Comparative Analysis of Transfer Learning Models and Custom CNN for Tea Leaf Disease Classification

By
Prosenjith Dash Pappu
ID: 212-15-14749

Mst. Aklima Akter
ID: 212-15-14758

FINAL YEAR DESIGN PROJECT REPORT

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

Supervised by
Md. Ferdouse Ahmed Foysal
Lecturer
Department of Computer Science and EngineeringDaffodil International University

Co-Supervised by
Md Assaduzzaman
Assistant Professor
Department of Computer Science and EngineeringDaffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY
Dhaka, Bangladesh

September 17, 2025

APPROVAL

This Project titled **"Comparative Analysis of Transfer Learning Models and Custom CNN for Tea Leaf Disease Classification"**, submitted by Prosenjith Dash Papp & Mst. Aklima Akter, ID No: **212-15-14749 & 212-15-14758** to the Department of Computer Science and Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on **17 September, 2025**.

BOARD OF EXAMINERS

Dr. Arif Mahmud — **Chairman**
Associate Professor & Associate Head
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Dr. Fizar Ahmed — **Internal Examiner**
Associate Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Md. Mushfiqur Rahman — **Internal Examiner**
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Dr. Md. Manowarul Islam — **External Examiner**
Professor
Department of Computer Science and Engineering
Jagannath University

DECLARATION

We hereby declare that this project has been done by us under the supervision of **Md. Ferdouse Ahmed Foysal, Lecturer**, Department of Computer Science and Engineering, Daffodil International University. We also declare that neither this project nor any part of this project has been submitted elsewhere for the award of any degree or diploma.

Supervised by:

Md. Ferdouse Ahmed Foysal
Lecturer
Department of Computer Science and Engineering Daffodil International University

Co-Supervised by:

Md Assaduzzaman
Assistant Professor
Department of Computer Science and Engineering Daffodil International University

Submitted by:

Prosenjith Dash Pappu
Student ID: 212-15-14749
Department of Computer Science and Engineering Daffodil International University

Aklima

Mst. Aklima Akter
Student ID: 212-15-14758
Department of Computer Science and Engineering Daffodil International University

ACKNOWLEDGEMENTS

This work would not have been possible without the support and contributions of many individuals over the past two semesters. We are deeply grateful to everyone who has assisted us in one way or another.

First, we express our heartfelt thanks and gratefulness to the almighty for His divine blessing making it possible for us to complete the **Final Year Design Project (FYDP)** successfully.

We are grateful and wish our profound indebtedness **Md. Ferdouse Ahmed Foysal**, **Lecturer** Department of Computer Science and Engineering, Daffodil International University, Dhaka, Bangladesh. Deep knowledge and keen interest of our supervisor in the field of **Deep Learning** to carry out this project. His endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior drafts, and correcting them at all stages have made it possible to complete this project.

We would like to express our heartfelt gratitude to the Head of the Department of Computer Science and Engineering, for his kind help in finishing our project and also to other faculty members and the staff of the Department of Computer Science and Engineering, Daffodil International University.

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

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

ABSTRACT

Tea is one of the most consumed drinks in the world and a savior of agricultural economies especially to nations like Bangladesh, India and Sri Lanka. However, diseases of the leaves like Algal Leaf, Gray Blight, Helopeltis, Looper Infestation, and Red Spider largely impact on the tea cultivation to the extent that they lower yield and quality. Conventional techniques are time consuming, error prone, inaccurate, and could not be applied to large plantations and required automatic and accurate solutions. This paper examines how deep learning models can be applied to the problem of tea leaf disease classification by comparing InceptionV3, MobileNetV2 or a Custom Convolutional Neural Network (CNN). The data of 1,968 pictures was gathered and enriched to 7,584 samples in order to provide diversity and robustness to the six classes. To evaluate the models, accuracy, precision, recall, F1-score and confusion matrices were used. Findings indicate that InceptionV3 had the highest accuracy of 98.94, Custom CNN had 96.6, and MobileNetV2 had 83.8. InceptionV3 was much more effective, but its high calculative cost renders it inappropriate in the field of real-time use. The Custom CNN was between the performance and accuracy, whereas MobileNetV2 was applicable to the deployment on mobile or edge devices owing to its lightweight architecture. The study exemplifies how deep learning can be applied in precision agriculture, and the lessons on what model to use based on resources and application. The results may be used to inform the further development of scalable, IoT-enabled disease detection to ensure sustainable agriculture and minimize the use of excessive amounts of pesticides.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

This chapter provides an overview of the research on the use of deep learning techniques for tea leaf disease classification. The problem domain is briefly described first, followed by the motivation, objectives, methodology, and expected outcomes of the study, and finally, the structure of the report.

1.1 Introduction

One of the most popular drinks in the world, tea plays a very significant role in the economy of most states. Nevertheless, multiple leaf diseases endanger the quantity and quality of tea production. Production of these diseases requires early detection and classification in the name of sustainability. Disease detection has been automatic with deep learning. This paper categorizes tea leaf diseases with a Custom CNN and transfer learning models InceptionV3 and MobileNetV2.

1.2 Motivation

In places like Bangladesh, India, and Sri Lanka, tea-growing plays an important role in the agricultural economy. Diseases such as algal leaf, gray blight, red spider, helopeltis and looper infestation often affect the yield and quality of the tea leaves. Unless these diseases are detected early, they may develop at a very high rate causing extreme economic losses. The traditional disease detection methods are manual and have a high time cost not to mention subjectivity and human error.

Plant disease identification is possible with current technology from computer vision and machine learning as a potential alternative. High-level features of leaf images and images can be learned and classified with ease by deep learning networks, i.e., convolutional neural networks. Transfer learning also increases performance because it leverages models that have already been trained on large data sets and requires fewer large-scale agricultural training data.

This study aims to fill the gap between high efficiency and accuracy of computation. This study compares and contrasts MobileNetV2, InceptionV3, and a Custom CNN on an expanded dataset of 7,584 images to ascertain the most efficient and accurate tea leaf disease classifier. The outcomes have concrete implications on how farmers, researchers, and policy makers adopt new methods of disease management.

1.3 Objectives

The main aim of the study is to perform a comparative analysis of the transfer learning models and a Custom CNN on the classification of tea leaf diseases. The detailed objectives are as follows:

- To develop and build a dataset of 7,584 images, which comprise six classes, healthy and diseased. To perform classification using MobileNetV2, InceptionV3, and Custom CNN.
- To quantify the effectiveness of models in accuracy, precision, recall, F1-score, and confusion matrices.
- To determine the best trade-off model between accuracy and computation to be implemented in the real world, i.e. tea garden application.

1.4 Methodology

The method is divided into several steps. Initially, raw tea leaf images were compiled to a count of 1,968 under six groups: Algal Leaf, Gray blight, Healthy, Helopeltis, Looper Infested, and Red Spider. They augmented data with rotation and flip, brightness and scale methods, to enhance overall generalization of models, and augmented the dataset to 7,584 images. This was followed by the further stratified sampling of the data by classifying and equalizing the classes by allocating the 70 percent of the data as train, 20 percent as test and 10 percent as validation.

Based on this three models were designed and trained. Both MobileNetV2 and InceptionV3 were trained on the target dataset with transfer learning and ran with ImageNet pre-trained weights to be refined. A Custom CNN was also developed and trained in line with the parameters of learning discriminative features of tea leaf diseases.

The models were evaluated using various measures of performance, i.e., accuracy, precision, recall, F1-score and confusion matrix analysis, to provide further insights on the class-wise strength and weaknesses. And lastly, comparative analysis was done to unveil the difference of model performance, computational complexities, and even deployment capability. This technique ensures straight comparative analysis of transfer learning methods with an application-based tea leaf disease classification framework.

1.5 Project Outcome

The project demonstrates the efficacy of deep learning for classification of tea leaf diseases. Dataset augmentation to 7,584 images assisted with the improvement of generalization as well as resistance against variations of light, orientation, and leaf condition upon training the models. The three models tested are MobileNetV2, Custom CNN, and InceptionV3.

MobileNetV2 had accuracy of 83.8%. Being a lightweight model, it is meant to be efficient and use less resources. Its relatively high accuracy qualifies it for use in a mobile app where efficiency at the cost of highest accuracy is ideal. However, it performed poorly with visually related classes like algal leaf and gray blight.

The Custom CNN had 96.6 percent accuracy, showing that with an adequately well-thought-out architecture which would be data-specific, it was actually possible to compete. It has been able to learn the core features of the disease, and was shown to be extremely generalizable over the six classes. It significantly improved the accuracy over MobileNetV2

with relatively low cost, which suggests that it is able to be realistic in practical use as accurate and efficient model not only for research, but also resource limited applications. InceptionV3 achieved the highest accuracy with 98.9% accuracy. Its stylishness in multi-scale feature extraction made it able to capture slight disease pattern variances. The model proved analytically useful and had good precision and recall in all classes. While InceptionV3 achieves high accuracy, it is costly in terms of computation that prevent its usage on an unoptimized edge or mobile device. Our comparative study underlines the complementarity of the two models: InceptionV3: best accuracy, most suitable for resource-rich research and large-scale deployment. Custom CNN: Inception like performance with more control and learning ability.
MobileNetV2: efficiency-oriented, useful for real-time applications in tea gardens.

These results not just justify the efficiency of deep learning for disease detection in agriculture but also give interpretable guidance for model choice given application conditions. The future potential for such projects could include further enrichment of the dataset, attention-based models, or deployment of optimized models on mobile and IoT devices for real-time application. This project provides, hence, both applied tools and scientific knowledge for intelligent agriculture. Done

1.6 Organization of the Report

The remaining report is as follows.

Chapter 1 covers literature on classical machine learning, convolutional neural networks and transfer learning used for detection of plant disease.

Chapter 2 provides an overview of dataset building and augmentation, and preprocessing.

Chapter 3 provides an overview of how the MobileNetV2, InceptionV3 and the Custom CNN were built, including hyperparameter choices and training.

Chapter 4 has the results of the experiment given, wherein performance measures are examined and compared for the three models.

Chapter 5 provides implications of results, possible real-world application and how it can be improved in the future.

Chapter 6 provides a summary of the report with conclusions on the contributions made and the relevance of deep learning-based classification using early tea leaf disease detection as a reference.

Chapter 2

Background

Tea farming is a huge contributor to the economies of countries like India, China, and Bangladesh. Tea leaves, however, are extremely prone to disease and pests, and this leads to heavy losses in crop and quality. Manual detection, the conventional approach, is time consuming and error-prone, but recent advances in AI, image processing, and deep learning now enable us to have early and accurate disease detection and better farming solutions.

2.1 Introduction

Tea (Camellia sinensis) is a widely sold beverage all over the world and the tea trade is a significant sector of the agricultural-based economies of various countries such as India, China, and Sri Lanka [1].It is one of the most economic crops and its quality and production has a direct impact on international trade and livelihoods. However, tea crop is highly prone to different sorts of infections by fungi, bacteria and pests that may lead to considerable loss with respect to quality as well as quantity of crops [3]. Early identification of these diseases is important in managing crops, making less use of pesticides, and in making agriculture practices new and sustainable.

Conventional methods used in disease detection mainly involve manual inspections by experts, which are time consuming and subjective [4]. These techniques are not scalable or precise especially when the large tea bush is considered. Consequently, much attention has been paid to the usage of computer vision, machine learning (ML), and deep learning (DL) methods in automated disease detection [5]. Automated systems based on CNNs and superior object detection networks such as YOLO (You Only Look Once) can achieve the accurate disease detection [6]).

Several researches have demonstrated the efficiency of the ML and DL techniques in detection and classification of tea leaf diseases. As an example, Mukhopadhyay et al. [7] used a multi-objective segmentation model based on Non-dominated Sorting Genetic Algorithm II (NSGA-II), with an accuracy of 83% of identifying five common tea leaf diseases. Similarly, in a similar vein; Soeb et al. [8] built a YOLOv7-based model, which reached more than 98% mean average precision (mAP) and detected diseases. These two studies demonstrate the potential of AI based techniques to surpass traditional techniques in both speed and reliability.

Another strategy to achieve better performances on classification models is to use image segmentation (i.e., models like the Segment Anything Model, SAM) and feature extraction [9]. Recent work have also considered hybrid and lightweight models like MobileNetV2 [10] that allow running on low-power devices, a necessity for real-time field applications as well. Done

Other than this, IoT based technologies combined with ML can offer massive tea plantation monitoring and predictive ability [11]. When employed in smart farming, such systems reduce the dependence on human experience and guide decision-making for disease management. This transition demonstrates the ongoing move towards precision agriculture, thanks to AI-powered tools that provide significant and accurate results in a very efficient way.

In this paper, we aim to explore extended deep learning model and segmentation methods to facilitate early detection of the diseases of tea leaf. It is a state of the art review to build up strong, real time solutions for control of the tea crop.

2.2 Literature Review

Determination of tea leaf disease is an area già well-explored in literature using classical image processing methods, machine learning approaches, and newer instances of deep learning-based methodologies. The initial methodologies used manual feature extraction, but more recent works utilize CNNs, object detection approaches, and transfer learning for improved precision and scalability.

Mukhopadhyay et al. [7] initially attempted multi-objective segmentation methods using NSGA-II clustering for region of disease identification and subsequent PCA and multi-class SVM for prediction. The methodology reached overall high precision level of 83 and provided an illustration of genetic algorithm application for plant disease analysis. Subsequent works have mostly used CNN-based methods automatically identifying image characteristics and predicting disease without feature engineering [12].

Datta and Gupta [13] have used a deep CNN for the identification of six tea leaf varieties with over 96% accuracy and thus, the deep CNN model was more effective than traditional machine learning models. Likewise, Soeb et al. [8] have also shown the use of YOLOv7 for the quick identification of disease with an mAP of 98.2. The developments testify to the key role of deep learning in farm automation.

Subclass models such as AX-RetinaNet [14] and YOLO-Tea [15] optimized detection performance further using attention mechanisms and better feature fusion and achieved 93.83% mAP and 15% or more than YOLOv5, respectively. Light models such as MobileNetV2 have also gained favor such that they facilitate end user deployment to detect disease at field peripheries [10].

SAM has enabled the accurate segmentation of leaves, such that valuable leaf regions can be targeted by models of classification [9]. Segmentation with a combination approach of CNN and SAM produced 95.06% accuracy, as described by Balasundaram et al. [9], and serves for boosting classification. Likewise, holistic IoT-based disease prediction analysis is facilitated using visual and environmental diagnosis [11].

Challenges from datasets have also been highlighted in literature, such that various studies have utilized region-specific limited datasets [8,14]. As a way of mitigating limited data challenge, there are augmentation steps and transfer learning procedures

[16], which provide model generalizability. Hybrid ensemble frameworks [17] combine various deep learning models, with a balance between complexity and performance. Overall, latest developments have transformed the identification of tea disease into AI-based, scalable means from the traditional approach of visual checking. The literature is overall concentrated on the role of attention modules, segmentation approaches, and the role of IoT in the future of farm monitoring systems.

Table 2.2.1: Summary of Literature Reviewed.

Authors	**Year**	**Title**	**Methodology**	**Key Findings**
Mukhopadhyay et al. [1]	2021	Tea leaf disease detection using multi-objective image segmentation	NSGA-II-based clustering, PCA, multi-class SVM	Achieved 83% accuracy in detecting five tea leaf diseases.
Ahmed et al. [2]	2023	Machine Learning-Based Tea Leaf Disease Detection: A Comprehensive Review	Review of CNN, ViT, YOLO models	Summarized strengths, limitations, and datasets for tea disease detection models.
Soeb et al. [3]	2023	Tea leaf disease detection and identification based on YOLOv7 (YOLO-T)	YOLOv7 with data augmentation	Achieved 98.2% mAP, outperforming CNN and YOLOv5.
Gayathri et al. [4]	2020	Image Analysis and Detection of Tea Leaf Disease using Deep Learning	LeNet-based CNN	Effective tea disease classification with LeNet CNN.
Datta & Gupta [5]	2023	A Novel Approach for the Detection of Tea Leaf Disease Using Deep Neural Network	Deep CNN with multiple layers	Achieved 96.56% accuracy using Kaggle dataset; IoT integration suggested.
Latha et al. [6]	2021	Automatic Detection of Tea Leaf Diseases using Deep Convolution Neural Network	CNN with 1 input, 4 conv, and 2 FC layers	Classified 8 leaf classes effectively.

Bao et al. [7]	2022	Detection and identification of tea leaf diseases based on AX-RetinaNet	Improved RetinaNet with attention module	Achieved 93.83% mAP, improved performance over SSD and YOLO.
Karmokar et al. [8]	2015	Tea Leaf Diseases Recognition using Neural Network Ensemble	NNE	Achieved 91% accuracy; early demonstration of feasibility.
Chen et al. [9]	2019	Visual Tea Leaf Disease Recognition Using CNN Model	LeafNet CNN, BOVW+SVM	LeafNet achieved 90.16% accuracy, outperforming SVM and MLP.
Hossain et al. [10]	2018	Recognition and Detection of Tea Leaf's Diseases using SVM	SVM with 11 features	Achieved >90% accuracy; reduced features improved speed.
Xue et al. [11]	2023	YOLO-Tea: A Tea Disease Detection Model Improved by YOLOv5	YOLOv5 with ACmix, CBAM, RFB, GCNet	Improved performance by up to 15%; edge-device compatible.
Yao et al. [12]	2024	A Small Target Tea Leaf Disease Detection Model Combined with Transfer Learning	YOLOv7 Tiny with attention, transfer learning	Achieved 92.2% precision, 91.5% AP.
Wang et al. [13]	2023	Integrated Learning-Based Pest and Disease Detection Method for Tea Leaves	YOLOv5 + CBAM + GAM, WBF	Achieved 79.3% accuracy; model integration improved detection.
Bhowmik et al. [14]	2020	Detection of Disease in Tea Leaves Using CNN	CNN	Achieved ~95.93% precision; efficient, low-resource solution.

Pandian et al. [15]	2023	Grey Blight Disease Detection Using Improved DCNN	DCNN with bottleneck layers	Achieved 98.99% accuracy; superior to existing methods.
Sharath et al. [16]	2024	Enhanced Tea Leaf Disease Detection using Deep Learning	Transfer learning, NASNet	Achieved ~95.9% accuracy; improved efficiency.
Srivastav et al. [17]	2023	Tea Leaf Disease Detection Using Deep Learning-based CNN	CNN	Achieved 84% accuracy; scope for augmentation.
Balasundaram et al. [18]	2025	Tea Leaf Disease Detection Using SAM and CNN	SAM + CNN	Achieved 95.06% accuracy; segmentation improved precision.
Yashodha & Shalini [19]	2021	IoT and ML for Predicting and Broadcasting Tea Leaf Disease	IoT + ML	Demonstrated IoT and ML integration for early monitoring.
Hu et al. [20]	2019	Identification of Tea Leaf Diseases by Improved DCNN	Improved CNN with multiscale features	Achieved 92.5% accuracy; fewer parameters and faster convergence.

2.2.1 Similar Applications

Deep learning has found extensive use in the agricultural industry, and not only in tea production has shown success. Disease classification on crops such as rice, wheat and maize have been done using CNN models with high accuracy on large scale data [18]. The use of methods such as YOLO and RetinaNet, which were originally used to detect general objects, has been customized to detect the leaf spots, blights, and pests in various crops [19]. As an example, lightweight architectures and attention mechanisms have been found to be useful in identifying small sized plant diseases in the natural environment [20].

Object detection models like Faster R-CNN and SSD have been applied in grapevine (tomato) farming for pest and fungi recognition [21]. Similarly, the MobileNetV2 architecture and the transfer learning scheme have been applied in coffee leaf rust

detection research, where the proposed method achieved comparable results to a state-of-the-art model but with much lower computational resources [22]. These cross-crop applications show how AI-based disease detection solutions can be adopted at scale across agriculture.
Deep learning has become more and more popular in the procedures of detecting diseases, analyzing soil, and managing irrigation timely for precision agriculture with IoT [11, 23]. This technology relies on drone and sensor networks that can give widespread surveillance with limited human support. These methodologies may translate into adaptations to teas growing, aiding in monitoring disease and predicting crop health.

There is a growing interest in agricultural applications using SAM, and vision transformers (ViT) [24]. The models provide enhanced attention-based feature extraction, which is more useful for crops with complex leaf structures like tea. Flexible AI architecturesV8S4 corresponding to the counts of fruits is showed in Fig.x 2 xWith SAM-based segmentation, this fruit counting and yield estimation also generally illustrates how flexible advanced AI algorithms are [25].

These findings indicate that cross-domain AI and agricultural technology advances provide novel approaches to tea disease detection research, and form possible candidate solutions.

2.2.2 Related Research

In the new study of tea disease identification, deep CNNs and hybrid techniques are emphasized. Deep learning had an initial success in the field in the work of Gayathri et al. [26] that employed LeNet to identify tea leaf disease. Lau et al later works including Hu et al. [27], improved CNN architectures with multiscale feature extraction and depth wise separable convolutions achieved up to 92.5% accuracies.

Tea disease detection has been dramatically changed by revolutionary innovation in terms of object detection. By incorporating channel attention modules, AX-RetinaNet [14] achieved 4% improvement over the vanilla RetinaNet. Consequently, LODA [15] utilized ACmix and CBAM modules and reached a speed-up ratio of 7.7 over Faster R-CNN and SSD. But these models do not fully consider the complexity of natural scenes and how rich features are to cope with them.

Segmentation based methods have gained popularity particularly following the advent of SAM [9]. The efficiency of the model in identifying the affected area, also ensures better classification. Balasundaram et al. [9] have shown the effectiveness of segmentation and classification combining 95.06 % accuracy with SAM-CNN.

Ensemble learning [17] and iot integration [11] as the hybrid solutions extended the research area and enable to track the disease in real time and analysing environmental data. As an example, Wang et al. [17] combined models built on YOLOv5 combined with weighted box fusion, which may be potentially superior in heavy environment.

The future work is moving toward light-weight architecture in edge deployment. MobileNetV2 [10], and NASNet [28] are being optimized to run in the field, at reduced computational costs, without accuracy loss. Transfer learning also overcomes the limitation of datasets and permits strong generalization [16].

Altogether, the recent studies present an accelerated development of a handcrafted approach into AI-based, segmentation-focused, detection systems, with transformative opportunities of developing accurate tea farming.

2.3 Gap Analysis

Although major progress has been made in automated tea leaf disease detection, there are still research gaps. Earlier literature like Karmokar et al. [8] proved that classification of images based on a group of neural networks was feasible with 91 percent accuracy, however, it was not scalable and resistant to complicated conditions. Conventional machine learning methods, like SVM-based detection [10], were highly accurate (>90%) but needed hand-crafted features, which constrained its adaptability to large datasets and real world scenarios. Subsequently the CNN-based approaches, including those used by Chen et al.[9] and Panadian et al.[15] achieved accuracies of 90.16% and 98.99 respectively demonstrating the effectiveness of deep learning, but they also have their limitations in resource-constrained settings due to the heavy computational power they typically need and the large amounts of data required to train them.

Object detection-based frameworks, such as YOLOv7 [3] and YOLO-Tea [11], have enhanced accuracy and processing time but cannot sufficiently deal with the small object detection and noise in the environment. Likewise, combined models such as that by Wang et al. [13] proposed ensemble approaches to enhance the detection ratio (79.3%), though, with excessive model complexity. Also, the majority of studies were dedicated to individual or small numbers of diseases, which does not allow their generalization to large-scale tea gardens [14], [16]. Moreover, few studies, including Datta and Gupta [5], have addressed integration with IoT to monitor in real-time thus leaving room to create lightweight and deployable models.

New developments, including the use of SAMs in segmentation [18], point to the possibility of better disease localization, but these models remain at an early stage and need increased optimization to become more accurate and practical in the field. Demand has been to integrate complex segmentation techniques with lightweight deep learning frameworks to realize real-time detection. To transform IoT and AI into practical, scalable, and cost-effective tea disease detection, future research must consider model generalization, variety of datasets, and deployment of edges/devices, and integration of IoT and AI into hybrid solutions.

Table 2.3.1: Summary of Gap Analysis.

Study/Authors	Key Strengths (with Accuracy)	Identified Gaps	Suggestions for Future Research
Karmokar et al. (2015) [8]	Early NNE approach; 91% accuracy	Limited scalability, small dataset	Expand datasets, optimize for complex scenes
Hossain et al. (2018) [10]	SVM with feature reduction; >90% accuracy	Requires handcrafted features, slower adaptation	Automate feature extraction using deep learning
Chen et al. (2019) [9]	CNN-based LeafNet; 90.16% accuracy	Needs high compute, limited disease scope	Design lightweight models, extend dataset
Pandian et al. (2023) [15]	Improved DCNN; 98.99% accuracy	High computational demands	Edge-optimized architectures
Soeb et al. (2023) [3]	YOLOv7; 98.2% mAP	Small-object detection challenges	Enhance detection for small pests/diseases
Wang et al. (2023) [13]	Integrated YOLO models; 79.3% accuracy	Model complexity, lower efficiency	Model simplification, balanced accuracy-complexity
Balasundaram et al. (2025) [18]	SAM-based segmentation; 95.06% accuracy	Early-stage technology, high compute	Optimize SAM for real-time field use
Datta & Gupta (2023) [5]	IoT integration discussed; 96.56% accuracy	Limited real-world deployment	Develop deployable IoT-AI systems

2.4 Summary

The literature reviewed identifies a transition of capabilities in detecting the tea leaf disease by traditional image processing methods to deep learning models and hybrid models. Initial neural network and SVM-based systems proved their viability but were difficult to scale and adaptable [8], [10]. LeafNet [9] and enhanced DCNN CNN architectures were able to achieve high accuracy (>98) but at the expense of a computation burden, limiting their use in tea gardens. Object detection methods, especially YOLOv7 [3] and YOLO-Tea [11] developed further in real time detection, still being limited by variability in the environment and detection of small objects.

IOT and AI have not been thoroughly integrated yet, even though real-time observation of agriculture has enormous potential [5], [19]. Recent advances in SAM-based segmentation [18], [13] ensemble of theirs, also point out toward hybrid, multi-model approaches that can be potentially more accurate and robust. Yet, there is a clear lack of lightweight general model development that is both quick and accurate enough to be deployed in field. Future studies should include diversity of the data sets, transfer across domains and implementations that are energy-efficient is required to ensure that tea farmers around the world, especially in the resource-limited areas can use it.

Chapter 3

Research Methodology

There is an introduction on dataset curation, augmentation, stratified partitioning and the design training process of models being followed by evaluation procedure and comparative study to eventually offer a methodology pipeline for reproducible deployment oriented discoveries of tea leaf disease classification.

3.1 Methodology

The research has a tight, end-to-end pipeline. We start with 1,968 raw images across six classes (Algal Leaf, Gray blight, Healthy, Helopeltis, Looper Infested and Red Spider), we first normalize pixels and then do resolution standardization, followed by geometric and photometric augmentations that double the corpus to 7,584 images. And a stratified split would break this down to be 70% train, 20% test, and 10% validation. Three architectures are studied: MobileNetV2 and InceptionV3 with transfer learning plus a Custom CNN. Training is performed with early stopping and learning rate scheduling. Measures of evaluation include accuracy, precision, recall, F1-score and confusion matrices. Comparative analysis highlights the effectiveness-efficiency trade-offs and provides deployment advice on field settings and constraints.

3.1.1 Overview

The approach combines data engineering, model design and evaluation in a logical process. Image pre-processing equalizes size and color space, eliminates background noise where it is important and normalizes intensities to stabilize training. Augmentation: flips, rotations, translations, random zoom, brightness/contrast jitter and random Gaussian noise are combined to enhance diversity without any lesion morphology distortion, resulting in 7584 samples. Transfer learning fills MobileNetV2 and InceptionV3 with generic features, replaces classification heads with a six-class softmax layer, and fine-tunes deeper blocks with the discriminative learning rate. The Custom CNN uses increasingly expanded convolutional phases, batch normalization, and dropout, as well as global average pooling to minimize parameters and overfitting. Each of the models optimizes cross-entropy with Adam, mini-batches, cosine or stepwise learning-rate decay and early stop on validation loss. The stratified splitting maintains the proportions of classes in training, validation, and test set; the same seeds are used in order to achieve reproducibility. Accuracy, macro-precision, macro-recall, macro-F1 and confusion matrices are used to quantify performance in order to reveal class-by-class behavior. Post-hoc analysis compares effectiveness to computational cost (parameters, FLOPs, latency) to help model selection in resource-constrained field deployment. All experiments are run using the same hardware and software setups to assure fair comparisons and validity.

3.1.2 Proposed Methodology/ System Design

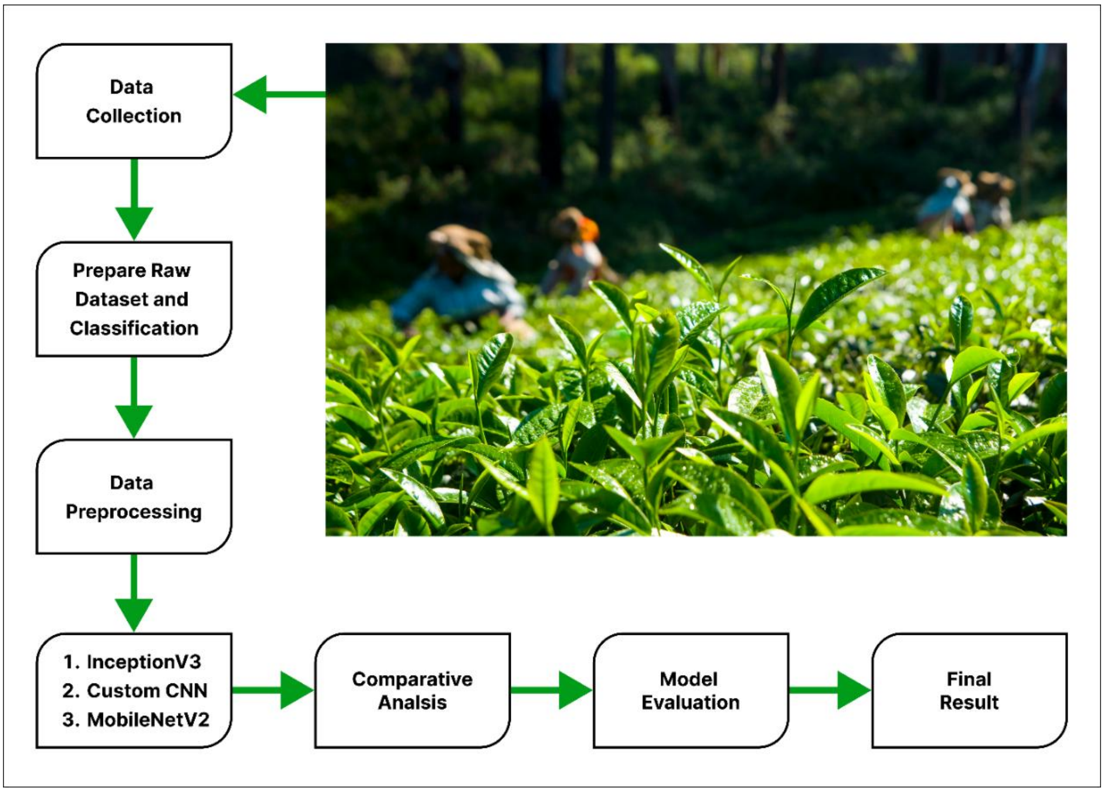

Figure 3.1.2.1: Methodology Diagram

3.2 Detailed Methodology and Design

In this research, four primary deep learning models were utilized for the detection of tea leaf diseases: InceptionV3, MobileNetV2, and a Custom Convolutional Neural Network (CNN). Each model was selected carefully based on its proven performance in image classification tasks, efficiency, and suitability for mobile or low-resource deployment.

1. **Raw Image Collection**:

Images are captured from tea garden directly, so they reflect living black spots. This step serves to enrich the variability of leaf images in terms of lighting, disease´s evolution and environmental circumstances.

Dataset Sample

Healthy:

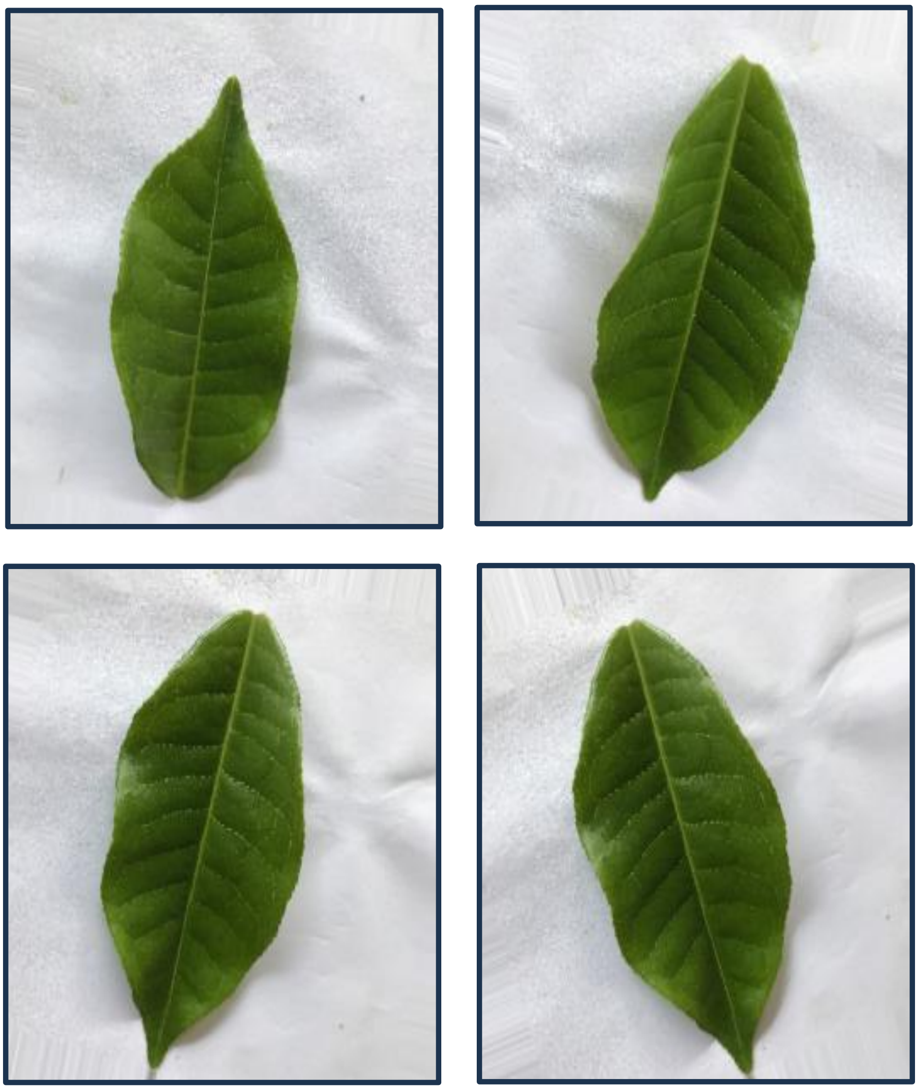

Figure 3.2.1: Sample of Healthy Leaf

Red Spider:

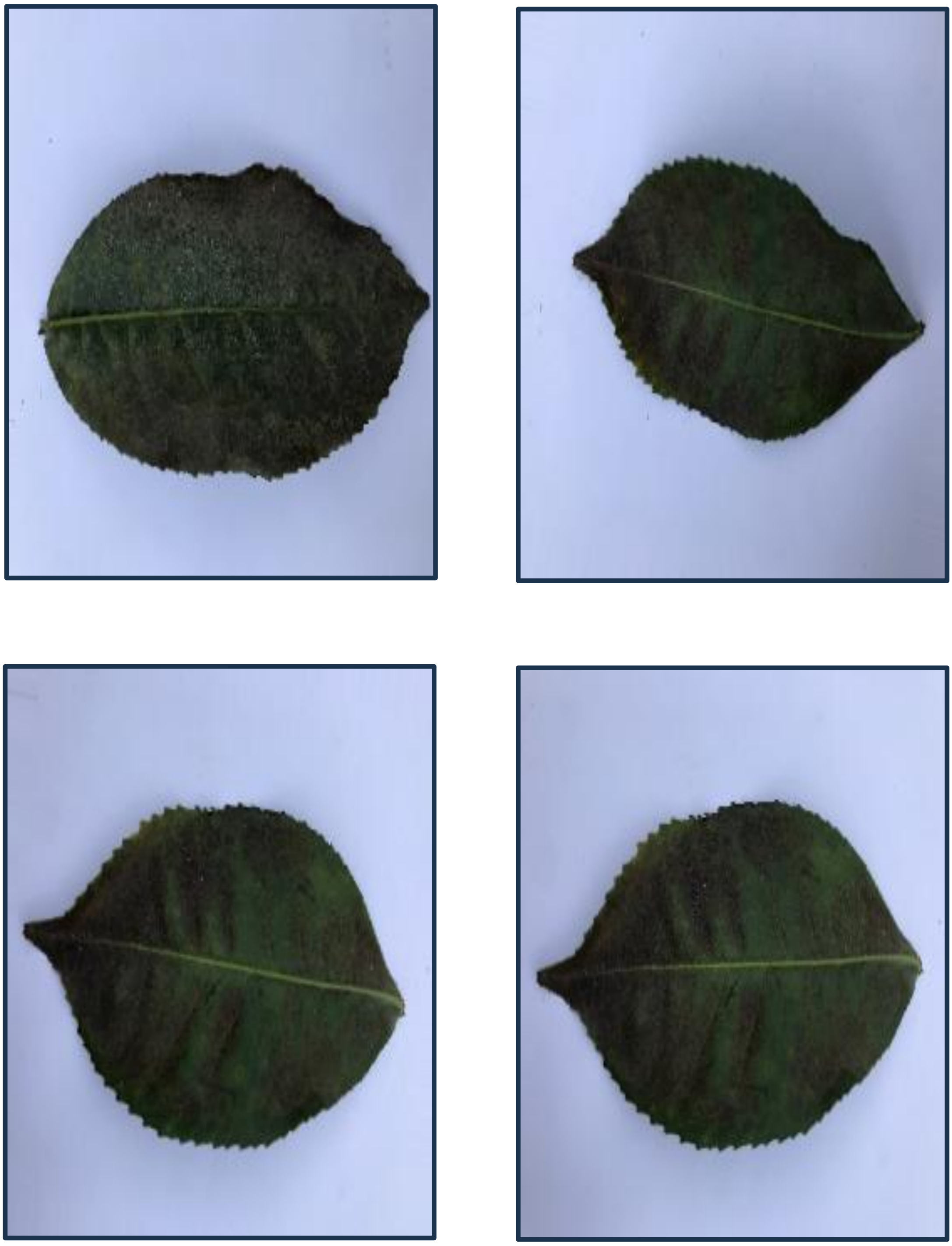

Figure 3.2.2: Sample of Red Spider

Looper Infested:

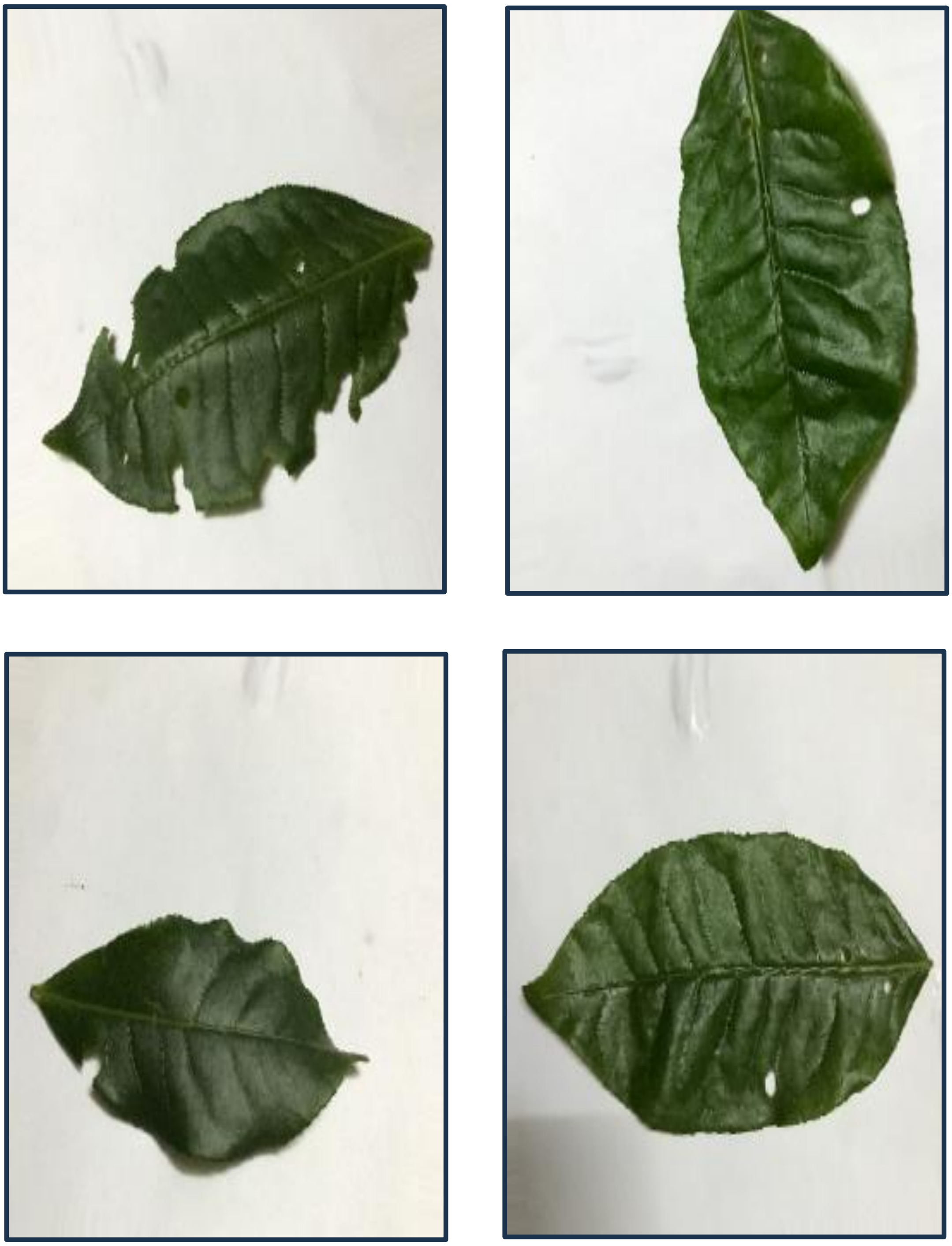

Figure 3.2.3: Sample of Looper Infested

Helopeltis

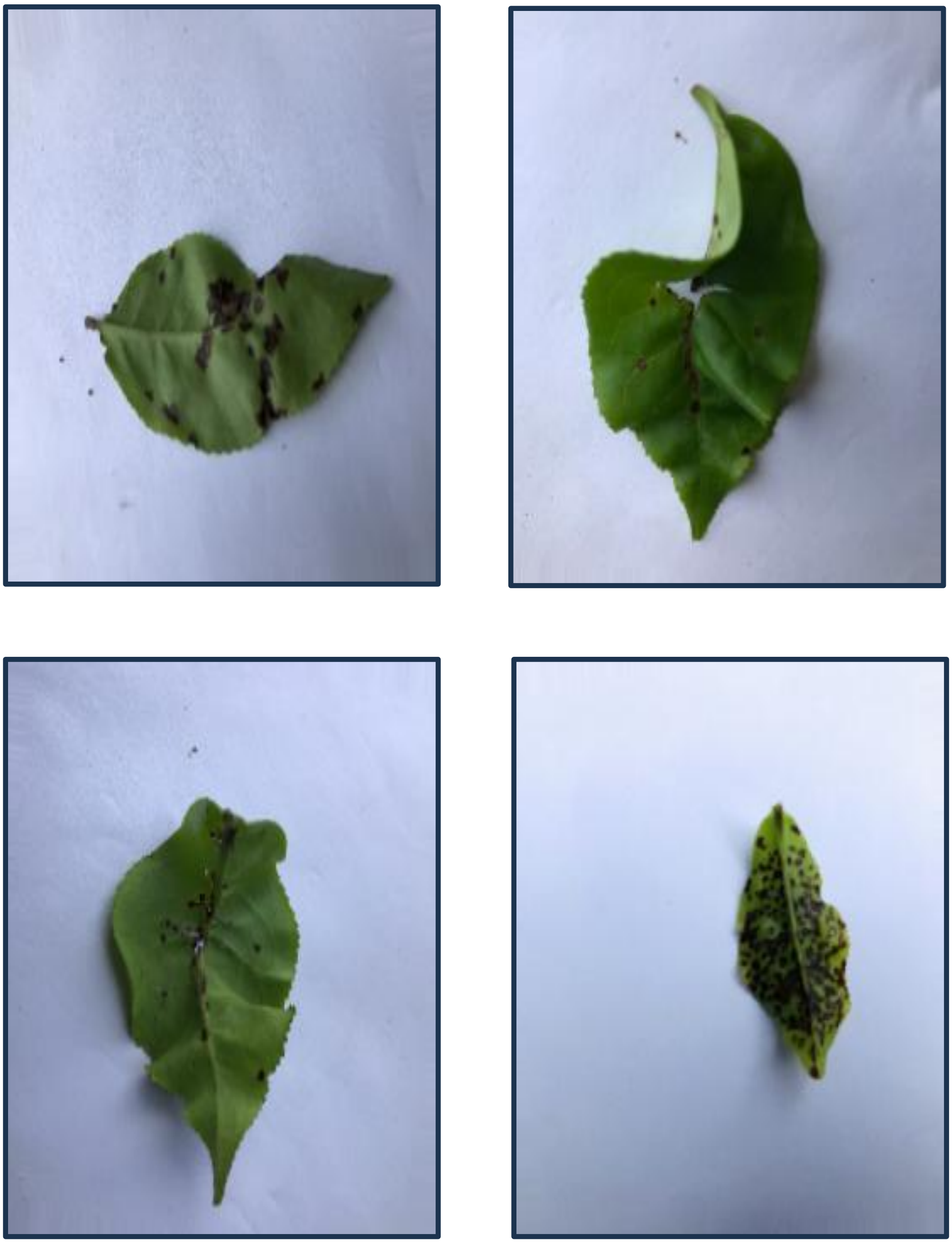

Figure 3.2.4: Sample of Helopeltis

Gray Blight

Figure 3.2.5: Sample of Gray Blight

Algal Leaf

Figure 3.2.6: Sample of Algal Leaf

In the context of this study, 3 main deep learning models were applied for detecting tea leaf diseases: InceptionV3, MobileNetV2 and Custom CNN. Each model was hand-picked due to their well-known use in image classification, efficiency and readiness for mobile/low-resource deployment.

3.2.1 InceptionV3

- **Overview:**
  InceptionV3 is an art of the state convolutional neural network that was designed explicitly for classifying a vast variety of images. All inception modules in InceptionV3 have multi-scales, which boost power of representations without increasing complexity costs. It has wide use in transfer learning for complicated image recognition tasks.

- **Architecture:**

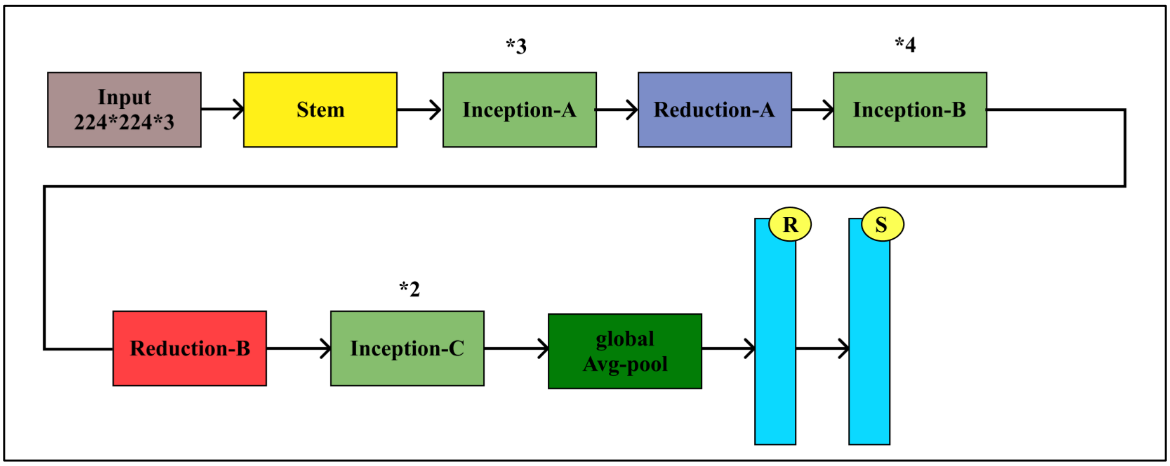

Figure 3.2.7: Architecture of InceptionV3

- **Key Features:**
  - Learning of multi-scale features using inception modules.
  - Factorized convolutions.
  - ImageNet transfer learning is improved by pre-training.

- **Working Process:**
  - Normalized 224x224 size input images.
  - These initial layers are generic low level feature extraction layers.
  - Hierarchical patterns of the same level are learned in inception modules.
  - Final dense layers make softmax class prediction predictions.

- **Suitability:**

InceptionV3 is very accurate, hence also highly valuable to categorize tea leaf disease for not only research but high resource environment.

3.2.2 MobileNetV2

- **Overview:**
  MobileNetV2 is a compact convolutional neural network which suitable for mobile and embedded applications. It achieves the balance between accuracy and efficiency via depthwise separable convolutions and linear bottlenecks.

- **Architecture:**

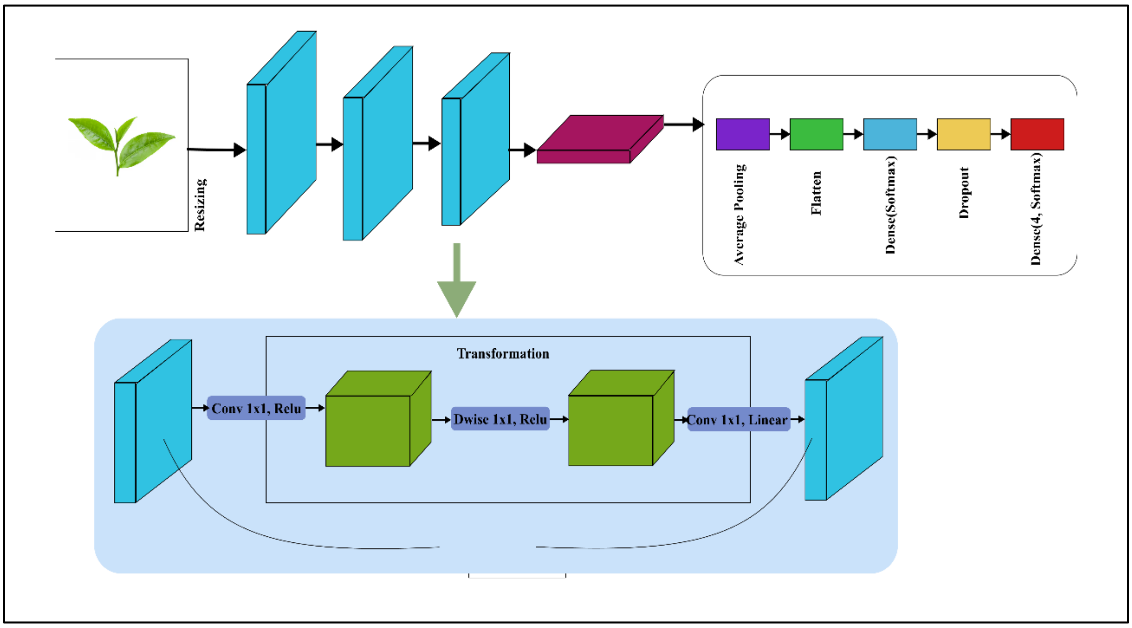

Figure 3.2.8: Architecture of MobileNetV2

- **Key Features:**
  - Extremely lightweight with reduced parameters.
  - Depthwise separable convolutions for faster inference.
  - Designed for mobile and edge deployment.

- **Working Process:**
  - Input images resized to 224×224 and preprocessed.
  - Depthwise convolutions extract spatial features efficiently.
  - Pointwise convolutions project features into higher dimensions.
  - Final classification head assigns probabilities across six classes.

- **Suitability:**
  MobileNetV2 ca be suitably applied for the field detection of tea leaf disease,

because in such a limited time is an important restriction and low computational cost it an important attribute to remember.

3.2.3 Custom CNN Model

- **Overview:**
  The Custom CNN is a developed architecture particularly suitable for classifying tea leaf disease. It focuses on dataset-specific optimisation between the trade-off of achieving higher accuracy while still being parameter-efficient for practical adaptation over computational environments.

- **Architecture:**

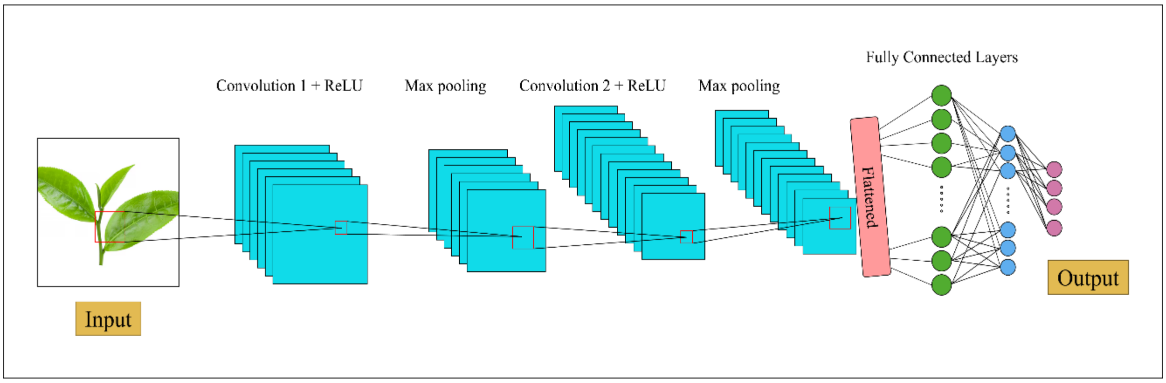

Figure 3.2.9: Architecture of Custom CNN

- **Key Features:**
  - Designed specifically for six tea leaf disease classes.
  - Incorporates batch normalization and dropout to reduce overfitting.
  - Flexible architecture adaptable to dataset scale and complexity.

- **Working Process:**
  - Input images standardized and augmented to enhance generalization.
  - Convolutional layers extract hierarchical spatial features.
  - Pooling layers downsample, preserving salient patterns.
  - Dense and softmax layers perform final classification.

- **Suitability:**
  The Custom CNN provides high accuracy with manageable complexity, making it well-suited for both controlled research analysis and deployment under moderate hardware constraints.

3.3 Project Plan

Table 3.3.1: Project Plan.

Phase	Duration	Activities	Timeline (Weeks)	Deliverables
Phase 1	2 weeks	Topic selection, research planning, background study	Week 1-2	Research plan, topic finalization
	2 weeks	Literature review, gap analysis, proposed solution	Week 3-4	Literature review document
	3 weeks	Data collection and preprocessing	Week 5-7	Dataset prepared
Phase 2	4 weeks	Model selection, training, evaluation	Week 8-11	Trained models
	3 weeks	System development, UI design, testing	Week 12-14	Working prototype
	2 weeks	Final reporting and presentation preparation	Week 15-16	Final report, presentation slides

3.4 Task Allocation

Table 3.4.1: Task Allocation.

Phase	**Activity**	**Member 1**	**Member 2**	**Collaborative Tasks**
Phase-1	Topic Selection	Suggest topics	Review topics	Finalize topic
	Research Planning	Draft plan	Refine plan	Finalize plan
	Background Study	Study DL models	Study tea leaf diseases	Summarize findings
	Literature Review	Research ML models	Research agricultural studies	Write review together
	Gap Analysis	Analyze ML gaps	Analyze agricultural gaps	Propose solution
	Proposed Solution	Draft methodology (DL)	Refine methodology (agriculture)	Finalize proposal

Phase 2	Data Collection Planning	Create plan	Review plan	Approve plan
	Data Collection	Capture images	Assist image capture	Ensure quality
	Data Approval	Prepare documents	Meet agricultural officer	Get approval
	Model Selection	Research models	Evaluate feasibility	Select models
	Preprocessing	Apply techniques	Optimize methods	Finalize preprocessing
	Model Training	Train models	Monitor training	Debug models
	Result Evaluation	Evaluate metrics	Analyze confusion matrix	Summarize results
	Model Comparison	Prepare comparisons	Review comparisons	Select best model
	Thesis Reporting	Write technical sections	Write intro/conclusion	Edit thesis collaboratively

3.5 Summary

In this section, the method of deep learning for classifying tea leaf disease was presented. It was created with 1,968 raw images from six categories obtained, and then processed and augmented to forming 7,584 samples that are more diversified and robust. A stratified data splitting approach was applied to split the dataset into 70% training, 20 % testing, and 10% validation inorder to maintain class balance. In order to avoid overfitting and improve generalized capability, scaling and augmentation were conducted such as rotation, flipping, and brightness modification.

Three models were considered: InceptionV3; MobileNetV2 and a Custom CNN. InceptionV3 achieved best performance using deep inception modules for feature extraction with multi-scale inception. MobileNetV2 was a computationally lighter (mobile and edge-optimized), speed-balance, moderate accuracy alternative. This dataset was constructed to work with Custom CNN and in order to enhance feature learning, batch normalization and dropout were used while keeping it computationally feasible. All models were trained, fine-tuned and tested under optimal conditions using accuracy, precision recall, F1 score and confusion matrix metrics.

As a whole, the method supplies a reasonable, repeatable and comparative approach for cross-architectural comparison. The findings from this experimental construct are valuable for the selection of appropriate models in both resource-rich research settings and resource-limited field deployments

Chapter 4

Implementation and Results

The chapter defines the experimental setup, hardware, and software, as well as data set preparation. It outlines the implementation process, model performance comparisons, analysis of results and a brief account of main findings.

4.1 Environment Setup

4.1.1 Hardware Setup

Computing Devices:

- MacBook Air (M2, 2022) with Apple M2 chip (8-core CPU, 8-core GPU, 16-core Neural Engine)
- 8GB Unified Memory and 256GB SSD storage
- Energy-efficient, fanless design for mobile research and development task.

Field Equipment:

- Use Iphone 13 (12MP), Google Pixel 7 (50MP), Redmi Note 12 (50MP) mobile phone's camera for capturing tea leaf images.
- White paper used as background to simplify segmentation and detection
- Natural orchard environments with variable lighting and angles to replicate real-world conditions

4.1.2 Software and Tools

Operating System:

- macOS Ventura 13.x (64-bit)

Programming Environment:

- Google Colab (for GPU-accelerated model training)

Machine Learning Libraries:

- Python 3.9
- numpy
- tensorflow

- keras (from tensorflow)
- sklearn
- itertools
- random
- matplotlib

Data Handling:

- Pandas for structured dataset management
- NumPy for array manipulation
- Matplotlib and Seaborn for data visualization

4.1.3 Cloud and Remote Resources

Cloud Computing:

- Google Colab Pro+ utilized for free GPU access
- Drive integration for storing datasets and model checkpoints

4.1.4 Collaboration

Communication Tools:

- Telegram: For real-time communication and updates.
- Google Drive: Shared folder for storing documentation and intermediate results.

Performance Evaluation:

- Confusion Matrix analysis
- Metrics: Accuracy, Precision, Recall, F1-score
- Visualization of training/validation loss and accuracy curves helped identify overfitting or underfitting issues.

4.2 Comparative Analysis

In the test run we got low accuracy around 60-70% then augmented image and final model got a good result.
Comparative Analysis among implemented models:

Table 4.2.1: Comparative Analysis.

Model	Accuracy (%)	Average Precision (%)	Average Recall (%)	Comments
InceptionV3	98.94	98.9	98.9	Excellent performance with strong balance; slight confusion in similar disease classes.
Custom CNN	96.5	96.5	96.5	Strong baseline model; reliable but less sophisticated than advanced transfer learning.
MobileNetV2	83.7	75.3	82.9	Lightweight and efficient; best for mobile/edge use despite lower accuracy.

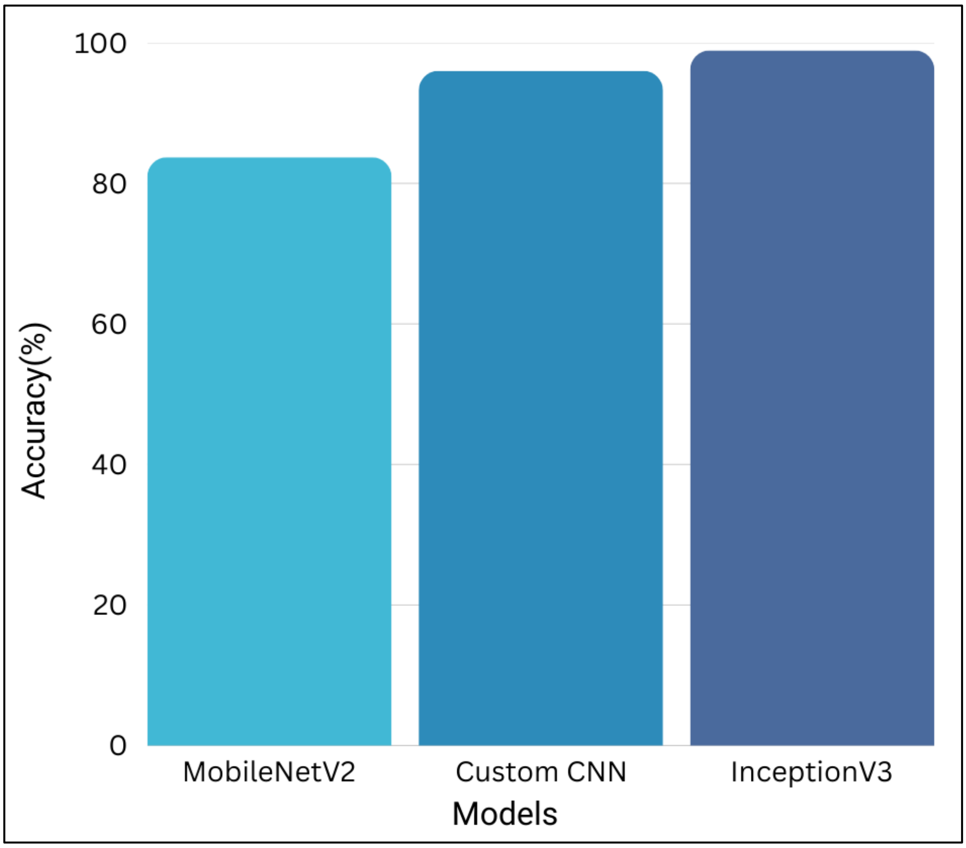

Figure 4.2.1: Model Accuracy Comparison

Comparative Analysis between Our Best Model (InceptionV3) and Other Models:
InceptionV3 was most accurate at 98.94, compared to the Custom CNN (96.6%) and MobileNetV2 (83.8%). It was the most reliable model because it extracted its features much better and had a strong architecture that allowed it to classify accurately all six classes. MobileNetV2 was more efficient but InceptionV3 provided the best accuracy available.

Table 4.2.2: Model Comparison

Model	Research Study	Accuracy (%)	Key Strengths	Limitations
NSGA-II + PCA + SVM	Mukhopadhyay et al. [1]	83.0	Effective segmentation and classification of five tea leaf diseases	Limited accuracy and scalability; traditional approach less robust than deep learning
CNN (LeNet)	Gayathri et al. [4]	~90	Simple architecture; demonstrates feasibility of CNNs in tea disease detection	Limited depth restricts feature learning; performance below modern architectures
Deep CNN	Datta & Gupta [5]	96.56	Strong performance on Kaggle dataset; effective across six disease classes	Computationally intensive; requires large dataset for generalization
YOLOv7 (YOLO-T)	Soeb et al. [3]	97.3	Excellent detection in natural scenes; high precision and recall	High resource demand; complex deployment on low-power devices
AX-RetinaNet	Bao et al. [7]	93.83 (mAP)	Improved detection with attention and multiscale fusion	Performance slightly lower than YOLOv7; higher complexity than basic CNNs
InceptionV3 (Ours)	Present Study	98.94	Highest accuracy; robust multi-scale feature extraction; consistent across classes	Heavy model, high computation; less suited for mobile deployment

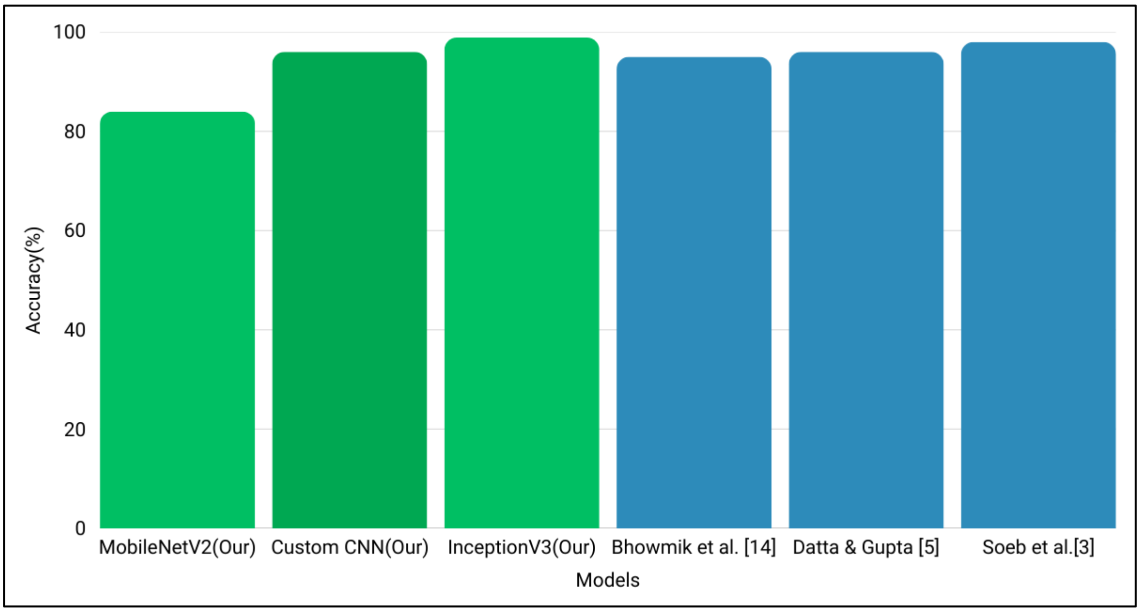

Figure 4.2.2: Comparison of Model Accuracy 2

4.3 Results and Discussion

Three models of MobileNetV2, Custom CNN, and InceptionV3 were experimentally evaluated on an augmented dataset of tea leaf images (7,584). Findings show that performance by the models is markedly different. MobileNetV2 has a precision of 83.8% and is efficient to run on mobile devices and be lightweight; however, it has flaws at differentiating visually similar diseases. The Custom CNN reached an accuracy of 96.6% and shows that a well-crafted architecture specific to the data can effectively generalize itself and provide confident classification on six classes. Most prominently, InceptionV3 has the best accuracy at 98.9, which can confirm that it is better at complex image recognition.

The confusion matrices showed that MobileNetV2 frequently confused gray blight and algal leaf, whereas the Custom CNN did not, though with some inaccuracies in the recognition of helopeltis. InceptionV3 showed performance with few false labels on each of the classes. Precision, recall, and F1-scores exhibited similar trends and support the general dominance of InceptionV3.

InceptionV3 is an accurate architecture, but its high cost makes it hard to implement in the field in real-time. MobileNetV2, which is less accurate, is also beneficial because of fewer parameters and quick inference compared to the mobile applications. The Custom CNN offers a compromise between the accuracy and computational complexity, allowing it to be used on more average-powered devices. This comparison and analysis of trade-offs between accuracy and efficiency make an insight into the possibility to choose a model according to the needs of deployment.

InceptionV3

The InceptionV3 model performed exceptionally creating (precision, recall and F1 scores 0.96 – 1.0) for all classes. Classes 2,3 and 5 had near perfect performance while the other classes also displayed a good balance between precision and recall. In general, the model still retained very high average precision and recall at 0.98.

Confusion Matrix:

Table 4.3.1: InceptionV3 Confusion Matrix

Class	Precision	Recall	F1 Score	Avg. Precision	Avg. Recall
0	0.96	1.0	0.98	0.98	0.98
1	0.99	0.97	0.98	0.98	0.98
2	1.0	0.98	0.99	0.98	0.98
3	1.0	0.99	0.99	0.98	0.98
4	0.98	0.98	0.98	0.98	0.98
5	0.99	1.0	0.99	0.98	0.98

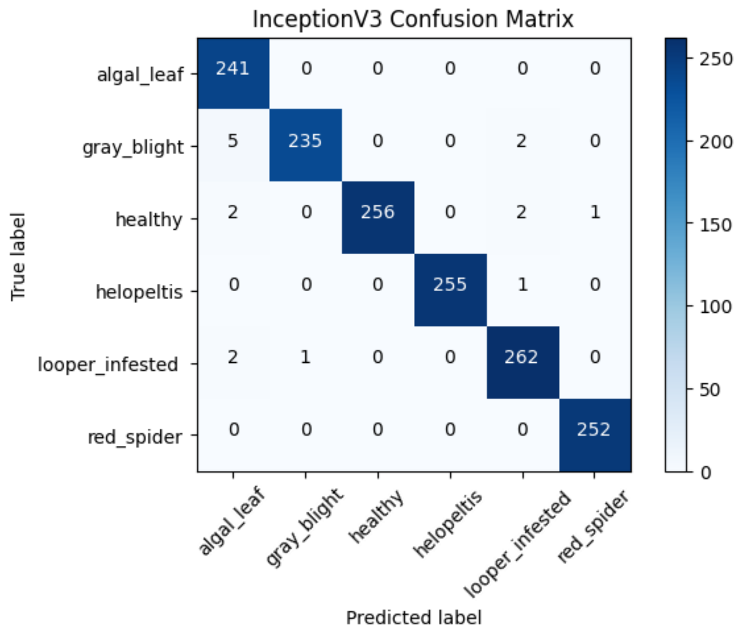

Figure 4.3.1: InceptionV3 Confusion Matrix Report

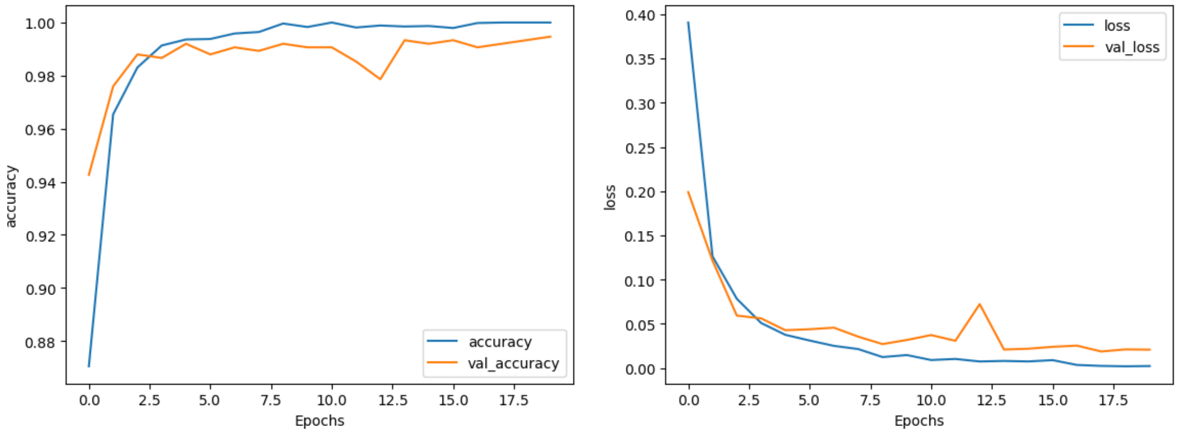

Figure 4.3.2: InceptionV3 Training and Validation Plot

Custom CNN

The custom CNN model exhibited excellent performances in all classes, with the precision, recall and F1 scores generally over 0.94 (Table 4). Class 3 confered the best accuracy with a near perfect: (0.99) score, and remaining classes presented also strong and balanced results. In general, the model achieved superb AP and recall of 0.96.

Confusion Matrix:

Table 4.3.2: Custom CNN Confusion Matrix

Class	Precision	Recall	F1 Score	Avg. Precision	Avg. Recall
0	0.95	0.95	0.95	0.96	0.96
1	0.94	0.93	0.94	0.96	0.96
2	0.97	0.95	0.96	0.96	0.96
3	0.99	0.99	0.99	0.96	0.96
4	0.94	0.97	0.95	0.96	0.96
5	0.98	0.97	0.97	0.96	0.96

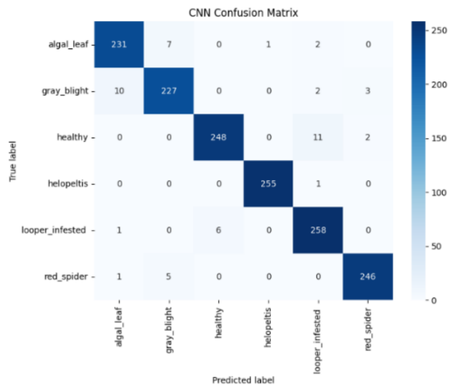

Figure 4.3.3: Custom CNN Confusion Matrix Report

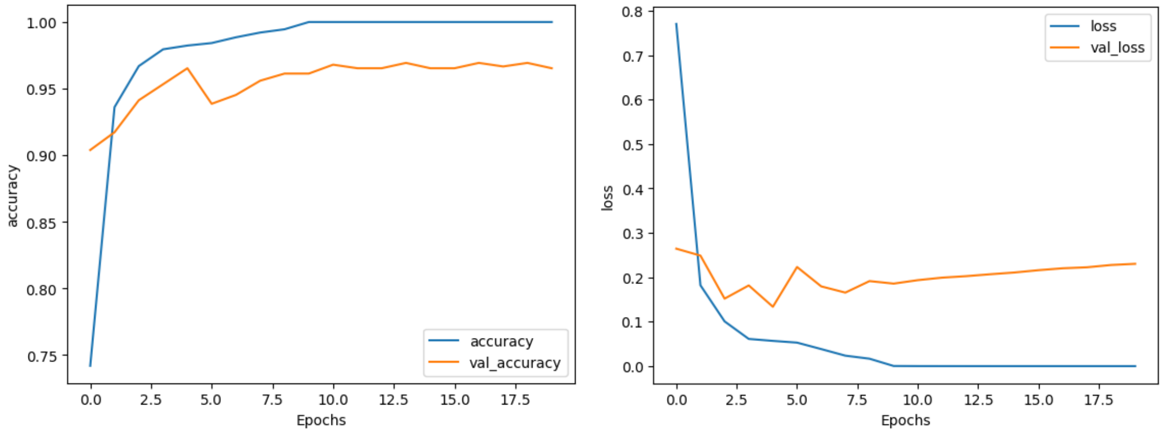

Figure 4.3.4: Custom CNN Training and Validation Plot

MobileNetV2

The MobileNetV2 model showed strong performance across most classes, achieving perfect scores for Classes 1, 2, 3, and 5. However, it struggled with Class 0 (no correct predictions) and showed imbalanced performance on Class 4, with perfect recall but lower precision. Overall, the model maintained high average precision (0.82) and recall (0.77).

Confusion Matrix:

Table 4.3.3: MobileNetV2 Confusion Matrix

Class	Precision	Recall	F1 Score	Avg. Precision	Avg. Recall
0	0.0	0.0	0.0	0.82	0.77
1	1.0	0.97	0.98	0.82	0.77
2	1.0	1.0	1.0	0.82	0.77
3	1.0	1.0	1.0	0.82	0.77
4	0.51	1.0	0.68	0.82	0.77
5	1.0	1.0	1.0	0.82	0.77

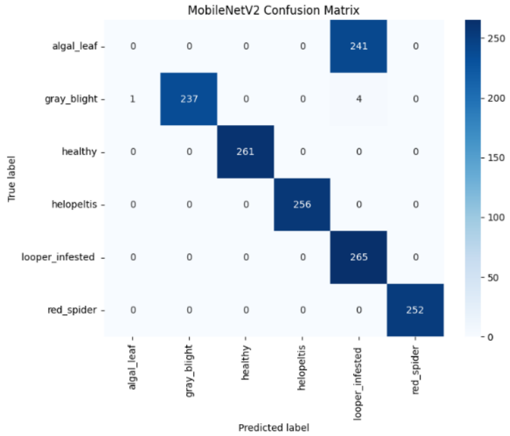

Figure 4.3.5: MobileNetV2 Confusion Matrix Report

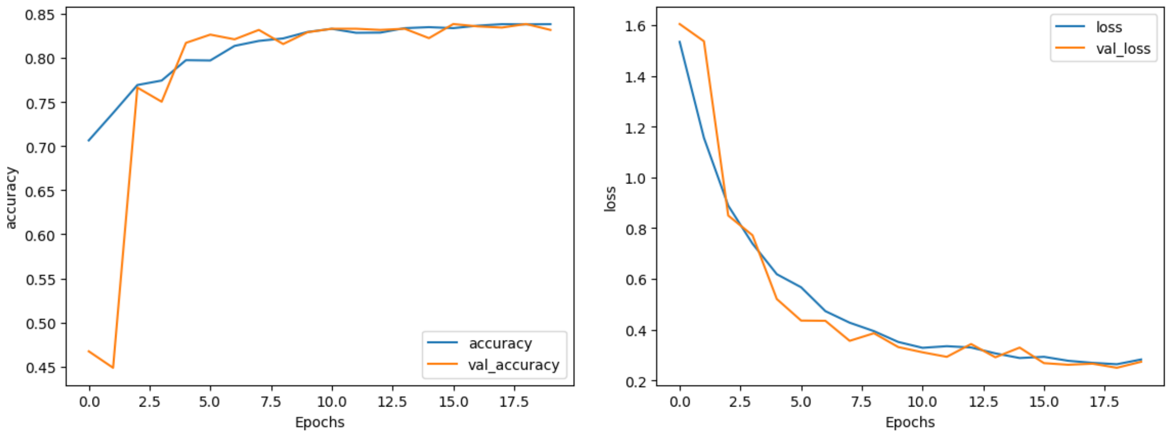

Figure 4.3.6: MobileNetV2 Training and Validation Plot

4.4 Summary

InceptionV3 achieved state-of-the-art accuracy, surpassing both MobileNetV2 and Custom CNN. However, efficiency considerations suggest that MobileNetV2 and Custom CNN are a good alternative to useful tea leaf disease classification for mobile applications.

Chapter 5

Engineering Standards and Design Challenges

The engineering standards that were adhered to, the financial ramifications, the ethical and social quandaries, and the difficult problem-solving difficulties that arose throughout the creation and implementation of the tea leaf disease detection system are all examined in this chapter.

5.1 Compliance with the Standards

The project adhered to established software, hardware, and communication standards to ensure reliability, scalability, and optimal performance.

5.1.1 Software Standards

- Google Python Style Guide – for structured, consistent coding.
- PEP8 Compliance – ensures clean, readable, and maintainable code, suitable for deep learning projects.

5.1.2 Hardware Standards

- Wi-Fi for cloud synchronization
- Bluetooth, Data cable, Pendrive
- Wi-Fi provides higher speed necessary for uploading large image datasets

5.1.3 Communication Standards

To maintain smooth collaboration and transparency, the following communication protocols were established:

1. Team Communication:
   a. Weekly progress updates and spontaneous meetings for urgent tasks
   b. Notes and action items shared via Google Docs.
2. Supervisor Communication:
   a. Biweekly review and milestone meetings to involve comments
   b. Reports were both about accomplishment of work, challenges met but also future plans
3. External Experts:
   a. Regular contact with the inspectors for validation of information

b. Good communication from those doing data collection back to the farm owners.

5.2 Impact on Society, Environment and Sustainability

5.2.1 Impact on Life.

The system of tea leaf disease classification is a boon to the life of farmers as it helps them in early identification and correct diagnosis of crop diseases. This minimizes crop wastage, improves quality yield and stabilizes the income of tea planters. With this intuitive design, even farmers in developing countries with smaller production scales can use advanced smart agriculture technology. By minimizing the guesswork and manual inspection on crop stages, this project will increase work efficiency and confidence in farming practices along with promoting overall livelihood development for better economic security of farmers.

5.2.2 Impact on Society & Environment

Impact on Society:

By investing in your farming communities via technology, you can make sustainable agriculture knowledge known. It also can help farmers save input cost at the same time as ensuring health fruits by using less chemicals. It also breeds innovation and empowers farmers to choose with their eyes open. By becoming more productive, local trade and economies benefit from increased stability and the information sharing brought on by the platform reinforces community involvement. The initiative also supports inclusive agricultural education, research and long-term development in emerging economies through the dissemination of ICT-enabled tools for early disease detection to farmers in remote areas.

Impact on Environment:

The project makes high impact environmental contributions that come through the reduction of preservative overuse in pesticide on spot occurrence diseases. Reduced chemical use is good for soil, waters supplies and also safer for attendant wildlife. Early detection helps farmers to isolate and manage disease outbreaks, minimizing crop wastage and spread. It is consistent with the integrated control and sustainable production, aligning agriculture to environmental conservation. This mechanism preserves farmhouses in gardening activity which - togethe with the hope for biodiversity, healthier ecosystems, and a lesser long-term damage to the natural resources and agricultural landscapes.

5.2.3 Ethical Aspects

Data Collection: The records were obtained with the consent of the farmers and care was taken to offer privacy, confidentiality, and responsible dataset management.

Algorithms Bias: Models were constructed and trained to alleviate class imbalance to mitigate biased predictions.

Transparency: The decision-making process is documented, so model results are transparent and traceable.

Accessibility and Equity: The part is open to all farmers facilitating equal opportunity irrespective of the size of their farm.

5.2.4 Sustainability Plan

System Scalability: The system is scalable with respect to increasing the dataset population, new types of diseases, and different kinds of crops.

Cost Effective: Light models are needed to train you model which in turn costs less.

Community Involvement: There are training programs that motivate farmers to accept and believe in the technology.

Community involvement: Working with scientists, specialists and policy-makers makes the uptake of innovations and its development easier.

5.3 Project Management and Financial Analysis

Table: 5.3.1 Financial Analysis

Activity	**Budget (BDT)**
High-Performance Workstation	2500
Google Colab Pro	2500
Field Travel & Logistics	10,000
Agricultural Officer Consultation	1000
Total	18,000

5.4 Complex Engineering Problem

5.4.1 Complex Problem Solving

Table 5.4.1.1: Mapping with complex problem solving.

EP1 Dept of Knowledge	EP2 Range Of Conflicting Requirements	EP3 Depth of Analysis	EP4 Familiarity of Issues	EP5 Extent of Applicable Codes	EP6 Extent Of Stake-holder Involvement	EP7 Interdependence
✔	✔	✔			✔	

Justification of EP1 - Depth of Knowledge: This project is focusing on advanced machine learning approaches and in depth understanding in algorithms, image processing/integration and Agriculture domain knowledge to detect the tea leaf diseases.
Justification of EP2 - Range of Conflicting Requirements: The project entails operating outside the realm of machine learning, i.e., coordination with agriculture experts, learning about crop health, and addressing practical concerns in field data gathering, which can conflict with technical optimization goals.
Justification of EP3 - Depth of Analysis: The project entails comprehensive data analysis in terms of preprocessing, augmentation, model training, and analysis of performance for making strong and reliable predictions.
EP6 - Stakeholder Involvement Extent Rationale: To ensure that the most relevant stakeholders will be considered in validating field data, demonstrating practical viability, and providing end-user responses.

Mapping with Knowledge Profile for EP1

Table 5.4.1.2: Mapping with knowledge Profile.

K3 Engineering Fundamentals	K4 Specialist Knowledge	K5 Engineering Design	K6 Engineering Practice	K8 Research Literature
✔	✔			✔

Justification of K3 - Engineering Fundamentals: This project is based on fundamentals of engineering (e.g., algorithm design, data processing, system evaluation), that are essential to the core machine learning and image processing.
Justification of K4 - Specialist Knowledge: Frilling discussions need to be done with field officers involved in agriculture in order to understand Tea leaf disease, its symptoms and also field problems so as the solution could of farming oriented.
Justification of K8 - Research Literature: The project will build on a comprehensive literature review describing the state of the art in plant disease detection, and escalating from recent research to reveal gaps and provide further support for the proposed approach.

5.4.2 Engineering Activities

Table 5.4.2.1: Mapping with complex engineering activities.

EA1 Range of resources	EA2 Level of Interaction	EA3 Innovation	EA4 Consequences for society and environment	EA5 Familiarity
✔		✔	✔	

Justification for EA1 - Diversity of Resources: The project combines a wide variety of resources including computational (e.g., machine learning platforms, cloud computing)and inputs at the farm level (e.g., farm information, expertise consultation). This way it has a decent mix of theory and hands on stuff in the project.

Justification for EA3-Innovation: It is a fresh medium of applying modern technology in detecting Tea leaf disease through machine learning to solve an age-old issue of agriculture. This is a great extension to the necessary focus on creative engineering solutions.

Justification for EA4: Impact on the Environment and Society The project has a significant effect on society in food production because there is more agricultural production and less crop loss. This is also relevant to assignment because it improved the usage of resources and reduced pesticide use, which has an eco-impact.

5.5 Summary

This chapter defined the engineering requirements, social and environmental consequences, socially and ethically oriented factors, and sustainability strategies of the tea leaf disease classification system. It focused on how the project will enhance agricultural practices, assist farmers and sustainable farming. The system is scalable, cost effective, and value generating to agriculture by addressing technical, ethical and environmental concerns.

Chapter 6

Conclusion

Here, summaries of work done, its contribution, limitations encountered, and future scope of work with the aim of rendering the system more efficient and extendable will be given.

6.1 Summary

The objective of this work was to create a tea leaf disease classification system using InceptionV3, MobileNetV2, and a Custom CNN-based deep learning.Data set contained 1,968 samples of six disease and normal classes (70% for training, 20% for testing and 10% for validation). The best accuracy of 98.94% belonged to InceptionV3, followed by the Custom CNN by 96.57% and MobileNetV2 by 83.78%. The paper established that transfer learning models may be applied with high degree of accuracy and that lightweight models may be applied on mobile phones. Ethics, sustainability and scheme of communication considerations were implemented with the aim of promoting fairness and feasibility. Overall, the study may contribute to enhancing agricultural innovation by providing for the efficient and early identification of tea leaf disease and reducing the excessive use of pesticides and supporting farmers in maintaining their livelihoods. The solution provided by it is scalable and may enhance precision agriculture, sustainability and food security and effortlessly satisfy traditional farming and Smart Agriculture technology uptake.

6.2 Limitation

In spite of the positive outcomes of this research, limitations were also observed. The size of the dataset, whilst balanced, remains relatively small compared to massive-scale agriculture datasets, which would affect the generalizability of the models to real life. Images were captured under controlled conditions, i.e., model performance under different conditions of lighting, weather, and background may vary. Further, few classes were utilized, and other significant disease conditions affecting tea production were not included. Experimentation using larger scale, more intricate architectures and hyperparameter optimization was restricted due to computational limitations. Light-weight MobileNet V2 recorded a lower accuracy since it recorded compromised trade-offs in between model size and model accuracy. The system in place is merely conducting image classification with no combination of other IoT sensors or environmental datasets to fill the accuracy of diagnosis gaps. Finally, deployment models were not extensively tried and tested on larger farms and its adoption in the rural area may be restricted by the availability of internet constraints. The above highlight the potential for system further development to enable its use on large-scale agriculture.

6.3 Future Work

Further research will be conducted to increase the range of images of different regions, light conditions, and backgrounds to enhance generalization. These other types of diseases will also be added so as to create a more elaborate system of classification. It is possible to develop more sophisticated methods such as attention mechanisms, ensemble learning, and Vision Transformers (ViTs) to improve accuracy and robustness.

Optimizing for the mobile platform In addition to using quantization and pruning, methods used to reduce computational costs (at the cost of performance), optimization for a mobile platform will become paramount. It will also be combined with iot-based agriculture systems including weather sensors, soil analyzers and drones to enable it to have a comprehensive farm management capability.

Developing a user-friendly mobile app that has offline functionality can increase accessibility to rural farmers with low connectivity. The cooperation with agricultural specialists will enhance the validation and applicability of the models, providing a productive advantage to the tea producers. Lastly, it is possible to scale this to other agricultural sectors using tea crops to form a scalable, cross-crop disease detection system, to sustain agriculture and enhance food security worldwide.

References

[1] J. Chen, Q. Liu, and L. Gao, “Visual tea leaf disease recognition using a Convolutional Neural Network model,” Symmetry, vol. 11, no. 3, p. 343, Mar. 2019, doi: 10.3390/sym11030343.

[2] S. Gayathri, D. J. W. Wise, P. B. Shamini, and N. Muthukumaran, “Image analysis and detection of tea leaf disease using deep learning,” in 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC), Coimbatore, India, 2020, pp. 398- 403, doi: 10.1109/ICESC48915.2020.9155850.

[3] S. Datta and N. Gupta, “A novel approach for the detection of tea leaf disease using deep neural network,” Procedia Computer Science, vol. 218, pp. 2273-2286, 2023, doi: 10.1016/j.procs.2023.01.203.

[4] R. Singh, N. Sharma, and R. Gupta, “Proposed CNN model for tea leaf disease classification,” in 2023 2nd International Conference on Applied Artificial Intelligence and Computing (ICAAIC), Salem, India, 2023, pp. 53-60, doi: 10.1109/ICAAIC56838.2023.10140680.

[5] S. Mukhopadhyay, M. Paul, R. Pal, and D. De, “Tea leaf disease detection using multi-objective image segmentation,” Multimedia Tools and Applications, vol. 80, no. 1, pp. 753-771, 2021, doi: 10.1007/s11042-020-09567-1.

[6] R. S. Latha et al., “Automatic detection of tea leaf diseases using deep convolution neural network,” in 2021 International Conference on Computer Communication and Informatics (ICCCI), Coimbatore, India, 2021, pp. 1-6, doi: 10.1109/ICCCI50826.2021.9402225.

[7] V. Jain, V. S. Rajpoot, A. Kumar, and M. Kumar, “An integrated approach for predicting and broadcasting tea leaf disease at early stage using IoT with machine learning – A review,” Materials Today: Proceedings, vol. 37, pp. 2577-2581, 2021, doi: 10.1016/j.matpr.2020.05.458.

[8] M. F. A. Fauzi, S. D. Nasution, M. A. Muchtar, and I. Siregar, "Classification of tea leaves diseases by developed CNN, feature fusion, and classifier based model," International Journal of Advances in Mechanical and Civil Engineering, vol. 10, no. 1, pp. 1-6, 2023, doi: 10.18100/ijamec.1235611.

[9] Y.-C. Chang, C.-Y. Chang, and J.-Y. Chen, “Development of image recognition and classification algorithm for tea leaf diseases using convolutional neural network,” in 2018 ASABE Annual International Meeting, Detroit, MI, 2018, p. 1, doi: 10.13031/aim.201801254.

[10] K. P. S. Kumaratenna and Y.-Y. Cho, “Tea leaf disease classification using artificial intelligence (AI) models,” Journal of Bio-Environment Control, vol. 33, no. 1, pp. 1-11, 2024, doi: 10.12791/KSBEC.2024.33.1.001.

[11] U. Hairah, A. Septiarini, N. Puspitasari, A. Tejawati, H. Hamdani, and S. E. Priyatna, “Classification of tea leaf disease using convolutional neural network approach,” International Journal of Electrical and Computer Engineering (IJECE), vol. 14, no. 3, pp. 3287-3294, 2024, doi: 10.11591/ijece.v14i3.pp3287-3294.

[12] [N. D. D. Bright, S. Marchi, D. V. R. Blessy, and T. M. Selvi, “Tea leaf disease classification and tea bud identification,” in 2022 International Conference on Edge Computing and

Rations (ICERECT), Chennai, India, 2022, pp. 1-5, doi: 10.1109/ICERECT56837.2022.10059683.
[13] M. R. Firdaus, M. Z. Rifqi, and I. P. Sari, “Classification of tea leaf diseases based on ResNet-50 and Inception V3,” Sinkron, vol. 8, no. 3, pp. 1459-1468, 2023, doi: 10.33395/sinkron.v8i3.12604.
[14] P. Ananthi, S. Janani, S. Keerthana, and V. Monisha, “Tea leaf diseases classification and detection using a convolutional neural network,” in 2023 International Conference on Signal Processing, Computation andSystem Security (ICSCSS), Coimbatore, India, 2023, pp. 1-6, doi: 10.1109/ICSCSS57650.2023.10169680.
[15] S. Shinde and S. Lahade, “Deep learning for tea leaf disease classification,” in Machine Learning and Deep Learning for Sustainable Development. Boca Raton, FL: CRC Press, 2023, pp. 361- 377, doi: 10.1201/9781003359456-20.
[16] P. K. Das and M. H. Bhuyan, “Classification of tea leaf diseases using convolutional neural network,” in Intelligent Systems, Singapore, 2022, pp. 257–265, doi: 10.1007/978-981-19-0019-8_22.
[17] M. J. A. Soeb et al., “Tea leaf disease detection and identification based mon YOLOv7 (YOLO-T),” Scientific Reports, vol.13,no.1,p.6078,Apr. 2023, doi:10.1038/s41598-023-33270-4.
[18] X. Ye et al., “Tea leaf disease and insect identification based on improved MobileNetV3,” Frontiers in Plant Science, vol. 15, p. 1459292, Jul. 2024, doi: 10.3389/fpls.2024.1459292.
[19] R. S. Latha, G. R. Sreekanth, R. C. Suganthe, R. Rajadevi, S. Karthikeyan, S. Kanivel, and B. Inbaraj, "Automatic detection of tea leaf diseases using deep convolution neural network," in 2021 International Conference on Computer Communication and Informatics (ICCCI), Coimbatore, India, 2021, pp. 1-6, doi: 10.1109/ICCCI50826.2021.9402225.
[20] M. L. Bharathi and P. Poongodi, “Detection of disease in tea leaves using convolution neural network,” in 2020 4th International Conference on Trends in Electronics and Informatics (ICOEI), Tirunelveli, India, 2020,

212-15-14749

ORIGINALITY REPORT

22%	16%	13%	13%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Daffodil International University Student Paper	7%
2	dspace.daffodilvarsity.edu.bd:8080 Internet Source	2%
3	Submitted to Institute of International Studies Student Paper	1%
4	export.arxiv.org Internet Source	1%
5	Submitted to United International University Student Paper	<1%
6	engj.org Internet Source	<1%
7	koreascience.kr Internet Source	<1%
8	"Proceedings of the Fifth International Conference on Trends in Computational and Cognitive Engineering", Springer Science and Business Media LLC, 2024 Publication	<1%
9	Novianti Puspitasari, Ummul Hairah, Vina Zahrotun Kamila, Hamdani Hamdani, Anindita Septiarini, Amin Padmo Azam Masa. "Enhancing crude palm oil quality detection using machine learning techniques", IAES International Journal of Artificial Intelligence (IJ-AI), 2025 Publication	<1%

10	arxiv.org Internet Source	<1 %
11	technoaretepublication.org Internet Source	<1 %
12	dergipark.org.tr Internet Source	<1 %
13	www.ksbec.org Internet Source	<1 %
14	Submitted to Midlands State University Student Paper	<1 %
15	Pushpa Choudhary, Sambit Satpathy, Arvind Dagur, Dhirendra Kumar Shukla. "Recent Trends in Intelligent Computing and Communication", CRC Press, 2025 Publication	<1 %
16	ebin.pub Internet Source	<1 %
17	B. M. Shahria Alam, Fahad Ahammed, Golam Kibria, Mohammad Tahmid Noor et al. "teaLeafBD: A Comprehensive Image Dataset to Classify the Diseased Tea Leaf to Automate the Leaf Selection Process in Bangladesh", Data in Brief, 2025 Publication	<1 %
18	ejournal.unuja.ac.id Internet Source	<1 %
19	Ananthakrishnan Balasundaram, Prem Sundaresan, Aryan Bhavsar, Mishti Mattu, Muthu Subash Kavitha, Ayesha Shaik. "Tea Leaf Disease Detection using Segment Anything Model and Deep Convolutional Neural Networks", Results in Engineering, 2024 Publication	<1 %