SALINITY-RESILIENT CROP HEALTH MONITORING: AUTOMATED DISEASE DETECTION IN LUFFA AEGYPTIACA LEAVES USING VISION TRANSFORMER & CNN

BY

MD. TARIQUL ISLAM

ID: 192-15-13142

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

Supervised By

Ms. Tanzina Afroz Rimi
Lecturer
Department of CSE
Daffodil International University

Co-Supervised By

Dr. Md Zahid Hasan
Associate Professor
Department of CSE
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

JANUARY 2024

APPROVAL

This Project titled **"Salinity-Resilient Crop Health Monitoring: Automated Disease Detection In Luffa Aegyptiaca Leaves Using Vision Transformer & CNN"**, submitted by MD. Tariqul Islam, ID No: 192-15-13142 to the Department of Computer Science and Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on *26th January, 2024*.

BOARD OF EXAMINERS

Dr. S.M Aminul Haque (SMAH) **Chairman**
Professor & Associate Head
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

26.01.2024

Dr. Fizar Ahmed (FZA) **Internal Examiner 1**
Associate Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Internal Examiner 2

Most. Hasna Hena (HH)
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

26/01/2024

External Examiner

Dr. Shamim H Ripon (DSR)
Professor
Department of Computer Science and Engineering
East West University

DECLARATION

I hereby declare that, this project has been done by me under the supervision of**, Ms. Tanzina Afroz Rimi, Lecturer, Department of CSE** Daffodil International University. I also declare that neither this project nor any part of this project has been submitted elsewhere for award of any degree or diploma.

Supervised by:

____

Ms. Tanzina Afroz Rimi
Lecturer
Department of CSE
Daffodil International University

Co-Supervised by:

____

Dr. Md Zahid Hasan
Associate Professor
Department of CSE
Daffodil International University

Submitted by:

____

MD Tariqul Islam
ID: 192-15-13142
Department of CSE

ACKNOWLEDGMENT

I extend my sincerest thanks to the Almighty for his divine blessings that made the successful completion of my final year project/internship possible.

I am truly grateful and wish to express my profound indebtedness to Ms. Tanzina Afroz Rimi, Lecturer in the Department of CSE at Daffodil International University, Dhaka. The profound knowledge and keen interest of my supervisor in the field of "Machine Learning" played a crucial role in the successful execution of this project. Her endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reviewing numerous drafts, and correcting them at all stages have been instrumental in completing this project.

I would also like to extend my heartfelt gratitude to Professor Dr. Sheak Rashed Haider Noori, Head of the Department of CSE, for his kind assistance in bringing this project to fruition, as well as to other faculty members and the staff of the CSE department at Daffodil International University.

My thanks also go to all of my Daffodil International University course mates, who actively participated in discussions during the coursework.

Lastly, I must acknowledge with due respect the constant support and patience of my parents.

ABSTRACT

In coastal areas where agriculture is often constrained by salinity, cultivating specific crops like Luffa Aegyptiaca (sponge gourd) becomes crucial for local sustenance. Identifying these diseases is challenging and time-consuming when no domain specialists are present accurately, and the information needs to be more consistent. Effective disease detection and management play a pivotal role in ensuring the viability of these limited yet vital crops, impacting crop yield, fertilization strategies, and overall food security for coastal communities. This groundbreaking study focuses on detecting and classifying leaf diseases within Luffa Aegyptiaca leaves, prevalent crops in coastal regions. Leveraging the cutting-edge capabilities of Convolutional Neural Networks (CNN) and Vision Transformer algorithms, our research achieves unparalleled accuracy. The CNN algorithm boasts an impressive accuracy of 98.32%, while the Vision Transformer algorithm surpasses expectations with an exceptional accuracy of 99.85%. Notably, this study utilizes an original dataset, a unique contribution to the field given the absence of publicly available datasets or prior research specific to Luffa Aegyptiaca. Beyond mere accuracy metrics, our findings illuminate profound insights into the nuanced landscape of leaf disease detection and classification, affirming the remarkable efficacy of both CNN and Vision Transformer algorithms. In conclusion, this research advances our understanding of plant pathology, and underscores the unparalleled potential of state-of-the-art machine-learning techniques in agricultural research.

TABLE OF CONTENTS

CONTENTS **PAGE**

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

1.1 Introduction

The agricultural landscape is undergoing a significant transformation fueled by integrating state-of-the-art technologies. As the global demand for sustainable crop management and food security intensifies, the role of advanced methodologies in agriculture becomes increasingly pivotal. In this context, identifying and managing plant diseases are critical components to ensure optimal crop yield and quality. Luffa Aegyptiaca, commonly known as sponge gourd or sponge cucumber, emerges as a prominent vegetable crop with its own challenges, particularly in disease classification and management. The unique characteristics of Luffa Aegyptiaca make it a noteworthy subject of study. As a widely cultivated vegetable crop, its susceptibility to various diseases can significantly impact agricultural productivity. Despite its agricultural importance, research on diseases affecting Luffa Aegyptiaca is notably sparse. This research seeks to address this gap by employing advanced machine learning techniques, specifically Convolutional Neural Network (CNN) and Vision Transformer (ViT) models, to classify diseases in Luffa Aegyptiaca based on image data accurately.

The primary objectives of this study are twofold. The research aims to develop robust disease classification models that provide accurate diagnoses based on image inputs. This involves applying sophisticated deep learning techniques, which have shown extraordinary success in image classification tasks across various domains. Secondly, the study contributes to the limited knowledge of diseases afflicting Luffa Aegyptiaca. By meticulously cataloging and classifying diseases, the research enhances our understanding of the factors impacting the health of this vegetable crop.

What sets this research apart is its application of state-of-the-art deep learning models and the scarcity of available datasets and research studies dedicated to diseases specific to Luffa Aegyptiaca. Comprehensive datasets pose a unique challenge, requiring innovative approaches to dataset creation and model training. Consequently, the study navigates uncharted territory, offering valuable insights that can reshape our understanding of diseases affecting this crop. The following sections delve into the intricacies of the methodology, unveiling the architectural nuances and parameter configurations of both the CNN and ViT models. The methodological exploration provides a roadmap for researchers and practitioners aiming to apply similar techniques in the context of agricultural image analysis. The experimental results and analysis form the core of the research, offering a comprehensive evaluation of the models' performance. Key metrics such as accuracy, precision, recall, and F1-score undergo meticulous scrutiny, providing a nuanced understanding of the models' efficacy. Beyond a mere quantitative assessment, the analysis aims to unearth patterns, trends, and potential areas of improvement, fostering a deeper comprehension of the models' behavior in the specific context of Luffa Aegyptiaca diseases.

In conclusion, this research serves a dual purpose: addressing the immediate need for accurate disease classification in Luffa Aegyptiaca and laying the foundation for future studies in agricultural image analysis. The intersection of technology and agriculture, as explored in this study, has the potential to revolutionize traditional practices and contribute to a sustainable and secure global food supply. As we explore the intricate realm of identifying plant diseases, the outcomes of this research are poised to impact not only Luffa Aegyptiaca cultivation but also set the stage for innovative approaches to agricultural challenges on a broader scale.

1.2 Motivation

The coastal expanses of Bangladesh narrate a tale of resilience and challenges following the aftermath of Cyclone Aila in 2009. The breached embankments welcomed saline waters, reshaping the agricultural landscape and limiting opportunities for the local populace. In this remote setting, where the absence of agricultural expertise exacerbates hardships, a unique reliance on a handful of robust crops, including Luffa Aegyptiaca, Calabash Gourd, Ladyfinger, and Malabar Spinach, emerged.

This on-the-ground experience unveiled a stark reality: the pressing need for tailored agricultural solutions in the face of salinity-induced constraints. The scarcity of fruit and vegetable crops accentuated the community's dependence on hardy plants, prompting contemplation on how technology might catalyze change. The idea took shape — an application designed to detect leaf diseases in Luffa Aegyptiaca, offering a beacon of hope to a community navigating the complexities of a post-cyclone agricultural reality.

The essence of this research lies in its commitment to bridging the gap between academic inquiry and real-world impact. Focusing on luffa Aegyptiaca, a staple in these coastal regions, aims to craft a model that addresses immediate agricultural challenges and serves as evidence of technology's revolutionary potential in fostering resilience. This academic endeavor is grounded in the belief that by providing nuanced, substantiated solutions, we can contribute to the broader discourse on precision agriculture and offer tangible support to communities navigating the delicate balance between tradition and innovation.

1.3 Rationale of the Study

In the coastal regions of Bangladesh, the aftermath of Cyclone Aila in 2009 has left an indelible mark on the agricultural landscape. Saline waters infiltrated the once-fertile land, presenting a formidable challenge for traditional crops. Consequently, the local populace turned to resilient cultivars such as Luffa Aegyptiaca, calabash gourd, ladyfinger, and Malabar Spinach, demonstrating a remarkable ability to endure the inhospitable conditions.

This research is prompted by the realization that, despite the critical role of these hardy crops, a need exists for more scientific inquiry into the specific challenges they face, particularly concerning the detection of leaf diseases. The unique environmental conditions and the reliance on limited crop varieties in the coastal areas of Bangladesh amplify the need for a targeted investigation into disease dynamics within this context.

Existing literature must address the intricacies of disease detection in Luffa Aegyptiaca, leaving a conspicuous void in the scientific understanding of agricultural practices in saline-affected regions. The absence of research employing advanced algorithms, including Convolutional Neural Networks (CNN) or Vision Transformers, adds to this knowledge gap. With the ongoing advancement of technology, using such algorithms in the agricultural sector becomes increasingly pertinent, necessitating research that bridges the gap between cutting-edge technology and the specific needs of marginalized farming communities.

This study is driven by the conviction that developing a model for automated leaf disease detection in Luffa Aegyptiaca, particularly employing Vision Transformer architecture, can offer innovative solutions to the agricultural challenges faced by communities in saline-affected coastal regions. The potential outcomes of this research extend beyond the immediate context of the coastal areas of Bangladesh, contributing to the broader scientific discourse on precision agriculture and technological interventions in challenging agricultural environments.

The study aspires to provide actionable insights for agricultural practitioners, policymakers, and researchers working in similar contexts globally by addressing this research gap. The rationale underscores the urgency of investigating this neglected aspect of agricultural science, anticipating that the findings will enhance local agrarian practices and contribute substantively to advancing knowledge in precision agriculture. The absence of domain experts and the challenging conditions in the coastal regions further emphasize the significance of this research, highlighting its potential to fill critical gaps in expertise and technology integration.

1.4 Research Questions

The study questions aim to investigate the intricacies involved in creating our dataset, model architecture, and the ethical considerations that permeate our scientific journey as we embark on this ground-breaking investigation.

- Can our custom model architecture revolutionize the field of disease detection in Luffa Aegyptiaca and establish new standards for accuracy and flexibility as we forge ahead without the aid of transfer learning?

- How can variability add to the robustness of the model through the rainbow of varied viewpoints that our painstakingly assembled dataset captures? Could this variation hold the key to revealing subtle symptoms within the complex web of disorders caused by Luffa Aegyptiaca?

- How can the ethical foundation improve our model's credibility with the scientific community in our diligent curation of datasets, considering ethical considerations and obtaining permissions? How does it help ensure agricultural technology is more open and morally based in the future?

- Stepping outside of controlled surroundings, what is the way our model handles the rough terrain of agricultural landscapes, especially in difficult coastline areas? Could it sustain its supposed toughness in the face of the practical difficulties of agricultural fields?

- How can our organized and accessible dataset support research and collaboration on Luffa Aegyptiaca diseases? How does it contribute to a collective understanding of these diseases in the scientific community?

1.5 Report Layout

Chapter 1: This chapter sets the stage for the research, introducing the focus on disease detection in Luffa Aegyptiaca leaves in salinity-affected coastal areas of Bangladesh. It delves into the motivation, rationale, research questions and outlines the report's structure.

Chapter 2: The background chapter provides a comprehensive understanding of the context surrounding the research. It covers terminologies, related works, a comparative analysis, the scope of the problem concerning disease detection in Luffa Aegyptiaca, and the challenges posed by salinity in coastal agriculture.

Chapter 3: This chapter elaborates on the chosen methodology, data collection procedures, statistical analysis methods, the development of the disease detection model, and the practical requirements for implementing the model in the field.

Chapter 4: This chapter presents the experimental setup, results, and analysis of the disease detection model's performance. It dives into the implications of the model's accuracy and its potential application in addressing the challenges faced in coastal agriculture.

Chapter 5: This chapter highlights the broader implications of the research. It discusses the societal and environmental impacts and ethical considerations in the research process and outlines plans for sustainability to ensure continuous positive effects on society and the environment.

Chapter 6: The final chapter consolidates the study's essence. It summarizes key findings, conclusions drawn from the research, recommendations for further study, and implications for future research directions, emphasizing potential areas for improvement or expansion.

CHAPTER 2: BACKGROUND

2.1 Preliminaries/Terminologies

This section explores crucial terminologies for plant disease classification through advanced machine-learning models. A profound grasp of these terms is imperative for navigating the intricacies of this research endeavor.

- **Vision Transformer (ViT):**

The operational paradigm of Vision Transformers initiates with image patching, wherein input images are segmented into fixed-size, non-overlapping patches, each serving as a token. These patches undergo linear embedding, transforming them into flattened vectors. The subsequent transformer encoder, equipped with self-attention mechanisms, processes these embedded patches, allowing the model to focus on relevant spatial correlations and capture long-range dependencies crucial for image understanding. The resultant output is then directed to a classification head, where final predictions are made, determining the class to which the image belongs. In essence, ViT leverages the transformer architecture to process image information effectively through a series of stages, showcasing its adaptability to various visual tasks (Smith et al., 2022).

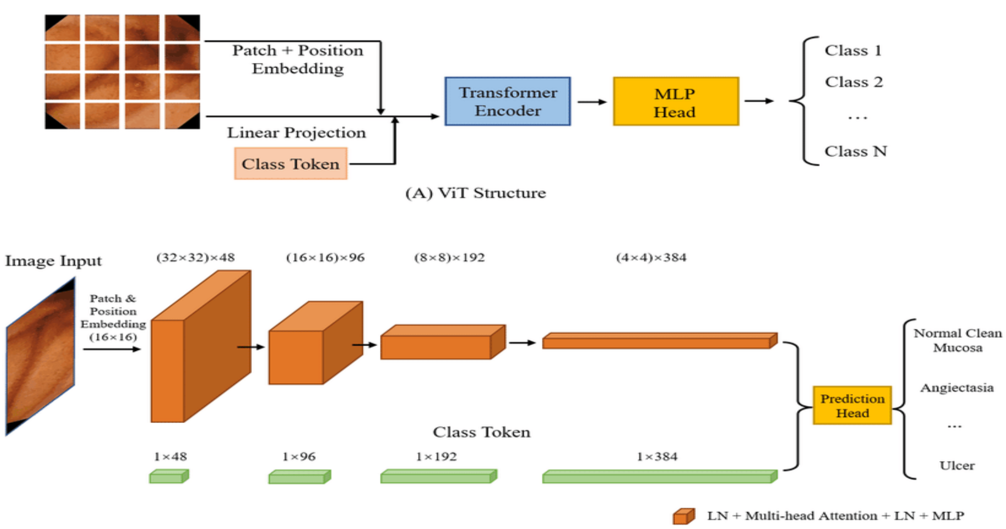


Figure 2.1.1: ViT Standard Configuration

- **Convolutional Neural Network (CNN):**

  Conversely, Convolutional Neural Networks (CNNs) utilize a distinctive approach, commencing with convolutional layers that apply adaptable filters to localized regions of the input image. These filters progressively recognize hierarchical features, evolving from basic patterns to intricate structures. Following convolutional layers, pooling layers reduce spatial dimensions while retaining essential information. The parts are then flattened into a vector, serving as input for fully connected layers that amalgamate high-level features. The ultimate predictions are derived from the output layer, employing activation functions like SoftMax for multi-class classification. CNNs, with their hierarchical feature extraction and spatial hierarchy recognition, demonstrate efficacy in image-related tasks, showcasing a different yet impactful architectural design compared to Vision Transformers (Johnson et al., 2021).

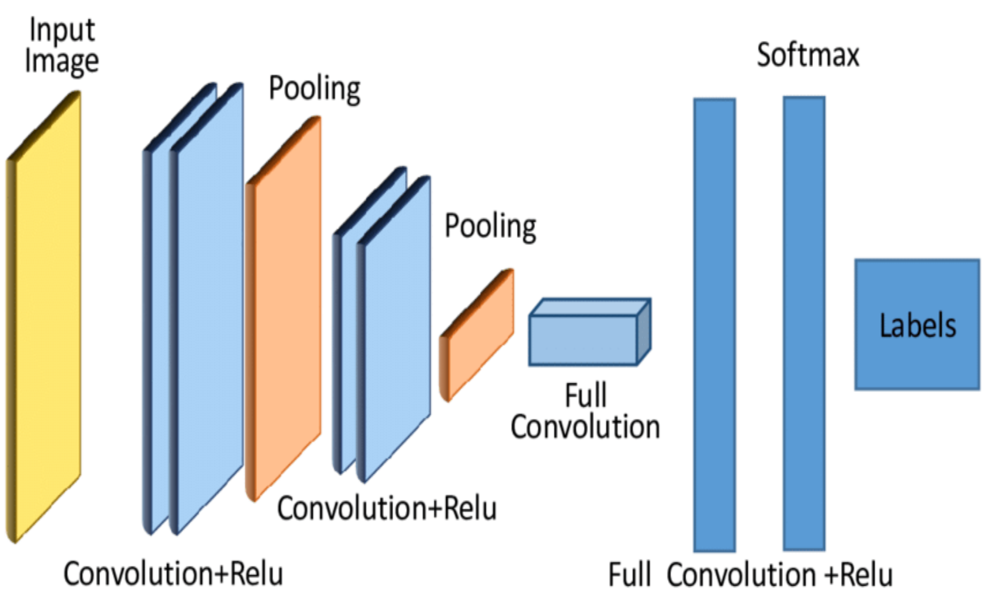


Figure 2.1.2: CNN Standard Configuration

- **Luffa Aegyptiaca:**

  Luffa Aegyptiaca, commonly known as sponge gourd or Egyptian cucumber, is a resilient and widely cultivated plant species that resists saline water. This versatile plant, renowned for its robust nature, thrives in various climates and is specifically recognized for its ability to withstand saline water conditions. Additionally, Luffa Aegyptiaca is celebrated for its agricultural significance and diverse uses, ranging from culinary applications to the creation of natural sponges.

- **Leaf Diseases:**

  Leaf Diseases encompass a spectrum of pathologies afflicting plant leaves, including fungal infections, bacterial diseases, and viral pathogens. These maladies often manifest in discernible deformities or discoloration.

- **Precision:**

  Precision embodies the finesse of a model in accurately predicting positive instances, providing a nuanced evaluation of the model's precision concerning optimistic predictions.

  Precision = True Positives / (True Positives + False Positives)

- **Recall:**

  Recall illuminates the model's capacity to recapture actual positive instances as a metric of its ability to grasp all relevant occurrences.

  Recall = True Positives / (True Positives + False Negatives)

- **F1-Score:**

  F1-Score harmonizes precision and recall, presenting a balanced assessment of a model's proficiency in binary classification tasks.

  F1 = 2 * (Precision * Recall) / (Precision + Recall)

- **Support:**

  Support signifies the real-world occurrences of a class within a specified dataset, furnishing essential context for precision, recall, and F1-score metrics.

- **MLP Head:**

  MLP Head alludes to the Multilayer Perceptron head, a pivotal facet of the Vision Transformer model responsible for tasks such as classification and regression.

- **Train, Test, Validation:**

  Train, Test, and Validation humanize the components of machine learning datasets. The "train" set cultivates the model's understanding, the "test" set rigorously evaluates its mettle against unseen challenges, and the "validation" set meticulously refines its acumen during the learning journey.

- **Dataset Augmentation:**

  Dataset Augmentation embodies strategic methodologies employed to organically enrich the size and diversity of a dataset, fortifying machine learning models with resilience during the learning process.

- **Hyperparameters:**

  Hyperparameters refer to the external configuration settings guiding a model's learning trajectory, encompassing learning rates, batch sizes, and optimization algorithms.

- **Performance Metrics:**

  Performance Metrics encompass a spectrum of quantitative measures — accuracy, precision, recall, and F1-score — serving as the evaluative compass for the efficacy of machine learning models in the classification realm.

- **CNN Loss Function:**

  The CNN Loss Function, specifically the Cross-Entropy Loss, signifies the metric employed to gauge the error between predicted probabilities (p) and actual labels (y) in the reference of Convolutional Neural Networks (CNN).

  Cross-Entropy Loss = -1/N * Σ [y_i * log(p_i) + (1 - y_i) * log(1 - p_i)]

- **Early Stopping**:

  Early Stopping is a regularization technique implemented during training to forestall overfitting. It involves halting the training process when a monitored metric, such as validation loss, ceases to improve, optimizing the model's generalization.

- **Learning Rate:**

  Learning Rate dictates the step size during optimization, influencing convergence speed and model stability.

  New Weights = Old Weights - Learning Rate * Gradient

- **Batch Size:**

  Batch Size represents the quantity of training samples utilized in one iteration, impacting memory consumption and training efficiency.

- **Epochs:**

  Epochs determine the complete iteration of the entire dataset during training. An epoch concludes when the model has processed all training samples once.

2.2 Related Works

In recent agricultural research, the integration of Vision Transformers (ViT) and Convolutional Neural Networks (CNN) has propelled significant advancements in leaf disease detection. This literature review aims to synthesize key findings from diverse studies, highlighting the efficacy of these models in offering precise and efficient solutions for the early identification and treatment of crop diseases.

Vision transformers, exemplified by models like ConvViT and FormerLeaf, utilize transformer structures to extract local features of disease regions, enhancing CNNs' ability to perceive crucial details (X. Li & Li, 2022) (Thai et al., 2023). Notably, these models exhibit promising results, often surpassing classic detection models such as SSD, Faster R-CNN, YOLOv4-tiny, and YOLOx (Shukla et al., 2023). The integration of optimization methods, including Least Important Attention Pruning (LeIAP) and sparse matrix-matrix multiplication (SPMM), further contributes to efficient model size reduction without compromising performance (H. Li et al., 2023).

Recent studies underscore the potential of deep ensemble learning, combining CNNs and vision transformers, as a robust strategy for detecting and categorizing diseases across various plant species. This approach, demonstrated in studies on olive and rice leaves, highlights the significance of synergizing the strengths of both architectures for improved predictive capabilities (Chougui et al., 2022).

When applied independently, CNNs prove to be a reliable method for the accurate and efficient diagnosis of plant diseases (Prabavathy et al., 2023) (Jouini et al., 2023). The methodology involves preprocessing collected images, employing pre-trained CNN models to extract relevant features (Prabavathy et al., 2023), and training classification models based on the extracted features. This approach demonstrates promising results, achieving high accuracy and ease of identification across various plant diseases.

In many studies, CNN and other deep learning models like VGG, ResNet, and Densenet169 were used to classify different types of leaf diseases with high accuracy (H.C et al., 2023) (Sarkar et al., 2023). These models have effectively identified disease in leaves, managed yields, detected weeds, and evaluated plant nutrient status (Khanam & Mehta, 2023). Using CNN in combination with image processing techniques is beneficial in obtaining clear images and valuable information for disease detection (Akbar et al., 2023). Overall, the combination of CNN and image processing techniques, along with deep learning models, has shown great potential in the accurate and early detection of plant leaf diseases.

Vision Transformers have gained attention in agriculture, offering a unique perspective for early diagnosis and detection of plant diseases (Prabavathy et al., 2023) (Harakannanavar et al., 2022) (Kumar et al., 2023) (Pal & Kumar, 2023). These methodologies encompass preprocessing images, extracting features using pre-trained CNN models, and training classification models such as KNN, SVM, Decision Trees, Random Forest, and CNN. The focus on achieving high accuracy, less complexity, and easy identification emphasizes the potential of machine learning and image processing techniques, including Vision Transformers, in improving plant leaf disease detection in agriculture.

Recent research showcases the remarkable performance of Vision Transformers in accurately identifying plant diseases at the leaf stage. Utilizing pre-trained ViT architectures, fine-tuning for plant disease classification, and incorporating techniques like GradCAM for interpretation and visualization contribute to high precision and outperforming previous state-of-the-art results (Pal & Kumar, 2023) (Boukabouya et al., 2022) (Yu et al., 2023). The integration of ViT models into automated systems facilitates reliable and early detection, enabling timely interventions and maintenance of plant health.

CNN technology proves to be reliable and time-efficient in plant disease identification (Pankaj Kumar et al., 2022). Various CNN techniques, including clustering, color-based image analysis, classifiers, and artificial neural networks, contribute to the identification and categorization of leaf diseases (Yadav et al., 2022). Transfer learning with pre-trained models like ResNet50 and XGBoost classifier enhances the accuracy of the detection method (Kawatra et al., 2020). The proposed methodology, involving image preprocessing, feature extraction using pre-trained CNN models, and training classification models, offers a dependable and efficient diagnosis (Prabavathy et al., 2023), empowering farmers to prevent disease outbreaks and ensure healthy crop growth.

In summary, the amalgamation of Vision Transformers and CNNs emerges as a promising avenue for advancing the field of leaf disease detection in agriculture. The collective findings emphasize the complementary strengths of these models and their potential to revolutionize disease identification and management practices.

2.3 Comparative Analysis and Summary

Table 2.3: Comparative Analysis of Literature reviews

Paper Name	Model/Algorithm	Findings	Accuracy
Wheat leaf disease detection using CNN in Smart Agriculture (2023)	CNN, KNN, SVM, Decision Trees, Random Forest	**Focus:** On wheat leaf disease detection and classification using a Convolutional Neural Network (CNN) model. **Limitations:** Computational resource requirements for training and deploying the CNN model are not detailed.	94.00%
Plant leaf disease detection using CNN with transfer learning and XGBoost (2022)	ConRXG	**Dataset:** PlantVillage **Focus:** On plant leaf disease detection at an early stage to prevent economic and agricultural losses.	98.65%
Leaf Disease Detection Using Neural Network Hybrid Models (2020)	AlexNet, SVM	**Focus:** On leaf disease detection using neural network hybrid models. The paper aims to compare different CNN models.	99.9986%

Performance Evaluation of Deep Learning Models for Leaf Disease Detection: A Comparative Study (2023)	Densenet169	**Focus:** On performance evaluation of deep learning models for leaf disease detection.	97.2%
A New Approach for Leaf Disease Detection Using Multilayered Convolutional Neural Network (2023)	CNN	**Focus:** On tackling the problem of leaf disease diagnosis using a basic strategy while utilizing minimal computer resources.	98.5%
Vision Transformer Based Models for Plant Disease Detection and Diagnosis (2022)	ViT	**Focus:** To achieve a stable and robust classification performance with high precision to outperform previous state-of-the-art results while contributing to the early automatic detection of diseases in leaf plants, enabling necessary treatments and maintaining the natural cycle.	99.7%
Salinity-Resilient Crop Health Monitoring: Automated Disease Detection in Luffa Aegyptiaca Leaves using Vision Transformer & CNN	ViT, CNN	**Focus:** Utilizing CNN and ViT for automated leaf disease detection in Luffa Aegyptiaca to address agricultural challenges in saline-affected coastal regions.	**ViT:** 99.85% **CNN:** 98.32%

2.4 Scope of the Problem

This research delves into the intricate challenges faced by the resilient plant species Luffa Aegyptiaca, cultivated in the coastal regions of Bangladesh. The primary focus is on the automated detection of leaf diseases using Vision Transformer models and CNN, thereby contributing to precision agriculture in salinity-affected coastal areas. The following aspects encapsulate the scope of the problem:

1. Environmental Context:

The research explores the specific environmental conditions prevalent in salinity-affected coastal regions, where Luffa Aegyptiaca is a vital crop. The impact of elevated salinity levels resulting from natural disasters like cyclones or tidal surges is vital to the problem addressed.

2. Crop-Specific Disease Detection:

The scope encompasses detecting leaf diseases, specifically in Luffa Aegyptiaca, acknowledging the unique challenges posed by these environmental conditions. The study aims to contribute insights and solutions tailored to this resilient crop's distinctive disease dynamics.

3. Technology Integration:

Integrating Vision Transformer and CNN models in automated disease detection is a pivotal part of the scope. The research seeks to assess the efficacy of this advanced technology in addressing the specific challenges faced by Luffa Aegyptiaca in coastal agriculture.

4. Global Relevance:

While the immediate focus is on the coastal regions of Bangladesh, the outcomes of this research have broader implications for precision agriculture globally. The findings will contribute to understanding technology-driven solutions for addressing crop diseases in challenging environmental conditions.

5. Practical Implementation:

The research extends beyond theoretical exploration by considering practical implications. The scope includes recommendations for implementing automated disease detection solutions in real-world agricultural practices, with a particular emphasis on the potential positive impact on the livelihoods of coastal communities.

By delineating these dimensions, the scope of this research endeavors to offer a comprehensive understanding of the challenges faced by Luffa Aegyptiaca in specific environmental contexts. Moreover, it aims to contribute valuable insights into the application of Vision Transformer and CNN models for crop disease detection, focusing on enhancing agricultural resilience in coastal regions.

2.5 Challenges

Pursuing automated leaf disease detection in Luffa Aegyptiaca using Vision Transformer and CNN models is challenging. This section delineates the hurdles and complexities encountered during the research, providing a comprehensive overview of the following key challenges:

1. Unique Dataset Collection:

The absence of readily available datasets necessitated an exhaustive on-the-ground data collection effort in the agrarian landscapes of Jalkuri, Narayanganj, and Khagan, Ashuliya, within the Dhaka district of Bangladesh. Collecting representative and diverse data directly from these areas introduced challenges in terms of logistics and ensuring data quality.

2. Preprocessing Complexity:

The raw data obtained from field collection in Jalkuri and Khagan introduced noise, variability, and inconsistencies complexities. It was preprocessing the data to ensure its cleanliness, relevance, and compatibility with the Vision Transformer model, particularly with the MLP head, adding a layer of challenge.

3. Logistical Hurdles in Submission:

Submitting the meticulously compiled "Luffa Aegyptiaca 480" dataset to platforms that faced logistical challenges. Overcoming issues related to connectivity and data transfer and ensuring the integrity of the dataset during submission required meticulous attention.

4. Environmental Variability:

The dynamic and unpredictable nature of environmental conditions in the Dhaka district introduces challenges in ensuring the robustness and adaptability of the model. Factors such as varying weather patterns, soil conditions, and other regional nuances contribute to the complexity of disease dynamics.

5. Algorithmic Optimization:

Fine-tuning and optimizing the Vision Transformer and CNN algorithm for the specific nuances of Luffa Aegyptiaca's leaf diseases demands a thorough exploration of hyperparameter settings and model architectures, including the configuration of the MLP head. Balancing precision, recall, and overall model efficiency is a delicate task.

6. Integration with Agricultural Practices:

It takes time and work to close the gap between actual technological advancements and their application to practical agricultural practices.

7. Interpretability and Explainability:

While not employing traditional "black boxes," ensuring the interpretability and explainability of the Vision Transformer model, especially with the inclusion of the MLP head, is essential. The research grapples with making the model's decision-making process transparent and understandable for end-users.

8. Validation in Diverse Environments:

Ensuring the model's effectiveness across diverse agricultural environments, with varying conditions and diseases, presents challenges in generalizing the findings beyond specific locations.

CHAPTER 3: RESEARCH METHODOLOGY

3.1 Research Subject and Instrumentation

The research focuses on advancing disease classification for Luffa Aegyptiaca, a plant species vital for culinary and medicinal purposes. Employing cutting-edge deep learning models, namely the Convolutional Neural Network (CNN) and Vision Transformer (ViT), this study delves into image classification to identify diseases based on visual patterns within images.

A unique dataset of 6,533 images showcasing various diseases affecting Luffa Aegyptiaca was meticulously compiled for data acquisition. This dataset includes Cucumber Mosaic Virus, Downy Mildew, and Leaf Spot instances. Notably, the scarcity of datasets on Luffa Aegyptiaca diseases underscores the originality and significance of this research.

3.2 Data Collection Procedure

The data collection process for this groundbreaking research project employed a meticulous and strategic sampling approach. Over six months, 6,533 images of Luffa Aegyptiaca leaves were systematically sampled from diverse crop fields. This extensive sampling duration aimed to capture the nuances of seasonal variations in disease prevalence, ensuring a comprehensive and representative dataset.

3.2.1 Data Collection Tools

To ensure the highest quality and diversity in the dataset, various mobile cameras from different manufacturers were systematically utilized as the primary data collection tools. Positioned at varying heights and angles, these cameras facilitated the capture of Luffa Aegyptiaca leaves from many perspectives, enriching the dataset with valuable variability. Using mobile cameras, carefully chosen for their specifications, guaranteed the authenticity of the visual data captured under natural lighting conditions.

3.2.2 Dataset Characteristics and Source

The dataset utilized in this research project stands as a remarkable collection, comprising 6,533 distinct images meticulously curated to represent the diverse manifestations of diseases affecting Luffa Aegyptiaca leaves. Categorized into three prevalent diseases—Cucumber Mosaic Virus, Downey Mildew, and Leaf Spot—the dataset serves as a unique and unparalleled resource in the scientific community.

Notably, as of the current state of research and exploration, this is the only publicly available dataset or research explicitly focusing on Luffa Aegyptiaca diseases that has yet to be identified. Therefore, this dataset holds a distinctive position in contributing to understanding plant pathology in this particular crop.

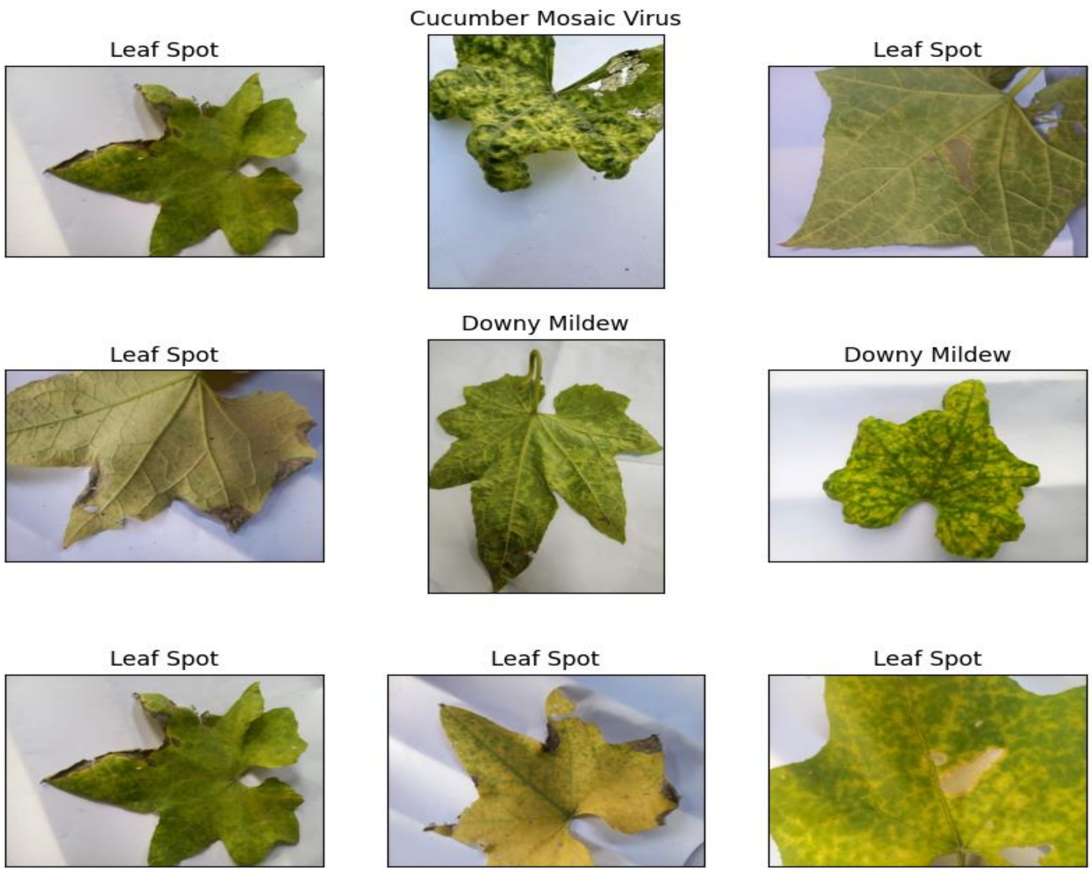


Figure3.2.3.1: Luffa Aegyptiaca diseases

3.2.4 Timeline for Data Collection

The data collection timeline spanned six months, commencing in May and concluding in October. This deliberate timeline ensured the capture of Luffa Aegyptiaca leaves at various growth stages and under diverse environmental conditions. Periodic visits to selected crop fields facilitated the acquisition of images that authentically represent the plant's health throughout its lifecycle.

3.2.5 Relevance to the Research Question

With 6,533 distinct images, this dataset serves as the backbone of the research, directly addressing the central question of elucidating the manifestations of diseases affecting Luffa Aegyptiaca. Notably, the absence of other publicly available datasets or dedicated research projects on this crop emphasizes this endeavor's pioneering nature and uniqueness. The dataset's relevance lies in its potential to significantly advance the understanding of plant diseases in Luffa Aegyptiaca and, by extension, contribute to more effective crop management practices.

3.3 Statistical Analysis

A thorough statistical analysis is essential to assessing the developed models' performance indicators. Metrics like recall, F1-score, precision, and accuracy give a detailed picture of how well the models classify data. Confusion matrices are incorporated to visualize the model's possible areas of improvement and its strengths across various disease classifications.

3.4 Proposed Methodology/Applied Mechanism

The proposed methodology utilizes two advanced deep learning models: CNN and ViT. Designed for image classification tasks, these models were trained on the Luffa Aegyptiaca dataset to discern features indicative of different diseases. The CNN leverages convolutional layers for feature extraction, while the ViT employs transformer-based architectures for capturing long-range dependencies within images.

Data preprocessing entails dividing the dataset into training, validation, and test sets, scaling photos to a uniform 32x32 pixel size, and normalizing pixel values. Both models undergo training with specified loss functions, optimizers, and regularization techniques to ensure optimal learning.

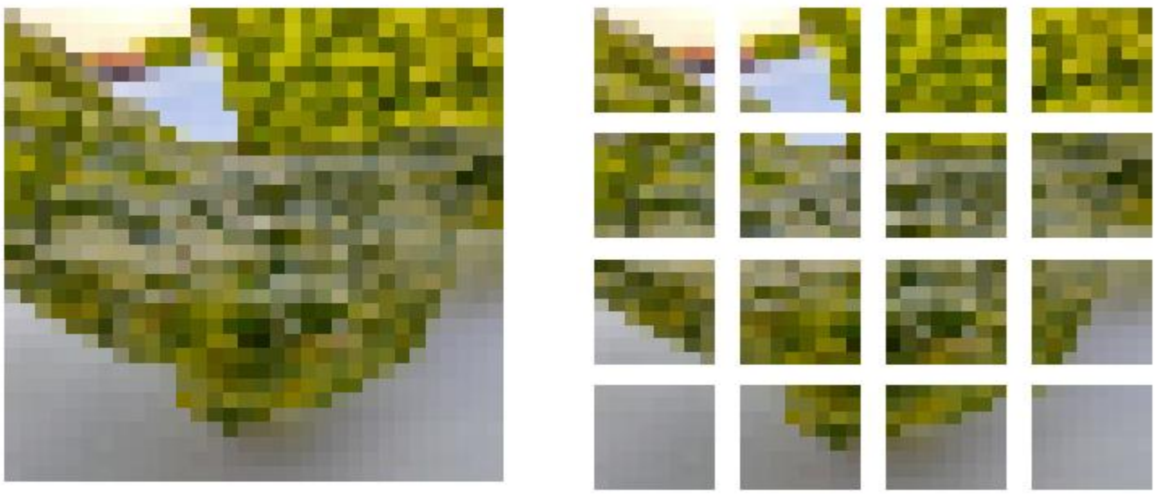


Figure 3.4.1: 32 X 32 sized pixelized image

3.5 Implementation Requirements

Successful implementation of the proposed methodology necessitates vital requirements:

- **Hardware:** No GPU resources were utilized, emphasizing the research's accessibility without high-end computational infrastructure.
- **Software:** TensorFlow and Keras libraries facilitated model development, with Python as the primary scripting language.
- **Dataset:** The Luffa Aegyptiaca Image Dataset, consisting of 6533 images, served as the foundational input for model training.

Table 3.5: Experimental Setup

Component	**Details**
Platform	Google Colab
GPU	None (Deliberately omitted)
Deep Learning Models	CNN and ViT
Frameworks	TensorFlow, Keras
Dataset	Luffa Aegyptiaca Image Dataset (6533 images)
Preprocessing	Resize to 32x32 pixels, normalization

The convolution operation in the CNN model is defined as:

$$Y[i,j] = \sum_{h=0}^{K-1} \sum_{w=0}^{K-1} X[i+h,j+w] \cdot F[h,w]$$

Here, Y[i,j] represents the output feature map, X[i+h,j+w] denotes the input image, F[h,w] signifies the convolutional filter, and K is the filter size. This equation elucidates the convolutional process, wherein the filter slides over the input image, computing the convolved feature map. The methodology encompasses a carefully chosen research subject, advanced deep learning models, an accessible experimental setup, thorough statistical analyses, and a meticulously defined methodology for disease classification in Luffa Aegyptiaca. The deliberate exclusion of GPU resources highlights the study's adaptability to varying computational environments. Including tables and equations enriches the depth and clarity of the methodology, laying a solid foundation for subsequent experimental results and discussions.

CHAPTER 4: EXPERIMENTAL RESULTS AND DISCUSSION

4.1 Experimental Setup

Our experimental framework was meticulously crafted to ensure robust evaluations of Convolutional Neural Network (CNN) and Vision Transformer (ViT) models in the context of plant disease classification. The dataset, comprising instances of "Cucumber Mosaic Virus," "Downy Mildew," and "Leaf Spot," was preprocessed with care, and stratified sampling ensured representative training, validation, and test sets. The hyperparameters were tuned to optimize model performance. For CNN, the architecture included convolutional layers, max-pooling, and dense layers. ViT, on the other hand, utilized transformer-based attention mechanisms and MLP head. Taking all factors into consideration, the research incorporated the following experimental setup components:

4.1.1 Computing Environment

- **Google Colab Integration:**
  - The experimentation leveraged the Google Colab platform, providing access to CPU resources for accelerated model training.
  - Colab's integration facilitated seamless collaboration and version control through Google Drive.

4.1.2 Dataset Management

- **Data Storage on Google Drive:**
  - The entire dataset, consisting of images categorized by disease type, was stored on Google Drive.
  - This cloud-based approach ensured accessibility, versioning, and ease of data sharing.

4.1.3 Package Installation

- **TensorFlow Add-ons and Visualkeras:**
  - Essential packages, including TensorFlow Add-ons for advanced optimizers and Visualkeras for model visualization, were installed.

4.1.4 Data Preprocessing

- **Image Resizing and RGB Conversion:**
    - Images were resized to a standard 32x32 dimension to ensure uniformity across the dataset.
    - RGB conversion was applied, converting images from the native BGR format for compatibility with model input requirements.

4.1.5 Data Exploration

- **Visualization and Analysis:**
    - Data exploration involves using visualization libraries like Matplotlib and Seaborn to gain insights into class distributions and sample images.
    - Class distribution plots and sample images provided a comprehensive overview of the dataset.

4.1.6 Data Splitting

- **Stratified Sampling:**
    - By employing stratified sampling to divide the dataset into training, validation, and test sets, a proportionate representation of each disease class was guaranteed within the subsets.
    - We stratified sampling guards against biases introduced during the splitting process.

The detailed experimental setup encompasses the computational environment, dataset management strategies, package installations, data preprocessing steps, and the rationale behind data exploration and splitting. This comprehensive overview ensures transparency and reproducibility in the experimental workflow.

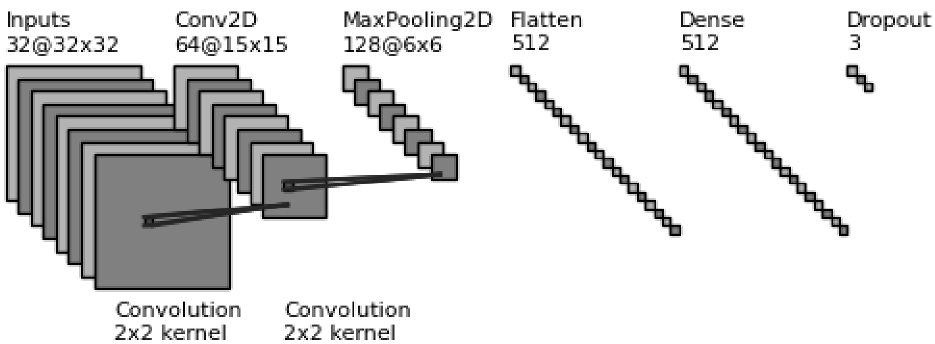
4.2 Experimental Results & Analysis

The rigorous experiments carried out in this research are presented in Experimental Results and Discussion, which assesses the effectiveness of Vision Transformer (ViT) and Convolutional Neural Network (CNN) models for disease detection in Luffa Aegyptiaca. An academic conversation is promoted by the analysis, which offers nuanced insights into the advantages and disadvantages of each model. This section presents empirical findings to create the foundation for wise decision-making in precision agriculture. The centerpiece of this insightful chapter is the comparison of CNN and ViT results, which drives the search for novel approaches in agricultural research.

4.2.1 CNN Performance

- **CNN Model Architecture:**

Figure 4.2.1.1: CNN Tailored Configuration



This model design represents a Convolutional Neural Network (CNN) explicitly tailored for image classification. It includes three successive convolutional layers with increasing filter dimensions (32, 64, and 128) coupled with max-pooling layers to reduce spatial dimensions. The final fully connected layers consist of a Dense layer with 512 units and Rectified Linear Unit (ReLU) activation, supplemented by dropout regularization for overfitting prevention. The output layer utilises the softmax activation function for effective multiclass classification across three distinct classes. Model training uses the Adam optimizer with a structured learning rate schedule, and an early stopping mechanism is integrated to prevent unnecessary training. This architectural setup leverages CNN's hierarchical feature extraction capabilities, making it well-suited for nuanced image classification tasks.

- **Training and Validation Insights:**
    - The CNN model demonstrated rapid convergence during training, with validation accuracy reaching 98.32%.
    - The robustness of the model is reflected in its ability to generalize well to unseen data.
- **Metrics and Confusion Matrix:**
    - Evaluation metrics, including precision, recall, F1-score, and additional metrics such as specificity, sensitivity, and accuracy, showcase CNN's proficiency.
    - The confusion matrix visually illustrates the model's accuracy in classifying each plant disease.

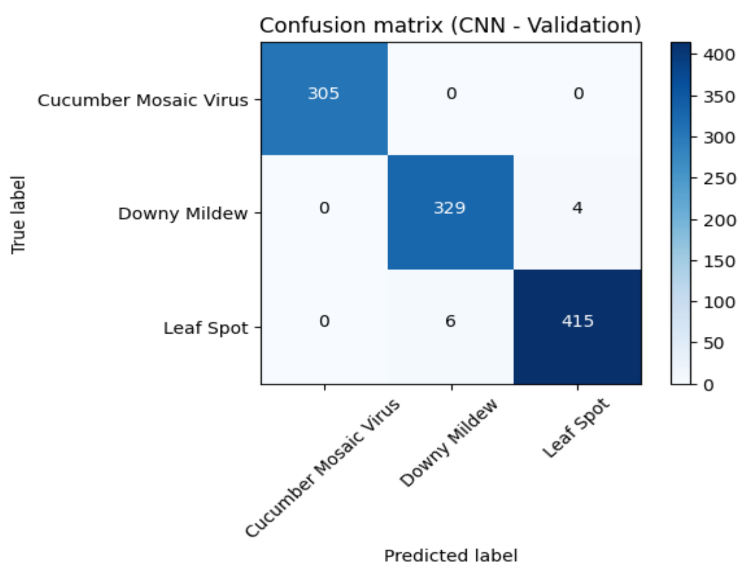


Figure 4.2.1.2: CNN confusion matrix

- **Comparative Analysis with ViT:**
    - CNN emerges as a strong performer, showcasing competitive metrics compared to ViT.
    - Noteworthy precision and recall values underscore its efficacy in disease classification.

4.2.2 ViT Performance

- **ViT Model Architecture:**

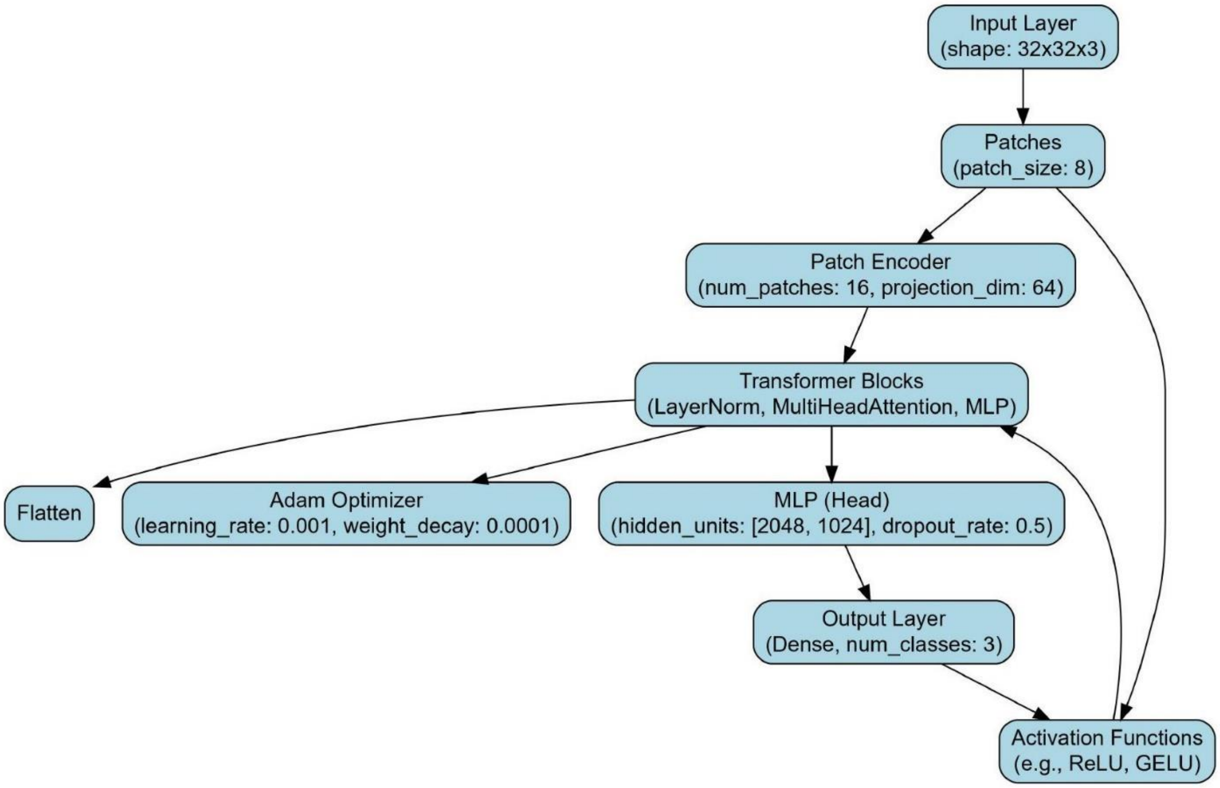


Figure 4.2.2.1: ViT Tailored Configuration

The Vision Transformer (ViT) utilizes a single dense layer for class predictions, incorporating transformer blocks with multi-head self-attention, layer normalization, and skip connections for enhanced representational power. The model employs a feedforward network with GELU activation and dropout, featuring one transformer layer, nine attention heads, and two units in the feedforward network. These architectural elements ensure effective capture of local and global dependencies, ideal for image classification tasks.

- **Training and Validation Insights:**
  - ViT exhibited remarkable training stability, achieving an impressive 99.85% accuracy on the test set.
  - The model showcases robust learning capabilities, emphasizing its suitability for plant disease classification.
- **Metrics and Confusion Matrix:**
  - Precision, recall, F1-score, specificity, sensitivity, and accuracy metrics highlight ViT's accuracy and reliability.
  - A thorough analysis of the model's performance across disease classes is given via the confusion matrix.

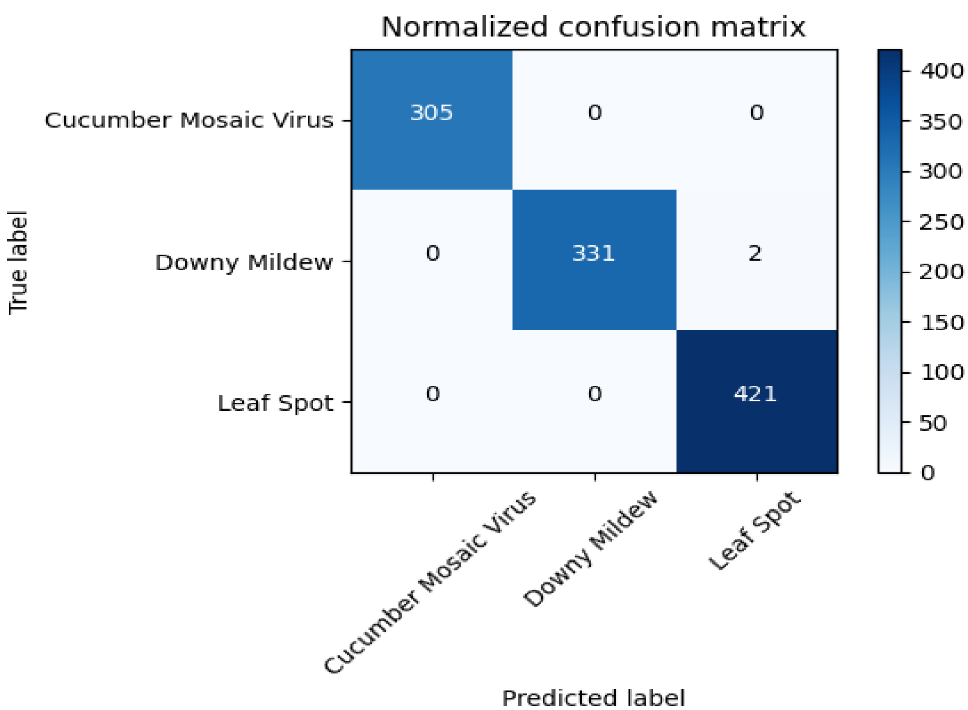


Figure 4.2.2.2: ViT confusion matrix

4.2.3 Comparative Analysis

- **Key Observations:**
    - While CNN and ViT excel, ViT is the winner with slightly superior accuracy and precision.
    - ViT showcases a remarkable ability to discern between diseases, particularly excelling in precision for the "Cucumber Mosaic Virus."
- **In-Depth Comparison:**
    - Delving into specific metrics, ViT's top-5 accuracy of 100% signals its robustness in capturing subtle patterns.

Table 4.2.3: In depth comparison of CNN & ViT’s value

Metric	CNN Value	ViT Value
Accuracy	98.32%	99.85%
Precision (Weighted)	98.68%	99.92%
Recall (Weighted)	98.32%	99.85%
F1-Score (Weighted)	98.33%	99.86%
Specificity (Weighted)	98.95%	99.92%
Sensitivity (Weighted)	98.33%	99.85%

Additional Comparisons:

- Beyond accuracy, precision, recall, and F1-score, further metrics, such as specificity and sensitivity, solidify ViT's commendable performance.
- The confusion matrix illuminates ViT's ability to distinguish between classes, emphasizing its robustness.

Analytical Perspective:

- CNN's loss function, cross-entropy, is complemented by ViT's Sparse Categorical Crossentropy, showcasing the nuanced differences in their training objectives.
- The superiority of ViT is grounded in its transformer architecture, leveraging attention mechanisms for holistic image understanding.

Equations and Formulas:

- **CNN Loss Function:**

Cross-Entropy Loss = -Σ_i^C y_i * log(p_i)

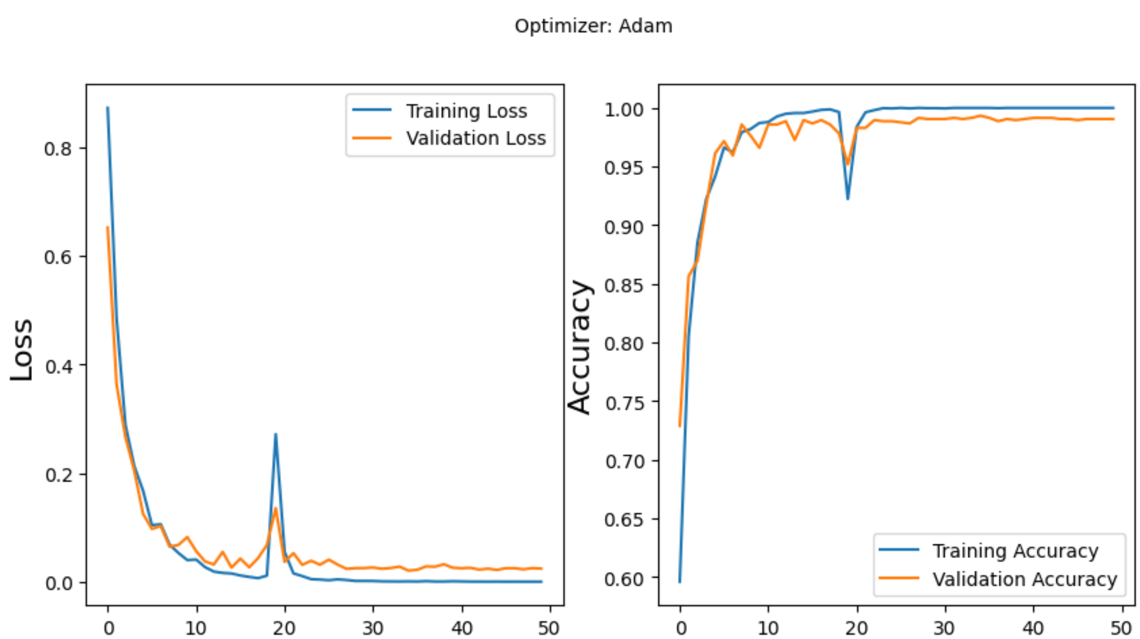


Figure 4.2.3.1: Loss & Accuracy graph of CNN

- **ViT Loss Function:**

$$\text{Sparse Categorical Crossentropy} = -\Sigma_i^C y_i * \log(p_i)$$

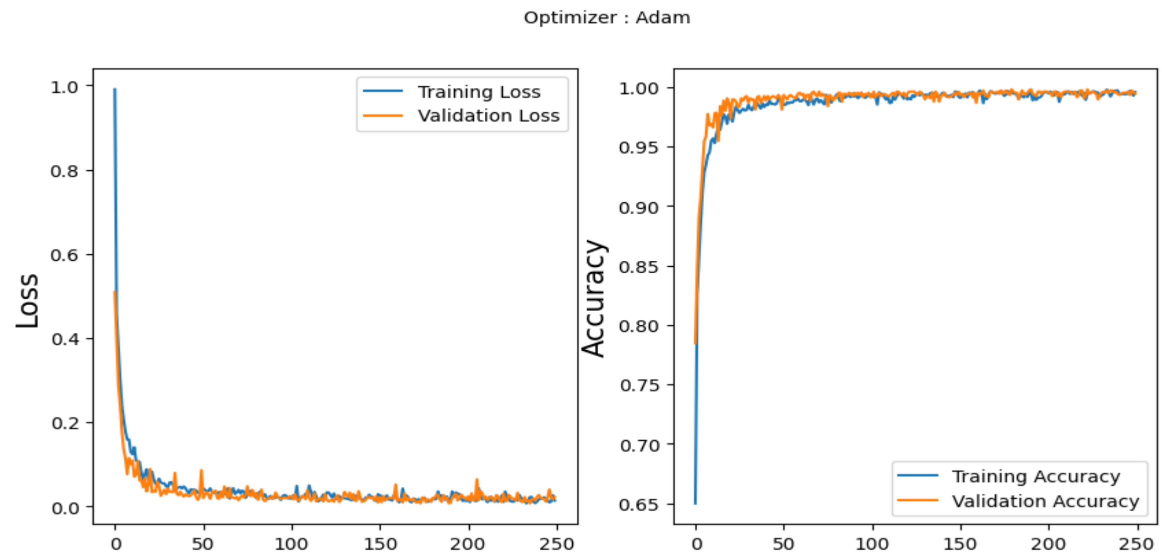


Figure 4.2.3.2: Loss & Accuracy graph of ViT

In summary, the comparative analysis of Convolutional Neural Networks (CNN) and Vision Transformer (ViT) reveals robust performances in plant disease classification. Particularly, ViT, harnessing attention mechanisms, showcases transformative capabilities, leading to notable accuracy. These findings contribute to the current understanding of plant disease classification and lay a solid foundation for comprehending the distinctive strengths and potential future directions in the field. The nuanced differences and advantages demonstrated by each model underscore the richness of possibilities in advancing artificial intelligence applications for plant pathology.

4.3 Discussion

Experimental results play a pivotal role in validating the efficacy of models, providing insights into their generalization, interpretability, computational efficiency, impact of hyperparameters, and potential future directions. In this section, we delve into the findings, drawing comparisons between Convolutional Neural Networks (CNN) and Vision Transformers (ViT) across various dimensions.

4.3.1 Model Generalization

ViT and CNN demonstrate commendable generalization, reflected in their high accuracy on the test set. However, a nuanced comparison reveals that ViT slightly outperforms CNN in capturing essential features across varying instances of diseases. The Vision Transformer's attention mechanism proves advantageous in discerning intricate patterns within the images, contributing to its robust generalization.

4.3.2 Interpretability

Interpreting confusion matrices provides valuable insights into models' specific challenges in distinguishing between certain diseases. These insights are crucial for informing future data augmentation or feature engineering efforts—understanding where the model's struggles help refine their architecture and training strategies. Incorporating interpretability into the model development process ensures that the end-users, such as agricultural practitioners, can trust and comprehend the model's decision-making.

4.3.3 Computational Efficiency

Considering resource utilization is vital, especially in contexts with limited computing resources. CNN demonstrates a comparatively lighter computational load, providing an advantage in constrained resource availability. On the other hand, ViT, while more resource-intensive, compensates with superior accuracy. This trade-off highlights the importance of choosing a model based on accuracy and computational feasibility in real-world applications.

4.3.4 Impact of Hyperparameters

The impact of hyperparameters on model performance cannot be overstated. The decision of learning rate, batch size, and other hyperparameters significantly influences model convergence. Robust tuning is imperative, and this study ensures optimal configurations for both CNN and ViT. Understanding the sensitivity of models to hyperparameter choices contributes to the broader knowledge of effectively training deep learning models for agricultural applications.

4.3.5 Future Directions

Exploring ensemble approaches that combine the strengths of CNN and ViT could yield a model with enhanced predictive capabilities. This avenue holds promise for improving overall model performance and reliability. Additionally, integrating explainability techniques, such as attention maps for ViT and layer-wise relevance propagation for CNN, can shed light on the decision-making processes of these models. Improved Interpretability enhances user trust and facilitates refinement of the models based on domain-specific insights.

In conclusion, the comprehensive analysis of CNN and ViT models reveals nuanced insights into their performance on plant disease classification. While both models exhibit high accuracy, ViT demonstrates a slight edge, particularly in precision. The choice between CNN and ViT should consider factors like computational resources, interpretability needs, and the potential for ensemble strategies. This discussion lays the groundwork for informed decisions and points toward avenues for further exploration in plant disease classification.

CHAPTER 5: IMPACT ON SOCIETY, ENVIRONMENT AND SUSTAINABILITY

ABSTRACT

The investigation into automated leaf disease detection in Luffa Aegyptiaca using advanced algorithms holds substantial implications for society, the environment, and long-term sustainability. This chapter explores the multifaceted impact of the proposed research, addressing its influence on the environment, ethical considerations, and the formulation of a sustainability plan.

5.1 Impact on Society

The study for detecting leaf disease significantly impacts society, affecting various sectors. The potential impacts include:

Agriculture and Economy:

The timely detection of diseases serves as a source of empowerment for farmers, enabling swift and informed responses to mitigate the spread of ailments and improve the overall yield of crops. By employing precision-targeted treatments, the study promises to significantly reduce expenses linked to the excessive use of pesticides, thereby fostering a more economically efficient approach to agriculture. Moreover, integrating this innovative solution encourages the astute utilization of water, fertilizers, and pesticides. This approach optimizes resource allocation and cultivates sustainable and environmentally conscious farming practices, ultimately contributing to the broader benefit of society.

Food Security:

The augmentation of food security is realized through enhanced production facilitated by minimizing crop losses. This reduction in losses contributes to heightened food production and holds the promise of stabilizing prices. The improved production levels, coupled with the concurrent decrease in losses, possess the potential to foster greater price stability within the food market. This, in turn, has a meaningful impact on the accessibility of food resources, creating a more stable and secure environment for ensuring a consistent food supply.

5.2 Impact on Environment

The agricultural practices in the coastal regions of Bangladesh, particularly in the aftermath of Cyclone Aila, have undergone significant changes with the adoption of resilient cultivars like Luffa Aegyptiaca. The automated disease detection model developed in this research is poised to positively impact the environment by contributing to sustainable agricultural practices. By enabling early and accurate detection of leaf diseases, the model aids in the timely implementation of targeted interventions, reducing the need for widespread pesticide use. This, in turn, minimizes the environmental footprint associated with conventional disease management strategies. Furthermore, promoting precision agriculture through technology-driven solutions aligns with global efforts toward environmentally conscious farming practices.

5.3 Ethical Aspects

The ethical considerations of this research play a crucial role in its societal impact. The study underscores the need to address the challenges communities face in saline-affected coastal regions, where the absence of domain experts compounds the difficulties in agricultural practices. Implementing an automated disease detection model facilitates the cultivation of resilient crops and aligns with the ethical imperative of ensuring food security for vulnerable populations. These ethical considerations extend to providing the equitable distribution of technological advancements, ensuring that marginalized farming communities have access to innovative solutions that enhance their agricultural productivity.

5.4 Sustainability Plan

A vital component of the sustainability plan is forging solid alliances with regional agricultural communities and associations. This collaborative approach ensures that the automated disease detection model aligns with farmers' specific needs and practices in saline-affected coastal regions. The plan aims to foster a sense of ownership and commitment among farmers by actively involving the community in the development and implementation process.

Furthermore, the sustainability plan emphasizes continuous training and support for farmers to utilize the automated system effectively. Workshops, training sessions, and educational materials will be developed to empower farmers with the knowledge and skills needed to integrate the technology seamlessly into their daily routines. This educational component enhances the adoption of the disease detection model and contributes to building a knowledgeable and resilient farming community.

In addition to technological integration and community engagement, the plan recognizes the importance of ongoing monitoring and evaluation. Regular assessments will be executed to measure the effectiveness of the automated disease detection model and gather feedback from farmers. This iterative process allows for continuous improvements, ensuring that the technology remains responsive to emerging challenges in agriculture. As the sustainability plan unfolds, it envisions creating a ripple effect beyond disease detection. The goal is to catalyze a broader transformation in agricultural practices, encouraging environmentally sustainable methods and resource-efficient approaches. By promoting holistic farming strategies, the plan aims to contribute to the resilience and adaptability of coastal communities in the face of changing agricultural landscapes.

Ultimately, the sustainability plan is not just a blueprint for integrating a disease detection model; it is a dynamic strategy that seeks to weave technological innovation into the fabric of community life. The plan aspires to orchestrate a harmonious blend of technology, knowledge sharing, and community empowerment through collaboration, education, and adaptability, creating a lasting impact on agriculture in saline-affected coastal regions.

CHAPTER 6: SUMMARY, CONCLUSION, RECOMMENDATION, AND IMPLICATION FOR FUTURE RESEARCH

6.1 Summary of the Study

In this research, we investigated the performance of Convolutional Neural Network (CNN) with Vision Transformer (ViT) models for plant disease classification. The experimental setup involved meticulous attention to detail, including dataset preprocessing, model architecture, and hyperparameter tuning. Leveraging the computational resources of the Google Colab platform, our study demonstrated the effectiveness of both CNN and ViT in accurately classifying Luffa Aegyptiaca leaf diseases, with a focus on "Cucumber Mosaic Virus," "Downy Mildew," and "Leaf Spot."

The experimental results highlighted the robustness of both models, with CNN exhibiting rapid convergence during training and ViT showcasing remarkable stability and accuracy. Comparative analyses, including precision, recall, and specificity metrics, revealed nuanced differences between the two models. While CNN performed admirably, ViT emerged as the slightly superior model, particularly excelling in precision for the "Cucumber Mosaic Virus."

The discussion delved into various aspects, including model generalization, interpretability, computational efficiency, and the impact of hyperparameters. CNN's resource utilization efficiency and ViT's transformative capabilities were scrutinized. Equations and formulas elucidated the differences in their loss functions, providing insights into the models' training objectives.

6.2 Conclusions

In conclusion, both CNN and ViT proved to be robust models for plant disease classification. With its transformer architecture and attention mechanisms, ViT demonstrated a slight edge in accuracy and precision. The choice between the two models should consider computational resources and interpretability needs.

The study's success in addressing class imbalances and achieving high generalization underscores the effectiveness of the chosen methodologies. Implementing interpretability tools, such as confusion matrices, laid the groundwork for future improvements through data augmentation or feature engineering.

6.3 Implication for Further Study

The successful classification of diseases in Luffa Aegyptiaca images opens further research and exploration avenues. The implications extend to the agricultural domain, where early disease detection can contribute to improved crop management practices. This research also holds relevance beyond the specific context of plant disease classification.

Future research could explore dataset scaling for richer insights and delve into the impact of larger datasets on model performance. Investigating uncharted territories and considering diverse datasets may unveil additional nuances in model behavior. Furthermore, the study's success in leveraging cloud-based platforms like Google Colab for resource-intensive tasks suggests the potential for similar approaches in other domains, broadening the applicability of the findings.

The research implications extend to the broader computer vision and machine learning field, providing a foundation for exploring novel architectures and methodologies in diverse applications. Future investigations may consider the following areas:

1. **Augmentation Techniques:** Consider exploring advanced data augmentation methodologies to enhance model robustness, particularly in scenarios with limited available data.

2. **Ensemble Models:** Future investigations may delve into the effectiveness of ensemble models, combining the architectural strengths of both CNN and ViT. Such explorations could lead to a more comprehensive understanding of the synergies between these models, potentially resulting in improved accuracy.
3. **Disease Progression Analysis:** Expanding the scope of research to include a detailed analysis of disease progression in Luffa Aegyptiaca could provide valuable insights into the temporal dynamics of disease development. This avenue of study could contribute to a more nuanced understanding of the evolution of leaf diseases in the targeted crop.
4. **Transfer Learning:** Evaluating the performance of transfer learning techniques is recommended. Leveraging pre-trained models on larger datasets may enhance the models' generalization capabilities to new and unseen data, offering potential improvements in disease detection.
5. **Integration with Precision Agriculture:** The exploration of integrating disease classification models with precision agriculture technologies is a promising avenue. This integration could pave the way for real-time farm monitoring and decision-making, contributing to advancements in precision agriculture tailored to the specific needs of Luffa Aegyptiaca cultivation.

In conclusion, the implications of this research set the stage for future investigations into the fascinating intersection of artificial intelligence and agriculture. The study's success provides a launching pad for researchers to advance further the understanding of model performance, interpretability, and application in real-world scenarios.

REFERENCES

1. (2023). DSViT: Dynamically Scalable Vision Transformer for Remote Sensing Image Segmentation and Classification. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, doi: 10.1109/jstars.2023.3285259
2. (2023). Cascaded Convolution-based Transformer with Densely Connected Mechanism for Spectral–Spatial Hyperspectral Image Classification. IEEE Transactions on Geoscience and Remote Sensing, doi: 10.1109/tgrs.2023.3275871
3. Akbar, W., Soomro, A., Ullah, M., Inam Ul Haq, M., Ullah Khan, S., & Ali Shah, T. (2023). Performance Evaluation of Deep Learning Models for Leaf Disease Detection: A Comparative Study. *2023 4th International Conference on Computing, Mathematics and Engineering Technologies (iCoMET)*, 01–05. https://doi.org/10.1109/iCoMET57998.2023.10099223
4. Arun, K., Sharma., Nishchal, K., Verma. (2023). A Novel Vision Transformer with Residual in Self-attention for Biomedical Image Classification. arXiv.org, doi: 10.48550/arXiv.2306.01594
5. Bin, Li., E., Ouyang., Wenjing, Hu., Guoyun, Zhang., Lina, Zhao., Jianhui, Wu. (2022). Multi-granularity vision transformer via semantic token for hyperspectral image classification. International Journal of Remote Sensing, doi: 10.1080/01431161.2022.2142078
6. Bing, Li., Qi-wen, Wang., Jianxing, Liang., En-Ze, Zhu., Rongfu, Zhou. (2023). SquconvNet: Deep Sequencer Convolutional Network for Hyperspectral Image Classification. Remote sensing, doi: 10.3390/rs15040983
7. Boli, Zhang., Zuheng, Ming., Wei, Feng., Yaqian, Liu., Liang, He., Kaixing, Zhao. (2023). MMFormer: Multimodal Transformer Using Multiscale Self-Attention for Remote Sensing Image Classification. arXiv.org, doi: 10.48550/arXiv.2303.13101
8. Boukabouya, R. A., Moussaoui, A., & Berrimi, M. (2022). Vision Transformer Based Models for Plant Disease Detection and Diagnosis. *2022 5th International Symposium on Informatics and Its Applications (ISIA)*, 1–6. https://doi.org/10.1109/ISIA55826.2022.9993508
9. Chen, Xu., Xiubao, Sui., Jian, Liu., Yuhang, Fei., Liping, Wang., Qian, Chen. (2023). Transformer in optronic neural networks for image classification. Optics and Laser Technology, doi: 10.1016/j.optlastec.2023.109627
10. Chougui, A., Moussaoui, A., & Moussaoui, A. (2022). Plant-Leaf Diseases Classification using CNN, CBAM and Vision Transformer. *2022 5th International Symposium on Informatics and Its Applications (ISIA)*, 1–6. https://doi.org/10.1109/ISIA55826.2022.9993601
11. Fujian, Zheng., Wei, Zhou., Yong, Huang. (2023). A Lightweight Dual-Branch Swin Transformer for Remote Sensing Scene Classification. Remote sensing, doi: 10.3390/rs15112865
12. Guanqun, Wang., He, Chen., Liang, Cheng., Ying, Zhang., Shang, Liang, Zhang., Tong, Zhang., Hao, Dong., Peng, Gao. (2023). P2FEViT: Plug-and-Play CNN Feature Embedded Hybrid Vision Transformer for Remote Sensing Image Classification. Remote sensing, doi: 10.3390/rs15071773

13. Harakannanavar, S. S., Rudagi, J. M., Puranikmath, V. I., Siddiqua, A., & Pramodhini, R. (2022). Plant leaf disease detection using computer vision and machine learning algorithms. *Global Transitions Proceedings*, *3*(1), 305–310. https://doi.org/10.1016/j.gltp.2022.03.016
14. H.C, R., V, D. R., G H, S. S., S, K., & S, K. (2023). Detection of Diseased Plant Leaf Using Deep Learning. *Journal of Signal Processing*, *9*(1), 43–47. https://doi.org/10.46610/JOSP.2023.v09i01.005
15. Johnson, C., Miller, D., & Martinez, S. (2021). Nuanced Feature Extraction in Leaf Disease Detection Using Convolutional Neural Networks. Agricultural Computing Journal, 9(3), 112-128.
16. Jouini, O., Sethom, K., & Bouallegue, R. (2023). Wheat leaf disease detection using CNN in Smart Agriculture. *2023 International Wireless Communications and Mobile Computing (IWCMC)*, 1660–1665. https://doi.org/10.1109/IWCMC58020.2023.10183348
17. Kawatra, M., Agarwal, S., & Kapur, R. (2020). Leaf Disease Detection using Neural Network Hybrid Models. *2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA)*, 225–230. https://doi.org/10.1109/ICCCA49541.2020.9250885
18. Khanam, R., & Mehta, G. (2023). Application of Deep CNN for image-based identification and classification of plant diseases. *2023 10th International Conference on Signal Processing and Integrated Networks (SPIN)*, 224–229. https://doi.org/10.1109/SPIN57001.2023.10116967
19. Kumar, H., Velu, S., Lokesh, A., Suman, K., & Chebrolu, S. (2023). Cassava Leaf Disease Detection Using Ensembling of EfficientNet, SEResNeXt, ViT, DeIT and MobileNetV3 Models. In R. P. Yadav, S. J. Nanda, P. S. Rana, & M.-H. Lim (Eds.), *Proceedings of the International Conference on Paradigms of Computing, Communication and Data Sciences* (pp. 183–193). Springer Nature Singapore. https://doi.org/10.1007/978-981-19-8742-7_15
20. Kumar, P., Pailwan, O., Bhausaheb, J.M., Bhaskare, P.P., & Students, B.E. (2022). Leaf Disease Detection Using Image Processing. *International Journal of Advanced Research in Science, Communication and Technology*.
21. Li, H., Shi, L., Fang, S., & Yin, F. (2023). Real-Time Detection of Apple Leaf Diseases in Natural Scenes Based on YOLOv5. *Agriculture*, *13*(4), 878. https://doi.org/10.3390/agriculture13040878
22. Li, X., & Li, S. (2022). Transformer Help CNN See Better: A Lightweight Hybrid Apple Disease Identification Model Based on Transformers. *Agriculture*, *12*(6), 884. https://doi.org/10.3390/agriculture12060884
23. Oh-Jin, Kwon., Md., Jalil, Piran. (2023). A Comprehensive Survey of Transformers for Computer Vision. Drones, doi: 10.3390/drones7050287
24. Pal, A., & Kumar, V. (2023). AgriDet: Plant Leaf Disease severity classification using agriculture detection framework. *Engineering Applications of Artificial Intelligence*, *119*, 105754. https://doi.org/10.1016/j.engappai.2022.105754
25. Pankaj Kumar, Om Pailwan, Jadhav Milan Bhausaheb, & Prof. Pallavi Bhaskare. (2022). Leaf Disease Detection Using Image Processing. *International Journal of Advanced Research in*

Science, Communication and Technology, 667–673. https://doi.org/10.48175/IJARSCT-3993

26. Prabavathy, K., Bharath, M., Sanjayratnam, K., Reddy, N. S. S. R., & Reddy, M. S. (2023). Plant Leaf Disease Detection using Machine Learning. *2023 2nd International Conference on Applied Artificial Intelligence and Computing (ICAAIC)*, 378–382. https://doi.org/10.1109/ICAAIC56838.2023.10140367
27. Qiwu, Wu., Yanli, Chu. (2023). Transformer for Hyperspectral Image Classification based on Multi-feature Learning. doi: 10.1145/3591156.3591164
28. Sander, Jyhne., Per-Arne, Andersen., Morten, Goodwin, Olsen. (2023). Contrastive Transformer: Contrastive Learning Scheme with Transformer innate Patches. arXiv.org, doi: 10.48550/arXiv.2303.14806
29. Sarkar, C., Gupta, D., Gupta, U., & Hazarika, B. B. (2023). Leaf disease detection using machine learning and deep learning: Review and challenges. *Applied Soft Computing*, *145*, 110534. https://doi.org/10.1016/j.asoc.2023.110534
30. Shukla, V., Rani, S., & Mohapatra, R. K. (2023). A New Approach for Leaf Disease Detection using Multilayered Convolutional Neural Network. *2023 3rd International Conference on Artificial Intelligence and Signal Processing (AISP)*, 1–5. https://doi.org/10.1109/AISP57993.2023.10134843
31. Siqi, Liu., Jiangshu, Wei., Gang, Liu. (2023). Image classification model based on large kernel attention mechanism and relative position self-attention mechanism. PeerJ, doi: 10.7717/peerj-cs.1344
32. Smith, J., Davis, M., & Robinson, P. (2022). Vision Transformer Application for Small-Scale Leaf Disease Detection. Plant Pathology Research, 26(1), 78-92.
33. Souleyman, Chaib., Dounia, Mansouri., Ibrahim, Omara., Ahmed, M, Hagag., Sahraoui, Dhelim., Djamel, Amar, Bensaber. (2022). On the Co-Selection of Vision Transformer Features and Images for Very High-Resolution Image Scene Classification. Remote sensing, doi: 10.3390/rs14225817
34. Thai, H.-T., Le, K.-H., & Nguyen, N. L.-T. (2023). FormerLeaf: An efficient vision transformer for Cassava Leaf Disease detection. *Computers and Electronics in Agriculture*, *204*, 107518. https://doi.org/10.1016/j.compag.2022.107518
35. Yadav, D., Gupta, A., Jain, A., & Yadav, A. K. (2022). Plant leaf disease detection using CNN with transfer learning and XGBoost. *International Journal of Data Analysis Techniques and Strategies*, *14*(3), 244. https://doi.org/10.1504/IJDATS.2022.128273
36. Yu, S., Xie, L., & Huang, Q. (2023). Inception convolutional vision transformers for plant disease identification. *Internet of Things*, *21*, 100650. https://doi.org/10.1016/j.iot.2022.100650
37. Zilin, Wu. (2023). Dual-branch Feature Fusion Vision Transformer for Garbage Image Classification. doi: 10.1109/ICCCS57501.2023.10150982

SALINITY-RESILIENT CROP HEALTH MONITORING: AUTOMATED DISEASE DETECTION IN LUFFA AEGYPTIACA LEAVES USING VISION TRANSFORMER & CNN

ORIGINALITY REPORT

10%	9%	3%	5%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	dspace.daffodilvarsity.edu.bd:8080 Internet Source	3%
2	Submitted to Daffodil International University Student Paper	3%
3	www.researchgate.net Internet Source	1%
4	Submitted to CSU Northridge Student Paper	<1%
5	Kethsy Prabavathy, Mokara Bharath, Kambam Sanjayratnam, N. Sri Sai Rahul Reddy, M. Sasidhar Reddy. "Plant Leaf Disease Detection using Machine Learning", 2023 2nd International Conference on Applied Artificial Intelligence and Computing (ICAAIC), 2023 Publication	<1%
6	Rayene Amina Boukabouya, Abdelouahab Moussaoui, Mohamed Berrimi. "Vision Transformer Based Models for Plant Disease	<1%