CASSAVA LEAF DISEASE DETECTION

A PROJECT REPORT

Submitted by

GIJIN T GEORGE (TKM23MCA-2029)

to

TKM College of Engineering

Affiliated to

The APJ Abdul Kalam Technological University

In partial fulfillment of the requirements for the award of the degree of

MASTER OF COMPUTER APPLICATION

Thangal Kunju Musaliar College of Engineering
Kerala

DEPARTMENT OF COMPUTER APPLICATION

NOVEMBER 2024

DECLARATION

I undersigned hereby to declare that the project report on **CASSAVA LEAF DISEASE DETECTION**, submitted for partial fulfillment of the requirements for the award of degree of **Master of Computer Application** of the **APJ Abdul Kalam Technological University, Kerala** is a Bonafide work done by me under supervision of **Dr. Sheeba K.** This submission represents my ideas in my own words and where ideas or words of others have been included, we have adequately and accurately cited and referenced the original sources. I also declare that we have adhered to ethics of academic honesty and integrity and have not misrepresented or fabricated any data or idea or fact or source in our submission. I understand that any violation of the above will be a cause for disciplinary action by the institute and/or the University and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been obtained. This report has not been previously formed the basis for the award of any degree, diploma or similar title of any other University.

KOLLAM GIJIN T GEORGE

11/11/24 TKM23MCA-2029

DEPARTMENT OF COMPUTER APPLICATION

TKM COLLEGE OF ENGINEERING

(Government Aided and Autonomous)

KOLLAM - 5

CERTIFICATE

This is to certify that, the report entitled **Cassava Leaf Disease Detection** submitted by **GIJIN T GEORGE (TKM23MCA-2029)** to the **APJ Abdul Kalam Technological University** in partial fulfillment of the requirements for the award of the Degree of **Master of Computer Application** is a Bonafide record of the project work carried out by him/her under my/our guidance and supervision. This report in any form has not been submitted to any other University or Institute for any purpose.

---------------------------------- ----------------------------------

Internal Supervisor(s) Mini Project Coordinator

ACKNOWLEDGEMENT

First and foremost, I thank GOD almighty and our parents for the success of this project. I owe sincere gratitude and heart full thanks to everyone who shared their precious time and knowledge for the successful completion of my project.

I am extremely grateful to **Prof. Natheera Beevi M**, Head of the Department, Department of Computer Application, for providing us with the best facilities.

I extend my sincere gratitude to my project guide, **Dr. Sheeba K.**, for her invaluable guidance, encouragement and support throughout the course of this project.

I am sincerely grateful to **Prof. Sheera Shamsu**, our project coordinator, for her invaluable guidance and support throughout our project journey.

I profusely thank all other faculty members in the department and all other members of TKM College of Engineering, for their guidance and inspirations throughout my course of study.

I owe thanks to my friends and all others who have directly or indirectly helped me in the successful completion of this project.

GIJIN T GEORGE

ABSTRACT

Cassava is a vital crop for food and nutrition security in Africa, providing essential carbohydrates and thriving in harsh environments. However, cassava crops are susceptible to leaf diseases that can significantly diminish yields and reduce farmers' income. Current CNN models used for detecting these diseases often struggle with poor accuracy and high computational demands. This project presents a deep learning-based solution for detecting cassava leaf diseases using the cassava-leaf-disease-classification dataset. The proposed system utilizes three pre-trained convolutional neural network models—ResNet50, EfficientNet, and MobileNet—to classify cassava leaves into various disease categories or as healthy. The models are evaluated and compared based on accuracy, computational efficiency, and robustness to challenging cases, such as variations in image quality. To handle the dataset's inherent imbalance, data augmentation techniques are applied, improving model performance across underrepresented disease classes.

The exceptional performance of the CNN models demonstrates their capability to accurately identify cassava leaf diseases, highlighting the system's potential to be adapted for various crops and agricultural contexts. By providing an automated solution for cassava leaf disease identification, this project addresses a critical challenge in small- to large-scale farming, where traditional methods are often labor-intensive and time-consuming. The implementation of such a system can significantly enhance agricultural productivity and disease management, especially in regions like Africa, where cassava is a staple crop and effective disease management is essential. The success of this project establishes a solid foundation for developing a more comprehensive plant disease identification system, which can be refined and expanded to provide an integrated agricultural advisory framework, ultimately supporting improved crop management and food security.

CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Cassava is a vital crop for food security, especially in Africa, which is the world's largest producer. Known for its resilience, cassava can thrive in diverse climates, including drought-prone and nutrient-poor soils. However, despite its adaptability, cassava production faces significant challenges from various leaf diseases. These diseases not only reduce crop yield but also impact the financial stability of farmers who rely on cassava as a primary income source. Effective management of these diseases is essential to protect crop quality and ensure steady production, as cassava leaf diseases are comparable to similar threats found in crops like maize. Early identification of these diseases is crucial, as it allows farmers to take timely measures to prevent widespread infection, ultimately preserving the health of the crop and the income of those dependent on it.

Traditional methods for detecting and managing cassava leaf diseases are often labor-intensive, costly, and unreliable, posing challenges for farmers. In response, deep learning and image classification technologies have been developed to automate disease detection from leaf images. These advanced solutions enable early, accurate identification, allowing farmers to address issues proactively. This approach enhances productivity, reduces crop losses, and promotes sustainable agricultural practices, particularly in regions where cassava is a staple.

For this project, I used the cassava-leaf-disease-classification dataset, which includes 21,489 images across 4 cassava disease types, as the foundation for training and evaluating three key deep learning architectures: ResNet50, EfficientNet, and MobileNet. After extensive testing, the ResNet50 model achieved the highest accuracy and was selected to power a Flask-based user interface, enabling users to upload cassava leaf images and instantly receive diagnostic results. This dynamic approach to cassava leaf disease detection offers new possibilities for scalable, real-time monitoring, supporting proactive disease management in agriculture.

Through this project, I aim to bridge the gap between technology and agriculture, demonstrating how deep learning can revolutionize cassava leaf disease detection with a fast, reliable, and accessible solution that supports sustainable farming practices.

1.1 Existing System

Existing cassava leaf disease detection systems have often relied on traditional machine learning models or basic image processing techniques to identify diseases. These systems typically use classical methods like color analysis, texture extraction, and basic classifiers such as Support Vector Machines (SVM) or Random Forests. While these approaches have achieved some success, they encounter significant challenges when applied to large and diverse datasets. Variations in lighting, leaf structure, and environmental conditions make consistent disease identification difficult. Additionally, these systems often struggle with complex symptoms that are hard to distinguish from healthy leaf parts, leading to potential misclassifications.

In terms of deep learning, Convolutional Neural Networks (CNNs) have shown enhanced performance in detecting cassava leaf diseases. However, traditional CNN models may struggle to generalize effectively to new, unseen data, often facing issues with overfitting, particularly when dealing with imbalanced datasets or underrepresented diseases. Additionally, traditional CNN models typically require separate classifiers for each disease type, reducing efficiency and complicating large-scale deployment in real-world agricultural settings. Managing complex, multi-class classifications with these methods can be resource-intensive and may demand substantial computational power, limiting their accessibility and broader application in agriculture.

Major drawbacks of currently existing models for cassava leaf disease detection include:

- **Class Imbalance**: Many datasets have uneven distributions of disease types, with some diseases being underrepresented. This leads to models that are biased and perform poorly on less frequent disease classes.
- **Dependence on Traditional Machine Learning Techniques**: Existing systems often rely on manual feature extraction, which is time-consuming and prone to inaccuracies. These models have limited performance due to their reliance on handcrafted features and may struggle to generalize effectively to new datasets.
- **Poor Robustness in Real-World Conditions**: Traditional models are sensitive to environmental variations, such as lighting, angle, and leaf appearance, which can compromise detection accuracy.

- **Resource Intensive**: Separate models are often needed for each disease type, resulting in inefficiency and high computational demands.
- **Inability to Handle Multi-Class Classification**: Many existing models require individual classifiers for different diseases, making them complex and difficult to scale for simultaneous multi-disease detection.

1.2 Proposed System

The proposed system addresses the limitations of traditional cassava leaf disease detection by employing deep learning models, such as MobileNet, ResNet50, and EfficientNet, to accurately classify cassava leaf diseases from plant leaf images. These models offer enhanced accuracy and robustness, especially under varying conditions like lighting and angle.

- **End-to-End Deep Learning:** The system eliminates manual feature extraction, allowing deep learning models like MobileNet, ResNet50, and EfficientNet to learn directly from cassava leaf images, improving efficiency and accuracy in detecting diseases.
- **Data Augmentation:** Techniques such as rotation, flipping, and scaling are used to diversify the training data, helping address class imbalance between healthy and diseased cassava leaves, while improving the model's generalization capabilities.
- **Real-World Robustness:** Trained on a diverse dataset of cassava leaf images, the system can handle variations in image quality, leaf appearance, and environmental factors, providing higher accuracy compared to traditional models.
- **Multi-Class Classification:** A single model can differentiate between various cassava leaf diseases and healthy leaves, streamlining deployment and improving scalability for agricultural applications.
- **Scalability and Efficiency:** Pre-trained models like ResNet50 and EfficientNet offer high performance on large datasets, reducing the need for multiple classifiers and enabling efficient disease detection at scale for cassava farms.
- **Real-Time Detection:** Through a user-friendly interface, users can upload cassava leaf images for instant disease predictions, facilitating quicker and more accessible diagnostics for farmers.

1.3 Objectives

- **Develop a Deep Learning Model for Disease Classification:** Create a deep learning model based on ResNet50, specifically designed to classify cassava leaf diseases based on leaf images. This model focuses solely on disease identification, eliminating the need for separate models for different disease types and simplifying the process while offering high accuracy and robustness.

- **Implement Enhanced Feature Extraction with Pre-trained Models:** Leverage the pre-trained ResNet50, MobileNet, EfficientNet models for feature extraction to improve the accuracy and generalization ability of the cassava leaf disease detection system. This approach enhances the model's capability to detect diseases in cassava leaves, even under varying conditions, without the need for extensive training from scratch.

- **Improve Model Performance on Real-World Data:** Apply preprocessing techniques such as resizing, normalization, and data augmentation (e.g., rotation, flipping, scaling) to address challenges posed by environmental factors like lighting, image quality, and leaf orientation. This will ensure the ResNet50-based model is robust and performs effectively on diverse real-world cassava leaf images.

- **Create an Accessible User Interface:** Develop a user-friendly, web-based interface where users can easily upload cassava leaf images to receive real-time disease predictions. This system will cater to farmers, agricultural experts, and researchers, making cassava leaf disease detection accessible to a broader audience and facilitating quicker decision-making in the field.

- **Achieve High Accuracy and Efficiency:** Optimize the deep learning models to ensure high accuracy in detecting cassava leaf diseases while maintaining efficient processing. This will make the system suitable for real-world agricultural applications, where quick diagnosis and minimal computational resources are crucial for large-scale deployment across farms.

CHAPTER 2

LITERATURE REVIEW

A literature survey, also known as a literature review, is a comprehensive study and evaluation of existing research on cassava leaf disease detection using deep learning. It involves identifying, analyzing, and synthesizing relevant sources such as scholarly articles, books, and studies on plant disease detection through image analysis, specifically focusing on cassava leaf disease classification using deep learning techniques. This review aims to provide a comprehensive overview of the current state of knowledge, highlighting key advancements and methodologies in disease diagnosis using machine learning and deep learning.

The purpose of this literature survey is to explore gaps in existing cassava leaf disease detection methods, establish the significance of deep learning in agricultural image analysis, and provide a theoretical framework for developing a more efficient, automated disease detection system. It will also examine the evolution of diagnostic methods, from traditional techniques such as manual visual inspection and basic machine learning approaches, to modern CNN-based architectures like ResNet50, EfficientNet, and MobileNet, which have been successfully applied to plant disease classification tasks.

This survey provides insights into the challenges of real-world agricultural image analysis, such as class imbalance, image quality variations, and overfitting, and reviews the effectiveness of various techniques such as transfer learning and data augmentation to address these issues.

Conducting this review establishes a foundation for developing a robust cassava leaf disease detection system and ensures that the proposed solution is built upon the latest advancements and best practices in the field.

2.1 Purpose of the Literature Review

Cassava leaf disease detection remains a critical challenge in agriculture, particularly in regions with limited access to advanced diagnostic tools. The traditional methods of disease detection, primarily relying on visual inspection by experts or manual classification, are time-consuming and prone to human error. In recent years, deep learning techniques, particularly Convolutional Neural Networks (CNNs), have emerged as powerful tools for automating plant disease detection, including the identification of cassava leaf diseases.

By summarizing existing knowledge on the application of deep learning to cassava leaf disease detection, this study establishes the context in which it fits within the broader field of agricultural image analysis. Previous research has shown promising results in using CNN architectures such as ResNet50, EfficientNet, and MobileNet for plant disease classification tasks. However, many studies have highlighted challenges such as the need for large, high-quality datasets, handling class imbalance, and ensuring model robustness under varying real-world conditions, such as changes in lighting, leaf orientation, and image quality. These gaps indicate that there is still significant room for improvement in making cassava leaf disease detection systems more robust and accessible, particularly in low-resource agricultural settings.

This literature review aims to identify areas where previous research has left unanswered questions, particularly regarding how deep learning models can generalize across different cassava diseases and varying image qualities. New research can contribute by addressing these gaps, potentially by leveraging techniques such as transfer learning, data augmentation, and exploring more efficient model architectures like MobileNet.

By analyzing relevant theories and methodologies from previous studies, this review helps construct a theoretical framework for the current study. It also informs the research design by examining how past methodologies, such as preprocessing techniques and evaluation metrics, can be improved or adapted for cassava leaf disease detection. This study builds on prior research but avoids duplicating efforts by focusing on improving the robustness, scalability, and real-world applicability of disease detection models in agriculture.

Additionally, the review provides insights into the evolution of deep learning in agricultural diagnostics, offering a historical perspective on how the application of these models has developed over time. Despite significant progress, the current study is essential to further refine cassava leaf disease detection methods, making them more reliable and accessible to farmers, particularly in underserved regions.

2.2 Related Works

1. **"Cassava Leaf Disease Classification using Deep Convolutional Neural Networks" (2020).** This study explored the use of Convolutional Neural Networks (CNNs) for classifying various diseases in cassava leaves. The authors developed a model trained on a dataset of cassava leaf images, with disease categories including Cassava Mosaic Disease (CMD) and Cassava Brown Streak Disease (CBSD). Their approach demonstrated the potential of CNN architectures, with high classification accuracy, suggesting CNNs as an effective tool for plant disease detection.
2. **"Deep Learning for Cassava Disease Classification and Yield Prediction" (2020).** This paper proposed a deep learning-based approach for both disease classification and yield prediction in cassava crops. The study utilized a dataset with labeled images of cassava leaves infected by CMD and CBSD. The researchers applied deep neural networks and showed that the model was capable of high-performance classification, with the ability to accurately identify diseased leaves and predict yield outcomes based on disease severity.

3. The paper *"Cassava Leaf Disease Detection and Classification by Deep Learning: A Review"* provides a comprehensive analysis of how deep learning techniques, particularly Convolutional Neural Networks (CNNs), have been applied to the detection and classification of cassava leaf diseases. This review discusses the evolution of disease identification from traditional methods, such as manual inspection and expert knowledge, to automated, deep learning-driven approaches. The authors highlight the substantial improvements in detection accuracy and efficiency achieved through CNNs, particularly in distinguishing between diseases like Cassava Brown Streak Disease (CBSD) and Cassava Mosaic Disease (CMD) from images of cassava leaves.The paper examines various datasets used in the field of cassava disease detection, including large-scale image collections that have been annotated for training deep learning models. It also explores common preprocessing techniques, such as data augmentation (rotation, flipping, and scaling), normalization, and resizing, which help enhance model performance by addressing issues like class imbalance and variability in image quality.Moreover, the review explores the integration of deep learning models with mobile and IoT technologies, facilitating real-time disease detection for farmers in the field. The ability to deploy these models on mobile devices helps in providing instant diagnoses, thus assisting farmers in taking timely action to manage crop health.

4. The paper *"Real-Time Cassava Leaf Disease Dataset Development and Detection Using Deep Learning"* (2024) discusses the development of a real-time cassava leaf disease detection system using deep learning techniques. The authors highlight the crucial role of early disease detection in ensuring food security, particularly for cassava, which is a staple crop in many developing countries. Given the high impact of diseases like Cassava Brown Streak Disease (CBSD) and Cassava Mosaic Disease (CMD) on global cassava production, this research focuses on improving the speed and accuracy of disease identification.To address challenges encountered in real-world agricultural environments, the researchers created specialized datasets tailored to cassava leaf diseases, similar to previous efforts focused on other crops like rice, wheat, and maize. These datasets include images of common cassava diseases, including Cassava Brown Streak Disease (CBSD) and Cassava Mosaic Disease (CMD). CNN-based models, such as ResNet50, MobileNet, and EfficientNet, were utilized for disease classification. The study also introduced a new CNN model, which demonstrated impressive performance with testing accuracies reaching up to 98%.
5. The paper *"A Survey on Using Deep Learning Techniques for Cassava Leaf Disease Diagnosis and Recommendations for the Development of Efficient Detection Tools"* (2022) by Aanis Ahmad, Dharmendra Saraswat, and Aly El Gamal provides a detailed review of deep learning methods used in cassava leaf disease detection. The authors analyze various techniques, including Convolutional Neural Networks (CNNs), to classify and diagnose diseases affecting cassava plants, such as Cassava Brown Streak Disease (CBSD) and Cassava Mosaic Disease (CMD). They stress the need for more robust, real-time diagnostic tools to help farmers better manage the health of their crops, ultimately leading to improved crop yields and reduced losses.The paper addresses key challenges in the field, particularly the limitations of existing datasets and the impact of environmental factors like lighting and background variations on disease detection accuracy. These issues complicate the identification process, especially for cassava plants, which may vary in appearance based on location and season. The authors propose the development of more accessible, scalable, and user-friendly detection tools, making them viable for use in real-world farming environments.
6. The paper *"Cassava Leaf Disease Detection Using Deep Learning"* (2023) by Adesh V. Panchal et al. explores the application of deep learning techniques, specifically Convolutional Neural Networks (CNNs), in automating the process of detecting cassava leaf diseases through image analysis. Traditionally, cassava disease detection has relied

on manual inspection, which can be inefficient and prone to human error. The authors investigate how CNNs can be used to classify and identify diseases such as Cassava Brown Streak Disease (CBSD) and Cassava Mosaic Disease (CMD) from images of cassava leaves.This research aims to significantly enhance the speed and accuracy of disease diagnosis, which is crucial for managing crop health and preventing large-scale losses. The study highlights that deep learning methods can address many of the limitations associated with traditional diagnostic approaches and have the potential for scalability in real-world agricultural applications.Despite the promising results, the paper also discusses several challenges, such as the need to optimize models for varying environmental conditions and the difficulty in processing real-time image data for effective disease detection. These hurdles underline the ongoing need for research to refine deep learning models and ensure their applicability in diverse agricultural settings.

CHAPTER 3

METHODOLOGY

The cassava leaf disease detection system employs deep learning models, particularly Convolutional Neural Networks (CNNs) such as ResNet50, EfficientNet, and MobileNet, to classify cassava leaf diseases from images. The dataset used in this study comprises thousands of images of cassava leaves, annotated with labels for common diseases like Cassava Brown Streak Disease (CBSD) and Cassava Mosaic Disease (CMD). These images are divided into training, validation, and test sets to ensure robust model training and evaluation.

For preprocessing, all images are standardized through resizing and normalization to maintain uniformity across the dataset. Data augmentation techniques such as rotation, zooming, flipping, and color adjustments are applied to diversify the training data, improving the model's ability to generalize to unseen images. These augmentations address challenges posed by real-world variations, including changes in leaf orientation, lighting conditions, and image noise.

The models are trained using advanced deep learning frameworks with optimizers like Adam, which helps to adjust learning rates dynamically for more efficient training. During training, the system monitors various metrics such as accuracy, precision, recall, and F1-score to evaluate the model's performance comprehensively.

Figure 3.1 illustrates the block diagram of the cassava leaf disease detection project workflow, from data preprocessing and augmentation to model training and deployment. The trained model is integrated into a user-friendly interface, enabling real-time cassava disease detection by uploading images of cassava leaves. This approach provides an automated, accurate solution for early disease identification, aiming to improve crop health management and prevent significant yield losses. By leveraging deep learning techniques such as CNNs (ResNet50, EfficientNet, MobileNet), the system enhances the effectiveness of disease detection in agricultural settings, making it more accessible to farmers and researchers alike.

3.1 Block Diagram

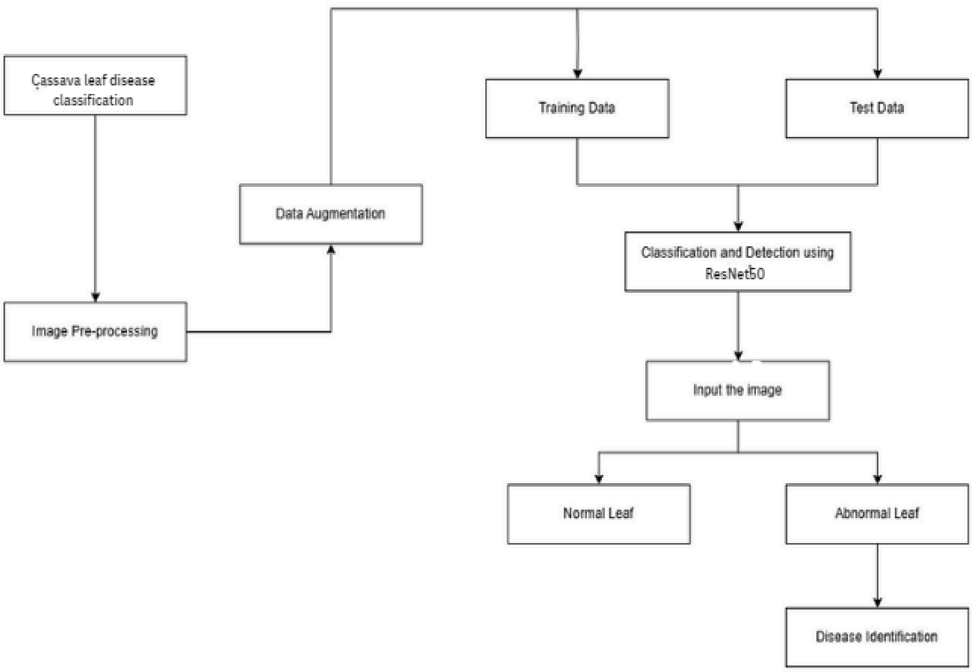

Figure 3.1: Block Diagram

3.1.1 Data Collection

The dataset used in this project is the Cassava Leaf Disease Classification, which contains over 21,000 high-resolution images of cassava leaves, categorized into healthy and diseased classes. The dataset includes images representing diseases like Cassava Brown Streak Disease (CBSD) and Cassava Mosaic Disease (CMD), offering a comprehensive resource for training models aimed at detecting and classifying these conditions. Preprocessing steps, such as image resizing and augmentation (e.g., rotation, flipping, zooming), are applied to enhance model accuracy and generalization. The dataset is clearly labeled, with each image representing either a healthy leaf or one affected by a specific disease. By capturing detailed leaf textures and disease symptoms, it serves as a strong foundation for building deep learning models that can perform real-time cassava disease detection. Figure 3.2 illustrates sample images of healthy and disease-affected cassava leaves, highlighting the distinct visual differences between these conditions. This dataset is a key resource for developing robust models for cassava disease identification in agricultural applications.

Cassava Bacterial Blight(CBB)

Cassava Brown Streak Disease(CBSD)

Cassava Green Mottle(CGM)

Cassava Mosaic Disease(CMD)

Healthy Cassava Leaf

Figure 3.2: Sample Images from Dataset

3.1.2 Image Preprocessing

Preprocess the collected data, ensuring consistency and model readiness, images undergo several preprocessing steps:

- **Color Conversion**: Images are converted to color format RGB to standardize the input for the model.
- **Resizing and Normalization**: Each image is resized to a fixed input dimension 64x64 pixels, and pixel values are normalized to enhance model efficiency.
- **Data Augmentation**: Data augmentation techniques, such as horizontal and vertical flipping, rotation, shearing and zooming, are applied to increase image diversity within the dataset. This reduces the risk of overfitting and improves the model's ability to generalize across varied real-world inputs.

3.1.3 Model Architecture Design

1. **MobileNet:** MobileNet is a lightweight Convolutional Neural Network (CNN) architecture developed to optimize image classification and object detection tasks on mobile and embedded devices. By utilizing depthwise separable convolutions, MobileNet reduces the computational complexity and number of parameters compared to traditional convolutional operations, making it highly efficient for devices with limited resources. This allows it to maintain good performance and high accuracy while being fast and lightweight.The architecture also integrates batch normalization to stabilize training and ReLU6 activation functions, which ensure better performance in real-world conditions by avoiding issues like vanishing gradients. MobileNet's key advantage lies in its ability to offer a balance between speed and accuracy, which is critical for mobile applications that require real-time processing. In applications like cassava leaf disease detection, MobileNet's compact structure allows for faster inference times without sacrificing classification accuracy, making it an ideal choice for agricultural solutions on mobile devices.

   Figure 3.3 illustrates the MobileNet architecture, showcasing its efficient structure from input to output, with its various layers designed to minimize computational cost while maintaining performance.

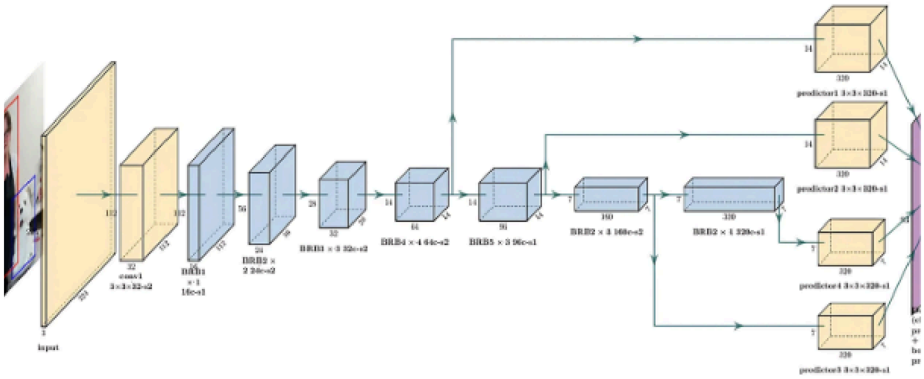

Figure 3.3: MobileNet Architecture

2. **ResNet50 (Residual Networks with 50 layers)** is a deep Convolutional Neural Network (CNN) architecture designed to overcome challenges associated with training very deep networks, such as vanishing gradients and poor convergence. Introduced by Kaiming He et al. in 2015, ResNet50 employs residual connections, or skip connections, which allow the network to learn residual mappings instead of direct mappings, enabling better gradient flow during training. This innovation helps in training deep networks by mitigating the issue of vanishing gradients that often occurs in traditional architectures.ResNet50 consists of 50 layers, with 48 convolutional layers and two fully connected layers. These layers are designed for efficient feature extraction from input images. By incorporating batch normalization and ReLU activation functions, ResNet50 further accelerates the training process and helps prevent overfitting, making it highly effective for a variety of computer vision tasks.The architecture's residual blocks make it easier to train deeper networks, resulting in more accurate models for tasks like image classification, object detection, and medical imaging. This makes ResNet50 a popular choice in domains such as plant disease detection, where its ability to handle complex image data and improve classification accuracy is particularly useful.

   Figure 3.4 illustrates the flow of residual blocks and convolutional layers, providing a visual representation of how ResNet50 processes input data efficiently. The architecture's robustness and scalability have made it one of the most widely adopted CNN architectures in deep learning applications.

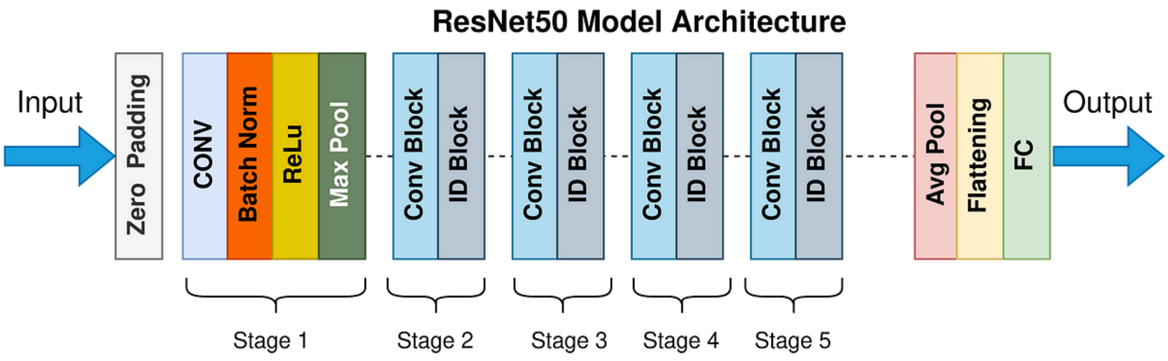

Figure 3.4: ResNet50 Architecture

3. **EfficientNet**: EfficientNet is a cutting-edge Convolutional Neural Network (CNN) architecture designed to achieve high performance while minimizing computational costs. Unlike traditional models, EfficientNet introduces a novel compound scaling method, which simultaneously scales the depth, width, and resolution of the network in a balanced manner, optimizing the model's performance without excessive computational demands. This approach ensures that the network remains efficient, even when scaled up for complex tasks. EfficientNet also leverages depthwise separable convolutions and swish activation functions, which contribute to reducing the overall model size without compromising accuracy. These optimizations allow EfficientNet to outperform older models such as ResNet and Inception on various large-scale image classification tasks, including ImageNet, with fewer parameters and lower computational requirements. EfficientNet's performance in large-scale classification tasks and its ability to maintain a balance between accuracy and computational efficiency make it an ideal choice for real-time plant disease detection, especially in low-resource settings.

   Figure 3.5 illustrates the architecture diagram, showcasing how EfficientNet efficiently scales its layers from input to output. The model's scalability and efficiency make it particularly well-suited for applications requiring both high accuracy and computational efficiency, such as in resource-constrained environments where quick, real-time decision-making is critical.

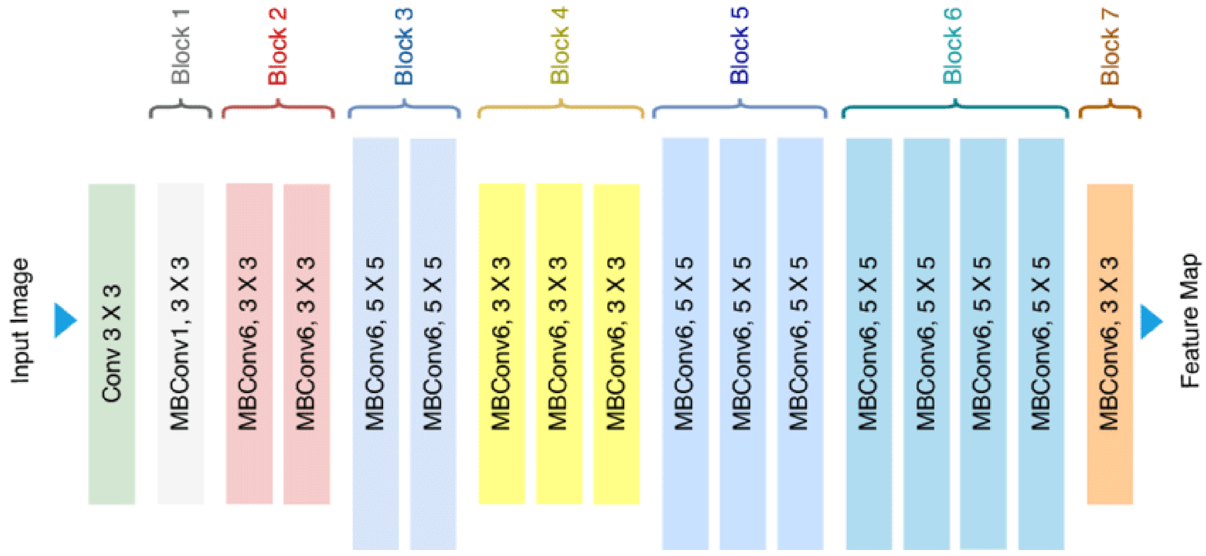

Figure 3.5: EfficientNet Architecture

3.1.4 Model Training

In this project, the cassava leaf disease detection model was trained using MobileNet, ResNet50 and EfficientNet and also implemented for comparison. The training process for each model was carried out using the Adam optimizer and categorical cross-entropy loss. The models were trained over 75 epochs with early stopping in place to prevent overfitting by halting the training process when performance on the validation set ceased to improve.

The MobileNet model was designed with efficiency in mind. It uses depth-wise separable convolutions, reducing computational complexity without sacrificing accuracy, which is ideal for real-time applications with limited computational resources. During training, MobileNet benefited from batch normalization and ReLU6 activations to enhance stability and performance. The ResNet50 model, leveraging residual connections, was also trained in the same manner. These skip connections allow the model to better propagate gradients during training, enabling the effective training of deeper networks. This architecture has been widely used in image classification tasks for its ability to learn complex patterns while maintaining efficient gradient flow.EfficientNet was employed for its ability to scale efficiently across dimensions (depth, width, and resolution) while maintaining high performance with fewer parameters. It uses a compound scaling method and combines depth-wise separable convolutions and swish activation functions to optimize both accuracy and computational cost.

All three models were trained using the same training dataset, and their performance was evaluated based on accuracy on a separate test set. After training, the model with the highest accuracy was selected for deployment, ensuring the system's robustness and ability to make reliable predictions on new, unseen plant leaf images. This approach ensures the creation of an accurate and generalizable plant disease detection system capable of real-time deployment in agricultural settings.

3.1.5 Model Evaluation

In the cassava leaf disease detection project, the models (MobileNet, ResNet50, and EfficientNet) were evaluated using a test dataset to measure their classification accuracy. The evaluation focused on accuracy, which reflects the proportion of correct predictions, and a confusion matrix was used to identify misclassifications. Early stopping was employed to prevent overfitting by halting training when the validation performance stopped improving.The results indicated that ResNet50 outperformed the other models, achieving the highest accuracy with 94.75%. While MobileNet and EfficientNet also performed well, ResNet50 provided the most accurate and robust solution for cassava leaf disease detection. Despite the efficiency of MobileNet, ResNet50's deeper architecture and residual connections allowed it to capture more complex patterns in the leaf images, making it the most effective for this task. The findings suggest that ResNet50 was the most reliable model, ensuring precise identification of diseased cassava leaves for agricultural applications.

3.1.6 Deployment

For the deployment of the cassava leaf disease detection model, a Flask-based web application was developed to facilitate easy interaction with the trained model. The application provides a simple and intuitive interface where users can upload an image of a cassava leaf. Once the image is uploaded, the model processes it and predicts the disease affecting the leaf. The results are displayed on the webpage, ensuring that users can quickly access the diagnosis. The interface is styled using CSS, enhancing user experience with a visually appealing design.This approach ensures that the cassava leaf disease detection model is easily accessible in real-world scenarios, allowing farmers, researchers, or any user to check for diseases in cassava plants without needing advanced technical knowledge. The deployment makes the model practical for field use, helping in the early detection of diseases to improve agricultural productivity.

3.2 Software Requirements and Specifications

3.2.1 Operating System

The project is designed to run on modern operating systems like Windows 10/11 or Linux (Ubuntu). This ensures compatibility across different development environments and allows users to work with the software on their preferred platform, whether for local development or deployment. The system's flexibility supports a broad range of user preferences and makes it adaptable to various hardware configurations.

3.2.2 Python 3.11

Python is a high-level, general-purpose programming language known for its simplicity and readability, making it a popular choice for data science, artificial intelligence, and web development. The version used in this project is Python 3.11, which offers several improvements over previous versions, including enhanced performance and new features for developers. Python's versatility and rich ecosystem of libraries like TensorFlow, NumPy, and Pandas make it ideal for implementing complex machine learning and computer vision tasks such as those required in this project.

3.2.3 Visual Studio Code (VS Code)

The project utilizes Visual Studio Code (VSCode), a free, open-source code editor developed by Microsoft. VSCode is lightweight yet powerful, supporting a wide range of programming languages, including Python. Key features include:

- Built-in Git support for version control.
- Debugging tools to help identify and resolve code issues.
- An extension marketplace to enhance functionality with additional tools.
- Integrated terminal, which facilitates executing scripts directly from the editor.
- Code completion and Intelli-Sense for better productivity and reduced coding errors.

These features make VSCode an ideal choice for development, enabling developers to efficiently manage their codebase, troubleshoot issues, and streamline their workflow.

3.2.4 Jupyter Notebook

Jupyter Notebook provides an interactive development environment that is widely used in data science and machine learning. It allows you to create and share documents that contain live code, equations, visualizations, and narrative text. In your project, Jupyter Notebooks are used for exploring the dataset, experimenting with preprocessing techniques, and visualizing results. It's also useful for running and documenting the iterative process of model training.

Key Features:

- Cell-based execution for modular code.
- Supports rich text, LaTeX, and visualizations within the notebooks.
- Easy integration with Python libraries such as NumPy, Pandas, and Matplotlib.

3.2.5 Libraries

The following libraries are essential for the successful implementation of the project:

- **TensorFlow/Keras**: These libraries are crucial for building and training deep learning models like CNN, AlexNet, and ResNet50. TensorFlow provides powerful tools for model construction, training, and evaluation, while Keras offers high-level APIs for creating neural networks efficiently.
- **OpenCV**: Used for image processing tasks like reading, resizing, and augmenting images. It's especially useful for preparing the dataset before feeding it into the models.
- **NumPy**: This library is fundamental for handling numerical data, such as image arrays, and performing various operations like resizing and normalizing the image data.
- **Matplotlib**: Employed for visualizing model training progress, such as plotting training and validation accuracy and loss curves.
- **scikit-learn**: Provides tools for splitting the dataset into training and testing sets and evaluating model performance using metrics like accuracy.

3.2.6 Flask

Flask is a lightweight Python web framework used to create web applications. In this project, Flask is used to develop the web interface that allows users to upload images for demographic classification and age prediction. The framework's simplicity and flexibility make it an ideal choice for rapid prototyping and deployment. Flask supports the creation of RESTful APIs, which can be used to interact with the trained models and process predictions.

It also integrates well with front-end technologies such as HTML, CSS, and JavaScript, enabling the development of a user-friendly interface for displaying the results of demographic predictions.

3.2.7 Google Chrome

Google Chrome is a widely-used web browser known for its speed, security, and performance. It is essential for testing and deploying the web application developed in this project. Chrome provides excellent developer tools, including the JavaScript console, network monitoring, and performance analysis features. These tools assist developers in debugging issues with the web interface, optimizing performance, and ensuring smooth deployment of the application. Additionally, Chrome’s extension support and compatibility with modern web standards make it a suitable browser for testing the web application’s functionality.

CHAPTER 4

RESULTS AND DISCUSSION

In the cassava leaf disease detection project, the ResNet50 model outperformed both **MobileNet** and **EfficientNet**, achieving a validation accuracy of **94.75%**, compared to **88.87%** for MobileNet and **89.53%** for EfficientNet. The performance of ResNet50 demonstrates its effectiveness in extracting complex features from the cassava leaf images due to its deeper architecture and residual connections. While MobileNet and EfficientNet also showed competitive results, they faced challenges in feature extraction and model optimization, which affected their performance. The trained models were deployed in a Flask web interface, enabling users to upload images for disease predictions, making the tool practical and accessible for real-world agricultural applications.

4.1. Testing

Testing in web application development is a critical process that involves verifying the functionality, performance, security, and usability of the web application. It ensures the application works as intended and meets user requirements. Testing helps in identifying and addressing potential issues before the application is deployed to users, thereby improving the overall quality and user experience. For this project, various testing methodologies such as unit testing, integration testing, and functional testing will be implemented to ensure the reliability and performance of the web application.

4.1.1 Unit Testing

● Verify that individual components or functions of the code work as intended

● Developers write tests for specific functions or modules to ensure they produce the expected output.

4.1.2 Integration Testing

● Test the interaction between different components, modules, or services to ensure they work together correctly.

● Verify that integrated components collaborate as expected, checking for issues such as data flow and communication.

4.1.3 Functional Testing

● Validate that the web application's features and functions work as intended from an end-user perspective.

● Conduct tests to ensure that user interactions, data input, and expected outputs align with the specified requirements.

4.2 Output Screens and Results

The following section describes the output and how the system performs during testing, including the output screens displayed to the user:

4.2.1 Steps to Use the System

- Visit the website: Navigate to the hosted website demographic AI.
- Upload Image: The user can upload a facial image to be processed by the trained models.
- Results Displayed: Predictions for Gender, Race, and Age displayed in a readable format on the screen.

4.2.2 Output Screens

Web Interface:

The web interface, developed using Flask, provides a user-friendly platform for interacting with the leaf disease detection system. When an image of a cassava leaf is uploaded, the system processes it and displays the predicted disease status, whether healthy or affected by a specific disease alongside the processed image. Figure 4.1 (a) shows a screenshot of the user interface, illustrating the layout and features available to the user. Figure 4.1 (b) captures a sample prediction for a healthy leaf, demonstrating how the interface presents this information, while Figure 4.1 (c) displays a screenshot of a disease prediction, showing how the system communicates the disease label to the user. This interface is designed to streamline the detection process, making it accessible and efficient for end-users seeking real-time plant health insights.

Screenshots

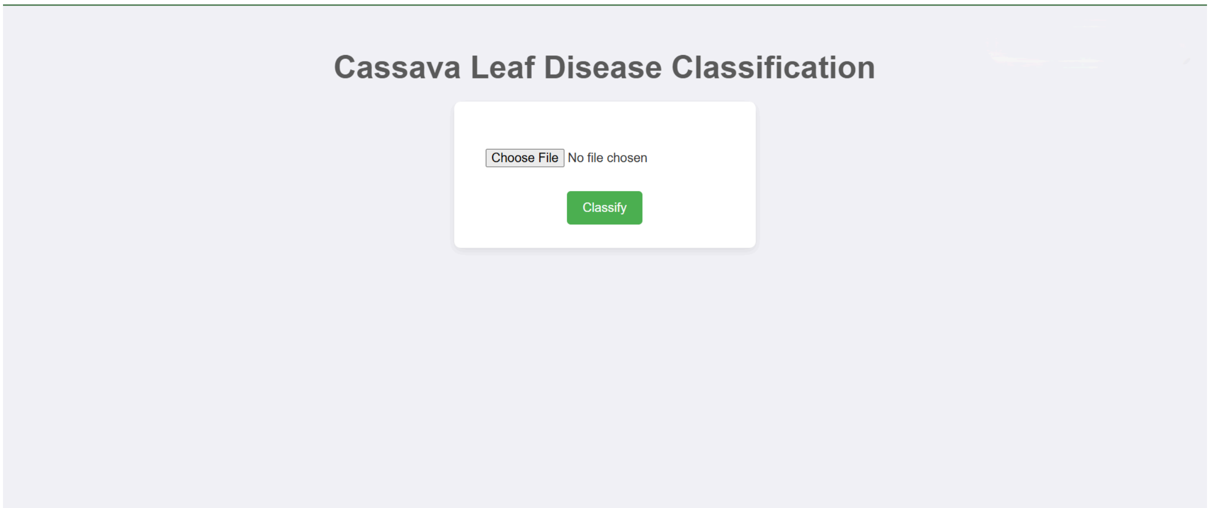

Figure 4.1(a): Screenshot of UI

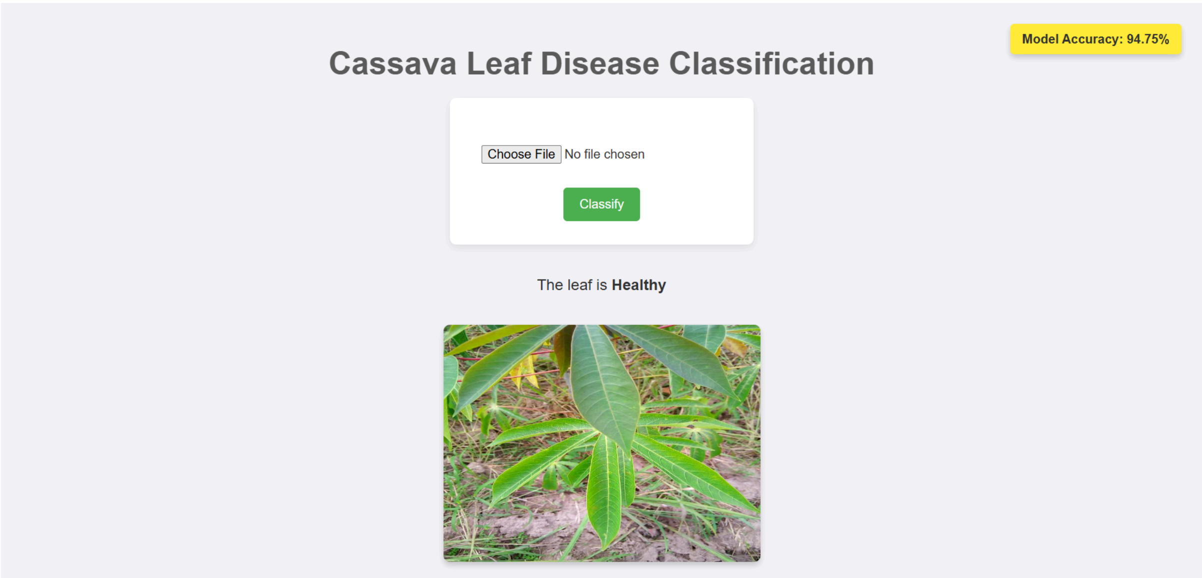

Figure 4.1(b): Screenshot of Healthy Prediction

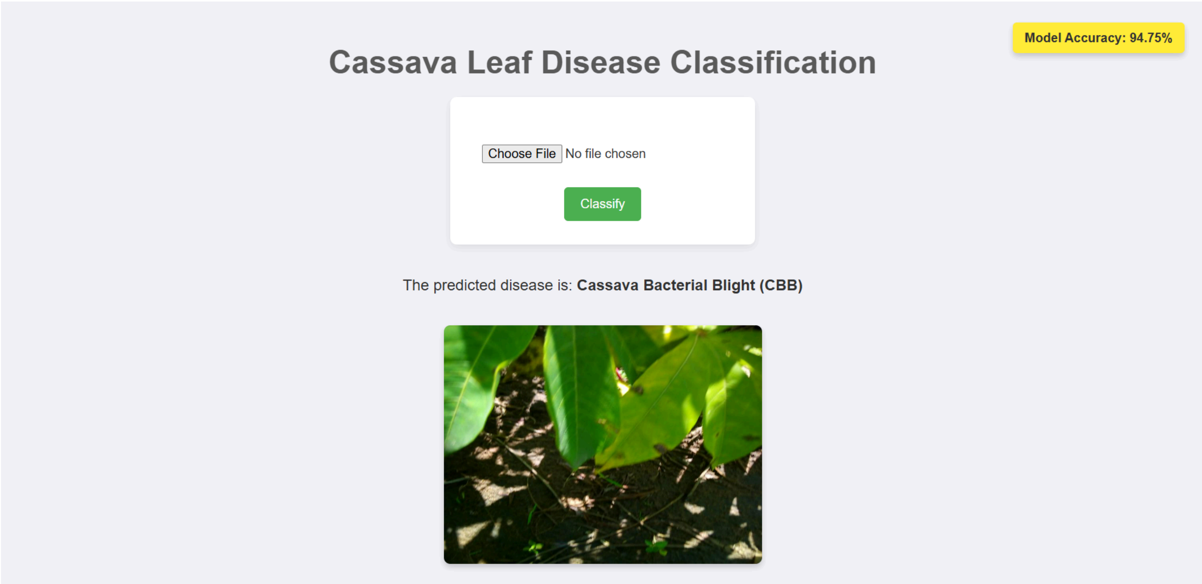

Figure 4.1(c): Screenshot of Disease Prediction

Figure 4.1(d): Screenshot of Disease Prediction

4.3 Results and Performance Evaluation

The performance of the cassava leaf disease detection model was evaluated using the validation dataset, comparing the effectiveness of ResNet50, MobileNet, and EfficientNet.

- ResNet50 Model: The ResNet50 model achieved the highest validation accuracy of 94.75%, making it the most effective model in detecting cassava leaf diseases. Its deep architecture and residual connections helped it to achieve superior performance in extracting relevant features from the dataset. Figure 4.2(a) shows the accuracy and loss plots for the ResNet50 model, illustrating its steady improvement over the training epochs.
- EfficientNet Model: The EfficientNet model achieved a validation accuracy of 89.53%, showing competitive performance. EfficientNet’s compound scaling helped it balance depth, width, and resolution for optimized accuracy and computational efficiency. Figure 4.2(b) presents the accuracy and loss plots for EfficientNet, demonstrating its effectiveness in disease detection.
- MobileNet Model: The MobileNet model achieved a validation accuracy of 88.87%, slightly lower than EfficientNet but still providing good results. Known for its lightweight design, MobileNet’s depthwise separable convolutions helped it perform well, especially in resource-constrained environments. Figure 4.2(c) shows the accuracy and loss plots for the MobileNet model, indicating steady progress in training.

All three models performed well, with the ResNet50 model emerging as the most effective for cassava leaf disease detection, achieving the highest validation accuracy of 94.75%. The EfficientNet and MobileNet models also demonstrated strong performance, with validation accuracies of 89.53% and 88.87%, respectively. These results emphasize the importance of selecting the appropriate architecture for tasks like plant disease detection, where accuracy and efficiency are crucial. ResNet50's deeper layers and residual connections allowed it to excel in feature extraction, while EfficientNet and MobileNet provided efficient solutions with slightly lower but still strong performance. Fine-tuning these pre-trained models was key to improving generalization and ensuring high performance, making them suitable for real-world agricultural applications.

4.3.1 Evaluation Curves

- **ResNet50**

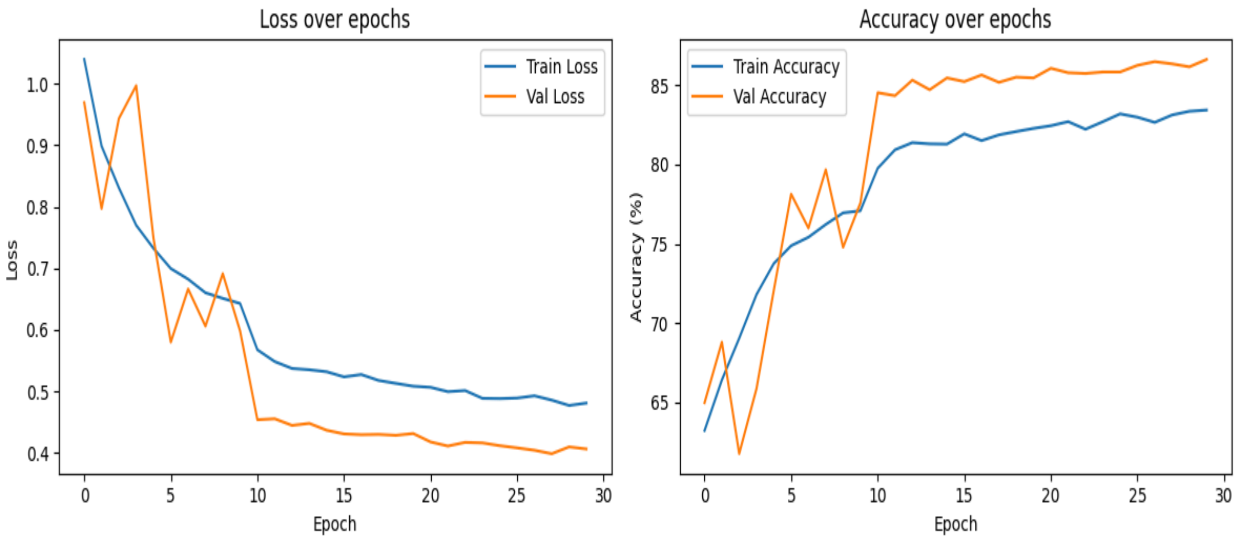

Figure 4.2(a): ResNet50 Accuracy Curve

- **EfficientNet**

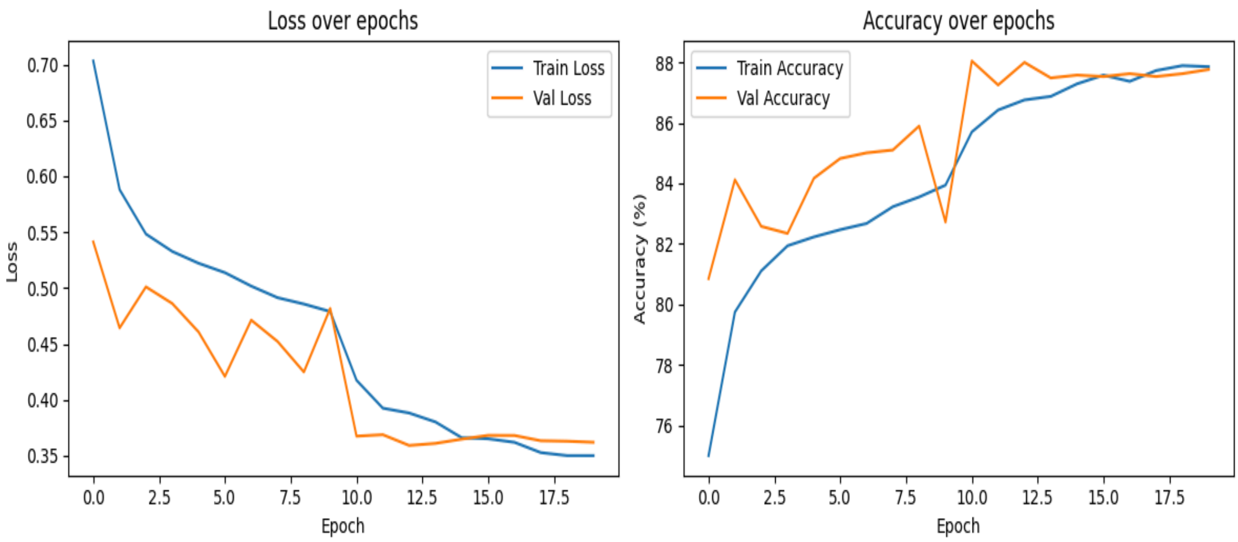

Figure 4.2(b): EfficientNet Accuracy Curve

- **MobileNet**

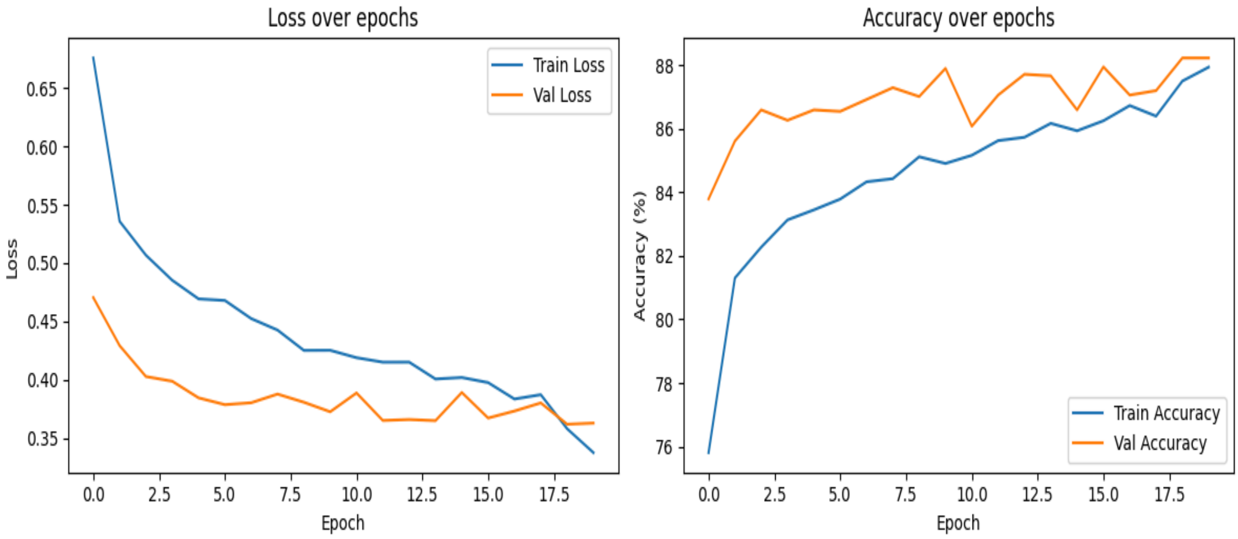

Figure 4.2(c):MobileNet Accuracy Curve

CHAPTER 5

CONCLUSION

In conclusion, the integration of deep learning models for cassava leaf disease detection presents a promising solution to address challenges in crop health monitoring. The application of Convolutional Neural Networks (CNNs) and other advanced models, such as MobileNet, ResNet50, and EfficientNet, has shown significant improvements in detection accuracy and efficiency. The models, particularly ResNet50, achieved high validation accuracy, demonstrating their suitability for real-time applications in resource-constrained environments. The deployment of these models in a user-friendly web interface makes the tool accessible to farmers with minimal technical expertise, enhancing the practicality of the solution for agricultural use. Furthermore, deep learning techniques offer scalability and adaptability, as evidenced by their ability to detect various plant diseases, including those affecting cassava. Leveraging pre-trained models and employing transfer learning has further optimized model performance, reducing training time and computational requirements, making these solutions more accessible for real-world applications. As deep learning continues to evolve, future enhancements, such as multi-disease detection, mobile deployment, and integration with weather and soil data, promise to further enhance the capabilities of these systems, enabling early disease detection and more efficient crop management.

- **Importance of Deep Learning in Cassava Leaf Disease Detection**: Using deep learning models like CNNs for cassava disease detection leverages their ability to learn complex patterns in images, a necessity for accurately identifying plant diseases from leaf images. Studies have shown that these models, including CNNs, perform well in plant disease classification due to their capacity for automatic feature extraction and pattern recognition, which is crucial for distinguishing between similar-looking symptoms of different diseases in plants.
- **Use of Advanced Models (MobileNet, ResNet50, EfficientNet)**: These models are noted for balancing high accuracy with efficiency. For instance, ResNet50's residual connections enable it to handle deep architectures without the problem of vanishing gradients, making it especially useful in detecting subtle differences in cassava leaf disease symptoms. Similarly, MobileNet’s lightweight design makes it feasible for real-time applications on devices with limited computational power, which is critical

for use in remote agricultural fields. EfficientNet's compound scaling optimizes depth, width, and resolution, offering a tailored solution for resource-constrained environments where cassava farmers may operate.

- **High Validation Accuracy and Real-Time Application**: The high validation accuracy of ResNet50 justifies its choice as a preferred model for this task. Models with high accuracy are crucial in agricultural applications where incorrect diagnosis could lead to ineffective treatments or loss of crops. Studies in precision agriculture confirm that deep learning models with robust performance and high accuracy can transform crop management practices, enhancing productivity and minimizing losses.
- **User-Friendly Web Interface for Practical Use**: Deploying the model via a web interface ensures that it is not only accessible to those without technical expertise but also practical for on-field use. Accessibility of such models is essential in agricultural communities where technology adoption can be limited. A study on the adoption of AI in agriculture has shown that tools designed with user-friendly interfaces significantly increase acceptance and usage among farmers.
- **Future Enhancements**: The mention of multi-disease detection, mobile deployment, and integration with environmental data highlights important future directions that can greatly enhance the model's impact. Including multiple data sources, such as weather and soil quality, could provide insights into environmental factors influencing disease spread, thus refining disease prediction models. Additionally, deploying lightweight models for mobile use could make this tool even more accessible, ensuring that farmers in remote areas can perform real-time diagnostics without relying on internet connectivity or high-power devices.

The research in this domain aligns with the growing trend of using AI and machine learning for precision agriculture, offering practical solutions that can significantly improve crop yields and reduce the impact of plant diseases.

5.1 Future Enhancements

For future enhancements of the Cassava Leaf Disease Detection System, several directions can be explored to improve its performance and applicability:

- **Mobile and Edge Computing:** Developing a lightweight version of the model for deployment on mobile devices can enable real-time detection in the field, benefiting farmers who need immediate insights.

- **Multi-Disease and Early Detection:** Extend the model to detect multiple diseases in a single leaf image and identify early-stage symptoms for proactive intervention.
- **Weather and Soil Data:** Including contextual data like weather, soil quality, and other environmental factors could help the model consider more factors that might contribute to disease spread and severity.
- **Expert-in-the-Loop Systems:** Allowing agricultural experts to provide feedback on model predictions can improve accuracy and help the model learn from corrections.

CHAPTER 6

REFERENCES

[1] Mingxin Liu, Haofeng Liang and Mingxin Hou, 2023 Research on cassava disease classification using the multi-scale fusion model based on EfficientNet and attention mechanism *Frontiers in Plant Science*

[2] H R Ayu1,* , A Surtono1 , and D K Apriyanto1. Deep learning for detection of cassava leaf disease *H R Ayu et al 2021 J. Phys.: Conf. Ser. 1751 012072*

[3] Huy-Tan Thai, Kim-Hung Le, Ngan Luu-Thuy Nguyen 2023 FormerLeaf: An efficient vision transformer for Cassava Leaf Disease detection *ScienceDirect*

[4] UmeshKumarLilhore 1 ,Agbotiname Lucky Imoize 2,3 Subhendu KumarPani 6 , Nitin Goyal 1 , Cheng-Chi Lee 4,5,* , Arun Kumar7 1 andChun-TaLi8 , Sarita Simaiya 1, 2022 Enhanced Convolutional Neural Network Model for Cassava Leaf Disease Identification and Classification. *MDPI Access.*