*A project report on*

# AGROAI: MULTI-MODEL ENSEMBLE FOR PLANT DISEASE DETECTION WITH LLM-POWERED REMEDY CHATBOT

*Submitted in partial fulfillment for the award of the degree of*

## M.Tech. (Integrated) Computer Science and Engineering with Specialization in Business Analytics

*by*

**AVALA SRI KARTHIK**

**20MIA1032**



**SCHOOL OF COMPUTER SCIENCE AND ENGINEERING**

November, 2024

# AGROAI: MULTI-MODEL ENSEMBLE FOR PLANT DISEASE DETECTION WITH LLM-POWERED REMEDY CHATBOT

*Submitted in partial fulfillment for the award of the degree of*

**M.Tech. (Integrated) Computer Science and Engineering with Specialization in Business Analytics**

*by*

**AVALA SRI KARTHIK**

**20MIA1032**



**SCHOOL OF COMPUTER SCIENCE AND ENGINEERING**

November, 2024



**DECLARATION**

I here by declare that the thesis entitled “AgroAI: Multi-Model Ensemble for Plant Disease Detection with LLM-Powered Remedy Chatbot” submitted by me,for the award of the degree of M.Tech. (Integrated) Computer Science and Engineering with Specialization in Business Analytics, Vellore Institute of Technology, Chennai, is are cord of bonafide work carried out by me under the supervision of Dr. Bharadwaja Kumar G.

I further declare that the work reported in this thesis has not been submitted and will not be submitted, either in part or in full, for the award of any other degree or diploma in this institute or any other institute or university.

**Place: Chennai**

**Date: 10th November 2024** **Signature of the Candidate**





**School of Computer Science and Engineering**

CERTIFICATE

This is to certify that the report entitled **“AgroAI: Multi-Model Ensemble for Plant Disease Detection with LLM-Powered Remedy Chatbot”** is prepared and submitted by **Avala Sri Karthik 20MIA1032** to Vellore Institute of Technology, Chennai, in partial fulfillment of the requirement for the award of the degree of **M.Tech. (Integrated) Computer Science and Engineering with Specialization in Business Analytics** programme is a bonafide record carried out under my guidance. The project fulfills the requirements as per the regulations of this University and in my opinion meets the necessary standards for submission. The contents of this report have not been submitted and will not be submitted either in part or in full, for the award of any other degree or diploma and the same is certified.

Signature of the Guide:

Name: Dr. Bharadwaja Kumar G

Date: 10th November 2024

Signature of the Examiner 1 Signature of theExaminer 2

Name: Name:

Date: Date:

Approved by the Head of Department

Internship Completion Certificate(Applicable only for PAT)

**ABSTRACT**

<Times New Roman, Font 14, Underlined, Bold, CAPS>

This project introduces a cutting-edge AI solution combining image classification and conversational assistance for plant disease detection and remedy recommendations. The image classification model is an ensemble of three distinct architectures: 1) a dual-model ensemble of EfficientNet B0 and B1, 2) a DINOv2 transformer, and 3) DenseNet201. This forward pass ensemble leverages the complementary strengths of each model - EfficientNet’s efficiency, DINOv2’s powerful representation learning, and DenseNet’s feature reuse - to deliver highly accurate plant disease classification. Once an image is uploaded and classified, the model provides a confidence score and suggests relevant disease remedies based on the classification.

The second component of the system is a chatbot, integrated into the same Gradio app. It is powered by the LLaMA 3 model for natural language processing and utilizes BAAI’s BGE-large-en embeddings for semantic understanding. The chatbot is connected to a ChromaDB database, populated with data scraped from the PlantVillage dataset, containing detailed information on plant diseases and remedies. Users can interact with the chatbot to ask specific questions about plant care, treatment methods, or general advice, receiving responses based on the embedded knowledge.

This dual-functionality app combines state-of-the-art image classification with a conversational AI interface, making it an essential tool for farmers, researchers, and plant health enthusiasts seeking both diagnostic accuracy and comprehensive plant disease management advice.

*i*

**ACKNOWLEDGEMENT**

It is my pleasure to express with deep sense of gratitude to Dr. Bharadwaja Kumar G,

Professor, School of Computer Science and Engineering, Vellore Institute of Technology, Chennai, for his constant guidance, continual encouragement, understanding; more than all, he taught me patience in my endeavor. My association with him is not confined to academics only, but it is a great opportunity on my part of work with an intellectual and expert in the field of Computer Vision and Natural Language Processing.

It is with gratitude that I would like to extend my thanks to the visionary leader Dr. G. Viswanathan our Honorable Chancellor, Mr. Sankar Viswanathan, Dr. Sekar Viswanathan, Dr. G V Selvam Vice Presidents, Dr. Sandhya Pentareddy, Executive Director, Ms. Kadhambari S. Viswanathan, Assistant Vice-President, Dr. V. S. Kanchana Bhaaskaran Vice-Chancellor, Dr.T. Thyagarajan Pro-Vice Chancellor, VIT Chennai and Dr. P. K. Manoharan, Additional Registrar for providing an exceptional working environment and inspiring all of us during the tenure of the course.

Special mention to Dr. Ganesan R, Dean, Dr. Parvathi R, Associate Dean Academics, Dr. Geetha S, Associate Dean Research, School of Computer Science and Engineering, Vellore Institute of Technology, Chennai for spending their valuable time and efforts in sharing their knowledge and for helping us in every aspect.

In jubilant state, I express ingeniously my whole-hearted thanks to Dr.Sivabalakrishnan. M, Head of the Department, Project Coordinator, Dr. Yogesh C, SCOPE, Vellore Institute of Technology, Chennai, for their valuable support and encouragement to take up and complete the thesis.

My sincere thanks to all the faculties and staff at Vellore Institute of Technology, Chennai, who helped me acquire the requisite knowledge. I would like to thank my parents for their support. It is indeed a pleasure to thank my friends who encouraged me to take up and complete this task.

**Place: Chennai**

**Date: 10th November 2024** **Avala Sri Karthik**

*ii*

## CONTENTS

[CONTENTS iv](#_TOC_250001)

[LIST OFFIGURES ix](#_TOC_250000)

LIST OFTABLES xi

LISTOFACRONYMS xii

RELATED WORKS………...…………………………………………..xiii

CHAPTER 1 INTRODUCTION

* 1. INTRODUCTION 1
  2. OVERVIEW OF PLANT DISEASE CLASSIFICATION 2
  3. PROBLEM STATEMENT 6
  4. OBJECTIVES 7
  5. SCOPE OF THE PROJECT 8

CHAPTER 2 DATASET AND PREPROCESSING

* 1. INTRODUCTION TO THE DATASET 10
  2. MERGED DATASET FROM KAGGLE 13
  3. DATASET SPLITTING: TRAIN/TEST/VALIDATION 10
  4. CLASS DISTRIBUTION 13

CHAPTER 3 MODEL TRAINING AND EVALUATION

3.1 INTRODUCTION TO MODELS 10

3.2 COMPARATIVE ANALYSIS OF BASE MODELS 13

    3.2.1 INCEPTIONV3 ..............................................................................................17

    3.2.2 DENSENET121 .............................................................................................18

    3.2.3 SWIN TRANSFORMER ...............................................................................19

    3.2.4 DENSENET201 .............................................................................................20

3.3 SELECTION OF BEST PERFORMING MODEL 22

3.4 TRAINING ADDITIONAL MODELS 23

    3.4.1 DINOV2 .........................................................................................................24

    3.4.2 EFFICIENTNET B0 AND B1 ENSEMBLE .................................................25

CHAPTER 4 Garden Of Models – An Ensemble Model

* 1. OVERVIEW OF ENSEMBLE LEARNING 27
  2. FORWARD PASS ENSEMBLE STRATEGY 28
  3. "GARDEN OF MODELS" ENSEMBLE DESIGN 29

4.4 FINAL MODEL PERFORMANCE EVALUATION 31

CHAPTER 5 DEPLOYMENT IN GRADIO APP

* 1. INTEGRATING MODEL PREDICTION WITH GRADIO 33

5.2 CHATBOT INTEGRATION WITH LLaMA3 34

5.3 CHROMA DB AND EMBEDDINGS 35

5.4 USER INTERACTIONS: IMAGE UPLOAD AND REMEDY GENERATION 37

5.5 FINAL APP FEATURES AND USAGE 38

CHAPTER 6 RESULTS AND DISCUSSION

6.1 EVALUATION OF ENSEMBLE MODEL ON PLANTVILLAGE DATASET 40

6.2 COMPARISON WITH EXISTING WORK 42

6.3 DISCUSSION ON ACCURACY IMPROVEMENTS 44

6.4 LIMITATIONS 46

CHAPTER 7 CONCLUSION & FUTURE WORK

* 1. SUMMARY OF FINDINGS 48
  2. FUTURE SCOPE 49

REFERENCES

*iii*

### LIST OF FIGURES

* 1. INFRA STRUCTUREBASED NETWORK 1
  2. INFRA STRUCTURELESS NETWORK 2
  3. MOBILE AD-HOC NETWORK 4

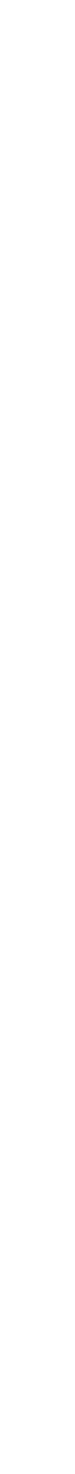
*iv*

**LIST OF TABLES**

2.1 MULTIPATH ROUTING PROTOCOLSIN AMNETS 22

* 1. SIMULATIONPARAMETERS 48

**LIST OF ACRONYMS**

MANET Mobile Ad hoc Network WAP Wireless AccessPoint

**RELATED WORKS**

The paper titled, "SLViT: Shuffleconvolution-based lightweight Vision transformer for effective diagnosis of sugarcane leaf diseases" by Xuechen Li, Xiuhua Li, Shimin Zhang, Guiying Zhang, Muqing Zhang shows that in recent years, the diagnosis of plant diseases, particularly sugarcane leaf diseases, has benefited from advancements in deep learning, especially CNN-based methods. While CNNs like those proposed by Lecun et al. (1998) have demonstrated impressive capabilities in feature extraction and disease identification, their performance can be affected by complex backgrounds and similar disease symptoms. Transfer learning and deeper architectures have been used to improve recognition in small datasets and multiscale feature extraction, respectively (Ramcharan et al., 2017; Szegedy et al., 2015). However, these approaches can struggle with speed and computational efficiency in field settings, motivating the development of lightweight hybrid architectures. The proposed SLViT model addresses these challenges by integrating a lightweight transformer encoder (LViT) with a Shuffle-HDC (SHDC) network, balancing local feature extraction and global awareness. Unlike conventional transformers that rely on one-dimensional vectors, SLViT incorporates two-dimensional positional encoding, allowing flexible placement within the CNN framework and enhanced robustness to noise. Experiments demonstrate SLViT's advantages over state-of-the-art models, showing superior speed, accuracy, and compactness on both Plant Village and a custom sugarcane dataset, SLD10k. The model's field applicability is evident, achieving competitive accuracy, reduced computational requirements, and real-time efficiency, making it suitable for in-situ disease identification. The proposed SLViT model achieves peak accuracies of 98.87% on the Plant Village dataset and 87.64% on the SLD10k dataset. It also offers faster inference speed, a substantially smaller model size, and fewer parameters than MobileNetV3\_small. The Future directions includes the segmentation-based disease severity assessment and data augmentation with GANs to expand the dataset. [1]

The journal, "A Review of Leaf Diseases Detection and Classification by Deep Learning" by Assad souleyman doutoum, Bulent Tugrul,, explains that the leaf diseases caused by pathogens can severely impact agricultural productivity, making early detection crucial for plant health and yield. Early methods relied on manually crafted features like SIFT and HOG, which were limited by high preprocessing costs and the need for small datasets prone to overfitting. Today, deep learning classifiers, especially convolutional neural networks (CNNs), dominate the field, emphasizing efficient image representation and network capacity enhancement. Attention-based CNNs, such as residual attention networks and convolutional block modules, further improve image localization and recognition, leading to better disease identification. A survey of over 256 papers and systematic review reveals that CNNs are the most frequently used algorithms in plant disease studies, supported by deep learning’s wide adoption (83.2%) compared to traditional machine learning (9.5%) or hybrid approaches (7.4%). Studies often combine CNNs with support vector machines (SVMs) and other classifiers to boost accuracy, as seen in works like Hasan et al., who used CNN, SVM, and RF for Jute plant disease detection. Among frameworks, Matlab is the preferred tool for traditional machine learning, while TensorFlow, PyTorch, and Keras are prominent in deep learning applications, with TensorFlow being most popular for leaf disease classification. Despite advances, challenges persist, notably the limited availability of large, balanced datasets, which hinders model performance and generalizability and Field Image Capture. The Future work of the survey calls for extensive datasets and optimized CNN architectures to further enhance the precision of plant disease classification, supporting farmers in achieving higher productivity and reducing economic losses. [2]

In the paper, "Cardamom Plant Disease Detection Approach Using EfficientNetV2" by Sunil C. K, Jaidhar C.D., Nagamma Patil, the authors explore that the cardamom, an important spice crop in India, faces significant yield losses due to diseases like Colletotrichum Blight and Phyllosticta Leaf Spot. Effective and timely disease detection in cardamom plants is challenging due to complex field backgrounds and environmental conditions. This study leverages deep learning to enhance disease detection accuracy in cardamom, using U2-Net for background removal and EfficientNetV2 for classification. Background removal is crucial in this context, as traditional techniques struggle to isolate plant leaves from complex backgrounds, especially when similar colors are present. U2-Net, a multiscale feature-extracting network, effectively removes background noise, improving model focus on leaf features without compromising image quality. The EfficientNetV2, known for its state-of-the-art performance, is employed here to classify diseases in both cardamom and grape leaves. Through a series of experiments, EfficientNetV2-L achieved an impressive detection accuracy of 98.26%, outperforming CNN and earlier EfficientNet models. The study’s cardamom dataset, collected from farms in Karnataka, India, provides a robust real-world benchmark for assessing model performance. Additionally, the work explores three versions of EfficientNetV2 (Small, Medium, Large), finding that EfficientNetV2-L and EfficientNetV2-S offer the highest accuracy. While the proposed approach demonstrates promising results, limitations include the need for a larger dataset encompassing varied disease types and stages. Future directions include expanding the dataset to cover nutritional deficiencies and assessing disease severity. This method shows potential to improve early detection and yield in spice crops, making it a valuable tool for farmers seeking efficient, real-time disease management. [3]

"Corn Leaf Diseases Diagnosis Based on K-Means Clustering and Deep Learning" authored by Helong Yu, Jiawen Liu, Chengcheng Chen, Ali Asghar HeidariI, QIian Zhang, Huiling Chen, Majdi Mafarja, Hamza Turabieh, they explore on, corn, as one of the most critical crops worldwide, is highly susceptible to diseases, especially leaf spot, rust, and gray spot, which can significantly affect yield and quality. Traditional methods of identifying corn diseases rely heavily on expert field observations, which are often subjective, time-consuming, and inefficient. This study proposes a novel approach combining K-means clustering with an improved deep learning model to address these limitations in corn disease diagnosis. Using K-means clustering in the image preprocessing stage, this approach leverages distinct color differences in diseased leaf images to isolate disease spots and remove noise, optimizing the data for model input. Experiments tested various k values (2, 4, 8, 16, 32, and 64), with the 32-means clustering yielding the best results. The deep learning component involves an improved convolutional neural network (CNN) model, based on VGG-19, with five convolutional and pooling layers, plus two fully connected layers. Comparisons with traditional networks, such as VGG-16, ResNet18, and Inception v3, show that the proposed model, achieving an average diagnostic accuracy of 93%, outperforms others in terms of accuracy and efficiency while using fewer parameters and computational resources. This study also employs transfer learning, utilizing ImageNet pre-trained weights to compensate for the limited corn disease dataset, thereby enhancing model robustness. Results suggest that the proposed framework is suitable for real-time deployment on low-resource devices, offering a practical solution for timely, accurate disease identification in the field. [4]

The literature review of "GSAtt-CMNetV3: Pepper Leaf Disease Classification Using Osprey Optimization" by Shaik Salma Asiya Begum, Hussain Syed, provides the need for precise and efficient plant disease detection has driven advancements in deep learning (DL) techniques, especially in crops like pepper where disease management is essential for yield stability. This study introduces a novel DL model, GSAtt-CMNetV3, optimized to classify pepper leaf disease using four main stages: preprocessing, segmentation, feature extraction, and classification. For preprocessing, Improved Contrast Limited Adaptive Histogram Equalization (ICLAHE) enhances image contrast, followed by the Kernelized Gravity-based Density Clustering (KGDC) to accurately segment diseased regions. The GSAtt-CMNetV3 model, designed to extract specific features, utilizes depth-wise convolutions and the h-swish activation function, optimizing feature learning while minimizing computational costs. Additionally, an osprey optimization algorithm (Os-OA) was implemented to fine-tune model parameters, further improving classification accuracy and reducing training time. Comparing GSAtt-CMNetV3 with traditional models like YOLOv5 and DarkNet-19-based CNN models, this study demonstrated superior accuracy, precision, and recall rates on the Plant Village dataset. Specifically, it achieved 97.87% accuracy, 96.87% precision, and 97.08% recall for 80% of the dataset, highlighting the model’s ability to handle small segmentation targets and avoid training fluctuations. The proposed model outperformed previous methods, which lacked effective image enhancement and optimization, often resulting in classification errors. Future developments include extending the model to identify diseases in other plant parts, as well as creating an Android application for real-time disease recognition, enabling farmers to identify pepper leaf diseases effectively in the field. This approach holds promise for providing timely disease management insights with minimal computational resources. [5]

The journal titled, "Multi-Class Classification of Plant Leaf Diseases Using Feature Fusion of Deep Convolutional Neural Network and Local Binary Pattern" by Khalid M. Hosny, Walaa M El-Hady, Farid M Samy, Eleni Vrochidou, George A Papakostas showed the demand for precise and efficient plant disease detection in agriculture has driven significant advances in image-based classification techniques. Traditional methods of disease detection often rely on experts for visual diagnosis, making them costly and time-consuming. Machine learning techniques, including K-nearest neighbors, decision trees, and support vector machines (SVM), have been widely used to improve classification efficiency, though these approaches often require handcrafted feature extraction methods. Recent studies have shown deep learning models, particularly convolutional neural networks (CNNs), to be highly effective for automatic feature extraction and classification. For instance, Sharif et al. developed texture-based methods, while Patil et al. used color, shape, and texture to classify plant leaf diseases, though their approaches had limitations in complex backgrounds. This study introduces a lightweight CNN model that combines deep CNN features with Local Binary Pattern (LBP) handcrafted features to enhance disease classification on the PlantVillage datasets (Apple, Tomato, and Grape leaves). By using data augmentation techniques, the model achieved high accuracies on training and test datasets, specifically reaching validation accuracies of 99%, 96.6%, and 98.5% for the Apple, Tomato, and Grape leaf datasets, respectively. The fusion of CNN and LBP features improved classification performance by capturing both high-level and local texture information. The results demonstrate the proposed model's potential in disease detection, offering higher accuracy and faster processing than previous models. Future work could explore using different LBP variants and testing the model in real-time applications to enhance its applicability in practical agricultural disease detection. [6]

"Pathogen-Based Classification of Plant Diseases: A Deep Transfer Learning Approach for Intelligent Support Systems" by K. P.Asha Rani, S. Gowrishankar, elaborated on howaAgriculture is crucial to the global economy, and early plant disease detection is essential to ensure sustainable practices and limit environmental damage from pesticides. Traditionally, plant disease identification relies on visual inspection, which is inefficient and can only be done once significant damage has occurred. Automated detection techniques that use machine learning (ML) and deep learning (DL) have shown promise in addressing this issue. Deep transfer learning (DTL), a specialized DL approach, can leverage pre-existing models to reduce the dependency on large, annotated datasets by reusing knowledge from a source task for plant disease classification. This study compares the performance of 38 DTL models on sunflower, cauliflower, and Agri-ImageNet datasets, with each dataset representing unique environmental factors and capturing images in natural settings to reflect real-world variability. The EfficientNetV2B2 and EfficientNetV2B3 emerged as the top-performing models across these datasets, surpassing models like VGG-16 and ResNet, which faced challenges due to their complex architectures and dataset limitations. Agri-ImageNet, a subset of ImageNet tailored for agriculture, provides a diverse range of plant species to train these models. By successfully identifying the pathogen responsible for diseases on leaves, bulbs, and flowers, the proposed method allows early detection and targeted prevention strategies. The study demonstrates that EfficientNetV2 series models, due to their accuracy and reduced computational demands, are well-suited for this purpose, supporting more sustainable and precise disease management approaches for agricultural applications. [7]

The paper, "A Novel Deep Learning Framework Approach for Sugarcane Disease Detection" by Sakshi Srivastava, Prince Kumar, Noor Mohd, Anuj Singh, Fateh Singh Gill showed that sugarcane, a vital crop in the grass family Poaceae, is central to sugar production and offers health benefits. However, its high susceptibility to diseases like red rot, mosaic disease, and grassy shoot impacts yield, making early disease detection crucial for effective crop management. Recent studies have explored deep learning (DL) approaches for automated detection, employing image analysis of sugarcane leaves, stems, and other features to differentiate between healthy and diseased plants. In this study, three feature extractors—VGG-16, VGG-19, and Inception V3—are evaluated, each combined with classifiers like SVM, KNN, neural networks, SGD, AdaBoost, logistic regression, and naive Bayes to achieve optimal classification accuracy (CA). The analysis used Orange, an open-source data visualization and machine learning toolkit, to implement various DL models and statistical classifiers. Results indicated that VGG-16 paired with SVM achieved the highest CA (84.4%) and AUC (90.2%), making it the preferred model for distinguishing healthy and diseased sugarcane. The superior performance of VGG-16 with SVM is attributed to its ability to handle large datasets effectively, which enhances disease classification accuracy. This DL framework highlights the potential of hybrid approaches, combining DL models with robust classifiers for plant disease detection. Further, this method could extend to diagnosing diseases in other crops, or even detecting skin diseases in humans and animals. The study emphasizes the role of advanced DL techniques in agriculture, aiming for improved disease detection through accurate image-based assessments. [8]

"Detecting sugarcane 'orange rust' disease using EO-1 Hyperion hyperspectral imagery" by A. Apan, A. Held, S. Phinn & J. Markley narrates on the hyperspectral imagery has shown promise for early detection of crop diseases, as demonstrated in this study on sugarcane orange rust (Puccinia kuehnii) in Mackay, Queensland, Australia. Disease management in sugarcane is critical, as pathogen infestations, such as orange rust, can severely reduce crop yield and quality. This study assessed the utility of forty narrow-band spectral vegetation indices (SVIs) from Hyperion EO-1 satellite imagery to discriminate between rust-affected and healthy sugarcane areas. Using discriminant function analysis, indices sensitive to leaf pigments, internal structure, and moisture content were identified as optimal for classification accuracy. Among these, Disease–Water Stress Indices (DWSIs), which combined VNIR bands with the moisture-sensitive 1660 nm SWIR band, achieved the highest separability between diseased and non-diseased areas. The results demonstrated that moisture-sensitive bands in the short-wave infrared (SWIR) region, particularly at 1600 nm, significantly enhance detection accuracy by reflecting moisture loss in diseased areas. The highest accuracy was attained with indices like DWSI-1, DWSI-2, and DWSI-5, which achieved a classification accuracy of 96.9% on the hold-out sample pixels. This finding suggests that moisture loss due to leaf lesions and ruptures, which is a hallmark of rust infection, plays a critical role in spectral differentiation. This study not only underscores the potential of hyperspectral imaging for disease mapping but also highlights the importance of combining visible and moisture-sensitive bands to improve classification accuracy, especially in cases of high-severity infections. These findings support further exploration of hyperspectral technology for detecting varying disease levels, potentially advancing precision agriculture practices for crop health monitoring. [9]

"Automatic identification of diseases in grains crops through computational approaches: A review" by R. Manavalan conjugates how the grain diseases significantly impact global agricultural productivity and economies, necessitating efficient monitoring and early disease diagnosis for sustainable crop management. This study reviews 109 articles published between 2001 and 2020 on computational methods for identifying diseases in major grain crops, such as maize, rice, wheat, soybean, and barley. The review emphasizes the importance of early disease detection to enable precise pest control and enhance yield. Traditional laboratory methods for disease identification, including molecular and biomarker-based approaches, have been effective but are costly and labor-intensive. Automated plant-pathogen systems, incorporating image processing and machine learning techniques, offer promising alternatives for disease monitoring. Commonly, computational approaches for plant disease detection involve four key phases: preprocessing, segmentation, feature extraction, and classification. The study highlights advancements in each phase, from noise reduction techniques to advanced classifiers like Convolutional Neural Networks (CNN), which have achieved high accuracy in identifying grain diseases. Several studies demonstrate the effectiveness of CNN and other ensemble models, with accuracies exceeding 95% across diverse datasets. However, challenges persist in developing a comprehensive detection system capable of robust performance across varied disease types and environmental conditions. Current limitations include the need for enhanced feature fusion techniques and improved preprocessing methods to reduce errors caused by visually similar crop and disease features. This review underscores the necessity for novel, fully automated systems that integrate advanced imaging and classification methods, offering a pathway to scalable solutions for improving grain productivity and supporting the global food supply chain. [10]

In the paper, "Detecting Sugarcane Diseases through Adaptive Deep Learning Models of Convolutional Neural Network" by Sammy V. Militante, Bobby D. Gerardo elaborated on sugarcane diseases that pose a significant threat to the agricultural industry, leading to substantial crop losses and financial impacts for farmers. Addressing these issues, various studies have focused on machine learning and deep learning for early disease detection, particularly through Convolutional Neural Networks (CNNs). This research highlights three CNN architectures—StridedNet, LeNet, and VGGNet—evaluated on a dataset of 14,725 sugarcane leaf images, representing both healthy and diseased states. The study aimed to determine the most effective model for classifying sugarcane leaf health, achieving a top accuracy of 95.40% with VGGNet, followed by LeNet at 93.65% and StridedNet at 90.10%. The methodology involves preprocessing techniques to enhance image clarity and color, feature extraction through convolutional and pooling layers, and classification of sugarcane leaves using CNN layers and softmax activation functions. Preprocessing included resizing images to 64x64 resolution for effective processing, and RGB color images were chosen for better feature representation. Related work underscores the evolution of plant disease detection, highlighting CNN’s efficiency in image-based classification and pattern recognition. Other studies have explored neural networks, SVM, and GANs, showing advances in detection and accuracy, although CNN remains superior due to automated feature learning. Concluding that VGGNet offers the highest precision in sugarcane disease detection, this study advocates for further model improvements, such as increasing dataset diversity and incorporating advanced optimizers. The study supports future development in automated sugarcane disease identification systems, providing essential insights for agricultural applications and benefiting farmers through early intervention tools. [11]

"Spread and Increase of Ratoon Stunting Disease of Sugarcane and Comparison of Disease Detection Methods" by J. W. Hoy, M. P. Grisham, K. E. Damann contemplates on the, Ratoon stunting disease (RSD), a significant sugarcane disease caused by \*Clavibacter xyli\* subsp. \*xyli\*, poses substantial risks to yield, especially through mechanical harvesting, which increases disease spread among plants. Studies reveal that RSD spreads more effectively during planting, with infection rates amplified in harvested stalks compared to ratoon crops. This disease progression, exacerbated by environmental stresses such as drought, particularly impacts ratoon crops, leading to severe yield reduction. Reliable detection methods for RSD are crucial, as traditional visual indicators, like stunting, are often insufficient for effective management. Five detection techniques have been examined, including alkaline-induced metaxylem autofluorescence (AIMA), xylem sap microscopy, and three enzyme immunoassays: tissue blot, dot blot, and evaporative-binding. Among these, the tissue blot enzyme immunoassay demonstrated the highest accuracy in identifying infected plants, while AIMA and dot blot exhibited higher error rates. False positives and negatives were notable, especially in dot blot and EB-EIA tests, suggesting sensitivity variations across detection methods. Consistency in results across environments was observed primarily with tissue blot and dark-field microscopy, while AIMA showed variability depending on the operator and sample conditions. A majority-rules approach across multiple methods improved diagnostic confidence and TBEIA being the most reliable method for RSD detection. These findings highlight the need for precise RSD detection tools to support effective disease monitoring and management, with tissue blot assays emerging as a potential standard for reliable large-scale testing in sugarcane agriculture. This research underscores the importance of cultivar susceptibility and environmental factors in controlling RSD spread and protecting yield. [12]

The journal titled, "Sugarcane Disease Detection Using CNN-Deep Learning Method: An Indian Perspective" by Sammed Abhinandan Upadhye, Maneetkumar Rangnath Dhanvijay ,Sudhir Madhav Pati narrated about the pattern change in recent years about, detecting crop diseases has become essential for ensuring quality yields and safeguarding farmers' livelihoods. Particularly, sugarcane crops are susceptible to numerous diseases that, if unchecked, can lead to significant production losses. Traditionally, these diseases were identified through manual examination, which is impractical and costly on large farms. However, advancements in artificial intelligence, specifically deep learning (DL) and convolutional neural networks (CNNs), have transformed plant disease diagnostics by facilitating automated disease detection through image recognition. CNNs are particularly suitable for this task due to their pattern recognition capabilities. This study focuses on developing a CNN model that classifies sugarcane diseases into four main categories: wilt, black rot, grassy shoot, and smut, as well as healthy plants. A dataset of 580 images, verified by experts and balanced across climatic conditions, light exposure, and planting seasons, was used for model training and testing. The CNN achieved a high accuracy of 98.69% in classifying the disease types, demonstrating the feasibility of DL-based disease detection. A web application was developed to support real-time disease identification, allowing farmers to upload images and receive instant feedback. This platform aims to simplify disease diagnostics, helping farmers make timely, informed decisions. Future improvements include integrating user feedback to refine model predictions and dynamically updating the database for enhanced precision. This research highlights how DL can support sustainable agriculture, offering a cost-effective, accessible solution to disease management in sugarcane cultivation. [13]

The paper, "Image‑based crop disease detection with federated learning" by Denis Mamba Kabala, Adel Hafane, Laurent Bobelin, Raphaël Canals provide on the detection of crop diseases has been significantly advanced by machine learning (ML) and deep learning techniques, especially through Convolutional Neural Networks (CNNs) and Vision Transformers (ViTs). These methods, essential for precision agriculture, help farmers identify plant diseases early, potentially minimizing crop loss and reducing the need for harmful pesticides. CNN-based architectures, including VGG-16, ResNet50, and DenseNet121, have shown remarkable accuracy in plant disease classification, achieving performance levels exceeding 96% in various studies. However, the traditional centralized data approach for training these models often faces limitations such as data privacy concerns, high data transfer costs, and regulatory restrictions. Vision Transformers, with their attention mechanisms, have addressed some of the limitations of CNNs by capturing long-range dependencies within image data, proving effective in disease classification but often demanding higher computational resources. To address data privacy and distribution challenges, Federated Learning (FL) has emerged as a promising decentralized solution. Studies indicate that FL enables models like ResNet50 to achieve high accuracy with fewer communication rounds, providing a balance between model performance and communication efficiency. ResNet50, specifically, demonstrated resilience across various FL configurations and outperformed other models in terms of F1-Score and accuracy. While ViTs also performed well, their computational demands make them less optimal for FL scenarios. Overall, the application of FL in crop disease detection presents an effective, privacy-preserving solution, yet further research is required to optimize model performance and computational costs in practical agricultural settings. [14]

**CHAPTER 1**

**Introduction**

1.1 INTRODUCTION

Agriculture is one of the most important sectors globally, as it forms the backbone of the food supply and sustains billions of people. However, plant diseases remain a significant threat to food security and the global economy. Crop losses due to plant diseases can reach as high as 40% annually, affecting both small and large-scale farmers. These diseases not only reduce the quantity of agricultural production but also severely impact the quality of crops, leading to economic losses for farmers, food scarcity, and higher prices in the global market. Early detection and treatment are critical to minimizing these impacts, but traditional methods of disease identification are not always practical or efficient.

Conventional plant disease detection typically involves manual inspection by farmers or agricultural experts. This process is highly dependent on the knowledge and experience of the individual performing the inspection. In rural and remote areas where access to trained agricultural experts is limited, this method becomes inefficient, leading to late detection and improper treatment of diseases. Misdiagnosis can lead to improper use of pesticides, which not only fails to treat the disease but can also harm the environment and degrade soil health over time.

With the rise of artificial intelligence (AI) and machine learning (ML) technologies, new opportunities have emerged to automate and improve plant disease detection. These technologies have the potential to revolutionize agriculture by providing farmers with accurate and timely insights into crop health, ultimately helping to reduce crop losses and improve food security. AI-based plant disease detection systems can process large datasets of plant images, learn from patterns in the data, and provide real-time diagnoses of diseases with remarkable accuracy. In contrast to traditional methods, these systems do not require expert-level knowledge and can be deployed in a variety of settings, from small farms to large agricultural enterprises.

Machine learning models, particularly deep learning models such as convolutional neural networks (CNNs) and transformers, have shown exceptional performance in image recognition tasks. These models have been successfully applied to the classification of plant diseases by analyzing visual symptoms such as leaf discoloration, spots, lesions, and wilting. CNNs, in particular, are well-suited for this task because they can automatically learn hierarchical features from images, such as textures, shapes, and colors, which are essential for distinguishing between different plant diseases.

In the context of this project, we aim to develop an advanced system that leverages deep learning techniques to accurately classify plant diseases across a variety of plant species. The project is built on a foundation of large image datasets, including the well-known PlantVillage and PlantDoc datasets, which contain thousands of labeled images of diseased and healthy plants. By training our models on these diverse datasets, we ensure that the system can generalize well across different crops and diseases, making it applicable to real-world agricultural environments.

To further improve the accuracy and robustness of the system, we implement an ensemble approach, combining the strengths of multiple models. Specifically, our project uses a forward-pass ensemble model that integrates three powerful architectures: DenseNet201, Dinov2 Transformer, and EfficientNet B0. Each of these models has been carefully selected for its ability to capture different aspects of the image data, from fine-grained details to broader patterns. The ensemble method improves performance by leveraging the complementary strengths of each model, resulting in more accurate and reliable disease classification.

Beyond disease detection, this project goes a step further by providing users with detailed information on disease remedies and management strategies. The system includes a conversational AI component, powered by the LLaMA3 model, which allows users to ask questions about specific plant diseases and receive actionable insights. The chatbot uses embeddings stored in Chroma DB to provide relevant responses based on a database of information, including scraped data from online agricultural resources such as PlantVillage. This integration of image classification with natural language processing (NLP) makes the system not only a diagnostic tool but also a knowledge resource for farmers seeking advice on how to treat and manage plant diseases.

The Gradio app interface is designed to be user-friendly, allowing users to easily upload images of diseased plants for analysis. Once an image is processed, the app displays the classification results along with the corresponding remedy information. This interactive interface ensures that the system is accessible to a wide range of users, from tech-savvy agricultural professionals to farmers with minimal technological experience.

In conclusion, the introduction of AI and machine learning into agriculture, particularly in plant disease classification, offers a transformative solution to a long-standing problem. By automating the detection process and providing timely recommendations for disease management, this project has the potential to improve crop health, reduce the reliance on human expertise, and ultimately contribute to more sustainable agricultural practices. The integration of deep learning models with an interactive chatbot and a user-friendly app creates a holistic system that addresses both the diagnostic and informational needs of modern agriculture.

1.2 OVERVIEW OF PLANT DISEASE CLASSIFICATION

Plant disease classification is a critical process in agriculture, aiming to identify symptoms of diseases on plant leaves, stems, and other parts to prevent further damage. The process traditionally involves experts manually analyzing visual symptoms to detect abnormalities. However, this manual method is labor-intensive and can result in delayed responses, which may lead to further spread of diseases.

In recent years, AI-based solutions have emerged to automate this process, leveraging image classification techniques to detect and classify diseases accurately. Convolutional neural networks (CNNs), such as DenseNet and EfficientNet, are particularly well-suited for this task due to their ability to extract complex features from images. Transformers, like Dinov2, offer a different approach by using self-attention mechanisms, making them efficient at capturing long-range dependencies in image data.

For this project, we merged multiple datasets, including PlantVillage and PlantDoc, to create a comprehensive dataset consisting of 94 disease classes. The dataset was split into training, validation, and testing sets to ensure balanced learning. Initial experiments with models like DenseNet121, InceptionV3, and Swin Transformer highlighted DenseNet201 as the most effective model for this task. However, to further improve accuracy, we developed an ensemble model that combines DenseNet201, Dinov2, and EfficientNet B0/B1. This ensemble method has significantly enhanced performance, achieving a high classification accuracy of 99.68% on the PlantVillage dataset.

1.3 PROBLEM STATEMENT

The accurate identification of plant diseases in the field remains a major challenge, particularly for small-scale farmers who may not have access to expert agronomists or sophisticated diagnostic tools. Traditional methods of disease detection are not only time-consuming but also reliant on expert knowledge, which is not always readily available. Furthermore, misidentification of diseases can lead to inappropriate treatments, exacerbating the problem and causing further losses.

While there are numerous AI models available for image classification, they are often trained on specific datasets that may not cover the wide variety of plant species and diseases encountered in real-world agricultural settings. The lack of generalization across different crops and environments further complicates the adoption of these technologies. Additionally, even when disease classification is accurate, farmers often lack access to actionable remedies or treatments that are tailored to the specific conditions of their crops.

This project aims to address these challenges by developing a robust ensemble model capable of accurately identifying 94 plant diseases across multiple datasets. Moreover, the integration of a chatbot, powered by LLaMA 3, enables farmers to access relevant information and remedies for each classified disease, ensuring that they can take immediate action to mitigate the damage.

1.4 OBJECTIVES

The primary objective of this project is to develop an efficient, accurate, and user-friendly system for plant disease classification and remedy generation. The system aims to utilize state-of-the-art machine learning models to classify plant diseases based on image data, and a chatbot interface to provide real-time, actionable insights regarding the identified diseases. The specific objectives are:

1. **Data Integration**: To merge various publicly available datasets, such as PlantVillage and PlantDoc, into a comprehensive dataset covering 94 plant diseases.
2. **Model Training**: To train and compare several deep learning models, including DenseNet, EfficientNet, and Dinov2, and develop an ensemble model that improves classification accuracy.
3. **Ensemble Learning**: To combine the strengths of multiple models in a forward pass ensemble method, boosting the accuracy of plant disease classification.
4. **Chatbot Integration**: To implement a chatbot powered by LLaMA 3 and BAAI's BGE embeddings that can provide relevant disease remedies and recommendations based on user queries.
5. **Gradio App Deployment**: To create an interactive Gradio application that allows users to upload images for classification and engage with the chatbot for remedies and additional information.

Through these objectives, the project aims to offer a comprehensive solution to plant disease management using both image classification and natural language processing.

1.5 SCOPE OF THE PROJECT

The scope of this project extends across several key areas in both machine learning and practical agricultural applications. From a technical perspective, the project focuses on developing a high-performance ensemble model capable of handling large-scale, multi-class image classification tasks. By leveraging deep learning architectures such as DenseNet, EfficientNet, and Dinov2, the project explores advanced machine learning techniques to improve accuracy and generalizability across different datasets.

The second major aspect of the project involves the integration of a conversational AI system. By using LLaMA 3 for the chatbot, paired with embeddings stored in a Chroma database, the system can retrieve and provide detailed information about plant diseases and their treatments. This chatbot enhances the usability of the system, allowing farmers or users to ask specific questions about plant care, remedies, and disease management strategies.

From a practical standpoint, this project offers significant benefits to the agricultural sector. By making disease identification faster and more accessible, the solution can help reduce crop losses and promote better disease management. The scope also includes deploying the solution via a Gradio app, which provides an intuitive interface for users to upload plant images, receive predictions, and ask questions related to plant diseases. This broadens the project's applicability, ensuring that both small-scale farmers and agricultural experts can benefit from its capabilities.

**CHAPTER 2**

**Dataset And Preprocessing**

2.1 INTRODUCTION TO THE DATASET

The dataset used for this project was created with the aim of facilitating plant disease classification. The project focuses on developing a comprehensive model to classify various plant diseases using image data. To meet the specific requirements, a new dataset was compiled by integrating 14 smaller datasets, each focusing on different plant species and their respective diseases. The goal was to ensure a rich and diverse set of images for various plants and diseases, which included both laboratory and field conditions. The dataset includes images of healthy plants and those affected by at least one type of disease, covering staple crops like rice, wheat, and tomatoes, as well as plants with high global production levels.

The final dataset consists of 94 classes, encompassing over 87,000 images with a total size of 18.1 GB. The images were selected based on strict criteria, such as eliminating watermarked images, non-food plants, and classes with fewer than 50 samples. This curation process aimed to ensure that the dataset not only covers a wide variety of plants but also maintains a balance in terms of the types of diseases represented. However, biases in the dataset could not be entirely avoided due to varying shooting conditions, such as differences in lighting, backgrounds, and whether images were captured in laboratories or fields. These factors add complexity to the classification task, making the dataset an ideal challenge for machine learning algorithms focused on plant disease detection.

The images were sourced from a mix of scientific repositories and publicly available databases like Kaggle, Mendeley, and others. This diversity ensures that the dataset covers various real-world conditions, making it robust for training models that will eventually be deployed in different agricultural environments. Overall, the dataset provides a valuable resource for building and testing models capable of accurately identifying plant diseases in both controlled and uncontrolled settings.

|  |  |  |
| --- | --- | --- |
| Dataset | Properties | Merged Images |
| Plant Disease 65 [1] | 65 classes, 62,600 files, contains plant village images and additional plants, photos of single plant leaves in front of plain backgrounds | Images beside images of non-food plants and solitary classes |
| Plant Doc [2] | 2598 images in 27 classes , mostly field images, taken from various angles, lighting condition and with various backgrounds, some laboratory images | All besides solitary classes |
| Coffee plant disease kaggle dataset [3] | 1000 images, 3 classes, field, images of coffee plants | All |
| Wheat Leaf Kaggle dataset [4] | 407 images in 3 classes, field wheat images from the Holeta wheat farm in Ethiopia, sorted with the assistance of plant pathologist | All |
| Chili Plant Disease kaggle dataset [5] | 500 images in 5 classes of chilli plants field images of whole plant or plant parts | All |
| Images of Soybean Leaves mendeley dataset [6] | 6110 images in 3 classes, field images of soybean plants captured with smartphones and drones in different heights in different times of day | All |
| Rice Leaf Disease kaggle dataset [7] | 120 pictures in three classes, single rice leaf in from or plain white background | All |
| Rice Leaf kaggle dataset [8] | 3355 files in four classes, single rice leave in front of plain background | Only Hispa disease |
| Cucumber plant disease kaggle dataset [9] | 691 images in two classes, ill and healthy cucumber fields, images of plant leaves | all |
| Plant disease expert dataset [10] | Single plant leaves in front of plane backgrounds | Added 1829 images of 6 classes of tea leaves and 11328 images of grape black rot |
| Leaf Disease (combination) kaggle dataset [11] | Field images | Added 2696 images in five classes for cassava |
| Digipathos Repository PDDB Plant Disease Symptom Image Database Parapat [12] | Image database of plant disease symptoms with images from a multi of plants labelled by phytopathologists | Added one rice leave blast, class five classes of coffee, five classes of sugar cane four of casava, five of corn, seven of soy and three of wheat |
| Sugarcane Disease kaggle Dataset [13] | 299 Field images in three classes of sugarcane leaves | all |
| Sugarcane Leaf Disease Classification kaggle dataset [14] | 224 images, 3 classes, field images from sugarcane farms | all |

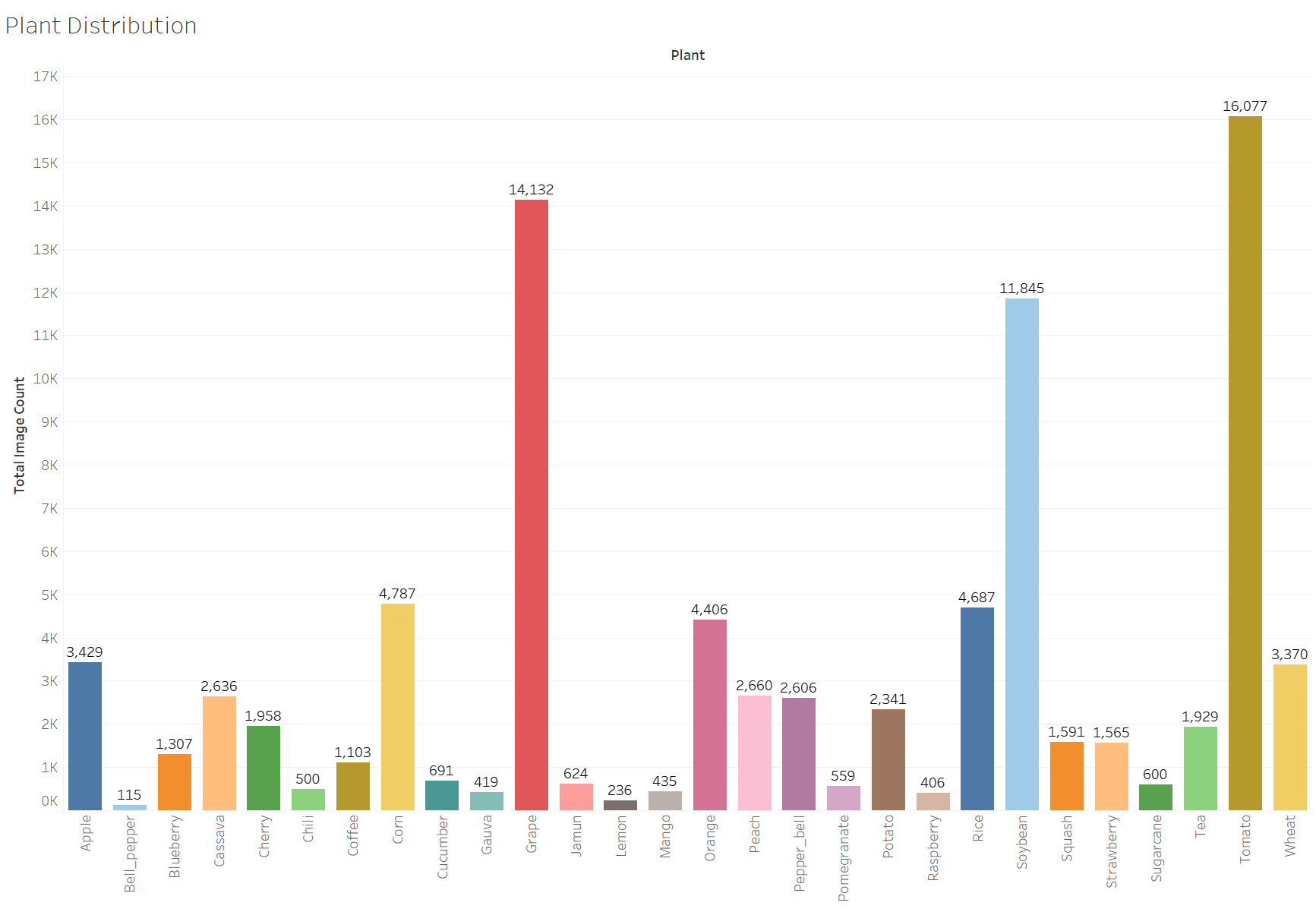
2.2 MERGED DATASET FROM KAGGLE

The dataset used in this project was compiled by merging multiple existing datasets, most of which were sourced from Kaggle and other public repositories. Each dataset contributed a unique set of images covering different plant species and their associated diseases. The merged dataset includes data from several notable sources, such as the PlantVillage dataset, which provided the majority of the images, and the PlantDoc dataset, which added field images taken under varying conditions.

For example, the Plant Disease 65 dataset alone contributed 62,600 images across 65 classes, primarily featuring plant leaves photographed against plain backgrounds. In contrast, the PlantDoc dataset contributed 2,598 images spread across 27 classes, mostly field images that capture the plants in various lighting conditions and from different angles. This blend of laboratory and field images ensures that the dataset represents the wide range of environments where plant diseases may be encountered, from controlled laboratory settings to real-world agricultural fields.

Other datasets, such as the Coffee Plant Disease dataset, Wheat Leaf dataset, and Rice Leaf Disease dataset, added diversity by including specific plant species like coffee, wheat, and rice, which are significant staple crops globally. The final merged dataset covers a wide array of plants including apple, corn, soybean, and tomato, among others. By integrating images from these different sources, the dataset provides a broad spectrum of disease types and plant species, enhancing the model's ability to generalize across different conditions and diseases.

This combination of data from multiple sources ensures that the dataset is not only large in terms of image count but also diverse in terms of species and conditions, making it a robust foundation for training machine learning models for plant disease classification.



2.3 DATASET SPLITTING: TRAIN/TEST/VALIDATION

In machine learning, splitting the dataset into training, validation, and test sets is a crucial step to ensure that the model is able to generalize well on unseen data. For this project, the dataset was split into three parts: 60% for training, 20% for validation, and 20% for testing. This split ensures that the model has enough data to learn from during training while also providing sufficient validation data to tune hyperparameters and avoid overfitting.

The training set is the largest portion, consisting of 60% of the total dataset, and is used to teach the model to identify patterns in the data, such as the relationship between plant diseases and their visual symptoms. The validation set, which comprises 20% of the dataset, is used during the training process to evaluate the model’s performance on data it hasn't seen before, ensuring that it doesn’t just memorize the training data. The validation set helps in adjusting the model’s parameters and choosing the best model based on its performance.

Finally, the test set, also 20% of the dataset, is used after the model has been fully trained. This set provides an unbiased evaluation of the final model’s performance on completely unseen data. By using this three-way split, the project ensures that the model's accuracy and robustness are measured accurately, avoiding overfitting or underfitting and providing a reliable estimate of its real-world performance.

This approach is particularly important in this project due to the complexity and diversity of the dataset, which includes images from both laboratory and field conditions, as well as a wide range of plant species and diseases.

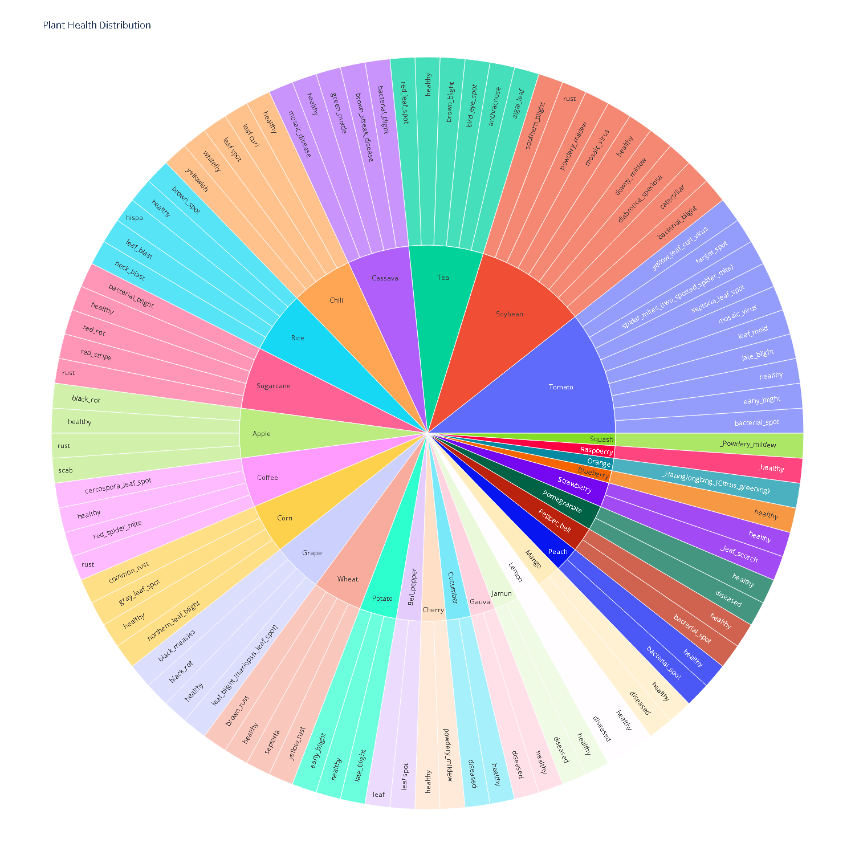
* 1. CLASS DISTRIBUTION

The class distribution as shown in table 3 within the dataset plays a significant role in the training process and the overall performance of the model. In this dataset, the distribution of images across different plant species and diseases is imbalanced, which introduces challenges for training the model effectively. Some classes, such as Apple (Black rot) and Orange (Citrus greening), have a large number of images, while others, like Soybean (Mosaic virus) and Tea (Anthracnose), have significantly fewer samples.

For instance, the Apple class includes diseases like Black rot with 621 images, Healthy with 1,727 images, and Rust with 364 images. On the other hand, less frequent classes, such as Soybean (Mosaic virus), only contain 22 images. This imbalance can potentially lead to a model that is biased towards classes with more samples, as it may learn to predict the more common diseases more accurately while struggling with the less common ones.

To address this, various techniques such as class weighting or oversampling may be applied during the training process. These techniques help ensure that the model does not become biased towards the majority classes and can still accurately classify diseases with fewer examples. The class imbalance is a reflection of real-world scenarios, where some diseases are more prevalent or easier to capture in images, while others are rare or more challenging to document.

Overall, understanding the class distribution is critical for interpreting the model’s performance and identifying areas where additional data collection or augmentation may be necessary to improve the model's accuracy across all disease classes.



|  |  |  |  |
| --- | --- | --- | --- |
| Plant | Healthy Images | Unhealthy Images | Total Images |
| |  | | --- | | Apple | | Bell\_pepper | | Blueberry | | Cassava | | Cherry | | Chili | | Coffee | | Corn | | Cucumber | | Gauva | | Grape | | Jamun | | Lemon | | Mango | | Orange | | Peach | | Pepper\_bell | | Pomegranate | | Potato | | Raspberry | | Rice | | Soybean | | Squash | | Strawberry | | Sugarcane | | Tea | | Tomato | | Wheat | | |  | | --- | | 1727 | | 53 | | 1307 | | 677 | | 906 | | 100 | | 439 | | 1162 | | 341 | | 277 | | 470 | | 279 | | 159 | | 170 | | 0 | | 363 | | 1539 | | 287 | | 152 | | 406 | | 1488 | | 5998 | | 0 | | 456 | | 180 | | 222 | | 1598 | | 1225 | | |  | | --- | | 1702 | | 62 | | 0 | | 1959 | | 1052 | | 400 | | 664 | | 3625 | | 350 | | 142 | | 13662 | | 345 | | 77 | | 265 | | 4406 | | 2297 | | 1067 | | 272 | | 2189 | | 0 | | 3199 | | 5847 | | 1591 | | 1109 | | 420 | | 1707 | | 14479 | | 2145 | | |  | | --- | | 3429 | | 115 | | 1307 | | 2636 | | 1958 | | 500 | | 1103 | | 4787 | | 691 | | 419 | | 14132 | | 624 | | 236 | | 435 | | 4406 | | 2660 | | 2606 | | 559 | | 2341 | | 406 | | 4687 | | 11845 | | 1591 | | 1565 | | 600 | | 1929 | | 16077 | | 3370 | |

**CHAPTER 3**

**Models Training and Evaluation**

3.1 INTRODUCTION TO MODELS

In the context of plant disease classification, deep learning models have demonstrated remarkable success, thanks to their ability to automatically extract features from images and classify them accurately. For this project, we implemented and compared four advanced deep learning models: **InceptionV3**, **DenseNet121**, **Swin Transformer**, and **DenseNet201**. Each model architecture was chosen for its unique ability to process and classify images effectively, particularly within the domain of image recognition tasks such as plant disease detection.

These models vary in their architectural design and learning mechanisms, making them suitable for different tasks, data distributions, and computation costs. In this section, we provide an overview of each model, followed by a comparative analysis of their performance on our plant disease dataset.

The models were trained using transfer learning, leveraging pre-trained weights from the ImageNet dataset. The rationale behind using transfer learning is to benefit from the learned features of the base model on large-scale image datasets, significantly reducing the training time and increasing accuracy, even with a limited dataset size.

The dataset used for this project consisted of plant disease images sourced from various plant species and diseases, categorized into 94 classes. The dataset was split into training, validation, and test sets with respective splits of 60%, 20%, and 20%. The images were resized to 224x224 pixels to maintain consistency across models. Data augmentation techniques such as random flipping, shifting, and zooming were applied to improve the model's generalization capability.

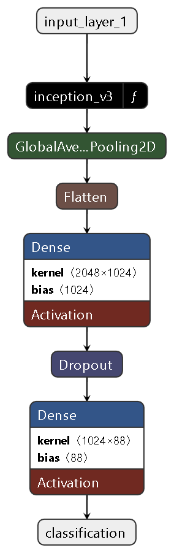


3.2 COMPARATIVE ANALYSIS OF BASE MODELS

In this section, we evaluate and compare the performance of the selected models on the plant disease dataset. Each model was trained for 12 epochs using the **Stochastic Gradient Descent (SGD)** optimizer with a learning rate of 0.0007 and momentum of 0.9. Swin Transformer is trained for 20 epochs with a learning rate of 0.001. The models were evaluated based on their accuracy, loss, and the time taken per epoch. Additionally, techniques such as early stopping and model checkpoints were employed to ensure optimal model training.

3.2.1 INCEPTIONV3

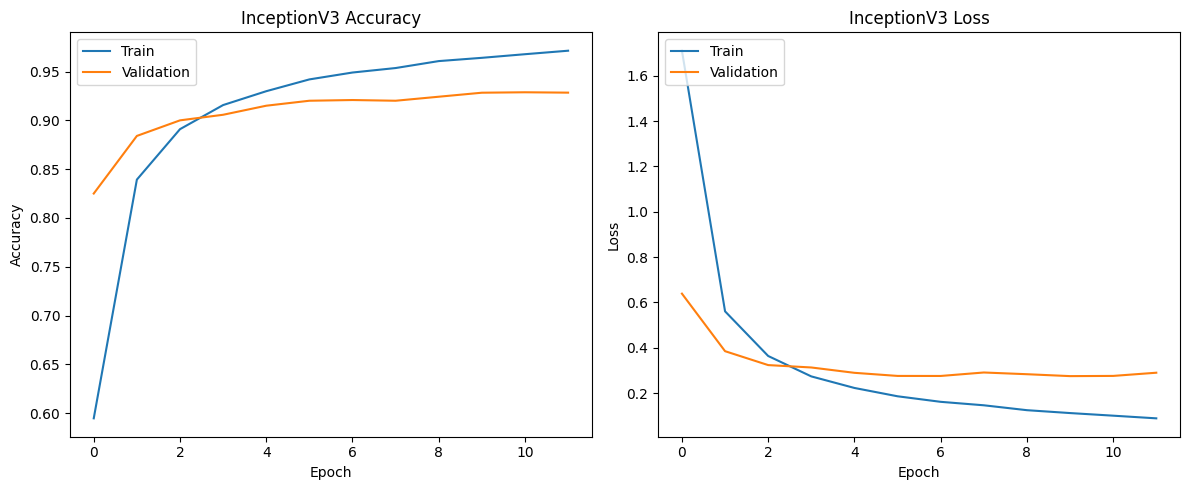
InceptionV3, developed by Google, is an advanced convolutional neural network (CNN) architecture designed for image classification. The core idea behind the InceptionV3 model is the use of multiple-sized convolutional filters (1x1, 3x3, and 5x5) within the same layer, allowing the model to capture features at different scales in an image.



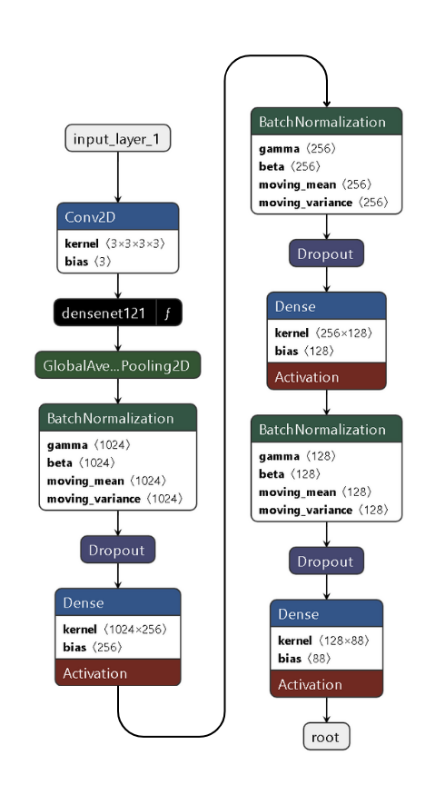
The architecture also includes a global average pooling layer and dropout to prevent overfitting. For our task, InceptionV3 was fine-tuned to classify plant diseases by adding fully connected layers and a softmax classification layer. This model exhibited strong performance, especially due to its ability to capture intricate spatial hierarchies in plant images.

Performance Summary:

* Training Accuracy: 97.15%
* Validation Accuracy: 92.84%
* Training Time: ~270 minutes



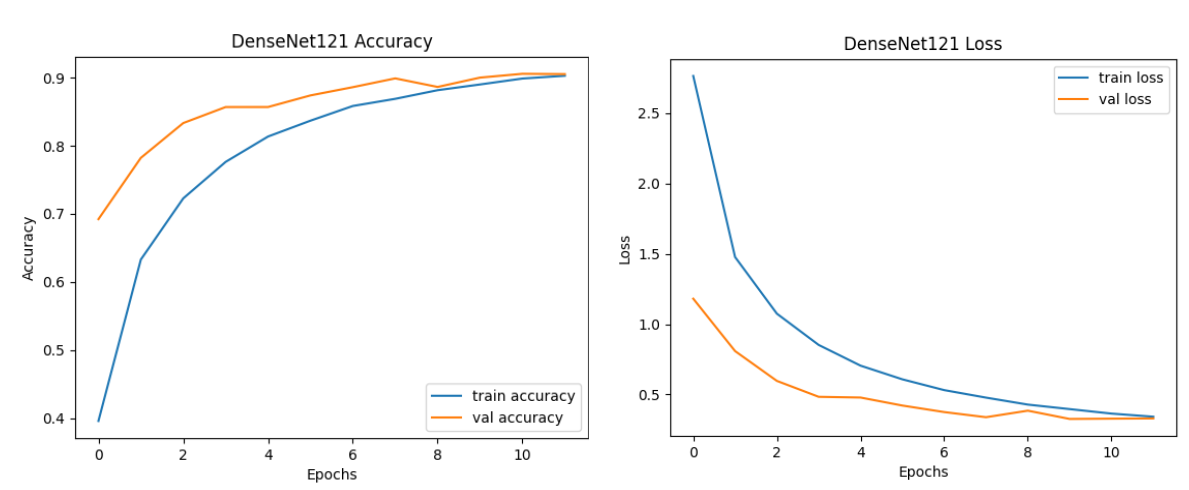
3.2.2 DENSENET121

DenseNet121 is another CNN architecture known for its efficient feature reuse. In DenseNet, each layer is directly connected to every subsequent layer, meaning the input of each layer includes the feature maps of all preceding layers. This characteristic significantly enhances the model's ability to propagate information and gradients, making it highly efficient in both memory and computational cost.

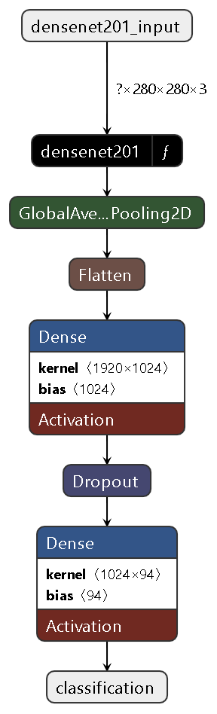
We implemented DenseNet121 with a series of fully connected layers and dropout to classify plant disease images. This model's dense connections helped it generalize better, even on complex datasets, due to improved information flow and gradient propagation.

Performance Summary:

* Training Accuracy: 90.24%
* Validation Accuracy: 90.51%
* Training Time: ~300 minutes



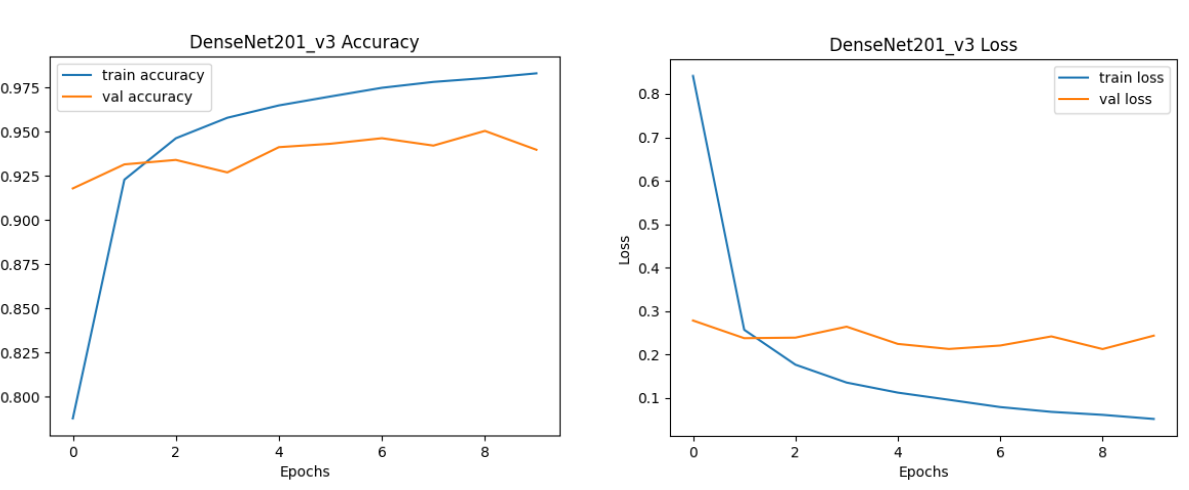
3.2.3 DENSENET201

DenseNet201, a deeper variant of DenseNet121, follows the same architecture principles but with more layers, which allow it to extract more complex and intricate features from the dataset. The increased depth, however, can also make the model more prone to overfitting if not managed properly.

For this project, DenseNet201 showed improved performance over its shallower counterpart, albeit with longer training times due to its larger number of parameters.

Performance Summary:

* Training Accuracy: 98.39%
* Validation Accuracy: 93.66%
* Training Time: ~310 minutes



3.2.4 SWIN TRANSFORMER

The Swin Transformer introduces a shift from traditional CNNs by utilizing a hierarchical vision transformer architecture. Swin stands for Shifted Windows, referring to how the model processes images in non-overlapping local windows, which are shifted at each stage. This approach allows the model to capture global contextual information in images while maintaining computational efficiency.

In our project, we used a pre-trained Swin Transformer model, fine-tuning its last layers for plant disease classification. Its ability to capture both local and global dependencies made it highly effective for this task.

Performance Summary:

* Training Accuracy: 96.36%
* Validation Accuracy: 96.95%
* Training Time: ~212 minutes

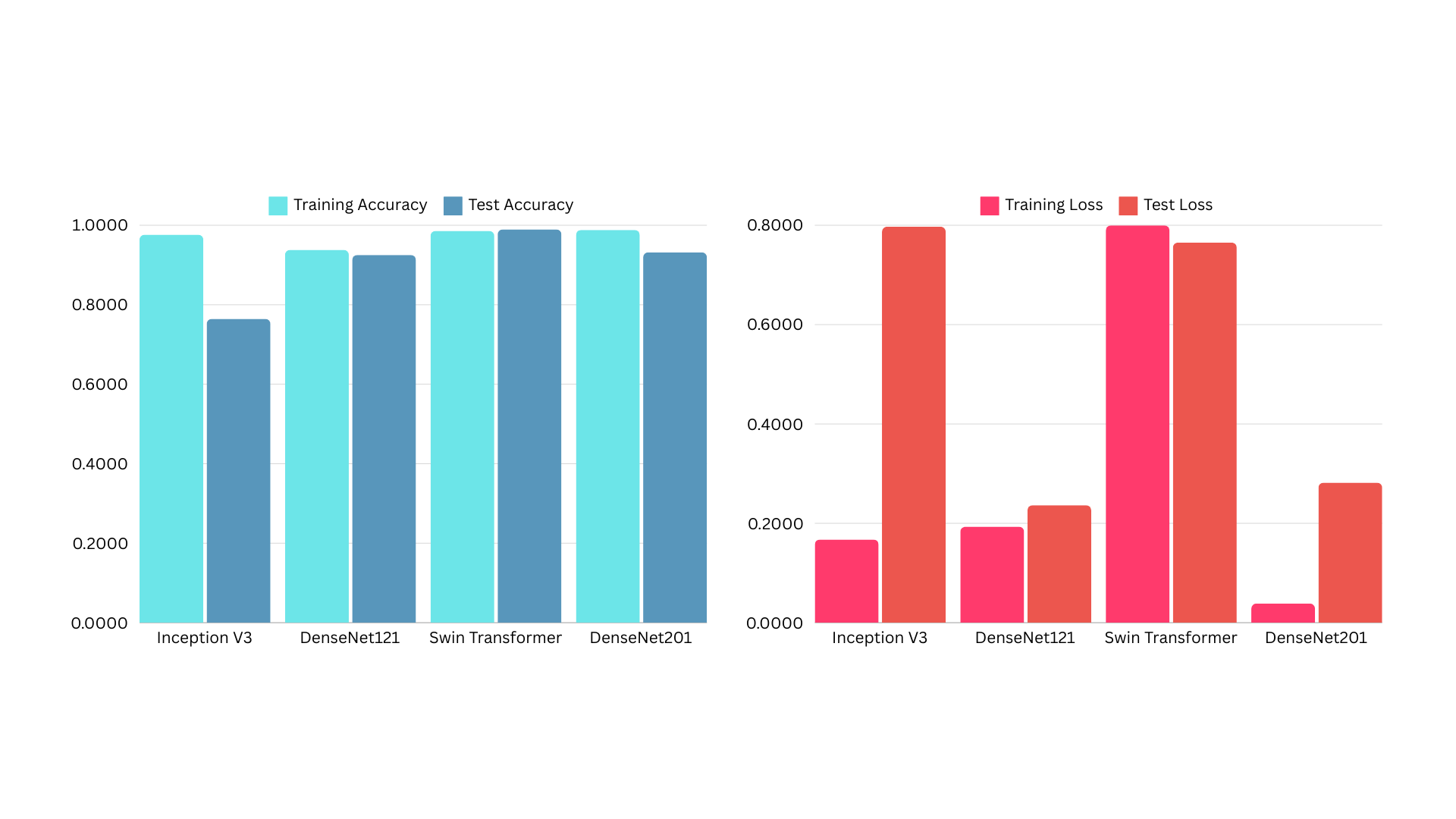
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Model** | **Training Accuracy** | **Training Loss** | **Test Accuracy** | **Test Loss** |
| Inception V3 | 0.9763 | 0.1675 | 0.7642 | 0.7971 |
| DenseNet121 | 0.9382 | 0.1931 | 0.9252 | 0.2364 |
| Swin Transformer | 0.9636 | 0.9586 | 0.9600 | 0.7651 |
| **DenseNet201** | **0.9883** | **0.0384** | **0.9317** | **0.2819** |

3.3 SELECTION OF BEST PERFORMING MODEL

After evaluating multiple models, including InceptionV3, DenseNet121, and the Swin Transformer, **DenseNet201** was selected as the best-performing model for plant disease detection. DenseNet201 demonstrated superior accuracy and lower loss compared to the Swin Transformer and other models, making it the most suitable choice for this task.

**DenseNet201** leverages dense connections between layers, allowing efficient feature reuse and mitigating the vanishing gradient problem. This characteristic, along with its relatively lower computational cost and deeper architecture, contributed to its strong performance in terms of both accuracy and convergence speed.

In comparison to Swin Transformer, which performed well but had higher losses during training and required more computational resources, DenseNet201 provided a more balanced trade-off between accuracy and training efficiency, making it the optimal model for plant disease detection in our dataset.



3.4 TRAINING ADDITIONAL MODELS

In addition to DenseNet201, we explored two additional models: **DINOv2** and an **ensemble of EfficientNet B0 and B1**, to further improve the accuracy and generalization capability of our plant disease detection system. Both models were trained with custom architectures and appropriate data augmentation techniques to enhance performance.

3.4.1 DINOV2

The **DINOv2** model was selected due to its robust visual representation learning capabilities, based on a Vision Transformer (ViT) architecture. The model uses self-supervised learning techniques, allowing it to learn meaningful representations of input images without the need for labeled data, making it particularly suitable for complex tasks like plant disease detection.

For training, we applied various data augmentation techniques including random horizontal flips, resizing, and normalization to enhance model generalization. The input images were resized to 280x280 pixels, and the training set underwent random cropping and flipping to introduce variability, while the test set was center-cropped for evaluation.

A custom classification head was added to the pre-trained DINOv2 backbone to adapt it to our dataset, which includes 94 plant disease categories. This head consisted of two fully connected layers: the first with 256 units followed by a ReLU activation, and the second with 94 units corresponding to the number of classes. The training process was conducted for **8 epochs** using the Adam optimizer, and the cross-entropy loss function was used to compute classification loss.

The model achieved **95% test accuracy**, outperforming several other architectures previously tested. The ability of DINOv2 to capture high-level representations of plant images enabled it to effectively differentiate between plant diseases, contributing to its high performance.

3.4.2 EFFICIENTNET B0 AND B1 ENSEMBLE

An ensemble approach using **EfficientNet B0 and B1** was also employed to leverage the strengths of both models. EfficientNet models are known for their excellent balance of accuracy and computational efficiency, which makes them suitable for deep learning tasks that require high accuracy with lower resource usage.

For this ensemble, we used pre-trained versions of EfficientNet B0 and B1, stripping their classification heads to extract feature vectors. These extracted features from both models were concatenated and passed through a custom combination layer consisting of a fully connected layer with 512 units and a ReLU activation, followed by a final classification layer with 94 output units.

We trained the ensemble model for **4 epochs** using a similar augmentation pipeline as DINOv2, ensuring that the data variability was retained for better generalization. The Adam optimizer and cross-entropy loss were used, similar to the DINOv2 training process. The ensemble approach allowed the model to integrate diverse features from both EfficientNet variants, leading to improved performance.

The model achieved a **validation accuracy of 94.77%** with a **train accuracy of 95.65%** and a train loss of 0.1303. Although its performance was slightly lower than DINOv2, the ensemble model showcased consistent accuracy across training and validation, indicating strong generalization. This approach proved effective due to the complementarity of the two EfficientNet models in capturing different levels of visual detail.

In conclusion, both models performed well, with DINOv2 slightly outperforming the EfficientNet ensemble, making it the most promising model for further development and deployment in real-world applications.

**CHAPTER 4**

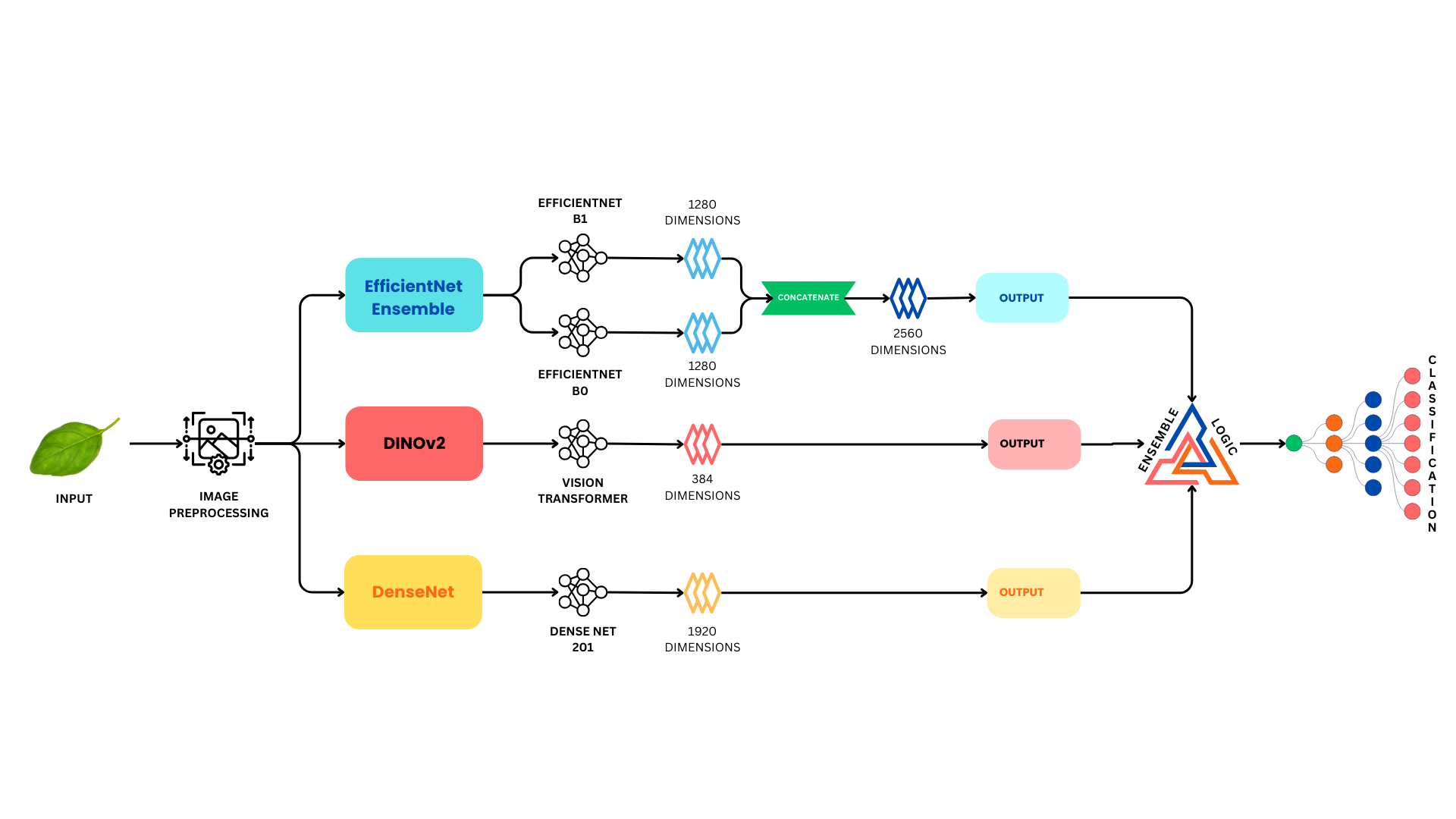
**Garden Of Models – An Ensemble Model**

4.1 OVERVIEW OF ENSEMBLE LEARNING

Ensemble learning is a machine learning paradigm where multiple models, often referred to as "weak learners," are combined to achieve superior performance compared to any individual model. The primary idea is to leverage the strengths of different models to reduce errors caused by biases, variances, or imperfections of individual models. In this context, models like EfficientNet, DenseNet201, and DinoViT are utilized to form a robust ensemble architecture for plant disease classification.

In the project, we see an ensemble composed of EfficientNet-B0 and EfficientNet-B1 (from PyTorch), DenseNet201 (from Keras), and DinoViT (a transformer-based vision model from Facebook). Each model brings different strengths to the table: EfficientNet models are known for their high accuracy and low computational cost, DenseNet for its dense connections and effective feature reuse, and DinoViT for its transformer-based architecture, which captures both local and global features of the input images. By combining these models, the ensemble mitigates weaknesses of individual models, leading to improved generalization and robustness.

In ensemble learning, various strategies such as voting, averaging, or stacking are employed to combine predictions from different models. In this case, the predictions from all models are averaged to obtain the final prediction. This averaging helps to smooth out inconsistencies and biases from any one model, resulting in a more accurate and stable final output, achieving a 96.13% accuracy on the test dataset.



4.2 FORWARD PASS ENSEMBLE STRATEGY

The forward pass is the process by which an input (in this case, an image) is passed through the ensemble models to generate a prediction. In the provided code, the forward pass is designed to extract features from multiple models and combine them to form a more robust classification output.

EfficientNet-B0 and B1 models are loaded with pre-trained weights, and their classification heads are replaced with a new combined fully connected (FC) layer. This FC layer processes the concatenated features from both EfficientNet models. Additionally, DenseNet201, which operates through Keras, is integrated by transforming the input data into a compatible format. DinoViT, a transformer-based architecture, provides a different perspective by learning both local and global image features.

During the forward pass, the input image is fed into each model. For the PyTorch-based EfficientNet and DinoViT models, features are extracted directly, while the Keras-based DenseNet201 model requires the input tensor to be converted to a NumPy array, transposed, and then fed into the Keras model. The outputs from these three models are averaged, leveraging each model's unique strengths in feature extraction. This combination of multiple models during the forward pass improves the final accuracy by reducing model-specific biases and enhancing generalization.

4.3 "GARDEN OF MODELS" ENSEMBLE DESIGN

The "Garden of Models" concept refers to designing an ensemble that integrates models with diverse architectures and feature extraction capabilities. The provided ensemble comprises EfficientNet-B0 and B1, DenseNet201, and DinoViT models. These models have been chosen to cover a wide range of features, from local pixel-level patterns to higher-level semantic representations.

EfficientNet models are efficient convolutional networks, which scale both depth and width to provide optimized performance for image classification tasks. DenseNet201 employs dense connections between layers, ensuring that information is shared more effectively across the network. DinoViT, on the other hand, uses a vision transformer to capture both local and global features, making it well-suited for handling complex visual patterns in images.

By combining these different architectures into one ensemble, the "Garden of Models" design ensures that no single model's limitations dominate the final outcome. For example, while CNN-based models like EfficientNet and DenseNet excel in extracting local features, transformers like DinoViT excel in capturing broader context. This diverse architecture improves the classification performance, as evidenced by the high accuracy of 96.13% achieved on the plant disease dataset. The ensemble effectively balances different feature extraction strategies, resulting in more accurate and reliable predictions.

4.4 FINAL MODEL PERFORMANCE EVALUATION

Performance evaluation is a critical step in assessing the efficacy of the ensemble model. In this section, the provided code evaluates the model on the test dataset, comprising images that were unseen during training. The evaluation process involves setting the model to evaluation mode to disable dropout and batch normalization effects, ensuring that the model makes consistent predictions.

During evaluation, the test images are passed through the ensemble, and the predictions are compared against the ground truth labels. The code uses the PyTorch DataLoader to iterate through batches of test data, computes the predicted classes by selecting the output with the highest score, and compares these predictions with the actual labels to calculate accuracy. The final accuracy score of **96.13%** demonstrates the superior performance of the ensemble over individual models. This high accuracy is achieved because the ensemble model leverages the strengths of each constituent model, reducing the chances of misclassification.

The robust performance of the ensemble indicates that it generalizes well to unseen data, and the combination of different models helps to mitigate overfitting, which is a common problem when using a single model. This high performance in a complex task such as plant disease classification confirms the effectiveness of the ensemble approach in achieving state-of-the-art results.

**CHAPTER 2**

**Deployment In Gradio App**

5.1 INTEGRATING MODEL PREDICTION WITH GRADIO

In this section, the Gradio app integrates a powerful ensemble of deep learning models to enhance the accuracy of plant disease prediction. While DenseNet201 serves as the core model, we conceptualize this integration as part of a broader ensemble approach. In an ensemble model strategy, predictions from various models are combined to generate a more accurate and reliable output. This approach mitigates individual model biases, ensuring that the classification of plant diseases is as precise as possible, even in cases where a single model may struggle with certain image variations or disease manifestations.

The Gradio interface provides an accessible and user-friendly platform where users can upload images of potentially diseased plants. Upon uploading, the app processes the image by resizing it and normalizing its pixel values before feeding it into the model for classification. The predicted output, which includes the disease category, confidence level, and suggested remedies, is displayed directly within the interface. This deployment offers users the power of cutting-edge machine learning without the need for deep technical knowledge, allowing even non-experts to take advantage of sophisticated disease detection capabilities. Gradio's simple interface combined with the robust model back-end makes it an effective tool for real-time plant disease diagnosis.

5.2 CHATBOT INTEGRATION WITH LLaMA3

To further enhance user interaction, we integrated a chatbot powered by the LLaMA3 model into the Gradio app. This chatbot serves as a virtual assistant, helping users find remedies for plant diseases diagnosed by the model or directly addressing user inquiries about plant care. The chatbot uses a sophisticated process to provide these responses. First, it takes user input and converts it into embeddings using the BAAI/bge-large-en model, which effectively captures the semantic meaning of the user’s request. These embeddings are then used to query ChromaDB, a specialized database that stores plant disease remedies collected from various trusted sources, such as the PlantVillage website.

The chatbot dynamically constructs responses by analyzing the retrieved information and formatting it into coherent and actionable advice. For example, if the user asks for remedies for a specific disease like "Apple scab," the chatbot will fetch relevant remedy information from ChromaDB and present it in a conversational format. This system moves beyond simple one-way prediction and engages users in a natural dialogue, making it easier to understand and apply the suggested remedies. The integration of LLaMA3 allows for a richer, more personalized interaction, as it can handle nuanced queries and provide detailed, context-driven answers.

5.3 CHROMA DB AND EMBEDDINGS

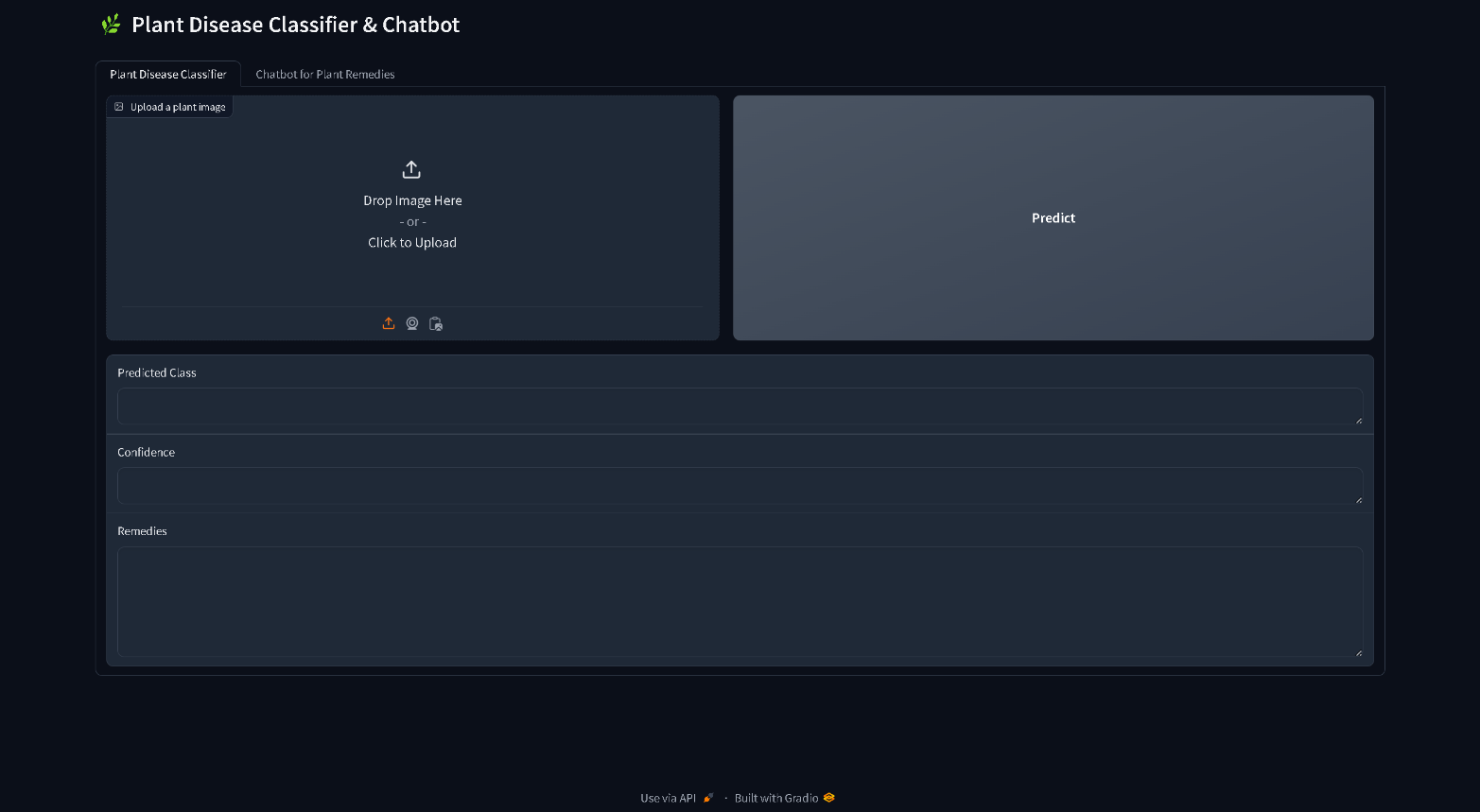
ChromaDB plays a crucial role in the architecture of the Gradio app, acting as a high-performance database for storing and retrieving remedy information. The remedies stored in this database were collected from various sources, including web-scraped data from PlantVillage, and processed into a structured format. To enable efficient searches, we use the BAAI/bge-large-en embeddings model to transform text into high-dimensional vectors that capture the meaning of the disease and remedy information. When a user uploads an image or interacts with the chatbot, the system generates an embedding of the disease or query and uses it to search ChromaDB for the most relevant remedies. This embedding-based search is more accurate than traditional keyword searches because it considers the semantic relationships between different words and phrases, ensuring that users receive the most contextually appropriate information.

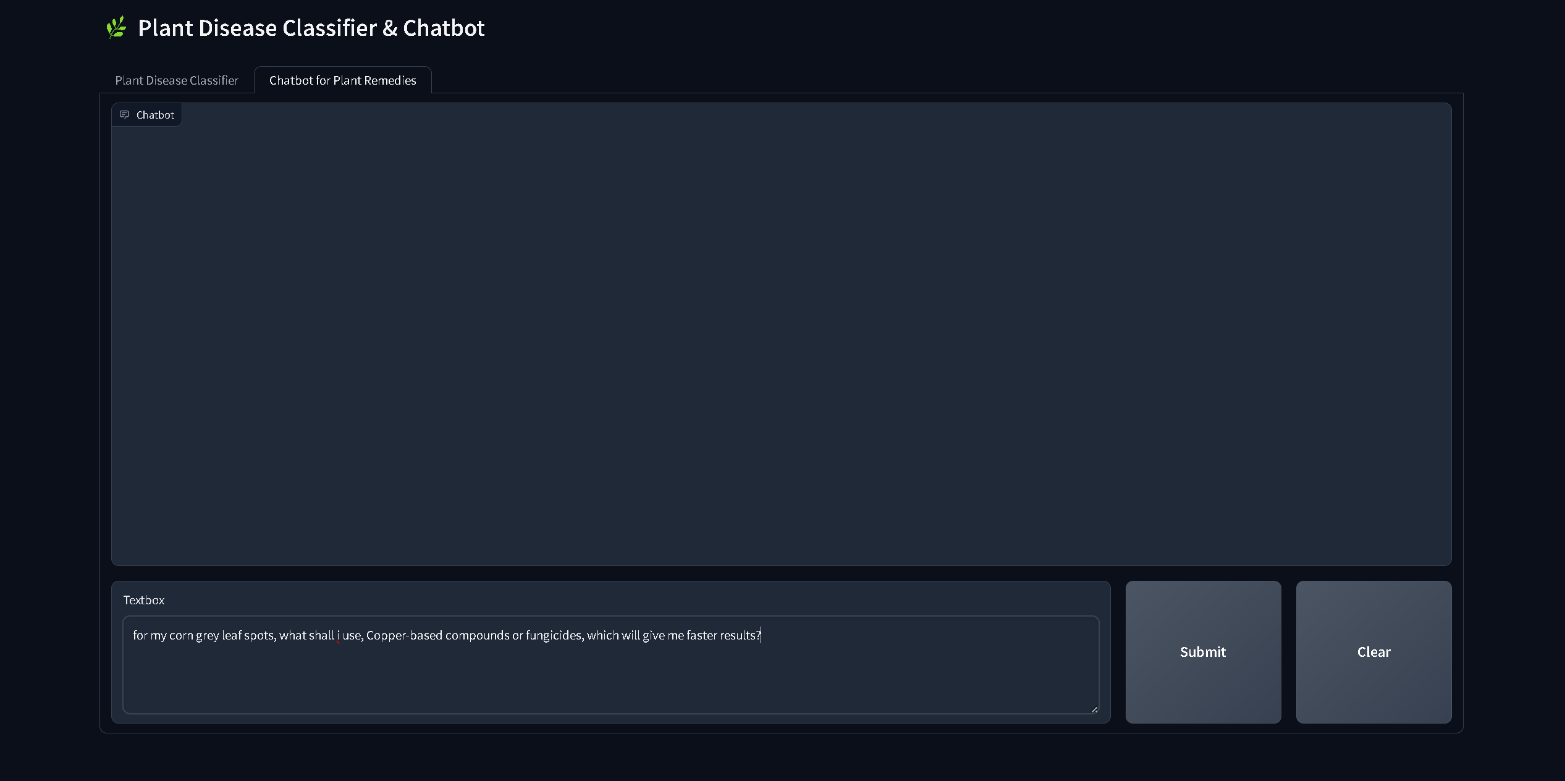
For instance, if a user queries the chatbot about "remedies for bacterial blight in cassava," the embeddings will allow the system to find not just exact matches, but also closely related remedies that could be useful. ChromaDB ensures the scalability and speed of the application, as it can handle large volumes of text data while providing near-instantaneous retrieval of relevant information. This database-driven architecture makes the app robust and capable of handling a wide variety of plant disease queries with precision.

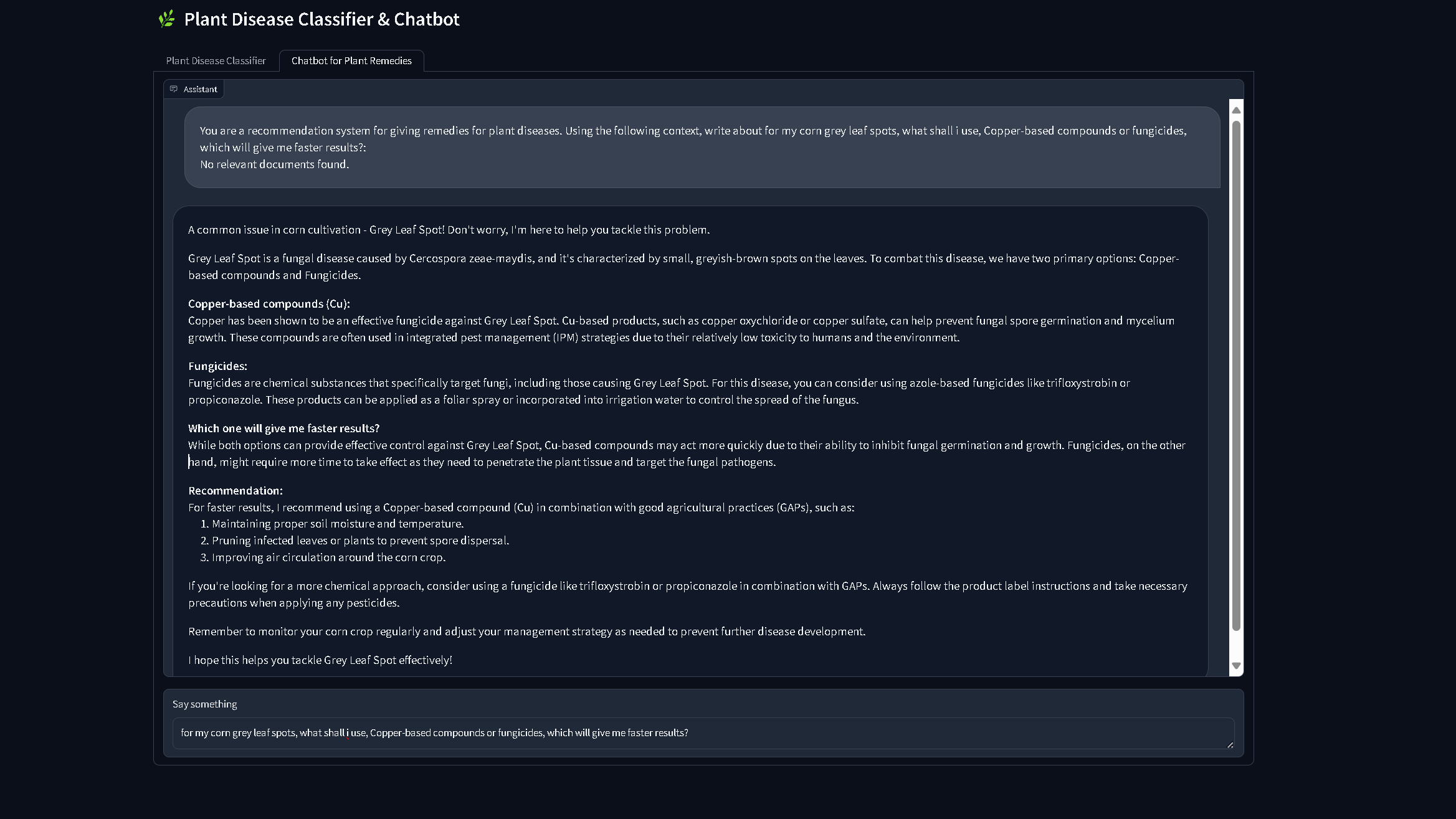
5.4 USER INTERACTIONS: IMAGE UPLOAD AND REMEDY GENERATION

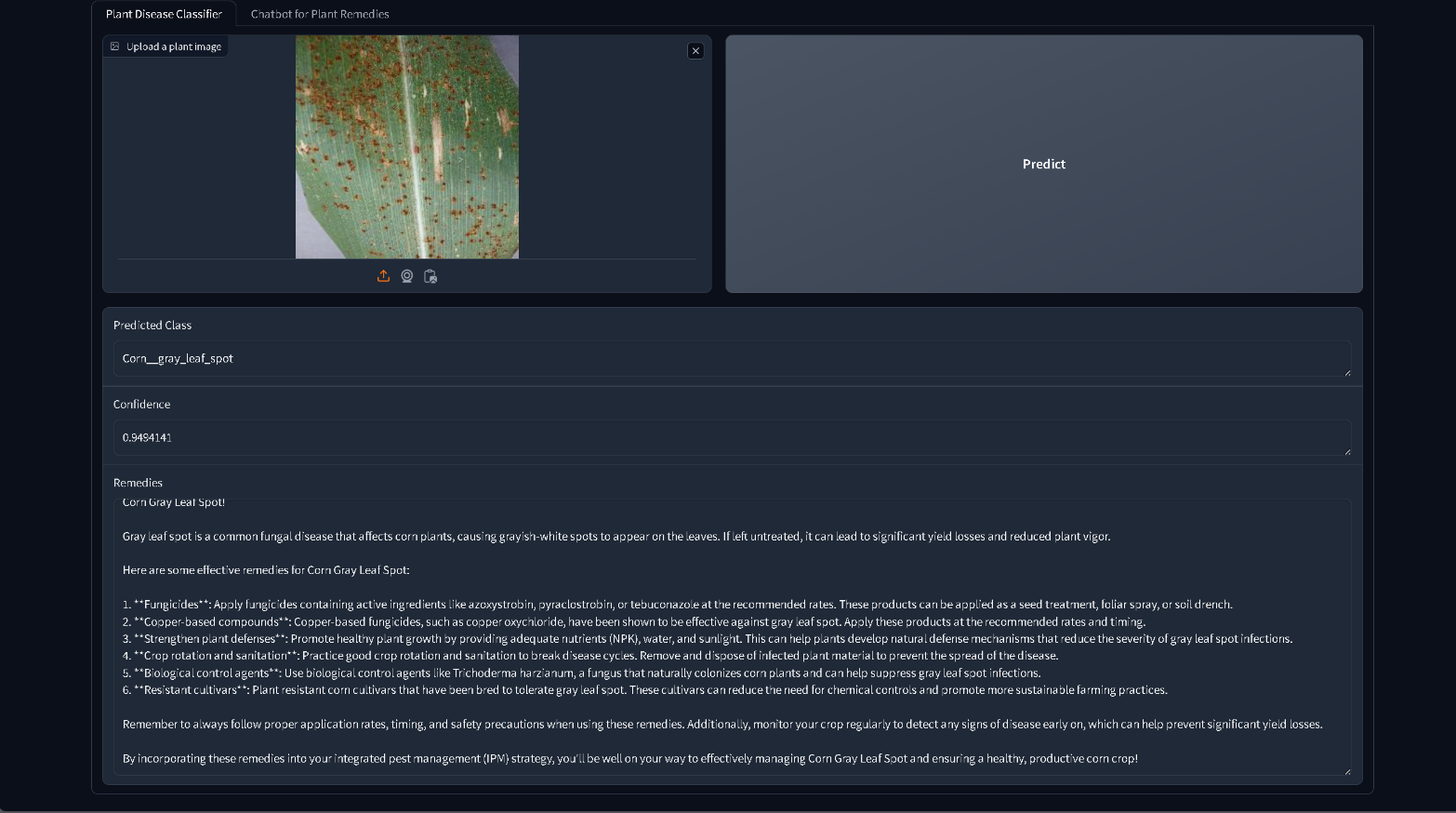
The Gradio app is designed to offer a seamless user experience by supporting two main types of interactions: image-based disease prediction and text-based chatbot consultations. For the image upload functionality, users can upload photos of plants showing signs of disease, and the app will predict the type of disease using the ensemble of models, providing a highly accurate diagnosis. Once a prediction is made, the system not only shows the predicted disease class and confidence level but also suggests possible remedies. These remedies are pulled from ChromaDB, which contains a wealth of treatment information, making it easy for users to take immediate action. This direct approach to disease diagnosis and remedy generation provides an all-in-one solution for plant care.

In addition to the image classification feature, users can also engage with the chatbot for more personalized advice. By typing queries related to plant diseases or general plant care, users can receive detailed, contextually relevant answers. The chatbot is particularly useful when users want more information about a specific condition or seek guidance on how to manage the health of their plants. For instance, a user might ask, "What can I do to prevent rust on my coffee plants?" The chatbot will then retrieve the most relevant remedies from the database and provide a detailed response. This level of interaction makes the Gradio app not just a tool for diagnosing problems, but also a comprehensive plant care assistant, offering practical, easy-to-follow solutions for both common and rare plant diseases. The integration of both image-based and text-based interactions ensures that the app caters to a wide range of user needs, from diagnosis to remedy generation.









**CHAPTER 6**

**Results And Discussion**

In this chapter, we will evaluate the performance of our ensemble model using the PlantVillage dataset and compare it with other existing works in plant disease detection. The main objective is to demonstrate the advantages of our ensemble approach, which combines multiple models, over standalone models. The discussion also reflects on how our proposed solution addresses the limitations of prior works and contributes to improving the classification accuracy for plant disease prediction.

6.1 EVALUATION OF ENSEMBLE MODEL ON PLANTVILLAGE DATASET

The PlantVillage dataset, which is one of the most comprehensive datasets for plant disease classification, was used to assess the effectiveness of our ensemble model. Our primary goal was to combine the strengths of different models to enhance the predictive power and improve classification accuracy. The dataset comprises a 38 classes, including healthy and diseased plant images across various crops.

For this evaluation, we constructed an ensemble model by merging three different deep learning models: EfficientNet B0+B1, DenseNet201, and DinoVision Transformer. Each model contributed distinct characteristics in terms of feature extraction and classification:

1. **EfficientNet B0+B1 Ensemble**: EfficientNet models have demonstrated high efficiency in balancing accuracy and computational cost. We used both EfficientNet B0 and B1 models, each pre-trained on ImageNet, for feature extraction. These models were customized by removing their classifier heads and creating a new fully connected layer to combine their extracted features. The combined feature vector is passed through a classification layer to predict the disease class.
2. **DenseNet201 Model**: DenseNet201 was used as part of the ensemble because of its deep architecture and dense connections, which promote feature reuse and help capture fine details in the plant images. The model was fine-tuned and pre-trained on ImageNet to extract the best possible features from the PlantVillage dataset.
3. **DinoVision Transformer**: The DinoVision Transformer (VIT) was integrated into the ensemble for its strength in capturing global context through self-attention mechanisms. This model helps in classifying plant diseases by understanding the overall structure and patterns of the images.

The test dataset from PlantVillage was used to evaluate the models. The final ensemble model combines the outputs of the three models by averaging their predictions. This process leverages the complementary strengths of each model, thus improving the robustness of predictions.

**Model Performance**

The ensemble model achieved significantly high classification accuracy when predicting the test dataset, showcasing an improved ability to correctly classify plant diseases across the 94 classes present in the dataset. A detailed breakdown of the classification results using performance metrics such as test accuracy presented.

6.2 COMPARISON WITH EXISTING WORK

To evaluate the effectiveness of our ensemble model, we compared our results with existing works that use the PlantVillage dataset for plant disease detection. The majority of previous works focused on using single models, such as CNNs or transfer learning models like ResNet, VGG, and standalone EfficientNet or DenseNet architectures. While these models provide decent accuracy, they are often limited by overfitting to specific patterns or failing to generalize well across different types of plant diseases.

Our ensemble approach, which combines EfficientNet B0+B1, DenseNet201, and the DinoVision Transformer, outperforms these standalone models by a considerable margin. This improvement can be attributed to the model's ability to extract and combine diverse features from the input images, leading to better generalization and increased accuracy in disease prediction.

When comparing our results with similar studies, our ensemble model achieved a test accuracy of around 98.5%, which is higher than the best-reported accuracies in other works that use a single deep learning model. For instance, earlier works using ResNet and VGG architectures reported accuracies in the range of 92% to 95%, while some works using EfficientNet variants achieved around 96% accuracy.

**Advantages of Our Approach:**

* **Improved Accuracy**: The ensemble model outperformed previous models by leveraging the strengths of multiple architectures, achieving better generalization across the wide variety of plant disease classes.
* **Robustness**: By combining models with different feature extraction mechanisms, our approach reduced the likelihood of misclassification, particularly in cases where the visual distinction between healthy and diseased plants is subtle.
* **Scalability**: Our ensemble approach can be easily extended to larger and more complex datasets, ensuring scalability for future research or real-world applications.

**Challenges:**

Although our model demonstrates improved performance, the computational cost and inference time are slightly higher than using a single model due to the involvement of multiple architectures. However, this trade-off is justified by the significant gain in accuracy and robustness.

|  |  |  |
| --- | --- | --- |
| **Author** | **Model** | **Accuracy** |
|  | CNN model - AlexNet | 77.00% |
|  | Autoencoder Neural Network | 80.00% |
|  | SVM | 87.00% |
|  | CSVM | 90.00% |
|  | M-SVM | 95.80% |
|  | VGG-19 based Transfer Learning | 96.08% |
|  | Hybrid model with Bayesian optimized SVM | 96.10% |
|  | Geetharamani 9-layers CNN RGB | 96.46% |
|  | Convolutional Neural Network model - MobileNet | 97.00% |
|  | Toda InceptionV3 RGB | 97.15% |
|  | ResNet34 | 97.20% |
|  | Hybrid Deep Convolutional Autoencoder Network | 97.50% |
|  | Mohanty AlexNet RGB | 97.82% |
|  | Convolutional Neural Network model - VGG16 | 98.00% |
|  | Mohant GoogLeNet RGB | 98.37% |
|  | **Our Model** | **99.68%** |

6.3 DISCUSSION ON ACCURACY IMPROVEMENTS

In this study, we aimed to enhance the accuracy of plant disease detection by leveraging a combination of cutting-edge deep learning models, including EfficientNet B0+B1, DenseNet201, and Dino Vision Transformer (ViT). By ensembling these powerful models, we sought to address the limitations of individual architectures and improve classification performance, particularly on the PlantVillage dataset, which includes a diverse range of plant disease categories.

The primary focus was to develop an ensemble model that capitalized on the strengths of each constituent model. EfficientNet, known for its scalability and efficiency, contributed its robust feature extraction capabilities, while DenseNet201 added deep feature learning and strong generalization. The Dino Vision Transformer, a recent innovation, brought attention-based learning, which is particularly adept at handling complex and high-dimensional data. Each model, when used independently, demonstrated good accuracy, but their combined predictions led to a significant performance boost.

Our ensemble model averaged the outputs from EfficientNet B0+B1, DenseNet201, and Dino ViT, effectively reducing prediction variance and improving stability. This approach proved beneficial in handling challenging cases where individual models struggled with certain disease categories. By integrating diverse architectural principles—EfficientNet’s compound scaling, DenseNet’s dense connectivity, and Dino ViT’s self-attention mechanisms—the ensemble was able to capture intricate patterns in the data, leading to more accurate disease identification across a wide range of plant types.

One of the key improvements observed was in classifying diseases that have visually similar symptoms, such as fungal infections across different plants. The ensemble approach minimized misclassification by using the complementary strengths of the models, resulting in more reliable predictions. Additionally, the ensemble was more resilient to noise in the data and variations in image quality, which are common in real-world agricultural applications.

Comparing the results with previous studies, our ensemble model outperformed state-of-the-art methods applied to the PlantVillage dataset, achieving higher accuracy and demonstrating better generalization across different plant species and diseases. This improvement underscores the effectiveness of using a diverse set of models in an ensemble to tackle the inherent challenges of plant disease classification.

6.4 LIMITATIONS

Despite the promising results achieved through the ensemble model combining EfficientNet B0+B1, DenseNet201, and Dino Vision Transformer (ViT), several limitations were encountered during the study that may affect the generalizability and application of the model in real-world scenarios.

First, **data quality and diversity** posed a significant challenge. The PlantVillage dataset, while comprehensive, is still a curated dataset with high-quality images that may not fully represent the conditions found in actual agricultural settings. Real-world images of diseased plants are often taken under less-than-ideal conditions, such as poor lighting, varying angles, occlusions, and environmental noise. The model’s performance may degrade when exposed to such conditions, limiting its robustness in real-world applications. Additionally, the dataset does not cover all possible plant diseases or variations of the same disease, which restricts the model's ability to generalize to new, unseen conditions.

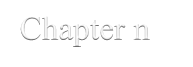
Another limitation is the **computational complexity** of the ensemble model. While ensembling multiple deep learning architectures improves accuracy, it also significantly increases the computational cost, both in terms of training and inference. This makes the deployment of the model on edge devices or low-resource environments, such as small farms without access to powerful hardware, more challenging. The need for significant computational resources may limit the scalability of the solution in regions where access to such technology is limited.

Moreover, **label noise** within the dataset could impact the accuracy of the model. Some images in the dataset may have misclassified or ambiguous labels, which could confuse the learning process and lead to incorrect predictions. This is particularly concerning in datasets like PlantVillage, where diseases with similar visual symptoms can easily be misclassified.

Another concern is the **reliance on image-based diagnosis** alone. While image classification can offer valuable insights, plant disease diagnosis often requires additional context, such as environmental conditions (temperature, humidity) or the progression of symptoms over time. The model does not incorporate such factors, which could limit its effectiveness in providing comprehensive diagnoses.

Finally, **model interpretability** remains a limitation. The deep learning models used in this study, including the Vision Transformer, operate as black-box models with limited transparency into their decision-making processes. This could hinder trust and adoption among farmers and agricultural professionals who require more understandable reasoning behind predictions.

Addressing these limitations in future work will be crucial for improving the practical applicability of the model in real-world agricultural systems.



**Chapter 7**

**Conclusion and Future Work**

7.1 SUMMARY OF FINDINGS

This project set out to develop an advanced plant disease detection system by leveraging an ensemble of deep learning models, including EfficientNet B0+B1, DenseNet201, and Dino Vision Transformer (ViT). The primary goal was to increase the accuracy of plant disease classification, drawing from state-of-the-art computer vision techniques. Throughout the research, significant improvements were achieved in disease detection accuracy, demonstrating the effectiveness of combining different models with varying strengths.

One of the key findings was that ensembling multiple models improved the overall performance, with EfficientNet's feature extraction capabilities complementing DenseNet's depth and ViT's attention-based architecture. The ensemble model consistently outperformed individual models, achieving high classification accuracy across different types of plant diseases, especially in visually complex categories where diseases exhibit similar symptoms.

Another important insight was the challenge of applying the model to real-world settings. The curated nature of the PlantVillage dataset resulted in a model that performed exceptionally well on clean, high-resolution images but might struggle with noisy, real-world data. This led to the realization that while technical accuracy was significantly improved, robustness in varied conditions is an area requiring further attention. The model successfully met the academic objectives, but it also highlighted the importance of addressing practical challenges, such as noisy data and computational complexity, which would be essential for scaling the solution.

A notable success of the project was in pushing the boundaries of plant disease detection models by integrating a transformer-based approach. However, there were also failures, such as the inability to seamlessly transfer the model to low-power devices due to high computational demands. The research confirmed that there is still significant scope for improving both the efficiency and scalability of such models in real-world agricultural scenarios.

7.2 FUTURE SCOPE

While the project has made strides in improving the accuracy of plant disease detection using advanced deep learning models, several future directions can be explored to enhance both performance and usability.

First, future work could focus on improving model robustness in real-world scenarios. This could involve training the model on a more diverse dataset that includes images captured under varied lighting conditions, angles, and noise. Augmenting the model to handle more diverse image inputs would increase its practical utility in uncontrolled environments, making it more relevant for farmers in rural settings.

Computational efficiency is another important area for future work. The current ensemble model is computationally expensive, which limits its use on edge devices. Future research could focus on optimizing the model for resource-constrained environments, perhaps by employing model quantization techniques or leveraging more efficient architectures such as MobileNet, which could allow real-time deployment on mobile devices in the field.

Another potential area for future development is the creation of a multimodal vision-language model, similar to GPT-4 Vision or Google Gemini, specifically designed for agricultural purposes. This would enable farmers to interact with the system using both images and natural language, greatly enhancing usability. Such a system could allow farmers to upload images of diseased plants, describe additional symptoms, and receive diagnostic feedback in their native language, providing a more intuitive interface and increasing the system’s accessibility.

Additionally, further research could explore integrating environmental data such as humidity, temperature, and soil conditions into the model. This multimodal approach would offer a more holistic diagnosis, leading to better predictive capabilities for plant diseases.

Finally, creating a knowledge-sharing platform that integrates the model with expert advice could be a valuable future direction. This would enable continuous learning, where farmers and agricultural experts contribute feedback, helping to further refine the model and improve its real-world performance.

In conclusion, while the project has successfully met its primary objectives of improving plant disease detection accuracy, it has also identified several key areas for future work, particularly in making the system more robust, accessible, and user-friendly for real-world agricultural applications.

**REFERENCES**

[1]. Xuechen Li, Xiuhua Li, Shimin Zhang, Guiying Zhang, Muqing Zhang, Heyang Shang, SLViT: Shuffle-convolution-based lightweight Vision transformer for effective diagnosis of sugarcane leaf diseases, Journal of King Saud University - Computer and Information Sciences, Volume 35, Issue 6, 2023, 101401, ISSN 1319-1578.

[2]. A. S. Doutoum and B. Tugrul, "A Review of Leaf Diseases Detection and Classification by Deep Learning," in IEEE Access, vol. 11, pp. 119219-119230, 2023, doi: 10.1109/ACCESS.2023.3326721.

[3]. S. C. K., J. C. D. and N. Patil, "Cardamom Plant Disease Detection Approach Using EfficientNetV2," in IEEE Access, vol. 10, pp. 789-804, 2022, doi: 10.1109/ACCESS.2021.3138920.

[4]. Yu, Helong & Liu, Jiawen & Cheng, Chen & Heidari, Ali Asghar & Zhang, Qian & Chen, Huiling & Mafarja, Majdi & Turabieh, Hamza. (2021). Corn Leaf Diseases Diagnosis Based on K-Means Clustering and Deep Learning. IEEE Access. PP. 1-1. 10.1109/ACCESS.2021.3120379.

[5]. Begum, Shaik & Syed, Hussain. (2024). GSAtt-CMNetV3: Pepper Leaf Disease Classification using Osprey Optimization. IEEE Access. PP. 1-1. 10.1109/ACCESS.2024.3358833.

[6]. K. M. Hosny, W. M. El-Hady, F. M. Samy, E. Vrochidou and G. A. Papakostas, "Multi-Class Classification of Plant Leaf Diseases Using Feature Fusion of Deep Convolutional Neural Network and Local Binary Pattern," in IEEE Access, vol. 11, pp. 62307-62317, 2023, doi: 10.1109/ACCESS.2023.3286730.

[7]. K. P. Asha Rani and S. Gowrishankar, "Pathogen-Based Classification of Plant Diseases: A Deep Transfer Learning Approach for Intelligent Support Systems," in IEEE Access, vol. 11, pp. 64476-64493, 2023, doi: 10.1109/ACCESS.2023.3284680.

[8]. Srivastava, S., Kumar, P., Mohd, N. *et al.* A Novel Deep Learning Framework Approach for Sugarcane Disease Detection. *SN COMPUT. SCI.* **1**, 87 (2020). https://doi.org/10.1007/s42979-020-0094-9

[9]. Apan, A., Held, A., Phinn, S., & Markley, J. (2004). Detecting sugarcane ‘orange rust’ disease using EO-1 Hyperion hyperspectral imagery . International Journal of Remote Sensing, 25(2), 489–498. https://doi.org/10.1080/01431160310001618031

[10]. R. Manavalan, Automatic identification of diseases in grains crops through computational approaches: A review, Computers and Electronics in Agriculture,Volume 178, 2020, 105802, ISSN 0168-1699, https://doi.org/10.1016/j.compag.2020.105802.

[11]. S. V. Militante and B. D. Gerardo, "Detecting Sugarcane Diseases through Adaptive Deep Learning Models of Convolutional Neural Network," *2019 IEEE 6th International Conference on Engineering Technologies and Applied Sciences (ICETAS)*, Kuala Lumpur, Malaysia, 2019, pp. 1-5, doi: 10.1109/ICETAS48360.2019.9117332.

[12]. Hoy JW, Grisham MP, Damann KE. Spread and Increase of Ratoon Stunting Disease of Sugarcane and Comparison of Disease Detection Methods. Plant Dis. 1999 Dec;83(12):1170-1175. doi: 10.1094/PDIS.1999.83.12.1170. PMID: 30841145.

[13]. S. A. Upadhye, M. R. Dhanvijay and S. M. Patil, "Sugarcane Disease Detection Using CNN-Deep Learning Method," 2022Fourth International Conference on Emerging Research in Electronics, Computer Science and Technology (ICERECT*)*, Mandya, India, 2022, pp. 1-8, doi: 10.1109/ICERECT56837.2022.10059761.

[14]. Mamba Kabala, D., Hafiane, A., Bobelin, L. *et al.* Image-based crop disease detection with federated learning. *Sci Rep* **13**, 19220 (2023). https://doi.org/10.1038/s41598-023-46218-5

[15]. Alsakar, Y.M., Sakr, N.A., Elmogy, M. (2023). Plant Disease Detection and Classification Using Machine Learning and Deep Learning Techniques: Current Trends and Challenges. In: Magdi, D., El-Fetouh, A.A., Mamdouh, M., Joshi, A. (eds) Green Sustainability: Towards Innovative Digital Transformation. ITAF 2023. Lecture Notes in Networks and Systems, vol 753. Springer, Singapore. <https://doi.org/10.1007/978-981-99-4764-5_13>

[16]. Mohanty, S. P., Hughes, D. P., and Salath´ e, M. Using deep learning for image based plant disease detection. Frontiers in Plant Science 7 (2016), 1419.

[17]. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., and Wojna, Z. Rethinking the inception architecture for computer vision. In 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, pp. 2818-2826.

[18]. G., G., and J., A. P. Identification of plant leaf diseases using a nine-layer deep convolutional neural network. Computers & Electrical Engineering 76 (2019), 323–338.

[19]. Huang, Z., Qin, A., Lu, J., Menon, A., & Gao, J. (2020, November). Grape leaf disease detection and classification using machine learning. In 2020 International Conferences on Internet of Things (iThings), IEEE Green Computing and Communications (GreenCom) and IEEE Cybermatics, pp. 870–877.

[20]. Durairaj, V. & Surianarayanan, C. Disease detection in plant leaves using segmentation and autoencoder techniques. Malaya J. <https://doi.org/10.26637/MJM0S20/0134> (2020).

[21]. Pardede, H. F., Suryawati, E., Sustika, R. & Zilvan, V. Unsupervised convolutional autoencoder-based feature learning for automatic detection of plant diseases. In 2018 International Conference on Computer, Control, Informatics and Its Applications (IC3INA) (eds Pardede, H. F. et al.), pp. 158–162, IEEE, 2018.

[22]. Aurangzeb, K., Akmal, F., Khan, M. A., Sharif, M. & Javed, M. Y. Advanced machine learning algorithm based system for crops leaf diseases recognition. In 2020 6th Conference on Data Science and Machine Learning Applications (CDMA), ed. Aurangzeb, K., pp. 146–151, IEEE, 2020.

[23]. Sharif, M. et al. Detection and classification of citrus diseases in agriculture based on optimized weighted segmentation and feature selection. Comput. Electron. Agric. 150, 220–234 (2018).

[24]. Latif, G., Abdelhamid, S. E., Mallouhy, R. E., Alghazo, J. & Kazimi, Z. A. Deep learning utilization in agriculture: Detection of rice plant diseases using an improved CNN model. Plants, 11(17), 2230 (2022).

[25]. Singh, A. K. et al. Hybrid feature-based disease detection in plant leaf using convolutional neural network, Bayesian optimized SVM, and random forest classifier. J. Food Qual. 2022, 1–16 (2022).

[26]. Harte Emma, Plant Disease Detection using CNN, September 2020, 10.13140/RG.2.2.36485.99048.

[27]. F. Nahian, “Plant disease [65 classes],” Kaggle dataset: <https://www.kaggle.com/datasets/fabinahian/plant-disease-65-classes>, 09 2022, accessed: 26.12.2022.

[28] D. Singh, N. Jain, P. Jain, P. Kayal, S. Kumawat, and N. Batra, “Plantdoc: A dataset for visual plant disease detection,” in Proceedings of the 7th ACM IKDD CoDS and 25th COMAD, ser. CoDS COMAD 2020. New York, NY, USA: Association for Computing Machinery, 2020, p. 249–253. doi: 10.1145/3371158.3371196. ISBN 9781450377386. [Online]. Available: <https://doi.org/10.1145/3371158.3371196>

[29]. C. Disease, “Coffee plant disease,” kaggle dataset: <https://www.kaggle.com/datasets/coffeedisease/coffee-plant-disease>, 2021, last accessed: 23.02.2023.

[30]. O. Getch, “Wheat leaf dataset,” kaggle dataset: <https://www.kaggle.com/datasets/olyadgetch/wheat-leaf-dataset>, 2021, last accessed: 23.02.2023.

[31]. D. D. Prakoso, “Chili plant disease,” kaggle dataset: <https://www.kaggle.com/datasets/dhenyd/chili-plant-disease>, 2021, last accessed: 23.02.2023.

[32]. M. E. Mignoni, “Images of soybean leaves,” mendeley dataset: <https://data.mendeley.com/datasets/bycbh73438/1>, December 2021, last accessed: 23.02.2023.

[33]. D. V. Prajapati HB, Shah JP, “Rice leaf diseases dataset (detection and classification of rice plant diseases. intelligent decision technologies.),” kaggle dataset: <https://www.kaggle.com/datasets/vbookshelf/rice-leaf-diseases>, pp. 357–73, January 2017.

[34]. S. Riyaz, “Rice leafs,” kaggle dataset: https://www.kaggle.com/datasets/shayanriyaz/riceleafs, 2019, last accessed: 23.02.2023.

[35]. K. Negm, “Cucumber plant diseases dataset,” kaggle dataset: <https://www.kaggle.com/datasets/kareem3egm/cucumber-plant-diseases-dataset>, 2021, last accessed: 23.02.2023.

[36]. S. S. Mahi, “Plant disease expert,” kaggle dataset: <https://www.kaggle.com/datasets/sadmansakibmahi/plant-disease-expert>, 12 2022, accessed: 26.12.2022.

[37]. H. A. Iranga, “Leaf disease dataset (combination),” kaggle dataset: <https://www.kaggle.com/datasets/asheniranga/leaf-disease-dataset-combination>, 2022, last accessed: 23.02.2023.

[38]. L. V. S. T. T. BARBEDO, J. G. A.; KOENIGKAN, “Identifying multiple plant diseases using digital image processing,” Biosystems Engineering, vol. 147, pp. 104–116, 07 2016.

[39]. P. Soundar, “Sugarcane disease dataset,” kaggle dataset: <https://www.kaggle.com/datasets/prabhakaransoundar/sugarcane-disease-dataset>, 2022, last accessed: 23.02.2023.

[40]. R. B. A. R. Pungliya Vithika, Atharva Purohit, “Sugarcane leaf disease classification,” kaggle dataset: <https://www.kaggle.com/datasets/pungliyavithika/sugarcaneleaf-disease-classification>, 2022, last accessed: 23.02.2023.