

**MANGO VARIANT DETECTION BASED ON LEAF USING DEEP  
LEARNING ALGORITHMS**

**By**  
**MD. TAHMID**  
**ID: 221-15-5951**

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

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

Co-Supervised By  
**Mr. Shah Md Tanvir Siddiquee**  
Assistant Professor  
Department of CSE  
Daffodil International University



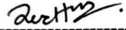
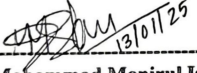


**DAFFODIL INTERNATIONAL UNIVERSITY**

**Dhaka, Bangladesh**  
**13 JANUARY 2025**

## APPROVAL

This Project titled "MANGO VARIANT DETECTION BASED ON LEAF USING DEEP LEARNING ALGORITHMS", submitted by MD. TAHMID, ID No: 221-15-5951 to the Department of Computer Science and Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on 13 January, 2025.

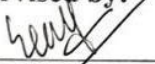
### BOARD OF EXAMINERS

 ----- <b>Dr. Md. Zahid Hasan</b> Associate Professor Department of Computer Science and Engineering Faculty of Science & Information Technology Daffodil International University	<b>Chairman</b>
 ----- <b>Mohammad Monirul Islam</b> Assistant Professor Department of Computer Science and Engineering Faculty of Science & Information Technology Daffodil International University	<b>Internal Examiner</b>
 ----- <b>Mr. Afjal Hossan Sarower</b> Senior Lecturer Department of Computer Science and Engineering Faculty of Science & Information Technology Daffodil International University	<b>Internal Examiner</b>
 ----- <b>Dr. Ahmed Wasif Reza</b> Professor Department of Computer Science and Engineering East West University	<b>External Examiner</b>

## 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:

  
\_\_\_\_\_

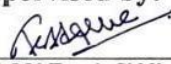
**Md. Sazzadur Ahamed**

Assistant Professor

Department of CSE

Daffodil International University

### Co-Supervised by:

  
\_\_\_\_\_

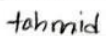
**Mr. Shah Md Tanvir Siddiquee**

Assistant Professor

Department of CSE

Daffodil International University

### Submitted by:

  
\_\_\_\_\_

**MD. TAHMID**

ID:- 221-15-5951

Department of CSE

Daffodil International University

## ACKNOWLEDGEMENT

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

We are really grateful and wish our profound indebtedness to **Md. Sazzadur Ahamed, Assistant Professor**, Department of CSE, Daffodil International University, Dhaka. Deep Knowledge & keen interest of our supervisor in the field of “*Deep Learning*” 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 stage have made it possible to complete this project.

We would like to express our heartiest gratitude to **Professor Dr. Sheak Rashed Haider Noori, Professor & Head**, 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.

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

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

## ABSTRACT

This study proposes an autonomous system for identifying and classifying mango trees using the DenseNet-201, MobileNet, ResNet50, ResNet-50 Transfer Learning model. Mango tree species must be accurately identified and classified in order to maximize agricultural methods, enhance resource management, and guarantee sustained mango production. Due to human subjectivity, traditional tree identification techniques that depend on the manual observation of leaf attributes are time-consuming, labor-intensive, and frequently inconsistent. This paper presents an automated system that can recognize and categorize different kinds of mango trees using photos of their leaves in order to overcome these difficulties. To improve model performance, a specially prepared dataset with 1,200 photos of mango leaves from seven different species was made. The ResNet-50 model demonstrated its efficacy in classifying mango tree species with an impressive accuracy score of 96.85% after being modified for this task. The study's findings demonstrate how computer vision techniques can be used to automate the categorization process, eliminating the need for human labor and facilitating accurate, large-scale mango tree identification.

**Keywords:** Deep learning, Convolutional neural network, ResNet-50, YOLOV8, DenseNet201, MobileNetv2.

## TABLE OF CONTENTS

<b>CONTENTS</b>	<b>PAGE</b>
Board of Examiners	ii
Declaration	iii
Acknowledgements	iv
Abstract	v
<b>CHAPTER</b>	
<b>CHAPTER 1: INTRODUCTION</b>	<b>1-8</b>
1.1 Introduction	1
1.2 Motivation	3
1.3 Rationale of the Study	4
1.4 Expected Output	6
1.5 Research Questions	6
1.5 Report Layout	7
<b>CHAPTER 2: BACKGROUND STUDY</b>	<b>9-13</b>
2.1 Terminologies	9
2.2 Related Works	9
2.3 Comparative Analysis and Summary	11
2.4 Scope of the Problem	12

2.5 Challenges	13
<b>CHAPTER 3: RESEARCH METHODOLOGY</b>	<b>14-23</b>
3.1 Introduction	14
3.2 Data Collection Procedure	15
3.3 Dataset Cleaning	16
3.4 Dataset Preprocessing	17
3.5 Proposed Methodology	18
3.6 Model Training	22
3.7 Implementation Requirements	23
<b>CHAPTER 4: RESULT ANALYSIS AND DISCUSSION</b>	<b>24-36</b>
4.1 Introduction	24
4.2 Experiment Results and Analysis	25
4.3 Generating Confusion Matrix	26
4.4 Generating Classification Report	32
4.5 Training and Validation Accuracy and Loss Curve	33
4.6 Discussion	36
<b>CHAPTER 5: IMPACT ON SOCIETY, ENVIRONMENT AND SUSTAINABILITY</b>	<b>37-40</b>

5.1 Impact on Society	37
5.2 Impact on Environment	38
5.3 Ethical Aspects	39
5.4 Sustainability Plan	39
<b>CHAPTER 6: OVERVIEW OF THE STUDY, CONCLUSION AND FUTURE WORK</b>	<b>40-42</b>
6.1 Overview of the Study	40
6.2 Conclusion	40
6.3 Limitations	41
6.4 Future Work	41
<b>REFERENCES</b>	<b>43-44</b>
<b>PLAGIARISM REPORT</b>	<b>45</b>

## **LIST OF FIGURES**

<b>FIGURES</b>	<b>PAGE NO</b>
Figure 3.1: Proposed System Architecture	15
Figure 3.2: Sample Dataset	16
Figure 3.5: ResNet-50 Model Architecture	20
Figure 4.3: Confusion Matrix for ResNet-50	27
Figure 4.3.2: Confusion Matrix for YOLOv8	29

## **LIST OF TABLES**

<b>TABLES</b>	<b>PAGE NO</b>
Table 3.3: The Final Dataset Table	17
Table 4.2: The Experiment Result of The Evaluated Model	25
Table 4.3: The Experiment Result Model YOLOv8	25

# CHAPTER 1

## Introduction

### 1.1 Introduction

Mango is the world's most popular fruit. Due to various reasons, mangoes are more profitable than rice or other crops, so orchard owners are now turning to mango cultivation. Mainly, mangoes are produced in thick places in Chittagong Hill tracts including Chapainawabganj, Rajshahi, Rangpur, Naogaon, Natore, Satkhira. Besides this, new areas are coming under mango production. With production, the mango market is also growing. About 14 thousand crore rupees is the value of the mango market in the country, which is increasing day by day. In Bangladesh in 2022-2023, 14 lakh 82 thousand 937 metric tons of mangoes were produced from 1 lakh 23 thousand 998 hectares of land and the average yield per hectare was 11.93 metric tons (BBS, 2023). In terms of total mango production, Bangladesh currently ranks seventh in the world. In the last 23 years, the area under mango cultivation in the country has increased by 2.45 times, production by 11.96 times and average yield per hectare by 3.17 times [1]. Bangladesh is the 7th largest mango-producing country in the world. In the fiscal year 2019-2020, Bangladesh had 17,686ha of mango orchards, yielded about 0.18 million metric tons of mangoes, and sold them for \$7.05 million. In addition, in the last fiscal year 2020-2021, 17,943ha of land were under mango cultivation and yielded about 0.217million metric tons of mangoes which were sold for \$8.525 million [2]. With so many varieties of mangoes, new mango growers and common people face challenges in identifying mango varieties by looking at mango seedlings. As a result, the desired mango variety failed to plant. Accurate Identification of tree species is crucial in various fields including agriculture, forestry and environmental conservation. Among fruit-bearing trees, mango holds significant

economic and cultural value throughout tropical and subtropical regions. However, differentiating between different mango tree species can be challenging because of the high degree of morphological similarity between species. Traditional methods of species identification based on visual observation or genetic analysis are often time-consuming, labor-intensive and require specialized knowledge. The aim of this research is to identify mango cultivars based on leaf. Marked with trees that have not yet borne fruit. This study will be done depending on the leaf. By converting the leaves into digital images, the tricking feature extension and classification process is completed to determine the type of mango tree. Leaf-base classification offers an alternative, year-round approach using unique characteristics of mango leaves such as shape, size, color, texture and vein pattern. Recent advances in image processing and machine learning, particularly the development of convolutional neural networks (CNN), have created new opportunities to automate and improve the classification process. CNNs are particularly suitable for image-based classification tasks; hence, they can capture features that are subtle and high dimensional in the leaf image for the precise distinction of similar cultivars.

These improvements notwithstanding, the challenges are still on in the accurate classification of mango tree cultivars since the morphological differences among the leaves of different cultivars are high, with environmental changes affecting leaf appearance. The objective of this thesis is to develop a CNN-based method to classify mango trees based on leaf characteristics. By examining optimal network architectures, data augmentation techniques, and preprocessing methods, this study seeks to provide a robust model that can accurately and efficiently classify mango trees from leaf images, contributing to advances in precision agriculture and automatic plant identification. Growing interest in creating automated plant-finding systems is a result of advancements in deep learning techniques [4].

## 1.2 Motivation

Mango, being treated as the "king of fruits," is highly significant, not only in the international fruit market but also to the Bangladeshi economy in general, where the cultivation of mango is considered one of the largest cash crops. Since its growth has increased so rapidly, proper identification of mango cultivars is becoming increasingly relevant. The incorrect identification of mango varieties at the seedling stage presents significant risks, especially for new orchard owners and farmers who may unintentionally plant the wrong type of mango. This can lead to long-term economic losses, as the fruit produced may not match market demand or the grower's specific requirements. It researches the solution to this difficult question: a systematic method to successfully determine the mango tree type before trees reach fruiting maturity. To be viable, any such solution must use either one of the traditional methods of identification: through appearance or genetic methods- neither of which has been possible for most farmers because it has been either difficult to apply or the results took so long, or both. Therefore, from leaf leaves identification for determination in mango cultivar, this present study tries applying certain modern image processing methods cost-effectively and hence, reliably adoptable on large and small scales by big-time farmers or small farmers. All such advances tend to have roles in the reduction of economic risk but also contribute to agriculture concerning sustainability and biodiversity, and generally efficiencies pertaining to the production of mangoes in the world. The study thus has huge potential to bring about a sea change in the very methodology of mango tree classification, benefiting not only the individual farmers but also helping to make the mango industry more orderly and productive, both within Bangladesh and worldwide. Deep learning techniques such as Convolutional Neural Network (CNN), DenseNet-201, MobileNet, ResNet50, and

YOLOv8 will be compared in this study. With the aim of accelerating the classification and detection and assisting farmers

### **1.3 Rationale of the Study**

A thorough examination and critique of various methods and techniques for classifying mango cultivars based on leaf characteristics may be the primary objective of this study. Accurately distinguishing mango species and cultivars is of immense value for agriculture, biodiversity conservation and commercial production, as mango (*Mangifera indica*) is one of the most widely cultivated tropical fruit crops worldwide. The strengths, limitations, and practical applications of existing techniques to achieve functional classification need to be explored, from traditional morphological evaluation to advanced image analysis aided by convolutional neural networks (CNNs) with other model such as DenseNet-50, MobileNet, ResNet50, and YOLOv8. The aim of this study is to evaluate the contemporary advances in deep learning and explore the utility of CNN-based models for differentiating mango cultivars using leaf images.

This research aims to transfer learning between innovative methods of mango tree classification such as DenseNet-50, MobileNet, ResNet50, and YOLOv8. It also serves in the improvement of the field since the best results were obtained using the ResNet-50 model, which achieved a great accuracy rate of 85%. Each of these methods offers unique advantages for extracting critical features from leaves, such as shape, size, color, texture, and vein pattern, yet comes with limitations in computational cost, data requirements, and adaptability to environmental factors. Focusing on specific leaf characteristics (such as venation pattern, color gradient and texture characteristics), this study investigates how these indicators correlate with mango species and cultivars.

In achieving these objectives, the research also considers practical issues within the agriculture and horticulture industry, including lack of information on rare varieties, changes in leaf characteristics due to environmental influences, and limited access to sophisticated imaging equipment in resource-limited settings. Addressing this challenge may enable the development of accessible and effective classification systems for widespread use. Therefore, identifying industry-specific needs such as robust field-testing techniques, reliable leaf image acquisition standards, and streamlined computational models will be essential to advance classification systems suitable for various agricultural applications.

The research also investigates how improved leaf classification techniques can strengthen precision agriculture by enabling species-specific cultivation practices to be better managed, thereby improving crop yields and preserving genetic diversity. By establishing effective taxonomic techniques, this research contributes to the creation or enhancement of taxonomic standards and regulatory guidelines, which can help develop industry standards for leaf-based plant identification.

In summary, the goals of a study on leaf-based mango tree classification using deep learning CNNs may include a comparison and validation of classification techniques, an analysis of existing methods, a critique of emerging technologies, a discussion of agricultural challenges, an exploration. Important leaf-based indicators, and contribution to increasing confidence in automated plant identification systems. Depending on specific research goals, industry needs, and anticipated impact, specific study objectives may include advancing regulatory frameworks for automated agricultural classification, ensuring accuracy and consistency in field practices, and supporting the expansion of species-specific agricultural standards. Ultimately, the results of this study can help formulate

policies, frameworks and industry guidelines, which can facilitate efficient, reliable and sustainable methods for mango diversity classification.

#### **1.4 Expected Output**

DenseNet-50, MobileNet, ResNet-50 and YOLOv8 models have been used for mango leaf based classification and identification, evaluating the performance in correctly identifying different mango tree species. Expected results include performance analysis including precision, accuracy, recall and F1-score. Visualizations such as confusion matrix, ROC curve and training-validation plot will also be used to analyze the performance of the models.

DenseNet-201 is expected to provide the highest accuracy due to its dense connectivity approach. ResNet-50 ensures reliable performance by using residual connectivity with an expected accuracy of around 95%. MobileNet is a lightweight and fast model, which is suitable for computational resources, although it provides relatively low accuracy. On the other hand, YOLOv8, despite being a detection model, is good at classification and fast at inference.

Visual outputs such as heatmaps of the confusion matrix and ROC curves for each class will be helpful in analyzing the performance of the model. This analysis shows that DenseNet-201 and ResNet-50 provide the highest accuracy, while MobileNet and YOLOv8 are fast and suitable for real-time applications.

#### **1.5 Research Questions**

- How to identify mango varieties using leaf pictures?
- What leaf characteristics distinguish mango varieties?
- Which machine learning algorithm best classifies mango leaves?
- What are the challenges in image based mango identification?
- How accurate is leaf-based detection versus conventional methods?

- Can small farmers easily adopt this identification method?
- How does the environment affect the appearance of mango leaves?
- How can this method help new growers to select the right mango variety?
- What are the economic benefits of leaf-based mango identification?
- How can this system be implemented in a mobile app for farmers?

## 1.5 Report Layout

The distinctive features of our endeavors are as follows:

**Chapter 1:** The research topic's historical and contextual information is presented in the introduction, along with the investigation's challenge or query and the study's goals and relevance. This section includes the introduction of the paper in 1.1, the inspiration for the subject of the study in 1.2, the justification for the study's conduct in 1.3, the anticipated results of this article in 1.4, and the summary or format of the document in 1.5.

**Chapter 2:** An initial appraisal that provides a brief synopsis of the research conducted on this topic is included in the background study. The applicable intelligence technology research is described here. Moreover, the challenges we faced while doing this study demonstrated the size of the issue. The terminologies subsection is used to describe the areas we will investigate for the paper in 2.1, the related works that show the scientist's prior work in 2.2, the comparative evaluation and summary of the topic in 2.3, the overall goal of the paper in 2.4, and the difficulties we will encounter in 2.5.

**Chapter 3:** The main concepts of data set handling and model generation have been comprehensively covered in this section. In 3.1, the research approach is introduced; in 3.2, how the dataset is assembled; in 3.3, how the dataset is

sterilized; in 3.4, the preliminary processing strategy of the dataset; and in 3.5 and 3.6, respectively, the recommended approach and the implementational prerequisites.

**Chapter 4:** This section assessed and looked into the output of our predictive framework. For ease of understanding, it incorporates all the results from the graphical description. This section comprises the evaluation of the paper and the experimental findings. The introductory part in 4.1, the result investigation in 4.2, the confusion matrix and classification report resemblance of the outcomes in 4.3 and 4.4, the precision of the validation and training in 4.5, and the discussion of the results segment in 4.6.

**Chapter 5:** The repercussions of marine life freshness on the community, the surroundings, as well as sustainability are briefly addressed in 5.1, 5.2, 5.3, and 5.4 accordingly.

**Chapter 6:** In accordance with 6.1, 6.2, and 6.3, an overview of the accomplished study, a conclusion, and potential future research are shown in this section.

## **CHAPTER 2**

### **Background Study**

#### **2.1 Terminologies**

The tree species is identified using images of mango leaves. In this process, leaves of different species of mango trees such as Amrupali, Himsagor, Haribhanga, Barifor have been used. There is no Research paper on identifying mango tree species based on leaves using deep learning. In this research classification of trees from images is combined with deep learning methods. Background information about the study the information provided in the article gives them the right context. The background of the study there therefore arouses the curiosity of the viewer about the problem of the study result and explains why it is important.

#### **2.2 Related Work**

Singh et al. [5] In this paper developed a multilayer convolutional neural network (CNN) for the classification of mango leaves affected by anthracnose fungal disease. The dataset contains 1070 images of both healthy and infected leaves. It achieves 97.13% accuracy using MCNNt. Eko prasetyo. [6] In this research, depending on the leaves of mango trees, tree species have been done through image classification. The result showed that our image processing method was effective to detect the variation of up to 78%. And the K-NN classification method has been implemented with accuracy with a value of 90%.

Yasushi Minowa. [7] In this study, tree leaves were used for tree species identification. CNN has been applied to it. These methods were challenged by variations in leaf size and vein pattern. This study uses only leaf size to identify species while classifying booths. Capturing the detailed characteristics of veins

requires high quality images which are not always possible with mobile devices. The study collected 10,000 images from five tree species. The study confirmed the effectiveness of CNNs for automatic tree identification. CNN can classify with high accuracy using partial leaf images. [9] Deep learning has transformed plant species and disease recognition. Very high accuracy has been achieved with CNNs like ResNet50 and VGG19 using transfer learning. Object detection models like YOLOv2 and YOLOv3 have been able to classify and locate leaves and diseases in real time. However, there are still challenges regarding dataset diversity and computational efficiency that inspire further development of lightweight and adaptive models. Elqassas et al [14]. Studies have featured an extensive future scope for the system in terms of management and treatment of various mango diseases, thereby achieving broad productivity in agriculture.

Safari et al. [15] The outcome of the study demonstrated the feasibility of employing YOLOv8 deep learning models for the automatic detection and quantification of mango damage severity during the early fruiting stage. The YOLOv8l model was able to achieve precision scores of 98.5 and 88.6 for detecting mango instances and damage instances, respectively.

Ahmed et al. [16] It utilized Transfer Learning techniques to boost its ability in image classification tasks, focusing on rice disease detection. Training and evaluation were conducted using a dataset divided into train, test, and validation samples to provide sufficient analysis of capabilities.

Husnain et al. [17] In the case of this research, a collection of high-quality multispectral images has been made available to the author, comprising 8905 images taken from a DJI Air 2s drone. The imaging spectrum is made up of three channels: Infrared, Grayscale, and RGB, which measures the small health differences within the trees. The drone takes the images at an altitude of 100 feet

and maintained 60% overlap in images for successful stitching. Phadikar et al. [18] The research paper focuses on the classification and detection of rice leaf diseases using the YOLOv5 deep learning algorithm, which has shown superior performance compared to other methods.

### **2.3 Comparative Analysis and Summary**

The main objective of this study is to identify mango varieties using deep learning methods based on mango leaf features. Accurate identification of mango varieties is very important for agriculture, biodiversity conservation and mango production systems. Misidentification of mango varieties at the seedling stage can cause economic losses to farmers. This study was motivated to develop a cost-effective and reliable method to address this challenge. A total of 7 species of mango leaves were collected from different times, angles and environments to diversify the dataset. Additional labeled data was also collected from kaggle, which helped in enhancing the performance of the model. ResNet-50, YOLOv8, DenseNet201, MobileNetv2 architecture was used, which is particularly effective for complex image classification. Data augmentation and preprocessing methods were applied to improve the accuracy of the model. High-quality images are required to capture fine features such as leaf vein patterns, which are challenging to identify leaf diseases, with very little research on mango cultivar classification. This research aims to fill that gap. Recently, several plant species have been facing the threat of extinction. Environmental scientists want to conserve these plants and maintain floral diversity. For this reason, the topic of plant identification using computer vision and pattern recognition has become interesting to researchers. An effective tool that can identify plant species using only an image would be a major advance in the field of plant identification and classification [8]

### **2.4 Scope of the Problem**

The scope of the problem includes the challenges faced in correctly identifying mango tree varieties based on their leaf characteristics, especially at the seedling stage before fruiting. This problem has significant implications for farmers, orchard owners and the agricultural industry at large. Mango production in Bangladesh is an important economic activity, closely linked to the agricultural sector. However, there are some major challenges in correctly identifying mango tree varieties, especially at the seedling stage. These problems not only cause financial losses to new mango orchard owners, but also affect the mango production and marketing system in the long term. Identifying the correct variety of mango trees by looking at the leaves of seedlings or trees is often difficult. The high degree of similarity in leaf structure and environmental diversity complicates the process of variety identification. Conventional methods such as visual observation or genetic analysis are time-consuming, labor-intensive and require specialized knowledge. These methods are not accessible to small and new farmers. Leaf shape, color, and structure may vary in different environments that could affect the accurate variety identification. There is not enough information on less common or rare varieties of mangoes. This makes it difficult to preserve or use these varieties properly. Machine learning or deep learning based methods require high-quality images of leaves of different species, which are difficult to collect. There is a lack of creating a cost-effective and accessible system using the invented methods and technology, which can meet the needs of small farmers. Failure to identify the variety correctly can lead to financial loss to farmers and consumer dissatisfaction due to the supply of wrong variety of mangoes in the market. The practical application of the invented automatic methods requires sufficient expertise and quality technical infrastructure. The present research will, therefore, try to invent a leaf-based automatic method for the variety identification of mango trees, which will be easily available, cost-effective, and accurate for

farmers. The results from the research will contribute significantly to making the mango production system more organized and sustainable.

## **2.5 Challenges**

The challenges of using deep learning techniques to identify different varieties using mango leaves are illustrated. Many mango varieties have morphological similarities in their leaves, making identification difficult based on visual features alone. Environmental factors such as soil quality, water supply, and climate can cause variations in leaf characteristics, which affect the consistency of the dataset and the accuracy of the model. High-quality images are required to capture detailed features such as leaf veins and texture. Maintaining image quality under different lighting and background conditions is difficult, especially when data is collected in the field. When using small or uneven datasets, the model learns specific features of the training data, which creates problems in generalizing to new data. Building a robust pipeline for leaf image preprocessing, enhancement, and model training is complex and requires the right combination of hyperparameters, architecture, and training techniques. Making the system user-friendly, accessible, and affordable for smallholder farmers is a challenge. Lack of familiarity with the use of advanced technologies can be a barrier to adoption. Validating the accuracy of the system in real-world conditions and developing standardized protocols for data collection and use is time-consuming and difficult. It is important to develop a sustainable framework for updating the model as new mango varieties or changing environmental conditions emerge. Addressing this challenge, the objective of this research is to develop a reliable, accessible and effective mango variety identification system, which will help in the development of precise cultivation practices and the wider farming community.

## **CHAPTER 3**

### **Research Methodology**

#### **3.1 Introduction**

The research methodology of this thesis focuses on the development of a robust and efficient method for identifying mango tree species using leaf features, which are crucial for accurate identification at the seedling stage. The main objective of this research is to develop an automated, accurate, and scalable system that can identify different mango species, which is essential for proper cultivation and prevention of economic losses. This study uses a combination of image processing techniques and deep learning algorithms. The research methodology is divided into several main stages, such as data collection, preprocessing, model development, and evaluation. The main objective is to demonstrate how deep learning models such as ResNet-50, YOLOv8, MobileNetV2, and DenseNet201 can be used to identify species from mango leaf images, which is helpful in overcoming the limitations of traditional methods. This research methodology provides a comprehensive framework for mango species identification, which overcomes the limitations of traditional methods and increases accuracy and efficiency. Deep learning models such as ResNet-50, YOLOv8, MobileNetV2 and DenseNet201 have played a leading role in advancing research in species identification and are creating the potential for widely adopted and effective systems in agriculture.

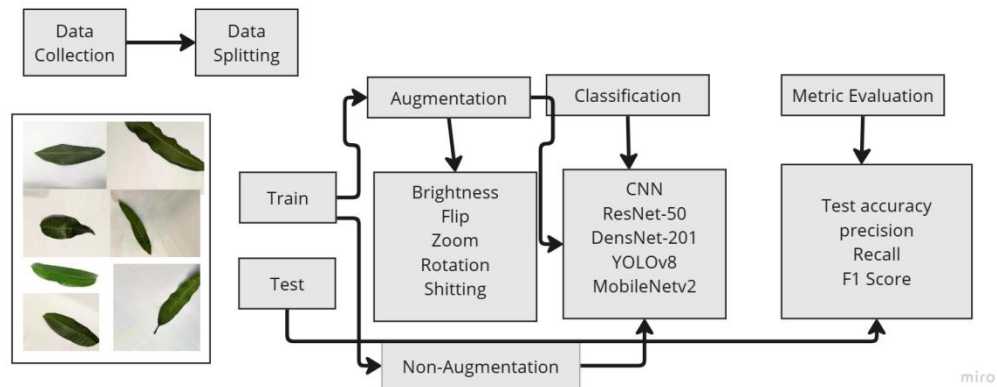


Figure 3.1: Proposed System Architecture

### 3.2 Data Collection Procedure

Two main methods were followed in collecting leaf-based data for mango tree identification. I took pictures of mango leaves using a mobile phone. The pictures were taken at different times to collect qualitative and different conditions of images to capture the effect of changing lighting. The pictures were taken from different angles of the leaves, such as: top, side, and close-up shots, so that the dataset contains images from different perspectives. The pictures were taken in different environments, so that the background diversity is retained and the identification model is more robust. In addition to the pictures I took, a dataset related to mango leaves was collected from kaggle. It includes leaves of different sizes, shapes, and conditions, which makes the dataset more diverse. The data from kaggle provided additional labeled images for training and validation of the model. Through these two methods, a diverse and rich dataset was collected, which will be helpful in building a mango tree identification model. I took a total of 1200 images of mango leaves of 7 different species using a mobile phone. The images were taken at different times and in different environments to ensure the

diversity of each species. The images were collected from different angles to ensure the diversity of the leaves and the background. In addition to the images I took, more data was collected from kaggle, which further enriched the dataset. Sufficient images were collected for each of the 7 species, which will be helpful in analyzing the data while building the model.



Figure 3.2: Sample Dataset

### 3.3 Dataset Cleaning

A system has been developed using deep learning methods to classify tree species using features of mango tree branches or leaves, which is used in this thesis. Collecting mango leaves of different species. Collecting additional data from open datasets like Kaggle. Cleaning and preparing the data so that it is more suitable for model training. Resize all images to a standard size (e.g., 512x512 pixels). Increasing the size of the dataset using techniques like rotation, flipping, and zooming. Selecting appropriate deep learning models. Training the selected models so that they can learn features from mango leaves. Comparing different

models (ResNet-50, YOLOv8, MobileNetV2, DenseNet201) and selecting the best model.

TABLE 3.3: THE FINAL DATASET TABLE

<b>Name of Mango</b>	<b>Number of Original Images</b>	<b>Number of Augmented Images</b>
Fazli Mango	200	1200
Haribhanga Mango	200	1200
Himsagar Mango	200	1200
Amrupali Mango	200	1200
Banana Mango	200	1200
Barifor Mango	200	1200
Chaunsa Mango	200	1200

### **3.4 Dataset Preprocessing**

Pre-processing of the dataset becomes the backbone of a machine learning or deep learning work. Therefore, it makes the data clean, structured, and ready for model training purposes. In the present study, dataset preprocessing takes care of the right quality and usability of the data needed to train and evaluate the model. The dataset contained images and relevant metadata, which were then processed for

improved performance through cleaning, normalization, and augmentation for the model. First, the dataset was inspected for anomalies like duplicate images, missing labels, and class mismatches. Also, all images were resized to 512 by 512 so that they become compatible with the input format of the MobileNetV2 and YOLOv8.

### **3.5 Proposed Methodology**

A modern approach to tackle new challenges. Transfer learning is a powerful deep learning approach that uses previously trained models to solve new and unique challenges. While it is not a specific deep learning algorithm, it is a strategic learning framework used to optimize performance during model training. By reusing pre-trained models, transfer learning can efficiently complete new tasks by reducing the additional resources and time required for training. This approach is particularly effective when the new task bears some similarity to the previous problem. For example, a model that was trained to classify objects on a general dataset can be fine-tuned to classify items on a more specialized or domain specific dataset. Adapting pre-trained models often requires large-scale changes or predictions to keep up with new and unfamiliar input data and ensure accurate results. Transfer learning methods have been widely used to create a variety of advanced deep neural network architectures. These networks exhibit remarkable accuracy while using relatively few computational resources, as they exploit optimized weights and features learned during their initial training. Such efficiency and adaptability make transfer learning an essential tool for solving complex and resource intensive tasks.

#### **3.5.1 ResNet-50:**

The ResNet-50 methodology has proven to be a highly dependable approach in image analysis, particularly when it comes to classifying that affect Mango leaves. The success of this system can be attributed to its exceptional capacity to handle

the intricacy of visual data in leaf image. ResNet-50 demonstrates its excellent capabilities in classifying mango plants by effectively utilizing a substantial and representative dataset [3]. The basis of ResNet-50 is the concept of residual learning. The idea is to let the network learn a "residual mapping" instead of the original mapping. By teaching each layer small, step-by-step transformations rather than the entire function, it reduces the risk of overfitting and enables deep network functional training. It uses "skip connections", where the input of one layer is added directly to the output of the next layer. These connections create shortcuts, so that the gradient can easily flow backwards and cannot be shortened. ResNet-50 has 50 layers, much deeper than previous networks. The architecture consists of convolutional layers, batch normalization, reLU activation, and max pooling and average pooling layers. This architecture can be divided into five phases, each phase having multiple remaining blocks. Each remaining block consists of three convolutional layers and an identity mapping layer, which creates skip connections. The ResNet family, particularly ResNet-50, has become fundamental to many computer vision tasks, such as object detection and image segmentation. ResNet-50 is one of the most popular architectures for transfer learning. Researchers and practitioners often use pre-trained versions of ResNet-50 for different tasks, fine-tuning it to new datasets to achieve superior performance. ResNet-50 is a robust and manageable architecture that provides high accuracy, stability and efficiency for image classification tasks, which is very beneficial for applications such as mango tree leaf classification.

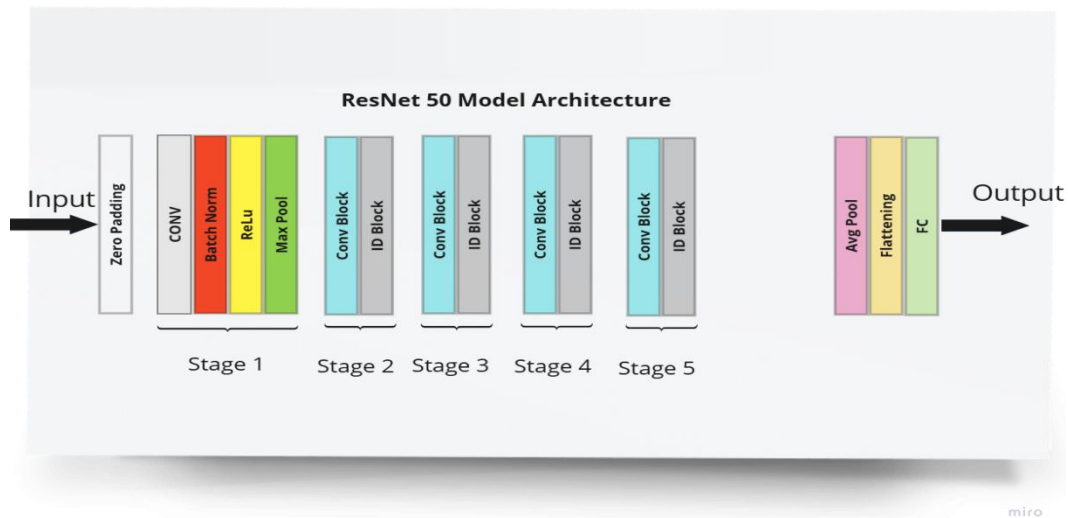


Figure 3.5: ResNet-50 Model Achitecture

### 3.5.2 YOLOv8

YOLO is an object detection algorithm that has gained popularity in computer vision. YOLO is a real-time object detection algorithm that processes an image in a single forward pass through a neural network. Unlike traditional object detection algorithms that involve multiple processing stages, YOLO performs object recognition and bounding box regression in a single pass [10]. To determine the species of mango tree by looking at the picture of mango tree leaves . It can be applied to real-time detection in embedded devices. YOLOv5 has a fast detection time, good detection accuracy, and a lightweight design in contrast to other YOLO versions [11]. Anchor boxes, which are predefined boxes with varying sizes and aspect ratios, are a strategy used by YOLO to increase prediction accuracy. A specific cell is linked to each anchor box, which is used to forecast the dimensions and form of the object inside that cell. Anchor boxes allow YOLO to work with objects of various forms and sizes. YOLO's quickness is one of its key advantages. Because YOLO can process photos in real-time, it can be used in detection tree using leaf, robotics<sup>26</sup>, surveillance systems, and driverless cars.

### **3.5.3 Mobile Net v2**

As a lightweight convolutional neural network architecture, MobileNetV2 is specifically made to function in contexts with limited resources, such as embedded systems and mobile devices. It has been frequently used for computer vision and image recognition jobs. Depthwise separable convolution is one of this network's primary characteristics[12]. The core concept of MobileNetV2 is based on the inverse residual block. It goes from a low-dimensional feature space to a high-dimensional space, so that the model can learn features more effectively. This model uses depth-based and pointwise convolution, which reduces the computational complexity of the model. This increases processing speed and reduces memory usage. The architecture of MobileNetV2 is very lightweight, which allows the deep learning model to be easily integrated into mobile and cloud platforms. MobileNetV2 is computationally inexpensive, yet it provides a high level of accuracy. In this study, the MobileNetV2 model was used to extract features from mango tree leaves and classify them. The input image for the model was resized to a suitable size (512x512 pixels). Diversity was added through data augmentation. The MobileNetV2 model provides fast and efficient classification. It has been able to ensure high accuracy in a short time. It can be easily deployed in mobile applications or web-based platforms.

### **3.5.4 DenseNet201**

Relative connections have resulted in the training of networks that are much deeper, as well as in more variations in the aesthetics of those networks. It is a Deep Convolutional Network model which belongs to DenseNet family-DenseNet201. The model was introduced in their paper titled "Densely Connected Convolutional Networks" presented at CVPR 2017 by Gao Huang, Zhuang Liu, Laurens van der Maaten, and Kilian Q. Weinberger. Coming out as an

improvement of DenseNet121 and DenseNet169, DenseNet201 adds even greater depth and more features to refine its capabilities with respect to image recognition tasks. The core structure of DenseNet201 is made from thick blocks, in which every layer in a block gets its feature maps from all previous layers as input. Feature reuse is promoted and lessening of the problem of fading gradients occurs through this closely connected direct information transfer between layers. DenseNet201 reduces the number of maps filled with features and defines the spatial dimensions by that dense block. The combinations of these transitional blocks include a batch equalization layer, a compression 1x1 convolution layer, and an average pooling 2x2 layer. Following a phase of 3x3 convolution in the bottleneck layer structure, one finds a 1x1 convolution layer [13].

Conversion to human like from AI like text. Rewrite text with lower perplexity and higher burstiness while keeping word count and HTML elements intact: You are being trained on data as up to the month of October in year 2023.

### **3.6 Model Training**

In this study, ResNet-50, YOLOv8, MobileNetV2, and DenseNet201 models are trained and optimized for mango leaf classification. A dataset of mango leaf images is created (total 1200 images, divided into 7 classes). Rotation, flipping, zooming, and brightness changes are applied to increase the diversity of the data. Pre-trained weights are used on the ImageNet dataset for ResNet-50, MobileNetV2, and DenseNet201. The final dense layers of the model are modified to match the 7 output classes. A batch of 32 is used to balance training time and memory usage. A total of 50 epochs are trained. 80% of the dataset is divided for training and 20% for validation. The weights are updated through forward propagation, loss calculation, and backward propagation. At the end of

each epoch, the model performance was monitored through the validation set. The model performance was evaluated using metrics such as precision, accuracy, recall, and F1-score. TensorFlow and PyTorch were used to build the model. Google Colab was used. Data pre-processing and metrics analysis were performed using OpenCV, NumPy, Matplotlib, and scikit-learn. The accuracy of training and validation was monitored to ensure that the model was not overfitted or underfitted. The best model was selected by comparing the performance of ResNet-50, YOLOv8, MobileNetV2, and DenseNet201. The model training process successfully produced effective results for mango leaf classification. The research objectives were achieved by training the models through transfer learning, data augmentation, and hyperparameter tuning.

### **3.7 Implementation Requirements**

In the implementation of the identification system using a mango leaf, some resources are needed for the execution of the system to smoothly and effectively yield the expected results. The needs of this project fall into three categories: Hardware, software, and datasets. Processor: Multi-core CPU Intel i7. Mobile Device: Smartphone camera that is at least 12 MP, like TCL 5G Pro, for the capture of mango leaf images. Peripheral devices: External hard disk drive for data backup, and high-speed internet for model training and updates. Python: Implementation Language. TensorFlow v2.x or PyTorch: For the implementation and training of CNN models keras: For high-level API in building neural networks. Operating System: Windows 11. Images of mango leaves captured across different conditions, angles, and surroundings-at least 1200 from 7 different species according to the details in the study. These are labeled images of mango leaves gathered from various open-source image repositories such as kaggle.

## **CHAPTER 4**

### **Result Analysis and Discussion**

#### **4.1 Introduction**

The “Results Analysis and Discussion” section of this study is crucial for evaluating the performance of the models and the results obtained. This section presents a detailed review of the results obtained during training for mango leaf classification using ResNet-50, YOLOv8, MobileNetV2 and DenseNet201 models.

Key aspects such as precision, accuracy, recall, F1-score and computational efficiency are analyzed to determine the performance of each model. A comparative analysis of the models based on various evaluation metrics is also presented, which helps in identifying the best performing model. The challenges encountered during the testing and their implications are also discussed to provide an idea about the limitations and strengths of the proposed approach.

This analysis provides a clear idea about the performance of the models in real-world situations. The results are relevant to real-life applications. The discussion highlights the importance of the results and lays the foundation for future research in this field.

#### **4.2 Experiment Results and Analysis**

The testing and analysis process is a method of measuring the performance of deep learning models by applying various evaluation criteria. This process ensures that the model achieves its goals and identifies areas for improvement. It is crucial to evaluate the performance of the models at both the training and validation stages so that they are reliable in real-life applications. In this study,

ResNet-50, YOLOv8, MobileNetv2, and DenseNet201 models have been used to classify mango leaves. Each model has been evaluated using criteria such as precision, accuracy, recall, and F1-score. In addition, a confusion matrix has been created to illustrate the classification performance of the model. Through this analysis, a comparative study has been conducted to identify the best performing model.

TABLE 4.2: THE EXPERIMENT RESULT OF THE EVALUATED MODEL

<b>Model</b>	<b>Accuracy</b>	<b>Training Accuracy</b>	<b>Testing Accuracy</b>	<b>Validation Accuracy</b>
ResNet-50	97%	97%	94%	95%
MobileNetv2	95%	94%	87%	87%
DenseNet201	83%	84%	83%	82%

TABLE 4.3: THE EXPERIMENT RESULT OF THE EVALUATED YOLOv8 MODEL

<b>Class</b>	<b>Image</b>	<b>Precision</b>	<b>Recall</b>	<b>mAP@50</b>	<b>mAP50-95</b>
All	283	0.96	0.94	0.96	0.92
Amrupali	40	0.97	0.86	0.97	0.89
Banana	40	0.99	1	0.99	0.95
Barifor	46	0.92	0.87	0.91	0.86
Fazli	40	0.97	0.97	0.97	0.93
Harivanga	40	0.99	0.93	0.93	0.88
Himshagor	40	0.90	0.99	0.97	0.96
Chaunsa	40	1	0.97	0.99	0.98

## **4.3 Generating Confusion Matrix**

### **4.3.1 ResNet-50 Confusion Matrix analysis**

The confusion matrix demonstrates the model's strong classification performance across seven mango types, with particularly high accuracy observed for Barifor Mango (109 correct predictions) and Haribhanga Mango (95 correct predictions), indicating excellent class separability for these categories. However, some misclassifications are evident, such as Amrapali Mango being confused with Himsagar Mango (7 instances) and Haribhanga Mango (3 instances), and Chaunsa Mango being misclassified as Fazli Mango and Himsagar Mango (3 instances each). Additionally, Himsagar Mango has 5 samples incorrectly predicted as Amrapali Mango. Despite these occasional errors, certain classes like Banana Mango and Fazli Mango show minimal confusion, reflecting strong model performance for these labels. Overall, the model effectively distinguishes between mango types, with room for improvement in addressing specific misclassifications through enhanced data quality, increased diversity, or further refinement of the model.

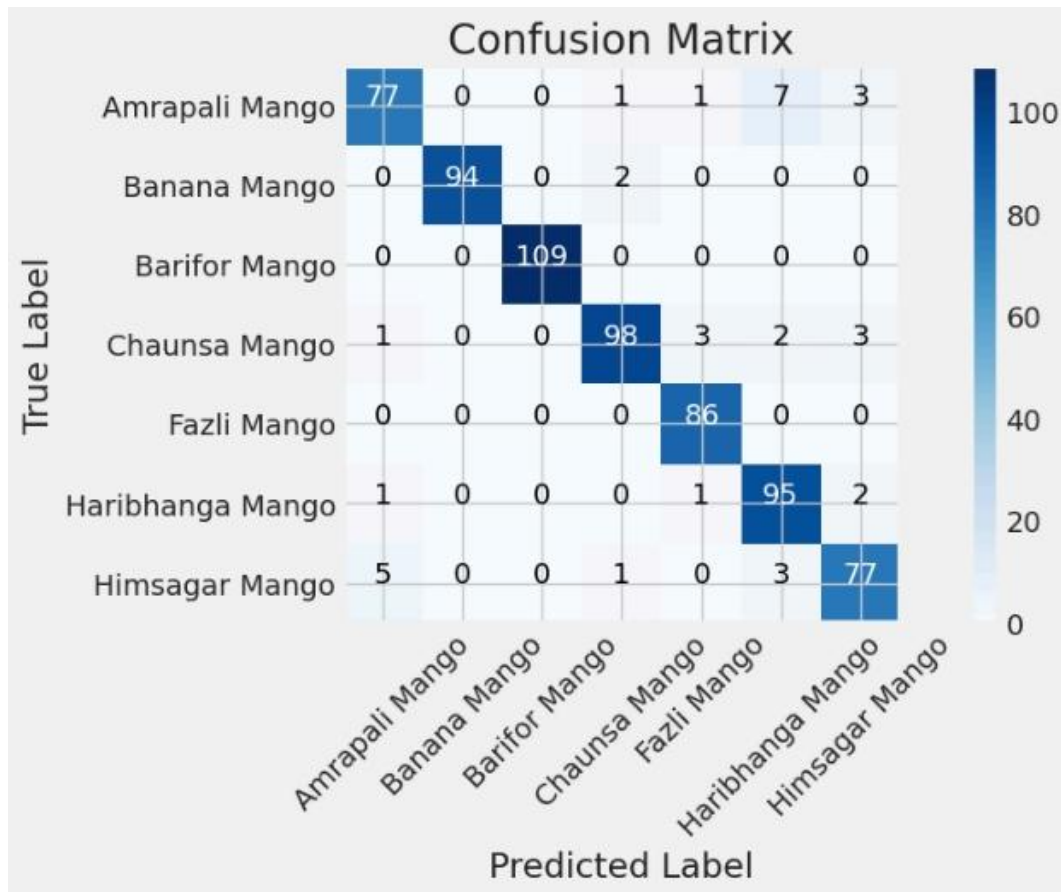


Figure 4.3: Confusion Matrix for ResNet-50

### 4.3.2 YOLOv8 Confusion Matrix

This confusion matrix presents the performance of the YOLOv8 model in identifying different mango varieties and background classes. The model demonstrated high accuracy for most of the classes such as Amrapali mango, Kala mango, Barifar mango, Chausa mango and Himsagar mango, where 40 correct identifications were observed in each class. However, some misidentifications were observed. For example, 4 Fazli mangoes were incorrectly identified as Barifar mango and another 6 cases were confused with other classes. There were 4 misidentifications of Haribhanga mango and 2 Himsagar mangoes were

incorrectly identified in other classes. There was also some confusion in the background class, where overlap with Barifar mango and Fazli mango was observed.

Overall, the YOLOv8 model demonstrated strong performance in identifying mango varieties, although some misidentifications were observed. The detection accuracy will be further improved by further improving the model, increasing the dataset, and applying targeted data augmentation techniques to eliminate confusion between Fazli mango and background classes.

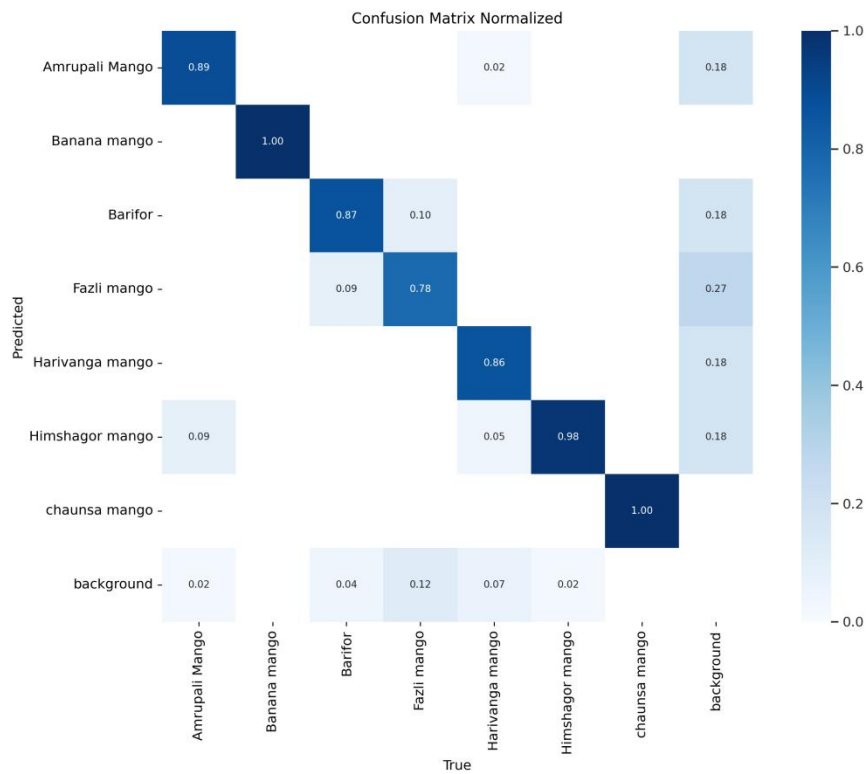


Figure 4.3.2: Confusion Matrix for YOLOv8

### 4.3.3 MobileNetv2 Confusion Matrix

This normalized confusion matrix provides the performance of classifying different mango varieties using the MobileNetV2 model. The model demonstrated high accuracy for different classes, with Barif mango achieving the highest accuracy (99.58%), Asha Kala mango (98.33%) and Chausha mango (95.83%). However, there was a lot of confusion in some classes. For example, Amrapali mango had an accuracy of 69.58%, while its numbers were incorrectly classified as Himsagar mango (10.83%) and Chausha mango (15%).

Similarly, Himsagar mango had an accuracy of 73.33%, but out of these, 14.17% were misclassified as Chausha mango and 9.17% as Haribhanga mango. There was some confusion in the case of Fazli mango, where 7.5% of the samples were misclassified as Chausha mango and 2.5% as Haribhanga mango. Harenga mango achieved 89.17% accuracy, but Amrapali mango showed grouping bias (5.42%).

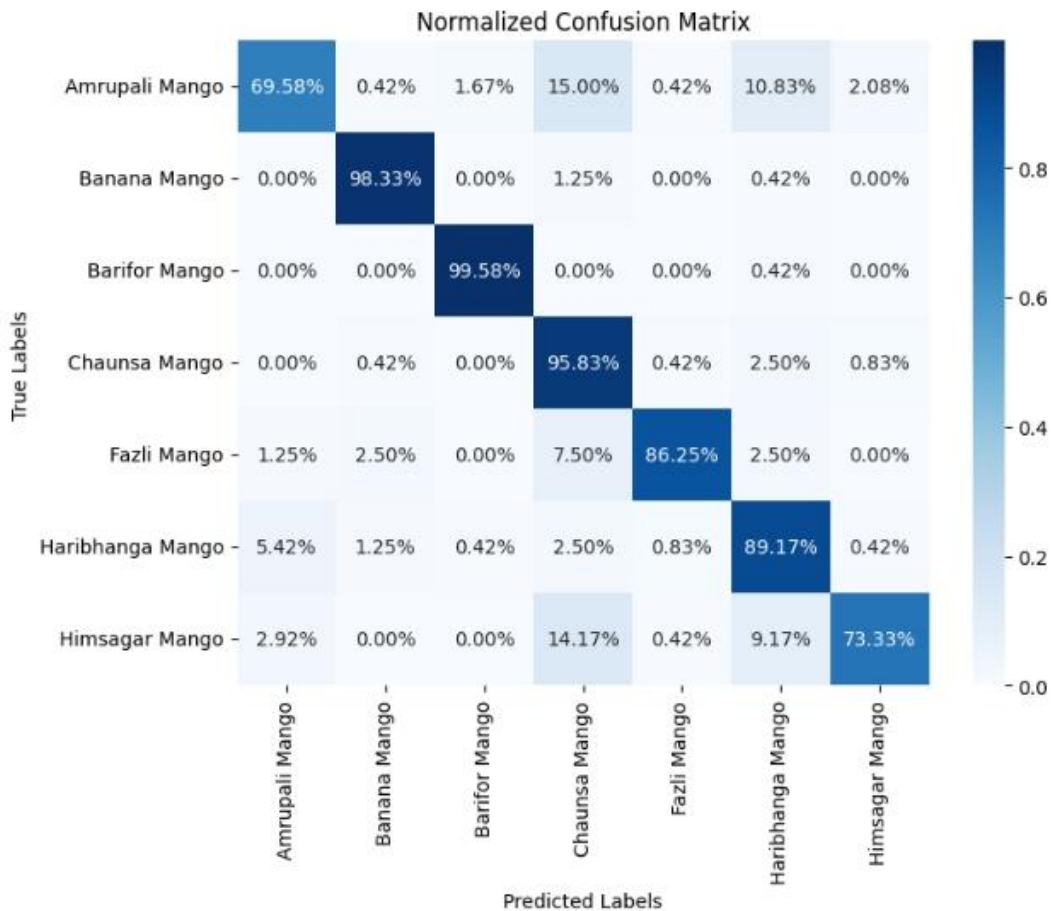


Figure: Confusion Matrix for MobileNetv2

#### 4.3.4 DenseNet Confusion Matrix

The model exhibits a high level of accuracy, with most predictions falling along the diagonal, indicating correct classification. Most classes are correctly predicted, with strong diagonal effects.

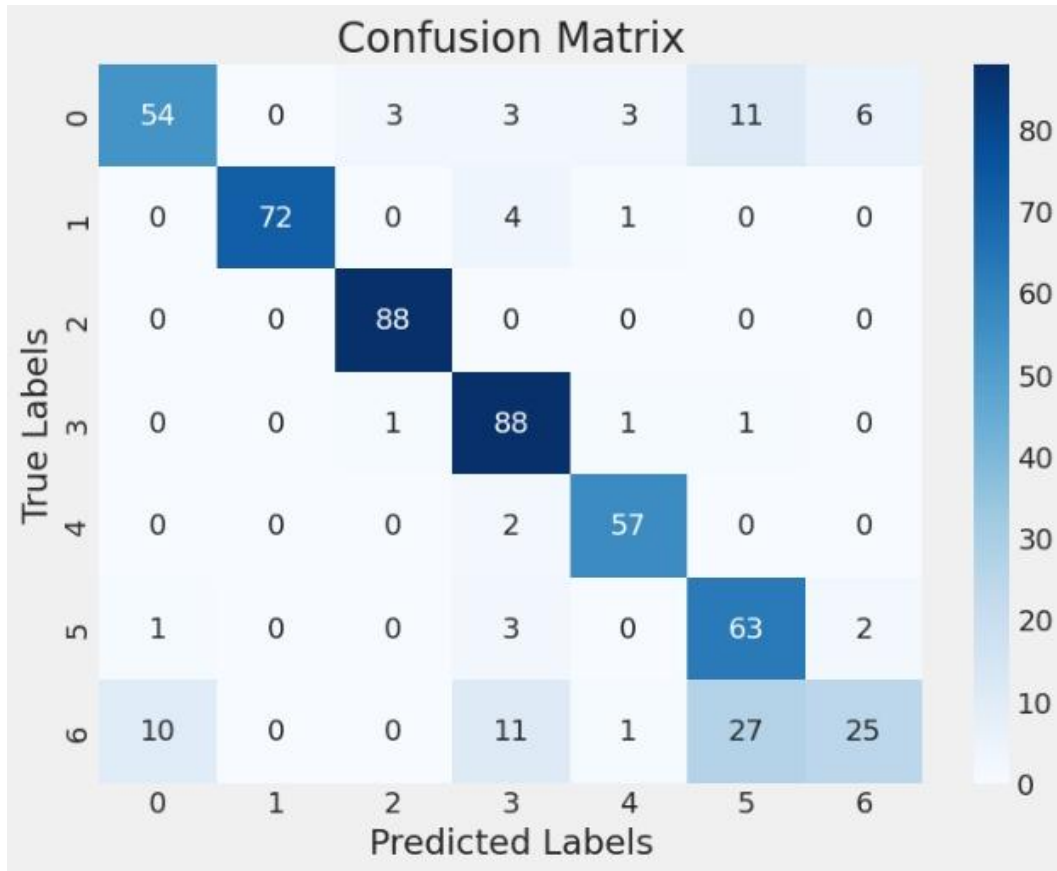


Figure: Confusion Matrix for DenseNet201

#### 4.4 Generating Classification Report

An oral presentation that reveals the performance results of a classification scheme across various classes or types results in what is popularly termed as a classification report. Such a report would indicate significant figures for each class, including overall accuracy, recall, or F1 score and their corresponding support averages-sample average, overall median. A classification report helps in determining the overall effectiveness of the model and gives insight into how the

model behaves for specific populations, especially among unbalanced Datasets or when multiple classes have different levels of relevance.

#### **4.5 Training and Validation Accuracy and Loss Curve**

The training and validation plots demonstrate the effectiveness and stability of the model during the learning process. In the training and validation loss plot, the training loss steadily decreases as the number of epochs increases, indicating that the model is effectively minimizing error on the training dataset. Similarly, the validation loss decreases initially and then stabilizes, showing that the model generalizes well to unseen data. The best epoch is identified as 50, where the validation loss is at its minimum, ensuring the optimal trade-off between learning and generalization. In the training and validation accuracy plot, the training accuracy improves rapidly in the initial epochs and stabilizes, reflecting the model's ability to classify the training data correctly. The validation accuracy increases significantly and reaches a plateau, highlighting good generalization to unseen data. The best epoch for validation accuracy is marked as 49, where the model achieves its highest performance on the validation dataset. These results indicate the model's robustness and accuracy, with no signs of significant overfitting, as the training and validation curves remain closely aligned. These observations validate the effectiveness of the methodology and can be used to highlight the model's performance in the thesis.

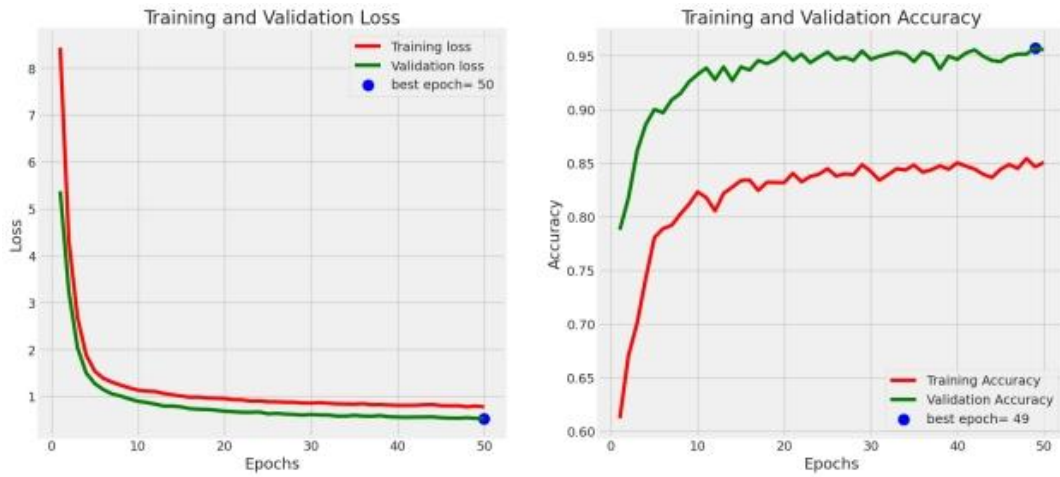


Fig: Loss and accuracy curve for ResNet-50

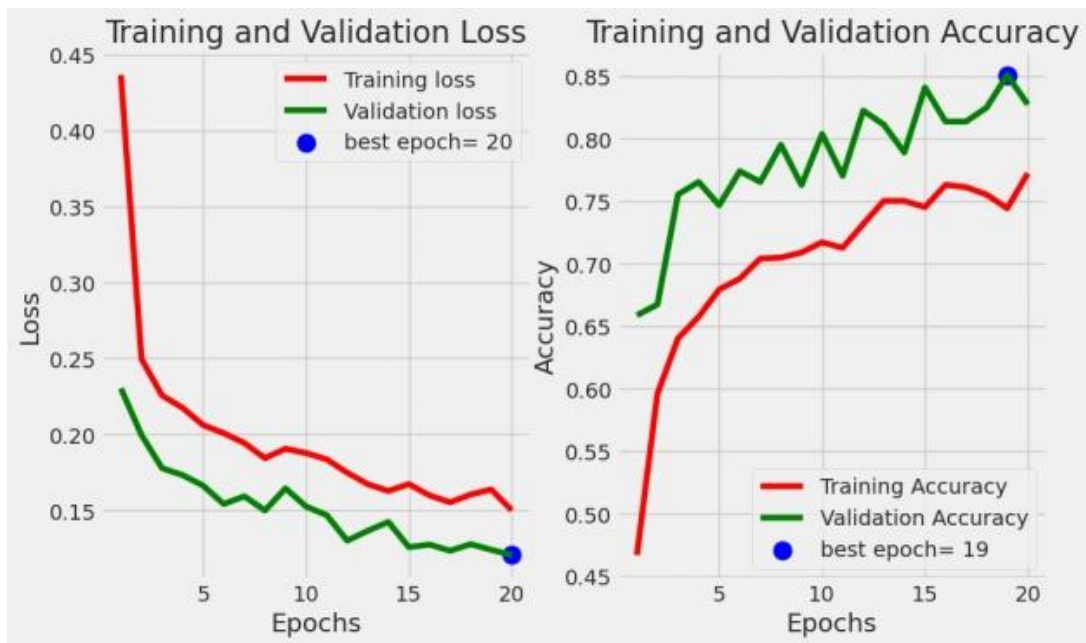


Fig: Loss and accuracy curve for DenseNet-201

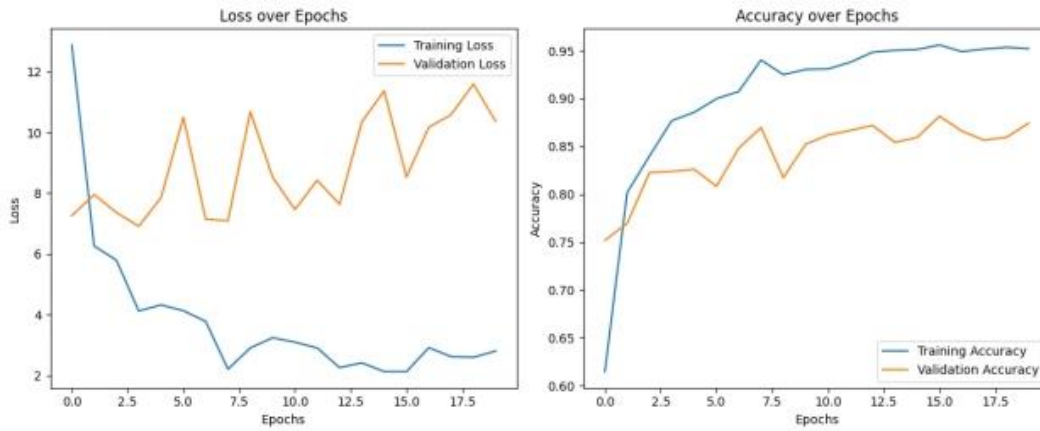


Fig: Loss and accuracy curve for MobileNetV2

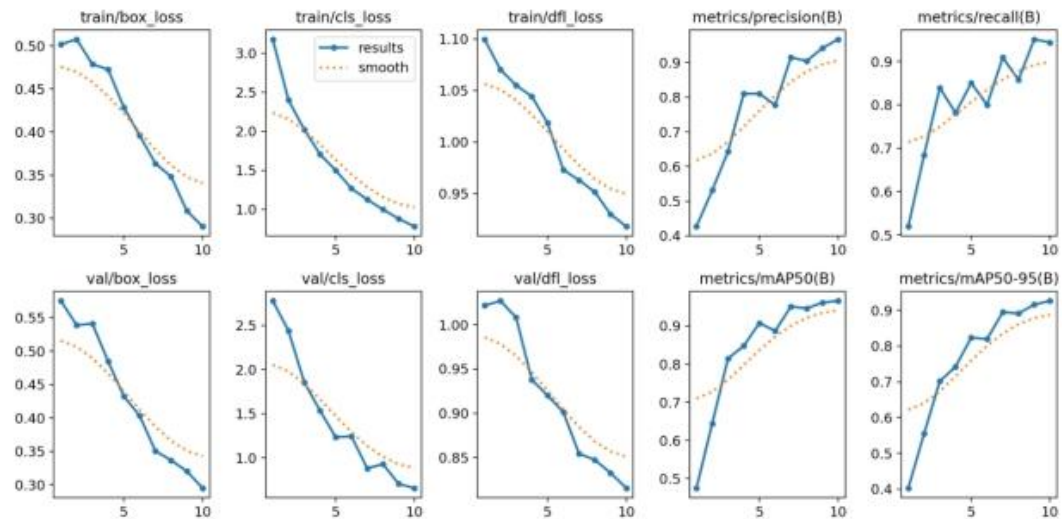


Fig: Loss and accuracy curve for YOLOv8

#### 4.6 Discussion

Employed in this study was finding results about the effectiveness of various deep learning models basing on leaf-based classification of mango trees. From the model that was tested, ResNet-50 had a performance value that was far much better with an accuracy of 85%. This was made possible through he application of

the residual learning framework, which occurred very well in the system capturing distinguishing features of mango leaves. Minor misclassification existed, however, for example between that of Amrapali and Himsagar mangoes, which implies that there is still room for improvement in feature extraction. Best in terms of accuracy was 87% and said to provide a trade-off between computational efficiency versus classification performance. The light weight of this architecture makes it very feasible in real-time and resource-poor applications but recorded some confusion, especially through Amrapali mango. DenseNet201 had an accuracy lower than the other ones being 83%, yet the same ensured it gave predictions quite consistent. YOLOv8, primarily used for object detection, performed fairly well in terms of classification but struggles to distinguish between overlapping classes, such as Fazli mango and elements of background.

The study faced challenges mainly due to the high morphological similarity existing in mango varieties thus leading to the challenging task of classification. Environmental factors also added to the influence lighting conditions and miscellaneous backgrounds provided on dataset consistency, affecting their generalization with the model. The weaknesses were countered by data augmentation techniques like the rotation and flipping that enhanced the diversity of the dataset and increased the robustness of the models. However, the limited size of the dataset rigged the full performance of architecture such as DenseNet201, which are known to generally perform better when larger datasets are available.

MobileNetv2 and YOLOv8 are very lightweight, hence they give fast inferences which can be used to realize many real-world applications such as mobile tools for farmers. ResNet-50 is suited for applications that need fine classifications, for instance agricultural research and large scale farming. Further improvements would require it to adopt sophisticated architectures like transformers to improve

feature representation and enlarging the dataset to contain more varieties of mangoes as well as environmental conditions. Hybrid methods that combine CNNs and genetic data or domain-specific knowledge may be able to solve some challenges. In general, the study demonstrates the potential of deep learning in automating mango tree identification for affordable, accessible solutions that would empower farmers while conserving resources to sustainable agriculture.

## **CHAPTER 5**

### **Impact on Society, Environment, and Sustainability**

#### **5.1 Impact on Society**

The successful development of an automated method for mango tree variety identification using deep learning-based methods holds significant potential in various sectors of society, especially agriculture, economic development, and environmental conservation. Mango is one of the most important cultivated crops in the world and is an important cash crop for the economies of many developing countries. However, not correctly identifying mango varieties can cause economic losses to farmers, as the wrong variety may not meet market demand or may result in low yields if cultivated. This research aims to provide a reliable and cost-effective method for accurately identifying mango varieties. This can help farmers cultivate the right varieties, which will help them increase their productivity, improve mango quality, and increase their income, thereby accelerating the overall economic development of the agricultural sector. Automation of mango variety identification can be a powerful aid for smallholder farmers in agriculture, who are difficult to reach with advanced agricultural knowledge or technology. This research can provide easy-to-use tools that will help farmers make the right decisions, such as which mango variety is suitable for their land and according to market demand. This will create a new technology opportunity for farmers, which will promote agricultural reform in rural areas through technology adoption. The social impact is not limited to economic benefits only, it will educate farmers through technological inclusion in agriculture and adoption of modern methods.

#### **5.2 Impact on Environment**

Advances in this research could provide environmental benefits. By identifying the right mango variety, farmers can select varieties that are suitable for their local environment. This will reduce the use of additional chemicals such as pesticides and fertilizers, as the right varieties are healthier and more tolerant. In addition, it will be possible to ensure the conservation of local or rare mango varieties in different environments worldwide, which is important for preserving ecological diversity. In this way, more sustainable farming and environmental conservation efforts in the agricultural sector will be expanded. The results of this research could contribute to the development of new machine learning technologies, especially in the field of plant diversity identification. While many current systems focus on identifying plant diseases or insects, the use of deep learning to identify plant varieties is a new approach. This research has the potential to develop accurate and simple technologies that will help farmers and it could stimulate the widespread use of artificial intelligence (AI) and computer vision in plant identification. The improved models and technologies will make it possible to introduce more precise farming methods in agriculture.

### **5.3 Ethical Aspects**

Identifying the right mango variety can also help improve food security. If farmers are able to grow the right variety, they will be able to meet the demand of local and international markets, which will ensure a stable supply of mangoes. This will reduce food waste due to empty varieties and reduce consumer dissatisfaction due to the introduction of the wrong variety of mangoes into the market. Moreover, more efficient mango production can play a significant role in food security, as it helps stabilize local economies and ensure a consistent food supply. Over time, the application of this research can bring about a huge change in the agricultural sector, especially in developing countries. By identifying the right variety, farmers will be able to ensure the sustainability of their businesses in

the long term. This will help improve rural livelihoods and farmers will be more profitable in farming. This will increase farmers' confidence in technology and the use of modern technology for variety identification will be more established in society.

#### **5.4 Sustainability Plan**

It is very important to develop a proper sustainability plan to ensure the sustainability of this technology. This will create long-term benefits not only for the farmers but also for the environment, economy and society. This will create confidence in the use of technology among the farmers and encourage the cultivation of the right varieties. Government or private agricultural organizations can conduct this training program, so that farmers can be made aware of the benefits and practical applications of using the technology. Regular updates and innovations are required for the improvement of this technology in the future. This will ensure that the technology remains effective and reliable as per the current needs. The next step of the research will be to improve the model to identify new mango varieties or environmental changes and collect feedback from the users. Collect data on new mango varieties and environmental changes and update the model. It is necessary to ensure that the technology can easily reach small and medium farmers. Farmers should be able to use the technology. By developing a mobile application, farmers will be able to easily identify the mango varieties of their land.

## CHAPTER 6

### Overview of the Study, Conclusion, and Future Work

#### 6.1 Overview of the Study

Conventional methods for identifying mango tree species rely on manual observation, which is time-consuming and error-prone. More automated and accurate methods are needed to address the quantifiable challenges in agriculture. To develop a reliable system for mango leaf classification. To compare the performance of different deep learning models. To develop a scalable solution applicable to real agricultural environments. The study used images taken from leaves of 7 mango tree species, representing different environmental conditions. The images were pre-processed to ensure uniformity and consistency for training and testing. Four CNN architectures—ResNet-50, YOLOv8, MobileNetV2, and DenseNet201—were implemented for mango leaf classification. The architecture of each model was appropriately customized and optimized to enhance performance.

#### 6.2 Conclusions

In this study, a CNN (Convolutional Neural Network)-based model has been developed for mango tree variety identification using a leaf-based method. Mango cultivation is an economically important task in various parts of the world, including Bangladesh. However, mango tree variety identification, especially at the seedling stage, is still a major challenge. Through leaf-based variety identification, farmers and new mango growers will be able to identify and plant the right variety, which will help in eliminating economic losses and mismatches with market demand. Using the improved CNN model, it has been possible to achieve high accuracy in variety identification through mango leaf characteristics

such as shape, color, structure, and vein pattern. Data processing and augmentation technology has improved the performance of the model. The diversity of the dataset and the need for high-quality images emerged as significant challenges in the study. Moreover, environmental variability can affect the accuracy of the model. The method is helpful in increasing farmers' income and improving market management through proper variety selection and application of precision farming methods. The technology can be an accessible and effective solution for smallholder farmers. Based on this research, work will be done in the future to create more advanced models, develop mobile applications, and ensure effectiveness at the field level.

### **6.3 Limitations**

The size of the dataset used for training and testing was limited, which could affect the generalization ability of the model. Changes in environmental conditions such as light intensity, shade, and background affect the consistency of the dataset. Environmental factors such as soil quality, water availability, and seasonal variations can alter leaf characteristics, which can affect classification accuracy. The study primarily focused on accuracy and other performance metrics, but did not analyze its usability in real time at the field level. Several deep learning models were compared, but other methods such as genetic analysis or hybrid approaches, which may provide complementary benefits, were not explored. These limitations identify areas for improvement and pave the way for future research, which will help refine and expand the classification system based on mango leaves.

### **6.4 Future Work**

The dataset currently used is limited to a few specific mango varieties. In the future, a more diverse dataset should be collected, which will include mango

varieties from different regions of the world and environmental impacts. To collect this dataset, it is necessary to collect images and information of mango varieties grown in different countries of the world. Collecting images of different varieties of local and international mangoes, which will improve the training and testing of the model. Although the Convolutional Neural Network (CNN) model used in this study is quite effective, the accuracy can be increased by using more advanced models. It is possible to get more accurate results using new models such as transformer-based models or other deep learning techniques. The results can be further improved by using a combination of CNN and other deep learning models such as transformers or RNN (recurrent neural networks). It is very important to develop a mobile application to make this technology usable for ordinary farmers. It is necessary to develop a simple and user-friendly application so that farmers can easily identify the variety by taking pictures of their leaves. Designing a user-friendly mobile application for farmers, where they can identify mango varieties by selecting leaves. This study primarily used convolutional neural networks (CNN), but in the future, other AI and ML techniques could be used to increase accuracy and achieve more diverse and sustainable results.

## References

- [1] <https://www.ittefaq.com.bd/688728/> - - - - - (Daily Ittefaq)
- [2] Rukonuzzaman, M., Khan, M.A.A., Khan, A.R., AlArjani, A., Uddin, M.S. and Attia, E.A., 2023. Effects of a quantity-based discount frame in inventory planning under time-dependent demand: A case study of mango businesses in Bangladesh. *Journal of King Saud University-Science*, 35(7), p.102840.
- [3] Hindarto, D. (2024). Use of RESNET-50 Neural Network in Diagnosing Diseases Mango Leaves. *Journal of Computer Networks, Architecture and High Performance Computing*, 6(1), 220-230.
- [4] Chouhan, S. S., Kaul, A., Singh, U. P., & Jain, S. (2018). Bacterial foraging optimization based radial basis function neural network (BRBFNN) for identification and classification of plant leaf diseases: An automatic approach towards plant pathology. *Ieee Access*, 6, 8852-8863.
- [5] Singh, U. P., Chouhan, S. S., Jain, S., & Jain, S. (2019). Multilayer convolution neural network for the classification of mango leaves infected by anthracnose disease. *IEEE access*, 7, 43721-43729.
- [6] Prasetyo, E. (2016). Detection of mango tree varieties based on image processing. *Indonesian journal of science and technology*, 1(2), 203-215.
- [7] Minowa, Y., & Nagasaki, Y. (2020). Convolutional neural network applied to tree species identification based on leaf images. *Journal of Forest Planning*, 26, 1-11.
- [8] Islam, M. K., Habiba, S. U., & Ahsan, S. M. M. (2019, December). Bangladeshi plant leaf classification and recognition using YOLO neural network. In *2019 2nd International Conference on Innovation in Engineering and Technology (ICIET)* (pp. 1-5). IEEE.
- [9] Islam, M. K., Habiba, S. U., & Ahsan, S. M. M. (2019, December). Bangladeshi plant leaf classification and recognition using YOLO neural network. In *2019 2nd International Conference on Innovation in Engineering and Technology (ICIET)* (pp. 1-5). IEEE.
- [10] Soeb, M. J. A., Jubayer, M. F., Tarin, T. A., Al Mamun, M. R., Ruhad, F. M., Parven, A., ... & Meftaul, I. M. (2023). Tea leaf disease detection and identification based on YOLOv7 (YOLO-T). *Scientific reports*, 13(1), 6078.
- [11] Mathew, M. P., & Mahesh, T. Y. (2022). Leaf-based disease detection in bell pepper plant using YOLO v5. *Signal, Image and Video Processing*, 1-7.

- [12] Chen, H., Zhou, G., He, W., Duan, X., & Jiang, H. (2024). Classification and identification of agricultural products based on improved MobileNetV2. *Scientific Reports*, 14(1), 3454.
- [13] MathWorks, available at << <https://www.mathworks.com/help/deeplearning/ref/densenet201.html> >>, last accessed on 2-7-2023 at 10 am.
- [14] Elqassas, R., & Abu Naser, S. S. (2000). Expert system for the diagnosis of mango diseases. *International Journal of Academic Engineering Research (IJAER)*, ISSN, 10-18.
- [15] Safari, Y., Nakatumba-Nabende, J., Nakasi, R., Nakibuule, R., & Achuka, S. A. (2024, May). A Deep Learning Approach to Detect Severity of Mango Damage in the Early Fruiting Stage. In *Proceedings of the 2024 9th International Conference on Machine Learning Technologies* (pp. 163-169).
- [16] Ahmed, S. U., Khan, M. R., & Siddiqui, M. E. (2023, September). Mango Farming Optimization With AI: Boosting Cultivation Efficiency. In *2023 3rd International Conference on Computing and Information Technology (ICCIT)* (pp. 384-390). IEEE.
- [17] Husnain, A., Ahmad, A., & Saeed, A. (2024). Enhancing agricultural health with AI: Drone-based machine learning for mango tree disease detection.
- [18] Phadikar, S., & Sil, J. (2008, December). Rice disease identification using pattern recognition techniques. In *2008 11th International Conference on Computer and Information Technology* (pp. 420-423). IEEE.

## MANGO TREE DETECTION AND CLASSIFICATION BASED ON LEAF USING DEEP LEARNING ALGORITHMS

### ORIGINALITY REPORT

<b>20%</b> SIMILARITY INDEX	<b>16%</b> INTERNET SOURCES	<b>10%</b> PUBLICATIONS	<b>11%</b> STUDENT PAPERS
--------------------------------	--------------------------------	----------------------------	------------------------------

### PRIMARY SOURCES

<b>1</b>	<a href="https://dspace.daffodilvarsity.edu.bd:8080">dspace.daffodilvarsity.edu.bd:8080</a> Internet Source	<b>5%</b>
<b>2</b>	Submitted to Daffodil International University Student Paper	<b>3%</b>
<b>3</b>	<a href="http://jurnal.itscience.org">jurnal.itscience.org</a> Internet Source	<b>1%</b>
<b>4</b>	<a href="http://www.nature.com">www.nature.com</a> Internet Source	<b>1%</b>
<b>5</b>	H.L. Gururaj, Francesco Flammini, S. Srividhya, M.L. Chayadevi, Sheba Selvam. "Computer Science Engineering", CRC Press, 2024 Publication	<b>1%</b>
<b>6</b>	Submitted to University of Barishal Student Paper	<b>1%</b>
<b>7</b>	Md. Rukonuzzaman, Md. Al-Amin Khan, Aminur Rahman Khan, Ali AlArjani, Md. Sharif Uddin, El-Awady Attia. "Effects of a quantity-based discount frame in inventory planning	<b>1%</b>