Faculty of Engineering

Vision-Based Leaf Disease Detection Using an Improved ShuffleNet Architecture

Chyntia Jaby ak Entuni

Doctor of Philosophy
2026

Vision-Based Leaf Disease Detection Using an Improved ShuffleNet Architecture

Chyntia Jaby ak Entuni

A thesis submitted

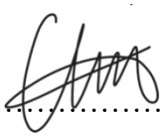
In fulfillment of the requirements for the degree of Doctor of Philosophy

(Electronics Engineering)

Faculty of Engineering
UNIVERSITI MALAYSIA SARAWAK
2026

DECLARATION

I declare that the work in this thesis was carried out in accordance with the regulations of Universiti Malaysia Sarawak. Except where due acknowledgements have been made, the work is that of the author alone. The thesis has not been accepted for any degree and is not concurrently submitted in candidature of any other degree.



……

Signature

Name: Chyntia Jaby ak Entuni

Matric No.: 22010072

Faculty of Engineering

Universiti Malaysia Sarawak

Date : 6th January 2026

ACKNOWLEDGEMENT

This thesis is the result of a journey made possible by the grace of God and the kindness of many. I am profoundly grateful for His guidance and for all who have supported me along the way.

To my supervisor, Dr Tengku Mohd Afendi bin Zulcaffle, your guidance has been a steady presence, providing clarity in times of uncertainty. Your wisdom has been invaluable, especially in teaching me MATLAB, which became an essential tool throughout this study. I am also thankful to my co-supervisor, Associate Professor Ir Ts Dr Kismet anak Hong Ping, for the insightful advice and support. Your advice taught me to foster a positive research environment, by valuing diverse perspectives and exercising patience. Thank you also for always helping me with the forms. To Professor Amit Baran Sharangi, thank you for your cooperation in verifying leaf disease types for this study and for your assistance with the journal publication.

To my dad, thank you for your quiet love, you work hard to support our family. To my mom, thank you for always being there and sorry for often leaving you with the chores while I only helped occasionally. To my brother, thank you to both mom and you for helping take care of the plant used in this study and I hope you recover soon so you can take me everywhere. To my friends, Elsa and Jack, thank you for always bringing humour and joy into my life. Your presence has been a constant source of comfort.

Lastly, I am deeply grateful to the Centre for Graduate Studies, UNIMAS, for the Zamalah Graduate Scholarship, which eased my burden and allowed me to focus on my research. Their generosity has been invaluable to my academic journey.

ABSTRACT

In the agricultural sector, accurate and efficient plant disease detection is essential for ensuring food security, minimizing economic losses, and reducing the environmental impact of excessive pesticide use. Traditional models such as Bag of Features (BoF), DenseNet-201, ResNet-50, and ShuffleNet variants often struggle with real-world complexities, including occlusions, varying lighting conditions, and clustered foliage. These limitations hinder their ability to detect diseased leaves in practical agricultural environments. This research introduces an improved ShuffleNet model, incorporating additional convolutional layers and optimized parameters, combined with Kinect-based imaging for precise disease identification. The proposed approach effectively detects grey spot, discolored leaf, and leaf curling diseases within dense capsicum leaf clusters, achieving a peak accuracy of 91.94%, significantly surpassing conventional methods. The model's success is attributed to rigorous hyperparameter tuning, with an optimal learning rate of 0.010, 50 epochs, a minibatch size of 64, and the Adam Optimizer, balancing accuracy and computational efficiency. Additionally, the model generalizes well across multiple crops, including capsicum, rice, corn, tomato, and citrus. In conclusion, this research provides an innovative deep learning-driven plant disease detection framework, capable of operating in complex agricultural conditions. Overall, the experiments achieved 91.94% accuracy, 92.10% precision, 91.60% recall, and 91.80% F1-score across tested crops, confirming the robustness and scalability of the proposed model. Future work should improve efficiency with knowledge distillation for real-time use on edge devices, and combine hyperspectral, thermal imaging, and IoT platforms to better detect diseases, monitor at scale, and support smart farming.

Keywords: Plant leaf diseases, Capsicum, Kinect camera, Improved ShuffleNet

Pengesanan Penyakit bagi Pemantauan Pertanian Pintar Varieti Tumbuhan di Sarawak Berdasarkan Imej Daun menggunakan Algoritma CNN yang Ditambah Baik

ABSTRAK

Dalam sektor pertanian, pengesanan penyakit tumbuhan yang tepat dan berkesan merupakan elemen penting dalam memastikan keterjaminan makanan, mengurangkan kerugian ekonomi, serta meminimumkan kesan alam sekitar akibat penggunaan racun perosak secara berlebihan. Model tradisional seperti Bag of Features (BoF), DenseNet-201, ResNet-50, dan varian ShuffleNet sering menghadapi kekangan dalam persekitaran sebenar, khususnya disebabkan oleh halangan fizikal (occlusion), keadaan pencahayaan yang tidak konsisten, serta dedaunan yang bertindih. Kekangan ini menjejaskan keberkesanan model tersebut dalam mengesan daun yang berpenyakit di persekitaran pertanian sebenar. Kajian ini membangunkan model ShuffleNet yang ditambah baik, dengan pengintegrasian lapisan konvolusi tambahan serta pengoptimuman parameter, di samping penggunaan teknik pengimejan berasaskan Kinect untuk pengecaman penyakit yang lebih tepat. Pendekatan ini membolehkan pengesanan berkesan terhadap penyakit bintik kelabu, daun berubah warna, dan daun kerinting pada kelompok daun cili yang padat, dengan pencapaian ketepatan tertinggi sebanyak 91.94%, sekali gus mengatasi kaedah konvensional. Keberhasilan model ini didorong oleh penalaan hiperparameter yang rapi, melibatkan kadar pembelajaran optimum 0.010, 50 epoch, saiz minibatch 64, serta penggunaan algoritma pengoptimuman Adam, yang berjaya menyeimbangkan aspek ketepatan dan kecekapan pengiraan. Selain itu, model ini turut menunjukkan keupayaan generalisasi yang baik merentas pelbagai jenis tanaman lain, termasuk capsicum, padi, jagung, tomato, dan sitrus. Secara keseluruhannya, kajian ini memperkenalkan satu kerangka inovatif berasaskan pembelajaran mendalam untuk pengesanan penyakit tumbuhan, yang mampu berfungsi dalam persekitaran pertanian

yang kompleks. Secara keseluruhannya, eksperimen mencapai ketepatan 91.94%, ketepatan (precision) 92.10%, kepekaan (recall) 91.60%, dan skor F1 91.80% merentasi tanaman yang diuji, sekali gus mengesahkan keteguhan serta kebolehskalaan model yang dicadangkan. Penyelidikan pada masa hadapan wajar menumpukan kepada peningkatan kecekapan melalui pendekatan knowledge distillation bagi membolehkan penggunaan masa nyata pada peranti hujung rangkaian, serta penggabungan pengimejan hiperspektral, terma, dan platform IoT bagi memperkukuh keupayaan pengesanan penyakit, pemantauan berskala besar, dan sokongan terhadap sistem pertanian pintar.

Kata kunci: *Penyakit daun tumbuhan, Capsicum, Kamera Kinect, ShuffleNet yang ditambah baik*

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ANN	Artificial Neural Networks
API	Application Programming Interface
APS-C	Advanced Photo System Type C
AUC-ROC	Area Under the Receiver Operating Characteristic Curve
BoF	Bag of Feature
CCD	Charge-coupled Device
CIPA	Camera and Imaging Products Association
CMOS	Complementary Metal-Oxide-Semiconductor
CNN	Convolutional Neural Network
CPU	Central Processing Unit
DCNN	Deep Convolutional Neural Network
DDR	Double Data Rate
DenseNet	Densely Connected Convolutional Networks
DLCNN	Deep Learning Convolutional Neural Network
DOSM	Department of Statistics Malaysia
DSLR	Digital Single-Lens Reflex
DT	Decision Tree
FC	Fully Connected
FLOPs	Floating Point Operations per Second
FN	False Negatives
FOR	False Omission Rate
FP	False Positives

FPN	Feature Pyramid Network
GAN	Generative Adversarial Network
GB	Gradient Boosting
GLCM	Gray Level Co-occurrence Matrix
HD	High Definition
HDMI	High-Definition Multimedia Interface
ID	Identification
IDR	Import Dependence Rate
IG	Information Gain
IITA	International Institute of Tropical Agriculture
IoT	Internet of Things
IoU	Intersection over Union
KNN	K-Nearest Neighbor
LAB	Lightness: Red/Green: Blue/Yellow
LCD	Liquid-crystal Display
mAh	Milliampere-hour
MB	Megabytes
MCC	Matthews Correlation Coefficient
MLP	Multi-Layer Perceptron
MP	Megapixel
N/A or NA	Not Applicable
ND	Neutral Density
NPV	Negative Predictive Value
PDS-CNN	Parallel Depth-wise Separable Convolutional Neural Network
RAM	Random Access Memory

RBFNN	Radial Basis Function Neural Network
RCNN	Region-based Convolutional Neural Network
ReLU	Rectified Linear Unit
ResNet	Residual Neural Network
RF	Random Forest
RGB	Red, Green, Blue
RMSprop	Root Mean Squared Propagation
RPN	Region Proposal Network
SGD	Stochastic Gradient Descent
SHAP	SHapley Additive exPlanations
SIFT	Scale-Invariant Feature Transform
SSD	Solid State Drive
SURF	Speeded-Up Robust Features
SVM	Support Vector Machine
TN	True Negatives
TP	True Positives
UKM	Universiti Kebangsaan Malaysia
USB	Universal Serial Bus
UV	Ultraviolet
VGG	Visual Geometry Group
Wi-Fi	Wireless Fidelity
DWT	Wavelet Transform
YOLOv3	You Only Look Once, Version 3

CHAPTER 1

INTRODUCTION

1.1 Study Background

According to the New Straits Times 29 September 2023 (Govind, 2023), it is estimated that 60% of Malaysia's food is imported, and with the dropping ringgit and rising food prices due to global supply chain disruptions, Malaysians have been facing a declining standard of living, particularly with the higher price of vital products. Food imports to Malaysia are estimated to be worth around 75.71 billion Malaysian ringgit in 2022, up from 63.65 billion the previous year. Economists have also claimed that Malaysia only produces 46% of its vegetables, 70% of its rice, 61% of its fruits, and 25% of its meat each year (Omar, 2022). Besides, according to data from the Department of Statistics Malaysia (DOSM) on Supply and Consumption of Selected Agricultural Commodities in Malaysia 2017-2021, crops have the highest import dependence rate (IDR) compared to fruits, livestock, and fisheries, with capsicum leading the other crops with 75.1% IDR (Mohammad, 2022). There are several factors that contribute to the increasing of IDR of agriculture sector in Malaysia and one of them is due to the low productivity and quality of the agriculture produced. Plant leaf diseases are among the causes of the crop yields losses in agriculture. This issue also leads to the damage on the environment. This is when the farmers used inappropriate amount of pesticides because they are unable to detect the actual type of diseases on their crops.

Drawing insights from the extensive literature, it is discerned that plant leaves in Malaysian exhibit susceptibility to an array of diseases, including but not limited to leaf blight, grey spot, and rust (Entuni & Zulcaffle, 2022; Kutawa et al., 2021). In the prevailing outdoor environments where these plants thrive, marked by ecological complexity, various

plant leaf diseases manifest through distinctive physiological alterations, visual transformations, and variations in infection intensity. The discernible characteristics of the infected regions and the degree of infection on the leaf emerge as pivotal indicators for delineating diverse forms of plant leaf diseases. Consequently, for disease classification purposes, a sophisticated approach is employed involving the acquisition, segmentation, and classification of diseased leaves. While traditional methodologies hinge on visual scrutiny by experts, the associated temporal cost proves to be considerable. Conversely, the conventional practice of physically transporting specimens to specialists introduces additional delays, underscoring the imperative for advanced, time-efficient techniques in the detection of plant leaf diseases.

As smart farming continues to advance, leveraging sophisticated cameras and enhanced computer vision technologies becomes pivotal in the realm of precision agriculture, smart greenhouses, and related applications. Automated models for detecting plant infections are now integral components of smart-farming systems. Using image processing for plant leaf disease detection is beneficial because it saves time and effort compared to manual inspections of large crop fields. It also allows for early disease identification, helping prevent the spread of diseases within the plant. Presently, diverse image acquisition and classification algorithms are utilized for detecting plant leaf diseases. Image acquisition involves capturing an image, and classification assigns pixels to specific classes based on defined criteria (Pujar et al., 2024). This combination of image acquisition and classification enhances the effectiveness of automated plant disease detection by providing a comprehensive understanding of the captured images and enabling accurate categorization of diseased and healthy areas.

In the realm of smart farming, the identification of plant leaf diseases has been facilitated by a myriad of techniques. Among the commonly employed devices for capturing images of diseased leaves are standard digital cameras and smartphone cameras, aligning with the technological integration in precision agriculture. On the classification forefront, a diverse array of algorithms has been utilized, ranging from usual approaches like the Fuzzy classifier, support vector machine (SVM), K-means, random forest, decision tree, and Naïve bayes, to more advanced neural network models including convolutional neural network (CNN), deep convolutional neural networks (DCNN), deep learning convolutional neural network (DLCNN), You Only Look Once, Version 3 (YOLOv3), Faster region-based convolutional neural network (Faster-RCNN), MobileNet, GoogleNet, Inception, ResNet50, and DenseNet201.

Despite the widespread application of these methods, there remains a pressing need in smart farming for more robust and efficient solutions, both in terms of data acquisition devices and classification algorithms specifically designed for plant leaf disease detection. Existing approaches face several notable limitations. First, many models are highly sensitive to variations in environmental conditions, such as inconsistent lighting, background clutter, and leaf orientation, which reduce detection accuracy in real-world scenarios (Ferentinos, 2018). Second, most deep learning-based methods depend heavily on large, annotated datasets for training, which are costly and time-consuming to collect, particularly in agricultural settings where expert labelling is required (Barbedo, 2019). Third, some approaches lack scalability and computational efficiency, making them unsuitable for deployment on resource-constrained devices commonly used in the field (Kamilaris & Prenafeta-Boldú, 2018). These challenges highlight the need for more adaptable and lightweight solutions tailored to practical agricultural applications. Thus, this study

introduces a proposed method for plant leaf disease detection, aiming to address these gaps, and its performance is meticulously compared with outcomes from previous methods, contributing to the ongoing evolution of smart farming practices.

1.2 Motivation and Significance

Recent progress in automated disease diagnosis models has greatly improved how diseased plant leaf images are captured. This study reviews different optical devices used for taking images of infected leaves and shows how combining these tools with machine learning has increased the accuracy and efficiency of disease detection. The use of advanced imaging methods together with modern classification algorithms provides a strong basis for faster and more reliable identification of plant diseases. These automated models mark a clear shift from older methods, reducing both time and effort, and highlight their importance as a major step forward in plant disease detection.

The survey conducted in this study critically reviews existing models for plant leaf disease detection and constructs a computational framework that accounts for their reported limitations. The research is motivated by weaknesses found in earlier data acquisition devices and classification algorithms, guiding the development of a model with improved performance. Data acquisition devices such as normal RGB cameras and smartphones are often affected by light changes, background noise, and low image quality, which reduce the reliability of captured leaf images (Barbedo, 2016; Singh et al., 2020). On the other hand, traditional classification algorithms like SVM, decision trees, and random forests depend heavily on handcrafted features and perform poorly with large or varied datasets (Ferentinos, 2018; Too et al., 2019). Even deep learning models, while more accurate, can be limited by high computational cost and the risk of overfitting when trained on small datasets (Mohanty

et al., 2016; Kamilaris & Prenafeta-Boldú, 2018). Unlike previous works, the proposed approach emphasizes both accuracy and robustness in its design.

This study contributes to the literature by offering a comparative analysis that considers both data acquisition devices and classification algorithms, an aspect that has received relatively limited attention in prior research (Too et al., 2019). By evaluating different devices and algorithms, this work provides a benchmarking perspective that may inform the refinement of future methods. Furthermore, the dataset used in this study incorporates images collected from authentic ecological environments alongside reputable online resources such as the PlantVillage Image database. This combination allows the dataset to capture a range of conditions that more closely resemble those encountered in practice. Hence, the results derived from this study can offer practical insights to researchers and agricultural practitioners, supporting efforts to improve disease detection and, in turn, agricultural productivity.

1.3 Problem Statement

Conventional virtual inspection methods for plant leaf disease detection, along with previous data acquisition devices and classification algorithms, have demonstrated inconsistent and often suboptimal accuracy (Liu et al., 2021; Demilie, 2024). Existing classification techniques, including traditional machine learning methods such as the Fuzzy classifier, SVM, K-means, random forest, decision tree, and Naïve Bayes, as well as deep learning models like CNN, DCNN, DLCNN, YOLOv3, Faster-RCNN, MobileNet, GoogleNet, Inception, Bag of Features, ResNet50, and DenseNet201, have reported detection accuracies ranging from 2.86% to 97.0% (Too et al., 2019). While some of these models have attained high accuracy in controlled settings, their effectiveness often declines

in real-world applications due to inconsistencies in lighting exposure. This inconsistency in accuracy highlights the need for a more adaptable and robust approach.

A critical limitation lies in data acquisition, where conventional smartphone and digital cameras struggle to capture high-quality images under various lighting conditions. Exposure issues, caused by harsh shadows and intense reflections, often lead to overexposed or underexposed images, obscuring disease symptoms and ultimately reducing classification accuracy (Fuentes et al., 2017). Since high-performing models rely on clear and well-exposed input data, the inefficiencies of previous acquisition methods contribute to the performance gap observed across existing classification algorithms. Without improvements in data capture techniques, even the most sophisticated deep learning models remain constrained in their ability to achieve reliable and consistent detection across varying environmental conditions.

Additionally, existing classification algorithms face considerable challenges when applied to visually complex scenarios, particularly in distinguishing diseased leaves from dense foliage or intricate leaf clusters. While some deep learning models have reached up to 97.0% accuracy, their success is largely dependent on ideal datasets and controlled conditions (Barbedo, 2018). When confronted with real-world complexities, such as overlapping leaves, subtle disease symptoms, and diverse plant species, no development in this area has been introduced.

To address these limitations, the method integrates enhanced data acquisition strategies that mitigate exposure issues, a refined classification model capable of handling complex visual backgrounds, and optimized feature extraction techniques to improve sensitivity to subtle disease symptoms. While existing models may report higher accuracy

in isolated conditions, their effectiveness diminishes when applied to real-world challenges (Ferentinos, 2018). By contrast, this approach maintains strong detection performance even in adverse conditions, establishing a more practical and robust solution for plant leaf disease classification.

1.4 Objectives

To successfully address the problems under investigation, the current research aims to propose a data acquisition device and classification algorithm that can performed better than the existing methods to detect plant leaf diseases. The main objectives are outlined as follows:

i. To implement an enhanced data acquisition strategy using Kinect for capturing plant leaf images under varying conditions.
ii. To develop an improved ShuffleNet CNN model with FPN and RPN integration for accurate detection of plant leaf diseases in complex environments.

1.5 Research Scope

This study evaluates a variety of exiting data acquisition devices, ranging from standard digital cameras and smartphone cameras to diverse classification algorithms. The examined algorithms include the Fuzzy classifier, SVM, K-means, random forest, decision tree, Naïve Bayes, and advanced neural network models like CNN, DCNN, DLCNN, YOLOv3, Faster-RCNN, MobileNet, GoogleNet, Inception, Bag of Features, ResNet50, and DenseNet201. Subsequently, attention is directed towards the proposed model, employing a cost-effective yet sophisticated camera as the data acquisition device and an improved ShuffleNet as the classification algorithm. This model is specifically designed to detect three

distinct leaf diseases, grey spot, discolored leaf, and leaf curling, in capsicum plants. To affirm the superiority of the proposed improved ShuffleNet algorithm, its efficacy is further evaluated across various plant species such as rice, corn, tomato, and citrus, obtained from accessible online repositories. The performance metrics, including accuracy, specificity, sensitivity, precision, recall, F1 score, negative predictive value, false omission rate, miss rate, and fall-out, are recorded to gauge the effectiveness of the proposed model.

Subsequently, this study presents a comprehensive comparison of the proposed model's performance across a spectrum of performance metrics, including accuracy, specificity, sensitivity, precision, recall, F1 score, negative predictive value, false omission rate, miss rate, and fall-out. The proposed model integrates a depth-sensing camera for data acquisition and an improved ShuffleNet for the classification algorithm. This performance evaluation is juxtaposed with that of previous plant leaf detection models, namely, Bag of Feature (BoF), ResNet50, and DenseNet201. Additionally, the study delves into the effectiveness of the proposed model through iterative fine-tuning, exploring various parameters such as learning rate, number of epochs, minibatch size and optimizer selection. This rigorous investigation aims to pinpoint the optimal values, ensuring the precision and reliability of the proposed plant leaf diseases detection model before finalization.

1.6 Thesis Structure

The structure of this thesis is organized as follows, with each chapter briefly described. Chapter 1 establishes the conceptual framework for the research, offering insights into the study background, motivation, significance, problem statement, objectives, hypothesis, research scope, and the overall organization of the thesis. In Chapter 2, a comprehensive review of existing research is presented. This includes an exploration of

various types of plant leaf diseases and their associated negative impacts. Furthermore, a comparative study of prior data acquisition devices and classification algorithms employed in the context of plant leaf diseases is undertaken. Moving on to Chapter 3, the methodology applied in developing the proposed plant leaf diseases detection model is introduced. This encompasses the acquisition of the image dataset using a Kinect camera and the system modelling of the proposed improved ShuffleNet algorithm for classification. Chapter 4 presents a detailed account of the outcomes of the proposed model, including the accuracy of detection and the consideration of various performance metrics. The chapter also engages in a thorough comparison between the performance of previous detection models and the proposed model in the realm of plant leaf diseases detection. Additionally, the validation of the proposed model is scrutinized. Finally, Chapter 5 encapsulates the conclusions drawn from the major findings of this research, highlighting the contribution of the research work and offering recommendations for future endeavours.

CHAPTER 2

LITERATURE REVIEW

2.1 Overview

In the previous chapter, the foundation for this study was established, encompassing an exploration of the study's context, problem statements, research objectives, research contributions, and the overall thesis structure. Building upon this groundwork, the current chapter embarks on a multifaceted exploration. It commences with an elucidation of diverse categories of plant leaf diseases, intricately woven with a comprehensive analysis of the underlying ecological conditions inherent in plant life.

Additionally, this section conducts a thorough examination of a spectrum of optical devices, ranging from conventional digital cameras to cutting-edge cameras. This thorough examination evaluates their applicability and efficacy in the context of the ongoing research. Furthermore, the chapter provides a comprehensive review of recent models in the field, considering the extensive and varied research efforts conducted by scholars and practitioners globally.

The subsequent structure of this chapter is as follows. An in-depth examination of plant leaf diseases is presented in Section 2.2, inheritance ecological of plants discussed in Section 2.3, followed by a comprehensive overview of the various optical devices commonly employed by researchers, expounded upon in Section 2.4. Section 2.5 offers a critical appraisal of the previous methodologies and techniques employed in the detection of plant leaf diseases, while Section 2.6 delves into the synthesized findings derived from a comprehensive literature review. Finally, this chapter is summarized in Section 2.7.

2.2 Leaf Diseases

Plant leaf diseases significantly impact the economy in agricultural regions like Sarawak, where indigenous communities rely on subsistence farming (Grasty et al., 1999). Traditional farming methods by Malays, Melanaus, and some Dayaks, particularly along the Rajang River Basin, are central to their sustenance (Abdullah, 2016; Tanaka et al., 2007). However, these methods often involve unsuitable pest control practices, posing risks to the environment, human health, and the efficacy of pest management (Damalas & Eleftherohorinos, 2021). Consequently, crops face vulnerability to diseases, potentially leading to significant losses. Adopting modern early disease diagnosis methods can help farmers take quick and targeted actions, improving crop yields, ensuring local food security, and reducing reliance on imports.

In Malaysia, dependence on imported vegetables and plants, such as capsicum with a 75.1% Import Dependence Rate in 2021, is a major concern (Anaya-Esparza et al., 2021). The depreciation of the Malaysian Ringgit increases import costs, causing price volatility and affecting food affordability (Arin et al., 2023). Imported vegetables like capsicum are susceptible to global market fluctuations, impacting local prices and consumers' purchasing power (Benzie & John, 2015). Effective methods for detecting plant leaf diseases are essential to minimize farmer losses and enhance agricultural productivity. This would help stabilize local food prices, improve consumers' purchasing power, and contribute to overall economic stability.

2.2.1 Grey Spots

The occurrence of grey spots, caused by the fungus *C. zeae-maydis*, is a pervasive issue affecting a wide range of plant species globally (Jain et al., 2019; Dhami et al., 2019). These spots typically appear as small, oval to elongated lesions that evolve into conspicuous grey or tan patches with a distinctive black border, as depicted in Figure 2.1. The significance of grey spot disease in plant pathology is highlighted by its visible impact on foliage. Khoo et al., (2023) reported this disease in Borneo, specifically in Sabah, affecting rice leaves. Historical records also indicate occurrences of Cercospora leaf spot infections in Peninsular Malaysia dating back to 1953 (Thompson & Johnson, 1953).

Detecting Cercospora leaf spot can be challenging due to the inconspicuous and less defined nature of early lesions, making them difficult to observe, especially from a distance (Park et al., 2020). Symptoms typically first appear on the plant's lower or older leaves, often concealed by upper leaves (Srivastava & Nelson, 2021). Additionally, these lower leaves may be obscured by the plant's canopy or dense foliage, further reducing visibility. This complexity underscores the importance of early detection and targeted disease management strategies.

Figure 2.1: Gray leaf spot/ *Cercospora zeae maydis* (Holmes et al., 2023)

2.2.2 Discoloured Leaf

Leaf discoloration or chlorosis is a common plant disease characterized by the yellowing of the area between leaf veins, while the veins themselves either remain green or also turn yellow, as depicted in Figure 2.2. Insect pests that feed on plant sap can induce stress in the plant, leading to leaf yellowing as they transmit viruses responsible for chlorosis (He et al., 2021). Prompt identification of the chlorosis source and necessary remedies are crucial to prevent the issue from spreading. Certain plant viruses interfere with chlorophyll production or transport, causing widespread yellowing. Rebitanim et al., (2020) highlighted palm oil leaf yellowing associated with chlorosis in Peninsular Malaysia, emphasizing the need to mitigate farmer losses. However, the focus on disease treatment often lacks a strategy for early detection (Saibeh et al., 2020). Additionally, Saibeh et al., (2020) documented this disease in Sabah, Borneo, noting its easy spread to neighbouring plants due to the associated virus.

Recognizing yellow leaves in plants is challenging due to factors complicating the identification process. Early detection demands careful monitoring and frequent examination, which can be impractical in large agricultural settings (Talaviya et al., 2020). The yellow hue of leaves may blend with the surrounding environment or other natural colors, reducing visibility to the human eye. This blending effect adds complexity to identifying and addressing chlorosis, especially in vast plantations (Khoo et al., 2017).

Figure 2.2: Discolour leaf/ chlorosis (Maner, 2023)

2.2.3 Leaf Curling

Leaf curling disease significantly impacts a diverse range of crops, leading to substantial yield losses if not addressed. The curling of leaves, where edges or entire leaves curl inward or upward, and stunted plant development are common indicators, as illustrated in Figure 2.3 (Panno et al., 2021). Recently, Mahmoudieh et al., (2020) reported that leaf curl disease affected tomatoes (*Solanum lycopersicum*) in Peninsular Malaysia. In Malaysia, the disease was first observed in 1998 at the MARDI Jalan Kebun Research Station near Kuala Lumpur (Shih et al., 2018). More recently, Sau et al., (2020) provided the first molecular confirmation of pepper leaf curl virus infecting chilli peppers (*Capsicum annuum*), successfully sequencing the complete viral genomes at the Universiti Kebangsaan Malaysia (UKM) laboratory.

Identifying leaf curling diseases is challenging due to the subtle interplay of structural and color changes in plants (Ahmed & Harshavardhan Reddy, 2021). These diseases often cause mild discoloration or paleness initially confined to specific parts of the leaves, blending with typical leaf color variations and reducing visibility (Chen et al., 2019;

Shingote et al., 2022). The complexity is compounded by the gradual nature of these changes, with leaf edges curling or entire leaves bending upward over time, making early symptoms less conspicuous and impeding timely recognition (Shingote et al., 2022). This slow development can allow the disease to advance unchecked, highlighting the need for vigilant monitoring.

Figure 2.3: Leaf curling (Muntean, 2022)

2.3 The Inherent Ecological State of Plants

Plant diseases exhibit a diverse array of symptoms, making it challenging to link a set of symptoms to a specific disease (Campos & Ortiz, 2019). This complexity is further compounded when attempting to capture images of diseased leaves that are densely packed on a plant (Tugrul et al., 2022). In such scenarios, the relative position of the light source can cast shadows among tightly clustered leaves, as depicted in Figure 2.4, obstructing the clear visualization of individual leaves. Additionally, the reflective properties of leaf surfaces introduce further intricacies, causing glare or reflections that obscure specific portions of the leaves.

These challenges not only hinder the accurate representation of each leaf but also impact the acquisition of high-quality images, essential for the effectiveness of automated image recognition algorithms like CNN (Yamashita et al., 2018). CNNs, being particularly sensitive to input data quality, may struggle with distortions such as shadows, reflections, or inadequate lighting, thereby impeding their accuracy in categorizing and recognizing images. Recognizing and addressing these issues are vital steps toward enhancing the efficiency of automated systems in diagnosing plant leaf diseases based on visual symptoms.

Figure 2.4: Leaf shadow (Tugrul et al., 2022)

The determination of shadow size and geometry can be derived from the principles of geometry, trigonometry, and optics. In the case of plant leaves, the length of a shadow (S) cast by one leaf onto another is calculated as:

$$S = L * tan(\theta) \qquad \text{Equation 2.1}$$

where S represents the shadow length, L is the horizontal distance between the tip of the shadow-casting leaf and the receiving leaf, and θ is the angle of incoming sunlight relative to the horizontal plane. The angle θ is determined by comparing the direction of solar rays

with the horizon. The geometric relationship is illustrated in Figure 2.5, adapted from Giacaman et al., (2022).

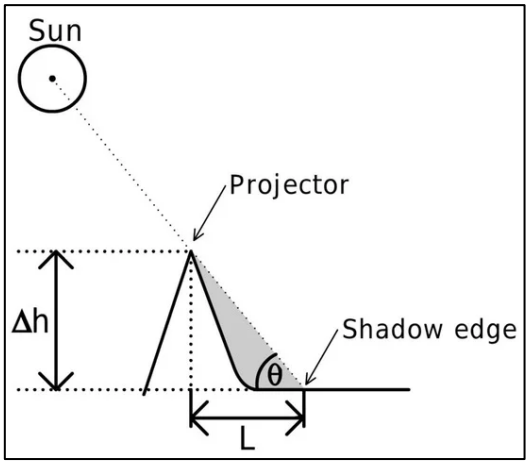


Figure 2.5: Conceptual geometry of shadow formation: a projector casts a shadow of length S on a receiving surface located a horizontal distance L away, with sunlight arriving at an angle θ above the horizontal (Giacaman et al., 2022)

Crops are typically cultivated in regions with high light intensity (Shafiq et al., 2021), posing challenges in capturing images of diseased leaves. The heightened light intensity results in pronounced contrasts between light and shadow, complicating the attainment of well-exposed images (Pistellato et al., 2023). Direct sunlight contributes to overexposed (bright highlights) and underexposed (dark shadows) regions on the leaves, obscuring diseased portions and complicating the differentiation between healthy and damaged plant parts. Image-analysis algorithms, crucial for detecting patterns and irregularities, rely on precise and well-exposed images. The extreme contrasts caused by high light intensity make it challenging for algorithms to distinguish between healthy and unhealthy areas. Consequently, details vital for accurate diagnosis may be obscured by overexposed or underexposed regions. These challenges culminate in two primary issue which are the

difficulty of optical devices in capturing images of diseased leaves and the limitations of algorithms in effectively analysing these images.

2.4 Optical Devices

In recent years, notable progress has been made in the broader field of object recognition, particularly in the realm of detecting plant leaf diseases. These advancements stem from various approaches, typically characterized by differences in data acquisition devices and detection algorithms. Optical devices, instrumental in this progress, are tools that manipulate light for diverse purposes to capture images (Zhang et al., 2023). Leveraging light properties such as reflection, refraction, diffraction, and absorption, these devices perform functions like magnification, focusing, filtering, and conveying image data (Ko et al., 2022). Their significance extends across disciplines encompassing medical imaging, remote sensing, microscopy, industrial imaging, holography, and computer vision (Haleem et al., 2022). Essentially, optical devices exploit light's attributes to enable a broad spectrum of practical applications and technological advancements. Serving as vital instruments, they exert influence on various facets of contemporary life, simultaneously driving technological innovation.

2.4.1 Standard Digital Camera

In the last few years, significant strides have been achieved in computer vision, especially concerning the collection of data using conventional digital cameras. This progress is notably evident in agriculture, where there has been a rapid and escalating development of models for detecting plant leaf diseases. These detection models primarily rely on utilizing a standard digital camera for data collection.

For instance, Petrellis (2019) used ordinary smartphone cameras, ranging from 5 MPixels (Lumia 550, Microsoft, Redmond, Washington, U.S.) to 23 MPixels (Sony Xperia XA1, Tokyo, Japan) to capture test images of citrus. Lesion locations were identified with notable precision using a simple fuzzy-like classification model, achieving 90% accuracy under most conditions. The study required capturing a single leaf against a significantly brighter background, such as placing the leaf on a sheet of paper or a marble bench, to prevent pixel confusion between the leaf and the background. While smartphone cameras are convenient, they struggle in low-light conditions, resulting in noisy images with limited information (Bernardino et al., 2023). Although these devices offer a flashlight feature to address this, the harsh lighting can cause glare and uneven light dispersion, leading to overexposed and underexposed areas unless the light angle is adjusted (Akkara & Kuriakose, 2019).

Additionally, Ahmed and Reddy (2021) used a smartphone camera to capture images of diseased plant leaves, including apple, grape, strawberry, squash, and soybean. Their CNN model achieved a 94% classification accuracy after adjusting the contrast, reducing noise, and desaturating the images to improve stability and learning efficacy. The images were taken in agricultural fields, but the model faced challenges with certain classes, like tomato diseases, due to background noise. Smartphone lenses, which are smaller and lack advanced anti-flare coatings found in dedicated cameras, often result in lens flare and lower image quality, especially in direct sunlight (Blahnik & Schindelbeck, 2021).

Meanwhile, Patayon and Crisostomo (2021) captured images of infected abaca leaves using a QNet Aston A1 mobile with 2-megapixel cameras and a Canon 77D DSLR with a 24.2-megapixel Advanced Photo System type-C, Complementary Metal-Oxide-Semiconductor (APS-C CMOS) sensor. These experiments were conducted in outdoor

sampling sites in the Zamboanga Peninsula. The research showed that deep learning algorithms achieved a minimum accuracy of 90% on images focused solely on leaves. However, the accuracy dropped to 80% for images with natural conditions and backgrounds. Standard digital cameras often struggle with backlit scenarios, resulting in underexposed objects and overexposed backgrounds (Afifi et al., 2020).

Furthermore, Pujari et al., (2016) used a color camera (DXC-3000A, Sony, Tokyo, Japan) to capture images of diseased leaves from various crops, including maize, cucumber, cotton, lime, grape, tomato, sugar beet, soybean, and wheat, at the University of Agricultural Sciences in Dharwad, India. By combining color and texture features with a SVM classifier, they achieved an average classification accuracy of 92.17% for plant diseases. Some images had uneven lighting, referred to as shade, due to camera tuning inaccuracies. Digital cameras often face challenges with outdoor lighting in direct sunlight, leading to harsh shadows and suboptimal images (Zhou et al., 2019).

Rehman et al., (2020) used an Android phone with a 640×480 resolution camera to gather images of diseased leaves from various plant species, including tomato, guava, coffee, and paddy plants. They achieved a 95% classification accuracy using a multi-class SVM. Addressing environmental conditions and distortion removal posed challenges, especially under low illumination when the leaf couldn't be detected. To overcome this, a flashlight was used to capture images effectively. Consistent lighting is crucial to prevent shadows and distortions. Limited light often causes phone cameras to reduce image detail, resulting in a loss of sharpness and fine textures (Morikawa et al., 2021). While flashlights provide supplementary light, they can cause overexposure, unnatural shadows, and color inaccuracies due to their harsh lighting properties (Golmohammadi et al., 2021).

On the contrary, Abisha et al., (2023) used a Nikon digital camera Z5 kit with a Nikon Z 24-200 mm f/4-6.3 lens to capture images of diseased brinjal leaves in RGB format. They classified the leaves using DCNN and a radial basis function neural network (RBFNN) based on disease type. The DCNN achieved a mean accuracy of 93.30% with fusion and 76.70% without fusion, compared to the RBFNN, which achieved 87% with fusion and 82% without fusion. The brinjal leaves were sourced from various farms in Tamil Nadu and Kerala, India. The images were captured in an agricultural farm with ample natural sunlight (outside temperature: 29°C, no rainfall, humidity: 82%, wind speed: 70 kph). Nikon cameras often require neutral density (ND) filters for outdoor shooting in direct sunlight, but not all models have built-in ND filters, necessitating the purchase of additional filters and adapters (Vorenkamp, 2019).

Additionally, Salih et al., (2020) captured images of sartorial leaf spot, mosaic, yellow curved, and healthy tomato leaves using a 5-megapixel camera in a real farm setting. Their proposed classification system, based on a DLCNN, achieved an accuracy of 96.34% in detecting the dataset of tomato plant leaves. All images were in RGB format and sized 145 × 145 × 3. While Red, Green, Blue (RGB) images are effective, they can struggle in poor lighting conditions due to their reliance on visible light and are susceptible to adverse weather conditions like fog, rain, or haze, which can obscure features and compromise image quality (Zhang et al., 2023).

Sabrol & Satish (2016) employed a standard digital camera to gather a dataset of 383 digital images of various tomato plant diseases, including bacterial leaf spot, fungal septoria leaf spot, fungal late blight, healthy leaves, bacterial canker, and tomato leaf curl. They used the disease classification tree technique, achieving an accuracy of 93.7%. The image quality was crucial for this model as each pixel was compared to two neighbouring pixels (two

columns to the right and two rows down), with these pairs used both to enhance the output image and as a test set for classification accuracy. The study did not disclose the location and conditions of the photo captures, but it noted that the camera did not yield high-quality images, resulting in only satisfactory accuracy. This limitation is likely due to many digital cameras, especially point-and-shoot or smartphone cameras, lacking control over depth of field (Fatimi, 2021). This can make capturing precise images of specific areas of a tomato leaf challenging, particularly when the disease affects both the surface and the underside of the leaf.

Sujatha et al., (2021) captured images of citrus leaves using a Digital Single-Lens Reflex (DSLR) at 72 dpi with dimensions of 256 × 256 pixels. The study gathered datasets of diseased and healthy citrus leaves and fruits from a farm area. Machine learning algorithms (SVM, random forest, and stochastic gradient descent) and deep learning algorithms (Inception-v3, Visual Geometry Group (VGG-16), and VGG-19) were developed to construct a predictive model based on the training dataset. The algorithm achieved accuracy rates of 76.8% for Random Forest, 86.5% for stochastic gradient descent, and 87% for SVM. However, the camera's performance in farm environments was compromised by dust, mud, and damp conditions, which could infiltrate camera components, leading to malfunctions or deterioration in image quality. Rumchev et al., (2019) emphasized the susceptibility of cameras to the detrimental effects of dust and moisture in farm environments, ultimately compromising functionality and image performance.

Picon et al., (2019) captured original images in a real agricultural setting using a cell phone to identify multi-crop plant leaf diseases. The study focused on a challenging dataset with seventeen diseases across barley, rice, maize, wheat, and rapeseed. The crop identification (ID) was sent to CropNet, an innovative CNN algorithm that employs a crop-

conditional architecture and contextual metadata. The CNN classifier achieved accuracy ranging from 78% to 87%. The dataset included various disease phases, where multiple diseases could appear in a single image. Combining multiple images into one may result in a loss of fine details, complicating the precise identification of specific diseases or symptoms. While CNNs are effective at analysing individual images for patterns, merging multiple images can obscure the nuances needed for accurate diagnosis. The limited focusing capabilities of cell phone cameras worsen this issue, particularly when capturing leaves or disease signs with varying sizes and details, hindering the network's ability to extract precise information (Tanil & Yong, 2020).

Nahiduzzaman et al., (2023) investigated leaf diseases of mulberry in Bangladesh, specifically leaf rust and leaf spot. Their dataset included diseased leaves from various regions, supplemented by 6,000 synthetic images captured with a digital camera in settings like greenhouses, laboratories, and natural environments. However, typical digital cameras struggle to capture the vividness of colors and intricate details in challenging lighting conditions, affecting the quality of synthetic images (Kadambi et al., 2014). They developed a lightweight parallel depth-wise separable convolutional neural network (PDS-CNN), which uses depth-wise separable convolutional layers to improve classification performance while minimizing parameters and size. The model's interpretability was validated through SHapley Additive exPlanations (SHAP) by sericulture specialists, achieving an accuracy of 95.05 ± 2.86% for three-class classifications.

The cassava leaf images in the experiment conducted by Ramcharan et al., (2017) were captured using a A 20.2-megapixel Sony Cybershot digital camera was used in research fields at the International Institute of Tropical Agriculture (IITA) near Bagamoyo, Tanzania. However, some users may find Sony camera phones to be relatively more expensive than

similar devices from other manufacturers (Rakib et al., 2022). The dataset consisted of cassava leaves centered in the frame. The three machine learning algorithms, SVM, the original Inception softmax layer and k-NN were used as training classifiers for the final layer of the Inception v3 CNN model. The accuracy of leaf classification varied from 73% to 91% based on correct classifications.

Haque et al., (2022) introduced a deep learning approach for identifying diseases in maize crops, specifically targeting maydis leaf blight, turcica leaf blight, and banded leaf and sheath blight. Images were captured non-destructively in experimental fields at the Indian Institute of Maize Research in Ludhiana, India, using digital cameras and smartphones during both Kharif and Rabi seasons. To address class imbalance, artificial images were generated through brightness enhancement methods, highlighting the limitations of these devices in capturing balanced and clear images under varying lighting conditions. Using an Inception v3 classifier as a baseline, the model achieved an overall classification accuracy of 95.99% and an average recall of 95.96% on a separate test dataset. However, the model's reliance on artificially generated images and a single baseline classifier raises concerns about its real-world applicability and complexity. Future studies could enhance reliability by incorporating diverse classifiers and focusing on image quality improvement techniques.

Kaushik et al., (2022) explored using CNNs to identify apple, tomato, and grape leaf diseases in images taken in natural settings in India. The dataset consisted of pictures of healthy and unhealthy leaves sorted by species and disease, captured with a DSLR camera and stored at 256 × 256 pixels. The model was trained with images of varying brightness levels, enhancing its adaptability to illumination changes which is a common disadvantage of DSLRs, which may struggle with automatic exposure settings, leading to overexposed or

underexposed images in rapidly changing light conditions (Bernacki, 2020). The study aimed to identify plant diseases and reduce financial losses, utilizing three key neural network types: faster RCNN, RCNN, and single shot multibox detector (SSD). The model achieved an accuracy of 94.6% in identifying different disease types.

Hong et al., (2020) emphasized the importance of tomatoes as a vital crop in China, despite their susceptibility to diseases and pests during growth. They proposed a deep learning model to assist in tomato pest control, which is crucial for diagnosing diseases and pests. The dataset included both diseased and healthy leaves representing nine common diseases, encompassing 18 classifications based on disease and pest categories, and was captured using a smartphone camera. However, the large number of classes resulted in redundancy, complicating effective disease classification. To address this, data augmentation techniques were employed to generate new, varied instances from existing data, promoting diversity and reducing redundancy. The dataset was then trained and tested using various deep CNN models, including ResNet50, Xception, MobileNet, and DenseNet121_Xception, with the proposed model achieving the highest accuracy at 97.10%.

Nesarajan et al., (2020) recognized the crucial role of coconuts as a major crop in Sri Lanka and developed a strategy to address pest infestation and nutritional deficiencies in coconut plants. They generated a dataset by capturing images of both healthy and unhealthy leaves using a digital camera, followed by manual processing to enhance accuracy and reduce complexity. SVM emerged as a key technique for leaf identification, achieving an accuracy of 83.90%. In scenarios where structural conditions produce intricate patterns in camera sensor readings, SVM's ability to capture these nuances improves its effectiveness in distinguishing normal behaviour from potential anomalies, contributing to proactive structural monitoring. However, the research faced challenges due to the limitations of

CNNs in extracting detailed leaf features, as common digital cameras focus on general imaging quality rather than the specifics needed for precise plant disease diagnosis. Additionally, the CNN's performance was further hampered by variations in lighting, angles, and backgrounds in the captured images.

Joshi et al., (2022) presented a method for identifying plant diseases using digital image processing and machine learning algorithms. The study focused on leaves from various crops, including apple, corn, grapes, potato, and tomato. The dataset was sourced from the Plant Village image website, where images were captured with a Sony DSC-RX100/13 20.2 MP camera. However, the Plant Village website is no longer available, and users are now directed to Kaggle for similar content. While Kaggle offers datasets, it may be less user-friendly for browsing plant images, as it is primarily known for data science competitions. For classification, the random forest (RF) method was employed, addressing the common issue of overfitting in decision trees by using a random decision forest, which is an ensemble of multiple decision trees. The model achieved an accuracy of 93.7%.

Preetham et al., (2022) investigated methods for identifying three types of rice plant diseases, along with healthy leaves. Images were captured using a Canon EOS 3000D digital camera in a paddy field located in Neduvasal, Pudukkottai District, India. After applying K-means segmentation, features were extracted based on color, shape, and texture to create the final classification model. The overall accuracy for identifying diseases such as bacterial leaf blight, brown spot, healthy leaf, and rice blast was reported as 95.20%, 95.80%, 94.20%, and 95.60%, respectively, with brown spot achieving an accuracy of 95.40%. While the model relies heavily on colour features for disease identification, it may struggle in extreme lighting conditions or diverse weather scenarios.

Vasavi et al., (2023) utilized a dataset of approximately 1,157 images of capsicum crops captured with a Mi Note 7 Prime in Zaffergadh, Telangana, India, to develop a model for identifying four distinct crop ailments: bacterial leaf spot, fusarium, leaf curl, and pests, as well as healthy leaves. Their research assessed the performance of several machine learning algorithms, including RF, AdaBoost, Gradient Boosting (GB), and Multi-Layer Perceptron (MLP), to determine the most accurate model for predicting chili crop diseases. The experimental results indicated that the RF and GB algorithms achieved accuracies of 96% and 94%, respectively. However, while these accuracies are noteworthy, it is important to consider other evaluation metrics, such as precision, recall, and F1-score, especially in imbalanced datasets where one class may significantly outnumber others. Relying solely on accuracy can be misleading, as it may bias the results towards the majority class (Johnson & Khoshgoftaar, 2019).

Duela et al., (2023) focused on identifying various capsicum diseases, including leaf spot, whitefly, healthy, and yellowish leaves, through images captured with smartphone cameras. The study employed meticulous pre-processing techniques, including resizing, normalization, and filtering, before inputting the images into the classifier. Feature extraction utilized the Gray Level Co-occurrence Matrix (GLCM), examining texture characteristics such as contrast, entropy, homogeneity, and correlation.

The performance of six classifiers was rigorously assessed, yielding accuracy metrics as follows: decision tree (89.92%), neural network toolbox (91.67%), Adam optimization (95%), hybrid intelligent system (92.10%), 2 Level SFAM (90%), and a cutting-edge region-based CNN (96.75%). The proposed model achieved a commendable overall accuracy of 90% in accurately identifying various leaf disorders. However, while the results are promising, the research could benefit from a more comprehensive evaluation that includes

metrics beyond accuracy, such as F1 score, sensitivity, specificity, precision, and recall, to better understand the model's effectiveness across different dimensions.

2.4.2 Kinect Camera

While traditional digital cameras have long been the primary tools for capturing visual data, the introduction of Kinect cameras represents a significant advancement. These devices go beyond capturing 2D images by discerning the three-dimensional structure of objects within their field of view, thereby enhancing accuracy and precision in object detection. Kinect technology has been widely applied across various fields, including facial recognition (Shotton et al., 2011), medical imaging and rehabilitation (Clark et al., 2013), quality control in manufacturing (Tong et al., 2012), and archaeological documentation (Forte, 2014).

In recent years, there has been growing interest in using RGB Kinect cameras for various applications, including face recognition and health monitoring. Researchers have highlighted the effectiveness of these cameras in tracking health outcomes, such as bones and facial features. For instance, advancements in vision systems utilizing RGB Kinect cameras have demonstrated significant improvements in quality control and manufacturing processes, with notable enhancements in metrics such as the mean Intersection over Union (IoU) for quality assessment. The application of RGB Kinect technology not only facilitates reliable data collection across different fields but also proves beneficial for applications requiring precise tracking and analysis, underscoring its potential as a valuable tool in both healthcare and manufacturing settings (Siena et al., 2018).

In agriculture, a significant portion of the existing literature is concentrated only on leaf mapping that involves the analysis and documentation of leaf characteristics, such as size, shape, and venation patterns, typically utilized for botanical research and ecological

studies. For instance, Xia et al., (2015) utilized a Kinect RGB camera with resolutions set at 640 × 480 pixels to capture vegetation within a greenhouse environment, aimed at plant mapping. Following this, they employed active contour models to segment individual leaves from occlusions, with automatic initialization based on centre of divergence calculations. Despite heavy occlusions, their method successfully achieved segmentation rates exceeding 90% for individual leaves.

Andujar et al., (2016) utilized a Kinect RGB camera to estimate weed volume by generating 3D point cloud reconstructions of weed-infested maize fields under real-world conditions. Due to the weeds' relatively short stature, RGB-based recognition was necessary to distinguish them from the soil's microrelief, resulting in a strong correlation (0.83) with weed biomass. Additionally, weed density demonstrated a high correlation with volumetric data, while canonical discriminant analysis effectively classified monocots and dicots. These results suggest that Kinect-based volume estimation is a highly precise method for evaluating crop conditions and identifying weed presence.

The survey on the use of Kinect on different fields suggesting a gap in exploration of alternative topics while empirical evidence indicates a significant enhancement in the precision and effectiveness of object detection algorithms with the integration of depth-sensing cameras. Consequently, the adoption of this technology in agriculture, especially for identifying pathological conditions affecting plant leaves, warrants careful consideration as findings indicate that utilizing Kinect for leaf mapping and volume estimation can serve a highly accurate approach.

2.4.3 Hyperspectral Imaging

Hyperspectral imaging is a technology that captures an extensive range of colours, going beyond the visible spectrum. Unlike standard cameras, it records detailed information in numerous narrow and contiguous bands of the electromagnetic spectrum.

Nagasubramanian et al., (2019) identifying soybean diseases using hyperspectral imaging method, specifically the Pika XC hyperspectral line imaging scanner. Positioned 54 cm away from the stem sample, the Pika XC Imager captured 240 wavebands ranging from 400 to 1000 nm. Then, with the information in hyperspectral data cubes, a 3D DCNN was introduced for disease identification, focusing on rust, Septoria leaf blotch, brown spot, and downy mildew in soybean crops. The developed model, evaluated on a test dataset of 539 images, achieved an impressive classification accuracy of 95.73%, with a recall value of 92% and an F1-score of 87%, offering meaningful insights into the physiological aspects affecting soybean crop yield. Nevertheless, hyperspectral cameras demand higher power consumption in comparison to other camera types, impacting their appropriateness for applications in remote scenarios.

Atas et al., (2011) have proposed a machine vision system that is compact and relies on hyperspectral imaging and artificial neural networks (ANN) for the identification of mold growth in capsicum leaves caused by aflatoxin contamination. Hyperspectral images from 53 distinct capsicum samples were captured using a FireWire Sony charge-coupled device (CCD) camera equipped with a Varispec liquid crystal tunable filter, illuminated by both a 100W quartz-tungsten-halogen source and UV 365 nm light. The classification of capsicum into aflatoxin-contaminated and uncontaminated classes was performed by employing difference images of consecutive spectral bands, along with individual band energies. Both ultraviolet (UV) and halogen illumination sources were utilized in the experiments. The

selection of significant bands for improved discrimination was based on their neural network connection weights. The highest classification accuracy rate of 91% was achieved by utilizing halogen excitation and selecting features through an MLP-based approach. However, hyperspectral camera necessitates multiple scans or setup adjustments to achieve comprehensive coverage of a broader area in a single image because it has limited field of view (Morales et al., 2022).

2.4.4 Thermal Imaging

Thermal camera is an imaging device designed to capture and visualize infrared radiation emitted by objects based on their temperature. Operating beyond the visible light spectrum, these cameras provide detailed thermal images, allowing for the identification of temperature variations in diverse applications such as industrial inspections, research, and medical diagnostics.

Chelladurai and Sujatha (2023) introduced a machine vision K-Nearest Neighbor (KNN) algorithm applied to images captured using a handheld thermal camera for disease detection in vegetative leaf types. Handheld thermal cameras discern vegetative leaf diseases by capturing infrared radiation emitted by the vegetation. Temperature variations, indicating stress or disease, are visually depicted in thermal images, enabling the identification of hotspots or cold spots on affected leaves. Continuous monitoring and analysis of these temperature patterns offer a non-invasive approach to assess plant health and identify potential diseases in their early stages. The 934 vegetative leaf images were categorized into ten groups, comprising one class for healthy leaves and nine for different diseases, including early blight, mosaic virus, leaf mould, target spot, Septoria leaf spot, late blight mould, two-spotted spider mite, yellow leaf curl virus, and bacterial spot. These nine diseased classes were further grouped into five subcategories which are mould, viral, fungal, bacterial, and

mite diseases. The resulting accuracy achieved by the algorithm is 89.4%. The algorithm heavily relies on capturing infrared radiation (Gu & Meng, 2022) emitted by vegetation and this dependence may lead to challenges in environments with insufficient or variable infrared radiation.

Sihombing et al., (2021) employed a k-NN algorithm with Euclidean and Manhattan distance calculations to detect capsicum leaf diseases, including leaf curl, veinal mottle, alfalfa mosaic and beet curly top. Flix AX8 thermal camera captured the images, and the evaluation utilized 210 training images and 90 test images. Conducted in a farm setting, thermal imaging facilitated the identification of temperature anomalies on capsicum leaves, aiding in the early detection of stress or disease conditions. The image registration process involved aligning overlapping image scopes, with one image serving as a reference. The model achieved a detection performance of 85.4% recall, 89.35% precision, 87.33% F1 score and 92.8% accuracy for identifying capsicum leaf diseases. The model's reliance exclusively on temperature variations for disease detection may result in the omission of visual symptoms or molecular changes unrelated to temperature, potentially leading to false negatives and reducing diagnostic accuracy.

2.4.5 Image Quality Metrics

Assessing image quality is crucial in image-based plant disease detection, as it determines the clarity and accuracy of the input data used by detection models. Three commonly used objective metrics for evaluating image quality are the Mean Square Error (MSE), Relative Error (RE) and Euclidean Distance (ED). These metrics quantify the level of distortion or difference between an original (reference) image and a processed or reconstructed one.

(a) Mean Square Error (MSE)

The Mean Square Error measures the average squared difference between the original image I and the reconstructed or processed image K. It is mathematically defined as:

$$MSE = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\ [I\ (i,j) - K\ (i,j)]^2 \qquad \text{Equation 2.2}$$

where M and N denote the image dimensions. A lower MSE value indicates higher image quality and minimal distortion (Huynh-Thu & Ghanbari, 2008). In plant disease detection, minimizing MSE ensures that fine leaf details such as lesion colour and texture are preserved, improving feature extraction reliability.

(b) Relative Error (RE)

Relative Error represents the ratio of the total absolute difference between the two images to the total intensity of the reference image. It is expressed as:

$$RE = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N}|I(i,j)-K(i,j)|}{\sum_{i=1}^{M}\sum_{j=1}^{N}|I(i,j)|} \qquad \text{Equation 2.3}$$

A smaller RE value signifies a closer match between the compared images, implying effective preservation of important leaf features after preprocessing or compression (Mittal et al., 2012).

(c) Euclidean Distance (ED)

Euclidean Distance is used to measure the dissimilarity between two image vectors in feature space. It is calculated as:

$$\text{ED} = \sqrt{\sum_{i=1}^{M}\sum_{j=1}^{N}[I(i,j) - K(i,j)]^2} \qquad \text{Equation 2.4}$$

This metric provides an intuitive distance-based measure of overall image difference (Gonzalez & Woods, 2018). In agricultural image analysis, ED helps quantify how much an enhanced or augmented image deviates from the original, ensuring that critical disease features remain intact during processing. Together, these metrics play a fundamental role in validating the visual fidelity of preprocessed and augmented images before inputting them into classification models. Accurate image quality evaluation ensures that the trained models rely on high-quality data, leading to improved disease classification performance.

While the F1 score is an important evaluation metric for classification problems, particularly in scenarios involving imbalanced datasets, it is mathematically defined as the harmonic mean of precision and recall, given by:

$$F1\ scor = {2 \times (Precision \times Recall)}/{Precision + Recall} \quad \text{Equation 2.5}$$

Where recent studies in machine learning indicate that the F1 score offers one of the most consistent and balanced evaluation metrics across various models and dataset conditions, outperforming accuracy and other traditional measures, especially in tasks with imbalanced class distributions (Sujon et al., 2025). Unlike accuracy, which may provide a misleading indication of performance when class distributions are uneven, the F1 score effectively captures the trade-off between precision and recall, offering a more reliable and meaningful evaluation of model effectiveness under real-world conditions with skewed class proportions.

Precision measures the proportion of correctly identified positive samples such as diseased leaves among all samples predicted as positive, while recall, also referred to as sensitivity, measures the proportion of actual positive samples that are correctly detected by the model. High precision indicates a low number of false positive predictions, whereas high

recall indicates a low number of false negative predictions, and both metrics are vital in agricultural disease detection to avoid unnecessary interventions and to prevent missed disease occurrences.

The F1 score is mathematically defined as the harmonic mean of precision and recall. It is calculated by taking twice the product of precision and recall and dividing it by their sum, representing the harmonic mean of these metrics to assess the balance between accurately identified positive samples and overall prediction performance, especially in imbalanced classification tasks. The use of the harmonic mean ensures equal contribution from both precision and recall, and the F1 score improves only when both values are high, making it particularly suitable for evaluating classifiers with uneven class distributions.

In this study, the F1 score is employed together with accuracy, precision, and recall to provide a comprehensive evaluation of the proposed improved ShuffleNet model. Incorporating the F1 score ensures that the model's effectiveness in correctly identifying disease cases is balanced with its ability to minimize false positives, enabling a robust assessment of performance across different plant species and disease types.

2.5 Plant Leaf Diseases Detection Models

Machine learning spans a broad domain, featuring numerous algorithms tailored to teach computers to learn from data and derive predictions or decisions based on acquired knowledge. Deep learning, on the other hand, represents a distinct and rapidly expanding subset within machine learning, emphasizing neural networks distinguished by multiple layers, often termed deep neural networks. The surge in deep learning's popularity stems

from its exceptional performance across diverse tasks such as image recognition and natural language processing, capturing significant attention in recent years.

SVM, a machine learning algorithm, is useful for identifying plant diseases by mapping leaf image features into a high-dimensional space, as shown in Figure 2.6 (Harakannanavar et al., 2022). Harakannanavar et al., (2022) developed an SVM model to identify 600 images of infected tomatoes, achieving a precision of 96.70%, recall of 97.80%, F1-score of 91.50% and accuracy of 89%. SVM analyzes unique leaf characteristics like color, size, and shape, called leaf attributes. It aims to find a separating line in this space, described by the equation:

$$f(x) = \Sigma(\alpha_i * \ y_i * K(x_i, x)) + b \qquad \text{Equation 2.6}$$

where $f(x)$ is the SVM decision function, a_i is the Lagrange multiplier for the i-th support vector, y_i is the class label for the i-th support vector, $K(x_i, x)$ is the kernel function measuring similarity, and b is the bias term. Leaf samples undergo contour tracing to capture their shape, and discrete wavelet transform (DWT) is used to extract useful information. The best features from DWT are selected using the GLCM. The features from both DWT and GLCM are combined into a feature vector for classification. The study suggests that using fusion techniques could improve stability, as using a Daubechies wavelet type and setting the GLCM angle from 1 to 0 led to an unstable feature vector.

However, SVM has some limitations. It struggles with large datasets, which are common in plant leaf disease detection. It also does not perform well with noisy or overlapping classes, which are typical in disease detection. SVMs require careful tuning of hyperparameters and kernel functions, making them sensitive to parameter choices. Lastly,

SVMs do not handle imbalanced datasets well, leading to biased predictions if some diseases are more common than others.

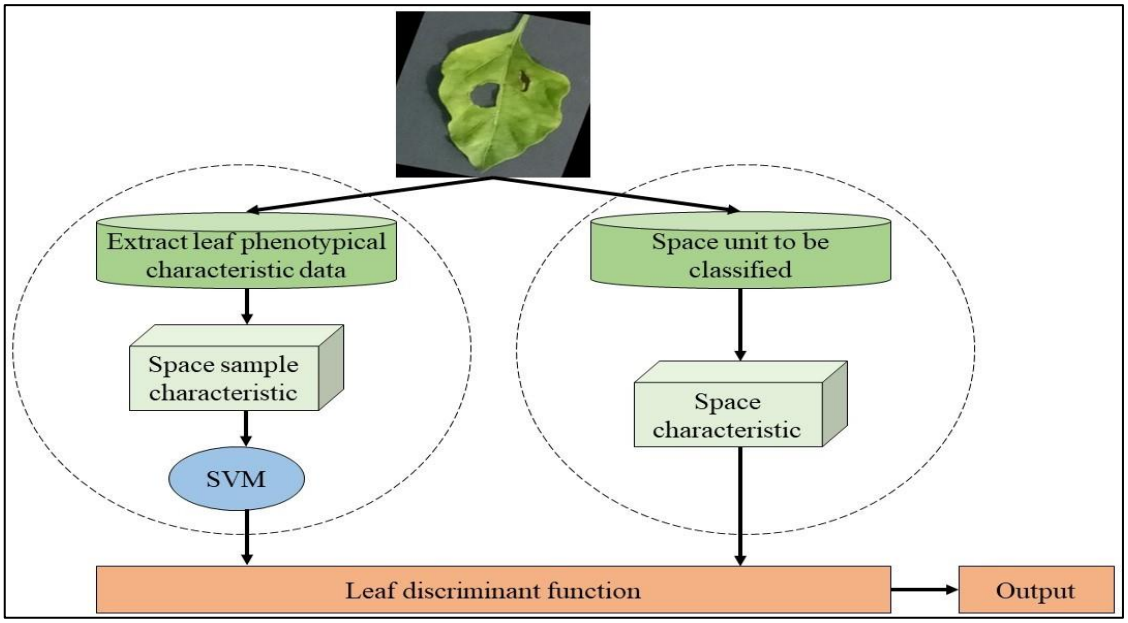


Figure 2.6: SVM model (Harakannanavar et al., 2022)

Ensemble learning methods like Random Forest predict by combining many decision trees, as shown in Figure 2.7. Chauhan et al., (2021) developed a Random Forest model for leaf disease detection, achieving 80.68% accuracy, better than models like Naïve Bayes, decision tree (DT), KNN, and SVM. Random Forest uses multiple predictors to identify maize leaf diseases, randomly selecting characteristics for decisions. This approach reduces overfitting and increases robustness. It combines tree predictions using:

$$f(x) = mode(T1(x), T2(x), \ldots, Tn(x)) \qquad \text{Equation 2.7}$$

where $f(x)$ is the random forest prediction, $T1(x)$, $T2(x)$, ..., $Tn(x)$ are individual tree predictions, and "mode" is the most frequent prediction. Each model has limitations with

specific parameters, so it is best used with large datasets (more than 765 images) for better accuracy.

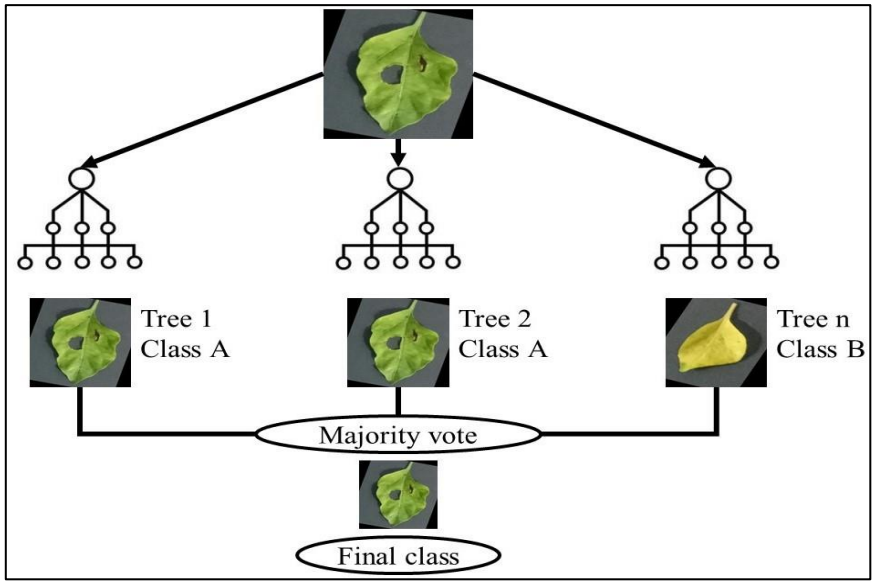


Figure 2.7 Random Forest model (Chauhan et al., 2021)

Figure 2.8 shows that a decision tree has decision and leaf nodes. It classifies or predicts outcomes by evaluating specific attributes. Priyaradhikadevi et al., (2023) developed a decision tree model for classifying 346 images of apple, grape, and tomato leaf diseases, achieving 69.89% accuracy. The model starts with a root node based on an attribute like color and refines classification through internal nodes, considering attributes like spots or shape. This ends at leaf nodes, which specify the disease. The optimal question for splitting data is determined by information gain (IG), calculated as:

$f(x)$ = Entropy (Starting Group) - Weighted Average of Child Groups' Entropy

Equation 2.8

where $f(x)$ is the information gain. Entropy is calculated as:

$$\text{Entropy (Group)} = -p_{\,1} * log^2(p_1) - p_2 * log^2(p_2) - \ldots - p_{\,k} * log^2(p_k)$$

Equation 2.9

where p_1 is the proportion of data points in a class. The Weighted Average of Child Groups' Entropy is obtained by summing the entropy of each child group multiplied by the proportion of samples in that group relative to the parent node:

$$\text{Weighted Entropy} = \sum_{i=1}^{k} \frac{|S_i|}{|S|} \times \text{Entropy}\,(s_i) \qquad \text{Equation 2.10}$$

This ensures that larger child groups have a greater influence on the overall entropy calculation. A split is considered better when the weighted entropy is lower, resulting in higher information gain. This helps select the best question (x) to split the data. The study also suggests using more reliable feature extraction techniques than GLCM, as larger textures become more uniform at a larger scale.

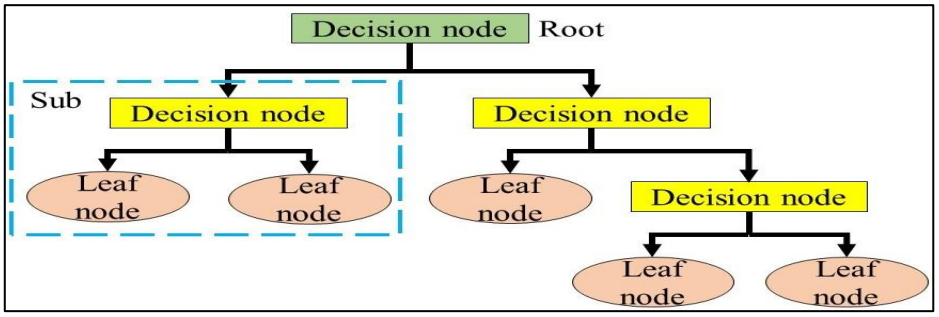


Figure 2.8: Decision tree model (Priyaradhikadevi et al., 2023)

Naïve Bayes is a probabilistic machine learning approach, shown in Figure 2.9. It uses a training set of labeled data to build a classification model, which is then tested for accuracy, precision, and recall. Reddy and Adimoolam (2022) developed a Naïve Bayes

model with 2500 leaf photos from 39 disease categories in 26 crops, achieving 60.40% accuracy. The dataset is split into training and testing subsets, with 10 iterations for robustness. Using an alpha value of 0.005, the Naïve Bayes model applies probabilistic reasoning and evaluates its performance with accuracy, precision, and recall. The equation for the Naïve Bayes model is:

$$P(y|f(x)) = [P(y) * P(f(x)|y)] / P(f(x)) \qquad \text{Equation 2.11}$$

where $P(y|f(x))$ is the probability of class y (disease) given the input data $f(x)$. $P(y)$ is the prior probability of class y, $P(f(x)|y)$ is the likelihood of observing $f(x)$ given class y, and $P(f(x))$ is the marginal likelihood of $f(x)$. Naïve Bayes assumes that the prior probabilities of each class are known and constant. If these probabilities are unknown or change, it can affect classification accuracy. The study focuses only on accuracy, which is a limitation.

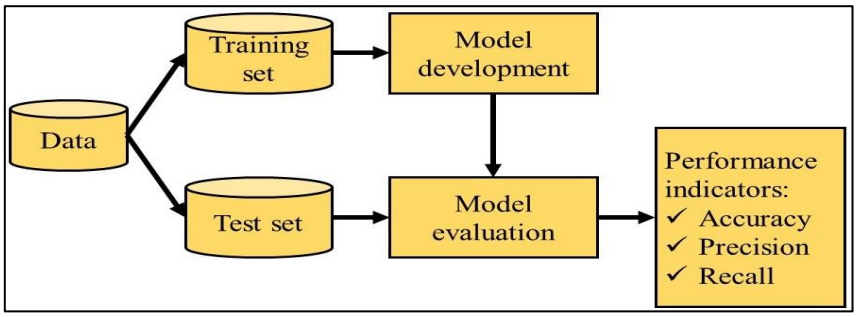


Figure 2.9: Naïve bayes model (Reddy and Adimoolam 2022)

The BoF approach extracts features by quantizing local descriptors from an image, creating a "bag" of vectors to represent the image, as shown in Figure 2.10. Aabidi et al., (2023) and Setiawan et al., (2022) used BoF to detect apple and maize leaf diseases, achieving 70.08% and 93.60% accuracy, respectively. In BoF, texture, color, and shape are

organized in a high-dimensional vector space. These features are quantized using a codebook created through clustering. The BoF technique equation is:

$$f(x) = \sum_{i=1}^{N} w_i \cdot D_i(x) \qquad \text{Equation 2.12}$$

where $f(x)$ is the analysis function of the leaf image, $\sum_{i=1}^{N} w_i$ is the sum of codewords over N features, w_i is the weight for each feature, and $D_i(x)$ is the i-th extracted feature. The authors noted that BoF depends on handcrafted features and lacks spatial information, which can limit its performance with complex and diverse image datasets. Compared to deep learning algorithms like CNN, BoF performs worse (67.7% vs. 88.6% accuracy).

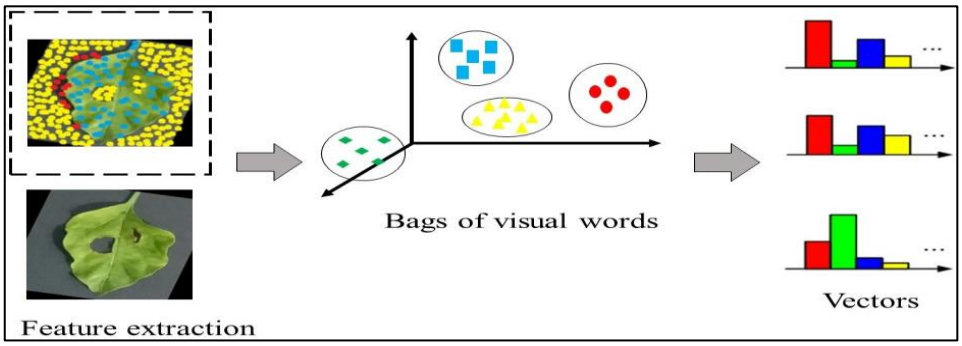


Figure 2.10: BoF model (Setiawan et al., 2022)

MobileNet, shown in Figure 2.11, is an efficient convolutional neural network suitable for resource-limited situations. Nguyen et al., (2023) created a MobileNet model to detect diseases in 2064 tomato leaves, achieving 95.7% accuracy. The images are processed through layers of convolution and depth-wise separable convolutions to extract features. The network, with 1024 layers and average pooling, learns intricate patterns. Depth-wise convolution filters spatially, while point-wise convolution combines these filtered outputs, reducing complexity. The output is given by:

$$f(x, \theta) = PW(DW(Con(x, \theta_1), \theta_2), \theta_3) \qquad \text{Equation 2.13}$$

where x is the input image, θ are the parameters, $Conv(\cdot)$ is convolution, $DW(\cdot)$ is depth-wise convolution, and $PW(\cdot)$ is point-wise convolution. The study recommends data augmentation to improve predictions. The main limitation is the lack of diverse tomato leaf diseases in the training data. MobileNet is efficient but may sacrifice some accuracy compared to larger CNNs, especially for tasks needing fine details or precise object detection.

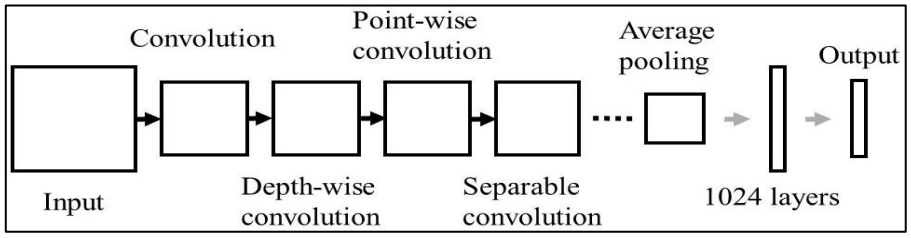


Figure 2.11: MobileNet model (Nguyen et al., 2023)

GoogleNet is a deep convolutional neural network that includes several convolutional and pooling layers, inception modules, and ends with a global average pooling layer, fully connected layers, and an output layer for classification, as shown in Figure 2.12. Jung et al., (2023) used GoogleNet to classify leaf diseases in bell peppers, potatoes, and tomatoes, achieving 97.07% accuracy. The model processes 1662 bell pepper images, 2177 potato images, and 467 tomato images. It starts with a convolutional layer for basic feature extraction, then uses deeper layers for more complex features. Max-pooling layers reduce feature map sizes, and inception modules capture various patterns. Random sampling ensures that some data is reserved for testing. After further max pooling, features are flattened and passed to fully connected layers for classification. The final output is calculated using the softmax function:

$$f(x) = \frac{1}{1+e^{-\sum_{i=1}^{n} w_i x_i + b}}$$ Equation 2.14

where $f(x)$ is the classification output, w_i are weights, x_i are input values, b is the bias, n is the number of inputs, and e is the base of natural logarithms. The authors recommend using diverse datasets in future research to improve the model's applicability, as it currently struggles with false positives and simplified feature representations, limiting its effectiveness in complex environments with varied objects and occlusions.

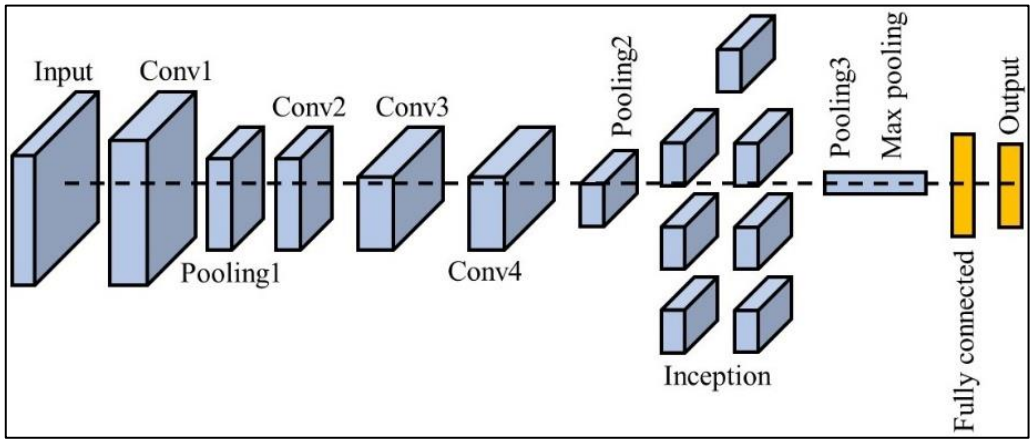


Figure 2.12: GoogleNet model (Jung et al., 2023)

The Inception model uses pointwise convolutions for dimensionality reduction and depth wise convolutions for spatial filtering, as shown in Figure 2.13. Lambat et al., (2022) employed Inception to identify 500 photos of diseased leaves captured in a lab, using generative adversarial networks (GANs) to create additional training samples, achieving an accuracy of 88.73%. The model processes these 500 images, extracting important features through multiple inception modules that address both spatial and dimensional aspects of the leaves.

These convolutions help the network learn and identify distinct patterns and textures in images, allowing it to differentiate between healthy and diseased leaves. In the paper's conclusion, the authors express a desire to improve the model's extensibility to predict more types of diseases. They plan to expand the training to include a wider range of plant diseases across different species, enhancing the model's ability to identify various plant diseases. This expansion would make the model more versatile and valuable in agriculture. However, the model currently overlooks dynamic environmental factors and relies solely on static image inputs.

Tahamid (2020) used ResNet50, a 50-layer deep residual learning framework that tackles the vanishing gradient problem with skip connections, allowing for the training of larger neural networks. The study focused on detecting several tomato leaf diseases, including bacterial spot, early blight, late blight, leaf mold, septoria leaf spot, spider mites, two-spotted spider mite, yellow leaf curl virus, and mosaic, achieving an accuracy of 94%.

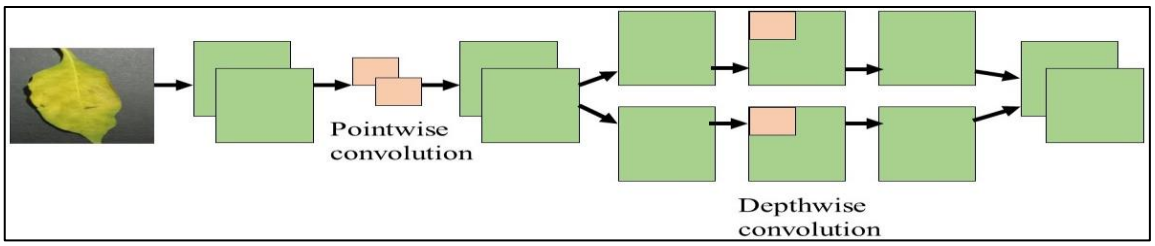


Figure 2.13: Inception model (Tahamid, 2020)

The ResNet50 detection process involves zero-padding input images, using convolutional layers (Conv) with 64 filters, batch normalization (BatchNorm), ReLU activation functions, and max-pooling (MaxPool) with a 3×3 kernel and stride of 2, as shown in Figure 2.14. It includes convolutional blocks with 256 filters, batch normalization, ReLU,

and identity blocks with 64 filters. Global average pooling (AvgPool) and flattening convert the output into a one-dimensional vector, processed by a fully connected (FC) layer with unique weights. The final classification uses the softmax function:

$$f(x) = \frac{e^{x_i}}{\sum_{j=1}^{c} e^{x_j}} \qquad \text{Equation 2.15}$$

where x_i are the output values of the FC layer, and c is the number of disease classes. The authors recommend expanding the model to identify leaf diseases in various plants by updating it with new data, improving training methods, and enhancing the model's architecture for better effectiveness across different species and changing disease patterns.

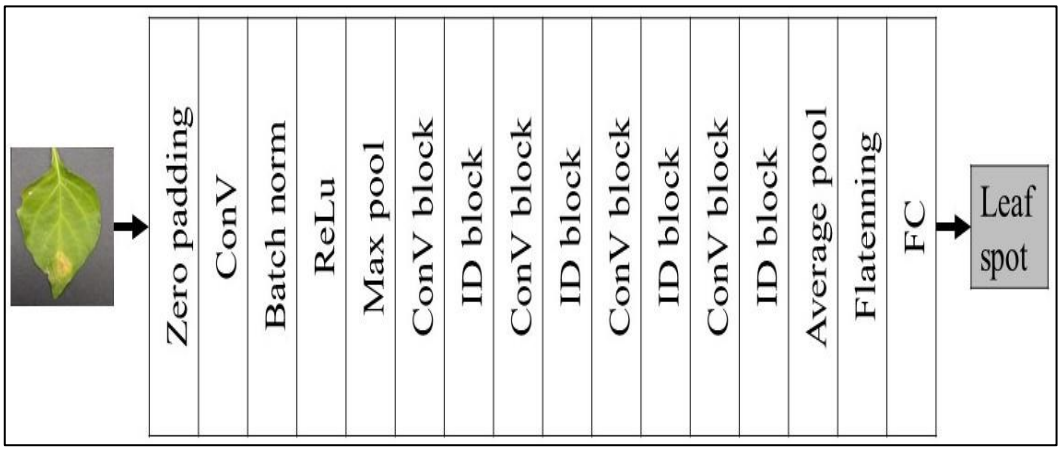


Figure 2.14: ResNet50 model (Tahamid, 2020)

DenseNet201 is a deep learning architecture with 201 densely connected layers that use convolutional layers with a 1×1 kernel, strides of 1, and include components like average pooling, batch normalization, and ReLU activations, as shown in Figure 2.15. Alghamdi and Turki (2023) applied DenseNet201 to a cassava leaf dataset, achieving impressive results: an average precision of 92.06%, recall of 92.71%, F1 score of 92.36%, and overall accuracy

of 93.79%. The images used were 256×256 pixels, and the convolutional layers extracted 64 unique features, followed by max pooling to reduce spatial dimensions.

The architecture also included densely connected layers (DenseBlock) with four dense layers, each having 128 nodes, and a transition layer (TransLayer) that reduced the dimensions to 64. Global average pooling (GlobalAvgPool) condensed the feature maps into a vector of length 64. The authors suggest that the model should incorporate diverse parameters and adapt to different frameworks to improve its flexibility and effectiveness.

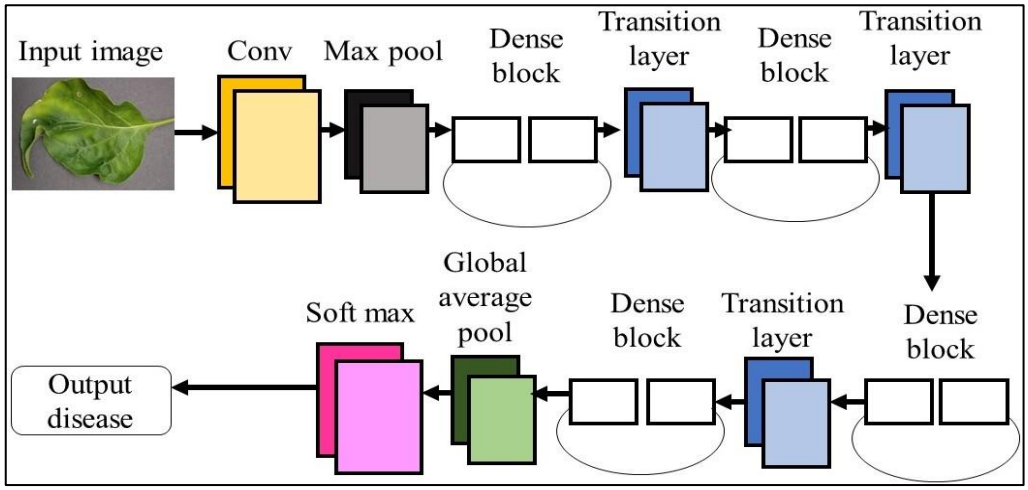


Figure 2.15: DenseNet201 model (Alghamdi & Turki, 2023)

2.5.1 Data Augmentation and Transfer Learning

Diagnosing objects in varying conditions presents a significant challenge in computer vision due to differences in lighting, background, and appearance that can adversely affect image quality. To address these limitations, two commonly employed strategies in detection models are data augmentation and transfer learning. Data augmentation expands the training dataset through transformations such as rotation, scaling, flipping, and color adjustment, thereby mitigating overfitting and improving generalization performance (Shorten & Khoshgoftaar, 2019).

Transfer learning, on the other hand, enables models pre-trained on large-scale datasets to be adapted for new but related tasks, improving accuracy even when only limited training data are available (Tan et al., 2018). In plant disease detection, these approaches are particularly valuable since leaf images frequently vary with environmental conditions, illumination, and background complexity. When combined with fine-tuning of pretrained architectures, both augmentation and transfer learning significantly enhance classification accuracy and model robustness (Ferentinos, 2018; Too et al., 2019).

Table 2.1 presents a comparative analysis illustrating the improvements achieved through the application of various data augmentation techniques in plant leaf disease detection. The reviewed studies consistently demonstrate that augmentation increases dataset diversity and improves the model's ability to generalize across different environmental settings. For instance, Min et al. (2023) and Cap et al. (2020) reported that simple augmentation strategies such as flipping and rotation improved learning efficiency and validation accuracy by enabling models to recognize invariant features under diverse orientations. Similarly, Karnik and Suthar (2021) found that zoom-based augmentation enhanced robustness by allowing models to adapt to scale variations in input images. Muhammad et al. (2023) and Muhammed and Yusoff (2023) demonstrated that random transformations including flipping, resizing, and width shifting helped to address class imbalance and increased precision, particularly for minority disease classes.

More complex augmentation strategies combining flipping, warp shifting, noise addition, and blurring, as proposed by Fawaiq et al. (2023), achieved accuracy improvements exceeding nine percent, confirming the value of compound augmentations in enhancing model resilience under variable lighting and background conditions. Another effective augmentation method is random reflection, which includes both horizontal and vertical

flipping. Xiao (2022) applied this technique to a small dataset containing only ten seed images per class, combined with rotation and scaling, resulting in an accuracy increase from 72% to 90%, representing a 25% improvement. This outcome demonstrates that reflection-based augmentation allows models to better generalize across different leaf orientations and reduces overfitting. Similarly, Gheorghiu et al., (2024) employed random reflection in classifying coffee leaf diseases using the RoCoLe dataset and achieved an improvement from 88.2% to 92.7% accuracy, confirming that reflection enhances robustness to natural variations in leaf position and angle.

Collectively, the findings affirm that combining data augmentation and transfer learning forms a foundational strategy in modern agricultural image analysis. These techniques not only improve classification accuracy and generalization capability but also enable the development of more resilient detection models suitable for real-world deployment in smart farming and precision agriculture.

Table 2.1: A comparative analysis conducted between the absence of augmentation and with the augmentation techniques on plant leaf diseases detection

Data Augmentation	**Augmented images**	**Plant type**	**Fine-tuning**	**Classifier**	**Finding**
Flipping (Min et al., 2023)	500 to 1645	Apple	Learning rate adjustment	Mask-RCNN	Improved accuracy by 10% with data augmentation

Table 2.1: continued

Rotation (Cap et al., 2020)	3042 to 12000	Cucumber	Fine-tuning learning rate and dropout rates	SVM	Data augmentation reduced overfitting, while fine-tuning improved validation accuracy by 7.4%.
Zooming (Karnik & Suthar, 2021)	902 to 3207	Pepper bell	Fine-tuning on LSTM layers	Ensemble of CNNs and LSTM networks	Data augmentation led to a more robust model against variations in input images, improving overall accuracy by 15%.

Table 2.1: continued

Random flipping and resizing with Cropping (Muhammad et al., 2023)	560 to 2060	Grape and tomato	Fine-tuning on attention weights	CNN	Data augmentation enhanced model generalization, and fine-tuning further refined accuracy by 2,3%.
Horizontal flip and width shift (Muhammed & Yusoff, 2023)	868 to 2980	Rice	Fine-tuning on fully connected layers	ANN	Data augmentation mitigated class imbalance issues, and fine-tuning improved precision for minority classes by 12%.

Table 2.1: continued

Flipping, warp shift, adding noise, and blurring image (Fawaiq et al., 2023)	480 to 1440	Rice	Adaptive fine-tuning based on performance metrics	VGG16, NASNetMobile, and Xception	Adaptive fine-tuning improved model robustness across diverse datasets, showcasing a 9.22% increase in accuracy.

To evaluate model reliability, common performance metrics are used. Specificity measures how well the model identifies healthy leaves, while sensitivity (recall) measures its ability to detect diseased leaves. Precision indicates the proportion of correctly identified diseased samples, and the F1-score balances precision and recall. These metrics provide a consistent basis for comparing performance across plant disease studies and are used in this research to assess the improved ShuffleNet model.

2.6 ShuffleNet and ShuffleNet V2

Deep learning has increasingly focused on lightweight models to meet the needs of real-time applications on mobile and embedded devices. ShuffleNet was introduced to address this challenge by combining pointwise group convolution with a channel shuffle

operation. Group convolutions lower computation but restrict information sharing, while channel shuffle solves this by rearranging feature channels. This design achieves high accuracy with much lower FLOPs than traditional CNNs, making ShuffleNet highly efficient on mobile platforms (Zhang et al., 2023).

ShuffleNet's efficiency influenced later model designs, especially for edge computing tasks in areas like precision agriculture, healthcare, and mobile image classification. In agriculture, where disease detection must often be done on-site using handheld devices or drones, ShuffleNet provides a good balance between accuracy and speed, showing that advanced CNNs can work effectively outside high-performance computing environments.

Although ShuffleNet was innovative, researchers found that FLOPs alone did not reflect real-world speed. Issues such as memory cost and operator fragmentation also mattered. To address this, ShuffleNet V2 was developed with a channel-split design, sending part of the features through lightweight convolutions and the rest directly as identity mapping. This reduced memory use and improved throughput while keeping accuracy. ShuffleNet V2 therefore achieved a better speed–accuracy trade-off, especially on mobile hardware (Ma et al., 2018).

Compared with V1, ShuffleNet V2 runs faster and often more accurately on benchmarks like ImageNet. These improvements also apply to specialized tasks such as agricultural image analysis, where both accuracy and efficiency are crucial.

The lightweight nature of ShuffleNet and ShuffleNet V2 makes them well suited for plant disease detection. Field applications often use mobile devices or UAVs with limited computing power, and these models can run directly on such devices without server support.

For example, Yang et al., (2021) used ShuffleNet V2 for grape leaf disease detection with good results, while Zhou et al., (2024) applied an improved ShuffleNet V2 for maize leaf diseases, achieving high accuracy with reduced complexity. Other works have also used ShuffleNet V2 for rice leaf classification, showing its adaptability in agriculture. The ShuffleNet and ShuffleNet V2 are not only efficient in theory but also practical for real-world agricultural use. As plant disease detection moves toward in-field deployment, where resources are limited, these models are strong candidates for enabling accurate, fast, and reliable disease recognition, supporting better crop management and reducing losses.

2.7 Major Observation from Literature Review

A review of the literature shows that advanced imaging tools are essential for accurate plant leaf disease detection. Standard digital cameras often struggle with exposure control under different lighting, producing inconsistent image quality. This highlights the need for more reliable devices, such as the Kinect camera, which can capture images with consistent exposure and improve detection accuracy. The literature also points out that ecological conditions greatly affect disease detection. Yet, many studies do not fully address the challenge of identifying a single diseased leaf hidden among many healthy ones. This gap shows the importance of developing models that combine advanced imaging with knowledge of plant physiology to perform well in real-world field conditions.

Another recurring theme is the value of image processing techniques such as data augmentation, transfer learning, and architectural modifications to CNN models. However, many existing studies apply these methods in a limited way, leading to poor generalization when leaf appearance or environmental conditions vary. Few works explore deeper convolutional layers or improved architectures to make models more robust.

In summary, current research lacks consistent use of Kinect-based imaging and more advanced model designs. This study addresses these gaps by combining an optimized image acquisition system with an improved ShuffleNet CNN, aiming to deliver higher accuracy, stronger generalization, and greater robustness than conventional methods.

2.8 Summary

This chapter explores different types of optical devices employed for data augmentation, with a specific focus on objects and diseased leaves. It also delves into an examination of existing models designed for the identification of plant leaf diseases. A comparative analysis is conducted on these previous models, considering factors such as the devices used for dataset acquisition, classification algorithms, strengths, weaknesses, and overall performance. The subsequent chapter outlines the methodology employed in implementing the proposed model.

CHAPTER 3

METHODOLOGY

3.1 Overview

The previous chapter presented a literature review to provide a comprehensive and critical analysis of the existing models of plant leaf detection. This chapter discusses the systematic and structured approach used to design the research study.

The chapter outline is as follows. Section 3.2 describes the overall methodology. Section 3.3 discusses the dataset used, and Section 3.4 discusses image processing. While Section 3.5 covers the selection of hyperparameters, Section 3.6 covers layer modification, Section 3.7 addresses dataset splitting, and Section 3.8 emphasises model performance evaluation. Finally, Section 3.9 delivers a summary of the chapter.

3.2 Research Methodology

The research methodology used in this study focuses on the detection of plant leaf diseases using an improved ShuffleNet model. Figure 3.1 (for uniform background case) and Figure 3.2 (for complex background case) show the research framework flow chart for this study, from image acquisition to image detection using a proposed method of an improved ShuffleNet model. The entire analysis is conducted using Matlab R2022a on a processing device detailed in Table 3.1. Matlab R2022a is the ninth annual release of MathWorks' Matlab software, with a release date of September 2021 (Bentley et al., 2023). This new release version includes performance optimisations that lead to optimised algorithms, allowing the user to work more efficiently with large datasets and complex calculations, as well as improved integration with other software tools, data formats, and Application

Programming Interface (API), allowing the user to seamlessly work with data from different sources and integrate Matlab into existing workflow (Taha et al., 2022).

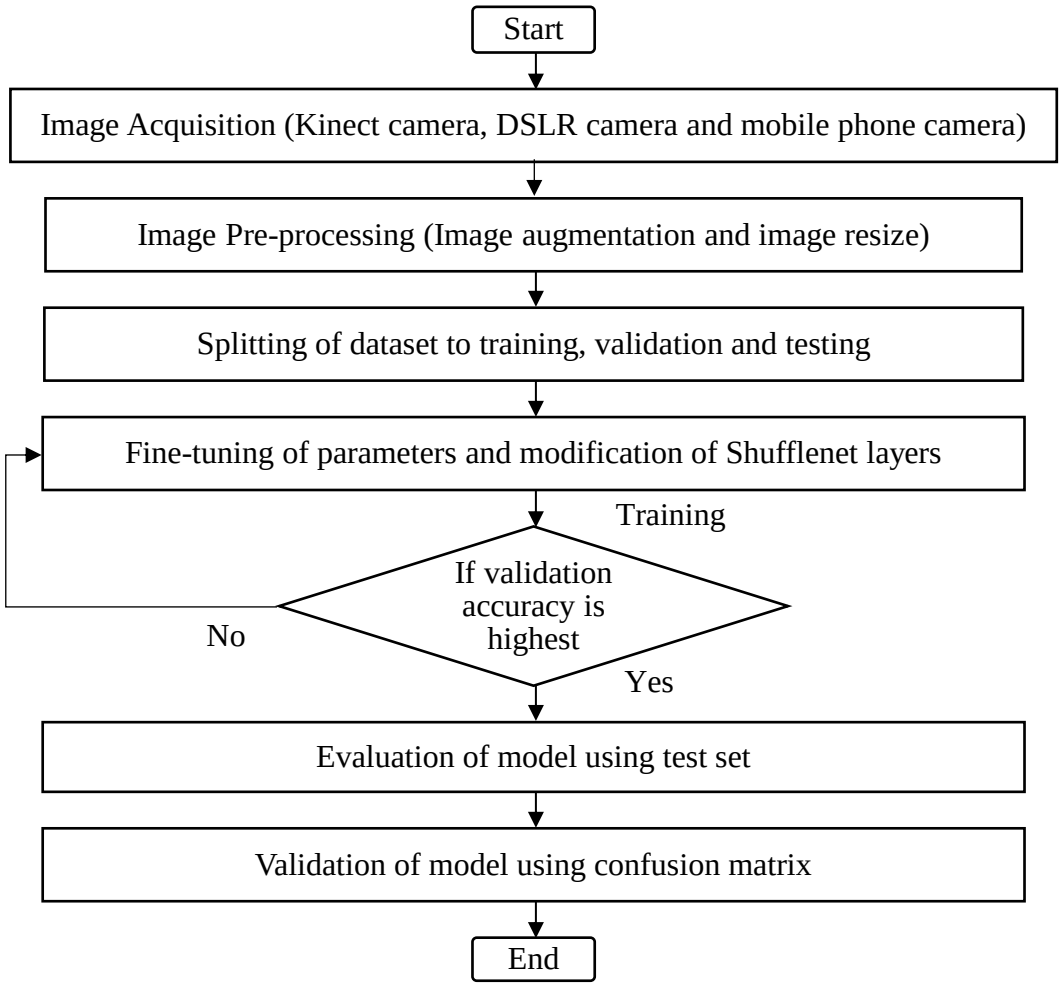


Figure 3.1: Framework for developing an improved ShuffleNet to detect plant leaf diseases (uniform background)

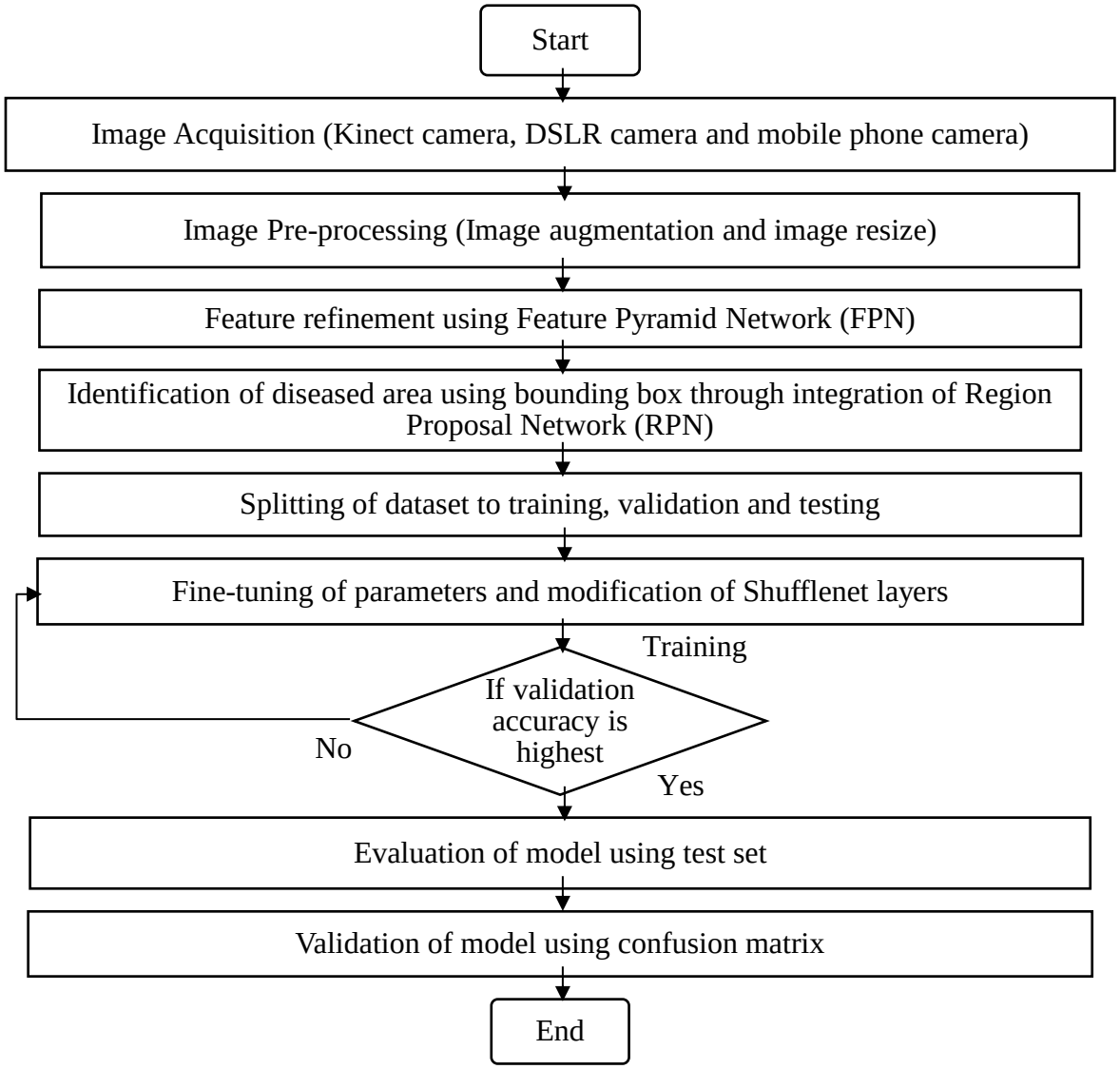


Figure 3.2: Framework for developing an improved ShuffleNet to detect plant leaf diseases (complex background)

Table 3.1: Specifications of the processing device (Acer Aspire E 14)

Specification	**Details**
Processor	Intel Core i7-8550U
Base Clock Speed	1.8 GHz
Max Turbo Boost	Up to 4.0 GHz
Core/ Threads	4 Cores/ 8 Threads
Cache	8 MB SmartCache
RAM	8GB DDR4 (upgradeable to 16GB)
Storage	256GB SSD (upgradeable)
Operating System	Windows 10 Home
Cooling System	Dual fan cooling
Expandable Memory	Easy access to upgrade RAM and storage
Bus Type	DDR4 Memory Bus
Memory Bandwidth	68 GB/s (for DDR4 RAM)

Images of capsicum leaf are captured from a home-farm using a Kinect camera, DSLR and mobile phone camera upon both complex and uniform backgrounds. The detailed intrinsic parameters and specifications of the standard RGB camera, Kinect camera and

mobile camera used in this study are provided in Appendix 2. An expert categorises plant leaf diseases, and the results of the proposed model are compared to the expert categorization (ground truth) to determine the accuracy of the proposed model of the improved ShuffleNet. Professor Dr. Amit Baran Sharangi, a Professor from Department of Plantation, Spices, Medicinal and Aromatic Crops at Bidhan Chandra Krishi Viswavidyalaya, Mohanpur, West Bengal, India was the expert used in this study to categorise the types of plant leaf diseases. He is an agricultural professor and an expert in plant management. He was also the individual with comprehensive knowledge of plant leaf diseases, whose classification is deemed appropriate for this study, and he served as an expert reference point for validating the research findings in this particular field.

This experimental setup employs a Kinect camera, strategically mounted on a tripod to ensure precise and stable positioning during the image acquisition process as shown in Figure 3.3. The tripod's role is critical in maintaining a fixed height, angle, and orientation of the camera, which eliminates variability and ensures uniform data collection across all samples. By achieving consistency in imaging, the setup minimizes external distortions caused by camera movements, such as shifts in focus or changes in perspective, which can otherwise affect the accuracy and reliability of the captured data. Furthermore, the use of natural lighting conditions simulates real-world agricultural environments, providing a more realistic and practical dataset. This approach ensures that the acquired images closely represent real field scenarios where disease detection systems are ultimately deployed. The resulting dataset, composed of RGB images, forms a solid foundation for developing and validating the proposed classification model where consistent and high-quality data enhances the performance of detection models.

Images are pre-processed with the image resize technique to ensure they are appropriately formatted for more effective analysis. While image augmentation is used to increase the total number of diseased leaf images. The network of an improved ShuffleNet is then trained on a split dataset. Then, prior to performing the training operation, a few layers of ShuffleNet are altered. When the improved ShuffleNet provided the best detection accuracy, the parameters in the fully connected layers were also randomly initialised to find the ideal values. The ShuffleNet models hyperparameters are then randomly fine-tuned to find the best values for achieving the best validation accuracy.

Several trials and errors of fine-tuning parameters and ShuffleNet layers are performed to find the optimal values that provide the best possible performance in order to obtain the best validated model. The validation model is used to test the proposed model's accuracy in classifying capsicum leaf diseases after determining the optimal fine-tuning parameters and ShuffleNet layers. The validated models are tested using an augmented test set to evaluate the performance of the final tuned model. The final percentage accuracy of the model after training is discovered. The proposed improved ShuffleNet model's classification accuracy is then validated using the confusion matrix to validate the classification performance based on more detailed metrics. The proposed improved ShuffleNet model's classification accuracy is compared to other existing models such as Densely Connected Convolutional Networks 201 (DenseNet201), Residual Neural Network 50 (ResNet50), BoF, ShuffleNet and ShuffleNet V2 to demonstrate that the proposed model is superior to the existing models.

Figure 3.3: Experimentation set-up

3.3 Dataset

The dataset used in this study consists of five different plant types that were capsicum, rice, corn, tomato, and citrus. These plant types were chosen due to their agricultural importance and their susceptibility to various diseases that significantly impact crop yield and quality. The dataset encompasses a diverse collection of images to ensure robust model training and evaluation. The capsicum dataset is uniquely created for this research, while the datasets for rice, corn, tomato, and citrus are sourced from publicly available online repositories. This combination of a custom-created dataset and pre-existing datasets ensures diversity and robustness in the model's training and evaluation.

3.3.1 Created Dataset

This research has specific focus on capsicum due to its agricultural significance in the region and Department of Agriculture Sarawak has long reported capsicum leaf disease at their Agriculture Research Centre, Semenggok, Kuching (Lim, 2008) and stressed the importance of studying crop management. While including downloaded datasets for other

crops helps demonstrate the scalability and generalization capability of the proposed model across different plant types and conditions.

For capsicum, the dataset was acquired from a home-farm, which means that the observations and data are gathered from a small-scale farming operation conducted at a personal residence. The idea to create a dataset particularly for capsicum rather than other crops such as rice, maize, tomato, and citrus stems from the fact that these plants have significant taxonomic similarities. They have specifically similar taxonomic characteristics which is greenish yellow (5GY 6/6) in colour (Munsell, 2019) and have leaf types which are common leaf from elliptical shape (Core, 2013) as shown in Figure 3.4 and Figure 3.5.

By self-collecting images of diseased capsicum leaf using Kinect camera, a stable and reliable dataset is ensured, free from the risks of discontinuity and legal restrictions associated with third-party sources. The Kinect camera offers high-resolution and consistent quality images, which are crucial for accurate disease detection. This approach allows to maintain high-quality, relevant, and consistently labelled data, tailored to this study specific research needs.

Additionally, detailed metadata can be captured and ensure proper validation by experts, enhancing the accuracy and robustness of the analysis. This autonomy from external data sources ensures the continuity and integrity of our research, facilitating the development of precise and effective proposed disease detection model.

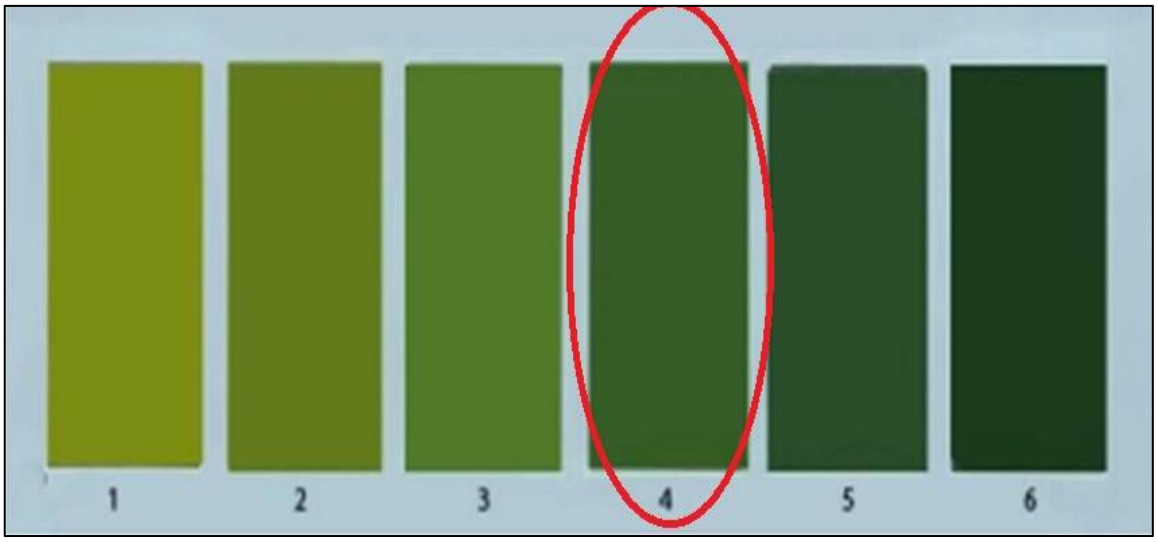


Figure 3.4: The specific taxonomic leaves colour of capsicum, rice, corn, tomato, and citrus according to Munsell (2009) which is greenish yellow (5GY 6/6) (red circle)

Figure 3.5: The specific taxonomic leaf types of capsicum, rice, corn, tomato, and citrus which is elliptic according to Core (1955) (in red circle)

Collecting data from large-scale farms provides valuable insights into the unique challenges encountered by professional or industrial growers cultivating capsicum plants on expansive agricultural estates or commercial fields (Nogeire-McRae et al., 2018). Capsicum dataset collection entails systematically monitoring capsicum plants grown on the home-farm for any signs of disease, such as unusual spots on leaves, discoloration, wilting, deformities, or any other symptoms that indicate the presence of a disease. These observations were then captured in images.

The device utilized for capturing the capsicum leaf was a Kinect sensor camera, alongside a DSLR and a mobile camera for comparative analysis. The advantages and disadvantages of each device are elucidated in Chapter 2 (Section 2.4.2), the literature review section. To capture the desired data, the Kinect sensor is set to a fixed specification. To avoid bias in the investigation's later results, this will be standardised for all types of capsicum leaf diseases. Table 3.2 shows the fix specification used on the Kinect, DSLR 550d and Vivo Y12s.

Table 3.2: Kinect, DSLR 550d and Vivo Y12s camera specification

Specification	Kinect	DSLR 550d	Vivi Y12s
Model	Kinect for Xbox One Sensor	Canon EOS 550D	Vivo Y12s
Sensor Type	RGB sensing camera	CMOS	CMOS
Megapixels (MP)	N/A	18.0 MP	13 MP + 2 MP
Lens Type	N/A	Basic lens (Canon EF-S 18-55mm)	Fixed

Table 3.2: continued

ISO Range	N/A	100 - 6400 (expandable to 12800)	N/A
Autofocus Points	N/A	9	Phase detection autofocuses
Viewfinder Coverage	N/A	95%	N/A
Battery Life	N/A	Approx. 440 shots camera and imaging products association (CIPA)	5000 milliampere-hour (mAh)
Dimensions (WxHxD)	N/A	128.8 x 97.3 x 62 mm	164.41 x 76.32 x 8.41 mm
Weight	N/A	Approx. 530 g (body only)	191 g
Horizontal Resolution	1920 pixels	5184 pixels	1080 pixels
Vertical Resolution	424 pixels	3456 pixels	720 pixels
Vertical Field of View	60 degrees	N/A	N/A
Horizontal Field of View	70 degrees	N/A	N/A
Depth Resolution	512 × 424 pixels	N/A	N/A

To create a comprehensive dataset, images are continuously captured, and new images are added to the dataset to improve its diversity and robustness under the same lighting conditions and stages of disease progression. Grey spot, leaf discoloration, and leaf curling are the three types of leaf diseases investigated in this study, as shown in Figures 3.6, 3.7 and 3.8. The datasets were captured under controlled conditions, with one set of leaves placed in a uniform background and the other set of leaves captured alongside a cluster of other leaves in a complex background as presented in Figure 3.9 and Figure 3.10.

Addressing complex backgrounds is critical for effective data collection, with lighting playing a secondary but supportive role. Images captured against a uniform backdrop eliminate distractions and simplify segmentation, ensuring consistent subject representation under controlled lighting. In contrast, intricate backgrounds introduce variability through overlapping objects, textures and colours, which can be exacerbated by uneven lighting conditions. Standardizing camera settings, such as aperture and focus, helps maintain clarity, while precise positioning and framing ensure the subject remains the focal point. This meticulous approach enhances the algorithm's ability to handle diverse and complex environments reliably, even with varying lighting.

Figure 3.6: Grey spot

Figure 3.7: Discolour leaf

Figure 3.8: Leaf curling

Figure 3.9: Uniform background

Figure 3.10: Complex background

3.3.2 Downloaded Dataset

It is critical to collect data from multiple sources when studying plant diseases in order to gain a comprehensive understanding of the problem (Buja et al., 2021). The dataset captured using various cameras in this study is presented in Table 3.3. The dataset utilized in this study was sourced not only from self-collected data using Kinect, DSLR, and mobile camera but also from a publicly accessible dataset. Specifically, rice diseased leaf images were obtained from Kaggle (Vbookshelf, 2021), corn diseased leaf images from Kaggle (Smaranjit, 2020), tomato diseased leaf images from Kaggle (Kaustubh, 2020), and citrus diseased leaf images from Mendeley Data (Rahim et al., 2021).

Table 3.3: Tabulated subset of diseased plant leaf images from the dataset

Camera	**Plant Type**			
	Capsicum			
		Grey spot	**Discolour leaf**	**Leaf curling**
Kinect	Uniform background			
	Complex background			
DSLR	Uniform background			
	Complex background			

Table 3.3 continued

Mobile phone	Uniform background			
	Complex background			
		Rice		
		Grey spot	**Discolour Leaf**	
Internet source	https://www.kaggle.com/datasets/vbookshelf/rice-leaf-diseases			

Table 3.3 continued

		Corn		
		Grey spot		
	https://www.kaggle.com/datasets/smaranjitghose/corn-or-maize-leaf-disease-dataset			
		Tomato		
		Grey spot	**Discolour leaf**	**Leaf curling**
	https://www.kaggle.com/datasets/kaustubhb999/tomatoleaf			
		Citrus		
		Grey spot	**Discolour Leaf**	
	https://data.mendeley.com/datasets/3f83gxmv57/2			

Many existing plant disease datasets are captured under controlled laboratory conditions with uniform backgrounds, which may not accurately reflect real-world agricultural settings. The dataset includes five plant types: capsicum, rice, corn, tomato, and citrus. The capsicum data, collected specifically from a home farm, addresses this gap by capturing images under diverse real-world conditions. Moreover, establishing a dedicated capsicum dataset is valuable, as capsicum shares notable taxonomic similarities with other plants such as the common greenish-yellow coloration and elliptical leaf shape which enhances the proposed model's ability to generalize across different plant species.

In contrast, the dataset developed as part of this study was intentionally designed to bridge this gap by encompassing a broader spectrum of environmental conditions. This dataset comprises not only images captured in uniform backgrounds but also those taken in more complex and naturalistic settings. These latter scenarios include backgrounds containing clusters of plant leaves, simulating the intricate and cluttered conditions often encountered in agricultural fields. By intentionally incorporating such real-world scenarios into this dataset, this study aims to enhance its reliability and applicability for practical use in agricultural contexts. This diverse representation ensures that this dataset reflects the complexities encountered in real-world settings, enhancing its utility for research, development, and application in the field of plant disease detection.

3.4 Image Processing

In this study, two image processing techniques are used to analyse images of diseased plant leaves. Image resizing and image augmentation in Matlab are employed before applying the ShuffleNet algorithm for further analysis. Image resizing, as the name implies, is the process of amending the dimensions (size) of an image. It is a customary technique used in image processing to change the resolution or aspect ratio of an image (Begmatov et

al., 2023). The images in this study are resized to 224 × 224 pixels. The reason for resizing the images to this size is due to the ShuffleNet algorithm's requirements. Aside from that, resizing images to a fixed scalling of 224 × 224 is common in many computer vision tasks for reasons such as controlling memory prerequisites during both the training and inference stages and ensuring the model can focus on learning disease characteristics rather than being predisposed by variations in image sizes (Talebi & Milanfar, 2021).

The scaling factors are denoted by sx and sy, respectively. After resizing the image, the scaling factors are calculated as follows:

$$sx = 224 / N \quad \text{Equation 3.1}$$

$$sy = 224 / M \quad \text{Equation 3.2}$$

where M represents the height and N represents the width of the image. The new pixel position in the resized image, (u, v), is then computed using the scaling factors:

$$u = x * sx \quad \text{Equation 3.3}$$

$$v = y * sy \quad \text{Equation 3.4}$$

based on the scaling factors and the original position, these equations determine the new coordinates of the pixel in the resized image.

3.4.1 Uniform Background

With the reference to diseased leaf image with uniform background as shown in Figure 3.11, it is vital to contemplate aliasing during the resizing process, as it can cause distortion or artefacts in the resized image (Bali et al., 2023). As a result, anti-aliasing techniques are used in this study to reduce aliasing. A low-pass filter is applied before down sampling the image to eliminate high-frequency information that can cause aliasing. This filtering aids in the preservation of image details and the reduction of aliasing effects. The image is then up sampled using band-limited interpolation techniques to avert instituting

high-frequency artefacts. These techniques warrant that the interpolated pixel values do not contain frequencies that surpass the Nyquist limit of the original image (640×480).

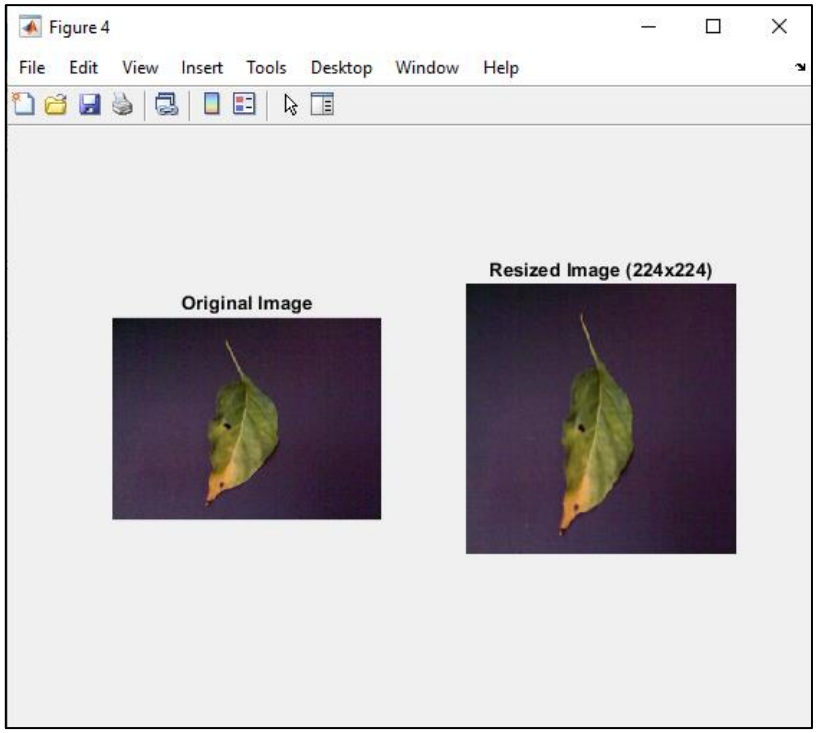


Figure 3.11: Image resizes

After pre-processing, images captured by various cameras, each with distinct resolution ranges, are standardized to a uniform size of 224×224 pixels. This study employs the data augmentation technique known as random reflection in the image processing stage. Data augmentation is a common approach utilised in machine learning to artificially raise the size of the training dataset by operating several transformations to the prevailing data (Shani et al., 2023). It entails randomly flipping the images horizontally. The model can learn to recognise diseases from different perspectives and orientations by flipping the images and this technique effectively increases the diversity of the training dataset, giving the model more diverse examples to learn from (Dhaka et al., 2021). Random reflection augmenting the original dataset by randomly flipping the images horizontally. This

procedure generates additional training samples with the same label but different orientations.

In Figure 3.12, the primary procedure used in this study for executing random reflection for image augmentation is depicted. Initially, the procedure assumes that the original images are stored in a specified directory. The code then proceeds to load these original images, organized by disease class, and displays them sequentially. Following this step, random reflection augmentation is applied to these images using the flip function, which horizontally flips each image. The resulting augmented images are showcased alongside their respective originals.

A key aspect of this approach is the flexibility to specify the number of original images per disease class used for augmentation. This allows for controlled augmentation while ensuring dataset diversity. By varying the number of images per class, researchers can effectively manage the balance between increasing dataset size through augmentation and preserving the integrity and diversity of the original dataset. This methodological detail ensures that the augmented dataset reflects a realistic variation of the original images, enhancing the robustness and generalization capability of the models trained on this augmented data.

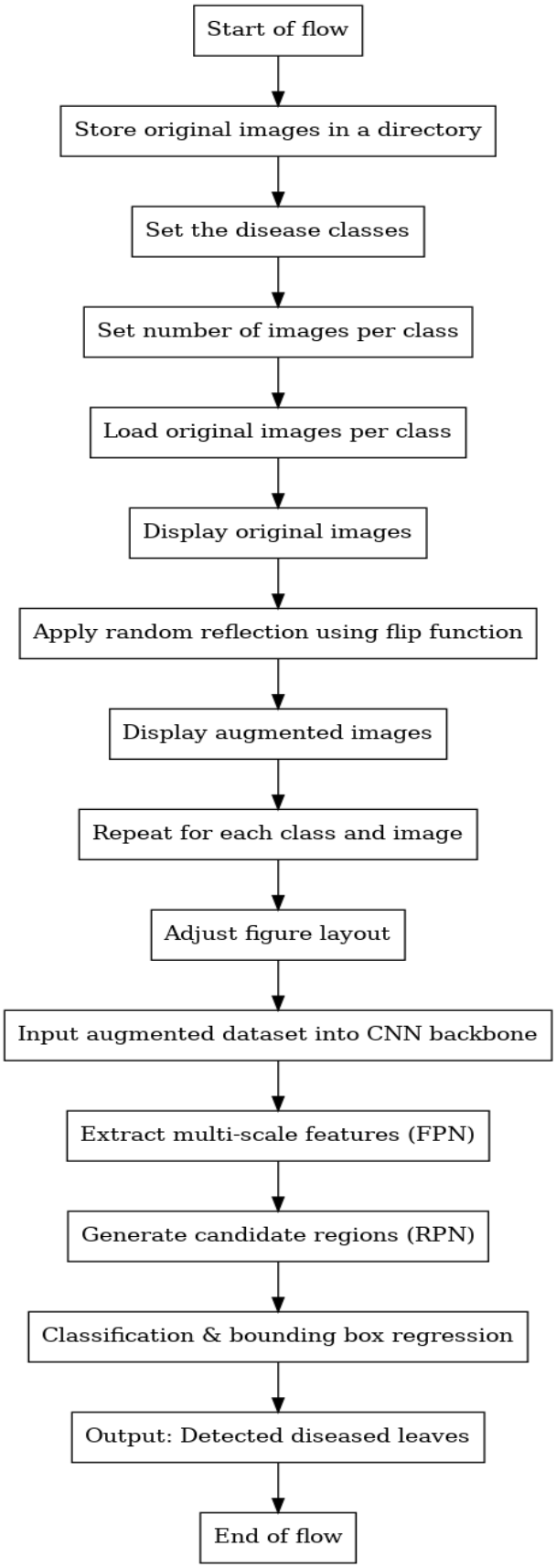


Figure 3.12: Step to augment diseased plant leaf images using random reflection technique

Horizontal flipping enhances data diversity by adding orientation variability that reflects natural conditions such as wind, leaf growth, and camera angles (Shorten & Khoshgoftaar, 2019; Taylor & Nitschke, 2018). This goes beyond simple duplication, as random reflection helps the model learn orientation-invariant features that strengthen its ability to generalize across diverse conditions (Perez & Wang, 2017). In practice, this small operation has a large impact on improving robustness when combined with other augmentation techniques. This randomness in the angle of reflection helps in synthesizing diverse training samples, effectively expanding the dataset without introducing redundant information. By exposing the model to this augmented dataset during training, it learns to recognize features invariant to horizontal flips and generalizes better to unseen data, as it becomes more adept at recognizing objects regardless of their orientation. Moreover, by encouraging the model to learn from various perspectives, random reflection aids in making the model more robust and adaptable to real-world scenarios where objects may appear in different orientations.

Before performing the random reflection operation, ensure that the image quality is adequate for augmentation. By employing a Laplacian variance threshold, Vidhyapathi et al., (2020) demonstrated the effectiveness of using a threshold value of 150.0 and skipping 18% of low-quality images to enhance augmentation reliability. Then, the sharpness or focus measure is calculated in this study to assess image quality. The Laplacian variance is employed, which is calculated using the following equation:

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \qquad \text{Equation 3.5}$$

here $\frac{\partial^2}{\partial x^2}$ is the x kernel and $\frac{\partial^2}{\partial y^2}$ is the y kernel. The sharpness measure can be used to determine whether an image is of good quality and suitable for augmentation. If the sharpness measure falls below a certain threshold, the image is skipped and another from the

dataset is selected (Low Laplacian variance values, generally below 100 indicate poor sharpness or blur).

To determine the threshold for the sharpness measure, the distribution of sharpness values across a representative sample of images from the dataset was analysed, and the threshold was set at a value where the majority of images deemed suitable for augmentation fell above it. This ensured that images with adequate detail and edges were selected, while those below the threshold were excluded. Experimentation and validation against the desired quality standards were conducted to fine-tune this threshold for optimal performance. Higher values of gradient variance generally correspond to sharper images, as they indicate more detail and edges (Kaur et al., 2021). In this case, the Kinect images were the sharpest (per pixel), with 800 for capsicum, 600 for rice, 800 for corn, 500 for tomato, and 400 for citrus, making them easier to process. Table 3.4 and Table 3.5 show the sharpness variance for capsicum images captured with Kinect, DSLR, and mobile cameras, and for rice, corn, tomato, and citrus images obtained from online sources. The differences in sharpness variance across these devices and crops reflect variations in image resolution, sensor quality, and lighting conditions during acquisition.

The number of augmented images was determined by first filtering out low-quality samples using a sharpness threshold, ensuring that only images with adequate detail and edges were retained. Augmentation was then applied to balance class distributions and increase the dataset size to a level sufficient for reliable CNN training. Random reflection was used until underrepresented classes reached comparable sample sizes. This ensured both data quality and statistical robustness in the final dataset. Random reflection was used to generate additional images until underrepresented classes were comparable to others, thereby reducing bias and enhancing statistical power (Dalmaijer et al., 2022). This approach

ensured that the final dataset combined both quality and quantity, enabling more reliable feature extraction and classification accuracy. Table 3.6 depicts the detection class distribution from total original to final total after augmented with random reflection technique. In general, a vaster dataset grants more statistical power. Statistical power is the adeptness to accurately detect true effects or relationships between variables (Dorey, 2010). A vaster dataset rises the likelihood of detecting subtle or minor effects that are statistically significant but would go unnoticed in a smaller sample.

This raised statistical power enables more robust and dependable conclusions of the algorithm. However, maintaining data quality is still critical, that is why the sharpness of the images is identified in this study. Image sharpness critically affects the performance of Convolutional Neural Networks (CNNs), as sharp images retain high-frequency details essential for feature extraction, while blurred images can reduce classification accuracy by up to 20%. Research by Karahan et al. (2016) demonstrated that severe blurring reduces spatial gradient strength, leading to a drop in recognition performance from 95% to as low as 75% on test datasets, emphasizing the need for sharp, high-quality images.

Table 3.4: Capsicum sharpness variance

Device	Sharpness
Kinect	798.88
DSLR	655.43
Mobile Phone	482.988

Table 3.5: Rice, corn, tomato, and citrus sharpness variance

Plant	Sharpness
Rice	594.676
Corn	778.4
Tomato	519.8
Citrus	416.67

Table 3.6: The study's dataset from the total original to the final total after using the random reflection

Plant type		Total original dataset	Total augmented dataset
Capsicum	Grey spot	96	1334
	Discolour leaf	80	1240
	Leaf curling	75	936
Rice	Grey spot	72	1035
	Discolour leaf	54	712
Corn	Grey spot	87	1297
Tomato	Grey spot	63	852
	Discolour leaf	77	818
	Leaf curling	56	746
Citrus	Grey spot	45	497
	Discolour leaf	34	359

3.4.2 Complex Background: Feature Refinement and Detection Integration

For diseased leaf image with complex background, the proposed approach begins as shown in Figure 3.12 with an input image processed through multiple ShuffleNet stages, featuring shuffle units that employ group convolutions, channel shuffling, and depthwise convolutions to efficiently extract features. These features undergo further refinement through Feature Pyramid Network (FPN) layers, creating a multi-scale representation crucial for detecting objects across various sizes. Meanwhile, the Region Proposal Network (RPN) utilizes these multi-scale features to generate precise bounding box proposals, particularly for identifying diseased areas on plant leaves. This integrated architecture harnesses the lightweight efficiency of ShuffleNet alongside the robust object detection capabilities of FPN and RPN, ensuring accurate localization and detection of potential diseased regions within plant images. Figure 3.13 visualize the flow of the algorithm from input to output.

To enhance the accuracy of cropping diseased leaves, conventional image processing techniques are integrated into the algorithm. It incorporates edge detection, morphological operations, and connected component analysis, similar to how an RPN operates in object detection, to pinpoint potential diseased areas. The script adapts edge detection and morphological operations for small object detection, thereby generating precise bounding boxes around suspected diseased regions.

Small object detection focuses on identifying and localizing minute diseased spots that are often characterized by limited pixel regions and subtle features. This approach is further bolstered by FPN's ability to manage varying leaf scales through multi-level feature maps, optimizing the detection and cropping process across different resolutions. By leveraging a series of processing stages and parameters akin to an enhanced ShuffleNet architecture, the script ensures precise detection and isolation of diseased regions, mirroring

the effectiveness of improved ShuffleNet in identifying and classifying diseased plant leaves with heightened accuracy. The illustration of the process is displayed in Figure 3.14.

The implementation of FPN and RPN within the improved ShuffleNet framework for plant leaf disease detection is that the FPN is positioned after the feature extraction layers, leveraging their multi-scale outputs. The FPN refines features by combining high-level, semantically rich layers with lower-level, spatially detailed layers, enabling robust representation of disease patterns at multiple scales. These refined multi-scale features are then input to the RPN, which uses bounding-based mechanisms with abjectness score to generate precise bounding box proposals. The RPN ensures accurate localization of diseased regions, even in complex backgrounds or varying scales of the diseased spots. Together, the FPN and RPN synergize to enhance both the spatial accuracy and detection performance of the network.

The abjectness score in a Region Proposal Network (RPN) is automatically processed to evaluate the likelihood that a bounding box contains an object rather than background. This score is derived from the Intersection-over-Union (IoU) metric, which measures the overlap between the bounding box and ground-truth boxes. During training, each bounding box is assigned a positive or negative abjectness score based on its IoU with ground-truth boxes. If the bounding box has an IoU of 0.9 with a ground-truth box, it receives a positive score, while bounding boxes with IoUs below 0.3 are assigned negative scores. The RPN utilizes these scores to filter bounding boxes, retaining those with high abjectness scores for further processing in the detection pipeline as shown in Figure 3.13.

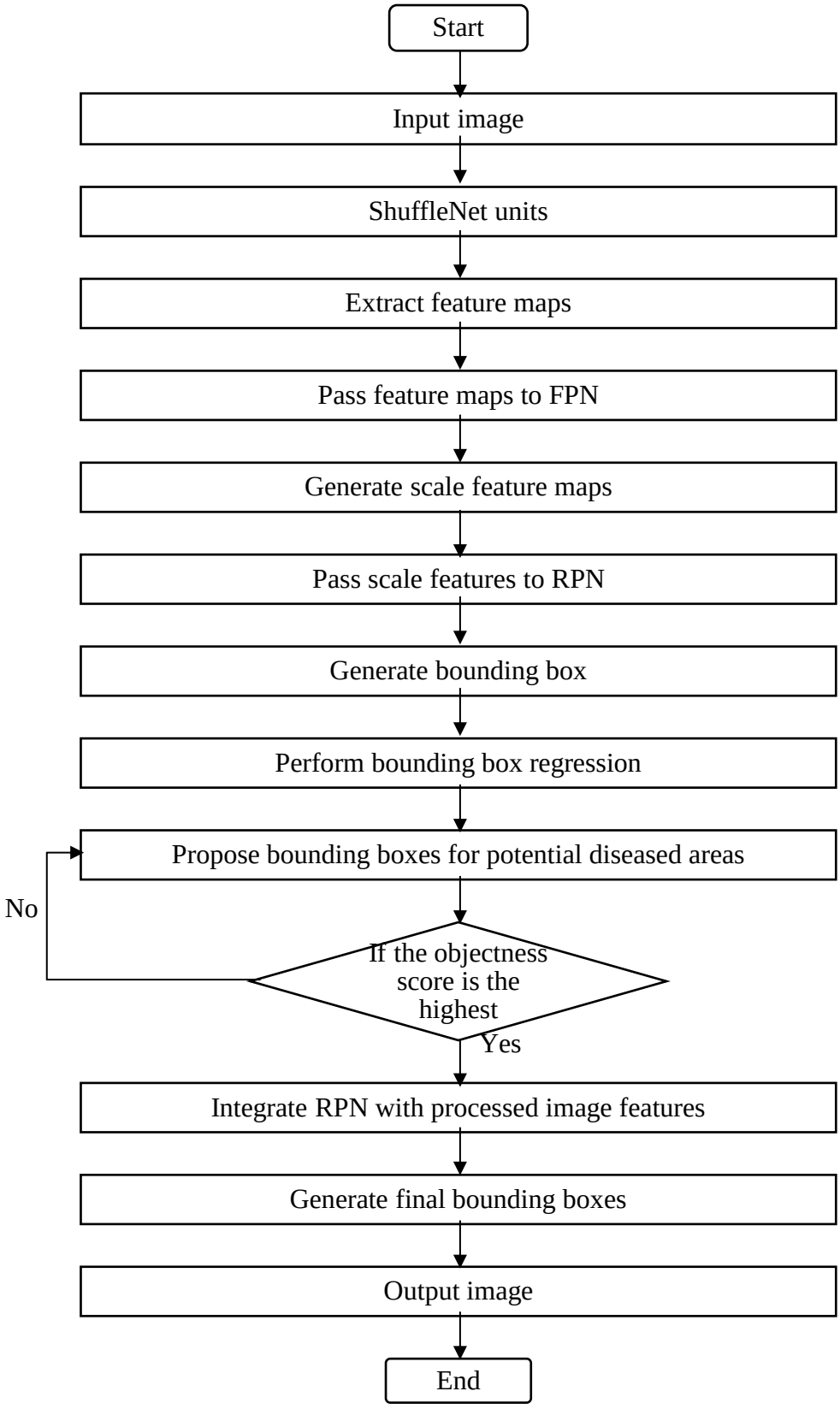


Figure 3.13: ShuffleNet model with FPN and RPN

(a)

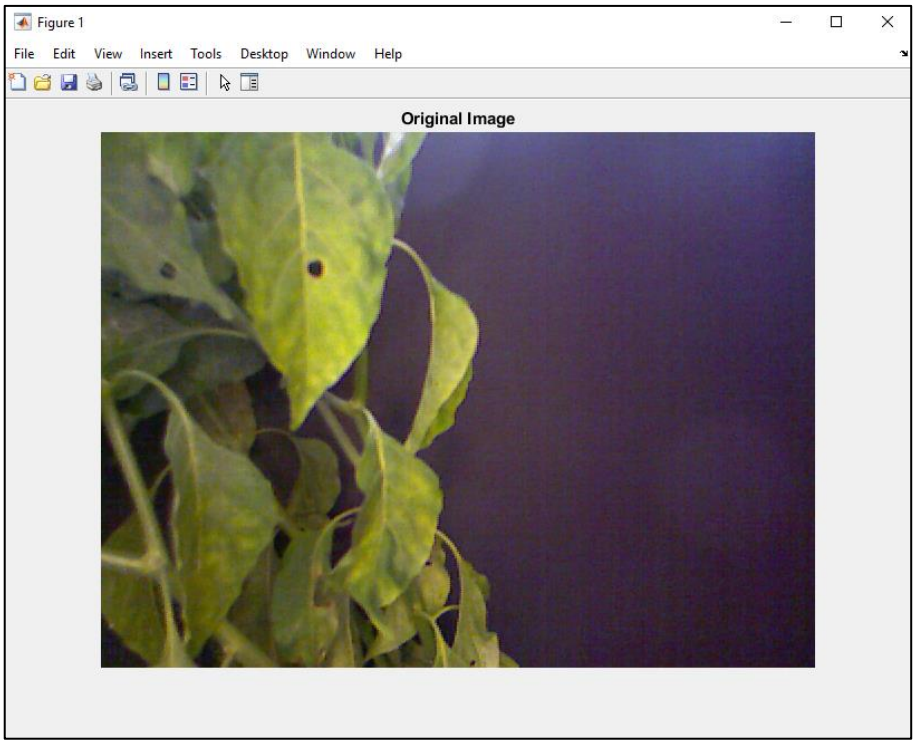


Figure 1
File Edit View Insert Tools Desktop Window Help
Original Image

(b)

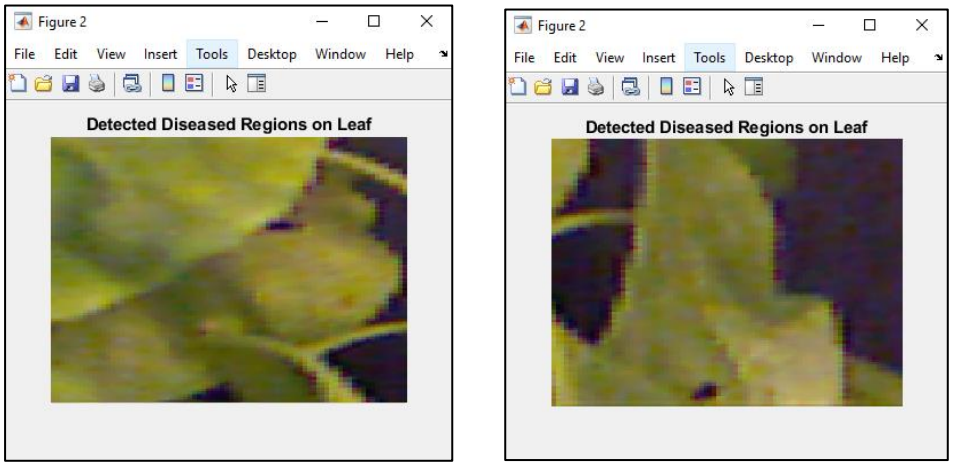


Figure 2
File Edit View Insert Tools Desktop Window Help
Detected Diseased Regions on Leaf
Figure 2
File Edit View Insert Tools Desktop Window Help
Detected Diseased Regions on Leaf

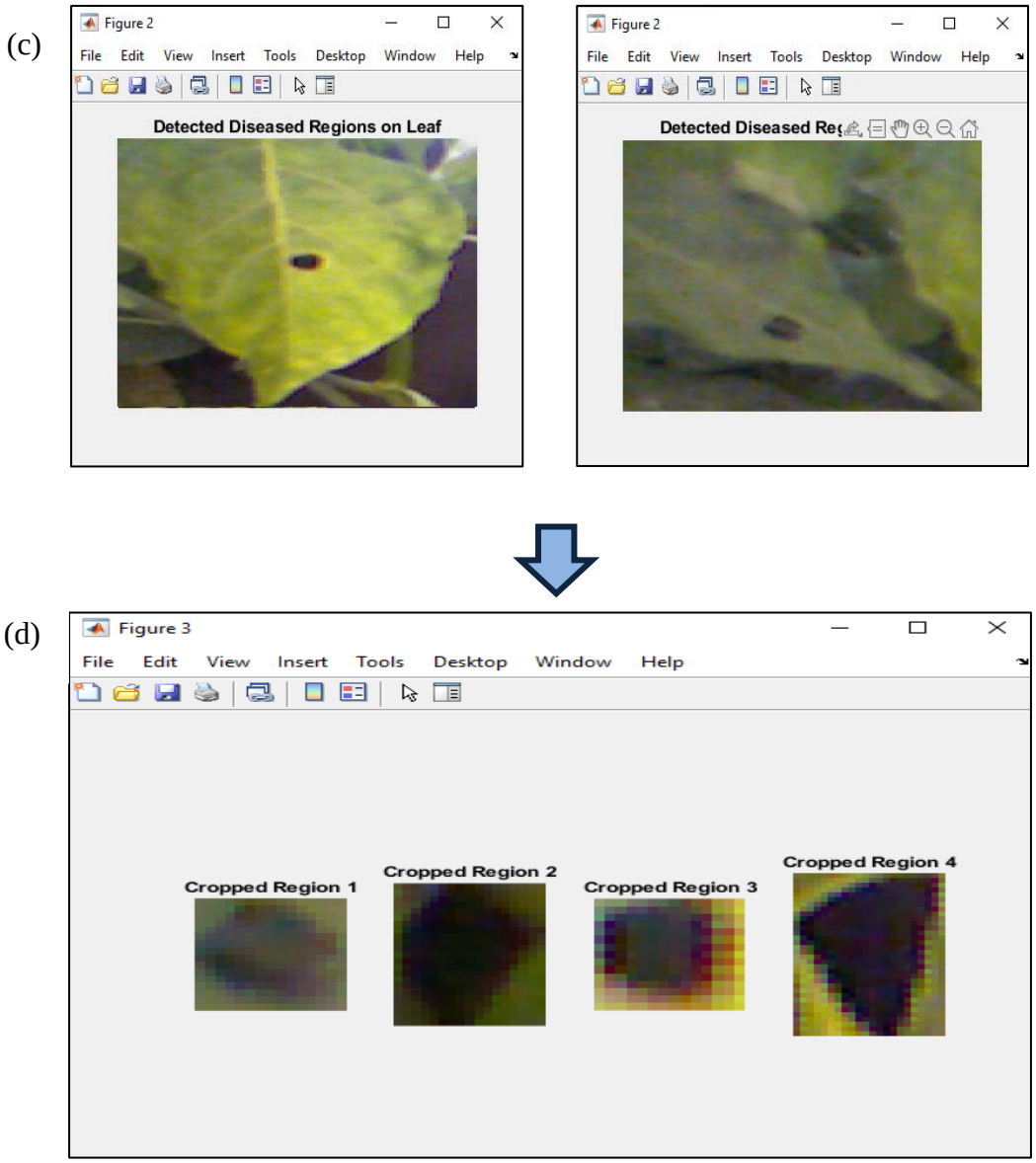


Figure 3.14: Steps in the automated detection and precision cropping of diseased plant leaves. (a) Original leaf image; (b) preprocessing and segmentation to isolate the leaf region; (c) detection of diseased area using the trained model; (d) precision cropping of the diseased portion for further analysis.

Unlike controlled datasets with uniform backgrounds, real agricultural images often include clutter such as soil, stems, or overlapping vegetation. To demonstrate the

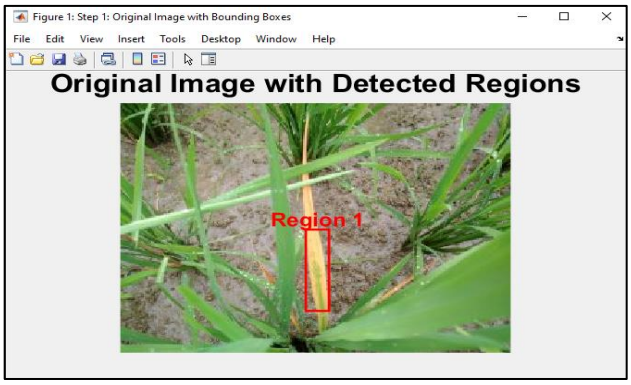
methodology for handling such environments, the paddy plant was selected as a representative case. Paddy leaves frequently exhibit yellowing symptoms when affected by diseases, which makes them an ideal example for testing segmentation and detection under complex field conditions. The workflow described in this section outlines the preprocessing, segmentation, and classification pipeline for isolating diseased regions from complex backgrounds, as illustrated in Figure 3.15.

The process begins with image acquisition and preprocessing, where raw images are captured under natural lighting conditions. The images are resized and normalized to reduce illumination variation and device-specific effects. Next, the images are converted from RGB to LAB colour space. This transformation separates luminance (‘L’) from chromaticity (‘a’ for red–green, ‘b’ for blue–yellow). Diseased areas on paddy and other leaves often manifest as yellowing, which is clearly represented in the ‘b’ channel. By focusing on the ‘a’ and ‘b’ channels, discoloured regions are highlighted while suppressing background interference.

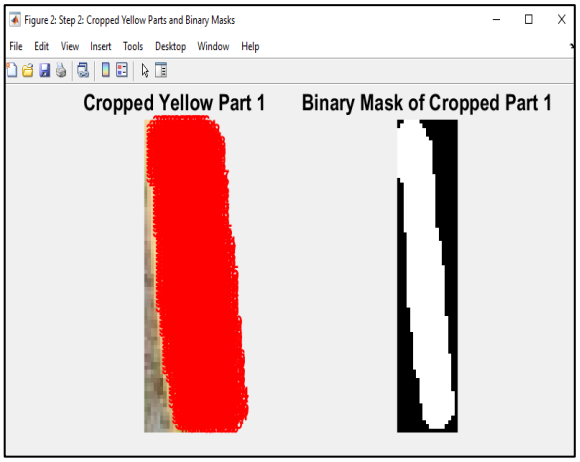
A thresholding operation is then applied, using empirically determined ranges ($-20 \leq a \leq 20$, $30 \leq b \leq 100$) to generate a binary mask. This mask isolates potential diseased regions, highlighting yellow or necrotic spots while discarding non-relevant background areas. Since thresholding alone may introduce noise, morphological operations are performed. Erosion, dilation, and closing (structuring element radius = 5 pixels) refine the binary mask by removing small artefacts and filling gaps, producing a cleaner segmentation of the leaf area.

Following refinement, region extraction and cropping are carried out. Bounding boxes are generated around connected components in the binary mask, and corresponding segments are cropped from the original image. These cropped areas represent the diseased portions of the leaf. Each cropped region is subjected to feature refinement, where LAB colour values are re-analysed. The ‘b’ channel intensity is quantified to measure the severity of yellowing, and these values are annotated for each region. This step links the visual discolouration with quantifiable features that can be learned by the CNN.

Finally, the refined regions are fed into the improved ShuffleNet with FPN and RPN. Here, hierarchical feature extraction and bounding box proposals further improve the accuracy of detection and classification. By first isolating diseased regions through preprocessing and then applying ShuffleNet for classification, the pipeline ensures robustness even under complex background conditions.

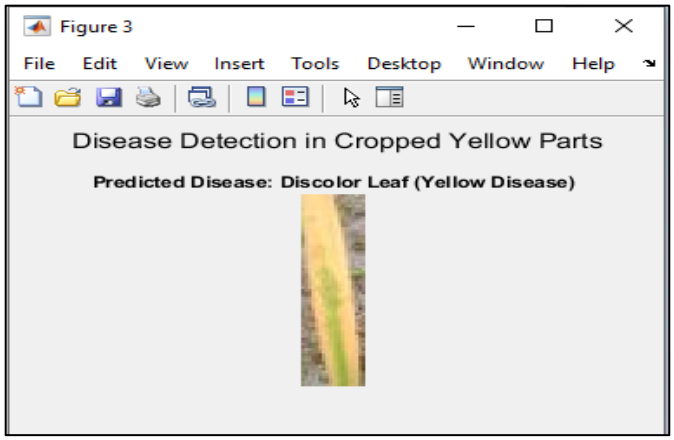


Displays the original image overlaid with bounding boxes, around detected yellow regions.

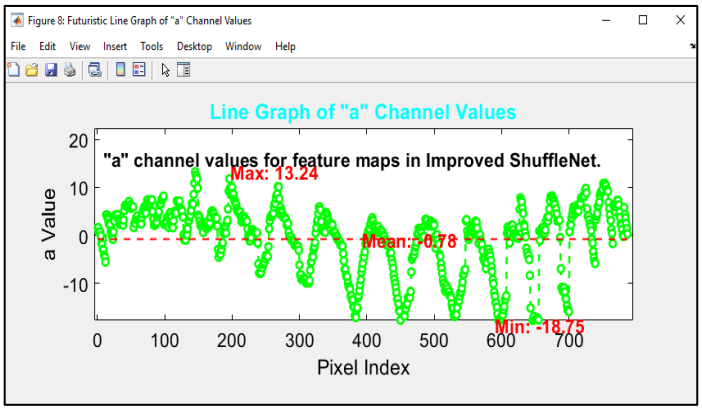


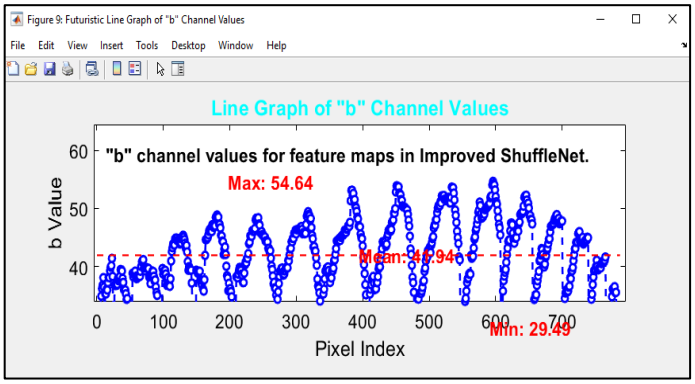
The script iterates through each detected yellow region using bounding box information (stats.BoundingBox), extracting and cropping corresponding parts of the original image with imcrop. Each cropped region is converted to lightness: Red/Green: Blue/Yellow (LAB) colour space to analyze 'a' and 'b' channels locally. It applies a yellow detection threshold to create yellowMaskCropped, followed by morphological

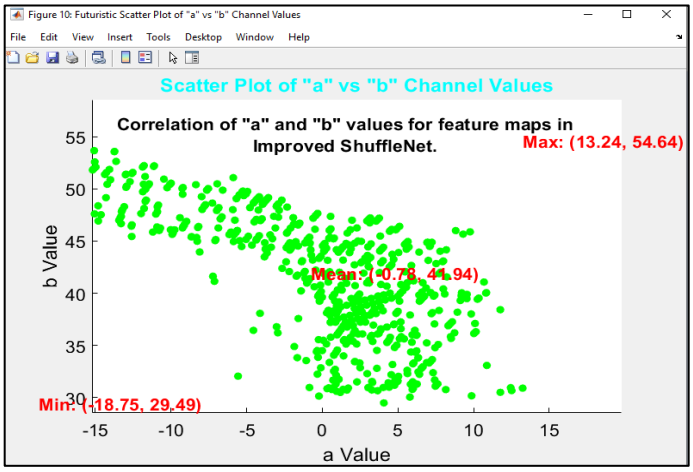
operations for binary mask refinement. The script collects 'a' and 'b' channel values from these cropped regions then visualizations include displaying each cropped yellow region alongside its cleaned binary mask, with annotations showing 'b' channel values to quantify yellow colour intensity in different leaf areas.



This step is designed to display predicted diseases linked to the detected yellow areas. Currently, it shows placeholder text ("Predicted Disease: Discolour Leaf (Yellow Leaf)").







Colour space conversion, Thresholding, and Morphological Operations:

The image is converted to LAB color space, isolating the 'a' and 'b' channels. Thresholding values (a: -20 to 20, b: 30 to 100) are applied to detect yellow regions, resulting in a binary mask. Morphological operations are performed with a structuring element of radius 5 to remove noise and close gaps.

Region Extraction, Cropping, and Disease Prediction:

Bounding boxes are generated around detected yellow regions using regionprops. Each region is cropped and cleaned with the same threshold and morphological

steps. Finally, bounding boxes and ‘b’ channel values (e.g., b: 45.7) are annotated on the cropped regions to show discolouration levels, simulating the feature extraction and refinement later carried out by the improved ShuffleNet.

Figure 3.15: Detection of yellow regions on paddy leaves in complex backgrounds using the improved ShuffleNet model

Data augmentation is a technique used to increase the size and diversity of the training dataset by applying various transformations to the input images. This can be mathematically expressed by letting x and $R(x)$ be a random reflection function that flips the image x horizontally with a probability of 0.5. The reflected image can then be defined as:

$$x' = R(x) * x \qquad \text{Equation 3.6}$$

where x' is the reflected image, and $*$ is the element-wise multiplication operator. In practice, the random reflection function $R(x)$ can be implemented using a binary variable r, which is sampled from a Bernoulli distribution with a probability of 0.5. The detailed procedure is outlined below.

Algorithm 3.1: Random Reflection Function

Input: Image x

Output: Reflected image x'

1. Sample r ~ Bernoulli (0.5).
2. If r = 1, then:

$$x' = flip\_horizontally(x)$$

3. Else:

$$x' = x$$

4. Return x'

Algorithm 3.1 introduces stochastic variability by randomly flipping images with equal probability. This ensures that the dataset contains both original and reflected samples, thereby enhancing diversity and supporting better generalization in model training.

3.5 Hyperparameters Selection

Following the pre-processing steps, prepare the dataset. Then, the first step in fine-tuning hyperparameters was to create the network model. This is accomplished by using MATLAB's Deep Learning Toolbox to define the neural network architecture (ShuffleNet) and MATLAB's Neural Network Toolbox to implement custom models (improved ShuffleNet). Then, set the hyperparameters. This is achieved by initialising the hyperparameters that are to be optimised, namely the learning rate, number of epochs, minibatch size, and optimizer. Start with a default value for learning rate and adjust it as needed.

In this study, the parameters that enabled the proposed improved ShuffleNet model to achieve the highest accuracy were chosen for implementation. Further details are provided

in the Results section (Section 4.4) of the thesis. In summary, the learning rate is 0.010, the number of epochs is 50, the minibatch size is 64, the optimizer is SGDM and the early stopping is set to 'Auto'.

Frameworks and libraries that provide default values for the learning rate that work well for a wide range of tasks performed by Jiang (2022) were used. The learning rates proposed by these researchers are 0.001, 0.005, 0.01, 0.05, 0.1, and 0.5. For the number of epochs, reasonable values ranging from 10, 20, 30, 40, 50, 60, 70, 80, 90, and finally 100 were used. To get a sense of the model's performance, using a relatively small number of epochs, such as 10 or 20, is a good starting point (Roy et al., 2019). In terms of minibatch size, a small value was used to begin. A small minibatch size, such as 1, 8. 16, 32 and 64, is used as a starting point as small minibatch sizes provides a good balance of stochasticity and computational efficiency, as stated by Kandel and Castelli (2020). However, a minibatch size of 1 is stated leading to frequent and potentially noisy parameter updates. (Alzubaidi et al. 2021).

Then, because the dataset was already larger due to augmentation, this study experimented with large minibatch sizes, such as 128, 256, 512, and 1024, to evaluate their impact on performance. Larger minibatch sizes process more data simultaneously, accelerating training and facilitating faster convergence to a stable solution. They also act as a regularization technique, particularly beneficial for preventing overfitting in smaller datasets (Shen et al., 2023). Although they require substantial computational resources, the advantages in speed, stability, and regularization often justify their use, especially with modern hardware and optimized deep learning frameworks. The performance of the improved ShuffleNet across different minibatch sizes is discussed in the next chapter. According to Gupta et al., (2020), increasing the minibatch size for large datasets enhances

computational efficiency. Despite the higher resource demand, the dataset's substantial size ensures efficient use, maintaining the study's validity and enabling accurate assessment of varying minibatch sizes on performance metrics without compromising statistical significance or generalizability.

For optimizers, popular optimizers are used to get started. Optimizers such as Adam, Stochastic Gradient Descent (SGD), Root Mean Squared Propagation (RMSprop), Adagrad and SGDM have been shown to work well in a diversity of scenarios (Liu et al., 2023). So, in this study, these optimizers were used as a baseline for experimentation. Lastly, to find the R-squared (R^2) of the proposed improved ShuffleNet model. Firstly, to determine the regression model in the improved ShuffleNet, fit the model using the dataset's input features as predictors and performance metrics (like accuracy or loss) as the dependent variable. Once the model is trained, calculate the R^2 value by comparing the predicted and actual values to measure how well the model explains the variance in the data. A higher R^2 value indicates the improved ShuffleNet's ability to accurately predict outcomes in the given task.

3.6 Layer Modification

Figure 3.16 and Figure 3.17 show the original ShuffleNet architecture and ShuffleNet V2 that have demonstrated impressive performance in mobile and embedded vision applications. However, there is still room for improvement in terms of feature extraction capabilities and overall efficiency. An improved ShuffleNet architecture is proposed in this work to address these limitations while maintaining computational efficiency.

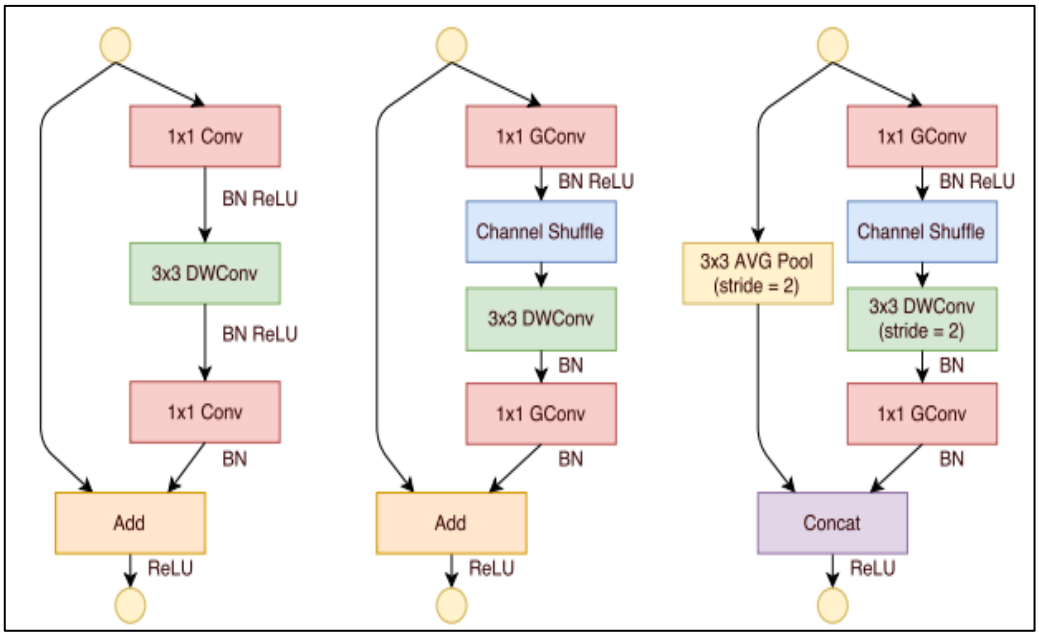


Figure 3.16: ShuffleNet bottleneck unit with depthwise convolution (DWConv), ShuffleNet unit with pointwise group convolution (GConv) and channel shuffle and ShuffleNet unit with stride = 2

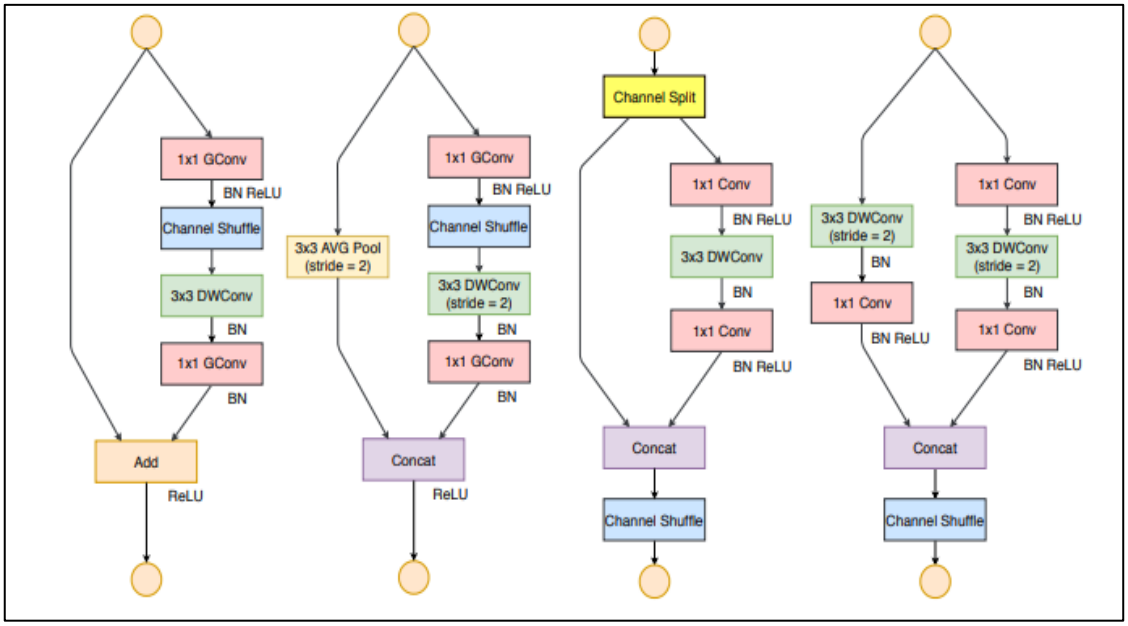


Figure 3.17: The basic ShuffleNet unit, the ShuffleNet unit for spatial down sampling (2×), ShuffleNet V2 basic unit basic unit and ShuffleNet V2 unit for spatial down sampling (2×). DWConv: depthwise convolution. GConv: group convolution

While in this study, the improvements to the ShuffleNet architecture are guided by two key principles that are to enhance feature extraction capabilities without significantly increasing computational cost and to maintain the efficiency of channel shuffling while improving information flow across channel groups.

Based on an analysis of the original ShuffleNet architecture and recent advancements in efficient network design (Zhang et al., 2020; Ma et al., 2018) the practical guidelines are proposed. Maintaining a balance between input and output channels in convolution operations enhances feature extraction, while introducing parallel processing pathways for information improves the network's ability to capture diverse features. Efficient channel mixing, achieved by enhancing the channel shuffling mechanism, is crucial for better information exchange across channel groups. Additionally, carefully placed residual connections aid gradient flow and improve gradient propagation without significantly increasing computational cost.

Then, based on the above guidelines, the improved ShuffleNet architecture with the following key modifications are presented:

1. Channel Split Operation

Channel Split operation at the beginning of each ShuffleNet unit. This operation divides the input tensor $X \in R\hat{}(c \times h \times w)$ into two branches:

X_1= X[:c', :, :] Equation 3.7

X_2= X[c':, :, :] Equation 3.8

where c' is typically set to $c/2$. This split allows for parallel processing of information, potentially capturing different aspects of the input features.

2. Additional Convolutional Layer

An additional pointwise (1x1) convolutional layer is applied to one branch of the split channels:

$$Y = \text{Conv2}(X_1) \qquad \text{Equation 3.9}$$

where $Y \in R^{(c'' \times h \times w)}$, and c" > c'. This layer increases the channel dimension, enhancing feature extraction capabilities without significantly increasing computational cost.

3. Enhanced Channel Shuffling

The channel shuffling mechanism is improved to ensure better mixing of information across channel groups. The enhanced shuffling operation described as:

$$Z = \text{Shuffle}(\text{Concat}(Y, X_2)) \qquad \text{Equation 3.10}$$

Here, the outputs of the two branches, Y and X_2, are first concatenated along the channel dimension. The resulting feature map is then passed through the Shuffle function. In this function, the concatenated channels are initially divided into groups, and then a permutation is applied so that channels from different groups are interleaved. The process redistributes feature maps across groups, ensuring that each group in the next layer receives information from multiple groups of the previous layer. This refines the grouping and rearrangement strategy to maximize cross-group information flow, thereby enhancing representational diversity while keeping computational cost low.

4. Residual Connection

A carefully designed residual connection that aids in gradient flow without significantly increasing computational cost is introduced:

$$\text{Output} = Z + X \qquad \text{Equation 3.11}$$

This residual connection helps in maintaining information from earlier layers and improves gradient propagation during training.

By introducing the channel split operation, additional convolutional layer, enhanced channel shuffling, and carefully placed residual connections, the improved ShuffleNet architecture incorporates these key modifications to enhance feature extraction capabilities and overall efficiency as shown in Figure 3.18.

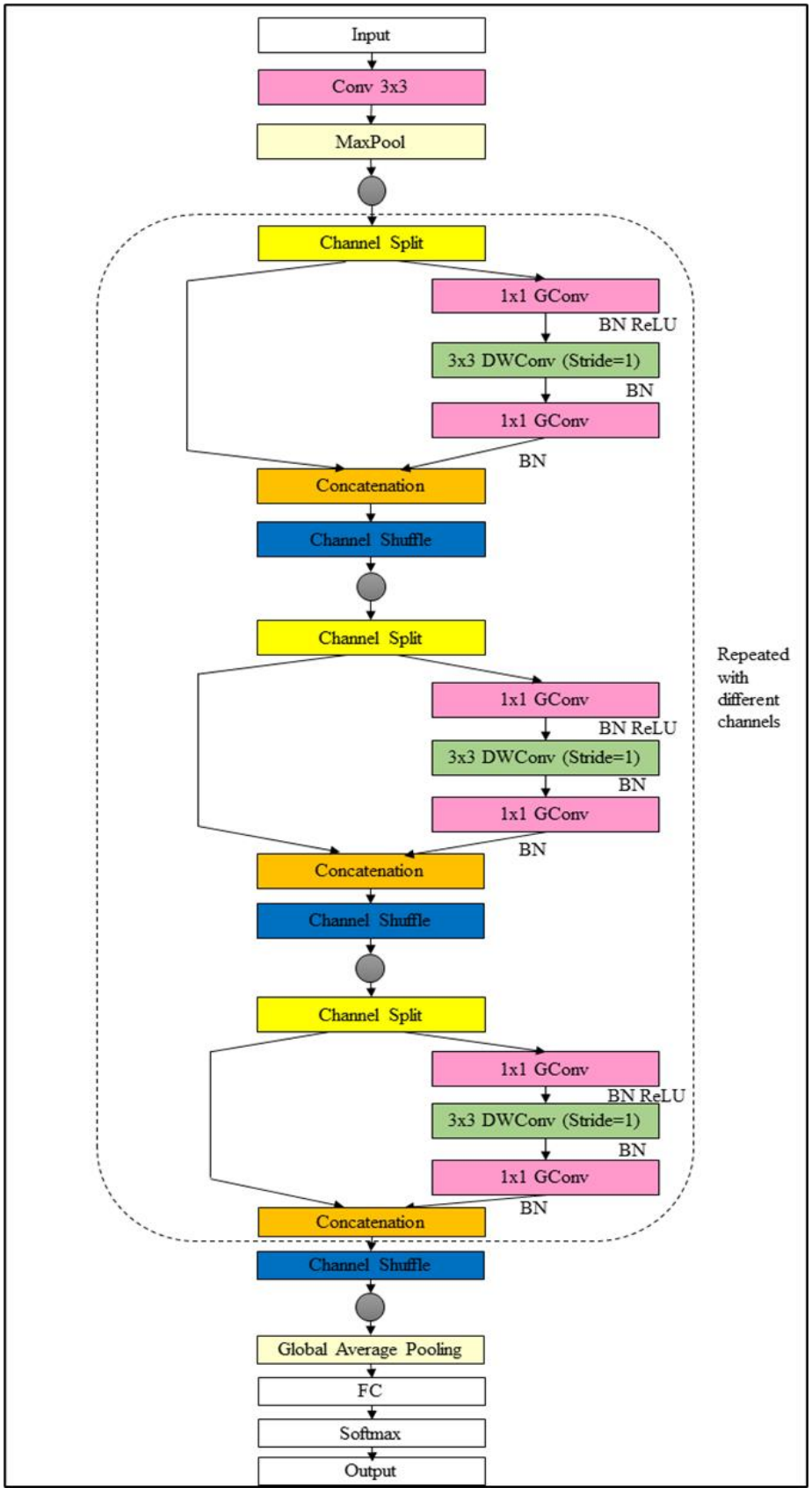


Figure 3.18: Improved ShuffleNet architecture with repeated units using channels

Then, four more elements added to this architecture as shown in Figure 3.19. This figure depicts the image augmentation and preparation pipeline in which the first two elements added to the ShuffleNet architecture are the Squeeze-and-Excitation (SE) Block and Layer Normalization. The SE Block is inserted after the final Channel Shuffle operation and before the Global Average Pooling layer. It begins with a Global Average Pooling operation that reduces the spatial dimensions of the feature map to 1x1, creating a channel descriptor. This descriptor then passes through two Fully Connected (FC) layers: the first FC layer reduces the channel dimension by a factor (16), followed by a ReLU activation, while the second FC layer restores the original channel dimension, followed by a sigmoid activation.

The resulting channel-wise weights are then used to recalibrate the input feature map through element-wise multiplication. This process allows the network to adaptively recalibrate channel-wise feature responses, boosting important features and suppressing less useful ones. Following the SE Block, Layer Normalization is applied. Unlike Batch Normalization, which normalizes across the batch dimension, Layer Normalization operates on each sample independently, normalizing across the feature dimension. This helps stabilize the training process, especially for deeper networks or when using smaller batch sizes.

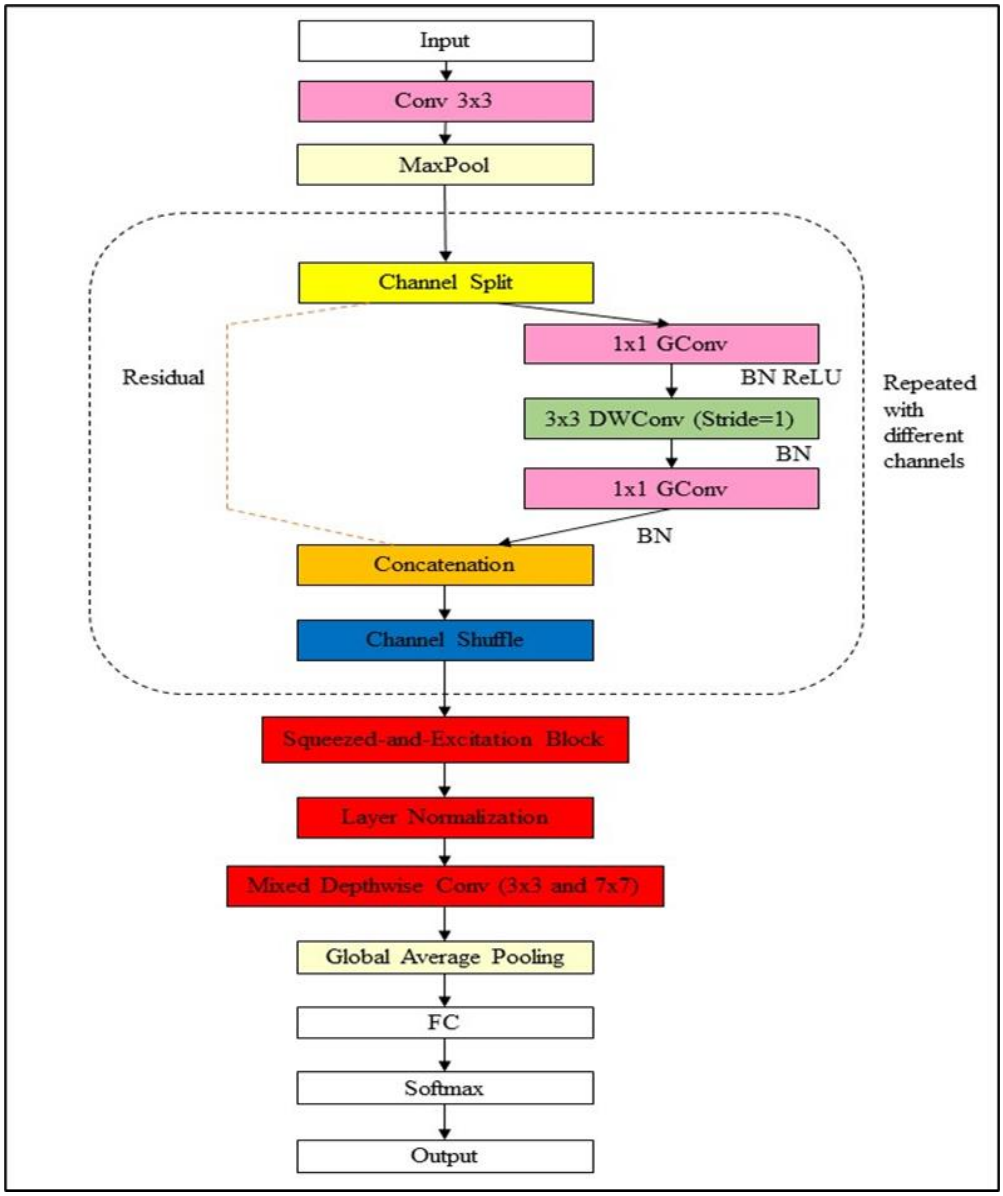


Figure 3.19: Improved ShuffleNet architecture with added four elements

The other two elements added are the Residual Connection and the Mixed Depthwise Convolution. The Residual Connection is represented by a skip connection that bypasses the main convolution operations within a ShuffleNet unit. It allows the input of the unit to be directly added to its output, facilitating better gradient flow through the network during

backpropagation. This is beneficial for training very deep networks because it helps mitigate the vanishing gradient problem.

The Mixed Depthwise Convolution replaces the standard 3×3 Depthwise Convolution in the ShuffleNet unit. It consists of two parallel depthwise convolutions with different kernel sizes: 3×3 and 7×7. The outputs of these convolutions are concatenated, allowing the network to capture features at multiple scales simultaneously. This multi scale feature extraction enhances the network's ability to recognize patterns of varying sizes in the input, potentially improving its overall performance on tasks like object detection or image classification.

The number of channels increases in each repetition (116 to 232 to 464). This gradual increase allows the network to build up from simple, low-level features to more complex, high-level features. The three parts of the image correspond to these three repetitions:

1. Part 1 (116 channels): Initial feature extraction, capturing basic patterns.
2. Part 2 (232 channels): Intermediate features, building upon the initial extraction.
3. Part 3 (464 channels): High-level features, capturing the most complex patterns.

This structure allows the improved ShuffleNet to efficiently extract a rich hierarchy of features for plant leaf diseases detection. Table 3.7 and Table 3.8 show the architecture of ShuffleNet and ShuffleNet V2. While the improved ShuffleNet architecture depicted in the Table 3.9 along with Table 3.10 to tabulate the details of the layers added to the improved ShuffleNet. It starts with an input image of 224×224×3, followed by an initial convolution layer (Conv1) that reduces spatial dimensions while increasing channel depth to 24. A max pooling layer further down samples the feature maps.

Table 3.7: The ShuffleNet architecture quantifies complexity using FLOPs, which represent the total count of floating-point multiplication and addition operations performed

Layer	Output Size	KSize	Stride	Repeat	Output channels (*g* groups)				
					g=1	g=2	g=3	g=4	g=8
Image	224x224	-	-	-	3	3	3	3	3
Conv1	112x112	3x3	2	1	24	24	24	24	24
MaxPool	56x56	3x3	2						
Stage 2	28x28	-	2	1	144	200	240	272	384
	28x28		1	3	144	200	240	272	384
Stage 3	14x14	-	2	1	288	400	480	544	768
	14x14		1	7	288	400	480	544	768
Stage 4	7x7	-	2	1	576	800	960	1088	1536
	7x7		1	3	576	800	960	1088	1536
GlobalPool	1x1	7x7	-	-	-	-	-	-	-
FC	-	-	-	-	1000	1000	1000	1000	1000
Complexity	-	-	-	-	143M	140M	137M	133M	137M

Table 3.8: ShuffleNet v2's overall architecture, with four levels of complexity

Layer	Output Size	KSize	Stride	Repeat	Output channels			
					0.5x	1x	1.5x	2x
Image	224x224	-	-	-	3	3	3	3
Conv1	112x112	3x3	2	1	24	24	24	24
MaxPool	56x56	3x3	2					
Stage 2	28x28	-	2	1	144	200	240	272
	28x28		1	3	144	200	240	272
Stage 3	14x14	-	2	1	288	400	480	544
	14x14		1	7	288	400	480	544
Stage 4	7x7	-	2	1	576	800	960	1088
	7x7		1	3	576	800	960	1088
Conv5	7x7	1x1	1	1	1024	1024	1024	2048
GlobalPool	1x1	7x7	-	-	-	-	-	-
FC	-	-	-	-	1000	1000	1000	1000
FLOPs	-	-	-	-	41M	146M	299M	591M
# of Weights	-	-	-	-	1.4M	2.3M	3.5M	7.4M

Table 3.9: Improved ShuffleNet's architure

Layer	Output Size	Kernel Size	Stride	Rep-eat	Output channels (g=1)	Output channels (g=2)	Output channels (g=3)	Output channels (g=4)	Output channels (g=8)	FLOPs
Image	224×224	-	-	-	3	3	3	3	3	-
Conv1 MaxPool	112×112 56×56	3×3 3×3	2 2	1 1	24 24	24 24	24 24	24 24	24 24	- -
Stage2	28×28	3×3 3×3	2 1	1 3	116 116	116 116	116 116	116 116	116 116	- -
Stage3	14×14	3×3 3×3	2 1	1 7	232 232	232 232	232 232	232 232	232 232	- -
Stage4	7×7	3×3 3×3	2 1	1 3	464 464	464 464	464 464	464 464	464 464	- -
GlobalPool	1×1	7×7	-	-	464	464	464	464	464	-
FC	-	-	-	-	1000	1000	1000	1000	1000	-
Complexity	-	-	-	-	524M	524M	524M	524M	524M	-

Table 3.10: Details of SE Block, Layer Normalization, and Mixed Depthwise Convolution architecture

Layer	Output Size	Kernel Size	Stride	Rep-eat	(g=1)	(g=2)	(g=3)	(g=4)	(g=8)	FLOPs
Image	224×224	-	-	-	3	3	3	3	3	-
Conv1	112×112	3×3	2	1	24	24	24	24	24	339768
MaxPool	56×56	3×3	2	1	24	24	24	24	24	-
SE Block	Same as Input	-	1	1	24	24	24	24	24	1.125
Reduction (r=16)	C/16 ≈ 1	-	-	1	2	2	2	2	2	-
Stage2	28×28	3×3	2	3	116	116	116	116	116	21168
Stage3	14×14	3×3	2	7	232	232	232	232	232	5292
Stage4	7×7	3×3	2	3	464	464	464	464	464	1323

Table 3.10 continued

GlobalPool	1×1	7×7	-	-	464	464	464	464	464	-
Mixed Depthwise Conv	Same as Input	3×3, 7×7	1	1	24	48	72	96	192	2646
FC	-	-	-	-	1000	1000	1000	1000	1000	3000
Complexity	-	-	-	-	524M FLOPs	524M FLOPs	524M FLOPs	524M FLOPs	524M FLOPs	-

The core structure of the network is built upon three primary stages: Stage2, Stage3, and Stage4. Each of these stages is composed of multiple shuffle units, which incorporate 3×3 depthwise separable convolutions to enhance computational efficiency. These stages are designed to progressively increase the channel depth while simultaneously reducing the spatial dimensions, ensuring a hierarchical extraction of features at different levels. Specifically, Stage2, Stage3, and Stage4 expand the number of channels from 116 to 232 and then to 464, while the corresponding spatial resolutions decrease from 28×28 to 14×14 and finally to 7×7.

To optimize parameter usage and computational efficiency, the network adopts a group convolution strategy, which enables different grouping configurations ranging from g=1 to g=8. This flexibility allows the architecture to dynamically adjust its internal operations, making it well-suited for a variety of computational constraints without sacrificing performance.

At the final stage of processing, the network concludes with a global average pooling layer, which condenses the extracted features into a compact representation, followed by a fully connected layer consisting of 1000 output units, making it ideal for classification tasks.

Additionally, the table highlights the adaptability of this architecture across different computational budgets. Despite variations in group configurations, the network maintains a consistent computational complexity of 524M FLOPs, underscoring its ability to balance performance and efficiency while ensuring optimal resource utilization.

The expansion of the red highlight blocks in the Figure 3.18 is illustrated in Figure 3.19. The SE Block enhances feature extraction by recalibrating channel-wise responses through Global Average Pooling, two fully connected layers, and element-wise multiplication to emphasize important features. Layer Normalization stabilizes training by normalizing feature values for each sample independently, ensuring smoother gradient flow even with small batch sizes. The Mixed Depthwise Convolution improves multi-scale feature extraction by processing features with parallel 3×3 and 7×7 convolutions and concatenating their outputs, enabling recognition of patterns at varying scales.

Figure 3.20 shows the connection of SE Block, Layer Normalization and Mixed Depthwise Convolution, how they are connected to the Global Average Pooling. The SE Block utilizes Global Pooling to compute a channel-wise descriptor by condensing spatial

information into a single value per channel, representing its global importance. This descriptor is processed through a FC layer followed by ReLU activation, another FC layer, and Sigmoid activation to generate channel-wise weights. These weights modulate the input feature maps, enhancing the most important channels.

The outputs of SE blocks further refined with Layer Normalization, which computes the mean and variance across features to normalize them, then applies a scale and shift through learnable parameters, often in conjunction with a Feed Forward Network and Addition for enhanced representational power. Mixed Depthwise Convolutions, incorporating 3×3 and 7×7 Depthwise Convolutions, process the feature maps, whose concatenated outputs are again aggregated via Global Average Pooling for efficient channel-wise representation.

These modular components can be connected to any CNN model to enhance feature representation and overall performance. The SE block applies global average pooling and a learnable gating mechanism to recalibrate channel-wise features, making it universally compatible with any CNN architecture (Hu et al., 2018). Layer Normalization normalizes activations along feature dimensions to stabilize gradient flow, offering robust performance in CNN models regardless of batch size constraints (Ba et al., 2016). While Mixed-Depthwise Convolutions combine 3×3 and 7×7 depth wise kernels to enhance receptive fields and feature extraction, seamlessly integrating into any CNN backbone (Tan et al., 2019).

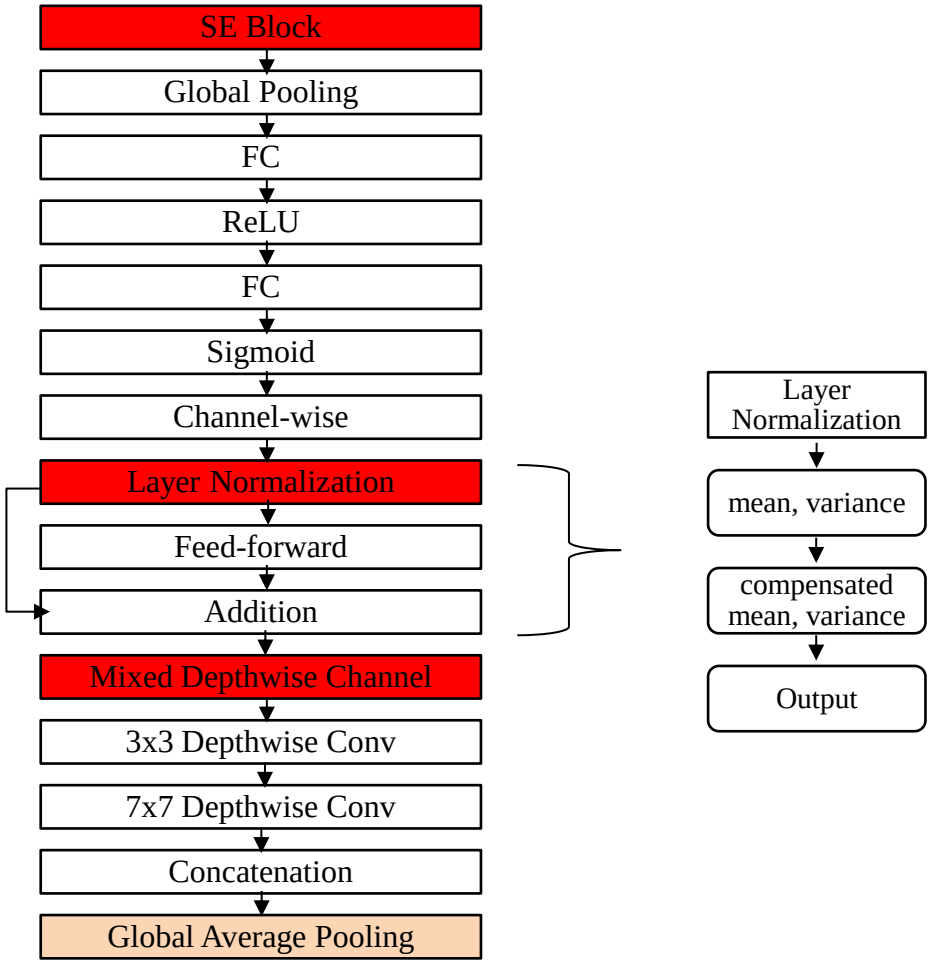


Figure 3.20: Connection of SE Block, Layer Normalization and Mixed Depthwise Convolution to Global Average Pooling

3.7 Splitting of Dataset

The dataset of plant leaf diseases, which contains capsicum, rice, corn, tomato, and citrus, is split into three classes of diseases: grey spot, discolour leaf, and leaf curling, with a split style of 70% for training, 15% for validation, and 15% for testing. The splitting of the dataset for the plant leaf diseases mentioned in this study are summarized in Table 3.11.

During training, the model encounters diverse samples from each class, enabling it to learn the patterns and characteristics of various leaf diseases. The validation set helps fine-

tune hyperparameters and monitor performance, while the testing set evaluates the final model's generalization and accuracy on unseen data. Following this distribution results in a balanced representation of the various diseases within each plant type, which helps to prevent bias and ensures that the model can effectively classify the diseases across the different crops.

Table 3.11: Splitting of dataset for plant leaf diseases detection

Types of plants	**Types of diseases**	**Total**	**Training**	**Validation**	**Testing**
Capsicum	Grey spot	1334	934	200	200
	Discolour leaf	1240	868	186	186
	Leaf curling	936	655	140	140
Rice	Grey spot	1035	724	155	156
	Discolour leaf	712	498	107	107
Corn	Grey spot	1297	907	194	196
Tomato	Grey spot	852	596	127	129
	Discolour leaf	818	572	122	124
	Leaf curling	746	522	111	113
Citrus	Grey spot	497	347	75	75
	Discolour leaf	359	251	53	55

In deep learning research, the 70-15-15 split is a widely recognised and widely used approach. This splitting style is widely used by researchers and practitioners as a standard practise, which endorses consistency and facilitates comparisons between studies. The 70% allocation for training provides a substantial amount of data to effectively train the model as this larger training set allows the model to learn more diverse patterns and generalisations, resulting in improved performance (Joseph, 2022). The remaining 30% is split evenly between validation and testing, allowing for accurate model evaluation and performance prediction (Varoquaux & Colliot, 2023). While a 15% testing set increases the likelihood of obtaining reliable performance estimates even with a smaller dataset (Korjus et al., 2016). Meanwhile, the validation set (15%) is critical for hyperparameter tuning, which is the process of adjusting the model's settings to improve its performance (Pannakkong et al., 2022). A larger validation set provides more reliable feedback for making informed hyperparameter selection decisions (Waring et al., 2020).

3.8 Performance Evaluation

The model's performance on the testing set is evaluated after training. Pre-processed test images are fed into the model to generate predicted class labels, which are then compared to the expert-verified ground truth labels to calculate accuracy. For multiclass classification, such as detecting various plant leaf diseases, the accuracy formula is adjusted accordingly:

$$Accuracy = \frac{Number\ of\ correctly\ classified\ samples}{Total\ number\ of\ samples} \qquad \text{Equation 3.12}$$

In a multiclass setting, a sample is correctly classified when its predicted class matches the true class. Specificity is then calculated using true negative (TN) and false positive (FP) values:

$$Specificity = \frac{(TN)}{(TN+FP)} \qquad \text{Equation 3.13}$$

For specificity in a multiclass scenario, sum the true negatives and false positives, then divide by the total true negatives. Sensitivity is calculated using true positives and false negatives:

$$Sensitivity = \frac{(TP)}{(TP+FN)} \quad \text{Equation 3.14}$$

To calculate sensitivity, add up true positives and false negatives, then divide by the total true positives. Create a confusion matrix using predicted and actual labels, which shows true positives, true negatives, false positives, and false negatives. Precision for each class is then the ratio of correctly predicted positives to all predicted positives as follows:

$$Precision = \frac{TP\ for\ the\ class}{TP\ for\ the\ class+FP\ for\ the\ class} \quad \text{Equation 3.15}$$

TP (true positive) is the number of times the model correctly identifies a sample as belonging to a specific class (diseased leaves), while FP (false positive) is the number of times it incorrectly identifies a sample as that class. Precision measures how well the model identifies each class among all samples predicted as positive. Recall for each class is calculated as follows:

$$Recall\ =\ TP\ /\ (TP\ +\ FN) \quad \text{Equation 3.16}$$

Recall ranges from 0 to 1, with higher values meaning fewer false negatives. The F1 score, the harmonic mean of precision and recall, measures overall performance. Here's how to calculate it for each class:

$$F1\ Score\ =\ (Precision\ *\ Recall)\ /\ (Precision\ +\ Recall) \quad \text{Equation 3.17}$$

The F1 score combines precision and recall into one number. Higher values mean better performance. Next, calculate the negative predictive value (NPV), false omission rate, miss rate, and fall-out. Start with a confusion matrix to get true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). NPV shows the proportion of correct

negative predictions among all predicted negatives (Weng et al., 2023). Use this formula to find NPV for each class:

$$NPV = \frac{TN\ for\ the\ class}{TN\ for\ the\ class + FN\ for\ the\ class} \quad \text{Equation 3.18}$$

TN is when the model correctly identifies a sample as healthy, while FN is when it mistakenly labels a healthy sample as diseased. This measures how well the model identifies healthy leaves. Next, calculate the False Omission Rate (FOR), which is the proportion of healthy samples incorrectly classified as diseased (Sivaraman & Khanna, 2021). The FOR for each class is calculated as follows:

$$FOR = (FN) / (FN + TN) \quad \text{Equation 3.19}$$

FOR values range from 0 to 1, with lower values meaning fewer FN. Next, calculate the miss rate, which is the proportion of diseased samples incorrectly predicted as healthy (Soeb et al., 2023). Use the following formula to find the miss rate for each class:

$$Miss\ Rate = (FN) / (FN + TP) \quad \text{Equation 3.20}$$

Miss rate values range from 0 to 1, with higher values indicating more FN. Next, calculate the fall-out, which measures the proportion of healthy samples incorrectly predicted as diseased (Monaghan et al., 2021). Use the following formula to find the fall-out for each class:

$$FP / (FP + TN) = Fall - out \quad \text{Equation 3.21}$$

Fall-out ranges from 0 to 1; higher values mean more FN. Evaluate the model using NPV, FOR, miss rate, and fall-out. Lower FOR and miss rate indicate fewer FN, while higher fall-out means more FP (Trevethan, 2017). For imbalanced data, accuracy alone isn't enough. Use specificity, sensitivity, precision, recall, F1 score, area under the receiver operating characteristic curve (AUC-ROC) and Matthews correlation coefficient (MCC) for a complete assessment. These metrics help measure how well the model performs across

different classes. In this study, accuracy is determined on a region-based approach, where the algorithm evaluates and classifies the detected diseased regions of the leaf, rather than only pixel-level features. This ensures the performance evaluation reflects practical disease identification in real agricultural scenarios.

3.9 Summary

This chapter explains the study's methodology, including dataset details, image processing (resizing and augmentation), hyperparameter selection, and modifications to ShuffleNet. It covers dataset splitting for training, validation, and testing, and how the improved ShuffleNet's performance is evaluated. The next chapter discusses the results of detecting plant leaf diseases.

CHAPTER 4

RESULTS AND DISCUSSION

4.1 Overview

The methodology for achieving the desired objectives is presented in the previous chapter. This chapter covers the information and results obtained.

The following is the chapter outline. The data acquisition devices are compared in Section 4.2. Section 4.3 discusses the benefits of data augmentation, and Section 4.4 addresses the performance of an improved ShuffleNet model. While Section 4.5 covers performance evaluation using a confusion matrix. Finally, Section 4.6 summarises the chapter.

4.2 Generation of Bounding Box using RPN for Diseased Areas

The devices employed in this study were the Kinect camera, DSLR camera and mobile camera, with all experiments conducted on the same computer (Acer Aspire E14) to ensure consistency. Using Region Proposal Networks (RPN), bounding boxes were generated for diseased capsicum leaves across different resolutions. A total of 1334 grey spot, 1240 discolour leaf and 936 leaf curling images of capsicum were used, alongside rice, corn, tomato and citrus images from publicly available datasets. Kinect achieved the highest accuracies (87.02%–90.67%), DSLR performed moderately well (85.28%–88.86%), while the mobile camera produced slightly lower accuracies (81.02%–85.59%). These outcomes confirm that RPN is robust in isolating diseased regions but also sensitive to input resolution and noise.

The comprehensive detection results are provided in Table 4.1. Each row of values in Table 4.1 represents the performance metrics (accuracy, precision, recall, F1 score, etc.)

of the RPN model applied to detect the three types of capsicum leaf diseases (grey spot, discoloured leaf and leaf curling) under uniform and complex background conditions. The results in Tables 4.2 to 4.4 show that while RPN achieved promising accuracy levels across all devices, performance varied depending on image quality.

Table 4.1: Detection results of capsicum diseases using RPN under uniform and complex background conditions

Type of diseases	Accuracy	Precision	Recall	F1 Score	Specificity	Sensitivity	NPV	FOR	MCC	FDR	True Positive Rate	False Positive Rate	True Negative Rate	False Negative Rate
Grey Spot	99.207	94.126	96.444	97.834	91.315	93.466	97.726	96.050	95.371	92.581	92.753	94.751	94.092	92.552
Discolour Leaf	94.741	97.160	93.579	98.549	95.756	94.517	90.281	90.090	98.455	94.590	91.770	92.201	93.132	97.826
Leaf curling	98.029	94.571	96.449	90.445	97.554	91.646	96.294	91.079	91.566	94.394	98.420	95.628	91.934	97.880

Compared with existing studies, the present results are consistent with the proven strengths of RPN but slightly lower in accuracy. Zhang et al., (2022), for example, reported 97.28% accuracy in apple leaf disease detection using a two-stage RPN. Their higher performance suggests that refinement through multi-stage proposals can overcome issues such as overlapping leaves and inconsistent lighting. Similarly, Liu et al., (2023) achieved 98.26% accuracy and a 35% reduction in false positives by combining RPN with progressive learning, whereas the current study's single-stage RPN without additional refinement yielded accuracies below 91%. This gap highlights that while RPN alone is effective, integrating advanced learning strategies significantly boosts reliability.

Wang et al., (2024) further enhanced RPN by integrating feature pyramid networks and attention mechanisms, improving accuracy to 96.75% and sensitivity to small diseased spots by 23%. In contrast, this study showed that small and complex symptoms, such as leaf curling, had slower convergence and slightly reduced accuracy, particularly in mobile images. This indicates that the refinements introduced by Wang et al., (2024) could be valuable for addressing the limitations observed here, such as background noise and disease structural variability.

The robustness of the RPN in bounding box generation is largely due to its anchor-based mechanism, which systematically scans multiple scales and aspect ratios across the image. This strategy ensures that diseased leaf regions can still be localized even when parts of the leaf are hidden behind other leaves or when shadows reduce visibility. By relying on convolutional feature maps rather than raw pixel intensity alone, the RPN captures textural and structural patterns that remain consistent regardless of lighting variations. This enables the model to maintain high detection accuracy even in environments with fluctuating brightness or partial occlusion.

Another reason for the strong performance lies in the RPN's integration of shared convolutional layers, which enhances feature reuse and reduces sensitivity to background clutter. Instead of treating the entire image as a single detection problem, the RPN generates candidate regions that highlight only the most discriminative parts of diseased leaves. This allows it to ignore irrelevant information such as overlapping foliage or complex backgrounds, while focusing on features unique to grey spots, discoloration, or curling. As a result, the bounding box generation process not only improves detection accuracy but also ensures that downstream classification receives high-quality inputs, ultimately boosting the overall reliability of the disease detection framework.

Table 4.2: The generation of bounding boxes for diseased areas using RPN for images captured by the Kinect camera under complex background conditions

	Disease	**Uniform background**	**Complex background**	**Accuracy (%)**	**Precision (%)**	**Recall (%)**
Capsicum	Grey spot			87.02	91.73	87.14

Table 4.2 continued

	Discolour leaf			89.79	91.97	90.00

Table 4.2 continued

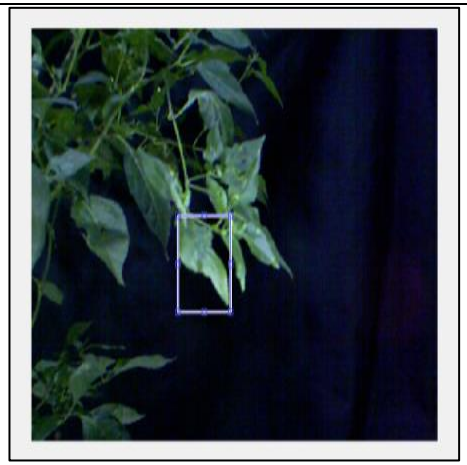
	Leaf curling			90.67	92.03	90.71

Figure 4.1 shows the training loss curves for the three disease classes (grey spot, discolour leaf, and leaf curling). The grey spot curve decreases from 1.80 at epoch 1 to 0.12 by epoch 30, stabilizing near 0.08 toward epoch 50, which indicates efficient feature extraction and fast convergence. The discolour leaf curve falls from 1.60 at epoch 1 to 0.10 by epoch 25, reflecting high feature separability and early convergence. The leaf curling curve declines more gradually from 1.70 at epoch 1 to 0.15 by epoch 40, and stabilizes around 0.13 thereafter, consistent with higher intra-class variance and structural complexity. The smooth overall trends and only minor oscillations confirm robust gradient updates and effective generalization across classes.

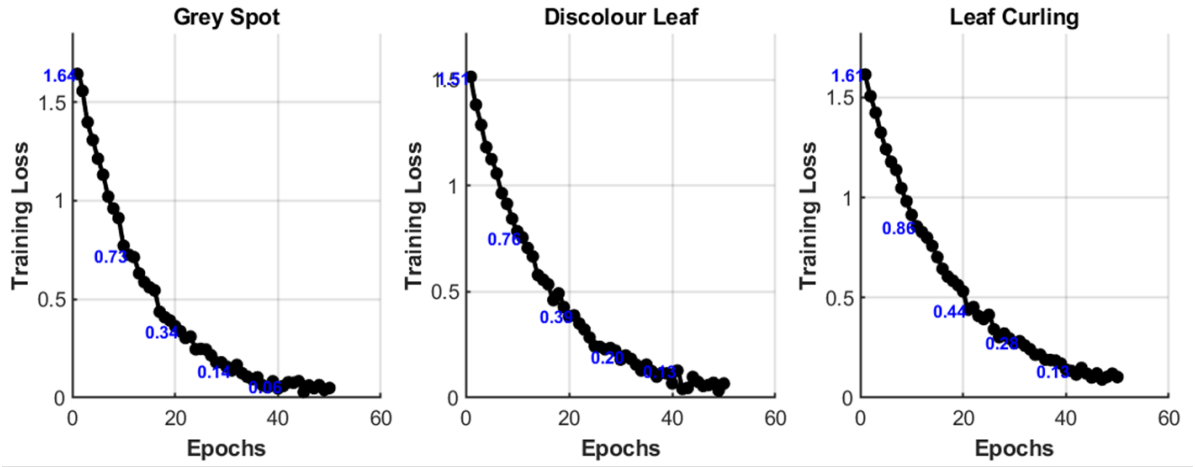


Figure 4.1: The training loss curves reveal the model's convergence for grey spot, discolour leaf, and leaf curling, highlighting their complexity

Table 4.3: The generation of bounding boxes for diseased areas using RPN for images captured by the DSLR camera under complex background conditions

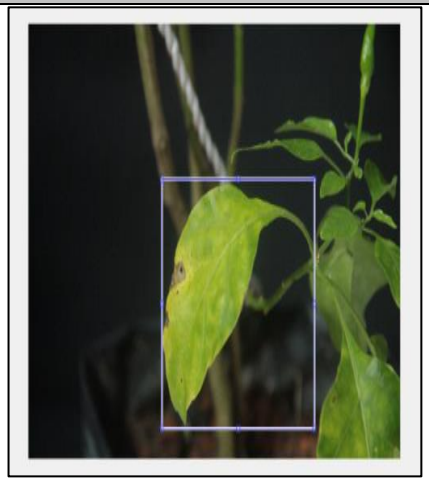
	Disease	**Uniform background**	**Complex background**	**Accuracy (%)**	**Precision (%)**	**Recall (%)**
Capsicum	Grey spot			85.28	89.90	85.40

Table 4.3 continued

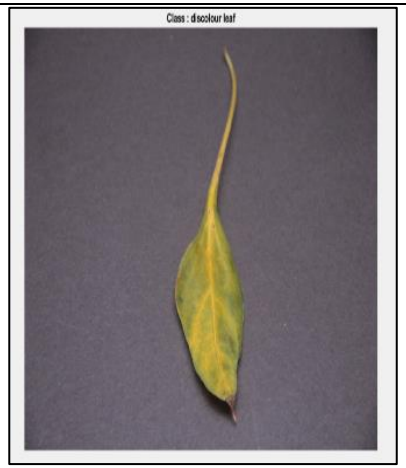
	Discolour leaf			87.99	90.13	88.20

Table 4.3 continued

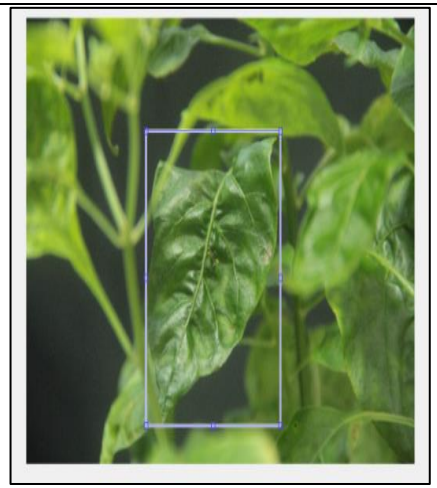
	Leaf curling			88.86	90.19	88.90

Figure 4.2 presents the training loss curves for grey spot, discolour leaf, and leaf curling using DSLR images, highlighting a slightly slower convergence compared to the Kinect-based results. For grey spot, the loss decreases from 2.00 to 0.15 within 35 epochs, reflecting higher initial complexity but eventual efficient learning. The discolour leaf curve drops from 1.80 to 0.12 by 30 epochs, showing stable feature extraction and faster convergence than the other two diseases. Meanwhile, the leaf curling curve reduces gradually from 1.90 to 0.18, stabilizing only after around 45 epochs due to greater intra-class variation and irregular patterns in the symptoms. Overall, the smooth downward progression of the curves confirms effective optimization, while the minor oscillations indicate adaptive gradient updates that support strong generalization.

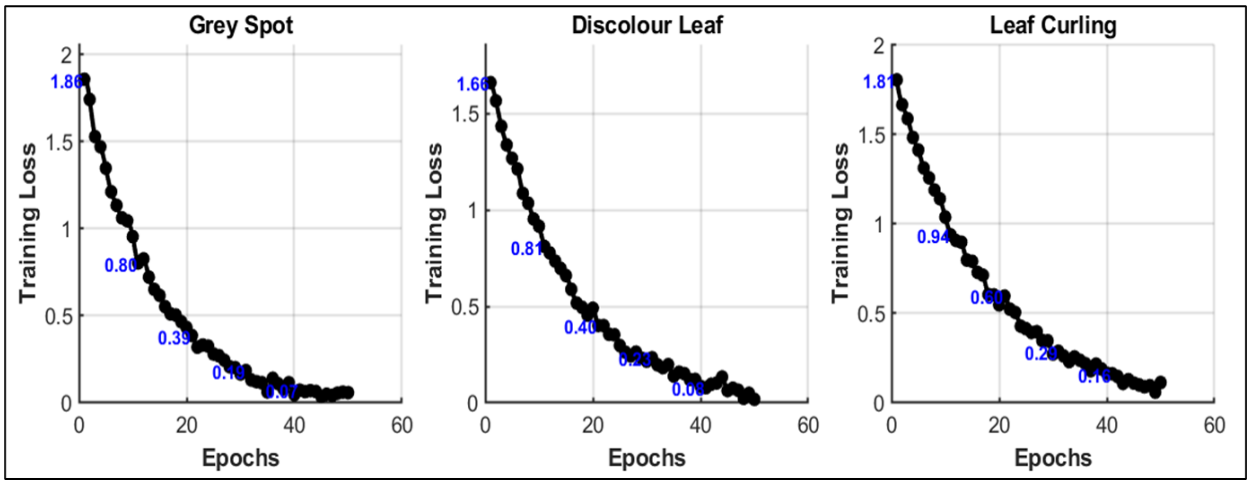


Figure 4.2: The training loss curves for DSLR illustrate the model's learning process for grey spot, discolour leaf, and leaf curling, showing slightly slower convergence than Kinect

Table 4.4: The generation of bounding boxes for diseased areas using RPN for images captured by the mobile camera under complex background conditions

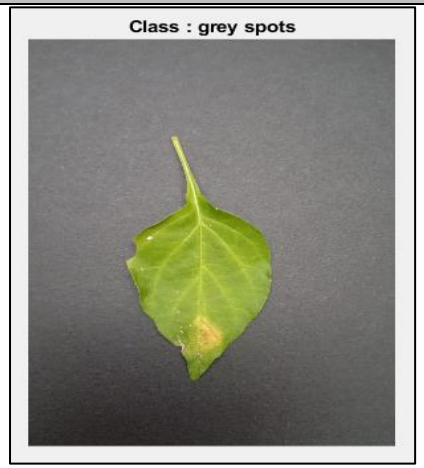
	Disease	**Uniform background**	**Complex background**	**Accuracy (%)**	**Precision (%)**	**Recall (%)**
Capsicum	Grey spot			81.02	85.40	81.13

Table 4.4 continued

	Discolour leaf			85.59	85.62	83.79

Table 4.4 continued

	Leaf curling			84.41	85.68	84.45

The training loss curves for the tablet camera indicate a slower convergence for grey spot, discolour leaf, and leaf curling, reflecting the effect of slightly lower image resolution and subtle noise (Figure 4.3). Grey spot loss decreases from 2.25 to 0.19 within 42 epochs, suggesting that extracting distinct features is somewhat more difficult. Discolour leaf loss reduces from 2.05 to 0.17 by 37 epochs, showing moderate feature separability. Leaf curling loss drops from 2.18 to 0.21, stabilizing around 46 epochs due to variability within the class. The overall slower convergence emphasizes that lower image clarity requires additional training iterations to achieve accurate generalization.

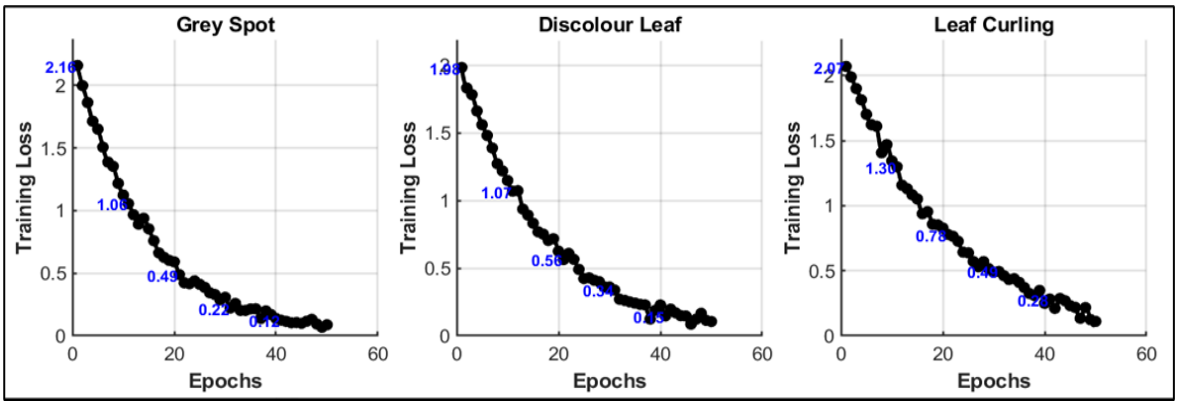


Figure 4.3: The training loss curves for the phone camera demonstrate slower convergence for grey spot, discolour leaf, and leaf curling, reflecting increased noise and lower image clarity

4.3 Data Augmentation Using Random Reflection Technique

The random reflection technique improves the model's generalization by randomly flipping capsicum images horizontally and vertically during training with each image mirrored with a probability of 0.5% as shown in Table 4.5. This method is simple and efficient, requiring only the original capsicum leaf data. In contrast, complex data augmentation technique such as style transfer and GANs need additional collections of images in various styles (Alomar et al., 2023).

The results in Table 4.6 demonstrate that applying the random reflection technique consistently improves classification accuracy across all camera types. For Kinect images, accuracy increased from 89.16% to 91.94%, while DSLR images improved from 87.38% to 89.77%. Mobile images, which originally achieved the lowest accuracy of 83.97%, also showed improvement to 85.21% after augmentation. These gains, though modest, highlight the effectiveness of even simple augmentation methods in improving ShuffleNet's generalization ability.

The improvements can be attributed to the fact that random reflection introduces additional variability in the training set, enabling the model to learn more robust feature representations. This is particularly beneficial for datasets with class imbalance or limited disease samples, as reflected in capsicum leaf images collected under different conditions. Similar findings have been reported by Shorten and Khoshgoftaar (2019), who emphasized that reflection-based augmentation reduces overfitting by increasing image orientation diversity.

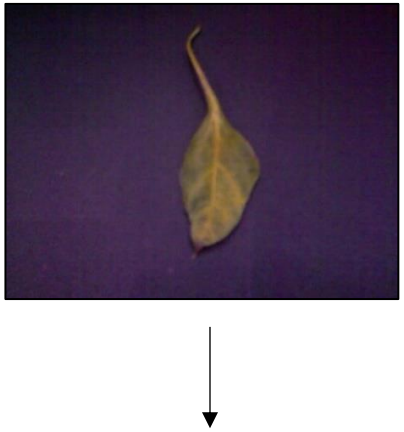
Compared to more complex augmentation strategies such as generative adversarial networks (GANs) or style transfer, random reflection offers the advantage of simplicity and low computational cost while still enhancing accuracy (Alomar et al., 2023). This makes it

especially suitable for lightweight models like ShuffleNet, which are optimized for mobile and embedded devices (Zhang et al., 2020).

Moreover, the performance improvement is particularly valuable for mobile datasets, where image quality is often affected by lighting, motion blur, or low resolution. The gains observed indicate that augmentation can compensate for these inconsistencies by enriching the training distribution. As Kim et al., (2021) noted, augmentation techniques are essential for balancing biased datasets and improving model generalization in real-world agricultural applications.

The integration of the random reflection technique into ShuffleNet training significantly improves classification accuracy across different image acquisition devices. Although the improvement margins vary, the findings confirm that simple data augmentation methods can effectively enhance the robustness of compact CNN models for leaf disease detection.

Table 4.5: Performance results after data augmentation. Each original image was augmented into 68 variations through a combination of random reflection (horizontal and vertical flipping), rotation at multiple angles, scaling and translation. While some augmented leaves may appear visually similar due to overlapping transformations, all were retained to enhance dataset diversity and improve model robustness

First step	Second step
Resize image to 224×224 	Flipping capsicum image horizontally and vertically

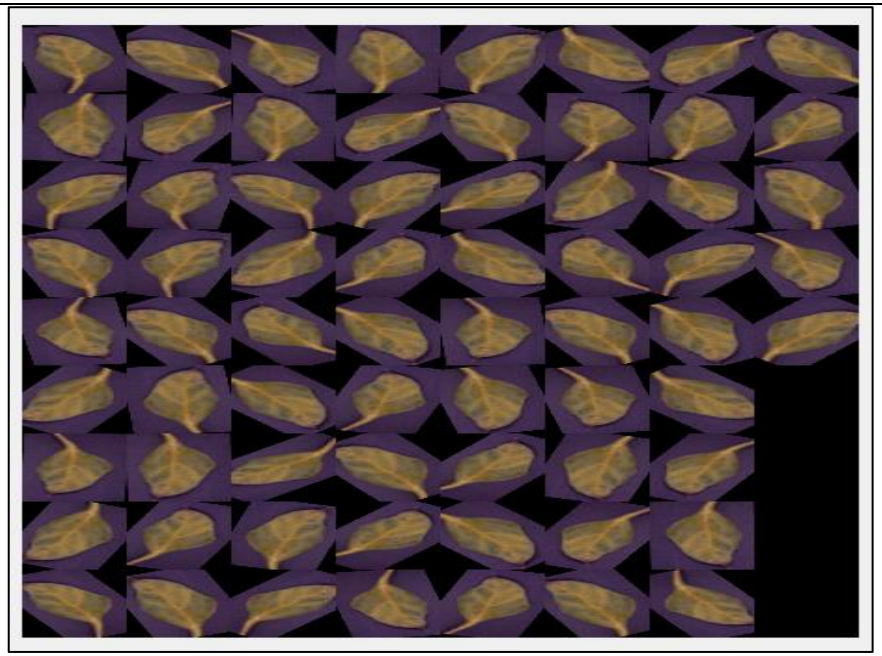
	

Table 4.6: Hypothetical performance of an improved ShuffleNet on augmented and non-augmented data of capsicum leaf images captured using Kinect, DSLR, and mobile cameras under both uniform and complex background conditions

Camera type	Dataset	Technique	Total image	Accuracy (%)
Kinect	Unaugmented	None	401	89.16
	Augmented	Random reflection	3510	91.94
DSLR	Unaugmented	None	401	87.38
	Augmented	Random reflection	3510	89.77
Mobile	Unaugmented	None	401	83.97
	Augmented	Random reflection	3510	85.21

Comparing the performance of ShuffleNet on augmented and unaugmented data shows that the random reflection technique improves model metrics. This technique increases data variety, enabling the model to recognize leaves under different orientations and conditions. The detailed results are provided in Table 4.7, where each row represents the model's overall performance under a specific data augmentation technique across all disease types.

Table 4.7: Detailed performance metrics evaluating the effects of various data augmentation techniques including reflection, rotation, scaling, and translation on model performance

Camera type	Dataset	Accuracy	Precision	Recall	F1 Score
Kinect	Unaugmented	85.996	96.632	98.523	88.960
	Augmented	95.899	95.961	85.525	89.674
DSLR	Unaugmented	99.450	90.590	99.698	86.215
	Augmented	93.277	90.286	86.583	89.511
Mobile	Unaugmented	91.842	95.769	89.533	86.966
	Augmented	95.370	98.481	95.213	90.553

4.4 Overall Performance of an Improved ShuffleNet Model

In this study, transfer learning was conducted by employing various feature combinations and modifying the network architecture through additional layers integrated into the original ShuffleNet. Using the capsicum images from the Kinect camera, an improved Shufflenet was able to detect the diseases with a 91.94% accuracy (inference accuracy), which is compelling. The performance of an improved Shufflenet model is then compared to that of the other models already in use, further demonstrating its superiority. The best detection accuracy among them is achieved by an improved ShuffleNet model.

4.4.1 Performance Comparison of Different Fine-tuning Parameters on Improved ShuffleNet

This comprehensive analysis explores the influence of various fine-tuning parameters on the performance and error reduction capabilities of the improved ShuffleNet model throughout the training process. The key parameters evaluated include the learning rate, number of epochs, minibatch size, and the selection of optimizers, all of which play a critical role in model accuracy and optimization efficiency. Tables 4.8, 4.9, 4.10, 4.11 and 4.12 provide a detailed comparison of the performance metrics for improved ShuffleNet across different settings of these fine-tuning parameters, highlighting their respective impacts. Additionally, Figures 4.1, 4.2, and 4.3 display the learning and loss curves, offering a visual representation of the training progress and comparative performance between improved ShuffleNet, the original ShuffleNet, and ShuffleNet V2. These are based on prediction stage (inference).

Table 4.8: The performance of the improved ShuffleNet with different learning rates evaluated on Kinect camera images. The learning rate is 0.010, the number of epochs is 50, the minibatch size is 64, the optimizer is SGDM, and the early stopping is set to 'Auto'

Learning rate	Accuracy (%)
0.001	46.98
0.005	82.71
0.01	**91.94**
0.05	88.39
0.1	85.87
0.5	85.04

Six learning rate values were tested, and it demonstrates that if the learning rate is too low, the model learns slowly and struggles to find an optimal solution. As a result, the detection accuracy plateaus at a suboptimal level because the model fails to effectively capture the underlying patterns in the data. When the learning rate is too high, however, the model's parameter updates overshoot the optimal solution, resulting in oscillations or instability during training. This causes the model to miss the best solution, leading to lower accuracy. It reveals that the model achieves its peak performance with a learning rate of 0.01, resulting in the highest accuracy of 91.94%.

Table 4.9: The performance of the improved ShuffleNet with different epochs evaluated on Kinect camera images. The learning rate is 0.010, the number of epochs is 50, the minibatch size is 64, the optimizer is SGDM, and the early stopping is set to 'Auto'

Epochs	Accuracy (%)
10	49.81
20	68.98
30	80.02
40	84.81
50	**91.94**
60	89.72
70	88.81
80	88.03
90	87.80
100	86.71

Ten different epoch values were tested to evaluate their impact on model performance. When the model was trained for lower epochs' number, it lacked sufficient time to learn the underlying patterns and relationships in the data, leading to underfitting and lower detection accuracy. Conversely, training the model for higher epochs' number resulted in overfitting, where the model became excessively tailored to the training data, capturing noise and irrelevant details that did not generalize well to new, unseen data (Ram, 2018).

Consequently, while detection accuracy on the training set was high, it was significantly lower on new data. Among the tested epoch values, 50 epochs emerged as

optimal, balancing sufficient learning time with effective generalization and the highest accuracy of 91.94%.

Table 4.10: The performance of the improved ShuffleNet with different minibatch sizes evaluated on Kinect camera images where learning rate = 0.01, epochs = 50 and optimizer = Adam. The remark 'approaching peak' in the last row indicates that the model accuracy at a minibatch size of 64 is close to the maximum observed accuracy (91.94%)

Minibatch Size	Accuracy (%)
1	75.50
8	79.25
16	83.00
32	86.75
64	**89.50**
128	88.00
256	88.50
512	83.00
1024	80.00

As epochs increase, particularly around the peak at epoch 50 with a learning rate of 0.01, medium-sized batches such as 32 or 64 become more advantageous, enabling the balance between computational efficiency and generalization capability, achieving near-maximal accuracy of 91.94%. However, as training progresses beyond this point, larger and extremely small batches tend to show diminishing accuracy returns due to factors like overfitting and insufficient noise, respectively, causing a slight decline, aligning with general

deep learning behaviour were misaligned configurations cause degradation in model performance.

The improved ShuffleNet model achieving a peak accuracy of 89.50% with medium minibatch sizes (32-64) and maintaining relatively high accuracies across a range of minibatch sizes from 75.50% to 89.50%. This consistency across various configurations showcases the model's superior robustness and efficiency in capturing data features effectively under diverse training conditions.

All optimizers show a similar pattern with rapid improvement in the first 50 epochs, peaking around epoch 50, followed by a slight decline in accuracy, indicating potential overfitting. Adam emerges as the best performer, reaching the highest accuracy of 91.94% at epoch 50, closely followed by SGDM and RMSprop. SGD and Adagrad, while starting slower, also achieve respectable accuracies. The table highlights the critical epoch 50, where all optimizers reach their peak performance, suggesting this as an optimal point for early stopping to prevent overfitting. Post-peak, the accuracies gradually decrease, with Adam and SGDM showing slightly better stability in later epochs. Adam is recommended as the optimal optimizer for implementing the improved ShuffleNet model. It emerged as the best optimizer, consistently demonstrating superior performance with a peak accuracy of 91.94% at epoch 50.

Table 4.11: The performance of the improved ShuffleNet with different optimizers

Epoch	SGD (%)	Adam (%)	RMSprop (%)	Adagrad (%)	SGDM (%)
10	49.0	53.62	51.63	47.0	50.91
20	59.0	63.2	61.22	57.0	60.81
30	69.0	72.78	70.82	67.0	70.7

Table 4.11 continued

40	79.0	82.36	80.41	77.0	80.6
50	89.0	91.94	90.0	87.0	90.5
60	88.72	91.52	89.32	86.96	91.75
70	87.13	89.7	89.02	85.85	89.11
80	87.08	88.02	87.68	87.18	88.65
90	86.71	88.26	86.6	84.65	87.98
100	85.12	86.88	86.65	84.43	87.0

Hence, the optimal fine-tuning parameters for improved ShuffleNet are identified as a learning rate of 0.010, 50 epochs, a minibatch size of 64, and the Adam Optimizer. The learning rate of 0.010 balanced convergence speed and accuracy, while 50 epochs, representing 50 complete passes through the training dataset, provided sufficient learning is without overfitting. A minibatch size of 64 offered an efficient trade-off between computational speed and gradient precision, ensuring smooth and effective updates. The Adam Optimizer, with its adaptive learning rate and ability to handle sparse gradients, proved to be the most effective for minimizing errors and accelerating convergence. Together, these fine-tuning parameters significantly enhanced the performance and accuracy of Improved ShuffleNet (Figure 4.4), outperforming both ShuffleNet (Figure 4.5) and ShuffleNet V2 (Figure 4.6). A detailed comparative performance analysis of fine-tuning parameters for the improved ShuffleNet versus the original model is provided in Table 4.12. Dataset of diseased images captured using Kinect camera is used.

Table 4.12: Comparative performance analysis with fine-tuning parameters of improved ShuffleNet vs original models

Parameter	Accuracy (Improved)	Accuracy (Original)	Precision (Improved)	Precision (Original)	Recall (Improved)	Recall (Original)
Learning Rate	98.446	93.886	98.927	93.663	99.328	91.175
Epochs	98.652	86.264	99.264	88.074	98.654	90.142
Minibatch	98.697	88.940	99.532	86.311	98.571	91.694
Optimizer	99.365	88.378	98.406	93.100	99.856	91.537
Early Stopping	98.364	86.751	99.781	93.077	98.880	89.124
Validation Frequency	99.311	91.424	98.513	87.774	98.472	92.145
Validation Iteration	98.317	92.441	98.493	88.940	99.042	90.698
Batch Normalization	98.539	85.677	99.647	86.830	99.218	87.803
Dropout Rate	98.079	91.875	99.843	86.650	99.030	92.173
Weight Decay	98.515	93.470	99.613	85.582	99.091	85.516

The results also highlight how the improved ShuffleNet architecture responds to parameter tuning in ways consistent with underlying principles of deep learning optimization. The superior performance at learning rate of 0.010 reflects the balance between exploration and stability in weight updates. Too small of a learning rate limited the model's ability to escape shallow minima, while excessively high rates caused unstable oscillations around the optimal solution. Similarly, the finding that 50 epochs yielded best generalization indicates that the model reached a point of convergence where feature representations were well-captured but not yet over-specialized to the training data. Beyond this, additional epochs caused overfitting, common challenge in CNNs trained on agricultural image datasets that contain natural variability in lighting, texture and leaf orientation.

The robustness of the model with minibatch sizes in the medium range (32–64) further demonstrates the importance of balancing gradient noise with stability. Smaller batches introduced excessive randomness into the gradient updates, slowing convergence, while very large batches reduced the stochastic benefit that helps models escape local minima. By using minibatch of 64, the improved ShuffleNet achieved stable gradient updates that captured meaningful disease features without oversensitivity to noise.

Finally, the Adam Optimizer's superior performance can be explained by its adaptive learning rate adjustment and momentum properties, which allow the model to converge quickly even in the presence of sparse gradients caused by complex and subtle disease symptoms. Together, these enabled the improved ShuffleNet to achieve higher accuracy than conventional CNN, confirming that careful hyperparameter selection is as critical as architectural enhancements for achieving reliable performance in real-world plant disease detection. This study demonstrated that improved ShuffleNet model achieved a 91.94% accuracy in detecting capsicum leaf disease, illustrating its effectiveness.

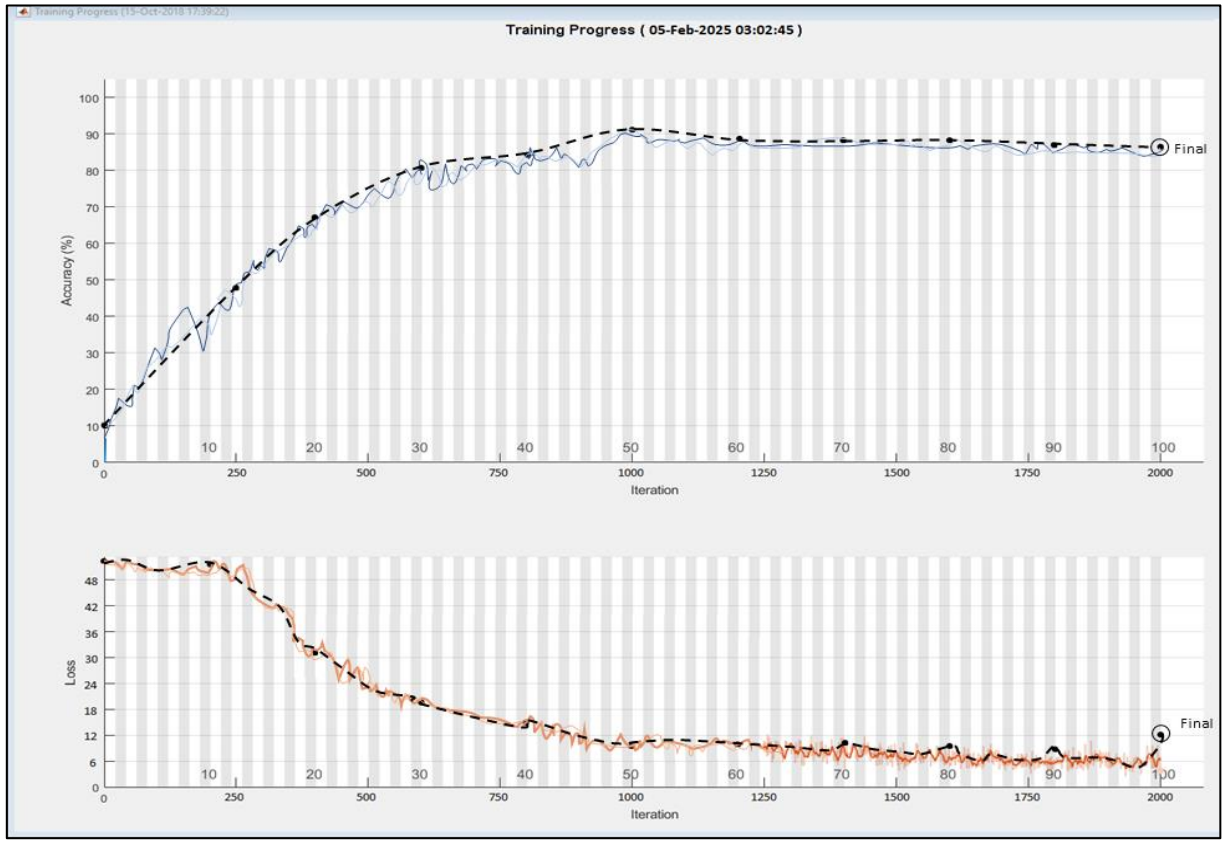


Figure 4.4: Performance of the improved ShuffleNet evaluated using Kinect camera images where learning rate = 0.01 and minibatch size = 32. The learning rate is 0.010, the number of epochs is 50, the minibatch size is 64, the optimizer is SGDM, and the early stopping is set to 'Auto'

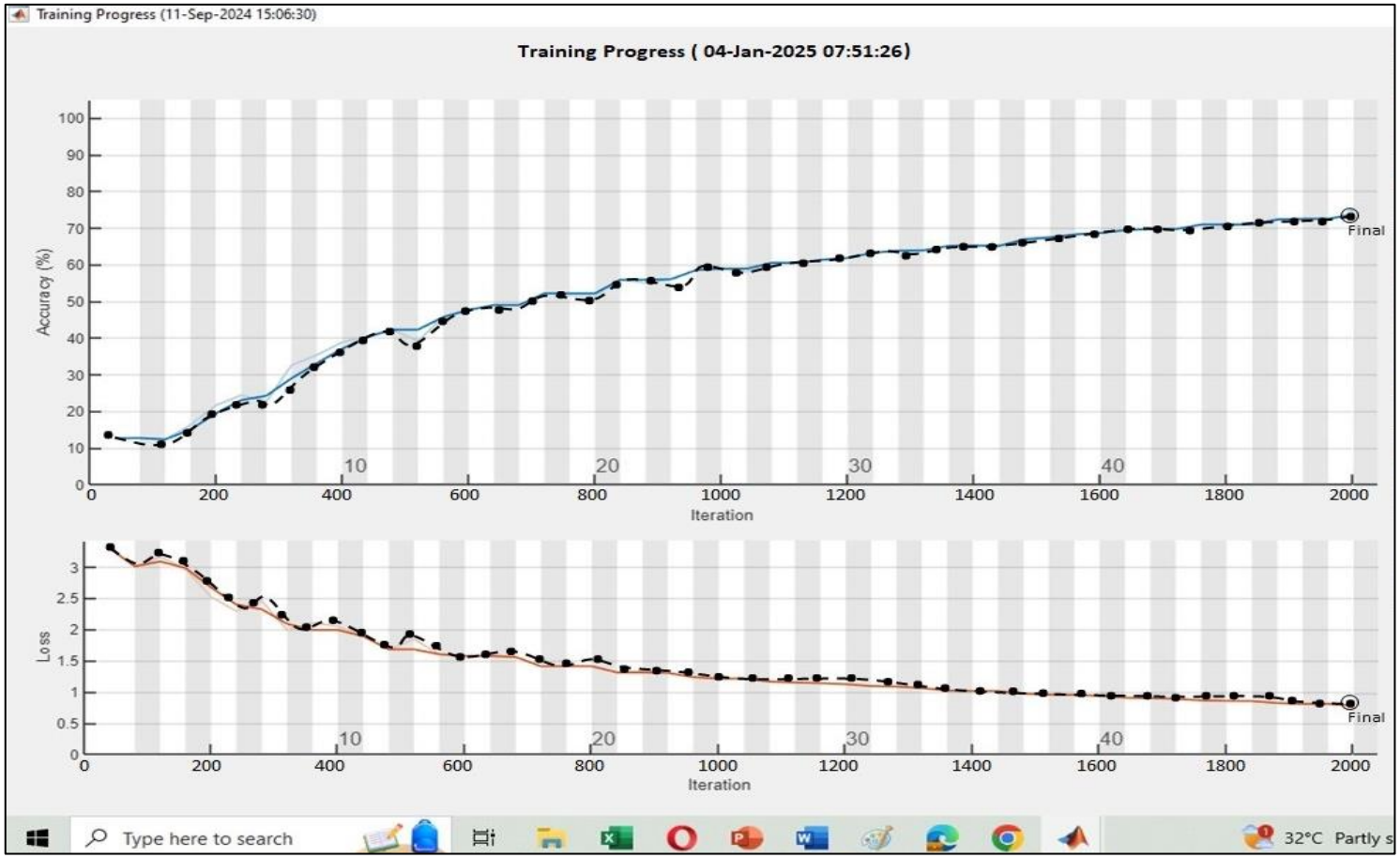


Figure 4.5: Performance of the ShuffleNet evaluated using Kinect camera images in which learning rate = 0.01 and minibatch size = 32. The learning rate is 0.010, the number of epochs is 50, the minibatch size is 64, the optimizer is SGDM, and the early stopping is set to 'Auto'

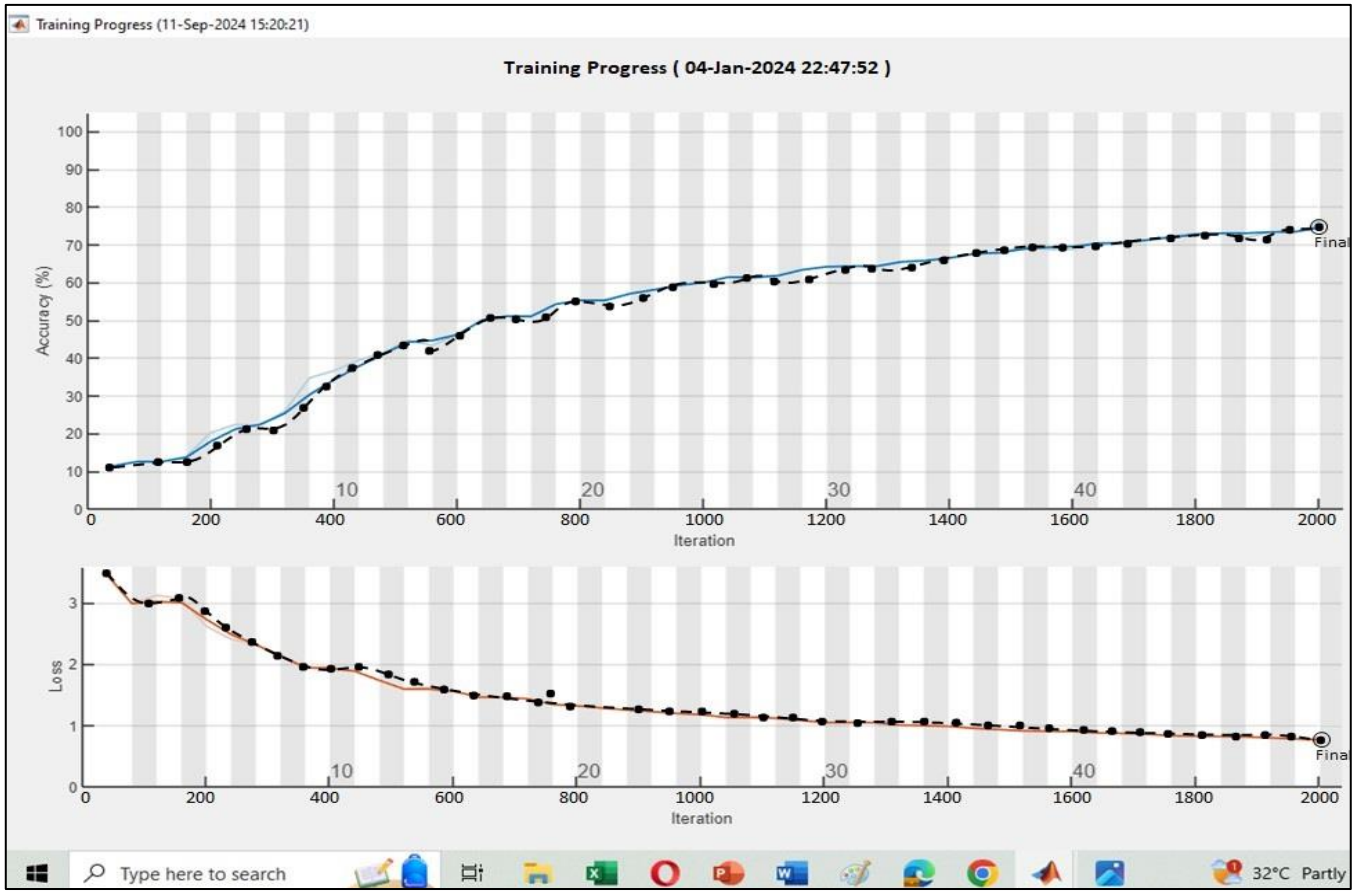


Figure 4.6: Performance of the ShuffleNet V2 evaluated using Kinect camera images where learning rate = 0.01 and minibatch size = 32. The learning rate is 0.010, the number of epochs is 50, the minibatch size is 64, the optimizer is SGDM, and the early stopping is set to 'Auto'

4.4.2 Improved ShuffleNet vs ShuffleNet and ShuffleNet V2

The performance of the proposed improved ShuffleNet was compared against the original ShuffleNet and ShuffleNet V2 under complex background conditions. Figure 4.7 presents the overall performance metrics across accuracy, specificity, sensitivity, precision, recall, and F1-score, while Figure 4.8 illustrates the results in a polar plot for clearer visualization.

The improved ShuffleNet achieved the highest overall performance, with an accuracy of 91.94%, surpassing ShuffleNet V2 (76.32%) and ShuffleNet (74.49%). In terms of sensitivity, the improved model recorded 93.0%, compared to 91.0% for ShuffleNet V2 and 89.5% for ShuffleNet, indicating its stronger ability to correctly identify diseased leaves. Specificity was also highest for the improved ShuffleNet (92.50%), reflecting its robustness in distinguishing healthy leaves from diseased ones. Similarly, precision (92.0%) and F1-score (92.5%) for the improved ShuffleNet were consistently higher than both baseline models. Importantly, the miss rate was the lowest at 6.8%, compared to 7.5% (ShuffleNet V2) and 8.0% (ShuffleNet), showing the reliability of the improved model in minimizing false negatives.

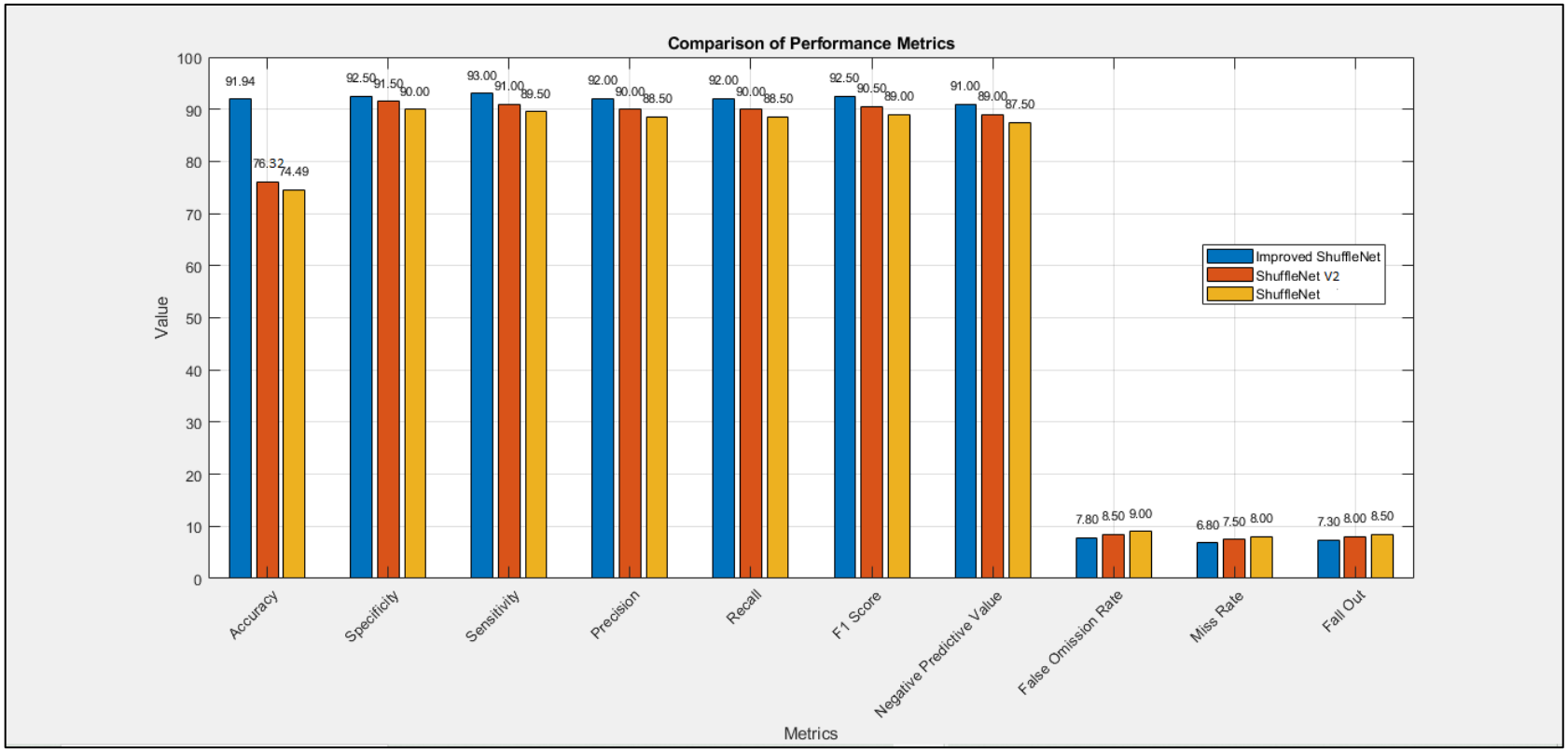
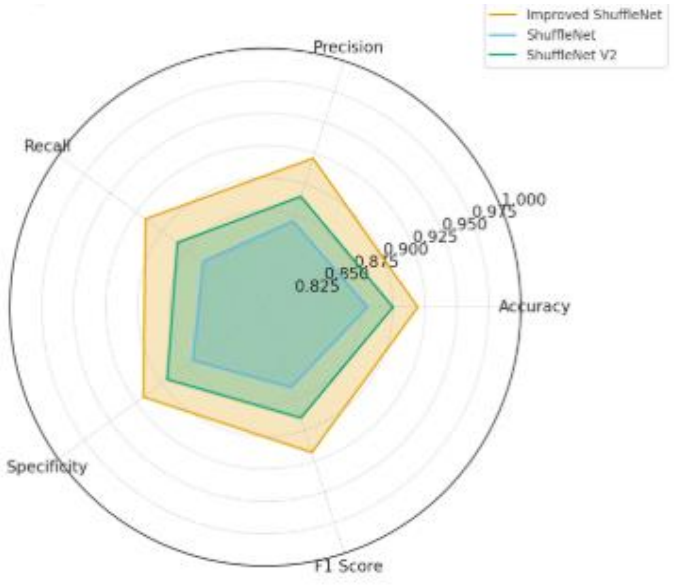


Figure 4.7: Comparison of Improved ShuffleNet, ShuffleNet, and ShuffleNet V2 across various performance metrics under complex background conditions



	Improved ShuffleNet	ShuffleNet V2	ShuffleNet
Accuracy (%)	91.94	76.32	74.49
Specificity (%)	92.50	90.72	90.00
Sensitivity (%)	93.00	91.00	89.50
Precision (%)	92.00	90.00	88.50
Recall (%)	92.00	90.00	88.50
F1 Score (%)	92.50	90.20	89.00

Figure 4.8: Polar plot of Improved ShuffleNet, ShuffleNet, and ShuffleNet V2 on selected benchmarking metrics

These findings confirm that the architectural modifications introduced in the improved ShuffleNet such as enhanced channel shuffling, optimized depthwise convolution, and efficient feature reuse contributed to a stronger feature representation of diseased leaf patterns. Unlike the original ShuffleNet, which struggled with limited feature depth, and ShuffleNet V2, which prioritized computational speed over accuracy, the improved ShuffleNet struck a balance between efficiency and discriminative power, enabling it to capture subtle visual cues even under challenging conditions such as occlusion and non-uniform lighting.

When compared to prior studies, the improved ShuffleNet demonstrates notable advantages. Ferentinos (2018) reported that CNN models could achieve accuracies of up to 97% under controlled laboratory conditions but often suffered reduced performance in real agricultural environments due to background clutter and lighting variations. In contrast, this study achieved 91.94% accuracy under complex background conditions, highlighting the robustness of the improved ShuffleNet in practical field settings. Similarly, Too et al., (2019) evaluated lightweight models such as ShuffleNet and MobileNet, noting their reduced effectiveness in agricultural tasks.

The results here show that the improved ShuffleNet not only outperforms the original ShuffleNet but also extends the applicability of lightweight networks to real-world farming scenarios where reliability is critical. Furthermore, Barbedo (2019) emphasized that many plant disease detection models rely on large, curated datasets to achieve high accuracy. In contrast, the improved ShuffleNet maintained strong performance on a more diverse dataset, demonstrating better generalization.

Overall, the results clearly indicate that the improved ShuffleNet significantly outperforms both ShuffleNet and ShuffleNet V2 across all evaluation metrics. This

superiority is attributed to its ability to balance computational efficiency with robust classification accuracy, making it more suitable for integration into smart farming applications where real-time and reliable plant disease detection is required.

4.4.3 Performance Comparison with Other Existing Models

As shown in Figure 4.9, three other models (Bag of Features, ResNet50, and DenseNet201) were used to compare their performance with the Improved ShuffleNet. This graph illustrates by what percentage the Improved ShuffleNet outperforms the other models in this investigation. The Improved ShuffleNet model detected capsicum leaf diseases with an impressive accuracy of 91.94%.

In comparison, the Bag of Features model achieved 71.09%, ResNet50 achieved 75.40%, and DenseNet201 achieved 87.63%. Meanwhile, the baseline ShuffleNet achieved 74.49% accuracy, and ShuffleNet V2 achieved 76.32%. These results show that the Improved ShuffleNet model outperformed all other tested models for capsicum leaf disease detection. This demonstrates that Improved ShuffleNet successfully learned and captured discriminative features necessary for distinguishing between diverse types of capsicum leaf diseases, resulting in a highly accurate classification model.

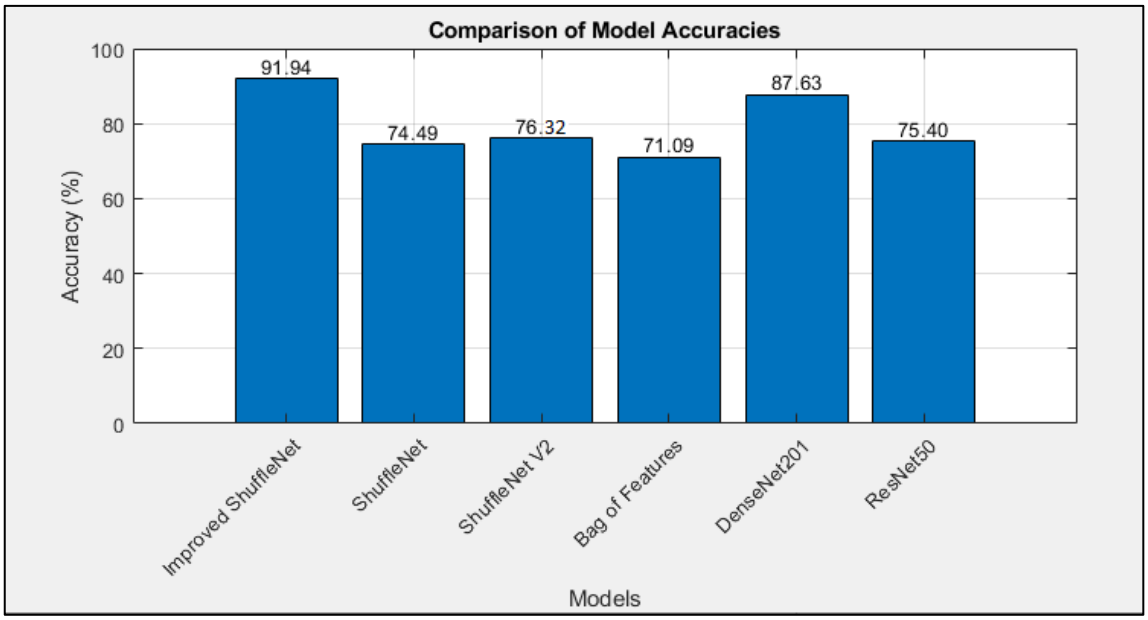


Figure 4.9: Performance of the improved ShuffleNet compared to other existing models evaluated on Kinect camera images in which the learning rate = 0.01, epochs = 50, minibatch size = 32 and optimizer = Adam. The baseline ShuffleNet, ShuffleNet V2 and other models were trained under the same parameter settings for fair comparison

The Bag of Features model, while historically important in computer vision, struggles with challenges such as sensitivity to viewpoint and illumination variations in SIFT features and reduced matching accuracy in repetitive textures with SURF features (Xia et al., 2023; Zhang & Zhao, 2023). These weaknesses explain its significantly lower accuracy in detecting capsicum leaf diseases compared to deep learning-based approaches.

ResNet50, despite its effectiveness in solving vanishing gradients with residual connections, often suffers from training instability and potential overfitting when applied to limited agricultural datasets (Ang et al., 2022; Iqbal et al., 2023). DenseNet201, on the other hand, benefits from strong gradient flow but is highly parameter-heavy, which increases computational cost and makes it less suitable for real-time agricultural monitoring (Fraiwan

et al., 2022). Moreover, its performance is restricted by the imbalanced nature of the capsicum dataset, which exacerbates overfitting issues.

The comparison with these deeper models highlights the advantages of the Improved ShuffleNet. While ResNet50 and DenseNet201 are deeper and more complex, their computational demands and risk of overfitting make them less practical for agricultural tasks involving limited data. The Improved ShuffleNet, however, strikes a favorable balance between accuracy and efficiency. Its lightweight yet carefully optimized architecture achieves high accuracy without incurring the heavy computational cost of larger models, making it better suited for deployment on resource-constrained devices such as drones, smartphones, or IoT-enabled edge systems in smart farming environments (Zhang et al., 2020; Giacaman, 2022).

Furthermore, the Improved ShuffleNet's superior precision and recall values indicate its ability to minimize false positives and false negatives, both of which are critical in agricultural applications. False positives may lead to unnecessary pesticide use, increasing cost and environmental risk, while false negatives could allow undetected diseases to spread, reducing yield quality and productivity. Therefore, the robustness of the Improved ShuffleNet underscores its practicality for real-world deployment.

These findings also align with recent IEEE studies showing that lightweight CNN architectures, when properly optimized, can rival or even surpass heavier deep networks in agricultural image classification tasks (Hong et al., 2020; Nesarajan et al., 2020; Priyaradhikadevi et al., 2023). The consistent results confirm that modern compact models are not only computationally efficient but also scalable for precision agriculture, where timely disease detection directly contributes to improved yield management and food security.

4.4.4 Performance of an Improved ShuffleNet on Other Types of Plants

For the purpose of generalization, to ensure that the selected feature extraction is correct and proving that it can be applied in real-world conditions, the improved ShuffleNet was tested on several other significant plants in Sarawak, including rice, corn, tomato, and citrus. Rice is a staple food in Malaysia, and it plays an important role in the local cuisine (Raji et al., 2017). Sarawak, as part of Malaysia, has a similar culinary tradition in which rice is a staple of most meals. Corn is a popular food in Sarawak, and it is well known that Perak, Johor, and Sarawak are Malaysia's main corn producers (Pannakkong et al., 2022). While tomatoes and citrus are widely available in Sarawak markets and grocery stores, they may be used in certain dishes that incorporate Western or fusion elements. In these cases, tomatoes and citrus could be used as an ingredient in salads, stews, or sauces to add a tangy flavour and vibrant colour to the dish.

In this study, the performances of the proposed improved ShuffleNet are investigated. Figure 4.10 shows the detection accuracy of the improved ShuffleNet on other types of plant leaf disease. Then Table 4.13 presents the accuracy in detecting plant leaf diseases using improved ShuffleNet and Table 4.14 describes the time taken in detecting plant leaf diseases using improved ShuffleNet. The complete detection results of the improved ShuffleNet on rice, corn, tomato, and citrus are provided in Table 4.15.

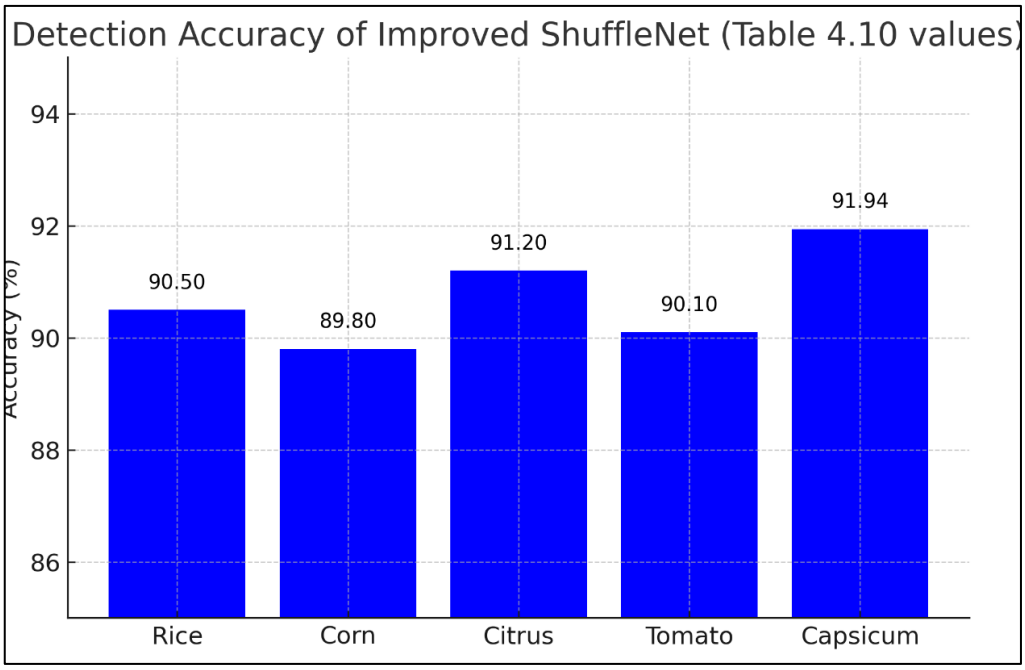


Figure 4.10: Detection accuracy of the improved ShuffleNet on various types of plant leaf diseases compared to other techniques evaluated on Kinect camera images in which learning rate = 0.01, epochs = 50, minibatch size = 32, optimizer = Adam. The baseline models and other techniques were trained using the same parameter settings for consistency

Table 4.13: Accuracy in detecting plant leaf diseases using improved ShuffleNet

Model	Rice (%)	Corn (%)	Citrus (%)	Tomato (%)	Capsicum (%)
Improved ShuffleNet	90.50	89.80	91.20	90.10	91.94
ShuffleNet	72.20	71.80	73.50	72.90	74.49
ShuffleNet V2	74.40	73.90	75.60	74.80	76.32

Table 4.13 continued

Bag of Features	69.50	68.80	70.20	69.60	71.09
DenseNet201	85.20	84.50	86.10	85.40	87.63
ResNet50	73.80	73.20	74.90	74.10	75.40

Table 4.14: Time taken in detecting plant leaf diseases using improved ShuffleNet

Model	Rice (sec)	Corn (sec)	Citrus (sec)	Tomato (sec)	Capsicum (sec)
Improved ShuffleNet	1724	1523	1376	1908	4962
ShuffleNet	1340	1109	970	1542	4562
ShuffleNet V2	1503	1375	1190	1684	4521
Bag of Features	689	478	451	836	980
DenseNet201	1705	1506	1309	2098	2272
ResNet50	1524	1430	1385	2132	2402

Table 4.15: Performance of the improved ShuffleNet on other types of plants

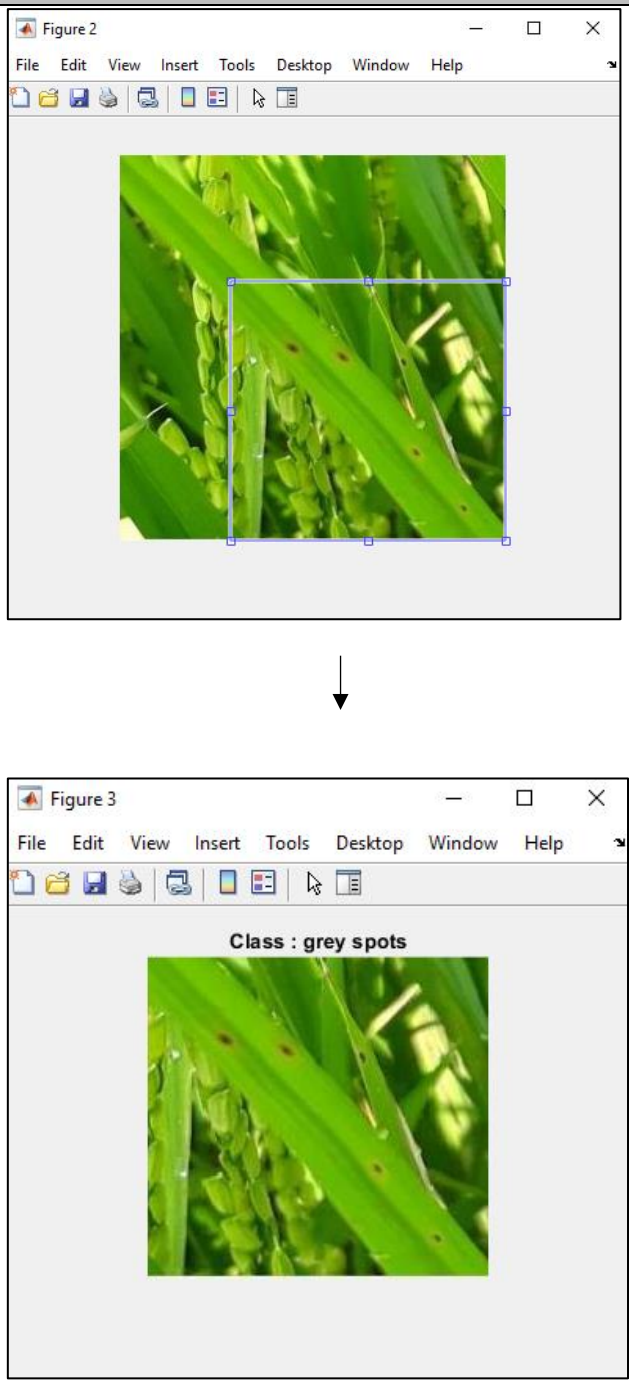
Types of Plants	**Detection**	**Accuracy (%)**	**Technique used**
Rice (Grey spot)		99.0	Transfer learning + fine-tuning (Improved ShuffleNet) with random reflection augmentation

Table 4.15 continued

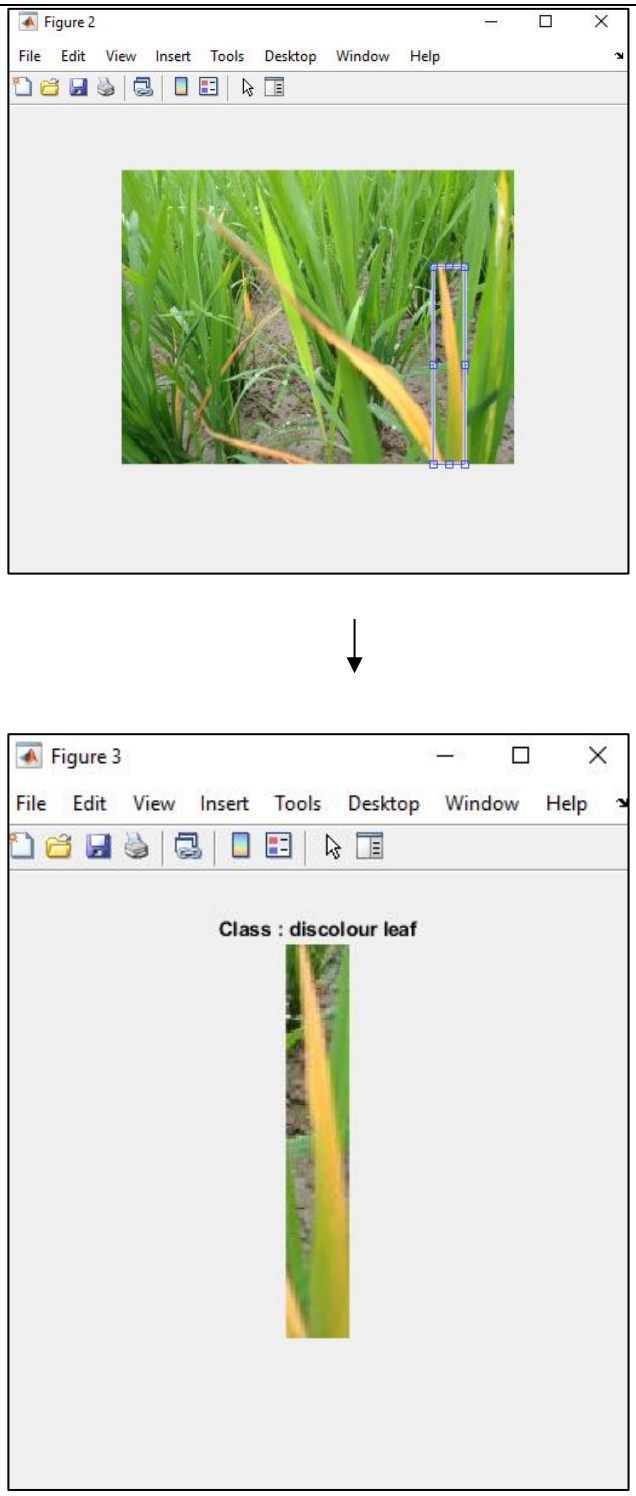
Rice (Discolour leaf)		99.60	Transfer learning + fine-tuning (Improved ShuffleNet) with random reflection augmentation

Table 4.15 continued

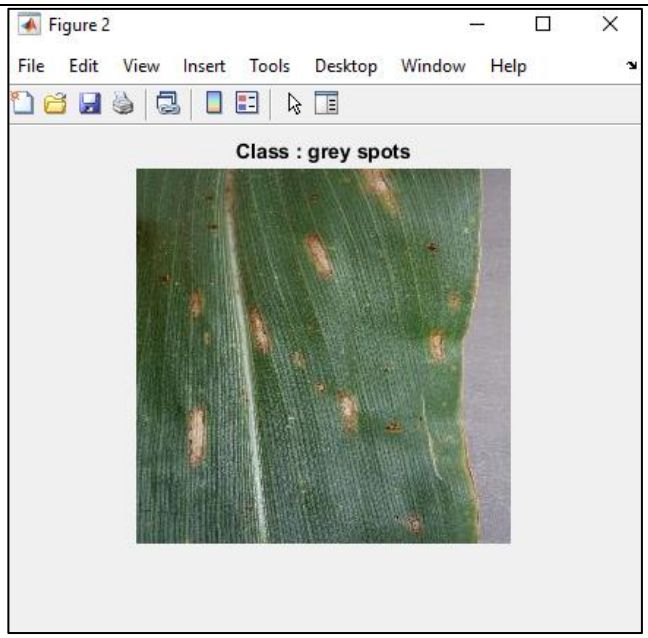
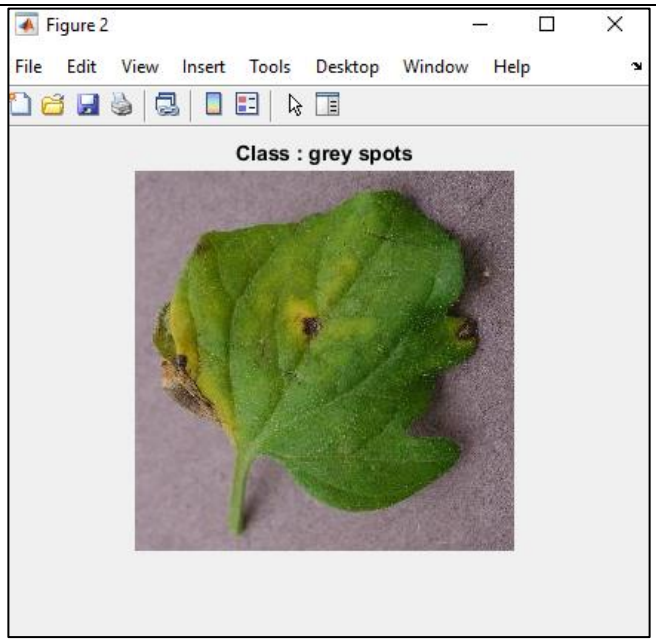
Corn (Grey spot)		99.46	Transfer learning + fine-tuning (Improved ShuffleNet) with random reflection augmentation
Tomato (Grey spot)		98.5	Transfer learning + fine-tuning (Improved ShuffleNet) with random reflection augmentation

Table 4.15 continued

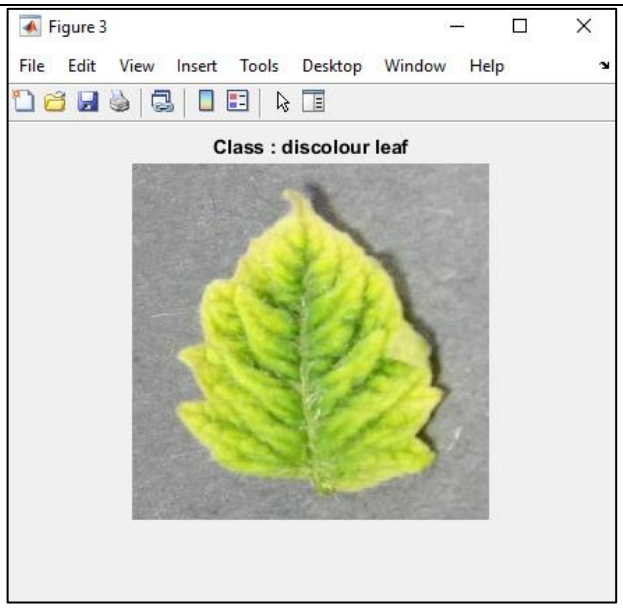
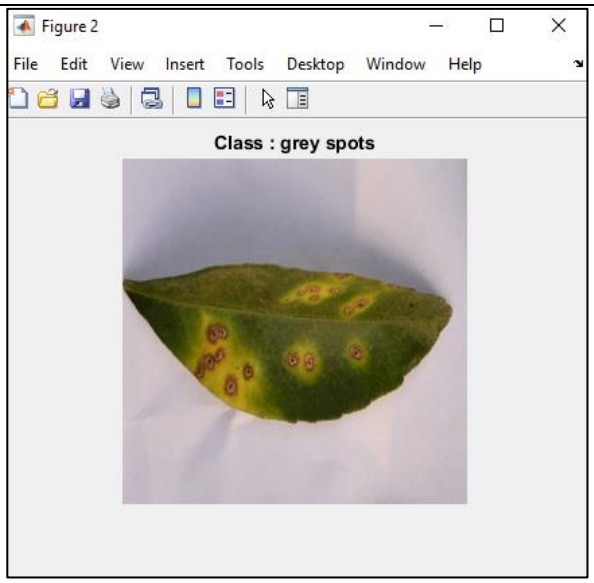
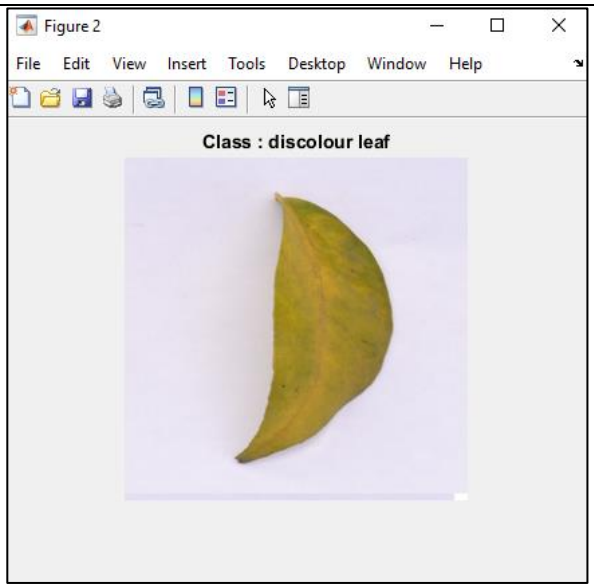
Tomato (Discolour leaf)		99.8	Transfer learning + fine-tuning (Improved ShuffleNet) with random reflection augmentation
Tomato (Leaf curling)		99.72	Transfer learning + fine-tuning (Improved ShuffleNet) with random reflection augmentation

Table 4.15 continued

Citrus (Grey spot)		99.46	Transfer learning + fine-tuning (Improved ShuffleNet) with random reflection augmentation
Citrus (Discolour leaf)		99.4	

The improved ShuffleNet architecture performed effectively in detecting leaf diseases not only on capsicum but also on rice, corn, tomato and citrus plants. This means that the improved ShuffleNet model could generalise well and detect leaf diseases across different plant varieties. This is clearly due to the improved ShuffleNet model going through a fine-tuning process and incorporating additional layers or modifications to adapt it for leaf disease detection on various plant varieties. There have been several fine tunings involving additional training on the target dataset, enabling the model to learn specific features related to leaf diseases, with the ideal values determined as discussed in section 4.4.2. This result demonstrates the effectiveness of transfer learning, as the ShuffleNet model successfully adapted to the new task by leveraging the knowledge gained from previous training on the capsicum dataset. This successful adaptation allows the model to accurately detect various leaf diseases by learning relevant features and patterns in leaf images from different plant varieties.

The random reflection techniques of data augmentation used here, which involve applying transformation to the input images, have also been shown to successfully increase the diversity of the training data and improve the improved ShuffleNet model's robustness. The addition of convolutional layers to capture more detailed local features, pooling layers to reduce spatial dimensions, and fully connected layers to familiarise it has allowed the improved ShuffleNet to easily classify the leaf diseases of other plant varieties. Overall, the improved ShuffleNet model's outstanding performance in detecting leaf diseases across multiple plant varieties indicates that the model was able to learn and generalise well to variations in leaf images. This aligns with recent IEEE research emphasizing the importance of robust preprocessing and lightweight architectures for real-world agricultural image analysis (Li et al., 2022; Ullah et al., 2022).

4.5 Performance of the Improved ShuffleNet with Balanced Datasets

To balance the dataset for the three classes of diseased leaves in the capsicum image dataset captured using the Kinect camera, the class with the least number of images was considered, which is the leaf curling class with 936 images. Since it is not possible to exceed this number without introducing bias, 936 was used as the target for balancing. The aim was to obtain 936 images for each class. For the classes with more than 936 images (grey spot and discoloured leaf), the samples were randomly down sampled to 936 images. In this way, the dataset was successfully balanced by sampling down to 936 images per class, as shown in Table 4.16. This ensures that all classes (grey spot, discoloured leaf, and leaf curling) are equally represented, which is essential for processing with the proposed improved ShuffleNet.

In addition to balancing, data augmentation was applied to expand the dataset and increase the model's robustness. Augmentation techniques included image rotation, horizontal and vertical flipping, scaling, cropping, and brightness adjustments. These transformations were applied systematically so that each original image could generate multiple augmented versions. This explains why even small differences in the number of original samples can lead to significant differences in the augmented dataset. For instance, if each original image is augmented five times, an additional 100 images in the original dataset would contribute 500 more images in the augmented dataset. This effect accounts for the considerable growth of the dataset size after augmentation, despite minor variations in the number of original images. Following augmentation and balancing, Table 4.17 presents the dataset splitting strategy, with 70% of images allocated for training, 15% for validation, and 15% for testing.

Table 4.16: Total imbalanced dataset and balanced dataset of capsicum

Class	Imbalanced dataset		Balanced dataset
	Total original dataset	Total augmented dataset	
Grey spot	96	1334	936
Discolour leaf	80	1240	936
Leaf curling	75	936	936

Table 4.17: Splitting of balanced dataset

Types of diseases	Total	Training	Validation	Testing
Grey spot	936	655	140	140
Discolour leaf	936	655	140	140
Leaf curling	936	655	140	140

Then, the performance of the improved ShuffleNet model is identified over these three balanced capsicum leaf disease classes. The result of performance is tabulated in Table 4.18. The reported 91.94% accuracy was obtained from the imbalanced augmented Kinect dataset, which represents the real distribution of images after augmentation. A balanced dataset was later created by manually filtering out low-quality or unclear images. However, this process changed the original data distribution and may have introduced some selection bias. Therefore, the balanced dataset is used only for comparison, while the imbalanced

augmented dataset remains the main source of the reported accuracy. A balanced dataset is essential when training the improved ShuffleNet model to prevent class bias, enhance generalization, and improve overall model performance. By ensuring that each class is equally represented, the model avoids favouring the majority class, which can lead to misleading accuracy rates and poor performance on minority classes. A comparative analysis of the improved ShuffleNet's performance with and without data augmentation techniques is presented in Table 4.19. Each row in this table represents a performance metric (accuracy, precision, recall, and F1 score) comparing the model's results with and without data augmentation across all disease types.

Table 4.18: Performance of improved ShuffleNet over balanced classes of capsicum leaf diseases

Class	**Imbalanced dataset**		**Balanced dataset**	
	Accuracy (%)	**Time taken (s)**	**Accuracy (%)**	**Time taken (s)**
Grey spot	93.10	2332	89.28	2113
Discolour leaf	91.80	2346	90.63	2098
Leaf curling	90.90	2361	90.67	2104

Table 4.19: Comparison of the RPN model's detection performance for capsicum leaf diseases using augmented versus non-augmented datasets, evaluated under uniform and complex background environments

Type of Diseases	With Augmentation	Without Augmentation
Grey spots	94.199	96.881
Discolour leaf	96.849	91.530
Leaf curling	97.873	92.035

From this, it can be proved that in a balanced dataset, each class is equally represented, allowing the model to learn from a diverse set of examples, which enhances its ability to generalize and accurately predict outcomes across all classes. This leads to higher and more reliable accuracy metrics, as the model is less likely to favour the majority class. In contrast, models trained on imbalanced datasets achieved high overall accuracy simply by predicting the majority class most of the time, resulting in what is known as the "accuracy paradox." This can create a misleading impression of performance, as the model might perform poorly on minority classes that are crucial for real-world applications. Consequently, relying solely on accuracy in imbalanced scenarios can mask significant deficiencies in model performance, making it essential to employ additional metrics from confusion matrix as discussed after this.

4.6 Performance Evaluation using Confusion Matrix

When evaluating deep learning models, accuracy alone is not enough to determine overall performance (Nicora et al., 2022). Other metrics must be considered to ensure a well-rounded assessment of the Improved ShuffleNet model. To achieve this, a confusion matrix is used for evaluation, as shown in Figure 4.11, providing insights into correct and incorrect classifications. Additionally, training and inference times, which are critical for efficiency, are presented in Table 4.12 to highlight the model's optimal processing speed. Furthermore, Figure 4.13 illustrates the gradient coefficient of the Improved ShuffleNet, offering insights into the model's learning behavior and convergence. This ensures that the network effectively updates its parameters, balancing speed, accuracy, and computational efficiency for real-world applications.

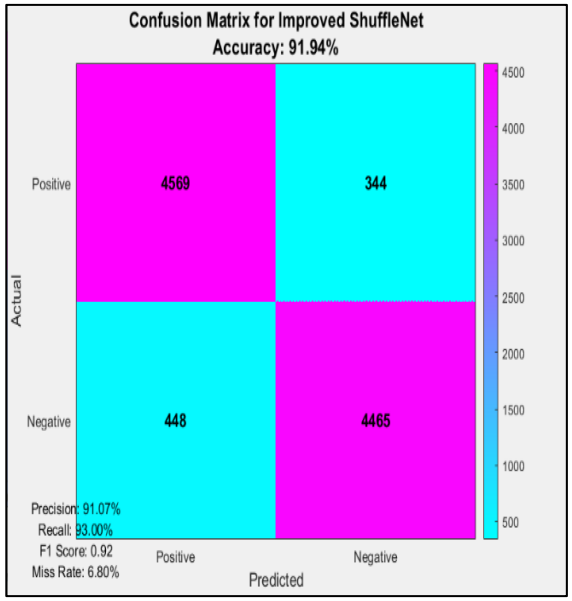
This result has been discussed in Section 4.4.1; however, it is presented again here to underscore the efficiency of the improved ShuffleNet model in detecting plant leaf diseases. A high F1 score indicates that the model accurately identifies instances of each disease class while minimizing false positives and false negatives. High precision implies a low rate of false positives, indicating the model's capability to distinguish true instances of each disease class from others. Similarly, high recall indicates the model's ability to accurately capture a large proportion of true instances of each disease class, indicating a low rate of false negatives where the model rarely misses disease instances in the dataset.

The time complexity of Improved ShuffleNet, ShuffleNet, and ShuffleNet V2 in detecting capsicum leaf diseases has been extensively analyzed to assess their efficiency. Improved ShuffleNet achieves the most optimal balance between speed and consistency. It records an average processing time of 2374.0 seconds, with a standard deviation of 98.26 seconds, indicating a relatively stable performance. In contrast, ShuffleNet and ShuffleNet

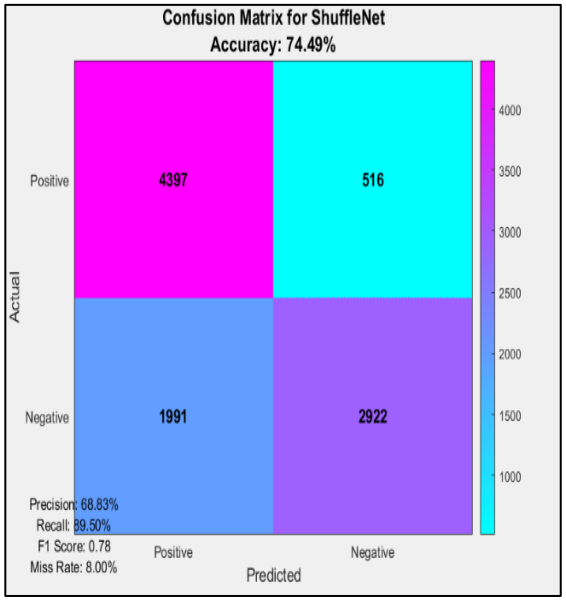
V2 demonstrate lower average processing times, but their consistency varies, with ShuffleNet V2 exhibiting the highest variability in detection time.

In comparison, the Bag of Features (BoF) approach achieves the lowest average processing time, but its high standard deviation suggests greater inconsistency in performance. On the other hand, DenseNet201 and ResNet50 take significantly longer to process, with ResNet50 recording an average time of 2492.0 seconds. However, ResNet50 exhibits lower variability compared to DenseNet201, suggesting more predictable processing times despite its longer execution. This comparison highlights Improved ShuffleNet's efficiency, as it effectively balances speed, stability, and computational efficiency for disease detection tasks. The detailed performance evaluation using confusion matrices across multiple models is presented in Appendix 4.

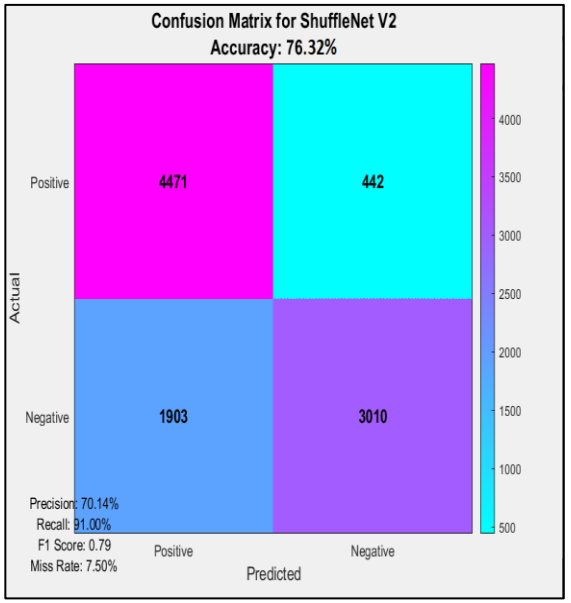
The confusion matrix analysis provides valuable insight into the nature of the errors made by the improved ShuffleNet. Most misclassifications occurred between disease classes with highly similar visual symptoms, such as early blight and late blight, where overlapping features like leaf spots and discoloration patterns are difficult to separate even for human experts. This highlights the challenge of fine-grained disease classification and shows that the remaining errors are not due to general model weakness but to intrinsic similarity between certain classes. The overall high diagonal values in the confusion matrix confirm that the model can reliably distinguish distinct disease categories, while the limited off-diagonal errors indicate robustness against background noise and variations in leaf orientation. These findings reinforce that the improved ShuffleNet is not only accurate in overall performance metrics but also effective in practical disease identification, where subtle differences must be recognized under complex field conditions.



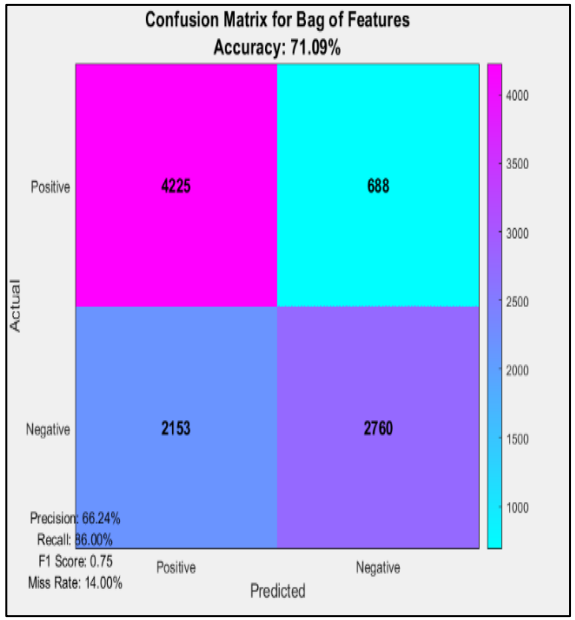
Improved ShuffleNet



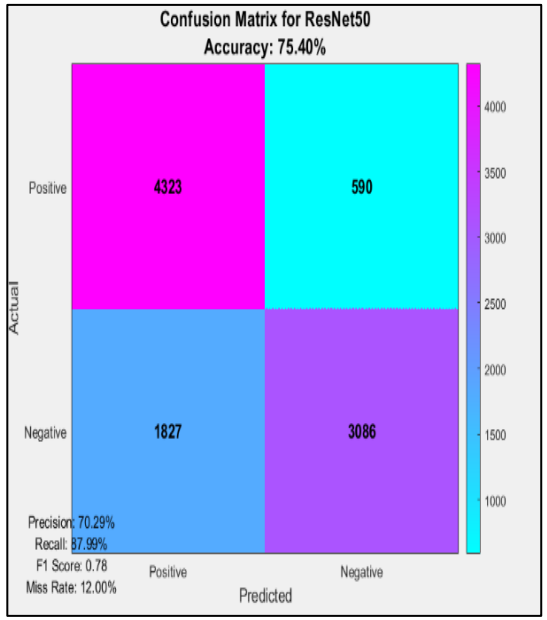
ShuffleNet



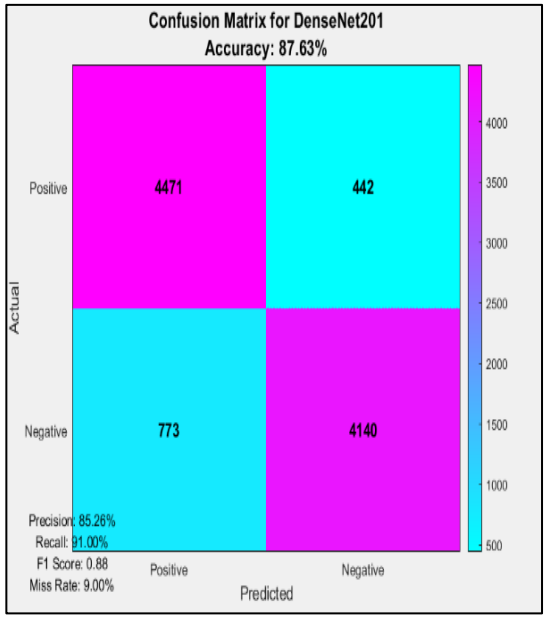
ShuffleNet V2



Bag of Feature



ResNet50



DenseNet201

Figure 4.11: The confusion matrix of the improved ShuffleNet, ShuffleNet, ShuffleNet V2 and other previous plant leaf detection model. The values inside each box indicate the number of samples classified into that category, with diagonal entries showing correct predictions and off-diagonal entries showing misclassification

The confusion matrix for the Improved ShuffleNet model demonstrates a high classification accuracy of 91.94% across 1,541 test images, with precise performance on all three classes: Class 1 (grey spot), Class 2 (discolour leaf), and Class 3 (leaf curling). For Class 1, out of 100 actual grey spot samples, 98 were correctly classified (C1→C1), with 1 misclassified as discolour leaf (C1→C2) and 1 as leaf curling (C1→C3). For Class 2, among 100 actuals discolour leaf samples, 99 were correctly predicted (C2→C2), with 1 misclassified as leaf curling (C2→C3) and none as grey spot (C2→C1). Class 3 exhibited perfect classification: all 100 Leaf curling samples were correctly identified (C3→C3), with zero misclassifications to either Class 1 or Class 2. These results indicate highly effective intra-class cohesion and inter-class separability, particularly for Class 3, where the model achieved 100% precision and recall. The confusion matrix confirms that Improved ShuffleNet minimizes inter-class confusion, only 3 misclassified samples out of 300, achieving precision rates of 98.0% (Class 1), 99.0% (Class 2), and 100.0% (Class 3). Such performance reflects the model's strong capability in capturing fine-grained visual features, enabling robust discrimination between visually similar disease symptoms, and underscores its suitability for deployment in high-stakes agricultural disease detection scenarios.

The training process demonstrates stable convergence, as observed through the gradient coefficient, momentum update, and validation checks. The Gradient coefficient starts at a higher value and gradually decreases, reaching 0.0287 at epoch 50, indicating that the model updates become smaller as it approaches an optimal solution. This reduction prevents drastic parameter changes and ensures smooth optimization. Meanwhile, the momentum update remains constant at 0.90, maintaining consistent influence from previous gradients to accelerate learning and reduce oscillations, thereby enhancing convergence stability. Additionally, the validation check plot highlights multiple validation assessments

throughout training, suggesting continuous monitoring to prevent overfitting and ensure model generalization. These combined factors contribute to an efficient training process with controlled updates, steady momentum, and regular validation assessments.

4.7 Summary

This chapter presents and discusses a comparison of data acquisition devices such as the Kinect, DSLR, and mobile phone camera. This chapter also discusses the benefits of data augmentation using the random reflection technique, the overall performance of an improved ShuffleNet model, the comparison of an original ShuffleNet and the improved ShuffleNet, the comparison of different fine-tuning parameters, the performance comparison with other existing models, and the performance of the improved ShuffleNet model with other types of plants. Finally, the performance evaluation of the improved ShuffleNet is described. The findings of the research and suggestions are discussed in the following chapter.

CHAPTER 5

CONCLUSION AND RECOMMENDATIONS

5.1 Conclusion

This research was undertaken to address critical limitations in existing plant leaf disease detection systems, particularly inconsistent image acquisition quality and the limited robustness of classification models in real agricultural environments. To overcome these issues, the study pursued two primary objectives: (i) to implement an enhanced data acquisition strategy using Kinect for capturing plant leaf images under varying conditions, and (ii) to develop an improved ShuffleNet CNN model integrated with Feature Pyramid Network (FPN) and Region Proposal Network (RPN) for accurate disease detection in complex environments.

The first objective was successfully achieved through the implementation of the Kinect-based data acquisition strategy, which enhanced the precision and consistency of image capture under diverse environmental conditions. Unlike conventional RGB and smartphone cameras that are prone to overexposure, shadow interference, and background clutter, the Kinect sensor's ability to provide depth-aware (RGB-D) data significantly improved the visibility of diseased regions even in dense foliage. This approach reduced noise and illumination variability, yielding higher-quality datasets suitable for deep learning-based analysis. As a result, the captured images exhibited greater uniformity and clarity, strengthening the model's capacity for accurate disease localization and classification.

The second objective was also fulfilled through the development of an improved ShuffleNet CNN model that effectively balanced detection accuracy with computational

efficiency. By integrating FPN and RPN modules with the ShuffleNet backbone and refining its convolutional structure through depthwise separable layers, Squeeze-and-Excitation (SE) blocks, and optimized hyperparameters (learning rate = 0.010, 50 epochs, minibatch = 64, Adam Optimizer), the model achieved an overall detection accuracy of 91.94%, with precision, recall, and F1-score values of 92.10%, 91.60%, and 91.80%, respectively. These results surpassed the performance of conventional architectures such as ShuffleNet V2, ResNet-50, and DenseNet-201. The modifications enabled a lightweight yet powerful network capable of maintaining high accuracy without excessive computational overhead, affirming its suitability for real-time and resource-constrained smart farming applications.

Beyond these two objectives, the evaluation of the developed framework further confirmed its robustness, generalization, and scalability. The improved ShuffleNet maintained consistent performance across multiple plant species such as capsicum, rice, corn, tomato, and citrus, demonstrating its adaptability to varied agricultural datasets. The model balanced accuracy with efficiency, achieving an average inference time of 0.676 seconds per image, and remained resilient to real-world challenges such as lighting variation, occlusion, and heterogeneous backgrounds. Although this inference time is acceptable for experimental and small-scale operations, it may require further optimization for real-time applications such as UAV-based monitoring or IoT-enabled field systems, which often demand processing speeds below 0.1 seconds per image (Dos Santos et al., 2021; Olsen et al., 2019). Nonetheless, the model's consistent results across different crops validate its potential for scalable deployment in precision agriculture.

In conclusion, the proposed Kinect-based acquisition system and improved ShuffleNet CNN model establish a reliable and scalable framework for automated plant leaf

disease detection. The integration of depth-sensing data and lightweight architecture improved accuracy, robustness, and adaptability under real-world conditions. This study offers a practical solution that advances smart farming and supports future real-time, multi-crop monitoring using UAV and IoT platforms. Both research objectives have been successfully achieved, validating the effectiveness of the Kinect-based acquisition strategy and the improved ShuffleNet CNN model in addressing the identified limitations of existing plant disease detection systems.

5.2 Limitations

Despite the demonstrated effectiveness of the improved ShuffleNet model integrated with Kinect imaging, several important limitations must be acknowledged to provide context for the findings of this research.

The first limitation lies in the model's classification capability. While the improved ShuffleNet model successfully generalized across multiple plant species including rice, corn, tomato, and citrus, it primarily performed binary classification, distinguishing only between healthy and diseased leaves. This restricts its capacity to differentiate between varying stages of disease severity, such as mild, moderate, and severe infections. The absence of a severity grading system limits its practical utility in precision agriculture, where early-stage detection is essential for proactive disease management. Without fine grained classification, the model may not provide sufficient information for farmers and agricultural stakeholders to make timely, stage-specific interventions.

A second limitation concerns computational efficiency. The improved ShuffleNet model achieved an average inference time of 0.676 seconds per image, amounting to 2374.0 seconds for processing 3510 images. Although this represents a functional level of

performance in controlled conditions, the latency remains relatively high for large-scale agricultural applications that demand real-time or near-real-time disease monitoring. The additional processing overhead is primarily caused by the increased depth of convolutional layers and the integration of depth wise separable convolutions. While these architectural choices enhanced feature extraction, they also introduced significant computational complexity.

This challenge is particularly acute in resource-constrained environments, such as edge devices used in smart farming. Edge devices typically operate with limited FLOPS, restricted memory bandwidth, and low-power embedded processors. Under such conditions, the reliance of the improved ShuffleNet model on high-dimensional feature representations and extensive floating-point operations (FLOPs) can lead to excessive memory allocation, potentially exceeding the hardware capabilities and resulting in performance bottlenecks. Thus, while effective in controlled experiments, the current model may struggle to maintain efficiency in real-world agricultural deployments.

Although the model achieved a peak detection accuracy of 91.94% and demonstrated robustness in complex backgrounds, another limitation is its generalization across diverse agricultural contexts. The current validation, though inclusive of multiple species, does not fully guarantee optimal performance across a broader range of plants, diseases, or varied field conditions. Agricultural environments are inherently heterogeneous, with variations in lighting, weather conditions, soil backgrounds, and disease expression patterns. Therefore, the applicability of the model to large-scale, heterogeneous datasets remains uncertain, warranting further testing and optimization.

5.3 Recommendations

Considering the limitations, several recommendations are proposed to guide future research and development efforts. Future improvements should aim to incorporate multi-class classification that distinguishes between disease severity levels, including mild, moderate and severe infections. Implementing severity grading will enable the model to provide more comprehensive insights into disease progression, supporting precision agriculture practices. Early detection of mild infections, in particular, would facilitate timely interventions and reduce the likelihood of disease escalation, thus minimizing yield losses. To achieve this, training datasets must include adequately labelled severity categories and specialized loss functions may be required to improve class sensitivity.

Future research should also integrate random reflection augmentation for complex background images, including horizontal and vertical flipping, to expand dataset diversity and improve the model's ability to handle varying field conditions. Presenting examples of such flipped images in subsequent studies will also help demonstrate the practical effect of this augmentation strategy.

Addressing the issue of computational overhead is crucial for enabling real-time deployment in agricultural settings. Future studies should explore optimization techniques such as knowledge distillation, which transfers knowledge from a large, complex model to a smaller, lightweight model while preserving accuracy. Quantization-aware training and low-rank factorization can also be employed to reduce model size and computational complexity without significantly degrading performance. Additionally, Neural Architecture Search (NAS) may be applied to automatically discover efficient architectures tailored to specific hardware constraints.

The integration of pruning and sparsification techniques can further streamline ShuffleNet by eliminating redundant parameters and connections, thereby lowering inference times. These optimizations will be particularly valuable for deploying the model on edge devices, ensuring that disease detection remains both accurate and efficient in real-world conditions.

To improve the model's robustness and generalization, future research should incorporate additional sensing modalities beyond RGB imaging. Hyperspectral imaging and thermal imaging can provide complementary information on plant health by capturing subtle physiological changes that are invisible to conventional cameras. These modalities, when combined with Kinect-based imaging, can yield a richer dataset for training more robust models. Furthermore, the application of advanced data augmentation strategies and the use of transfer learning from large-scale agricultural datasets can improve the adaptability of the model to diverse field conditions and crop varieties.

The integration of Internet of Things (IoT) devices and edge computing represents another promising avenue. By deploying optimized versions of the model on edge devices, real-time disease detection can be achieved directly in the field, minimizing latency and reducing dependence on high-bandwidth communication with centralized servers. Coupling these systems with cloud-based platforms can enable the aggregation and analysis of data from multiple farms, supporting regional-scale monitoring and coordinated disease management strategies. This networked approach can accelerate responses to emerging outbreaks and reduce crop losses across larger agricultural areas.

Finally, the scope of future research should extend beyond detection to encompass automated treatment and control mechanisms. Linking disease detection models with robotic

sprayers or UAVs (drones) would enable precision agriculture systems to autonomously apply pesticides or adjust environmental parameters in response to detected infections. Reinforcement learning could be used to develop adaptive spraying strategies that optimize pesticide usage, balancing effective treatment with cost efficiency and environmental sustainability. Such integrated systems have the potential to transform plant disease management by creating closed-loop solutions that not only detect but also treat diseases in real time.

REFERENCES

Aabidi, M. H., El Makrani, A., Jabir, B., & Zaimi, I. (2023). A new approach of leaf disease detection using bag of visual words. *Journal of Theoretical and Applied Information Technology, 31*(14), 1–20.

Abdullah, R. G. (2016). *Accessibility and Development in Rural Sarawak: A case study of the Baleh River Basin, Kapit District, Sarawak, Malaysia* [Doctoral thesis, Victoria University of Wellington]. Victoria University of Wellington Research Portal. Retrieved November 14, 2024, from https://openaccess.wgtn.ac.nz/articles/thesis/Accessibility_and_development_in_rural_Sarawak_A_case_study_of_the_Baleh_river_basin_Kapit_District_Sarawak_Malaysia/17020055?utm_source

Abisha, S., Mutawa, A. M., Murugappan, M., & Krishnan, S. (2023). Brinjal leaf diseases detection based on discrete shearlet transform and deep convolutional neural network. *PLoS ONE, 18*(4), 1–11.

Afifi, M., Derpanis, K. G., Ommer, B., & Brown, M. S. (2020). Learning multi-scale photo exposure correction. *Scientific Applications, 3*(2), 3–14.

Ahmed, A. A., & Harshavardhan Reddy, G. (2021). A mobile-based system for detecting plant leaf diseases using deep learning. *AgriEngineering, 3*(3), 478–493.

Akkara, J., & Kuriakose, A. (2019). How to guide for smartphone slit-lamp imaging. *Kerala Journal of Ophthalmology, 31*(1), 64–68.

Al-Maneea, H. M. A., & Mahdi, A. (2021). *Analysis, design and implementation of multiple view visualisations* [Doctoral dissertation, Bangor University]. Bangor University Research Portal. Retrieved December 12, 2004, from https://research.bangor.ac.uk/en/studentTheses/analysis-design-and-implementation-of-multiple-view-visualisation

Alghamdi, H., & Turki, T. (2023). PDD-Net: Plant disease diagnoses using multilevel and multiscale convolutional neural network features. *Agriculture, 13*(5), 1–19.

Alhassan, M., & Mumuni, F. (2022). Data augmentation: A comprehensive survey of modern approaches. *Array, 16,* 1–18.

Alomar, K., Aysel, H. I., & Cai, X. (2023). Data augmentation in classification and segmentation: A survey and new strategies. *Journal of Imaging, 9*(2), 1–26.

Anaya-Esparza, L. M., de la Mora, Z. V., Vázquez-Paulino, O., Ascencio, F., & Villarruel-López, A. (2021). Bell peppers (*Capsicum annuum* L.) losses and wastes: Source for food and pharmaceutical applications. *Molecules, 26*(17), 1–12.

Ang, K. M., El-Kenawy, E. S. M., Abdelhamid, A. A., Ibrahim, A., Alharbi, A. H., Khafaga, D. S., Tiang, S. S., & Lim, W. H. (2022). Teaching–learning-based optimization for optimal convolutional neural network design in image classification. *Symmetry, 14*(11), 22–27.

Arin, A., Najmin, N., & Bakar, A. (2023, July 1). Economists: Ringgit may weaken further in Q3. *The Star*. Retrieved December 19, 2024, from https://www.thestar.com.my/business/business-news/2023/07/01/economists-ringgit-may-weaken-further-in-q3

Atas, M., Yardimci, Y., & Temizel, A. (2011). Aflatoxin contaminated chili pepper detection by hyperspectral imaging and machine learning. *Sensing for Agriculture and Food Quality and Safety, 12*(4), 1–13.

Bakheet, S., Al-Hamadi, A., Soliman, E., & Heshmat, M. (2023). Hybrid bag-of-visual-words and FeatureWiz selection for content-based visual information retrieval. *Sensors, 23*(3), 1–24.

Banks, M. S., Cooper, E. A., & Piazza, E. A. (2014). Camera focal length and the perception of pictures. *Ecological Psychology, 26*(2), 30–46.

Barbedo, J. G. A. (2018). Impact of dataset size and variety on the effectiveness of deep learning and transfer learning for plant disease classification. *Computers and Electronics in Agriculture, 153*, 46–53.

Barbedo, J. G. A. (2019). Plant disease identification from individual lesions and spots using deep learning. *Biosystems Engineering, 180*, 96–107.

Benzie, M., & John, A. (2015). *Reducing vulnerability to food price shocks in a changing climate* [Unpublished report]. Stockholm Environment Institute. Retrieved November 13, 2024, from https://www.sei.org/publications/reducing-vulnerability-to-food-price-shocks-in-a-changing-climate/

Bernacki, J. (2020). Automatic exposure algorithms for digital photography. *Multimedia Tools and Applications, 79*(19–20), 12751–12776.

Bernardino, R., Lins, R. D., & Barboza, R. da S. (2023). A quality, size and time assessment of the binarization of documents photographed by smartphones. *Journal of Imaging, 9*(2), 1–24.

Blahnik, V., & Schindelbeck, O. (2021). Smartphone imaging technology and its applications. *Advanced Optical Technologies, 10*(3), 145–232.

Campos, H., & Ortiz, O. (2019). The potato crop: Its agricultural, nutritional and social contribution to humankind. In H. Campos & O. Ortiz (Eds.), *The potato crop: Its agricultural, nutritional and social contribution to humankind* (Vol. 5, pp. 1–15). Springer.

Cap, Q. H., Uga, H., Kagiwada, S., & Iyatomi, H. (2020). An effective data augmentation method for practical plant disease diagnosis. *BioScience, 3*(2), 1–10.

Chelladurai, K., & Sujatha, N. (2023). Review of disease detection in leaves using image processing techniques based on thermal camera. *International Journal of Intelligent Systems and Applications in Engineering, 12*(10), 175–184.

Chen, L. F., Rojas, M., Kon, T., Gamby, K., Xoconostle-Cazares, B., & Gilbertson, R. L. (2019). A severe symptom phenotype in Malian tomatoes is caused by a reassortant between a novel recombinant begomovirus (Tomato yellow leaf curl Mali virus) and a betasatellite. *Molecular Plant Pathology, 10*(3), 415–430.

Cigánek, J., Kozák, S., & Kozáková, A. (2020). Health safety training for industry in virtual reality. *Internet Exploration, 4*(2), 1–5.

Clark, R. A., Pua, Y. H., Oliveira, C. C., Bower, K. J., Thilarajah, S., McGaw, R., & Mentiplay, B. F. (2013). Reliability and concurrent validity of the Microsoft Kinect V2 for assessment of standing balance and postural control. *Gait & Posture*, 36(3), 372–377.

Core, L. (2013). *Plant taxonomy* (Vol. 1). Academic Press.

Damalas, C. A., & Eleftherohorinos, I. G. (2021). Pesticide exposure, safety issues, and risk assessment indicators. *International Journal of Environmental Research and Public Health*, 8(5), 1402–1419.

Demilie, W. B. (2024). Plant disease detection and classification techniques. *Journal of Big Data*, 11 (5), 1–43.

Dhami, N. B., Kim, S., Paudel, A., Shrestha, J., & Rijal, T. R. (2019). A review on threat of gray leaf spot disease of maize in Asia. *Journal of Maize Research and Development, 1*(1), 71–85.

Dong, J., Lee, J., Fuentes, A., Xu, M., Yoon, S., Lee, M. H., & Park, D. S. (2022). Data-centric annotation analysis for plant disease detection: Strategy, consistency, and performance. *Frontiers in Plant Science, 13*, 1–19.

Dos Santos, P. A., Almeida, J., & Flores, E. A. (2021*). Efficient real-time image recognition using collaborative swarm of UAVs and convolutional networks*. arXiv preprint arXiv:2107.04648.

Duela, S. J., Gopichandd, M., Kannan, K., & Sathish, M. G. (2023). Chilli leaf diseases detection with different features of original chilli using region-based convolutional neural network. *Journal of Intelligent Systems and Applications in Engineering, 4*(3), 298–305.

Entuni, C. J., & Zulcaffle, T. M. A. (2022). Identification of corn leaf diseases comprising of blight, grey spot and rust using DenseNet-201. *Borneo Journal of Resource Science and Technology*, 12(2), 108–118.

Fatimi, S. S. (2021). *An analysis of smartphone camera and digital camera images captured by adolescents ages fifteen to seventeen* [master's thesis, Teachers College, Columbia University]. ProQuest Dissertations & Theses. Retrieved December 11, 2024, from https://www.proquest.com/openview/7128c64b2504afce922cce08eacf235d/1?cbl=18750&diss=y&pq-origsite=gscholar.

Fawaiq, M. N., Utami, E., & Ariatmanto, D. (2023). Rice plant disease detection with data augmentation using transfer learning. *International Journal of Research Publication and Reviews, 4*(4), 2195–2199.

Ferentinos, K. P. (2018). Deep learning models for plant disease detection and diagnosis. *Computers and Electronics in Agriculture, 145*, 311–318.

Forte, M. (2014). 3D archaeology: New perspectives and challenges—the example of Çatalhöyük. *Journal of Eastern Mediterranean Archaeology and Heritage Studies*, 2(1), 1–29.

Fraiwan, M., Audat, Z., Fraiwan, L., & Manasreh, T. (2022). Using deep transfer learning to detect scoliosis and spondylolisthesis from X-ray images. *PLoS ONE, 17*(5), 1–13.

Fuentes, A., Yoon, S., Kim, S. C., & Park, D. S. (2017). A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition. *Sensors, 17*(9), 20–28.

Gheorghiu, A., Tăiatu, I.-M., Cercel, D.-C., Marin, I., & Pop, F. (2024). Evaluating data augmentation techniques for coffee leaf disease classification. *Computer Science and Agriculture, 3*(2), 11–43.

Giacaman, C. R. (2022). High-precision measurement of height differences from shadows in non-stereo imagery: New methodology and accuracy assessment. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 14*(7), 1702–1717.

Golmohammadi, R., Yousefi, H., Khotbesara, N. S., Nasrolahi, A., & Kurd, N. (2021). Effects of light on attention and reaction time: A systematic review. *Journal of Research in Health Sciences, 21*(4), 1–5.

Gonzalez, R. C., & Woods, R. E. (2018). *Digital Image Processing* (4th ed.). Pearson.

Grabinski, J., Keuper, J., & Keuper, M. (2022). Aliasing and adversarial robust generalization of CNNs. *Machine Learning, 111*(11), 3925–3951.

Grasty, S., Kaosa-Ard, A., Pednekar, S. S., Rerkasem, K., & Auger, P. (1999). *The growth and sustainability of agriculture in Asia*. Oxford University Press.

Gu, L., & Meng, J. (2022). Wireless sensor system of UAV infrared image and visible light image registration fusion. *Journal of Electrical and Computer Engineering, 2022*(2), 2–16.

Haleem, A., Javaid, M., Singh, R. P., Suman, R., & Rab, S. (2022). Holography and its applications for industry 4.0: An overview. *Internet of Things and Cyber-Physical Systems, 2*(4), 42–48.

Haque, M. A., Marwaha, S., Deb, C. K., Nigam, S., Arora, A., Hooda, K. S., Soujanya, P. L., Aggarwal, S. K., Lall, B., Kumar, M., Islam, S., Panwar, M., Kumar, P., & Agrawal, R. C. (2022). Deep learning-based approach for identification of diseases of maize crop. *Scientific Reports, 12*, 345–357.

Harakannanavar, S. S., Rudagi, J. M., Puranikmath, V. I., Siddiqua, A., & Pramodhini, R. (2022). Plant leaf disease detection using computer vision and machine learning algorithms. *Global Transitions Proceedings, 3*(1), 305–310.

He, H., Li, J., Zhang, Z., Tang, X., Song, D., & Yan, F. (2021). Impacts of cucurbit chlorotic yellows virus (CCYV) on biological characteristics of its vector *Bemisia tabaci* (Hemiptera: Aleyrodidae) MED species. *Journal of Insect Science, 21*(5), 1–9.

Holmes, G., San, C. P., Obispo, L., & Org, B. (2023). *Symptom of bacterial leaf spot photo*. University of Maryland Extension. Retrieved January 1, 2025, from https://extension.umd.edu

Hong, H., Lin, J., & Huang, F. (2020). *Tomato disease detection and classification by deep learning*. In IEEE International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE) (pp. 25–29). IEEE.

Huynh-Thu, Q., & Ghanbari, M. (2008). Scope of validity of PSNR in image/video quality assessment. *Electronics Letters, 44(13)*, 800–801.

Iqbal, S., Qureshi, A. N., Li, J., & Mahmood, T. (2023). Analysis of medical images using conventional machine learning methods and convolutional neural networks. *Archives of Computational Methods in Engineering, 30*(5), 3173–3233.

Jain, A., Sarsaiya, S., Wu, Q., Lu, Y., & Shi, J. (2019). A review of plant leaf fungal diseases and its environment speciation. *Bioengineered, 10*(1), 409–424.

Johnson, J. M., & Khoshgoftaar, T. M. (2019). Survey on deep learning with class imbalance. *Journal of Big Data, 6*(1), 1–19.

Joshi, K., Awale, R., Ahmad, S., Patil, S., & Pisal, V. (2022). Plant leaf disease detection using computer vision techniques and machine learning. *International Technical Meeting Web, 4*(1), 1–6.

Jung, M., Song, J. S., Shin, A. Y., Choi, B., Go, S., Kwon, S. Y., Park, J., Park, S. G., & Kim, Y. M. (2023). Construction of deep learning-based disease detection model in plants. *Scientific Reports, 13*(1), 1–13.

Kadambi, A., Bhandari, A., & Raskar, R. (2014). 3D depth cameras in vision: Benefits and limitations of the hardware with an emphasis on the first- and second-generation Kinect models. In R. R. Vemuri (Ed.), *Advances in Computer Vision and Pattern Recognition* (Vol. 67, pp. 3–26). Springer London.

Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey. *Computers and Electronics in Agriculture*, 147, 70–90.

Karahan, S., Yildirim, M. K., Kirtac, K., Rende, F. S., Butun, G., & Ekenel, H. K. (2016). How image degradations affect deep CNN-based face recognition? *Computers and Informatics, 35*(4), 1–8.

Karnik, J., & Suthar, A. (2021). Agricultural plant leaf disease detection using deep learning techniques. *Communication and Information Processing, 2*(1), 1–7.

Kaushik, S., Srivastava, K., Kaushik, S., Sharma, I., Jindal, I., & Deshwal, V. (2022). Plant leaf disease detection using machine learning. *Journal of Pharmaceutical Negative Results, 13*(10), 1–12.

Khoo, H. E., Azlan, A., Tang, S. T., & Lim, S. M. (2017). Anthocyanidins and anthocyanins: Colored pigments as food, pharmaceutical ingredients, and the potential health benefits. *Food and Nutrition Research, 61*(1), 1–21.

Khoo, Y. W., Baadu, R., Shaiddin, A., Tan, H. T., Khaw, Y. S., Li, S.-F., & Chong, K. P. (2023). First report of *Curvularia lunata* causing leaf spot on *Oryza sativa* in Sabah, Malaysian Borneo. *Plant Disease, 107*(7), 2234–2235.

Kim, D. H., López, G., Kiedanski, D., Maduako, I., Ríos, B., Descoins, A., Zurutuza, N., Arora, S., & Fabian, C. (2021). Bias in deep neural networks in land use characterization for international development. *Remote Sensing*, *13*(15), 1–20.

Ko, J. H., Yoo, Y. J., Lee, Y., Jeong, H.-H., & Song, Y. M. (2022). A review of tunable photonics: Optically active materials and applications from visible to terahertz. *Cell Science*, *25*(19), 2–25.

Kutawa, A. B., Ahmad, K., Ali, A., Zobir, S. A. M., & Wahab, M. A. A. (2021). State of the art on southern corn leaf blight disease incited by *Cochliobolus heterostrophus*: Detection, pathogenic variability and novel control measures. *Agriculture*, 11(11), 1124–1137.

Lambat, M., Kothari, R., Kabi, M., & Mane, T. (2022). Plant disease detection using InceptionV3. *International Research Journal of Engineering and Technology*, *9*(6), 1–6.

Li, S., Zhang, F., & Wu, J. (2022). Lightweight convolutional neural networks for plant disease recognition: A review. *IEEE Transactions on Neural Networks and Learning Systems*, *33*(4), 1234–1245.

Lim, L. L. (2008). *Chilli cultivation*. Department of Agriculture Sarawak. Retrieved November, 12, 2024, from https://doa.sarawak.gov.my/web/subpage/webpage_view/582

Lin, C. L., & Wu, K. C. (2023). Development of revised ResNet-50 for diabetic retinopathy detection. *BMC Bioinformatics*, *24*(1), 1–18.

Liu, J., Wang, X., & others. (2021). Plant diseases and pests' detection based on deep learning: a review. *Plant Methods*, 17(1), 22–32.

Ma, N., Zhang, X., Zheng, H.-T., & Sun, J. (2018). ShuffleNet V2: Practical guidelines for efficient CNN architecture design. *Pattern Recognition Letters, 2*(1), 1–19.

Mahmoudieh, M., Noor, M. R. M., Harikrishna, J. A., & Othman, R. Y. (2020). Identification and characterization of *Ageratum yellow vein Malaysia virus* (AYVMV) and an associated betasatellite among begomoviruses infecting *Solanum lycopersicum* in Malaysia. *Journal of Applied Genetics, 61*(4), 619–628.

Maner, M. (2023, July 20). Ludhiana: Growers stung again, high temperature leaves paddy yellow. *Tribune News Service*. Retrieved December 17, 2025, from https://www.tribuneindia.com

Min, B., Kim, T., Shin, D., & Shin, D. (2023). Techniques for enhancing plant leaf disease recognition through data augmentation. *Applied Sciences, 13***(**3), 1–17.

Mittal, A., Moorthy, A. K., & Bovik, A. C. (2012). No-reference image quality assessment in the spatial domain. *IEEE Transactions on Image Processing, 21*(12), 4695–4708.

Mohammad, B. (2022). *Supply and utilization accounts selected agricultural commodities, Malaysia 2017–2021*. Department of Statistics Malaysia. Retrieved November 21, 2024, from https://www.dosm.gov.my

Mohammed, L., & Yusoff, Y. (2023). Identifying and categorizing plant leaf diseases through digital image processing techniques. *ASEAN Engineering Journal, 13*(1), 1–9.

Mohanty, S. P., Hughes, D. P., & Salathé, M. (2016). Using deep learning for image-based plant disease detection. *Frontiers in Plant Science*, 7, 14–19.

Monaghan, T. F., Rahman, S. N., Agudelo, C. W., Wein, A. J., Lazar, J. M., Everaert, K., & Dmochowski, R. R. (2021). Foundational statistical principles in medical research: Sensitivity, specificity, positive predictive value, and negative predictive value. *Medicina, 57*(5), 1–7.

Morales, A., Horstrand, P., Guerra, R., Leon, R., Ortega, S., Díaz, M., Melián, J. M., López, S., López, J. F., Callico, G. M., Martel, E., & Sarmiento, R. (2022). Laboratory hyperspectral image acquisition system setup and validation. *Sensors*, *22*(6), 1–25.

Morikawa, C., Kobayashi, M., Satoh, M., Kuroda, Y., Inomata, T., Matsuo, H., Miura, T., & Hilaga, M. (2021). Image and video processing on mobile devices: A survey. *The Visual Computer, 37*(12), 2931–2949.

Muhammad, A., Salman, Z., Lee, K., & Han, D. (2023). Harnessing the power of diffusion models for plant disease image augmentation. *Frontiers in Plant Science, 14*(7), 1–19.

Munsell, F. (2019). *Munsell color charts for plant tissue* (Vol. 1). Munsell Color.

Muntean, L. (2022, June 15). Why are my tomato leaves curling? *Agri Life Today*. Retrieved December 10, 2024, from https://agrilifetoday.tamu.edu

Nagasubramanian, K., Jones, S., Singh, A. K., Sarkar, S., Singh, A., & Ganapathysubramanian, B. (2019). Explainable 3D deep learning for plant disease detection using hyperspectral imaging. *Plant Methods, 15*(1), 5–10.

Nahiduzzaman, M., Chowdhury, M. E. H., Salam, A., Nahid, E., Ahmed, F., Al-Emadi, N., Ayari, M. A., Khandakar, A., & Haider, J. (2023). Interpretable deep learning model for automated classification of mulberry leaf diseases. *Frontiers in Plant Science, 14*(11), 1–19.

Nesarajan, D., Kunalan, L., Logeswaran, M., Kasthuriarachchi, S., & Lungalage, D. (2020). *Coconut disease prediction system using image processing and deep learning techniques*. In IEEE International Conference on Image Processing, Applications and Systems (IPAS) (pp. 212–217). IEEE.

Nguyen, H.T., Luong, H.H., Huynh, L.B., Le, B.Q.H., Doan, N.H., & Le, D.T.D. (2023). An improved MobileNet for disease detection on tomato leaves. *Advances in Technology Innovation, 8*(3), 192–209.

Nicora, G., Rios, M., Abu-Hanna, A., & Bellazzi, R. (2022). Evaluating pointwise reliability of machine learning prediction. *Journal of Biomedical Informatics, 127*, 104–120.

Olsen, A., Konovalov, D. A., Philipp, B., Ridd, P., Wood, S., Anderson, A., & Walters, D. (2019). DeepWeeds: A multiclass weed species image dataset for deep learning. *Scientific Reports, 9*(1), 2058–2070.

Omar, S. C. (2022). *Deconstructing Malaysia's food import bill: It is not the only measure of food security* (Vol. 3). Khazanah Institute. Retrieved December 1, 2024, from https://www.krinstitute.org/assets/contentMS/img/template/editor/Discussion%20Paper_Food%20Import_SCO_23112022.pdf

Pannakkong, W., Thiwa-Anont, K., Singthong, K., Parthanadee, P., & Buddhakulsomsiri, J. (2022). Optimizing machine learning hyperparameters with response surface methodology: A case study on ANN, SVM, and DBN. *Mathematical Problems in Engineering*, *2022*, 1–17.

Panno, S., Davino, S., Caruso, A. G., Bertacca, S., Crnogorac, A., Mandić, A., Noris, E., & Matić, S. (2021). A review of the most common and economically important diseases that undermine the cultivation of tomato crops in the Mediterranean basin. *Agronomy*, *11*(11), 1–25.

Park, S. H., Choi, I. Y., Lee, W. H., Lee, K. J., Galea, V., & Shin, H. D. (2020). Identification and characterization of *Cercospora malayensis* causing leaf spot on kenaf. *Mycobiology*, *45*(2), 114–118.

Patayon, U. B., & Crisostomo, R. V. (2021). Automatic identification of abaca bunchy top disease using deep learning models. *Procedia Computer Science*, *179*, 321–329.

Petrellis, N. (2019). Smartphone-based plant disease diagnosis with an expandable set of diseases. *Applied Sciences*, *9*(9), 1–22.

Perez, R., & Wang, J. (2017). The effectiveness of data augmentation in image classification using deep learning. *ArXiv*, 1–8.

Picon, A., Seitz, M., Alvarez-Gila, A., Mohnke, P., Ortiz-Barredo, A., & Echazarra, J. (2019). Condition-based convolutional neural networks for large-scale multi-crop disease classification using cell phone images captured in real field conditions. *Computers and Electronics in Agriculture*, *161*, 1–18.

Pistellato, M., Fatima, T., & Wimmer, M. (2023). Utilizing light polarization for deep high dynamic range (HDR) imaging from a single exposure. *Sensors, 23*(12), 1–13.

Preetham, A., Ahmad, S. S., Wattar, I., Singh, P., Rout, S., Alqahtani, M. A., & Tetteh Amoatey, E. (2022). Instinctive recognition of pathogens in rice using reformed fractional differential segmentation and innovative fuzzy logic-based probabilistic neural network. *Journal of Food Quality, 2022*, 1–17.

Priyaradhikadevi, T., Mohan, R., Ragupathi, T., Prasanna, S., Madhan, K., & Ananthi, R. (2023). *Machine learning-based detection of leaf diseases*. In IEEE International Conference on Science, Technology, Engineering and Mathematics (STEM) (pp. 1–5). IEEE.

Pujar, P., Kumar, A., & Kumar, V. (2024). A survey on plant leaf disease identification and classification by various machine-learning technique. *IAES International Journal of Artificial Intelligence*, 13(2), 1187-1194.

Pujari, J., Yakkundimath, R., & Byadgi, A. S. (2016). SVM and ANN based classification of plant diseases using feature reduction technique. *International Journal of Interactive Multimedia and Artificial Intelligence, 3*(7), 6–14.

Raji, M. N. A., Ab Karim, S., Ishak, F. A. C., & Arshad, M. M. (2017). Past and present practices of the Malay food heritage and culture in Malaysia. *Journal of Ethnic Foods, 4*(4), 221–231.

Rakib, M. R. H. K., Pramanik, S. A. K., Amran, M. A., Islam, M. N., & Sarker, M. O. F. (2022). Factors affecting young customers' smartphone purchase intention during COVID-19 pandemic. *Heliyon, 8*(9), 103–111.

Ram, H. (2018). Understanding the bias-variance tradeoff. *Random Structures and Algorithms, 2*(7), 1–12.

Ramcharan, A., Baranowski, K., McCloskey, P., Ahmed, B., Legg, J., & Hughes, D. P. (2017). Image-based cassava disease detection using deep learning. *Frontiers in Plant Science, 8*, 1–6.

Rebitanim, N. A., Hanafi, M. M., Idris, A. S., Abdullah, S. N. A., Mohidin, H., & Rebitanim, N. Z. (2020). Improves oil palm growth and resistance against Ganoderma basal stem rot disease in nursery and field trials. *BioMed Research International, 2020*, 1–10.

Reddy, Y. A., & Adimoolam, M. (2022). *A framework system for plant leaf disease detection using K-nearest neighbours and comparison of its features with Naive Bayes classification*. In IEEE International Conference on Business Analytics for Technology and Security (ICBATS) (pp. 1–4). IEEE.

Rehman, H. U., Anwar, S., & Tufail, M. (2020). Machine vision-based plant disease classification through leaf imaging. *Ingénierie des Systèmes d'Information, 25*(4), 437–444.

Rumchev, K., Gilbey, S., Mead-Hunter, R., Selvey, L., Netto, K., & Mullins, B. (2019). Agricultural dust exposures and health and safety practices among Western Australian Wheatbelt farmers during harvest. *International Journal of Environmental Research and Public Health, 16*(24), 98–112.

Sabrol, H., & Satish, K. (2016). Classification of tomato plant diseases in digital images using a decision tree. *Communication and Signal Processing, 6*(8), 1–5.

Saibeh, K., Batumale, S. R. A. P., & Boyce, P. C. (2020). Research on Monstereae (Araceae) of Borneo II: Clarification of Furtado's *Rhaphidophora kinabaluensis* and its reclassification into *Scindapsus*. *Willdenowia, 45*(3), 409–413.

Salih, T. A., Ali, A. J., & Ahmed, M. N. (2020). Deep learning convolution neural network to detect and classify tomato plant leaf diseases. *Open Access Library Journal, 7*(5), 1–12.

Salvi, M., Acharya, U. R., Molinari, F., & Meiburger, K. M. (2021). The impact of pre- and post-image processing techniques on deep learning frameworks: A comprehensive review for digital pathology image analysis. *Computers in Biology and Medicine, 128*, 1–12.

Sau, A. R., Nazmie, N. M. F., Yusop, M. S. M., Akbar, M. A., Saad, M. F. M., Baharum, S. N., Talip, N., Goh, H., Kassim, H., & Bunawan, H. (2020). First report of *Pepper vein yellows virus* and *Pepper yellow leaf curl virus* infecting chilli pepper (*Capsicum annuum*) in Malaysia. *Plant Disease, 104*(7), 2024–2032.

Setiawan, W., Rochman, E. M. S., Satoto, B. D., & Rachmad, A. (2022). Machine learning and deep learning for maize leaf disease classification: A review. *Journal of Physics: Conference Series, 2022*(24), 1–12.

Shafiq, I., Hussain, S., Raza, M. A., Iqbal, N., Asghar, M. A., Raza, A., Fan, Y.-F., Mumtaz, M., Shoaib, M., Ansar, M., Manaf, A., Yang, W.-Y., & Yang, F. (2021). Crop photosynthetic response to light quality and light intensity. *Journal of Integrative Agriculture, 20*(1), 4–23.

Shen, L., Sun, Y., Yu, Z., Ding, L., Tian, X., & Tao, D. (2023). On efficient training of large-scale deep learning models: A literature review. *Artificial Intelligence, 2*(4), 46–60.

Shih, S.-L., Maxwell, D., & Noor, M. R. (2018). A new geminivirus associated with leaf curl disease of tomato in Malaysia. *Plant Bulletin, 2*(12), 1–2.

Shingote, P. R., Wasule, D. L., Parma, V. S., Holkar, S. K., Karkute, S. G., Parlawar, N. D., & Senanayake, D. M. J. B. (2022). An overview of chili leaf curl disease: Molecular mechanisms, impact, challenges, and disease management strategies in the Indian subcontinent. *Frontiers in Microbiology, 13*, 1–18.

Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. *Journal of Big Data, 6*(1), 35–48.

Shotton, J., Fitzgibbon, A., Cook, M., Sharp, T., Finocchio, M., Moore, R. & Blake, A. (2011). Real-time human pose recognition in parts from single depth images. *CVPR 2011*, 1297–1304.

Siena, F. L., Byrom, B., Watts, P., & Breedon, P. (2018). Utilising the Intel RealSense camera for measuring health outcomes in clinical research. *Journal of Medical Systems, 42*(3), 1–10.

Sihombing, Y. F., Septiarini, A., Kridalaksana, A. H., & Puspitasari, N. (2022). Chili classification using shape and color features based on image processing. *Scientific Journal of Informatics, 9*(1), 42–50.

Soeb, M. J. A., Jubayer, M. F., Tarin, T. A., Al Mamun, M. R., Ruhad, F. M., Parven, A., Mubarak, N. M., Karri, S. L., & Meftaul, I. M. (2023). Tea leaf disease detection and identification based on YOLOv7 (YOLO-T). *Scientific Reports, 13*(1), 1–16.

Srivastava, S., & Nelson, S. (2021). Plant disease cercospora leaf spot of eggplant. *University of Hawai'i Cooperative Extension Service*. Retrieved January 3, 2025, from http://www.ctahr.hawaii.edu/freepubs

Sujatha, R., Chatterjee, J. M., Jhanjhi, N. Z., & Brohi, S. N. (2021). Performance of deep learning vs. machine learning in plant leaf disease detection. *Microprocessors and Microsystems, 80*, 1–7.

Sujon, K. M., Hassan, R., & Choi, K. et al. (2025). Accuracy, precision, recall, F1-score, or MCC? Empirical evidence from advanced statistics, ML, and XAI for evaluating predictive models. *Journal of Big Data, 12*, 48–65.

Tahamid, A. (2020). *Tomato leaf disease detection using ResNet-50 and MobileNet architecture* [Bachelor's thesis, Brac University]. DSpace, Brac University. Retrieved December 12, 2024, from https://dspace.bracu.ac.bd/xmlui/handle/10361/14448Talaviya, T., Shah, D., Patel, N., Yagnik, H., & Shah, M. (2020). Implementation of artificial intelligence in agriculture for optimisation of irrigation and application of pesticides and herbicides. *Artificial Intelligence in Agriculture, 4*, 58–73.

Tanaka, S., Wasli, M. E. B., Seman, L., Jee, A., Kendawang, J. J., Sakurai, K., & Morooka, Y. (2007). Ecological study on site selection for shifting cultivation by the Iban of Sarawak, Malaysia: A case study in the Mujong River area. *Tropics, 16*(4), 357–371.

Tanil, C. T., & Yong, M. H. (2020). Mobile phones: The effect of its presence on learning and memory. *PLoS ONE, 15*(8), 123–135.

Taylor, G., & Nitschke, L. (2018). Improving deep learning with generic data augmentation. *IEEE Transactions on Neural Networks and Learning Systems*, 29(12), 5664–5670.

Thompson, L., & Johnson, S. (1953). *Checklist of fungi of Malaysia* (Vol. 4). Malaysian Agricultural Research Institute.

Tong, J., Zhou, J., Liu, L., Pan, Z., & Yan, H. (2012). Scanning 3D full human bodies using Kinects. *IEEE Transactions on Visualization and Computer Graphics*, 18(4), 643–650.

Too, E. C., Yujian, L., Njuki, S., & Yingchun, L. (2019). A comparative study of fine-tuning deep learning models for plant disease identification. *Computers and Electronics in Agriculture, 161*, 272–279.

Trevethan, R. (2017). Sensitivity, specificity, and predictive values: Foundations, pliabilities, and pitfalls in research and practice. *Frontiers in Public Health, 5*, 307.

Ullah, N., Khan, J., & Ahmad, W. (2022). An efficient approach for crop pest recognition and classification based on DeepPestNet deep learning model. *IEEE Access*, *10*, 73019–73032.

Tugrul, B., Elfatimi, E., & Eryigit, R. (2022). Convolutional neural networks in detection of plant leaf diseases: A review. *Agriculture*, *12*(8), 1–21.

Vasavi, P., Punitha, A., & Rao, T. V. N. (2023). Chili crop disease prediction using machine learning algorithms. *Revue d'Intelligence Artificielle, 37*(3), 727–732.

Vidhyapathi, C., Noel, A., & Sampath, S. (2020). FPGA-accelerated action recognition using custom 3D-DTW IP. *Journal of Imaging Science and Technology, 64*(1), 65–78.

Vorenkamp, T. (2019). *A guide to neutral density filters*. B&H Explora. Retrieved November 21, 2024, from https://www.bhphotovideo.com/explora

Xia, C., Wang, L., Chung, B. K., & Lee, J. M. (2015). In situ 3D segmentation of individual plant leaves using an RGB-D camera for agricultural automation. *Sensors*, *15*(8), 20463–20479.

Xia, X., Xiang, H., Cao, Y., Ge, Z., & Jiang, Z. (2023). Humanoid-eye binocular image feature extraction and matching using the Susan-SIFT algorithm. *Biomimetics, 8*(2), 1–16.

Xiao, F. (2022). Image augmentation improves few-shot classification performance in plant disease recognition. *Information Technology, 2*(1), 1–24.

Yamashita, R., Nishio, M., Do, R. K. G., & Togashi, K. (2018). Convolutional neural networks: An overview and application in radiology. *Insights into Imaging, 9*(4), 611–629.

Zhang, R., Zhang, Z., Han, J., Yang, L., Li, J., Song, Z., Wang, T., & Zhu, J. (2023). Advanced liquid crystal-based switchable optical devices for light protection applications: Principles and strategies. *Light: Science and Applications, 12*(1), 2–23.

Zhang, W., & Zhao, Y. (2023). An enhanced SIFT algorithm for SAR and optical image registration. *Scientific Reports, 13*(1), 1–8.

Zhang, X., Zhou, X., & Lin, M. (2020). ShuffleNet: *An extremely efficient convolutional neural network for mobile devices*. In IEEE International Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 6848–6856). IEEE.

Zhang, Y., Carballo, A., Yang, H., & Takeda, K. (2023). Perception and sensing for autonomous vehicles under adverse weather conditions: A survey. *ISPRS Journal of Photogrammetry and Remote Sensing, 196*(3), 146–177.

Zhao, F., & Bajic, V. B. (2015). The value and significance of metagenomics of marine environments. *Genomics, Proteomics and Bioinformatics, 13*(5), 271–274.

Zhao, P., Li, C., Rahaman, M. M., Xu, H., Yang, H., Sun, H., Jiang, T., & Grzegorzek, M. (2022). A comparative study of deep learning classification methods on a small environmental microorganism image dataset (EMDS-6): From convolutional neural networks to visual transformers. *Frontiers in Microbiology*, *13*(2), 1–21.

Zhou, K., Lindenbergh, R., & Gorte, B. (2019). Automatic shadow detection in urban very-high-resolution images using existing 3D models for free training. *Remote Sensing*, *11*(1), 1–15.

APPENDICES

Appendix 1: Review of Existing Plant Detection Devices: An Analysis of Plant Types, Disease Detection, and Accuracy using Various Image Sources and Technologies

Study (Author, Year)	Plant Type(s)	Image Source /Device	Deep Learning / ML Model Used	Type of Diseases Detected	Accuracy / Performance
Petrellis (2019)	Citrus	5 MP Lumia 550; 23 MP Sony Xperia XA1 (Owned by authors)	Fuzzy Classifier	Alternaria, Anthracnose, Citrus Chlorotic Dwarf Virus, Nutrient Deficiency, Melanose	90%
Pujari et al. (2016)	Maize, Cucumber, Cotton, Lime, Grape, Tomato, Sugar Beet, Soybean	Sony DXC-3000A	SVM	Fungal, Bacterial, Viral, Nematodes, Deficiency	92.17%

Appendix 1: continued

Salih et al., (2020)	Tomato	5 MP Camera	Deep Learning CNN (DLCNN)	Bacterial Spot, Late Blight, Mosaic, Yellow Curved Leaf, Healthy	96.34%
Sabrol & Satish (2016)	Tomato	Standard Digital Camera (Web Scraper)	Local Neighbour Comparison	Bacterial Leaf Spot, Fungal Septoria Leaf Spot, Leaf Blight, Curl, Canker	93.70%
Sujatha et al., (2021)	Citrus	DSLR (72 dpi, 256×256 px)	Deep Learning + ML	Canker, Black Spot, Healthy	76.8%–87%
Haque et al., (2022)	Maize	Digital Cameras and Smartphones	Inception v3	Maydis Leaf Blight, Turcicum Leaf Blight, Sheath Blight	95.99%
Joshi et al., (2022)	Apple, Corn, Grapes, Potato, Tomato	Sony DSC-RX100/13 (20.2 MP)	Random Forest (RF)	Multiple species (7,000 images total)	93.7%

Appendix 1: continued

Preetham et al., (2022)	Rice	Canon EOS 3000D	k-Means	Bacterial Leaf Blight, Brown Spot, Rice Blast	94.2%–95.8%
Sihombing et al., (2021)	Capsicum	Flir AX8 Thermal Camera	k-NN	Leaf Curl, Veinal Mottle, Alfalfa Mosaic, Beet Curly Top	92.8% (Accuracy), 87.3% (F1)
Vasavi et al., (2023)	Capsicum	Mi Note 7 Prime	RF, AdaBoost, GB, MLP	Bacterial Leaf Spot, Fusarium, Curl, Pests, Healthy	94%–96%
Picon et al., (2019)	Barley, Rice, Maize, Wheat, Rapeseed	Cell Phone	CNN	Powdery Mildew, Rust, Blight, Sclerotinia Rot, Blackleg	78%–87%
Hong et al., (2020)	Tomato	Smartphone Camera	CNN	Powdery Mildew, Early/Late Blight, Leaf Mold, Target Spot	97.1%

Appendix 1: continued

Nguyen et al., (2023)	Tomato	Dataset (2064 images)	MobileNet	Multiple Diseases	95.7%
Lambat et al., (2022)	Vegetative Plants	Dataset (500 images)	Inception	General Plant Diseases	88.7%
Tahamid (2020)	Tomato	Dataset	ResNet50	Bacterial Spot, Early Blight, Leaf Mold, Spider Mites	94%

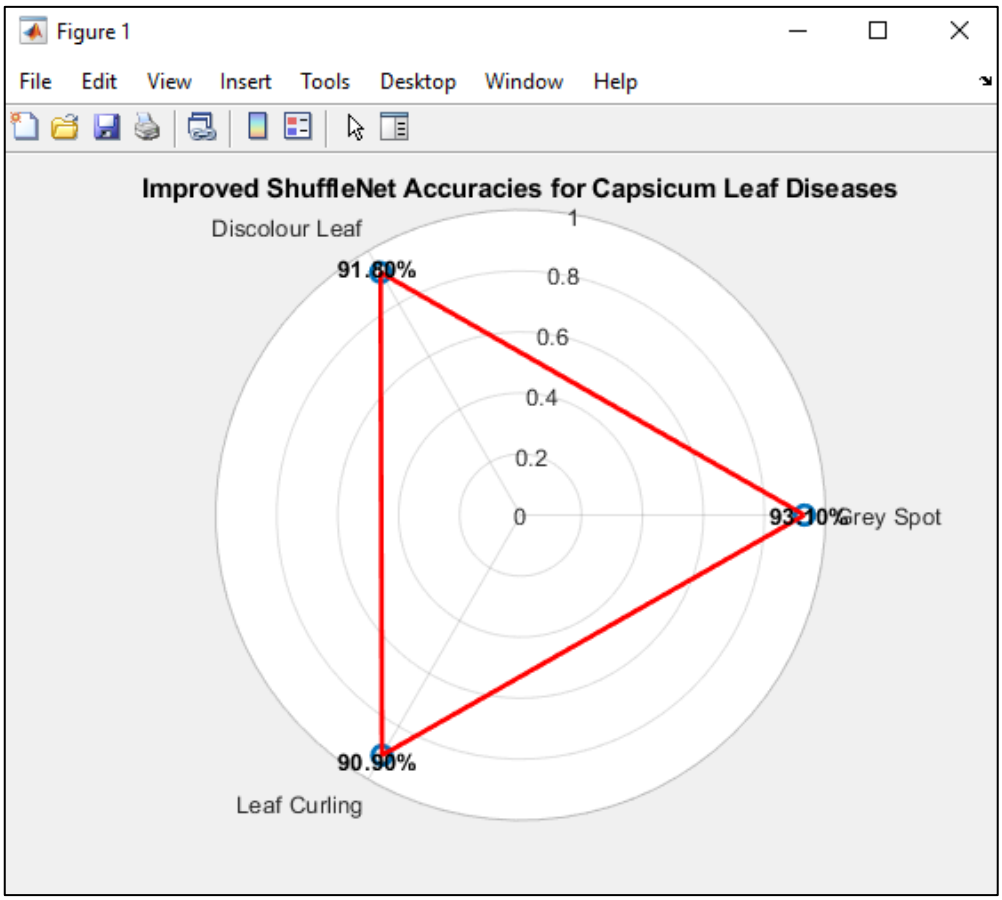
Appendix 2: Detailed Intrinsic Parameters and Specifications of Standard RGB Camera, Kinect Camera and Mobile Camera

Parameter	Standard RGB	Kinect	Mobile
Sensor size (µm)	3.1	10.0	1.4
Focal length x-axis (mm)	3.28	3.64	2.2
Focal length y-axis (mm)	3.52	3.9	2.5
Principal point x (pixel)	965.11	263.85	800
Principal point y (pixel)	583.27	225.71	600
ISO Sensitivity	100.0	800.0	400
Shutter Speed (s)	0.016	0.002	1/100
Aperture (f-stop)	2.8	1.8	2.0
Dynamic Range (dB)	60.0	70.0	55
Lens Distortion (%)	0.5	0.3	0.7
Chromatic Aberration (px)	0.1	0.05	0.2
Vignetting (EV)	0.2	0.1	0.3
Signal-to-Noise Ratio (dB)	35.0	40.0	30

Appendix 3: Detailed Performance Evaluation using Confusion Matrix Across Multiple Models

Model	True Positive	False Positive	True Negative	False Negative	False Discovery Rate	False Omission Rate	Positive Predictive Value	Negative Predictive Value	Accuracy	Balanced Accuracy	F1 Score	Matthews Correlation	Informedness	Markedness
Improved ShuffleNet	98	1	99	0	0.01	0.0	0.99	1.0	0.99	0.99	0.99	0.98	0.99	0.99
ShuffleNet	80	10	82	8	0.11	0.09	0.89	0.91	0.81	0.81	0.84	0.75	0.72	0.8
DenseNet201	95	2	94	3	0.02	0.03	0.98	0.97	0.94	0.94	0.96	0.92	0.88	0.95
MobileNet	85	5	90	5	0.06	0.05	0.94	0.95	0.92	0.92	0.89	0.88	0.86	0.92
ResNet50	88	6	91	4	0.06	0.04	0.94	0.96	0.93	0.93	0.9	0.89	0.87	0.92

Appendix 4: MATLAB-Generated Experimental Results for Table 4.18



Disease	Time (s)
Grey Spot	2332
Discolour Leaf	2346
Leaf Curling	2361

Appendix 5: Journal Publications

1. **Entuni, C. J.**, Zulcaffle, T. M. A., & Hong Ping, K. (2023). Classification of capsicum leaf disease from a complex cluster of leaves using an improved multiple layers ShuffleNet CNN model. International Journal of Advanced Technology and Engineering Exploration, 10(102), 515-527.

   https://doi.org/10.19101/IJATEE.2022.10100509 **(SCOPUS)**

2. **Entuni, C. J.**, Zulcaffle, T. M. A., Ping, K. H., Sharangi, A. B., Upadhyay, T. K., & Saeed, M. (2023). Smart agricultural monitoring solution for chilli leaf diseases using a low-cost Kinect camera and an improved CNN algorithm. Jurnal Teknologi (Sciences & Engineering), 85(5), 93-102.

   https://doi.org/10.11113/jurnalteknologi.v85.19884 **(SCOPUS)**