to this decline. The validation loss, as opposed to the training loss, which illustrates how accurately model 3 reproduces the initial data, demonstrates how accurately model 3 creates new data. In other words, it examines the efficiency with which model 3 generates new data. Figure 4.6 provides a visual representation of one example of the training versus validation process for loss model 3. Model 3 is superior to models 1 and 2 since it generated rather smooth zigzagging curves for both the guarantee losses and the activity losses. When the duration of the epochs is spread out across more extended periods of time, both the validation loss and the training loss are reduced. This is true for both the validation loss and the training loss. There is a rapid increase in knowledge, and there needs to be evidence to suggest that model 3 has been overfitting in any way to the data set that has been provided. The findings, on the other hand, indicate that the validation procedure resulted in a smaller amount of data being discarded than was initially anticipated. However, validation led to a less substantial loss of data, and as a consequence, it is an adequate model 3 for the dataset I have selected to use. The line of training loss represents the state. As a direct consequence of this, I made the decision to incorporate it into my implementation. Because the results demonstrate that it suffers less loss of data than the other two models do. As a result of all of the facts made above, it is reasonable to assert that model 3 is ideal for utilization in more complex activities.

4.3 Evaluation

Inside this section, I have focused on the implications that can be drawn from my findings. First, I have demonstrated that my algorithm can distinguish between fresh and rotting fruit. In order to accomplish this, I began by gathering the necessary training data and determining whether or not our implemented algorithm was accurate. Thus, the implementation model can readily identify those fruits with the name of the effect, whether fresh or rotting fruits, and it also creates the names of the fruits themselves. Here discovers the primary purpose of my project.

Fresh Apple

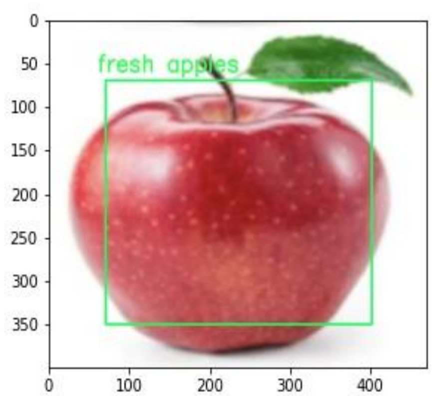

Figure 4.7 Fresh Apple Detect

The initial testing results of my implementation algorithm are displayed in Figure 4.7. I implemented the model that I had chosen after selecting a processing Fresh apple to have an effect on the Fruits. It achieves remarkable levels of success. It is able to distinguish freshly picked apples precisely. The actual inspection of the fruit is displayed in the box that is colored green. It is also fully capable of recognizing the names of fruits.

Rotten Apple

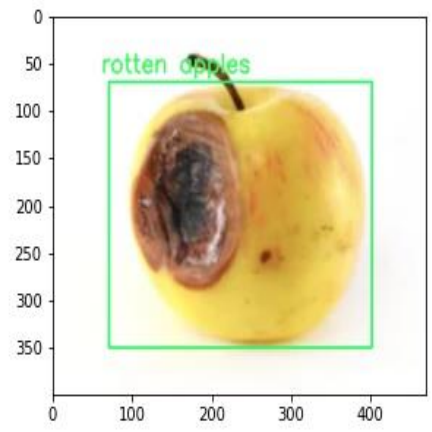

Figure 4.8 Rotten Apple Detect

Figure 4.8 provides a display of the results of the second test I ran on my implementation algorithm. Following the selection of a processing method that would have an impact on the Fruits, I put into action the model that I had decided to go with. It is incredibly successful on many different fronts. First, it is accurate in its identification of rotten apples that have been selected. The actual examination of the fruit is depicted in the green box that is located in the middle of the chart. Additionally, it is completely capable of identifying the names of various fruits.

Fresh Orange

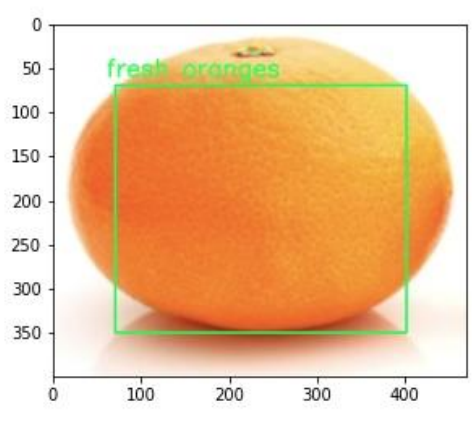

Figure 4.9 Fresh Orange Detect

My implementation algorithm's preliminary test results are shown in Figure 4.7. After deciding on a model that would have an effect on the Fruits throughout processing, I went with a Fresh Orange. Again, extremely impressive results are obtained. It can tell the difference between stale apples and those just picked. The red box represents the actual inspection of the fruit. It can also identify fruits by name with high accuracy.

Rotten Orange

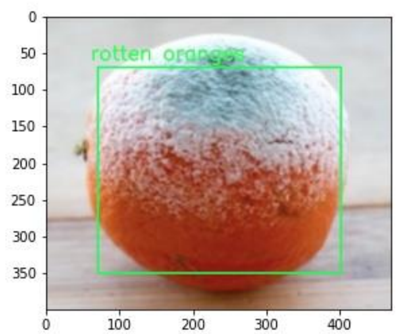

Figure 4.10 Rotten Orange Detect

In Figure 4.9, you can see the results of the second test I ran on my implementation method. As soon as I settled on a processing method that would have an influence on the Fruits, I put my chosen model into action. There are so many ways in which it succeeds that it would be tedious to mention them all. To begin, it can determine whether or not a particular Orange is rotten. The precise inspection of the fruit is depicted in the red box at the centre of the chart. Furthermore, it can correctly identify the names of various fruit.

CHAPTER 5

CONCLUSION AND FUTURE WORK

5.1 Summary

Finding out whether a particular image shows decaying fruit or not is one of the main objectives of our study. Our research has mainly focused on this issue. We did see CNN's meticulously precise image classification in operation. It's true. We began by compiling information from numerous sources for our investigation. After some kind of preliminary processing, the training and evaluation were divided into two independent stages. We shall succeed in achieving our objectives soon.

5.2 Conclusion

In the agriculture industry, sorting produce to differentiate between fresh and damaged produce is essential. This study presents a brand-new convolutional neural network (CNN) model for classifying rose and red fruits. The results demonstrated that the proposed model 3 could more precisely and firmly discern between fresh and decaying fruit. Using the proposed neural network model, model 3 could automate the brain imaging technique to tell the difference between fresh and rotten fruits and reduce the number of mistakes made when sorting red and mint fruits. The proposed CNN model was right about 98% of the time.

5.3 Implementation for future work

The following are the guidelines for the continuation of the work's production:

- Since CNN requires a sizable amount of data, I plan to enhance this model by incorporating more data.
- Another more nuanced use of classes is on the horizon.
- I plan to release a web app and an android app in the near.
- In the future I will add here AI base system that provides notification of the observation of the user.

- In this research only three fruits are classified and detected in the future I will add more fruits for detection
- In the future, I will add the percentage of any fruit that is rotten or good.

.

REFERENCE

[1] Azizah, L.M.R., Umayah, S.F., Riyadi, S., Damarjati, C., Utama, N.A. (2017). Deep learning implementation using convolutional neural network in mangosteen surface defect detection. In 2017 7th IEEE International Conference on Control System, Computing and Engineering (ICCSCE), pp. 242-246. https://doi.org/10.1109/iccsce.2017.8284412

[2] VenkataRamiReddy, C., Kishore, K.K., Bhattacharyya, D., Kim, T.H. (2014). Multi-feature fusion based facial expression classification using DLBP and DCT. International Journal of Software Engineering and Its Applications, 8(9): 55-68. https://doi.org/10.14257/ijseia.2014.8.9.05

[3] Wang, L., Li, A., Tian, X. (2013). Detection of fruit skin defects using machine vision system. In 2013 Sixth International Conference on Business Intelligence and Financial Engineering, pp. 44-48. https://doi.org/10.1109/bife.2013.11

[4] Emmanuel Karlo Nyarko, Ivan Vidović, Kristijan Radočaj, Robert Cupec, A nearest neighbor approach for fruit recognition in RGB-D images based on detection of convex surfaces, Expert Systems with Applications, Volume 114, 2018, Pages 454-466, ISSN 0957-4174, https://doi.org/10.1016/j.eswa.2018.07.048.

[5] Roy, K., Chaudhuri, S.S. & Pramanik, S. Deep learning based real-time Industrial framework for rotten and fresh fruit detection using semantic segmentation. Microsyst Technol 27, 3365–3375 (2021). https://doi.org/10.1007/s00542-020-05123-x

[6] Ni X, Li C, Jiang H (2020) Blueberry harvest ability trait extraction from 2D images and 3D point clouds based on deep learning and photogrammetric reconstruction. In 2020 ASABE Annual International Virtual Meeting (p. 1). American Society of Agricultural and Biological Engineers.

[7] Brahimi, M., Boukhalfa, K. & Moussaoui Deep learning for tomato disease: classifcation and symptoms visualization. Appl. Artif. Intell. 31, 299–315 (2017).

[8] M. Mehedi Hasan, M. Omar Faruk, B. Biswas Biki, M. Riajuliislam, K. Alam and S. Farjana Shetu, "Prediction of Pneumonia Disease of Newborn Baby Based on Statistical Analysis of Maternal Condition Using Machine Learning Approach," 2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence), 2021, pp. 919-924, doi: 10.1109/Confluence51648.2021.9377169.

[9] H. Qassim, A. Verma and D. Feinzimer, "Compressed residual-VGG16 CNN model for big data places image recognition," 2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC), 2018, pp. 169-175, doi: 10.1109/CCWC.2018.8301729.

[10] E. Rezende, G. Ruppert, T. Carvalho, F. Ramos and P. de Geus, "Malicious Software Classification Using Transfer Learning of ResNet-50 Deep Neural Network," 2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA), 2017, pp. 1011-1014, doi: 10.1109/ICMLA.2017.00-19.

[11] Mithu M.A., Momo S.I., Hasan M.M., Rahman K.M., Sattar A. (2022) Pumpkin Leaf Disease Detection: Convenience of CNN Over Traditional Machine Learning in Terms of Image Classification. In: Somani A.K., Mundra A., Doss R., Bhattacharya S. (eds) Smart Systems: Innovations in Computing. Smart Innovation, Systems and Technologies, vol 235. Springer, Singapore. https://doi.org/10.1007/978-981-16-2877-1_32

[12] Theckedath, D., Sedamkar, R.R. Detecting Affect States Using VGG16, ResNet50 and SE-ResNet50

Networks. SN COMPUT. SCI. 1, 79 (2020). https://doi.org/10.1007/s42979-020-0114-9

APPENDIX

The first problem we had when doing the analysis was establishing the analytical technique for our investigation. It wasn't a standard job, and little had been done in this subject previously. As a result, we weren't able to get much help from any source. We also started gathering data by hand. After a long time of hard labor, we might be able to achieve it.

PLAGIARISM REPORT

ROTTEN_FRUITS_CLASSIFICATION_USING_DEEP_LEARNING_A... 2.pdf

ORIGINALITY REPORT

18%	15%	9%	%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	dspace.daffodilvarsity.edu.bd:8080 Internet Source	12%
2	eprints.utar.edu.my Internet Source	1%
3	"Advances in Data Computing, Communication and Security", Springer Science and Business Media LLC, 2022 Publication	1%
4	medium.com Internet Source	<1%
5	Zhang Yibo, Liu Qi, Hao Peifeng. "Vehicle Type Classification System for Expressway Based on Improved Convolutional Neural Network", 2020 3rd International Conference on Artificial Intelligence and Big Data (ICAIBD), 2020 Publication	<1%
6	towardsdatascience.com Internet Source	<1%