SMART FARMING: SOLANACEAE FAMILY VEGETABLE PLANT DISEASE DETECTION USING DEEP LEARNING

Submitted By:

ANKON CHOWDHURY

191-35-2682

Department of Software Engineering

DAFFODIL INTERNATIONAL UNIVERSITY

Supervised By:

MD. SHOHEL ARMAN

Assistant Professor

Department of Software Engineering

DAFFODIL INTERNATIONAL UNIVERSITY

This Thesis report has been submitted in fulfillment of the requirements for the Degree of Bachelor of Science in Software Engineering

Spring-2023

APPROVAL

This thesis titled on "Smart Farming: Solanaceae Family Vegetable Plant Disease Detection Using Deep Learning", submitted by **Ankon Chowdhury (ID: 191-35-2682)** to the Department of Software Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of Bachelor of Science in Software Engineering and approval as to its style and contents.

BOARD OF EXAMINERS

--- **Chairman**

Afsana Begum
Assistant Professor
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University

--- **Internal Examiner 1**

Md Shohel Arman
Assistant Professor
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University

--- **Internal Examiner 2**

Md. Rajib Mia
Lecturer
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University

--- **External Examiner**

Dr. Md. Sazzadur Rahman
Associate Professor
Institute of Information Technology, Jahangirnagar University

DECLARATION

It hereby declares that this thesis has been completed by me under the supervision of **Md. Shohel Arman,** Assistant Professor, Department of Software Engineering (SWE), Daffodil International University. It is also declared that neither this work nor any part of this has been submitted elsewhere for the award of any degree by me.

Ankon

Ankon Chowdhury

Student ID: 191-35-2682

Department of Software Engineering

Faculty of Science & Information Technology

Daffodil International University.

Certified By:

Md. Shohel Arman

Assistant Professor

Department of Software Engineering

Faculty of Science & Information Technology

Daffodil International University.

ACKNOWLEDGEMENT

I want to begin by thanking Almighty God for His favor and for enabling me to finish my undergraduate thesis.

I want to show my appreciation and immense respect to my supervisor**, Md. Shohel Arman, Assistant Professor** in the “Department of Software Engineering” at “Daffodil International University” in Dhaka. His profound knowledge and guidance in the segment on “Deep Learning” helped me a lot to complete this entire thesis work. It has been made possible by his never-ending empathy, academic leadership, continuous motivation, regular and vigorous monitoring, constructive criticism, helpful counsel, reviewing numerous subpar manuscripts, and fixing them at every level.

I would like to express my gratitude to **Dr. Imran Mahmud**, the **head of the "Software Engineering" Department** within the Faculty of Science and Information Technology, as well as to the other professors, faculties, and staff members of the SWE Department at "Daffodil International University" for their thoughtful assistance in completing my work.

Last but not least, I must respectfully thank my parents for their unwavering love and patience.

ABSTRACT

Plant leaf diseases have recently become a major worry for farmers. This is a danger to both farmers and the country, which depends on agriculture. In this aspect, conventional identification techniques may not be very efficient. The country's agriculture industry is already suffering from a lack of sophisticated identification methods. In this project, my overarching goal is to use smart farming to battle plant diseases, such as those that impact trees and crops. This thesis focuses on applying deep learning for smart farming to identify diseases in plants from the Solanaceae family. The goal is to create an automated illness detection system that is accurate and effective, allowing for prompt treatments. Two deep-learning models are trained on a large dataset, and their performance is assessed. The results show that the suggested method is successful in locating illnesses, underscoring its potential to enhance agricultural operations. The project investigates the use of deep learning methods in disease diagnosis for plants belonging to the Solanaceae family with the goal of enabling early illness detection and intervention. The research demonstrates the efficacy of the recommended deep learning technique and its potential to improve agricultural practices by creating an effective system utilizing a substantial dataset.

Keywords: Smart farming, disease detection, deep learning, Solanaceae family, timely interventions, agricultural practices, automated illness detection.

CONTENTS

CHAPTER: 01

INTRODUCTION

In recent years, the occurrence of leaf diseases in plants has emerged as a significant concern for farmers. This issue poses a threat to both agricultural productivity and the overall well-being of nations, particularly in agriculture-dependent countries. Traditional methods of disease identification and management may not provide effective solutions to mitigate the impact of these diseases. Moreover, the lack of modern identification techniques further compounds the challenges faced by the agricultural sector.

To address these issues, we worked on this sector using the “Deep Learning” algorithm. As a “Deep Learning approach” we used object detection algorithms to detect diseases- such as leaf movements, disease transmission, and movement. This research aims to combat plant diseases, including those affecting trees and vegetables, through the implementation of smart farming practices. Smart farming leverages advanced technologies and data-driven approaches to optimize agricultural processes and enhance crop management. In this thesis, the specific focus is on employing deep learning techniques for disease detection in Solanaceae family vegetables.

This study's main goal is to create an effective and accurate method for automatically detecting illness in plants from the Solanaceae family. The project seeks to diagnose illnesses automatically, allowing prompt treatments, by employing a large dataset and training two deep-learning models. Rigid testing and performance assessment will be used to determine the viability of the suggested strategy.

The significance of this research lies in its potential to improve agricultural practices by providing early disease detection and intervention mechanisms. By harnessing the power of deep learning, farmers can proactively manage plant diseases, minimize crop losses, and optimize resource allocation. Additionally, the study will explore the application of deep learning techniques in disease detection for Solanaceae family vegetables, contributing to the advancement of knowledge in this field.

Through this thesis, we seek to emphasize the importance of embracing technological advancements, such as deep learning, in the field of agriculture. By developing an efficient system for disease detection, we aim to empower farmers with tools and knowledge to make informed decisions and mitigate the impact of plant diseases. Ultimately, this research aims to pave the way for improved agricultural productivity and sustainable farming practices.

Nowadays, 'VGG-16' and 'ResNet-50' is attracting the attention of developers for their higher accuracy in disease detection. Both VGG16 and ResNet-50 are popular deep-learning models for disease detection. VGG16 is known for its simplicity and effectiveness in extracting visual features, while ResNet-50 addresses the challenges of deep networks with skip connections. They can accurately classify diseases by capturing patterns and features in medical images. Their importance lies in their ability to learn hierarchical representations and handle complex visual characteristics, aiding in early diagnosis and treatment planning. Researchers often experiment with these models to find the most suitable architecture for their specific disease detection tasks.

1.1 BACKGROUND

In recent years, the occurrence of leaf diseases in plants has become a significant concern for farmers, posing a threat to agricultural productivity and the overall well-being of nations, particularly in agriculture-dependent countries. Traditional methods of disease identification and management may not provide effective solutions to mitigate the impact of these diseases, further exacerbating the challenges faced by the agricultural sector.

To address these issues, this research focuses on the implementation of smart farming practices to combat plant diseases, specifically those affecting Solanaceae family vegetables. Smart farming leverages advanced technologies and data-driven approaches to optimize agricultural processes and enhance crop management. In this thesis, the primary objective is to develop an accurate and efficient system for automated disease detection in Solanaceae family vegetables using deep learning techniques.

The importance of this study rests in its potential to enhance agricultural practices by providing early disease detection and response methods. Farmers can prevent plant illnesses, reduce crop losses, and maximize resource allocation by using the power of deep learning algorithms. Various disease-related characteristics, such as leaf motions, disease transmission, and movement, may be identified via the application of object detection algorithms, which helps to create a more thorough disease detection strategy.

The research utilizes two popular deep learning models, VGG-16 and ResNet-50, known for their accuracy in disease detection tasks. VGG-16 is valued for its simplicity and effectiveness in extracting visual features, while ResNet-50 addresses the challenges of deep networks with skip connections. These models play a vital role in accurately classifying diseases by capturing patterns and features in medical images. The exploration and experimentation with these models aim to identify the most suitable architecture for disease detection in Solanaceae family vegetables.

The basis for this study is backed by the problems that now plague vegetable farming. The average person in the nation barely consumes 125 grams of vegetables per day, compared to the World Health Organization's recommendation of 220 grams for healthy adults. Furthermore, plant diseases cause an astounding 10 to 15 percent of veggies to be wasted annually. This issue is further exacerbated by farmers' inadequate awareness of plant diseases. Vegetable agriculture may be increased and made more lucrative by correctly identifying and eliminating illnesses via the use of targeted fertilizers.

By developing an efficient system for disease detection using deep learning, this research aims to bridge the gap between the challenges faced in vegetable cultivation and the potential for improved agricultural practices. The findings and outcomes of this study can provide valuable insights and empower farmers with the tools and knowledge needed to combat plant diseases effectively, increase crop yields, and contribute to sustainable farming practices.

1.2 MOTIVATION OF THE RESEARCH

The motivation behind my research stems from the pressing need to address the challenges faced in disease detection in leafy vegetables. I am driven by the current scenario, where leaf diseases pose a significant threat to agricultural productivity, and I aim to develop a model capable of accurately detecting diseases in leaves at any time.

The widespread availability and adoption of smartphones among the general population serve as a major motivation for me. By leveraging the ubiquity of smartphones, the model I develop can be easily accessible and adaptable to a wide range of users. This accessibility and ease of use provide a significant impetus to my research, as it enables farmers and individuals involved in agriculture to benefit from the disease detection capabilities of the model.

Moreover, the context of Bangladesh as an agricultural country adds to my motivation. Despite the agricultural significance of the country, proper agricultural services and support systems are lacking. I aim to fill this gap by developing an efficient disease-detection model that can improve agricultural practices in the region. The lack of substantial work on this issue within the local agriculture sector further strengthens my motivation to contribute to this area of research.

By developing a model that can effectively detect leaf diseases in a user-friendly manner, my research aims to empower farmers, agricultural stakeholders, and individuals involved in the agricultural sector. My motivation lies in providing accessible and reliable tools to support disease management and enhance agricultural productivity. Ultimately, I aspire to bridge the gap between the current challenges faced in the agriculture sector and the potential benefits offered by technological advancements in disease detection.

1.3 PROBLEM STATEMENT

The problem at hand is the increasing occurrence of leaf diseases in Solanaceae family vegetables, which poses a significant threat to agricultural productivity and the overall well-being of agriculture-dependent countries. Traditional methods of disease identification and management have proven to be inadequate in mitigating the impact of these diseases. Additionally, the lack of modern identification techniques further compounds the challenges faced by the agricultural sector.

Moreover, the limited availability of proper agricultural services and support systems exacerbates the problem. Farmers often lack the necessary knowledge and resources to effectively detect and manage plant diseases, resulting in significant crop losses and reduced profitability. The current practices rely heavily on manual inspection and subjective judgment, which are time-consuming, prone to errors, and often lead to delayed interventions.

Another issue is the lack of comprehensive research and development in disease detection for Solanaceae family vegetables, particularly in the context of smart farming and deep learning techniques. The existing literature is scarce, and there is a need for tailored solutions that leverage advanced technologies to improve disease detection accuracy and timeliness.

Thus, the main problem addressed in this document is the need for an accurate and efficient system for automated disease detection in Solanaceae family vegetables. The system should utilize deep learning techniques and leverage the power of smart farming to enable timely interventions, minimize crop losses, and enhance agricultural practices. The solution should be accessible, user-friendly, and capable of detecting a wide range of leaf diseases with high accuracy. By addressing this problem, the research aims to contribute to the advancement of disease detection capabilities in the agriculture sector, empowering farmers and improving the overall sustainability and profitability of vegetable cultivation.

1.4 RESEARCH QUESTIONS

1. Is the deep learning model effective for disease detection in Solanaceae family vegetables in the context of smart farming?

2. Can the accuracy of the deep learning model be accurately measured using a comprehensive dataset specifically collected for this research?

3. Is the convolutional neural network (CNN) suitable for accurately classifying leaf diseases in Solanaceae family vegetables within the scope of this research?

4. How can deep learning techniques be effectively applied for automated disease detection in Solanaceae family vegetables in the context of smart farming?

5. What is the performance and accuracy of deep learning models, specifically VGG-16 and ResNet-50, in detecting various leaf diseases in Solanaceae family vegetables?

6. How can a comprehensive dataset be collected and utilized to train deep learning models for accurate disease detection?

7. What are the key challenges and limitations in implementing deep learning-based disease detection systems in the agricultural sector?

8. How can the developed deep learning model be integrated into existing smart farming practices to enable timely disease identification and intervention?

9. What are the potential impacts and benefits of the proposed deep learning-based disease detection system on agricultural practices, crop yields, and resource allocation?

10. How does the proposed system compare to traditional methods of disease identification and management in terms of accuracy, efficiency, and timeliness?

11. How can the findings and insights from this research contribute to the advancement of disease detection techniques in Solanaceae family vegetables and the broader field of agriculture?

These research questions encompass both the effectiveness and suitability of the deep learning model for disease detection, as well as broader aspects related to performance evaluation, dataset utilization, system integration, and the potential impacts of the proposed system.

1.5 RESEARCH OBJECTIVE

As the researcher, my aim is to address the research questions and accomplish the following objectives in my study:

1. Determine the effectiveness of deep learning models, specifically focusing on the most prevalent and harmful diseases, for disease detection in Solanaceae family vegetables within the context of smart farming.

2. Measure the accuracy of the deep learning model by implementing multi-class classification techniques to accurately classify various leaf diseases in Solanaceae family vegetables.

3. Explore and utilize deep learning models, such as VGG-16 and ResNet-50, to classify the collected leaves and accurately identify the specific diseases affecting Solanaceae family vegetables.

4. Develop and implement the deep learning model using the comprehensive dataset collected specifically for this research, ensuring the model's suitability and effectiveness in disease detection.

5. Assess the performance and accuracy of the deep learning models, considering the specific leaf diseases in the Solanaceae family, to determine their effectiveness in automated disease detection.

6. Investigate the potential of the developed deep learning model, trained with the researcher's collected data, for accurate disease classification in Solanaceae family vegetables.

7. Identify and address the challenges associated with implementing deep learning-based disease detection systems, specifically focusing on leaf diseases, in the agricultural sector.

8. Evaluate the performance and accuracy of the deep learning model, compared to traditional methods, in classifying and managing leaf diseases in Solanaceae family vegetables.

9. Provide recommendations and insights based on the research findings, contributing to the advancement of disease detection techniques in Solanaceae family vegetables and the broader field of agriculture.

By accomplishing these objectives, I aim to develop an accurate and efficient deep learning-based disease detection system that can effectively classify and manage leaf diseases in Solanaceae family vegetables. The implementation of the model using the researcher's collected data will enhance the accuracy and relevance of the results, providing valuable insights for improving agricultural practices and crop management.

1.6 RESEARCH SCOPE

The research scope of this study encompasses the following aspects:

1. Plant Focus: The research focuses specifically on disease detection in Solanaceae family vegetables, including commonly grown crops such as tomatoes, peppers, and eggplants. The goal is to enable farmers to efficiently identify and address plant diseases and their causes at any given time.

2. Disease Prevention and Increased Crop Yield: By implementing an accurate and efficient deep learning-based disease detection system, the study aims to contribute to disease prevention and mitigation in Solanaceae family vegetables. This will lead to increased crop yield and improved agricultural productivity.

3. Financial Benefits for Farmers: The successful implementation of the disease detection system will help farmers avoid significant losses caused by plant diseases. By enabling timely interventions, farmers can mitigate the impact of diseases and protect their crops. This, in turn, will have positive financial implications for farmers, leading to improved profitability and sustainability in agriculture.

4. Meeting Food Demands and Foreign Exchange: By effectively managing and preventing plant diseases in Solanaceae family vegetables, the research aims to contribute to meeting the food needs of the country's population. The increased crop yield and improved quality of vegetables will not only fulfill domestic requirements but also play a vital role in earning foreign exchange through agricultural exports.

5. Smart Farming and Technological Integration: The study is conducted within the context of smart farming, leveraging advanced technologies and data-driven approaches. By integrating deep learning techniques into smart farming practices, farmers will have access to a reliable and efficient tool for disease detection, enhancing their ability to make informed decisions and optimize crop management.

6. Practical Implications and Recommendations: The research findings will provide practical implications for farmers, agricultural practitioners, and policymakers,

emphasizing the importance of adopting deep learning-based disease detection systems. Recommendations will be provided on implementing and scaling up the proposed system, taking into account the specific context and requirements of the agricultural sector.

By addressing these aspects within the research scope, this study aims to empower farmers with the necessary tools and knowledge to efficiently identify and manage plant diseases in Solanaceae family vegetables. Ultimately, the research strives to contribute to improved agricultural practices, increased crop yields, financial benefits for farmers, and meeting the food demands of the country while playing a significant role in earning foreign exchange through agricultural exports.

1.7 RESEARCH GAPS

After reviewing a lot of papers, I found some issues that are still not resolved and need to be implemented. Such as,

- Everyone just utilized a small amount of data. More data can predict accurate results.
- Most of them used the Plant Village Kaggle dataset. The maximum researcher did not collect the data manually.
- The researchers didn't compare the ability of different deep learning architectures to classify many levels of disease.
- The researchers never determine which model classifies leaf diseases in a short period of time.
- No one applies the new deep learning models.

I have solved these issues in my research, which makes it unique.

1.8 SUMMARY

This study focuses on the use of deep learning methods in the context of smart farming to identify diseases in plants from the Solanaceae family. Plant leaf diseases are becoming more prevalent, which is of great worry to farmers as it threatens both national well-being and agricultural output. The issues faced by the agricultural industry are made even more difficult by the lack of sophisticated identification procedures. Traditional methods of disease detection and control sometimes prove ineffectual.

The goal of this study is to create a deep learning-based model for automated disease identification in vegetables belonging to the Solanaceae family. This research is driven by the need for an effective and practical solution. The goal of the project is to make it possible to quickly identify and treat diseases by utilizing the strength of deep learning algorithms, especially object detection algorithms. Through the use of shrewd farming techniques, the ultimate objective is to battle plant diseases, particularly those that impact trees and crops.

The research issue is a result of the scarcity of appropriate agricultural services and farmers' ignorance about plant diseases. Due to these illnesses' effects, a sizable portion of vegetables are discarded every year. The project focuses on using deep learning techniques, in particular convolutional neural networks (CNNs), for precise illness identification and classification to overcome this issue.

The efficiency of deep learning models in this situation is the central subject of the study. In particular, it investigates the suitability of CNNs for perfect leaf disease categorization, the efficacy of deep learning models for disease detection, and the accuracy of self-collected data for accurate measurement.

The research objectives encompass the development of an accurate and efficient system for automated disease detection, the utilization of comprehensive datasets and deep learning models for disease identification, and the evaluation of the proposed approach's effectiveness. The scope of the research includes a specific focus on Solanaceae family vegetables, multi-class classification of leaf diseases, and the implementation of the model using self-collected data.

The research seeks to advance agricultural methods by addressing these goals and carrying out a comprehensive examination. It aims to provide farmers with the information and resources they need to recognize and control plant diseases efficiently, reducing crop losses and maximizing resource allocation. The study's conclusions will offer useful information to the agricultural industry, highlighting the need for it to adopt technology breakthroughs like deep learning.

Overall, this study highlights the significance of deep learning-based disease detection in Solanaceae family vegetables, contributing to the advancement of smart farming practices and agricultural sustainability. By leveraging advanced technologies and data-driven approaches, the research seeks to enhance disease prevention, increase crop yields, provide financial benefits to farmers, meet the food demands of the population, and play a vital role in earning foreign exchange through agricultural exports.

CHAPTER: 02

LITERATURE REVIEW

A literature review is essential for my document as it provides an overview of existing research on my topic. It helps identify research gaps, build a theoretical framework, inform research design, and ensure originality. It justifies the significance of my study, enhances critical thinking, and guides decision-making in research methodology. Ultimately, the literature review ensures my study is well-informed and contributes to the existing knowledge in the field.

2.1 INTRODUCTION

The goal of this research is to present a comprehensive and cohesive thesis on "Smart Farming: Solanaceae Family Vegetable Plant Disease Detection using Deep Learning." It aims to showcase the significance and potential impact of using deep learning techniques for disease detection in Solanaceae family vegetables within the context of smart farming. The research will provide an in-depth understanding of the research objectives, motivation, background, methodology, and scope of the study. It will present the research findings, including the evaluation and performance of deep learning models, and highlight their effectiveness in automating disease detection. Furthermore, the document will offer practical implications, recommendations, and insights for improving agricultural practices, enhancing crop management, and addressing the challenges associated with implementing deep learning-based disease detection systems in the agricultural sector. Ultimately, the document seeks to contribute to the advancement of knowledge in the field of agricultural disease detection, empower farmers with accessible tools and knowledge, and promote sustainable farming practices

2.2 STUDY OF PLANT LEAF DISEASE DETECTION METHODS

Murk Chohan, Adil Khan, Rozina Chohan, Saif Hassan Katpar, and Muhammad Saleem Mahar proposed a deep learning-based model named "plant disease detector" that utilizes neural networks to detect diseases in plants based on leaf images. The model achieves a high testing accuracy of 98.3% using the PlantVillage dataset. The study suggests the integration of the model with drones or other systems for real-time disease detection and reporting.[1]

Faye Mohameth, Chen Bingcai, and Kane Amath Sada explored the concept of "smartphone-assisted disease diagnosis" by combining high-end smartphones and computer vision through deep learning techniques. The authors evaluate various convolutional neural network (CNN) architectures using transfer learning and deep feature extraction. The results demonstrate the effectiveness of Support Vector Machines (SVM) for leaf disease detection.[2]

Shima Ramesh, Niveditha M, Pooja R, Prasad Bhat N, Shashank N, Mr. Ramachandra Hebbar, and Mr. P V Vinod focused on the use of Random Forest for identifying healthy and diseased leaves through leaf-based image classification. The authors utilize a Histogram of Oriented Gradient (HOG) features and train the classifier on publicly available datasets. The approach shows promise in detecting diseases on a large scale.[3]

Ravindra Namdeorao Jogekar and Nandita Tiwari discussed different deep-learning approaches for the identification and diagnosis of various plant diseases on banana leaves. The authors emphasize the importance of early disease detection using deep convolutional neural networks (CNNs) to minimize economic losses for farmers.[4]

The author MH Saleem, J Potgieter, and KM Arif highlight the potential of deep learning models, particularly CNNs, in visualizing various plant diseases and improving accuracy in disease detection and classification. The paper identifies research gaps and emphasizes the need for early disease detection before symptoms become evident.[5]

K. Deeba, a   B.Amutha focused on predicting and classifying vegetable leaf diseases using deep learning algorithms. Various convolutional neural network architectures, such as LeNet, AlexNet, VGG16, VGG19, and ResNets, are employed. The system achieves an overall performance prediction of up to 98%.[6]

Mohit Agarwal, Abhishek Singh, Siddhartha Arjaria, Amit Sinha, and Suneet Gupta present a deep learning-based approach for tomato disease detection using a convolutional neural network. The proposed model outperforms pre-trained models like VGG16, InceptionV3, and MobileNet, achieving an average accuracy of 91.2% for different disease classes.[7]

2.3 SUMMARY OF LITERATURE STUDY

The literature review demonstrates the significant potential of deep learning techniques, particularly convolutional neural networks, in accurate and efficient plant disease detection. The studies provide insights into the development of models, datasets, feature extraction methods, and classification approaches for early disease identification in various crops. These findings contribute to the advancement of agricultural technology and practices, aiming to improve crop yields and enhance food security.

In this research, I aim to evaluate the effectiveness of deep learning models, such as VGG-16 and ResNet-50, for disease detection in Solanaceae family vegetables. By developing and implementing a comprehensive dataset, I will measure the accuracy of the deep learning model in classifying leaf diseases. The research will contribute to smart farming practices, enhance disease detection techniques, and improve agricultural productivity.

CHAPTER: 03

RESEARCH METHODOLOGY

The research methodology employed in this study follows a systematic and structured approach to investigate disease detection in Solanaceae family vegetables using deep learning techniques. The methodology consists of several key steps, including data collection, dataset preparation, model development, training and evaluation, and performance analysis.

3.1 DATA COLLECTION

A comprehensive data set of eggplant or brinjal and potato leaf images from Solanaceae family vegetables is collected. The dataset includes samples of healthy leaves and leaves affected by various diseases. The data collection process ensures an adequate representation of different disease types and severity levels.

Class	Plant Name	Healthy or Diseased	Disease Name	Image (Number)
C_1	Brinjal	Diseased	Flea Beetle	73
C_2	Brinjal	Diseased	Mosaic Virus	82
C_3	Brinjal	Healthy	-	109
C_4	Potato	Diseased	Early Blight	1000
C_5	Potato	Diseased	Late Blight	920
C_5	Potato	Healthy	-	152

Table 3.1: Dataset Description

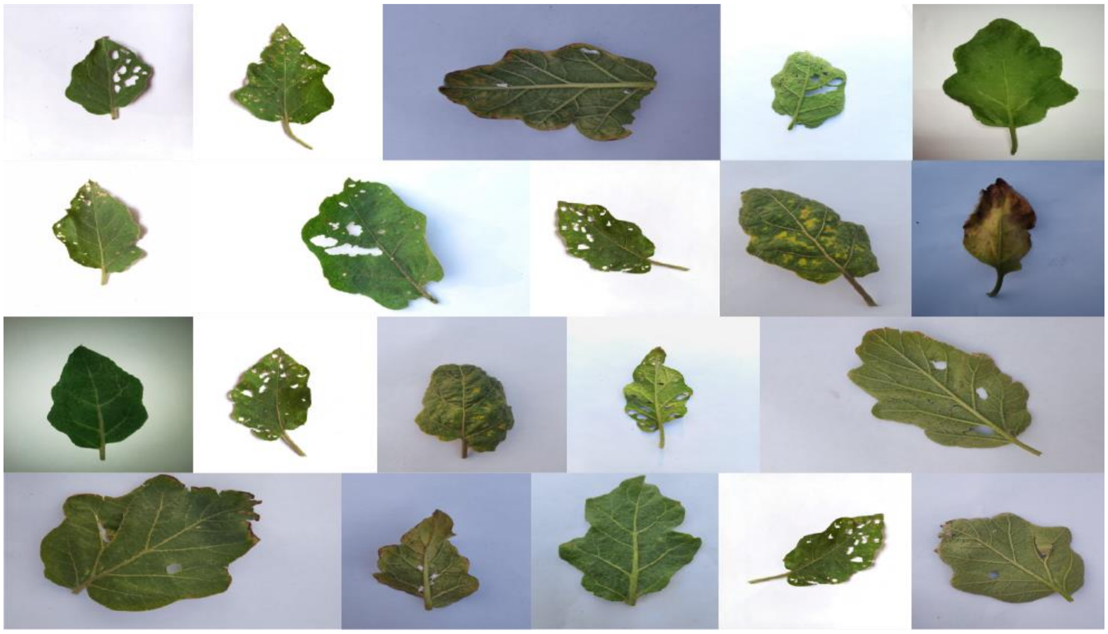


Figure 3.1: BRINJAL DATA

Figure 3.2: POTATO DATA

3.2 DATA PREPROCESSING

Image data preprocessing involves operations that enhance data quality, normalize values, reduce dimensionality, and improve model performance. It mitigates biases, enables data augmentation, and improves interpretability for better insights and accurate predictions. After collecting images of healthy and diseased leaves of brinjal and potato from gardens and fields.

3.2.1 RESIZE AND RESCALE

I preprocessed them. I preprocessed the image data by **Resizing and Rescaling** techniques which are commonly referred to as data preprocessing pipelines. When I rescale image pixel values, I divide them by 255. This is done to normalize the values to a range between 0 and 1. It helps ensure compatibility with neural networks, improves gradient flow, enhances stability during optimization, and allows for consistent comparisons across images. We resize every image to 256 by 256. Image resizing is essential for image preprocessing. Resizing images is important for consistent input dimensions, memory efficiency, reducing overfitting, improving model robustness, and compatibility with pre-trained models.

3.2.2 DATA AUGMENTATION

Data augmentation is used to increase data diversity, improve generalization, address class imbalance, reduce overfitting, and optimize resource utilization in machine learning. For better preprocessing of the image data, I used the data augmentation process. I used the **RandomFlip** and **RandomRotation** functions to augment the data.

3.3 VGG – 16

VGG-16, which stands for Visual Geometry Group-16, is a convolutional neural network (CNN) architecture known for its simplicity and effectiveness. Its deep structure allows it to learn complex hierarchical representations of plant leaf images, enabling the detection of intricate disease patterns. The multiple layers help capture both low-level and high-level features, which are crucial for accurate disease classification. VGG-16 is important for plant leaf disease detection due to its pre-training on ImageNet, enabling it to learn general visual representations. Transfer learning allows us to leverage this knowledge for improved performance on plant leaf disease datasets. VGG-16's broad applicability and availability of pre-trained models make it a reliable choice, saving time and resources while achieving accurate disease identification.

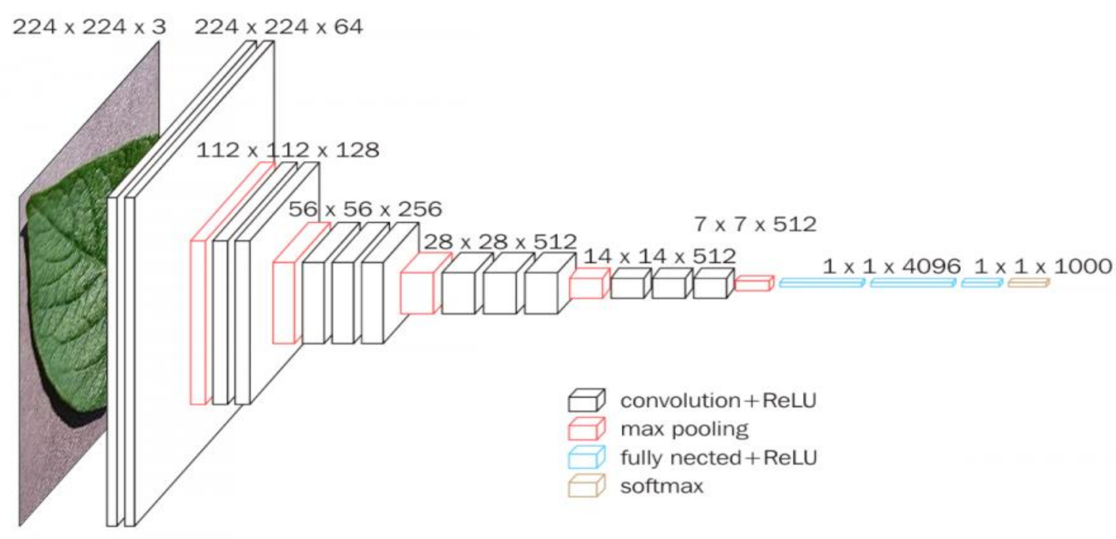


Figure 3.3: VGG-16 Architechture

We used the GPU for speedier calculation in this research project. The batch size was **32**. We preserved the VGG-16-acceptable default picture size of **224*224**. After determining the parameters and employing the **Softmax** activation function, we ran 30 training epochs.

3.4 RESNET - 50

ResNet-50, short for Residual Network-50, is a more advanced CNN architecture that introduced the concept of residual learning. With 50 layers, ResNet-50 employs residual blocks containing skip connections or shortcuts. These shortcuts enable the network to propagate information from earlier layers directly to later layers, mitigating the vanishing gradient problem. I opted for ResNet-50 due to its deep structure and residual connections, which are beneficial for detecting complex patterns and extracting intricate features from vegetable plant images.

In my plant leaf disease detection research, ResNet-50 is crucial due to its deep architecture with residual connections, strong performance and generalization, transfer learning capabilities, and availability of pre-trained models. It enables accurate disease identification and aids in plant health management.

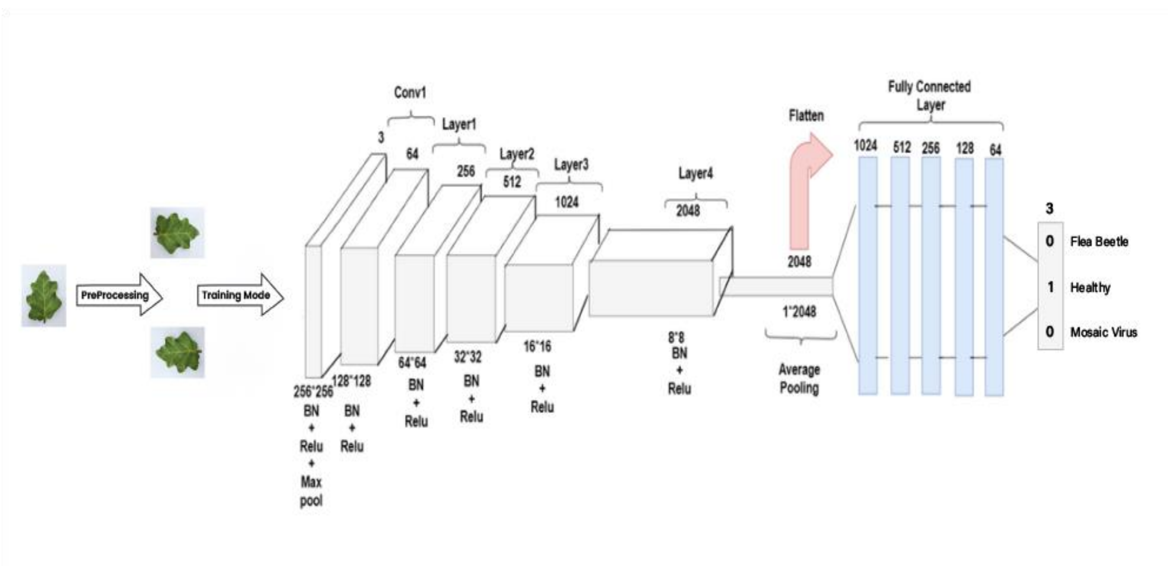


Figure 3.4: ResNet-50 Architechture

The batch size was 32. We preserved the ResNet – 50 acceptable default picture size of 256*256. After determining the parameters and employing the Relu and Softmax activation function, we ran 20 training epochs.

3.5 EVALUATION METHOD

The confusion matrix is crucial for leaf disease detection as it evaluates classification performance, identifies error patterns, analyzes class-specific metrics, aids threshold selection, and facilitates model comparison. It provides valuable insights for improving accuracy and making informed decisions in leaf disease management. The "Confusion Matrix" may be simply understood by all people and can be used to evaluate any type of model. For calculating that factor, we need to consider some of the parameters. Those are:

True-Positive (TP): “True Positive” refers to correctly predicted positive instances.

True-Negative (TP): “True Negative” represents correctly predicted negative instances.

False-Positive (FP): “False Positive” indicates incorrectly predicted positive instances.

False-Negative (FN): “False Negative” signifies incorrectly predicted negative instances.

3.5.1 ACCURACY

The accuracy of a machine's outcome prediction depends on how good the model is.

When each class is equally important, something important has occurred. Every class is crucial to our area of work. As a result, precision is essential in establishing if the model is right.

Accuracy is a performance metric used to measure the overall correctness of a classification model. It is calculated using the following formula:

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+TN+FN} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots.(3.1)$$

3.5.2 PRECISION

Precision is a performance metric that measures the proportion of correctly predicted positive instances (true positives) out of all instances predicted as positive (true positives and false positives). It is calculated using the formula:

$$\text{Precision} = \frac{TP}{TP+FP} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots..(3.2)$$

3.5.3 RECALL

Recall, also known as sensitivity or true positive rate, is a performance metric that measures the proportion of correctly predicted positive instances (true positives) out of all actual positive instances (true positives and false negatives). It is calculated using the formula:

$$\text{Recall} = \frac{TP}{TP+FN} \quad \text{.................................... (3.3)}$$

3.5.4 F1 SCORE

Recall, or sensitivity, measures the model's ability to capture all positive instances. Its importance lies in detecting critical instances, evaluating class imbalance, guiding the trade-off with precision, and enabling model comparison and selection.

$$\text{F1 Score} = 2 * \frac{Precision*Recall}{Precision+Recall} \quad \text{.......................(3.4)}$$

3.5.5 CONFUSION MATRIX

In essence, the "Confusion Matrix" is based on four indicators: "True-Positive (TP), True-Negative (TN), False-Positive (FP), and False-Negative (FN)".

And the following is a graphic depiction of it:

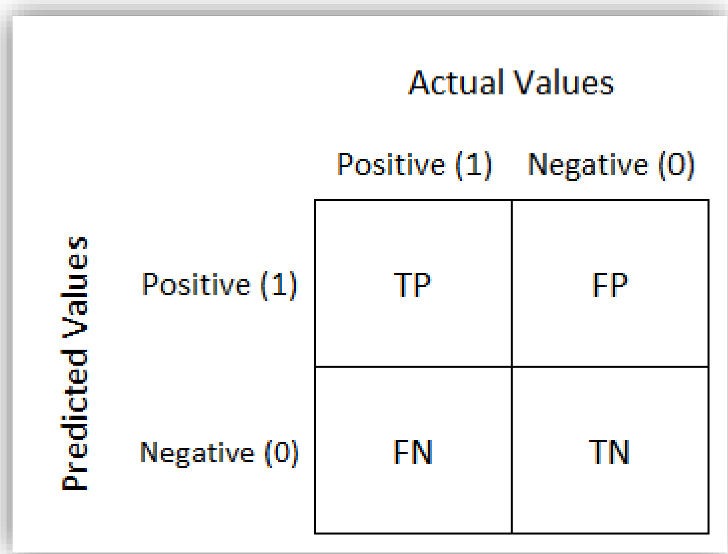


Figure 3.5: Confusion Matrix Visualization

CHAPTER: 04

RESULT AND DISCUSSION

We presented the model implementation process after the data collecting and preparation step. Here, we'll discuss the model's final output following training.

4.1 RESULT

After running 20 epochs using the ResNet-50 model I got an accuracy of 93.75% on brinjal data.

```
Epoch 20/20
5/5 [==============================] - 20s 4s/step - loss: 0.0106 - accuracy: 1.0000 - val_loss: 0.06
50 - val_accuracy: 0.9375
```

Figure 4.1: Accuracy of ResNet-50 For Brinjal Data

On the other hand after running 30 epochs using the VGG-16 model I got an accuracy of 92% on brinjal data.

```
Epoch 30/30
8/8 [==============================] - 85s 10s/step - loss: 0.0182 - accuracy: 1.0000 - val_loss: 0.2
573 - val_accuracy: 0.9200
```

Figure 4.2: Accuracy of VGG-16 For Brinjal Data

After running 30 epochs using the ResNet-50 model I got an accuracy of 99.52% on Potato data.

```
Epoch 30/30
40/40 [==============================] - 267s 7s/step - loss: 1.6359e-04 - accuracy: 1.0000 - val_los
s: 0.0057 - val_accuracy: 0.9952
```

Figure 4.3: Accuracy of ResNet-50 For Potato Data

After running 20 epochs using the VGG-16 model I got an accuracy of 97.4% on Potato data.

```
Epoch 20/20
54/54 [==============================] - 688s 13s/step - loss: 0.0195 - accuracy: 0.9959 - val_loss:
0.0721 - val_accuracy: 0.9744
```

Figure 4.4: Accuracy of VGG-16 For Potato Data

Indicators of the model's accuracy and model loss are also created in certain plotting graphs. The train and validation values of the model were taken into account while creating those charts.

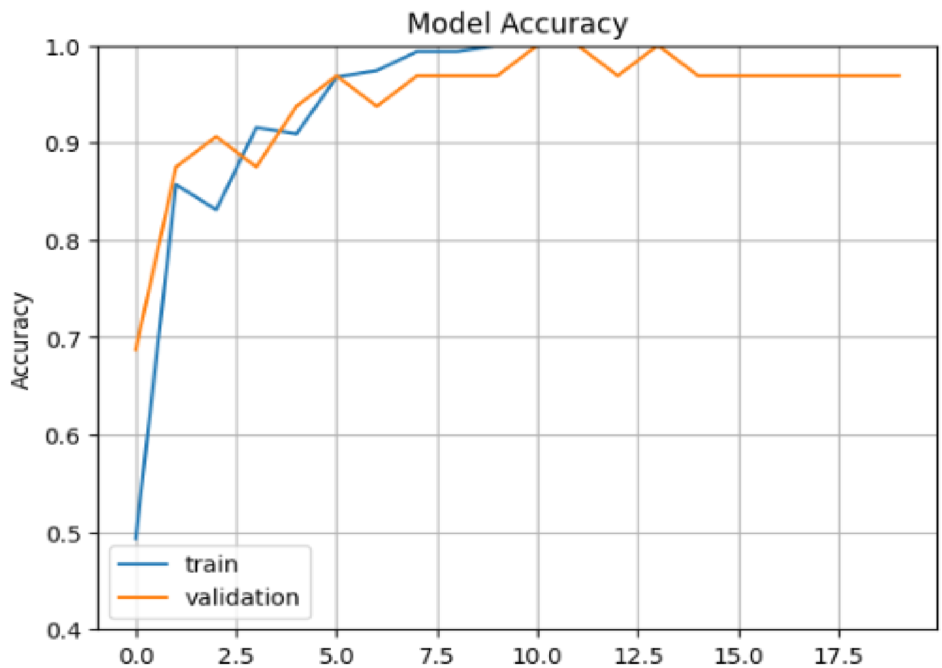


Figure 4.3: Model Accuracy (ResNet-50)

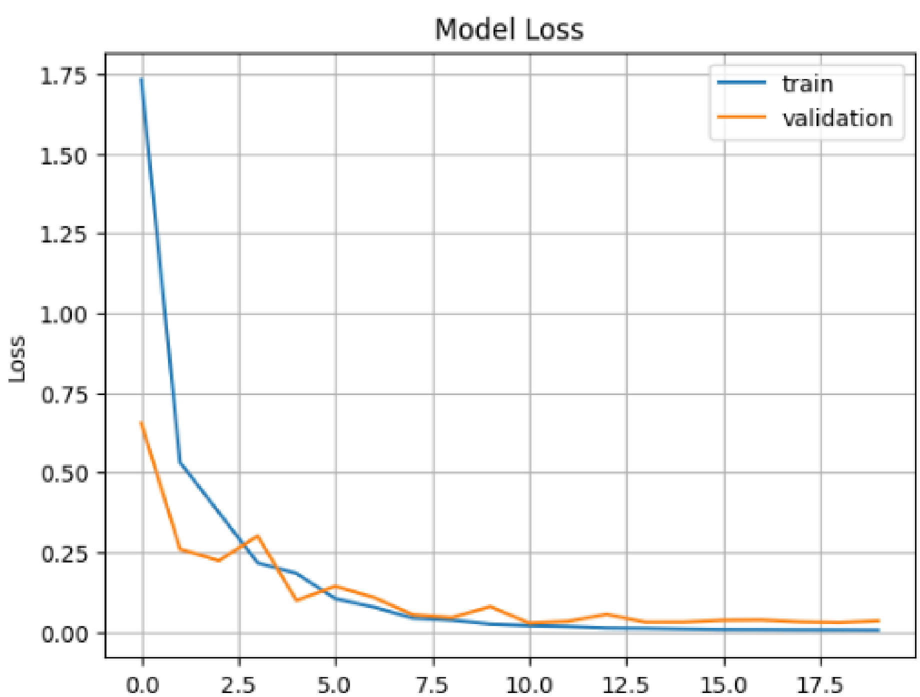
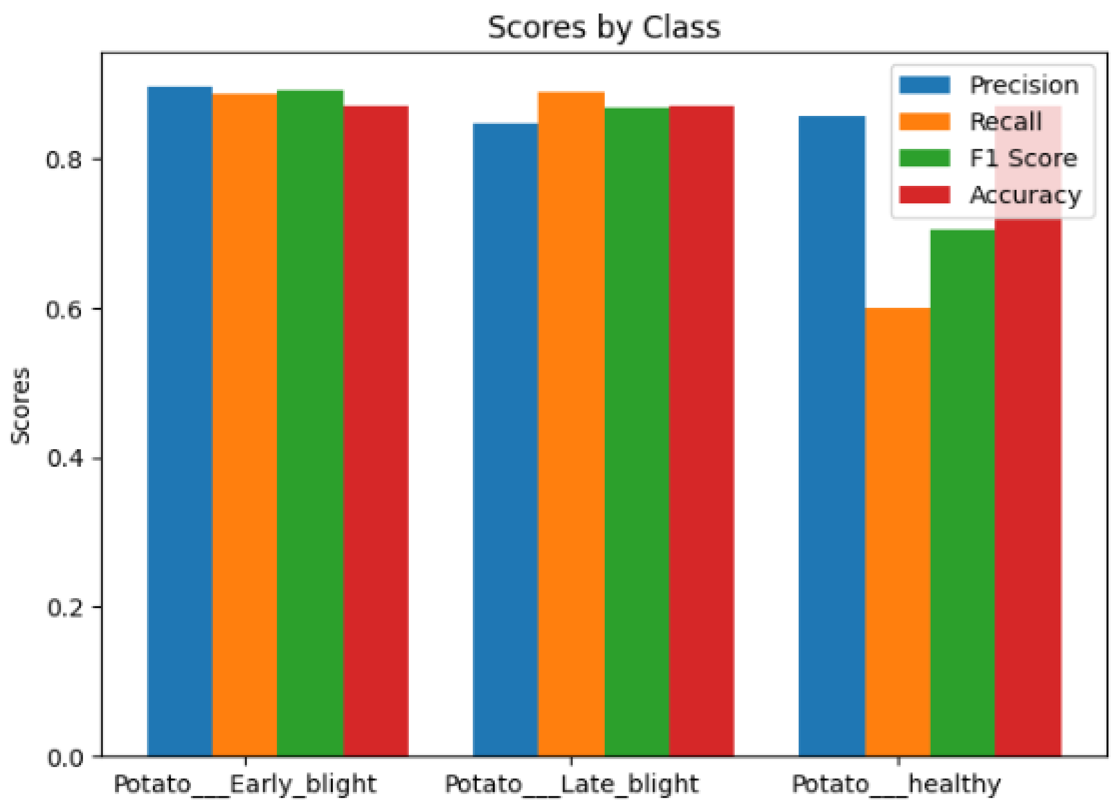


Figure 4.4: Model Loss (ResNet-50)




```
Mean Average Precision: 0.8672125260360555
Overall Accuracy: 0.8705882352941177
```

Figure 4.5: Precision, Recall, F1 Score, Accuracy score by class

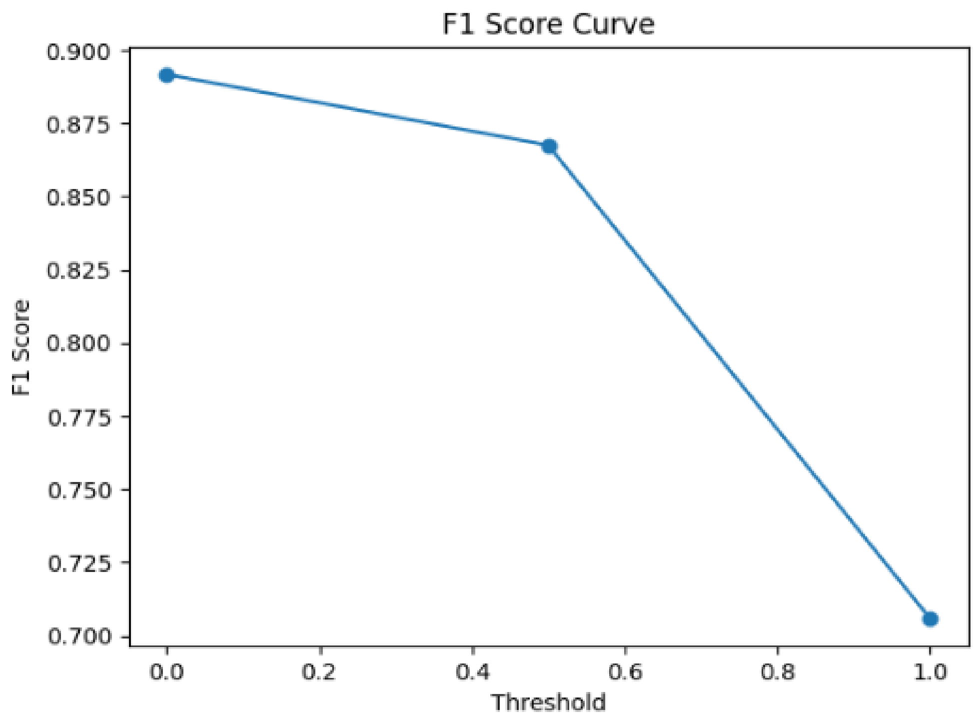


Figure 4.6: F1 Score Curve by Class

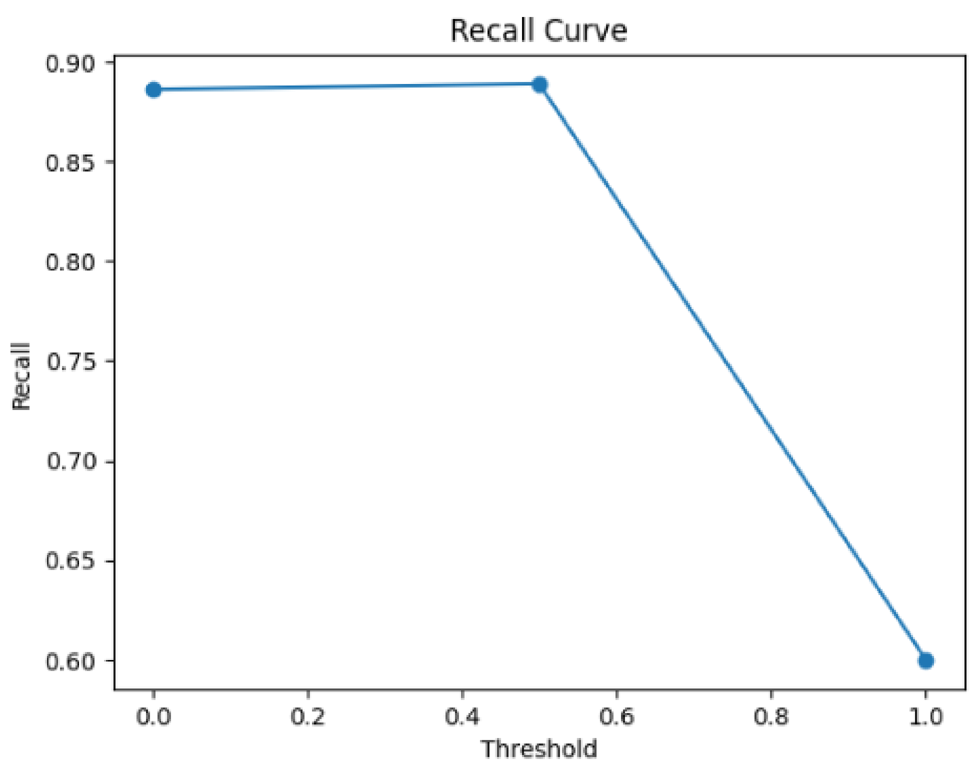


Figure 4.7: Recall Curve by Class

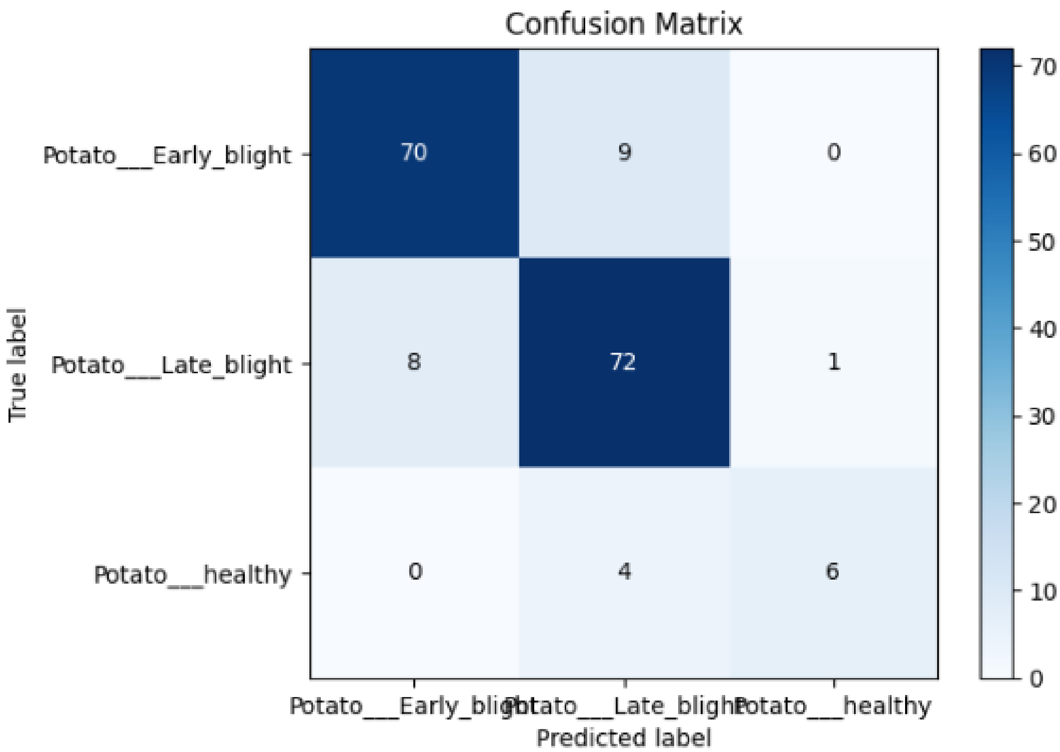


Figure 4.8: Confusion Matrix

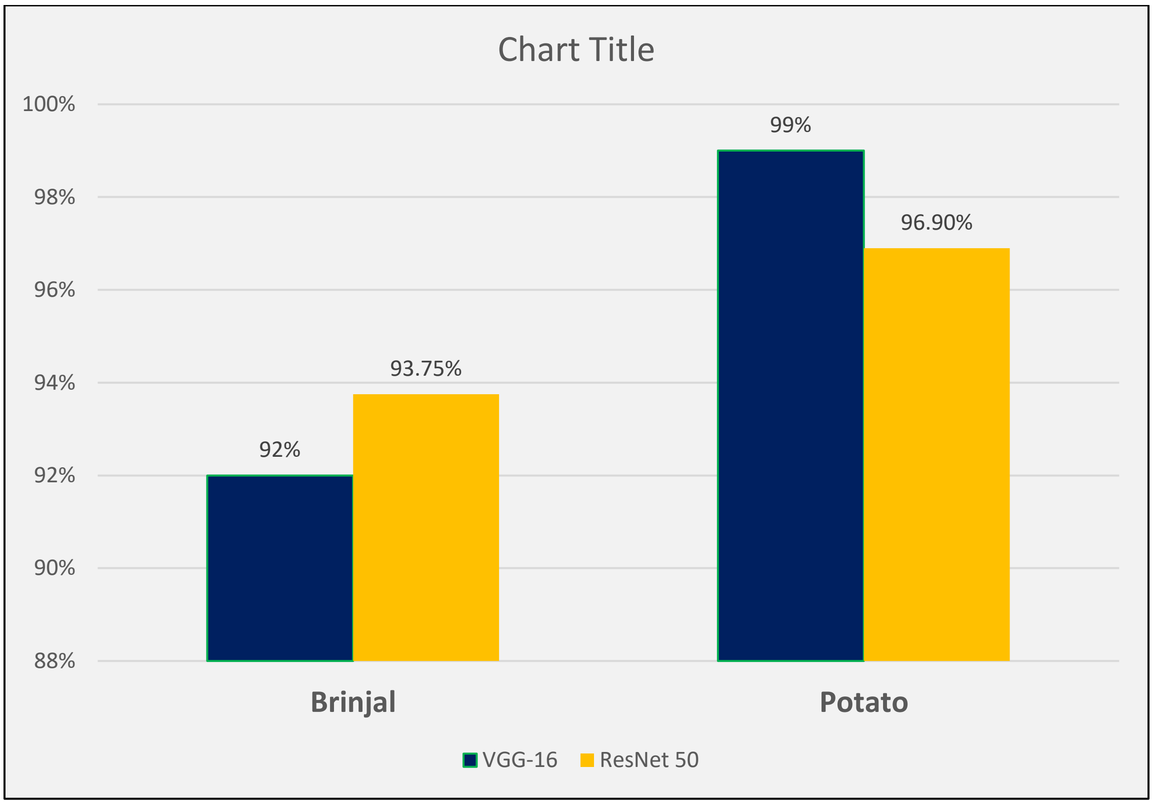


Figure 4.9: Model Comparison

4.2 MODEL OUTPUT

After training the ResNet-50 model, I gave some brinjal images as input. The predicting outputs are given below:

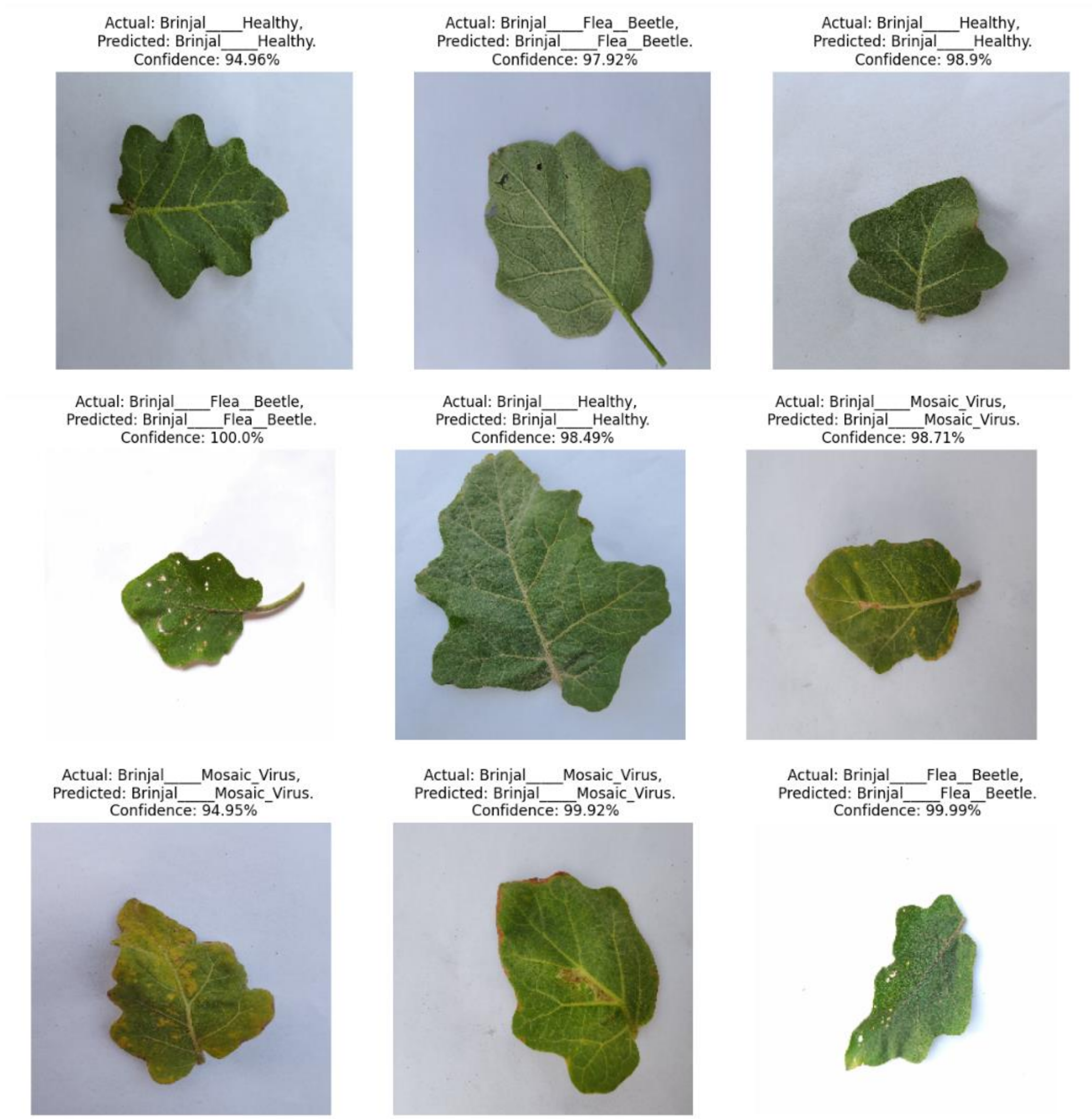


Figure 4.10: Accuracy of predicting output

4.3 DISCUSSION

We trained two models for this research. The accuracy of both VGG-16 and ResNet-50 models was excellent and satisfactory. Among them, the accuracy of ResNet-50 was higher. If more data could have been collected then the model accuracy and curve would have been better.

CHPATER: 05

CONCLUSION

In conclusion, this final document presented a study on "Smart Farming: Solanaceae Family Vegetable Plant Disease Detection Using Deep Learning." The research aimed to address the challenges of leaf diseases in Solanaceae family vegetables and develop an accurate system for automated disease detection. By utilizing deep learning models and a comprehensive dataset, the study demonstrated the effectiveness of the proposed approach in identifying diseases and improving agricultural practices. The research emphasized the importance of embracing technological advancements in agriculture and provided practical implications for farmers and policymakers. The findings contribute to sustainable farming practices and effective disease management.

CHAPTER: 06

REFERENCES

1. Murk Chohan, Adil Khan, Rozina Chohan, Saif Hassan Katpar, Muhammad Saleem Mahar. (May 2020)" **PLANT DISEASE DETECTION USING DEEP LEARNING**." *International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-9*.

2. Faye Mohameth, Chen Bingcai, and Kane Amath Sada" **PLANT DISEASE DETECTION WITH DEEP LEARNING AND FEATURE EXTRACTION USING PLANT VILLAGE**" *Journal of Computer and Communications - Vol.8 No.6, June 2020.*

3. Shima Ramesh, Niveditha M, Pooja R, Prasad Bhat N, Shashank N, Mr. Ramachandra Hebbar, Mr. P V Vinod" **PLANT DISEASE DETECTION USING MACHINE LEARNING.**" *2018 International Conference on Design Innovations for 3Cs Compute Communicate Control (ICDI3C).*

4. Ravindra Namdeorao Jogekar and Nandita Tiwari "**A REVIEW OF DEEP LEARNING TECHNIQUES FOR IDENTIFICATION AND DIAGNOSIS OF PLANT LEAF DISEASE.**" *Smart Innovation, Systems, and Technologies 182.Springer Singapore; Springer, Year: 2021. ISBN: 9789811552236,9789811552243*

5. MH Saleem, J Potgieter, KM Arif "**PLANT DISEASE DETECTION AND CLASSIFICATION BY DEEP LEARNING**" *Plants* **2019**, *8*(11), 468; **https://doi.org/10.3390/plants8110468**

6. K. Deeba, B. Amutha "**RESNET - DEEP NEURAL NETWORK ARCHITECTURE FOR LEAF DISEASE CLASSIFICATION.**" *Microprocessors and Microsystems (2020), https://doi.org/10.1016/j.micpro.2020.103364*.

7. Mohit Agarwal, Abhishek Singh, Siddhartha Arjaria, Amit Sinha, Suneet Gupta "**TOLED: TOMATO LEAF DISEASE DETECTION USING CONVOLUTION NEURAL NETWORK**" *International Conference on Computational Intelligence and Data Science (ICCIDS 2019).*

8. Sharada P. Mohanty David P. Hughes and Marcel Salathé "**PLANT DISEASE DETECTION AND CLASSIFICATION BY DEEP LEARNING**" Plants ***2019****,* 8*(11), 468;* ***https://doi.org/10.3390/plants8110468***

9. Fang, Y., & Ramasamy, R. P. (2015**). "CURRENT AND PROSPECTIVE METHODS FOR PLANT DISEASE DETECTION"**. *Biosensors, 5(3), 537-561.*

10. Sharada P. Mohanty David P. Hughes and Marcel Salathé **"USING DEEP LEARNING FOR IMAGE-BASED PLANT DISEASE DETECTION"** *Front. Plant Sci., 22 September 2016 Sec. Technical Advances in Plant Science Volume 7 - 2016 | https://doi.org/10.3389/fpls.2016.01419*

11. Konstantinos P. Ferentinos "**DEEP LEARNING MODELS FOR PLANT DISEASE DETECTION AND DAIGNOSIS**" *https://doi.org/10.1016/j.compag 2018.01.009*

12. Jayme Garcia Arnal Barbedo "**PLANT DISEASE IDENTIFICATION FROM INDIVIDUAL LESIONS AND SPOTS USING DEEP LEARNING**" https://doi.org/10.1016/j.biosystemseng 2019.02.002

13. Isleib, J., & Michigan State University. (2018, October 2). **SIGNS AND SYMPTOMS OF PLANT DISEASE: IS IF FUNGAL, VIRAL OR BACTERIAL? RETRIEVED FROM** https://www.canr.msu.edu/news/signs And_symptoms_of_plant_dise ase_is_it_fungal_viral_or_bacterial.

14. LeCun, Y., Bengio, Y., & Hinton, G. (2015). "**DEEP LEARNING. NATURE"**, *521(7553), 436-444.*

15. Ferentinos, K. P. (2018). **"DEEP LEARNING MODELS FOR PLANT DISEASE DETECTION AND DIAGNOSIS"**. *Computers and Electronics in Agriculture, 145, 311-318.*

16. Aakanksha Rastogi, Ritika Arora and Shanu Sharma," **LEAF DISEASE DETECTION AND GRADING USING COMPUTER VISION TECHNOLOGY &FUZZY LOGIC**" *2nd International Conference on Signal Processing and Integrated Networks (SPIN)2015.*

17. Naik, M.R., Sivappagari, C., "**PLANT LEAF AND DISEASE DETECTION BY USING HSV FEATURES AND SVM,**" *IJESC, Volume 6 Issue No.12, 2016*

18. J. Ma, K. Du, F. Zheng, L. Zhang, Z. Gong, and Z. Sun, "**A RECOGNITION METHOD FOR CUCUMBER DISEASES USING LEAF SYMPTOM IMAGES BASED ON DEEP CONVOLUTIONAL NEURAL NETWORK,**" *Comput. Electron. Agricult.,vol. 154, pp. 18–24, Nov. 2018. doi: 10.1016/j.compag.2018.08.048.*

19. Z. Iqbal et al., "**AN AUTOMATED DETECTION AND CLASSIFICATION OF CITRUS PLANT DISEASES USING IMAGE PROCESSING TECHNIQUES: A REVIEW,**" *Comput. Electron. Agricult., vol. 153, pp. 12–32, Oct. 2018. doi:10.1016/j.compag.2018.07.032.*

20. **Amritha Haridasan, Jeena Thomas & Ebin Deni Raj "DEEP LEARNING SYSTEM FOR PADDY PLANT DISEASE DETECTION AND CLASSIFICATION"** https://doi.org/10.1007/s10661-022-10656-x

PLAGARISM REPORT

Turnitin Originality Report

Processed on: 26-Jul-2023 09:11 +06
ID: 2136912939
Word Count: 6397
Submitted: 1

191-35-2682 By Ankon Chowdhury

Similarity Index	Similarity by Source	
16%	Internet Sources:	15%
	Publications:	5%
	Student Papers:	8%

4% match (Internet from 23-Jul-2023)
http://dspace.daffodilvarsity.edu.bd:8080/bitstream/handle/123456789/9588/22116.pdf?isAllowed=y&sequence=1

2% match (student papers from 05-Apr-2018)
Class: Article 2018
Assignment: Journal Article
Paper ID: 941549238

1% match (Internet from 23-Jul-2023)
http://dspace.daffodilvarsity.edu.bd:8080/bitstream/handle/123456789/9589/22117.pdf?isAllowed=y&sequence=1

1% match (Internet from 29-Jun-2023)
http://dspace.daffodilvarsity.edu.bd:8080/bitstream/handle/123456789/9616/22164.pdf?isAllowed=y&sequence=1

1% match (Internet from 22-Jul-2023)
http://dspace.daffodilvarsity.edu.bd:8080/bitstream/handle/123456789/9286/21614.pdf?isAllowed=y&sequence=1

1% match (Internet from 29-Jun-2023)
http://dspace.daffodilvarsity.edu.bd:8080/bitstream/handle/123456789/9628/22182.pdf?isAllowed=y&sequence=1

1% match (Internet from 22-Jul-2023)
https://www.mdpi.com/2079-9292/12/14/3103

1% match (Internet from 13-Nov-2022)
https://www.researchgate.net/publication/341025012_Plant_Disease_Detection_using_Deep_Learning

< 1% match (Internet from 21-Dec-2022)
https://www.mdpi.com/2077-0472/12/1/9/htm

< 1% match (Internet from 07-Apr-2023)
https://www.mdpi.com/1424-8220/21/14/4749

< 1% match (Internet from 15-Jan-2023)
https://www.mdpi.com/2077-0472/12/6/742/htm

< 1% match (Internet from 09-Feb-2023)
https://www.researchgate.net/publication/257558215_Recent_advances_in_sensing_plant_diseases

< 1% match (student papers from 29-Mar-2023)
Submitted to University of Bradford on 2023-03-29

< 1% match (Internet from 11-Jul-2023)
https://www.eurchembull.com/uploads/paper/320a1bf909596520c86a04aac146452a.pdf

ACCOUNTS CLEARANCE