VISION-BASED RIPE MANGO CLASSIFIER USING MACHINE LEARNING

A Capstone Project Presented to the Faculty of the College of Computer Studies St. Michael's College
Iligan City

By

SUMALPONG, JAYRALD M.
TONG, JOSHUA SHENG JI

December 2024

VISION-BASED RIPE MANGO CLASSIFIER USING MACHINE LEARNING

A Capstone Project Presented to the Faculty of the College of Computer Studies St. Michael's College Iligan City

By

SUMALPONG, JAYRALD M.
TONG, JOSHUA SHENG JI

December 2024

Title	: Vision-Based Ripe Mango Classifier using Machine Learning
Author	: Jayrald M. Sumalpong Joshua Sheng Ji U. Tong
No. of Pages	: 69
Host Institution	: St. Michael's College of Iligan, Inc.
Address, Region	: Iligan City, Region X

Abstract

Classifying the ripeness of mangoes, detecting diseases, and identifying mangoes from non-mango objects are important tasks in ensuring mango quality. This study developed a system using YOLOv5, a machine learning algorithm, to help farmers, consumers, and mango producers. The system can determine whether a mango is ripe, detect diseases like Anthracnose, Stem-end Rot, and Black Mould Rot, and check if an object is a mango or not. The system was tested using images and a live camera. For images, it achieved an accuracy of 83.33% in detecting mangoes, 52% in classifying ripeness, and 84% in detecting diseases. In live camera testing, the system performed better, with 100% accuracy for mango detection and 96.67% for ripeness classification. Although there were challenges in preparing the dataset, training the model, and combining different tasks, the system proved to be accurate, reliable, and easy to use. This project provides a helpful tool for improving the process of classifying mangoes and reducing errors in the agricultural field. Thus, the objectives of this study are met: detecting mangoes, classifying ripeness, and disease detection using machine learning. Though the models are proven to be lacking and inconsistent, it is doing what it is intended to be and delivering the needed functionalities for the users.

209 words

Keywords: YOLOv5, machine learning

DEDICATION

This journey truly has been a rollercoaster ride that is filled with unforgettable memories and moments to be cherished. I would like to dedicate this study to the individuals who have been my pillars of support throughout this journey.

I would like to express my gratitude to the Almighty God for accompanying me and to be there and guiding, giving me strength, and for keeping me sane throughout these 4 years of ups and downs and one of the difficult rides of my life.

To the CCS faculty, thank you for sharing the wisdom and strength everyday to teach us and shaping us to be a competent and becoming a stronger individual and to our adviser thank you for the incessant guidance and support.

I am sincerely grateful to my family for supporting my studies, especially to my mom who's giving me reason to work hard and supporting me financially. To my friends and super friends, thank you for always being there when I needed you most—for your advice, companionship, and countless coffee sessions that kept me motivated. I dedicate this study to all individuals that supported me.

This study is dedicated to all of you who stood by me, supported me, and believed in my abilities. Your contributions have made this accomplishment possible, and for that, I am forever grateful.

-Jayrald

DEDICATION

下半年無論是學術還是非學術，都是非常艱苦的。說到這裡，我想說幾句話。

我要感謝上帝陪伴我走過這段艱難的旅程，賜予我繼續前進的力量並引導我。我還要感謝幫助和激勵我的CCS教師。感謝我的家人的支援。感謝一直在我身邊的朋友，ML 和 Chicken Nuggets。還要感謝所有和我在一起的人，尤其是羽毛球和跑步（Slowdy）.

나의 동기 부여와 영감의 원천에 대한 특별한 언급 : NewJeans (5로 돌아오십시오). 영원히 토끼; Aespa - 겨울 영원히 편견; ITZY - 류진 - ahhhh; 트와이스 - 백만 명 중 한 명; 레드 벨벳 - SSW 최고의 가수; Day6 - 베스트 밴드; 그리고 Wicked Playlist. 끝없는 청취 세션에 감사드립니다.

With that being said, this is dedicated to everyone who has supported me, this study will be etched in my mind forever (due to its complexity and stress).

-唐聖吉

Acknowledgement

This study would not have been possible in its present form without the following persons who, in one way or another, contributed their support, resources, time, effort, and expertise:

To Maam Leonie A. Cajes, our supportive adviser; we thank you for your sustained guidance and encouragement that led to the completion of this study.

To Ma'am Fe, Sir Veldin and Ma'am MJ, we thank you for your comments and suggestions to make our system more efficient. Your valuable insights play an important role in shaping our system. We deeply appreciate your time and dedication in contributing to the improvement of this study.

To Sir Rhun, who is not only a mentor but also a friend of ours, thank you for your unwavering support, inspiring words, and for always believing in our potential. Your guidance and encouragement have made a lasting impact on us, and for that, we are truly grateful.

To the researchers family; we thank you for your whole hearted and untiring support-financially and emotionally-and for strengthening our patience and courage to strive harder for the fulfillment of this study,

To all of the researchers' friends and classmates who, in one way or another, have given their support while working in this study-thank you!

Above all, to our Almighty God, who is a sustainer in our endeavors, provider of our needs, and keeper of the bonds that unite us until the realization of the manuscript.

-Jayrald Sumalpong & Joshua Sheng Ji Tong

Table of Contents

List of Tables

List of Figures

LIST OF ACRONYMS AND ABBREVIATIONS

YOLO	You Only Look Once
AI	Artificial Intelligence
ML	Machine Learning
IU	Image Upload
LC	Live Camera

.

Chapter 1

THE PROBLEM AND ITS SETTING

Introduction

The Mango (Mangifera indica L.), member of family Anacardiaceae, is amongst the most important tropical fruits of the world. The opportunity for breeding improvement in mango is significant and challenging. There is a lot more varietal wealth available but certain inherent constraints are involved like: long juvenility, high clonal heterozygosity, one seed per fruit, recalcitrant seeds, polyembryony, early post-zygotic auto-incompatibility and large area requirement for assessment of hybrids [1].

Ripe mangoes are crucial for the economic, nutritional, and medicinal value. It supports livelihoods through agriculture and exports, with studies in Brazil, India, and Indonesia demonstrating the economic viability and profitability of mango cultivation, showing Benefit-Cost Ratios (BCR) ranging from 1.76 to 4.7, and strong profitability indicators like Net Present Value (NPV) and Internal Rate of Return (IRR) [2, 3, 4, 5]. Additionally, ripe mangoes provide essential vitamins and antioxidants, and are used in traditional medicine to treat various ailments. The demand for high-quality ripe mangoes also drives innovation in postharvest handling and marketing, enhancing their global importance.

Mango is subject to a number of diseases at all stages of its development. Some of these diseases cause heavy loss and are a limiting factor in mango cultivation in some regions.

Major pests and diseases are mango twig borer, scab, anthracnose, and stem-end rot. Scab affects young fruit, twigs, leaves, and blossoms, while anthracnose and stem-end rot are the most destructive, impacting almost all parts of the mango tree and limiting fruit storage and shelf life [6]. Mango production in the Philippines faces several challenges that hinder its potential. Issues include climate change, pests and diseases, poor nutrition, low adoption of new technologies, post-harvest losses, and lack of government support.

To overcome these challenges and enhance mango production, focused efforts in research and development are essential. Improvements in pest and disease management, cultivation techniques, and post-harvest practices can greatly increase yields and improve fruit quality. Increased government support and effective policies are crucial for sustainable growth in the mango industry.

AI technology helps farmers get high-quality agricultural crops. The essential idea of AI in agriculture is its flexibility, reliability, speedy performance, and applicability. AI technology improves enterprise performance and productivity by automating processes or tasks that once required human skill [7].

In agriculture, image processing and computer vision are used extensively for tasks like identifying, classifying, grading, and evaluating mango quality.

However, identifying defects and determining fruit maturity remain challenging tasks for these technologies [8].

A promising solution lies in the application of machine learning to detect and manage mango crop diseases. Machine learning algorithms can analyze large datasets of images and environmental conditions to identify early signs of diseases and recommend timely interventions.

Furthermore, machine learning supports the development of precision agriculture techniques. By combining data from sensors, drones, and satellite photos, farmers may optimize resource use and increase yields by making educated decisions about fertilization, irrigation, and pest management.

The researcher conducts this study because the mango is one of the most important fruits internationally and nationally. Mango also helps farmers earn for their daily needs and improves the economy of the country. However, mango farmers face many difficulties, including a variety of diseases at different parts of the mango's developmental stage. These difficulties have a negative impact on both the quantity and quality of mango production. Traditionally, detecting these diseases can be labor-intensive, prone to human error, and inefficient since not every farmer knows what kind of disease the fruit has, leading to a loss of income or revenue. With AI advancements over the years, including machine learning, it is now cost-effective and efficient to detect diseases from a given dataset. Machine learning can also be used to provide information and solutions that are beneficial and important to farmers. This study aims to use machine learning's capabilities in identifying mango diseases, addressing a basic yet crucial factor in the mango industry.

Statement of the Problem

Mango farmers, particularly in the Philippines, encounter significant challenges that affect their productivity and profitability. These challenges include:

1. Difficulty in accurately identifying the ripeness of post-harvested mangoes, which impacts market value and consumer satisfaction.
2. Inefficient tools and methods for detecting diseases in mangoes, leading to increased losses and reduced crop quality.
3. Some consumers have difficulty in classifying the ripeness of a mango which can be time-consuming and may lead to poor purchases decisions or food wastage.

Objectives of the Study

The objective of this study is to develop a Vision-based Mango Ripeness Classifier and Disease Detection System using Machine Learning. Specifically, it aims to:

1. To design a system using machine learning for mango farmers to detect and classify ripe and unripe post-harvested mango with real-time detection or image uploads

2. To develop a robust disease detection feature within the system that identifies common mango diseases, enabling early diagnosis, and intervention to reduce post-harvest losses.
3. To test and evaluate the system's performance in classifying the ripeness of a mango, as well as to assess its usability, accuracy, and effectiveness in providing reliable results for end-users.

Scope and Limitation of the Study

This study focuses on the Vision-Based Ripe Mango Classifier using Machine Learning system, a web-based platform utilizing Python for the machine learning model and YOLOv5 as the algorithm. YOLOv5 employs deep learning techniques, specifically convolutional neural networks, to perform real-time object detection and classification. This system is intended for farmers, mango farm owners, and consumers of mangoes throughout the Philippines.

The system will allow users to upload a photo, or use a real-time camera for detecting if the mango is ripe or unripe without the need for creating an account or logging in. Additionally, the system also has a feature that can detect mango diseases as part of the ripeness classification.

Regarding the detection, the system will identify if the mango is ripe or unripe, or if the mango has any disease/s, name of the disease, and characteristic features.

For ripeness classification, mangoes that are ripe and are color green will be identified as unripe as this is not part of the model's scope. Only yellow ripe mangoes, and green unripe mangoes will have an accurate result.

Diseases that the system can identify are only the following: Anthracnose, Stem-end Rot, and Black Mould Rot.

Additionally, the system will also have a resources page that showcases important information like: name, characteristic features, and recommendations of the 3 mango diseases available in the system.

It should be noted that the photos must be good quality, meaning they must have a dimension of at least 500x500. It's better if the dimensions are larger, though the process may take some time. The estimated processing time for image upload is 1.5 - 3 seconds, while the live camera is 1 - 2 seconds but due to the size of the model, when the camera detects a mango the processing makes the frame rate of the camera drop resulting in a lag. The image should be clear and not pixelated; the file size must not exceed 100 MB; the image format should be PNG or JPEG; it should have proper lighting (i.e., not dark); there should be no distractions such as blurriness or graininess that may affect the detection process; it should have proper color representation; and there should be no occlusions. This ensures that the system can work properly and provide the necessary results. Additionally, not every disease of mangoes will be identifiable by the system, only those included in the predefined dataset and for which the model has been trained. Following these notes will ensure an accurate result.

Significance of the Study

The following entities that will benefit from this study are the farmers, mango farm owners, mango consumers, retail and distributors, current researchers, and future researchers.

Farmers. Farmers would benefit from the study; this is because it deals with specific problems that have to do with them, and as a result, they will learn about proper mango management so as to classify ripeness, and avoid the diseases that affect the crop's growth and development. So, the farmers can correct their farm management according to this research's suggestion.

Owners of Mango Farm. By using a system for classifying the ripeness level, and diseases in mangoes, owners of farms can inspect the health of mangoes regularly by themselves and in case of any infection, they can take appropriate measures. Even without consulting professionals, farm owners can identify these very well. This assures their buyers that what they have bought is properly taken care of and is safe to consume.

Consumers of Mangoes. There are several advantages to consumers from this system which include: it sees to it that all consumer purchases come with a thorough checkup on them ensuring what they buy is what they need; customers may utilize the system on their phone where they could either upload photographs or use their phone camera in order to take images of the fruit or even examine the mango in real-time to assess its ripeness level, and if there are symptoms indicating illness like discoloration around its skin or rottenness at its bottom tops.

Retailers and Distributors. If retailers and distributors are utilizing this system, they are able to sell mangoes while at the same time protecting the business' image and reputation. The system helps in winning customer confidence that enhances sales and improves revenue.

Future Researchers. The future researchers can obtain lessons from the previous studies on vision-based ripeness classification that uses machine learning in many ways. Using the past research as a foundation, they can build on existing knowledge, identify the gaps that need to be addressed and further explore, refine machine learning methods, and provide valuable insight to inform agricultural practices and policies. With this process, they can contribute to the ongoing progress and development in the field of agricultural technology and crop disease management.

Definition of Terms

Classifier. Classification is the process of grouping information according to specific attributes. In machine learning, a classifier is an algorithm that automatically arranges or groups data into one or more of a set of "classes" (*Classifier Definition*, n.d.).

In the system, a classifier is used to identify if a mango is ripe, if it has a disease, and vice-versa.

Machine Learning. Machine learning (ML) is a branch of artificial intelligence (AI) and computer science that focuses on using data and algorithms to enable AI to imitate the way that humans learn, gradually improving its accuracy [9].

In the system, machine learning enables automated identification and classification of mango diseases by analyzing images that allows for early detection and effective crop management.

Mango. A tropical usually large ovoid or oblong fruit with a firm yellowish-red skin, hard central stone, and juicy aromatic pulp [10].

In the system, mangoes are the primary focus, with the goal of detecting and classifying diseases to ensure the quality.

Ripe. Completely grown or developed and prepared for consumption or use (*Ripe - Definition, Meaning & Synonyms*, n.d.).

In the system, ripe post-harvested mangoes with diseases are the primary focus, with the goal of detecting ripe mangoes.

Vision-based. Computer-based technologies known as vision-based systems replicate human vision to carry out a range of functions (*What Are Vision Based Systems? | IndMALL*, n.d.).

The system will be utilizing vision-based technology in order to classify ripe mangoes, and as well as mangoes with diseases.

Chapter 2

REVIEW OF RELATED LITERATURE AND STUDIES

This section provides a comprehensive analysis and evaluation of the existing literature and studies related to the research topic.

Related Literature

In recent years in the agricultural sector, machine learning and image processing have played a huge role in identifying fruit defects. It is said to be nondestructive, low-cost, fast, reliable, and real-time which is essential and plays a huge role in the agricultural sector [11].

Another way to look at this is that not all farmers are literate and know about mango diseases. For example, based on a study by Veling, not all farmers are literate in his country [12]. Which means they can't get the necessary information about a disease and would require an expert. However, it is not easy for experts or agricultural officers to reach every farmer. According to Tumang, in the Philippines, there are approximately 2.5 million farmers who rely on mangoes for their daily incomes, and one of the difficulties they face is the decrease in mango production due to infestation of pests and diseases [13]. Thus, one of the most basic yet important aspects in the fruit industry is identifying fruit diseases [14].

Machine learning (ML) can plan, process, and perceive data, whether small or big – helping different people in different sectors making sound decisions and gaining relevant information. With machine learning's broad range in agriculture, studies have been dedicated to weed detection, yield detection, crop recognition, and disease detection [15].

You Only Look Once (YOLO)v5 is the 5th version of the YOLO model. It is a computer vision model used for real-time object detection which is used in machine learning [16], and the model used in this system. For instance, it has been used among other uses for the detection of strawberry diseases [17] and real-time vehicle detection [18].

Many researches have indicated the potentiality of machine learning in detecting mango diseases [19]. With extensive progress made in machine learning and computer vision, there are various automated diagnostic solutions for analysis of mango diseases that are machine learning algorithm based [20]. It has helped farmers take precautionary actions minimizing loss of yield and preventing quality loss leading to increasing the income of farmers and the quality of mangoes [21]. Thus, the continuing advancement and improvement in computer vision techniques has helped various agriculture problems, like mango fruit disease detection, easier [22].

Mango is very rich in carotene, which fights against oral cavity and lung cancer apart from its distinct fragrance and flavor. It is a good source of vitamins, minerals, fibers, prebiotic dietary, and antioxidant compounds that improves human health. On the other hand, new studies show that eating mangoes can keep breast leukemia and prostate cancer away [23]. Therefore, to get the most out of these fruits people should fight off anything that compromises the fruit.

Related Studies

Farmers face many challenges including climate change and crop production. One issue is the lack of awareness about crop diseases that lead to the poor growth of the crop. By using the Internet of Things (IoT), it offers a solution to connect physical objects like sensors and actuators to the internet, that enables better collaboration, monitoring, and visibility of plant status and growth. IoT devices are placed in farm fields and can detect plant diseases, alert the farmers, provide training, and offer solutions to problems. A wireless sensor network (WSN) can monitor plant diseases and gather environmental data, while an automatic learning decision supports a system that can manage disease detection and provide reliable results. This technology-driven approach helps the farmers to make informed decisions, improves efficiency, and enhances crop productivity [24].

According to Polly and Devi, Agriculture is relevant in feeding the human and cattle populations globally. The increasing agricultural productivity requires fast and precise disease diagnosis. Farmers face many challenges such as water shortages, fertility issues, pests, diseases, low yields, and financial problems. The traditional methods for identifying the issues are often ineffective. Early disease detection can prevent crop loss, improve plant productivity, and enhance product quality. Image processing techniques, especially those that involve deep learning, offer advanced solutions for these agricultural problems [25].

The study analyzes the latest developments and challenges in using deep learning and advanced image processing technologies for ripeness classification and disease identification. It examines many agricultural problems that are being researched, the models used, data sources, and the accuracy that is achieved according to the researcher's performance criteria. The analysis aims to enhance the system's reliability and efficiency, enabling further research to explore the deeper capabilities of deep learning in the detection and categorizing of fruit ripeness and diseases [25].

The study investigates the machine learning techniques in agriculture and specifically for classifying two usual diseases in mango: Powdery Mildew and Sooty Mould. The study uses the MangoLeafBD dataset, the study applies a Gradient Boosting Classifier that enhances the mean shift image segmentation and Hue moments for the feature extraction. The model's performance was thoroughly evaluated using the 5-fold cross-validation, which provides insights into its accuracy, precision, recall, and F1 score. The results will indicate moderate success and with the highest accuracy and precision observed in the initial fold, it demonstrated the model's potential for important disease identification. The study reviews the challenge of differentiation between diseases with similar symptoms by introducing an innovative data-driven approach for managing different diseases in cultivating mango. The research contributes to precision in agriculture by showcasing the potential of machine learning to improve disease diagnosis and treatment strategies by supporting sustainable agricultural practices [26].

Kumar et al. stated that mangoes have a major impact on economic and environmental roles in India and are prone to a range of diseases that affect quality and economic growth [27]. The mango diseases have a significant impact, with over 1500 species and over 1000 commercial varieties that are being grown in India. A common fungal infection, anthracnose, can cause significant harm to mango trees. Timely identification of anthracnose is important for mango farm owners to provide early treatments, protect fruit quality, and increase yields. Recent developments in deep learning and computer vision provide precise computational capabilities to identify fungal and bacterial diseases in mango trees. The study introduces a new CNN structure that was created for detecting anthracnose in mango trees. It is based on real-time data collected from farms in Karnataka, Maharashtra, and New Delhi. The new algorithm attained a classification accuracy of 96.16%, surpassing the other modern methods [27].

The study analyzes the use of artificial intelligence to classify the different types of mango varieties in the Philippines based on their leaf characteristics, mainly Indian mango, Carabao, and Saperada mangoes. Due to the difficulty and inaccuracy of identification, AI techniques, especially fuzzy logic and K-nearest neighbor (K-NN), were employed. The process included image processing methods such as enhancement and morphological feature extraction from images. The classifiers demonstrated potential, indicating that AI can improve the automated classification accuracy of mango varieties [28].

The paper focuses on an innovative machine learning approach to enhance the mango cultivation management system. The study's objective is to include developing an automated mango grading system based on weight and color and predicting mango yield using historical data and weather patterns. The study aims to forecast the ideal water requirements for the various stages of mango growth and establish a system for early disease detection. The key methodologies involve transfer learning with VGG16 and ResNet50 for image tasks, SVM for grade protection, and linear regression for yield and water prediction. The study achieves notable results, with grading accuracy at 88%, yield prediction at 74.9%, and water level prediction showing potential [29].

The study focuses on an autonomous method for classifying and detecting mango diseases that combines histogram-oriented gradients (HOG) with convolutional neural networks (CNN). The shape and texture data is captured by HOG, while CNN is used for feature extraction in the system. By combining the best features of both approaches, the hybrid model improves performance and achieves 96.80% accuracy in illness identification and classification. The accuracy, precision, and recall measures are utilized to validate the model's performance. It provides a dependable and efficient method for managing diseases early on in mango farming [30].

The study investigated the problem of powdery mildew disease on mango leaves, which caused a considerable loss in production and quality. The paper provides a model that combines convolutional neural networks (CNN) and support vector machines (SVM) to categorize illnesses into four severity categories. The method involves data organization, CNN-based feature extraction, and SVM classification. Using a dataset of 2559 pictures, the hybrid model obtained an overall accuracy of 89.29% and a macro-average F1-score of 90.10. The model performs well, although it struggles in courses with lower support proportions [31].

Applications like mapping and autonomous yield estimation require the ability to recognize fruits in orchards. Conventional methods that depend on manually designed characteristics are vulnerable to changes in actual orchard settings. In contrast, deep learning-based one-stage object detection methods like YOLO offer superior detection accuracy but at the expense of higher computational complexity. The study introduces MangoYOLO5, an improved, faster, and more compact version of YOLOv5. It is designed for detecting mangoes in open orchards using the MangoNet-Semantic dataset. MangoYOLO5 incorporates several optimizations, including removing the feedback convolutional layer from the BottleneckCSP module and eliminating two convolutional layers from the focus module. These changes result in a lighter model that performs better than YOLOv5s in terms of precision, recall, and mean average precision (mAP), especially under challenging conditions like occlusion, distance, and lighting variations. Additionally, the reduced model complexity shortens the training time, enhancing its suitability for real-time applications [32].

Table 2.1 Research Synthesis Matrix

Related Studies	Specific Objectives	Methodology	Problem: Unknown/Gap	Remarks
Chourasiya et al., (2023)	Focuses on the challenges that farmers encounter in crop production specifically in crop diseases that leads to poor growth of the crops.	The use of Internet of Things (IoT) enabling better monitoring and visibility and wireless sensor network (WSN) to monitor diseases and gather environmental data.	IoT and machine learning technologies hold promise for addressing the issues, the integration for practical and user-friendly solutions for farmers is still limited.	Integrating IoT and machine learning offers a powerful solution for mango disease detection, providing farmers with timely, accurate information.
Polly and Devi (2022	To enhance the system reliability and efficiency to further research and explore the capabilities of deep learning in detecting crop disease.	Practices deep learning and image processing to identify and diagnose crop diseases.	Current agricultural practices for disease detection are manual and inefficient, often leading to delayed diagnosis and increased crop losses.	Traditional methods are slow and inefficient, leading to crop losses. Leveraging IoT and deep learning can provide farmers with real-time, accurate disease detection.

Giri et al., (2023)	The use of machine learning in agriculture, specifically for identifying two diseases in mango: Powdery Mildew and Sooty Mould.	The MangoLeafBD dataset and Gradient Boosting Classifier to enhance mean shift image segmentation and hu moments for feature extraction.	The machine learning advancements show promise for mango disease identification, practical, user-friendly solutions for farmers are still missing.	The study shows that machine learning, using a Gradient Boosting Classifier with advanced techniques, can classify mango leaf diseases. Despite moderate success, it promises reliable disease identification and advances precision agriculture by improving diagnosis and supporting sustainable farming.
Kumar et al. (2021)	Focuses on a common fungal disease that is affecting the growth of mango trees in India and it would be convenient for mango cultivators to identify the disease beforehand and apply treatment.	Deep learning convolutional neural network (CNN) architecture to identify disease on a mango. a real-time dataset captured in farms of Karnataka, Maharashtra and New Delhi.	Despite the importance of mangoes in India, current methods for diagnosing fungal disease on mango are inadequate for timely and accurate identification.	The system that can be used for disease on mango leaves a higher accuracy but it needs to develop a model like an IoT application, like a camera.

Delgado et al. (2019)	Focuses on a common fungal disease that is affecting the growth of mango trees in India and it would be convenient for mango cultivators to identify the disease beforehand and apply treatment.	The use of Artificial Intelligence (AI) and K-nearest neighbor (KNN) to classify the mango variety in the Philippines, specifically the indian, carabao and saperada mango.	Accurately classifying mango varieties by physical characteristics is challenging and often requires agricultural expertise, making it impractical for widespread use. Current methods rely on subjective and uncertain data that limit their effectiveness.	Classifying plants visually is difficult and requires expertise. This study shows that using fuzzy logic and K-NN classifiers for mango leaf identification is promising. These methods offer potential for fast, efficient mango variety identification, though further improvements are needed.
Withanaarachchi et al (2023)	The study focuses on an innovative approach using machine learning to revolutionize the mango cultivation process. It aims to enhance productivity, quality, and disease detection management.	It utilizes transfer learning with VCG16 and ResNet50 for image-based tasks, SVM for grade prediction, and linear regression for yield prediction.3	While the research showcases the potential of machine learning to enhance mango cultivation, practical implementation remains a challenge. The early disease detection models show good results but need further refinement in terms of accuracy and reliability.	The study shows the potential of machine learning to improve mango cultivation through automated grading, yield prediction, water forecasting, and early disease detection. While promising, further refinement is needed to enhance accuracy and develop user-friendly

				applications for farmers.
Sema et al.(2023)	To develop a hybrid automatic classification system for mango diseases.	Utilizing the Convolutional Neural Network (CNN) and Histogram Oriented Gradients (HOG) techniques.	Existing methods for mango disease detection have made progress, it often relies on single techniques that may not capture the complexity of disease symptoms.	The study investigates a hybrid approach to mango disease detection, integrating CNN and HOG techniques to enhance accuracy and reliability. By overcoming the challenges the research advances agricultural practices, providing a more effective solution for early disease identification and management.
Banerjee et al.(2020)	Aims to develop a system that can detect and classify mango leaf powdery mildew disease.	It uses the Convolutional Neural Networks (CNN) and Support Vector Machines (SVM).	Current methods for detecting mango lead are limited in accuracy and reliability.	Addresses the limitations of current mango leaf detection method by CNN-SVM model.
Chandana et al.(2023)	Refining the model by removing unnecessary convolutional	Develop and optimize using MangoYOLO5, enhanced version of	Current methods for fruit detection in orchards often rely on traditional	By optimizing YOLOv5s with targeted improvements to reduce

	layers and modules to reduce computational complexity and training time while improving detection accuracy.	YOLO 5vs and MangoNet-Semantic dataset.	hand-crafted features or deep learning models designed for general object detection, such as YOLOv5s.	computational complexity and training time while improving accuracy, MangoYOLO5 addresses critical challenges such as occlusion and varying environmental conditions.

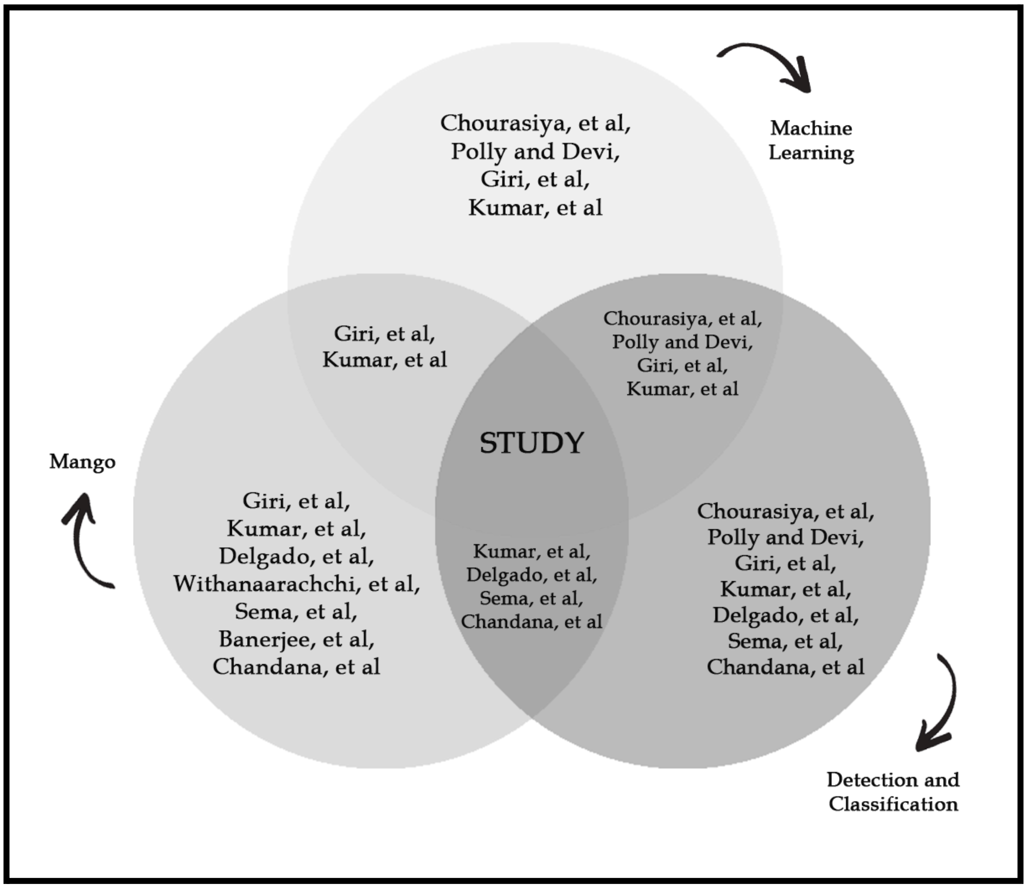


Figure 2.1 Summary of Related Studies

Chapter 3

RESEARCH METHOD

This section discusses the materials needed and the process of developing and deploying the system. In addition, it presents different visual diagrams illustrating how the system will function.

3. 1 Analysis Modeling

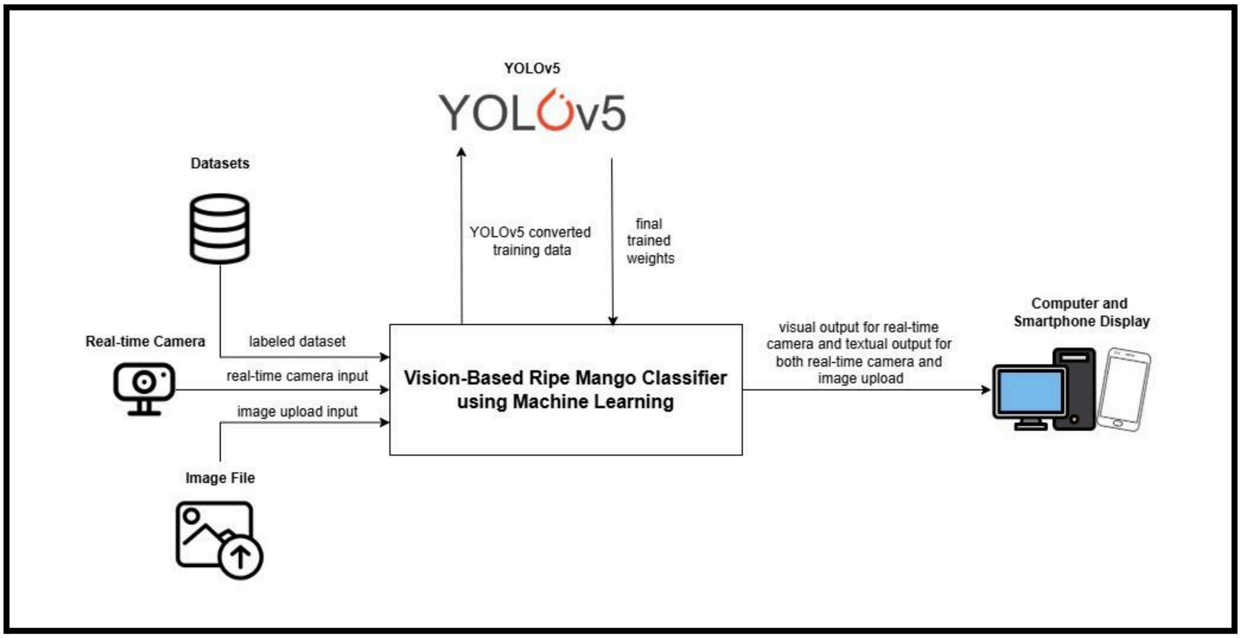


Figure 3.1 Conceptual Framework

Figure 3.1 shows the conceptual framework or the blueprint that will be used in the development of the system. It is a high level diagram that shows the way the system works in the most simple way. In essence, the user will load up the website. From there, they will be able to immediately use the service that the system provides and select which input type they want to use between real-time camera or image upload input. The input will then undergo training using the

YOLOv5 model which has been trained beforehand with the custom dataset. Lastly, the result will then be displayed on the website either through textual display for image upload inputs, and both visual and textual display for real-time camera.

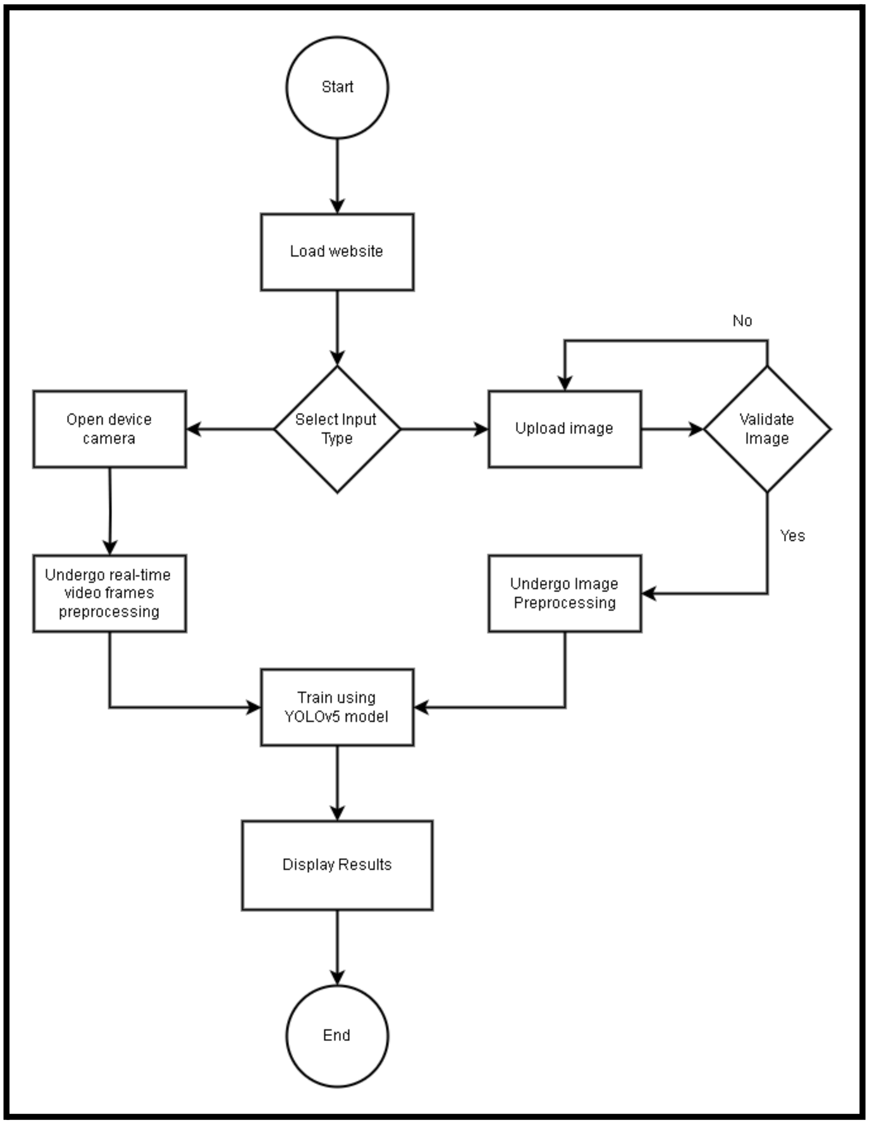


Figure 3.2 Process Flow Diagram

The process flow diagram, figure 3.2, shows a detailed visualization of the mango disease detection system. Showing the step-by-step workflow, starting from the user interface which is the website. It then accepts the mango image input by which the user can choose between a live camera or uploading an image of the mango. Both inputs will then undergo preprocessing to prepare for the YOLOv5 model. Results will then be displayed in the user interface to the user. By following this structured process, the system ensures accurate and efficient detection, providing valuable information to users.

3. 2 System Design

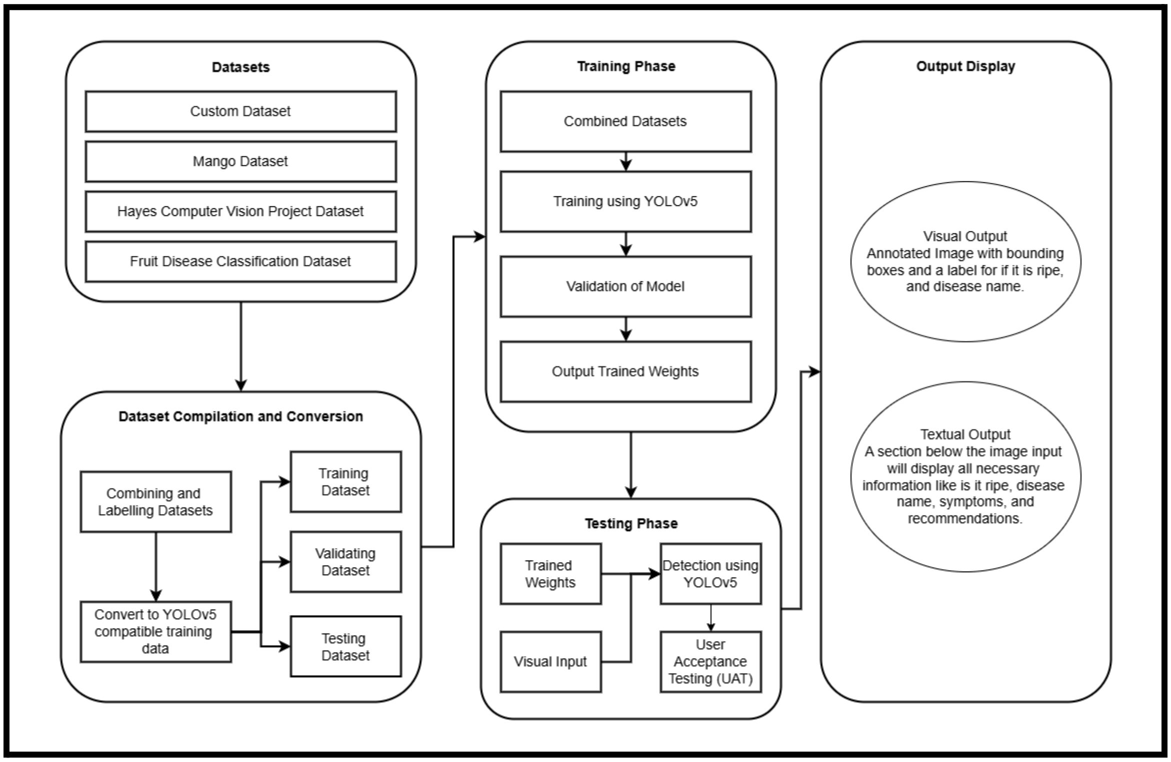


Figure 3.3 System Design

The system design serves as a more detailed blueprint for developing and implementing the system.

3.2.1 Datasets

The training datasets used in this study comprises of:

- **Custom Dataset:** a mango dataset gathered by the researchers for the purpose of this study
- **Mango Dataset:** a mango dataset with ripe and unripe classes [33]
- **Hayes Computer Vision Project Dataset:** a mango dataset with mango and not-mango classes [34]
- **Fruit Disease Classification Dataset:** mango with diseases dataset with Anthracnose, black mould rot, and stem end rot classes [35]

These datasets will serve as the basis for training the machine learning models, in this case the YOLOv5 model.

3.2.2 Dataset Compilation and Conversion

In order to ready the datasets for training, there are a few steps to take into account:

- **Labeling Datasets:** put any image with a given disease in a separate folder with the correct disease category and the name of the disease will be its label.
- **Convert to YOLOv5 Compatible Training Data:** this data set then gets converted into a better compatible data set for the model. This involves removing background from each image to ensure accuracy. Which is going to be used as a dataset for training, validation and testing for the YOLOv5 model.

3.2.3 Training Phase

The execution in the training phase includes:

- **Combined Datasets:** use compiled and converted dataset.
- **Trained using YOLOv5:** train the YOLOv5 model on combined dataset. Here the model learns to identify the patterns and features in each mango disease
- **Validation of Model:** the model's performance is validated using the validation dataset to ensure accurate results
- **Output Trained Weights:** after training, the model's trained weights are generated. These weights are the parameters that the model use to make predictions with
- **User Acceptance Testing:** this is where the intended user will test the system if it's working properly and providing the correct outputs, and give their feedback

3.2.4 Testing Phase

The testing phase involves:

- **Trained Weights:** the trained weights represent the knowledge the model has gained about the appearance of the mango, and is loaded into the model to analyze the image and make accurate results
- **Visual Input:** input from the user, either from real-time camera or uploaded photos

- **Detection using YOLOv5:** with the trained weights, the model detects diseases from the input

3.2.5 Output Display

The output display presents the results to the user:

- **Visual Output:** in real-time camera inputs, mangoes are annotated with bounding boxes with a label for if it is ripe, and disease name if there are any
- **Textual Output:** for both real-time camera and uploaded photo, a section below the input will display the necessary information like: disease name, characteristic features, and recommendations

The system design ensures a proper systematic way of how the detection system works and its approach.

3.3 Datasets

The dataset used by the researchers is a mix of custom dataset gathered by the researchers for the purpose of this study and available datasets online for a more balanced dataset. The researchers will be collecting ripe, unripe, healthy, and unhealthy mangoes.

3. 4 Training Procedure

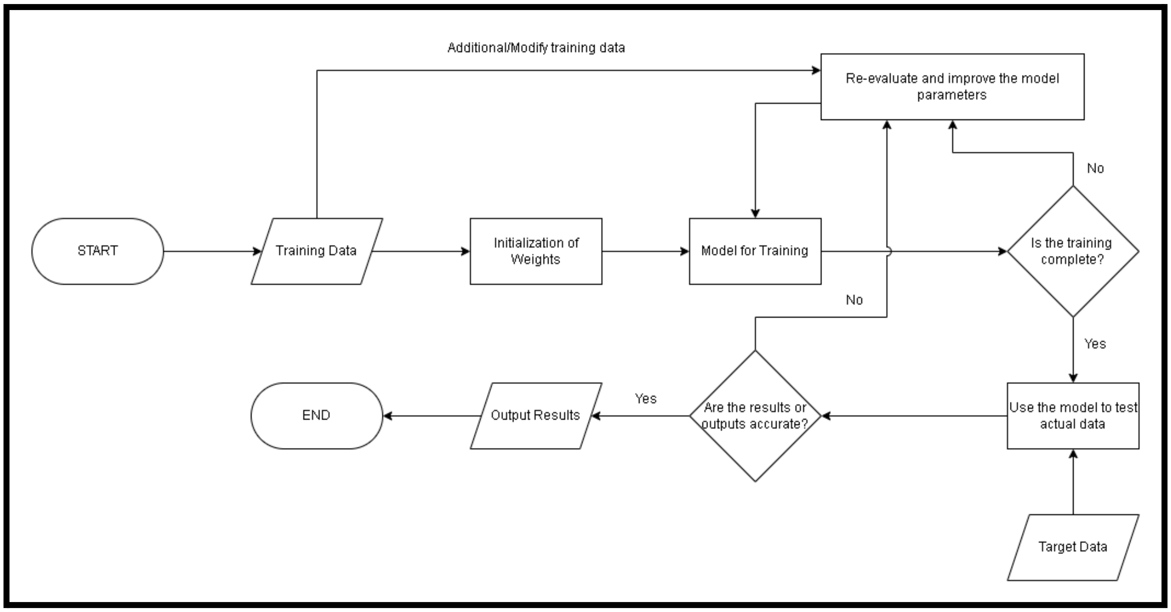


Figure 3.7 Flowchart for Training Procedure

Figure 3.7 shows the training process of the model. Firstly, the training data consists of the images of mangoes with diseases and file names with the corresponding disease. Then, YOLOv5 already has pre-trained weights so this is initialized. The initialized model is then trained using the training data. If the results are not satisfactory, the researchers will modify the training data and re-evaluate the model's performance. If the results are satisfactory, it can then be used to train with real-world data or images of mangoes that the model has not seen, and if the results are accurate, then the model is good to go.

3.5 Performance Metrics

To evaluate the performance of our Vision-Based Ripe Mango Classifier system, the researchers employed the following metrics:

Metric	Description
Accuracy	Proportion of correctly classified samples (ripe and diseased) out of the total samples.
Error Rate	Proportion of incorrectly classified samples out of the total samples.
Precision	Proportion of true positive predictions (correctly identified ripe or diseased) out of all positive predictions.

Table 3.1 Performance Metrics and Description

3.6 Equations Used

The researchers used the following formulas in calculating some of the results for this study. The formulas cover the accuracy of ripe mango and disease detection.

To calculate the prediction accuracy of the detection model, the researchers used the following equation:

$$Accuracy = \frac{CorrectDetections}{n} x\ 100 \quad (3.1)$$

For the error rate of the detection model, the researchers used the following equation:

$$ErrorRate = \frac{IncoorectDetections}{n} x\ 100 \quad (3.2)$$

Calculating the precision of detecting each classifier for all locations, the researchers used the following equations:

$$\text{Precision} = \frac{CorrectDetections}{CorrectDetections + IncorrectDetections}\ x\ 100 \quad (3.3)$$

3. 6 Hardware Specifications

This section talks about the hardware specifications required for the development and deployment, and as well as the user-end requirements of the mango disease detection system.

3.6.1 Development Hardware

Minimum Computer Specifications: Processor: Intel i5 or AMD Ryzen 5 (or higher); RAM: Minimum 8GB, recommended 16GB; GPU: NVIDIA GTX 1050Ti or higher; Storage: Minimum 256GB

Internet Connection: Stable internet connection, recommended 10 mbps or higher

Camera: Webcam or smartphone camera, recommended 12 MP or higher

3.6.2 Deployment Hardware

Smartphone: Android or iOS device with minimum 4 GB RAM and a high-resolution camera (12 MP or higher)

Laptop/Desktop: Minimum Intel i3 or better, 4GB RAM, with a webcam

Internet Connection: Stable internet connection, recommended 10 mbps or higher

3. 7 Software Specifications

This section talks about the software specifications required for the development and deployment, and as well as the user-end requirements of the mango disease detection system.

3.7.1 Development Software

Operating System: Linux (Ubuntu 20.04 or later recommended), Windows 10 or higher, or macOS

Development Environment: Python 3.10 or later

3.7.2 Deployment Software

Smartphone OS: Android 8.0 or later, iOS 11 or later

Laptop/Desktop OS: Windows 10 or later, macOS Mojave or later, any recent Linux distribution

Web Browser: Latest version of any browser available

Chapter 4

RESULTS AND DISCUSSION

This chapter covers the testing results and discussions including the graphical user interface of the developed system. Each part of the system will be discussed thoroughly in this chapter to have a clearer understanding of the system and the functions that it offers.

Description of Prototype

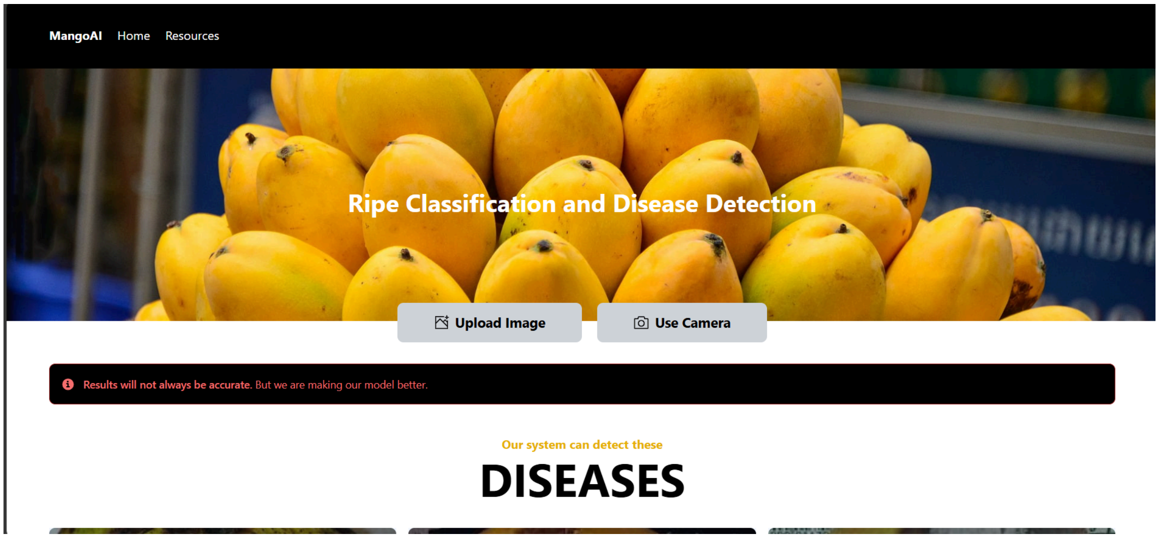


Figure 4.1 Landing Page

Figure 4.1 shows the landing page where the users can choose their inputs for detection and classification, and also where the disease card to show which diseases are available for detection.

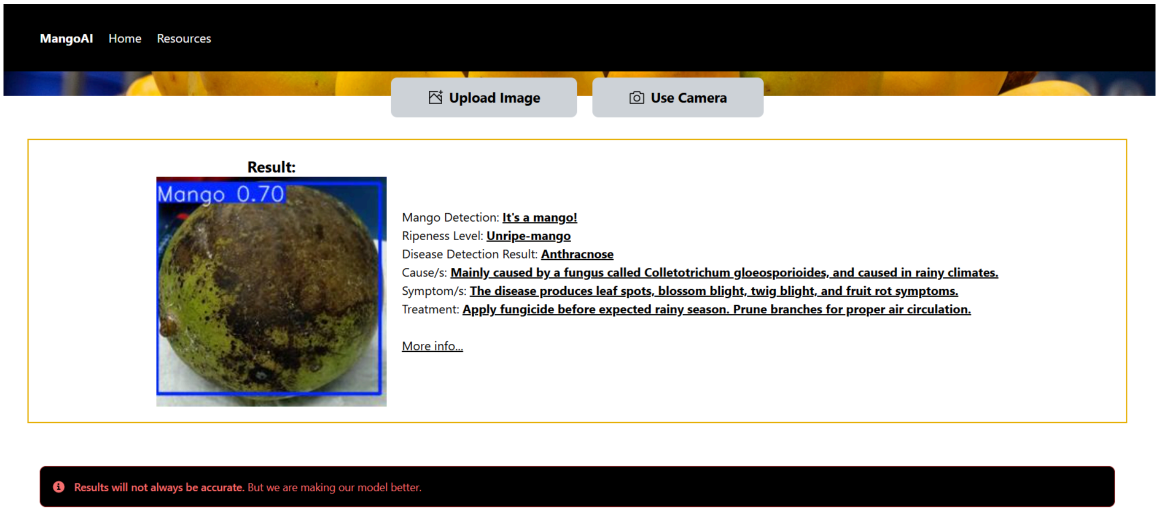


Figure 4.2 Upload Image Section

The upload image section, figure 4.2, shows the result of the mango ripeness classification and disease detection when a user uploads an image as their input. It will also display a container where the results will be displayed when a user uploads their image input and the model has finished scanning

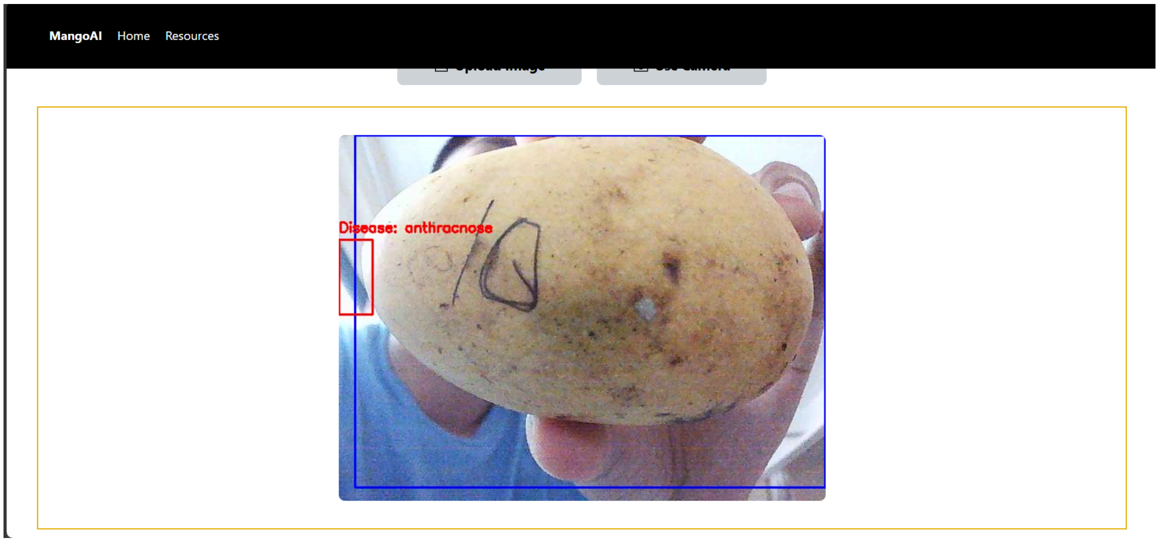


Figure 4.3 Live Camera Input Section

The live camera input section, Figure 4.3, shows up when the user clicks the open camera button. It shows your camera that is integrated with our model for classifying and detecting, and shows a bounding box when a mango is in view of the camera.

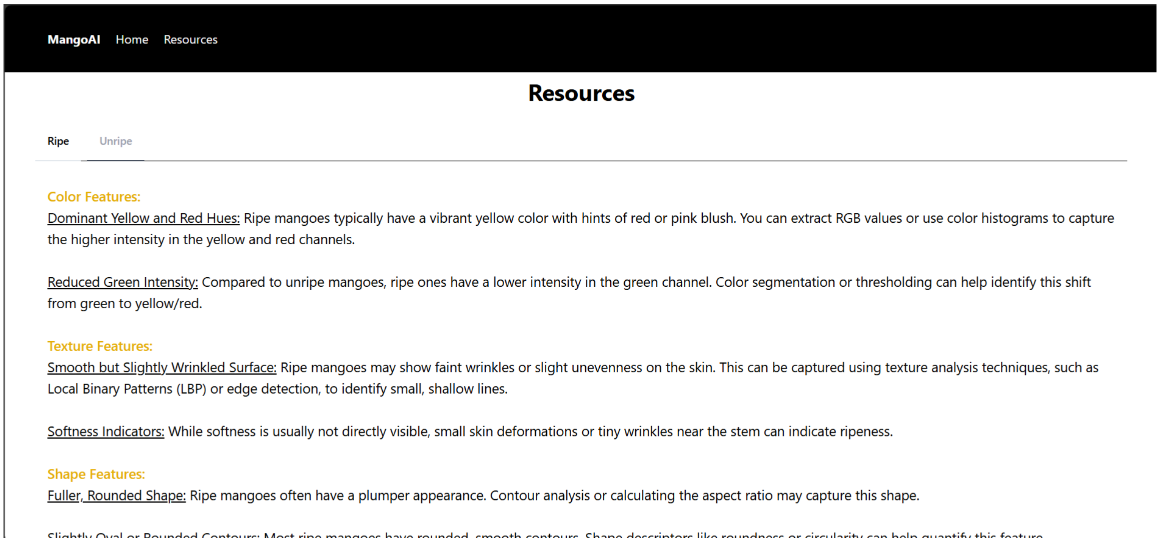


Figure 4.4 Resources Page

The figure 4.4 shows the resources page which shows information about ripe and unripe mangoes, and the 3 diseases: Anthracnose, Black-mould rot, and Stem-end rot, information.

System Testing

The system testing shows the outcomes of testing the system's trained model and how well it performs detecting mango, its ripeness, and disease if there are any.

Table 1.
Results for Image Upload Input

Test No.	Time (seconds, milliseconds)	Mango Detection	Ripeness Classification	Disease Detection (if any)
1	1.96	Y	N	Y
2	2.2	Y	Y	Y
3	1.78	Y	Y	Y
4	1.33	N	NA	NA
5	1.48	N	NA	NA
6	1.28	N	NA	NA
7	2.04	Y	Y	Y
8	1.84	Y	N	Y
9	1.58	Y	Y	Y
10	1.75	Y	Y	Y
11	1.71	Y	N	N
12	2.28	Y	N	Y
13	2.16	Y	N	Y
14	1.89	Y	N	Y
15	1.98	Y	Y	N
16	1.78	Y	Y	Y
17	2.13	Y	N	Y
18	1.94	Y	Y	Y
19	2.11	Y	Y	N
20	1.98	Y	Y	N
21	2.11	Y	Y	Y
22	1.56	Y	N	Y

Table 1. *(Cont.)*
Results for Image Upload Input

Test No.	Time (seconds, milliseconds)	Mango Detection	Ripeness Classification	Disease Detection (if any)
23	2.04	Y	N	Y
24	1.45	Y	N	Y
25	1.5	N	NA	NA
26	1.99	Y	Y	Y
27	1.94	Y	N	Y
28	2.56	Y	N	Y
29	1.5	N	NA	NA
30	1.88	Y	Y	Y

Average time: 1.86 seconds

Table 1 shows the test results of image upload inputs. It shows the time, and the results of the machine learning model. For context, Y means yes it got the correct result, N means no it didn't get the correct result, and NA is not applicable for those images that weren't detected as a mango because the system will not go on with the ripeness and disease model if it does not detect a mango.

With an average time of 1.86 seconds, it already has great processing time especially considering that the system handles mango detection, ripeness classification, and disease detection. It has 3 models but it still achieved a sub-2-second average response time.

Mango detection results

$$Accuracy_{IU} = \frac{25}{30} x\ 100 = 83.33\%$$

$Error\ Rate_{IU} = \frac{5}{30} x\ 100 = 16.67\%$

$Precision_{IU} = \frac{25}{25+5} x\ 100 = 83.33\%$

For mango detection, an 83.33% result shows a promising score, but it still has false positives and false negatives which can still be improved to achieve a more accurate detection.

Ripeness classification results

$Accuracy_{IU} = \frac{13}{25} x\ 100 = 52.00\%$

$Error\ Rate_{IU} = \frac{12}{25} x\ 100 = 48.00\%$

$Precision_{IU} = \frac{13}{13+12} x\ 100 = 52.00\%$

For ripeness classification with file upload input, the system got 13 out of 25 correct classifications. With the error rate just closely behind the accuracy rate, this means that the ripeness classification model still needs a lot to improve on.

Disease detection results

$Accuracy_{IU} = \frac{21}{25} x\ 100 = 84.00\%$

$Error\ Rate_{IU} = \frac{4}{25} x\ 100 = 16.00\%$

$Precision_{IU} = \frac{21}{21+4} x\ 100 = 84.00\%$

For disease detection, 84.00% is a good result. It only got 4 mistakes but it can still be improved by providing more high quality images, and images that show the physical characteristics of the diseased mangoes to improve the accuracy of the model.

Below are some of the tests done for image upload testing.

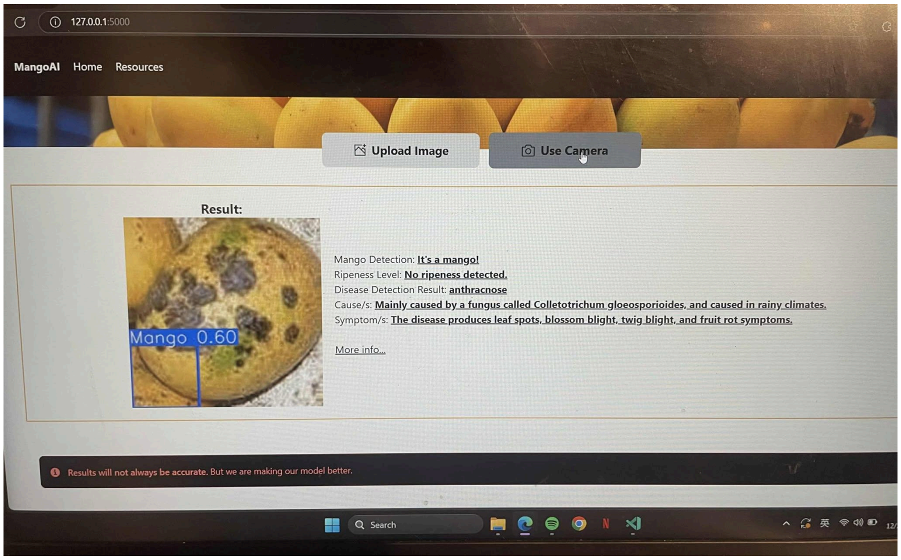


Figure 4.5 Ripe Anthracnose IU Test

Figure 4.5 shows a ripe mango with anthracnose. As seen, it detected a mango and with anthracnose which is right. However, it didn't detect the ripeness which is unusual as the color of the mango is clearly visible, but maybe it is because of the quality of the image that added to the factor.

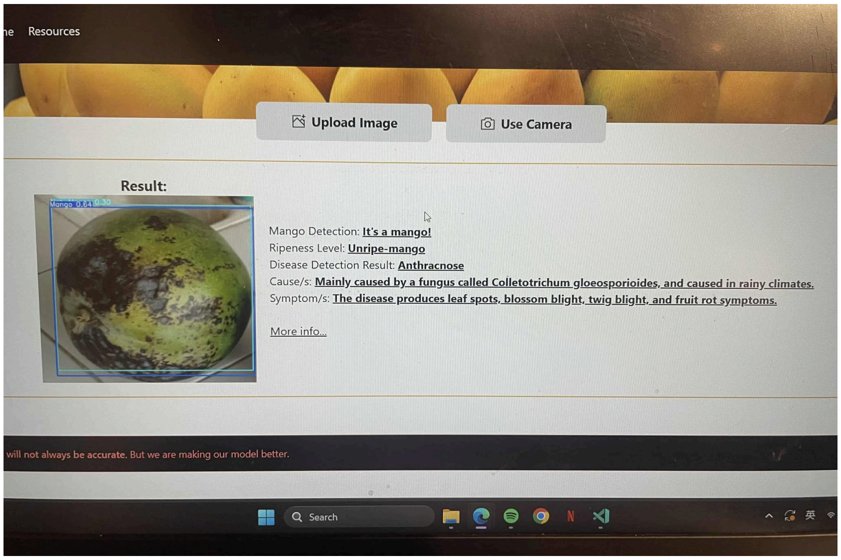


Figure 4.6 Unripe Anthracnose IU Test

Figure 4.6 test image is an unripe mango with anthracnose. As seen in the image, the result is correct identifying it as a mango that is unripe and with anthracnose.

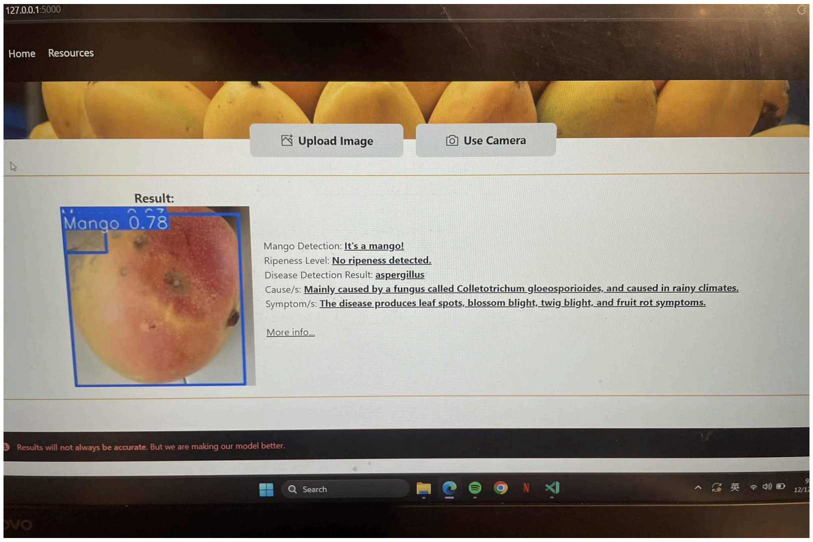


Figure 4.7 Ripe Stem-end Rot IU Test

Figure 4.7 testing is a ripe mango with lasiodiplodia or stem-end rot. The model didn't identify the ripeness, which could be due to the color which is not green nor yellow. But other than that, it got the disease correctly.

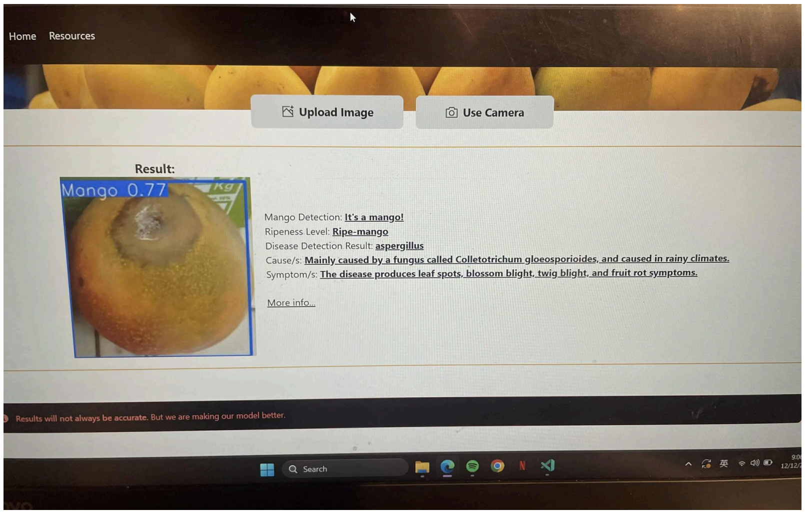


Figure 4.8 Ripe Stem-end Rot 2 IU Test

Figure 4.8 testing is a ripe mango with lasiodiplodia too. The results correctly identified all this even with slight physical appearance change which is surprising but good.

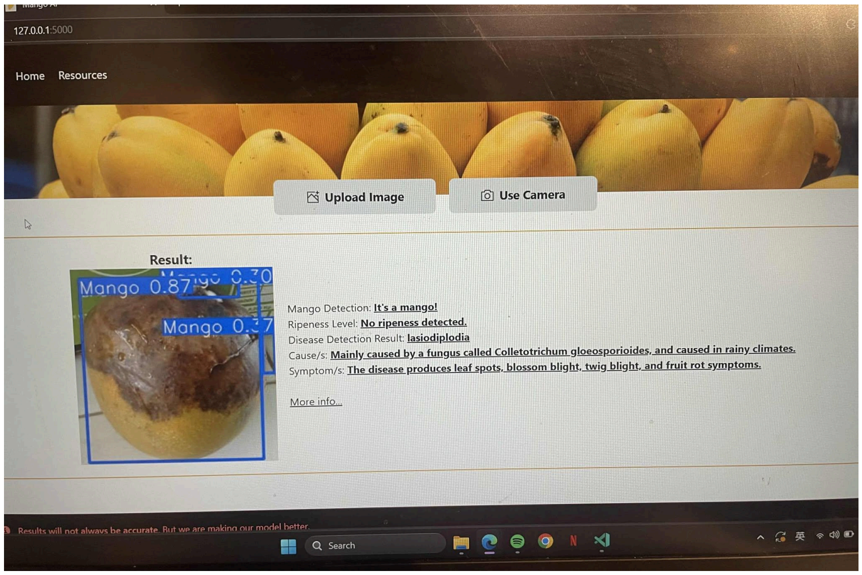


Figure 4.9 Ripe Black-mould Rot IU Test

Figure 4.9 testing is a ripe mango with aspergillus or black-mould rot. The model only didn't identify the ripeness, which could be due to the black spot covering the color of the mango.

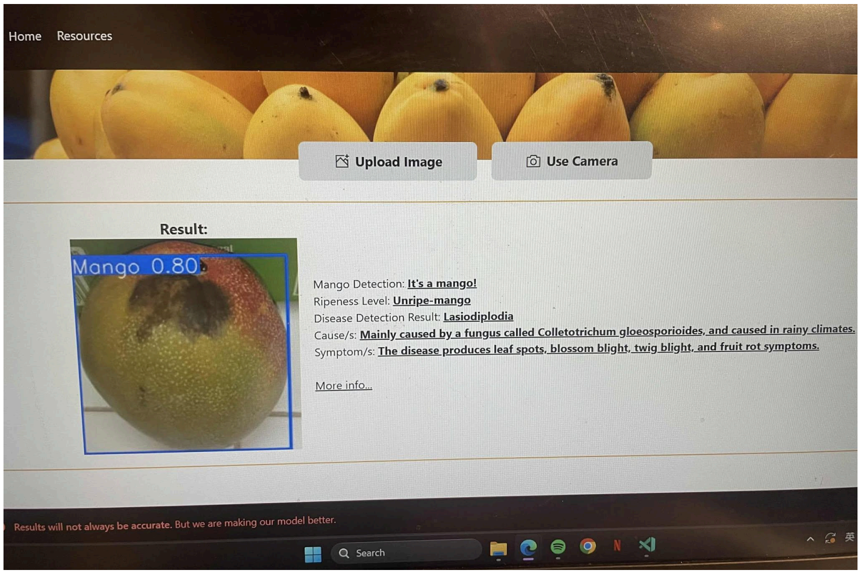


Figure 4.10 Unripe Black-mould Rot IU Test

Figure 4.10 shows an unripe mango with anthracnose. It detected everything right - mango, ripeness, and disease. The ripeness of the mango was identified because the black spot is not that big which shows the color of the mango.

Table 2.
Results for Live Camera Input

Test No.	Mango Detection	Ripeness Classification
1	Y	N
2	Y	Y
3	Y	Y
4	Y	Y
5	Y	Y
6	Y	Y
7	Y	Y
8	Y	Y
9	Y	Y

Table 2. *(Cont.)*
Results for Live Camera Input

Test No.	Mango Detection	Ripeness Classification
10	Y	Y
11	Y	Y
12	Y	Y
13	Y	Y
14	Y	Y
15	Y	Y
16	Y	Y
17	Y	Y
18	Y	Y
19	Y	Y
20	Y	Y
21	Y	Y
22	Y	Y
23	Y	Y
24	Y	Y
25	Y	Y
26	Y	Y
27	Y	Y
28	Y	Y
29	Y	Y
30	Y	Y

Table 2 shows the result of live camera input testing. It shows the number of tests, result of mango detection, and ripeness classification. Disease detection is not part of the table as the mangoes available for purchase do not have any diseases as told by the seller and as confirmed by opening and tasting the mangoes by the researchers. Also, test numbers 1-15 are unripe mangoes, and 15-30 are ripe mangoes.

Mango detection results

$Accuracy_{LC} = \frac{30}{30} x\ 100 = 100.00\%$

$Error\ Rate_{LC} = \frac{0}{30} x\ 100 = 0.00\%$

$Precision_{LC} = \frac{30}{30+0} x\ 100 = 100.00\%$

Testing using the camera showed promising results, with mango detection garnering a 100% accuracy. It means that the model is able to successfully detect the fruit well.

Ripeness classification results

$Accuracy_{LC} = \frac{29}{30} x\ 100 = 96.67\%$

$Error\ Rate_{LC} = \frac{1}{30} x\ 100 = 3.33\%$

$Precision_{LC} = \frac{29}{29+1} x\ 100 = 96.67\%$

For ripeness classification, getting 96.67% accuracy is also promising. Despite some unripe mangoes starting to lose its green color, the model was still able to classify it as unripe which shows good sign for the model. Though, the model still needs refinement and better dataset in order to better the results. But, classifying the ripeness of a mango has a lot of factors which are complex and will not always be perfect.

The images below show the result of the live camera testing and its equivalent heatmap. For the test result, a blue bounding box means the system identified the subject as a mango; a green bounding box is the ripeness classification (ripe or unripe), and; a red bounding box is for the disease if there are any.

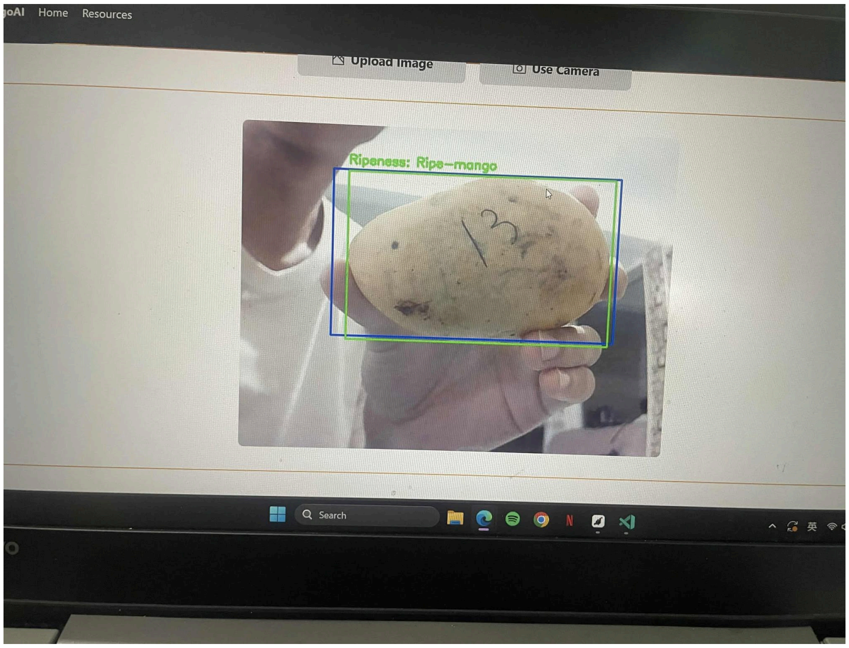


Figure 4.11 Ripe Mango No Disease LC Test

As seen in Figure 4.11, a ripe mango with no disease has been tested. The system got both of it correctly and not showing any red bounding box for the disease since there is no disease on the mango. And as seen in Figure 4.12 below, the heatmap is not working as it should be for all testing, showing a tint of purple for all the tests.

Figure 4.12 Ripe Mango No Disease LC Heatmap

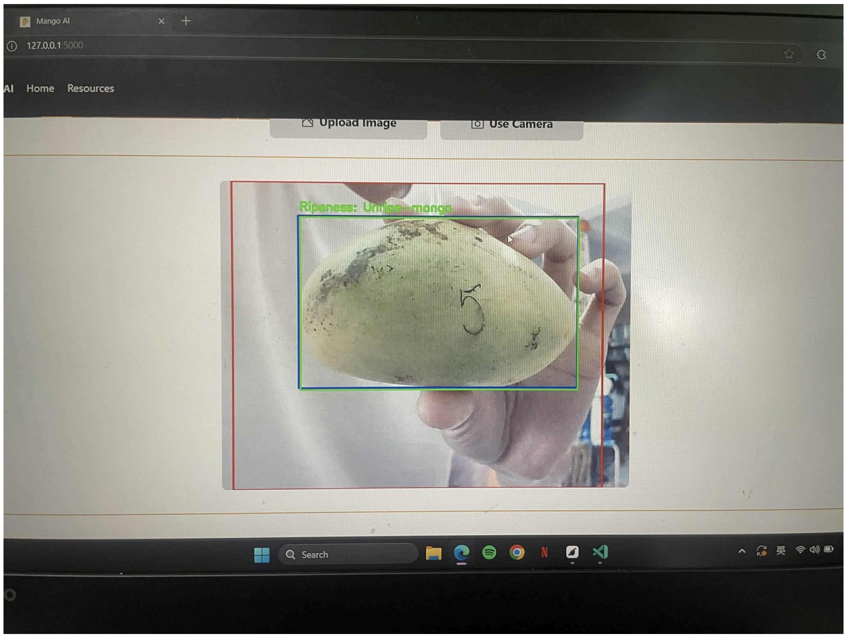


Figure 4.13 Unripe Mango No Disease LC Test

Figure 4.13 test is an unripe mango with no disease. Even with the test mango losing its green color as it is starting to become ripe, the model still gave the correct result. Good lighting is key especially in this context when the yellow color is starting to bloom. Also, a red bounding box can be seen, though, this is an abnormal reaction from the disease model.

Figure 4.14 Unripe Mango No Disease LC Heatmap

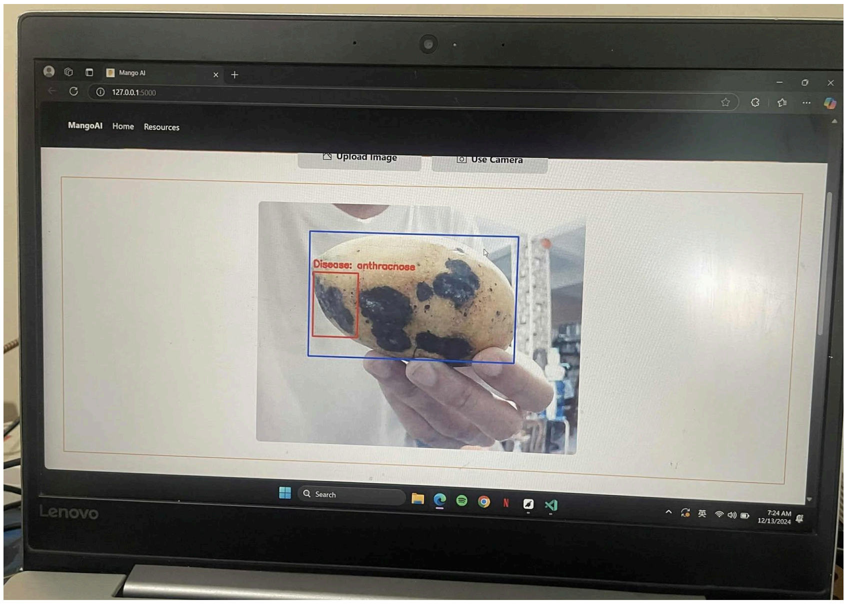


Figure 4.15 Ripe Mango with Disease LC Test

Figure 4.15 testing is a ripe mango that's starting to rot as it has been a few days since obtaining the fruit. As seen, it has detected a mango with the blue bounding box, and it also has a red bounding box saying anthracnose. But it didn't get the ripeness red bounding box which could be the cause of the black spots, or it needs to have better lighting. Furthermore, as the researchers did the user testing and tested the same mango, the expert said that this is indeed affected with anthracnose.

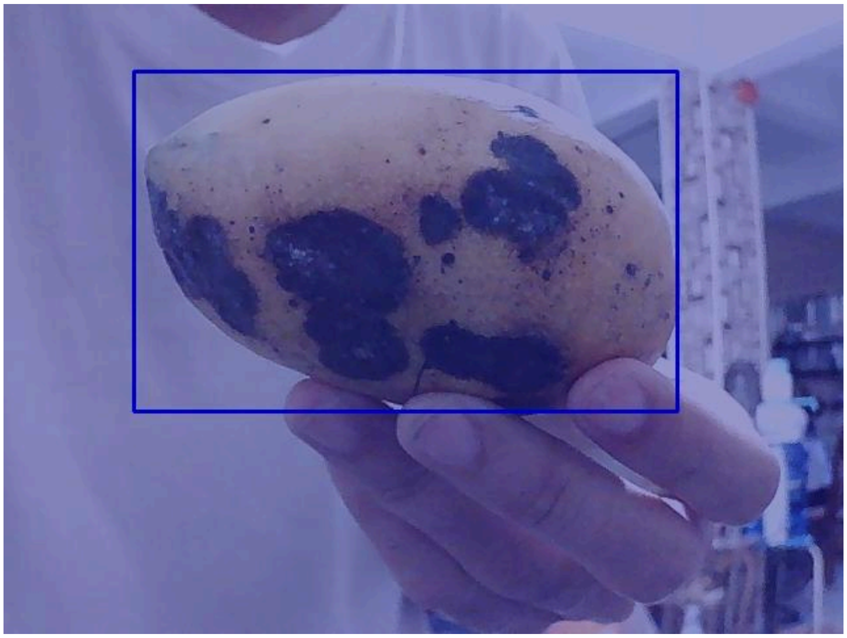


Figure 4.16 Ripe Mango with Disease LC Heatmap

Figure 4.17 User Testing 1

Figures 4.17 (above) and 4.18 (below) show the researchers testing the system with the expert on mangoes. Tests are done on mangoes that are ripe, unripe, and with disease. Before every test, the researchers ask the expert if the mango is ripe or unripe, and if it has diseases, and then test it afterwards. The results are what the researchers expected based on what the experts say, achieving an almost perfect accuracy. Mango model got every mango correctly. Ripeness classification got almost every test correct with the model having a hard time detecting the ripeness of the mango with a disease. And the disease model only identified the mango with anthracnose with disease.

Figure 4.18 User Testing 2

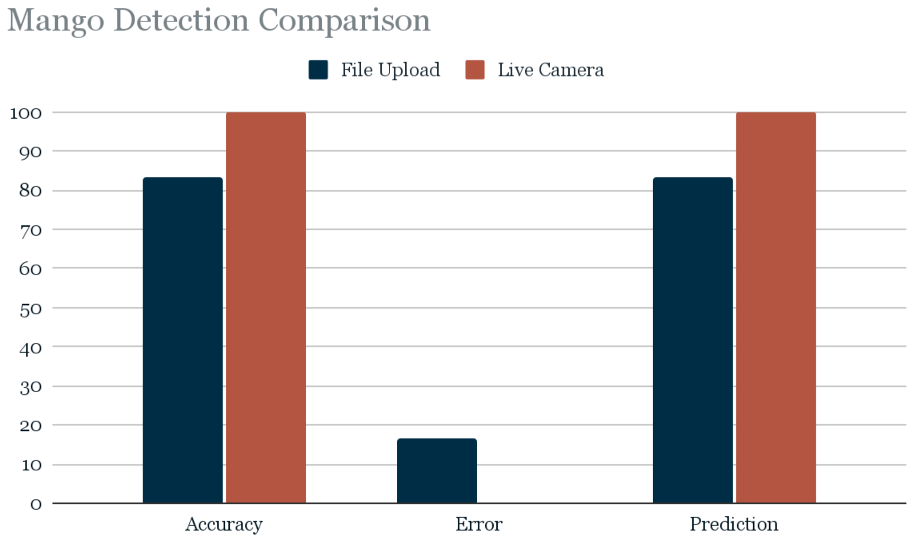


Figure 4.19 Mango Detection Result Comparison

As shown in figure 4.17, the live camera input shows better results in detecting mangoes than the file upload input but with just a small deficit.

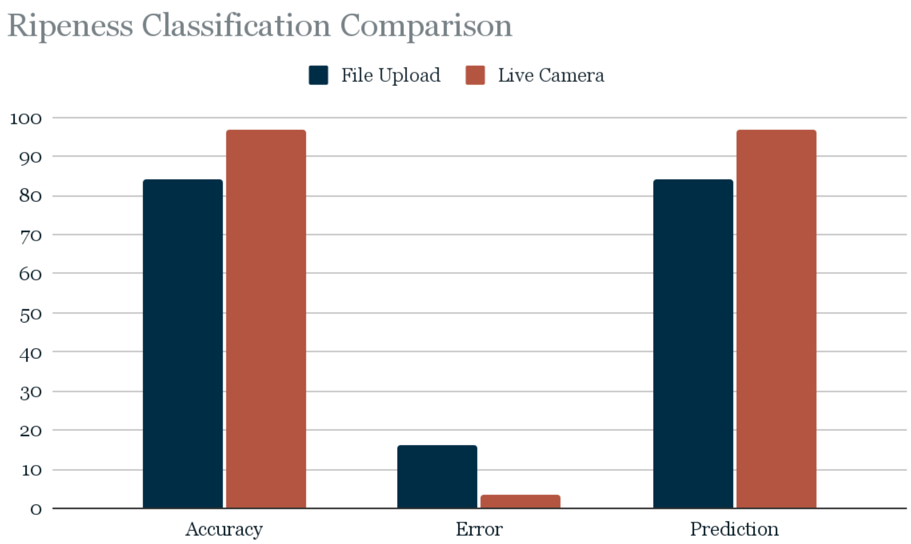


Figure 4.20 Ripeness Classification Result Comparison

As shown in figure 4.18, the live camera input also shows better results in ripeness classification.

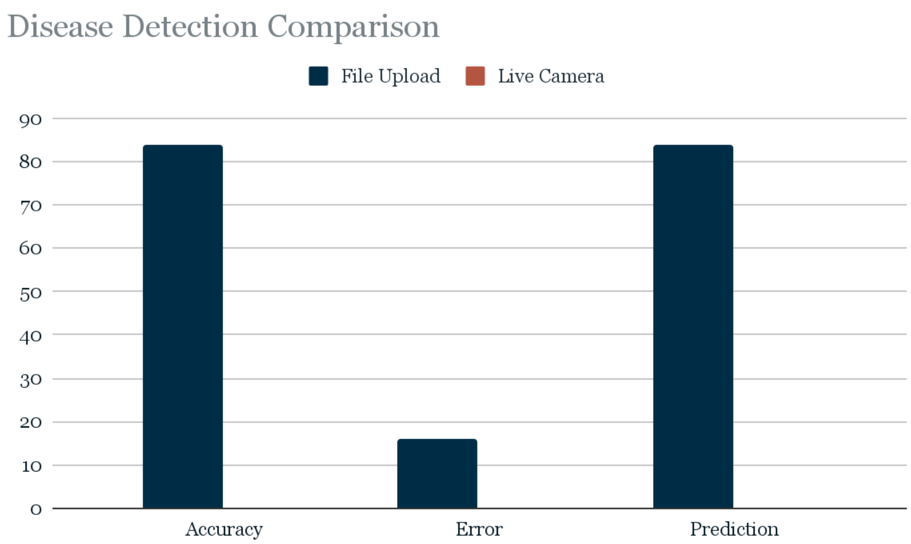


Figure 4.21 Disease Detection Result Comparison

As seen in figure 4.19, disease detection is not counted in the live camera input, but it shows that the file upload input has a high accuracy and prediction outcome in detecting the 3 diseases in our system.

Based on Equation (3.1), the mango detection accuracy for the 2 inputs - image upload and live camera, is as follows:

$$Accuracy_{IU} = \frac{25}{30} x\ 100 = 83.33\%$$

$$Accuracy_{LC} = \frac{30}{30} x\ 100 = 100.00\%$$

Average Mango Detection Accuracy 91.67%

As with above, the ripeness classification accuracy is as follows:

$$Accuracy_{IU} = \frac{13}{25} x\ 100 = 52.00\%$$

$$Accuracy_{LC} = \frac{29}{30} x\ 100 = 96.67\%$$

Average Ripeness Classification Accuracy 74.34%

Based on Equation (3.2), the error rate for each is as follows:

$$Error\ Rate_{IU} = \frac{5}{30} x\ 100 = 16.67\%$$

$$Error\ Rate_{LC} = \frac{0}{30} x\ 100 = 0.00\%$$

Average Mango Detection Error Rate 8.34%

As with above, the error rate for ripeness classification is as follows:

$$Error\ Rate_{IU} = \frac{12}{25} x\ 100 = 48.00\%$$

$$Error\ Rate_{LC} = \frac{1}{30} x\ 100 = 3.33\%$$

Average Ripeness Classification Error Rate 74.34%

Calculating the precision based on Equation (3.3), the precision rate for each input is as follows::

$$Precision_{IU} = \frac{25}{25+5} x\ 100 = 83.33\%$$

$$Precision_{LC} = \frac{30}{30+0} x\ 100 = 100.00\%$$

Average Mango Detection Precision 91.67%

As with above, the ripeness classification precision is as follows:

$$Precision_{IU} = \frac{13}{13+12} x\ 100 = 52.00\%$$

$$Precision_{LC} = \frac{29}{29+1} x\ 100 = 96.67\%$$

Average Ripeness Classification Precision 74.34%

Calculating the results for the testings done on the two inputs, the researchers have found that the mango model shows the best result with a precision of 91.67% for both the inputs. This means that the mango model is trained well and would identify a mango most of the time. For the ripeness classification, there are a lot of factors that need to be considered which is why the precision is much lower. Things like lighting, physical appearance, color, and image resolution needs to be considered for the model to correctly and surely classify the ripeness of the mango. And the disease model, though getting >90% precision with the file upload, still needs a lot to look into as diseases also have a lot of factors to be considered especially since the 3 diseases mentioned in this system have quite similar physical appearance which is the main basis of the model in identifying the disease.

Throughout the development of this project, the researchers have faced several challenges from learning how to train the models, dataset collection, constant training of the models, and testing. Identifying and addressing these challenges was essential to ensure the reliability and functionality of the system.

Understanding the complexities of machine learning model training was one of the first difficulties faced. It took a lot of time to learn how to train using python and the YOLOv5 model as this was a first for the researchers.

The dataset used in this project played a critical role in the system's performance. However, finding a suitable dataset that met the requirements for mango detection, ripeness classification, and disease identification proved difficult. The researchers faced issues such as:

- **Data Scarcity:** Limited availability of mangoes due to offseason especially for mangoes with diseases
- **Data Quality:** Existing datasets often contained mislabeled or low-resolution images, which affected training.
- **Balancing Classes:** Ensuring the dataset had an even distribution of samples for each class to prevent model bias.

  To address these issues, the researchers combined multiple datasets and manually annotated or cleaned data when necessary. This process was labor-intensive but critical for achieving reliable results.

Furthermore, the researchers have identified that diseases that distinctively change the physical appearance of the fruit can affect the ripeness models ability to correctly classify the ripeness of the fruit, hence, giving a “no ripeness detected” result. All 3 diseases have a chance of hugely affecting the physical appearance of the mango, showing big black spots, which is why tests with big black spots on the mango have a “no ripeness detected“ result.

Testing the disease model was also not enough, only 1-2 mangoes of the 30 mangoes bought got infected after a few days of testing which is why the researchers were still able to include a mango with disease test above. This is also the reason why disease detection wasn't included in table 2 as having only 2 mangoes with diseases for live camera testing is not proper and can outbalance the results.

Chapter 5

SUMMARY, CONCLUSION, AND RECOMMENDATION

This chapter presents the summary of findings, the conclusions of the study, and recommendations.

Summary of Findings

This study focused on the development and implementation of a vision-based system for ripe mango classification, mango detection, and disease detection using machine learning. The system utilized the YOLOv5 algorithm for object detection and classification tasks. The following are the key findings from the research:

1. The objective of the said system is to let users classify the ripeness of a mango, and it was successfully completed. Thanks to the work of python's machine learning capabilities and the YOLOv5 model, this was made possible.
2. Though YOLOv5 is an older model released in 2020, it still showed great results proving what it's known for - balance between speed and accuracy.
3. Detecting mangoes proved to have a high accuracy, with a success rate of 83.33% for file uploads, and 100.00% for live camera input during testing.
4. The system achieved 52.00% accuracy for image uploads with errors primarily attributed to factors like lighting conditions. However, during live camera testing, accuracy increased to 96.67%, indicating that real-time detection is more robust when environmental factors are controlled.

5. For disease detection, the model achieved 84.00% accuracy on uploaded images with test images coming from online dataset. It effectively detected the mango diseases the system has to offer. However, due to mango off-season and lack of availability during the making of this project, live camera testing for disease detection is not well tested, but during the testing, healthy mangoes were used and the system did not detect any diseases which is promising.
6. Positive outcomes were obtained from the system's testing which satisfies the third objective of helping users classify the ripeness level of a mango. This way users can make better decisions when purchasing and prevent food wastage as the mango can be eaten from the get go.

Conclusion

The Vision-based Ripe Mango Classifier System delivered its functionalities as what it is intended to be. It proves to be reliable, showing high accuracy results in classifying what's ripe and not. With a >90.00% overall score for the live camera input, this makes it easier for users as they can just use their phones in the instant they wish to buy mangoes. Testing shows promising results on the functionalities of the system. However, some features like disease detection are inconsistent, showing a need for big improvements. In conclusion, despite the optimistic results of the system, continuously improving and updating the dataset and the model will make the system useful in the bigger sense, reliable, and useful for all.

Recommendations

Based on the conclusions drawn, the following are the researchers' recommendations.

1. The researchers recommend using a newer model of the YOLO algorithm like YOLOv7 or YOLOv8 as it offers advantages, especially if the project demands higher accuracy, faster inference, or newer and better features.
2. The researchers recommend optimizing and fine tuning the dataset, model, and training with better hyperparameters (e.g. Epochs, Learning rate, Image size) for better results.
3. The researchers recommend optimizing the processing of the camera integrated with the models so lessen the lag. Having higher and smoother frames on the camera will definitely be a much better user experience.
4. The researchers recommend doing more tests for the disease model especially with the live camera input, so that the disease model will be tested properly and can be fine tuned accordingly.

REFERENCES

[1] B. R. Salvi, R. S. Varadkar, and N. V. Dalvi, "Recent Developments in Conventional Mango Breeding," 2019. Available: http://www.isasat.org/Vol-3-issue-1-Jan-2019/AARJ_III_1_2019_3_Salvi_pp30-42.pdf.

[2] A. G. S. Subrinho, J. L. P. Araújo, J. R. P. de Lima, and A. S. Neto, "Economic and Financial Determination of Mango Cultivation Under Different Irrigation Management, in the Curaçá Project, in Juazeiro - BA," 2021. Available: https://elicit.com/notebook/57828fc1-1109-4502-83a8-d7781d53a285#17ef520a1aee1a36e01c59e0d12d591f.

[3] S. Singh, "A Financial Viability and Relative Profitability of Mango Orcharding in Lucknow District of Uttar Pradesh," 2020. Available: https://elicit.com/notebook/57828fc1-1109-4502-83a8-d7781d53a285#17ef520a1aeba1b47e8dffc2245b243d.

[4] Vasuki. R1 , M. Soundariyan2, "Cost of Cultivation and Profitability of Mango Cultivation in Dharmapuri District, Tamil Nadu", 2022. Available: https://eprajournals.com/IJIR/article/9714/abstract

[5] Emmy Hamidah, "Analysis of Financial Feasibility and Business Economics Manalagi Mango Cultivation", 2023. Available: https://ouci.dntb.gov.ua/works/735awY0l/

[6] "Mango – Industry Strategic Science and Technology Plans (ISPs) Platform," 2022. Available: https://ispweb.pcaarrd.dost.gov.ph/isp-commodities/mango/.

[7] I. M. Hassoon, "Classification and Diseases Identification of Mango Based on Artificial Intelligence: A Review," *Journal of Quality and Communication Management*, vol. 14, no. 4, pp. 1–10, 2022. Available: https://doi.org/10.29304/jqcm.2022.14.4.1085

[8] D. Sahu and R. M. Potdar, "Defect Identification and Maturity Detection of Mango Fruits Using Image Analysis," 2019. Available: https://www.researchgate.net/profile/Ravindra-Potdar/publication/351945160_Defect_Identification_and_Maturity_Detection_of_Mango_Fruits_using_Image_Analysis/links/60c6c7ce92851ca6f8e9bba4/Defect-Identification-and-Maturity-Detection-of-Mango-Fruits-using-I

[9] "What Is Machine Learning (ML)?," *IBM*, Available: https://www.ibm.com/topics/machine-learning.

[10] "Mango Definition & Meaning," *Merriam-Webster*, August 2, 2024. Available: https://www.merriam-webster.com/dictionary/mango.

[11] R. Nithya, B. Santhi, R. Manikandan, and A. H. Gandomi, "Computer Vision System for Mango Fruit Defect Detection Using Deep Convolutional Neural Network," *MDPI*, 2022. Available: https://doi.org/10.3390/foods11213483.

[12] S. Veling, "Mango Disease Detection by Using Image Processing," 2020. Available: https://www.researchgate.net/publication/335816862_Mango_Disease_Detection_by_using_Image_Processing.

[13] G. S. Tumang, "Pests and Diseases Identification in Mango Using MATLAB," 2019. Available:

https://doi.org/10.1109/ICEAST.2019.8802579.

[14] V. Ashok and D. S. Vinod, "A Novel Fusion of Deep Learning and Android Application for Real-Time Mango Fruits Disease Detection," in *Advances in Artificial Intelligence and Data Engineering*, Singapore: Springer, 2020, pp. 811–818. Available: https://link.springer.com/chapter/10.1007/978-981-15-5400-1_74

[15] L. Benos, A. C. Tagarakis, G. Dolias, R. Berruto, D. Kateris, and D. Bochtis, "Machine learning in agriculture: A comprehensive updated review," *Sensors*, vol. 21, no. 11, pp. 1–45, 2021. Available: https://doi.org/10.3390/s21113758

[16] P. Jing, D. Ergu, F. Liu, Y. Cai, and B. Ma, "A Review of YOLO Algorithm Developments," *Procedia Computer Science*, vol. 199, pp. 1066-1073, 2022. Available: https://www.sciencedirect.com/science/article/pii/S1877050922001363.

[17] S. Chen, Y. Liao, and B. Huang, "An improved lightweight YOLOv5 algorithm for detecting strawberry diseases," 2023. Available: https://www.researchgate.net/publication/371259229_An_improved_lightweight_YOLOv5_algorithm_for_detecting_strawberry_diseases

[18] Y. Zhang, Z. Guo, J. Wu, Y. Tian, H. Tang, and X. Guo, "Real-Time Vehicle Detection Based on Improved YOLO v5," *Sustainability*, vol. 14, no. 19, 2022. Available: https://www.mdpi.com/2071-1050/14/19/12274.

[19] N. Rahaman, M. R. Chowdhury, A. Rahman, A. Ahmed, H. Hossain, M. Rahman, H. Biswas, S. Kader, A. T. Noyan, and M. Biswas, "A Deep Learning Based Smartphone Application for Detecting Mango Diseases and Pesticide Suggestions," 2023. Available: https://doi.org/10.12785/ijcds/1301104.

[20] D. Faye, I. Diop, and D. Dione, "Mango Diseases Classification Solutions Using Machine Learning or Deep Learning: A Review," *Journal of Computer and Communications*, vol. 10, no. 12, pp. 16–28, 2022. Available: https://doi.org/10.4236/jcc.2022.10120

[21] P. B. Jawade, D. Chaugule, D. Patil, and H. Shinde, "Disease Prediction of Mango Crop Using Machine Learning and IoT," 2020. Available: https://doi.org/10.1007/978-3-030-24322-7_33.

[22] S. Goel and K. Pandey, "A Survey on Deep Learning Techniques in Fruit Disease Detection," *International Journal of Distributed Systems and Technologies (IJDST)*, vol. 13, no. 8, pp. 1–16, 2022. Available: https://www.igi-global.com/gateway/article/307901

[23] M. Khaskheli, "Mango Diseases: Impact of Fungicides," 2020. Available: https://doi.org/10.5772/intechopen.87081.

[24] S. Chourasiya, A. A. Gadpale, P. S. Thethi, P. D. Nagdeve, P. Wakode, and M. Indorkar, "Apply machine learning and image processing to detect plant diseases," in *Proceedings of the 2023 IEEE Global Humanitarian Technology Conference (GHTC)*, 2023. Available: https://ieeexplore.ieee.org/document/10063139

[25] R. Polly and A. E. Devi, "A Deep Learning-based Study of Crop Diseases Recognition and Classification," 2022. Available: https://ieeexplore.ieee.org/document/9742950.

[26] A. V. M. Giri, I. A. Musdar, H. Angriani, and M. Taruk, "Enhancing Disease Management in Mango Cultivation: A Machine Learning Approach to Classifying Leaf Diseases," 2023. Available: https://www.researchgate.net/publication/377090439_Enhancing_Disease_Management_in_Mango_Cultivation_A_Machine_Learning_Approach_to_Classifying_Leaf_Diseases

[27] P. Kumar, S. Ashtekar, J. S. S., B. K. P., V. P. T., and R. M. Kumar, "Classification of Mango Leaves Infected by Fungal Disease Anthracnose Using Deep Learning," 2021. Available: https://ieeexplore.ieee.org/abstract/document/9418383.

[28] N. O. Delgado, E. R. Arboleda, J. L. Dioses Jr., and R. M. Dellosa, "Identification Of Mango Leaves Using Artificial Intelligence," 2019. Available: https://www.researchgate.net/publication/338047474_Identification_Of_Mango_Leaves_Using_Artificial_Intelligence/links/5dfb9a99a6fdcc28372ea1f0/Identification-Of-Mango-Leaves-Using-Artificial-Intelligence.pdf02

[29] S. P. Withanaarachchi, B. D. N. Jayamanne, A. Niroshani, M. Z. M. Aksham, M. M. Hansika, and S. Lokuliyana, "Smart Mango - Smart System for Mango Plantation," 2023. Available: https://www.semanticscholar.org/paper/Smart-Mango-Smart-System-for-Mango-Plantation-Withanaarachchi-Jayamanne/7a568b233f3b3a9d518bc04d87ee78fd1059688a.

[30] W. Sema, Y. Yayeh, and G. Andualem, "Automatic Detection and Classification of Mango Disease Using Convolutional Neural Network and Histogram Oriented Gradients," 2023. Available: https://www.researchsquare.com/article/rs-3073109/v1.

[31] D. Banerjee, V. Kukreja, S. Hariharan, and V. Jain, "Enhancing Mango Fruit Disease Severity Assessment with CNN and SVM-Based Classification," in *2020 IEEE 6th International Conference on Signal Processing and Communication (ICSC)*, Greater Noida, India, 2020. Available: https://ieeexplore.ieee.org/document/10126397

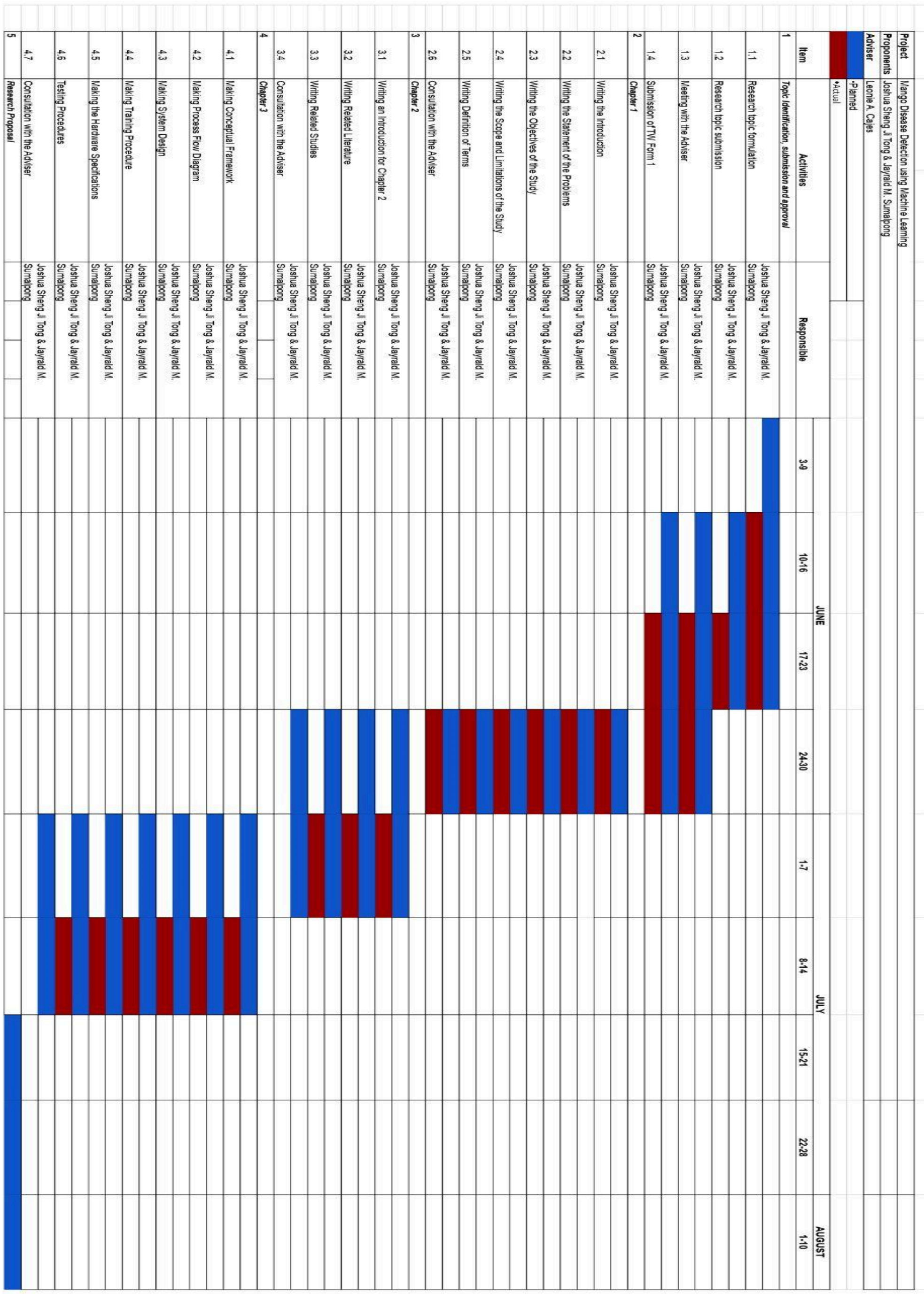
[32] P. H. Chandana, R. V. P. Yerra, and P. Subudhi, "MangoYOLO5: A Fast and Compact YOLOv5 Model for Mango Detection," 2023. Available: https://www.researchgate.net/publication/370566674_MangoYOLO5_A_Fast_and_Compact_YOLOv5_Model_for_Mango_Detection

[33] Huy, September 2024,“ Mango Dataset,” Roboflow [Online]. Available: https://universe.roboflow.com/huy-5srpj/mango-zda4q

[34] Mango, October 2024, “Hayss Computer Vision Project”, Robowlo [Online]. Available: https://universe.roboflow.com/mango-8g86x/hayss-yyxos

[35] Nikhil, 2024, “Fruit-disease-classification”, Roboflow [Online]. Available: https://universe.roboflow.com/nikhil-s7rhq/fruit-disease-classification/dataset/11

Appendix A. Gantt Chart



Appendix B. Permission Letter

St. Michael's College
College of Computer Studies
Quezon Ave., Iligan City, Philippines 9200

Date: 06/28/2024

RICARDO RESURRECCION
Fruit Farm Owner/Seller

Sir Resurrecion:

Praised be Jesus and Mary!

We are 4th year students of St. Michael's College and are currently undertaking a study entitled "Mango Disease Detection using Machine Learning" in partial fulfillment of the requirements for our Bachelor of Science in Information Technology (BSIT) course.

In line with this, we would like to ask your permission to allow us to conduct an interview to gather relevant information regarding common mango diseases.

We are confident that your contributions will significantly enhance the quality of our research. The interview will be scheduled at a time and place most convenient for you.

We are hoping for your favorable approval on this matter. Thank you very much and God bless!

Respectfully yours,

JOSHUA SHENG JI TONG
Researcher

JAYRALD M. SUMALPONG
Researcher

Noted by:

LEONIE A. CAJES
Research Adviser

EDSEL B. MONTEROLA, Ph.D.
College Dean

Received by:

Appendix C. User Testing

Curriculum Vitae

Tong, Joshua Sheng Ji U.

Pala-o, Iligan City

Mobile #: 09935216683

Email Address:
joshuashengjiuy.tong@my.smciligan.edu.ph

PERSONAL INFORMATION

Date of Birth:	December 6, 2001
Age:	23
Sex:	Male
Civil Status:	Single
Height:	5’11ft
Weight:	54kg
Citizenship:	Filipino
Religion:	Roman Catholic
Language:	Filipino, Cebuano, English, Chinese, Korean
Father’s Name:	Kei Hock L. Tong
Mother’s Name:	Jessica U. Tong

EDUCATIONAL BACKGROUND

- Tertiary

  Bachelor of Science in Information Technology
  St. Michael’s College of Iligan, Inc.
  Quezon Avenue, Iligan City
  S.Y. 2022-Present

Bachelor of Science in Information Technology
Xavier University-Ateneo de Cagayan
Corrales Avenue, Cagayan de Oro City
S.Y. 2020-2022

- Senior High

  Xavier University - Ateneo de Cagayan
  Corrales Avenue, Cagayan de Oro City
  S.Y. 2018-2020

- Secondary

  Lanao Chung Hua School
  Pala-o, Iligan City
  S.Y. 2014-2018

- Elementary

  Lanao Chung Hua School
  Pala-o, Iligan City
  S.Y. 2008-2014

SPECIAL SKILLS

- Proficient knowledge in Programming Languages (JavaScript, HTML, PHP, Python)
- Proficient knowledge in website developing using WordPress
- Basic knowledge in networking
- Basic knowledge in react mobile app development
- Proficient knowledge in graphics design and video editing

SEMINAR ATTENDED

- Google DevFest 2024 (Google Developer Group CDO) – Nov 16, 2024
- DICT Raising Awareness and Inspiring the Startup Ecosystem (RAISE) – Sept 5, 2024
- Adobe Illustrator Essentials - Sept 23, 2024

Curriculum Vitae

Sumalpong, Jayrald M.

Purok Matinabangon, Ubaldo D. Laya, Iligan City

Mobile #: 09762498586

Email Address: jayraldmontealto.sumalpong@my.smciligan.edu.ph

PERSONAL INFORMATION

Date of Birth:	June 11, 2003
Age:	21
Sex:	Male
Civil Status:	Single
Height:	5'8ft
Weight:	67.08kg
Citizenship:	Filipino
Religion:	Roman Catholic
Language:	Filipino, Cebuano, English
Father's Name:	Alquin C. Sumondong
Mother's Name:	Mae M. Sumalpong

EDUCATIONAL BACKGROUND

- Tertiary

  Bachelor of Science in Information Technology
  St. Michael's College of Iligan, Inc.
  Quezon Avenue, Iligan City

S.Y. 2021-Present

- Senior High

  Christians Horizon School Inc.
  Purok Mabinati-on, Ubaldo D. Laya, Iligan City, Philippines
  S.Y. 2019-2021

- Secondary

  Christians Horizon School Inc.
  Purok Mabinati-on, Ubaldo D. Laya, Iligan City, Philippines
  S.Y. 2015-2019

- Elementary

  Ubaldo D. Laya Elementary Central School
  Ubaldo D. Laya, Iligan City
  S.Y. 2009-2015

SPECIAL SKILLS

- Proficient in Django, Python, PHP, HTML and CSS
- Proficient in Adobe Photoshop for graphics design
- Energetic and firm, enthusiastics, and quick to assimilate concepts
- Good in problem solving skills

SEMINAR ATTENDED

- Google DevFest 2024 (Google Developer Group CDO) – Nov 16, 2024
- my.ComApps - A technology development program - April 2024
- Optimization Technique Prediction Model Through Modified Crossover Operator - October 2023
- Student Leadership Training - October 2023