GUAVA LEAF DISEASE DETECTION: A POWERFUL APPROACH WITH VISION TRANSFORMERS

BY

MD. ANAM RAYAD DOHA
ID: 201-15-3521

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

Supervised By

MD. SAZZADUR AHAMED
Assistant Professor
Department of CSE
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

JULY 2024

APPROVAL

This Project titled **"GUAVA LEAF DISEASE DETECTION: A POWERFUL APPROACH WITH VISION TRANSFORMERS"**, submitted by **Md. Anam Rayad Doha** 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 is held on **13 July 2024**.

BOARD OF EXAMINERS

Dr. Sheak Rashed Haider Noori (SRH) — **Board Chairman**
Professor & Head
Department of CSE
Faculty of Science & Information Technology
Daffodil International University

Dr. Md. Zahid Hasan (ZH) — **Internal Examiner 1**
Associate Professor
Department of CSE
Faculty of Science & Information Technology
Daffodil International University

Mr. Md Mohammad Masum Bakaul (MB) — **Internal Examiner 2**
Sr. Lecturer
Department of CSE
Faculty of Science & Information Technology
Daffodil International University

Dr. Md. Zulfiker Mahmud (ZM) — **External Examiner**
Professor
Department of CSE
Jagannath University

DECLARATION

I hereby declare that this project has been done by me under the supervision of **Md. Sazzadur Ahamed, Assistant Professor, 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:

Md. Sazzadur Ahamed
Assistant Professor
Department of CSE
Daffodil International University

Submitted by:

Md. Anam Rayad Doha
ID: 201-15-3521
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

Firstly, I would want to express my utmost gratitude and appreciation to the holy power for bestowing upon me His blessings, which enabled me to complete my final year project successfully.

I express my gratitude and acknowledge our deep gratitude to **Md. Sazzadur Ahamed, Assistant Professor** at the Department of CSE, Daffodil International University, Dhaka. The supervisor's profound expertise and strong passion for the topic of "Computer Vision" will be crucial in conducting this project. The project was made feasible by his unwavering patience, scholarly leadership, consistent encouragement, energetic supervision, constructive criticism, helpful counsel, thorough review of numerous subpar versions, and diligent corrections at every level.

I would like to extend my utmost appreciation to **Dr. Sheak Rashed Haider Noori, the Head of the Department of CSE**, for his generous assistance in completing our project. I would also like to thank the other faculty members and staff of the Department of CSE at Daffodil International University.

I would like to express my gratitude to all my fellow classmates at Daffodil International University who actively participated in this debate during the completion of our coursework.

Lastly, I would want to express my sincere gratitude for the unwavering assistance and forbearance of my parents.

ABSTRACT

Guava is one of the most economically important fruits in tropical and subtropical regions like Bangladesh, India, Pakistan and many more. However, its production is significantly impacted by various leaf diseases. In recent years, computer vision techniques have emerged as effective tools for automating disease detection in plants. In this study, we propose a novel approach for guava leaf disease classification using Vision Transformer (ViT), a state-of-the-art deep learning architecture known for its ability to capture long-range dependencies in images. Our dataset comprises images of guava leaves categorized into five classes: Healthy, Canker, Dot, Rust, and Mummification, totaling approximately 4749 images. We first preprocessed the images and then trained a Vision Transformer model on the dataset. The experimental results demonstrate the effectiveness of our approach, with a training accuracy of 97.42%, validation accuracy of 95.60%, and testing accuracy of 97.54%. Furthermore, we conducted comparative experiments with traditional transfer learning models, including ResNet50 model (Training accuracy – 98.76%, Testing accuracy – 95.00%, validation accuracy – 96.10%), VGG19 model (Training accuracy - 90.39%, Testing accuracy - 89%, validation accuracy - 88.77%), EfficientNetB5 model (Training accuracy - 96.49%, Testing accuracy - 94.97%, validation accuracy - 93.33%). While these models have shown promising performance in various computer vision tasks, our results indicate that the Vision Transformer outperforms them in terms of both accuracy and computational efficiency for guava leaf disease classification. The success of our proposed approach underscores the potential of Vision Transformers in agricultural applications, particularly in the early detection and management of plant diseases. Our findings contribute to the ongoing efforts to leverage artificial intelligence for sustainable agriculture, offering a powerful tool for farmers and researchers to combat guava leaf diseases and ensure global food security.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Introduction

Guava, a tropical fruit, is both easily cultivable and highly nutritious. In Bangladesh, it is alternatively referred to as peyara/pearah or goyaa. The plant is extensively grown in Southeast Asia, Latin America, and other areas characterized by warm weather. The fruit is highly valued for its delectable taste, abundant vitamins, and antioxidants, making a substantial contribution to worldwide food security and nutrition. However, the production of guava encounters a significant obstacle in the shape of several leaf diseases. These illnesses can result in substantial reductions in crop productivity, diminish the quality of the fruit, and eventually affect the livelihoods of farmers who depend on guava agriculture.

Step into the domain of artificial intelligence and computer vision, where creativity and practicality converge in unforeseen ways. In this day of advanced technology, where each pixel can achieve great things and every piece of computer programming is like a stroke of a paintbrush on the canvas of advancement, we are embarking on a voyage driven by the power of Vision Transformers. These transformers provide farmers with a promising solution to the challenges posed by guava leaf diseases, because of their exceptional capacity to analyse images with a high level of precision.

The categorization of guava leaf diseases represents a crucial area in agricultural research. Protectors of the guava orchards, farmers need to recognise the telltale signs of Canker, Dot, Rust, and Mummification, each of which poses a different risk to the well-being and productivity of the plants. However, even with years of practice, the human eye can become unreliable in the finer details of a disease's course.

For efficient disease management and long-term crop production, guava leaf diseases must be identified and categorized as soon as possible. Early detection of diseases enables farmers to promptly execute specific control strategies, such as targeted pruning, the use of fungicides or bactericides, and the adoption of cultural practices, to avoid the spread of the disease. To effectively reduce the spread of diseases and minimize yield losses while retaining crop quality, it is crucial to rapidly isolate sick plants and adopt suitable control techniques. Moreover, timely identification allows researchers and agricultural extension services to observe the occurrence and spread of diseases, hence aiding in the creation of customized therapies and disease prediction models. Timely categorization of diseases is equally vital as it offers valuable understanding of disease patterns, enabling the detection of developing risks and the choice of resilient cultivars. Furthermore, precise categorization aids in monitoring the effectiveness of control measures and evaluating the lasting effects of disease management policies on guava production systems. In general, the early identification and categorization of diseases form the basis for taking proactive measures to manage them. This, in turn, helps to improve the ability of crops to withstand diseases, promotes sustainable agricultural methods, and ensures food security.

The Vision Transformer (ViT) is an innovative model that offers a new method for picture classification tasks, such as identifying guava leaf illnesses. ViT diverges from conventional Convolutional Neural Networks (CNNs) by employing transformer architecture, initially devised for natural language processing tasks. ViT employs a transformer encoder to process fixed-size patches of the input image, thereby capturing spatial relationships among the patches, rather than depending on convolutional layers. ViT is able to effectively capture both local and global aspects of the input image, enabling rapid representation learning. The paper's methodology involves adapting and fine-tuning the ViT model using a dataset that consists of guava leaf photos that have been annotated with disease diagnoses. The dataset is partitioned into training, validation, and testing sets, with suitable data augmentation techniques implemented to improve the model's generalization. The model is trained using a supervised learning strategy, where classification performance is improved through the use of backpropagation and gradient descent optimization. Hyperparameters, such as the learning rate, batch size, and optimizer, are meticulously adjusted to attain the best possible performance. The trained Vision Transformer (ViT) model is subsequently assessed on the validation and testing datasets to determine its accuracy, precision, recall, and F1 score for each illness category. Furthermore, the efficacy of the ViT model in detecting guava leaf disease is assessed by comparing its performance with other transfer learning models, namely VGG19, ResNet50, and EfficientNetB5. The research demonstrates, via thorough experimentation and evaluation, that ViT is highly effective tool for precise and quick classification of guava leaf diseases. This highlights its potential to enhance disease management strategies in agriculture.

1.2 Motivation

As the demand for guava continues to rise, guava farmers face a significant challenge in mitigating the impact of diseases that threaten their orchards' health and productivity. Traditional methods of disease detection often prove laborious, time-intensive, and reliant on expert knowledge, leading to delayed intervention and substantial crop losses. In response to this critical issue, this study presents a pioneering approach harnessing the power of deep learning for the early detection of guava leaf diseases, particularly through the analysis of leaf imagery. By embracing cutting-edge technology, this method aims to streamline disease detection processes, offering enhanced accuracy and efficiency compared to conventional techniques.

At the heart of this proposed solution lies the Vision Transformer model, representing a revolutionary advancement in image analysis and pattern recognition. Vision Transformers have demonstrated exceptional capabilities in capturing intricate features and subtle nuances within visual data, making them ideally suited for the complexities of guava leaf disease detection. The adoption of Vision Transformers in this context heralds a transformative shift in disease management practices for guava farmers, empowering them with timely and precise insights into the health of their orchards.

Supplemented by complementary models such as VGG19, ResNet50, and EfficientNetB5, this approach amalgamates the strengths of diverse deep learning architectures to further augment detection accuracy and robustness. VGG19's deep architecture enables the extraction of hierarchical features, while ResNet50's residual connections enhance model convergence and performance. EfficientNetB5, renowned for its scalability and computational efficiency, ensures optimal resource utilization in guava leaf disease detection tasks. By leveraging these advanced deep learning models, this research endeavors to revolutionize disease management practices in guava farming, providing farmers with invaluable tools to preemptively identify and address diseases, thereby minimizing crop losses and promoting sustainable agricultural practices. Ultimately, this innovative approach promises to empower guava farmers.

with the knowledge and resources needed to safeguard their livelihoods and contribute to the continued growth and prosperity of the guava industry.

1.3 Rational of Study

The rationale behind this study on guava leaf disease detection stems from the imperative need to address the pressing challenges faced by guava farmers worldwide. Guava, being a popular fruit crop with increasing demand, is susceptible to various diseases that pose significant threats to orchard health and productivity. Traditional methods of disease detection often fall short in terms of accuracy, timeliness, and scalability, leaving farmers vulnerable to substantial crop losses. Recognizing the limitations of conventional approaches, this study seeks to introduce a novel and powerful solution leveraging the capabilities of deep learning, specifically focusing on the utilization of advanced models such as Vision Transformers, VGG19, ResNet50, and EfficientNetB5. The primary motivation for employing deep learning techniques lies in their unparalleled ability to extract intricate patterns and features from complex datasets, particularly in the domain of image analysis.

The selection of Vision Transformers as the main model for this study is driven by several factors. Vision Transformers have recently emerged as a breakthrough in computer vision, offering remarkable performance in tasks such as image classification and object detection. Their unique architecture, which incorporates self-attention mechanisms to capture global dependencies within images, makes them well-suited for the nuanced and multifaceted nature of guava leaf diseases. By focusing on Vision Transformers as the primary model, this study aims to explore their potential to revolutionize disease detection in guava orchards, paving the way for more efficient and accurate diagnostic processes.

Furthermore, the inclusion of supplementary models such as VGG19, ResNet50, and EfficientNetB5 serves to enrich the analytical capabilities of the study. Each of these models brings its distinct advantages, ranging from deep hierarchical feature extraction to enhanced computational efficiency. By leveraging a diverse ensemble of deep learning architectures, this study seeks to harness the collective strengths of these models to achieve optimal performance in guava leaf disease detection.

1.4 Research Question

1. How does the performance of Vision Transformer compare to traditional deep learning models such as VGG19, ResNet50, and EfficientNetB5 in detecting guava leaf diseases?
2. What is the impact of dataset size and diversity on the accuracy and generalization capability of the Vision Transformer model for guava leaf disease detection?
3. Can transfer learning techniques be effectively employed to fine-tune pre-trained Vision Transformer models for improved detection of specific guava leaf diseases?
4. How does the computational efficiency of Vision Transformer compare to other deep learning models in the context of guava leaf disease detection, particularly in resource-constrained environments?
5. What are the key factors influencing the interpretability and explain ability of predictions generated by the Vision Transformer model in guava leaf disease detection?

1.5 Expected Output

The purpose of this research study is to investigate the potential of utilizing Vision Transformers and other deep learning models, including VGG19, ResNet50, and EfficientNetB5, in conjunction with transfer learning techniques, to enhance the precision of guava leaf disease detection using leaf imagery. The primary objective is to identify distinctive features and patterns within guava leaf images that serve as reliable indicators of various diseases, thus facilitating the development of a robust classification model.

By employing transfer learning, this study aims to leverage pre-trained models' knowledge and adapt it to the specific task of guava leaf disease detection, thereby potentially improving the accuracy and efficacy of disease identification. The research will evaluate the effectiveness of this transfer learning methodology, considering metrics such as accuracy, precision, recall, and F1-score, and will conduct a comparative analysis with other available methods, including traditional machine learning approaches.

Furthermore, this study seeks to explore prospective avenues for future research in utilizing Vision Transformers and transfer learning for guava leaf disease detection. It will propose innovative methods to enhance and broaden existing approaches, including the integration of multi-modal data sources such as environmental factors and geographic information, to further improve disease diagnosis accuracy and robustness.

The findings of this research could have significant implications for guava farmers and agricultural practitioners by providing them with advanced tools and techniques for early and accurate disease detection, ultimately contributing to the preservation of guava orchard health and productivity. This study endeavors to advance the field of agricultural diagnostics and may pave the way for the development of scalable and efficient disease management strategies in the context of guava farming.

1.6 Project Management and Finance

In order to successfully design and implement the proposed Vision Transformer (ViT) model for guava leaf disease classification, it is crucial to carefully examine both project management and budgetary resources. The implementation of effective project management will be achieved by following a well-defined plan that includes specific milestones, strategies for allocating resources, and stated dates. This will guarantee a methodical and effective workflow across all the stages of research, starting from the gathering of data and extending to the construction and evaluation of models.

There aren't really any major costs involved in building this guava leaf disease detection system. We spent some money on a good camera and travel to take pictures of the leaves ourselves, but there were also free collections of data available online that we used instead. The real cost comes from using a powerful computer to train the models. We potentially use our own computer and hope it's powerful enough, but if not, we might need to rent processing power from a special online service. There are also free tools available to help us build and test the system. So overall, this project should be inexpensive. To keep costs down, we'll try to find free image collections and use our own computer whenever possible for future development.

Here Figure 1.1 shows that the research lasts one year. In this Gantt Chart, for three months, I collected the data. Then, I worked on the literature review part for two months. Data pre-processing worked for one month. After that, four months of work were spent on model implementation. Then the one-month analysis of the result. After completing those steps, I worked on report writing.

Figure 1.1: Gantt Chart of My Research

1.7 Report Layout

This thesis is intricately organized to systematically lead the reader through the study process, starting from defining the problem and ending with exploring the potential effects and future avenues. Every chapter has a specific function, which helps to create a clear and thorough knowledge of the suggested ViT-based approach for classifying guava leaf diseases.

Chapter 1: Introduction

The opening chapter establishes the context by emphasizing the economic significance of guava and the adverse effects of leaf diseases on its cultivation. Next, the text presents the idea of utilizing computer vision for disease diagnosis and provides a rationale for employing a Vision Transformer (ViT) model for this purpose. This chapter will explicitly establish the research goals and provide a concise overview of the overall structure of the study.

Chapter 2: Background and Relevant Studies

It explores the important foundational information. This encompasses the investigation of many ailments affecting guava leaves, their observable traits, and the current techniques employed for identifying these diseases. Moreover, a thorough examination of deep learning methodologies, with specific emphasis on Vision Transformers, will be provided. This chapter provides the theoretical basis for the proposed technique and highlights the constraints of current methodologies, thereby supporting the necessity of this research.

Chapter 3: Methodology of research

Here provides a comprehensive explanation of the specific methodological technique utilized in this investigation. The text will provide an explanation of the procedure of collecting data, which involves obtaining guava leaf photos from a specific source and categorizing them into several disease groups. The subsequent chapter will explore the intricate aspects of the ViT model's architecture, encompassing its training procedure, strategies for adjusting hyperparameters, and the evaluation metrics employed to gauge its performance. Furthermore, this document will provide an explanation of any data pre-processing strategies used to improve the performance of the model.

Chapter 4: Experimental Outcome

In this part showcases the empirical findings acquired by training and assessing the ViT model on the job of classifying guava leaf diseases. This will require thorough examination of the model's performance indicators, including accuracy, precision, recall, and F1-score. The chapter will additionally investigate the performance comparison between the proposed ViT technique and classic transfer learning models such as ResNet50, VGG19, and EfficientNetB5. It will emphasize the advantages of the ViT approach. Additional visualizations of the classification outcomes can be incorporated to offer deeper understanding.

Chapter 5: Analysis and Societal Influence

It explores the wider consequences and future prospects of the research. The text will analyze the importance of the findings for guava farming and their possible influence on enhancing crop output and food security. This chapter will additionally examine the constraints of the present study and pinpoint opportunities for future investigation. Possible areas for enhancement could involve examining the influence of various image pre-processing methods or evaluating the incorporation of ViT with other deep learning frameworks. Furthermore, the practicality of using the model in real-life agricultural environments will be examined.

Chapter 6: Conclusion, Summary, Suggestions, and Research Implications

The last chapter functions as the apex of the entire thesis. The purpose of this summary is to provide a concise overview of the main findings from the research and to reaffirm the conclusions made on the efficacy of the ViT model in classifying guava leaf diseases. This chapter will also highlight the possible impact of this research on the field of sustainable agriculture and end by restating the significance of the proposed approach in guaranteeing worldwide food security.

The organized arrangement of information in this layout enables the reader to easily comprehend the study method, its conclusions, and its possible impact on guava leaf disease detection and agricultural disease management.

CHAPTER 2

BACKGROUND STUDY

2.1 Preliminaries

This research explores the utilization of Vision Transformers (ViTs), an advanced deep learning framework, for the automatic categorization of guava leaf diseases. Visual Transformers (ViTs) are highly proficient in identifying intricate relationships between distant elements in images, which makes them exceptionally well-suited for tasks such as illness diagnosis. The analysis employs a dataset consisting of over 4749 guava leaf photos classified into five separate categories: Healthy, Canker, Dot, Rust, and Mummification. In order to achieve the best possible performance of the model, we employ image pre-processing techniques. Our approach primarily involves training a Vision Transformer (ViT) model using the provided dataset. This enables the model to utilize the natural visual features of healthy and sick guava leaves, ultimately resulting in higher classification accuracy. This research investigates the capabilities of ViTs, which could lead to the creation of reliable and effective instruments for detecting guava diseases. This contribution is highly valuable to the field of precision agriculture.

2.2 Related Works

Deep learning, a subset of artificial intelligence, has fundamentally transformed the domain of computer vision, namely in areas such as image categorization and identification of objects. This robust methodology has been effectively implemented in diverse fields, such as medical diagnosis, autonomous vehicles, and remote sensing. Deep learning is being actively investigated in the field of agriculture for the purpose of identifying and categorizing plant diseases.

This section of the thesis will explore the pertinent literature about the utilization of deep learning methods, specifically convolutional neural networks (CNNs) and Vision Transformers (ViTs), for the classification of plant diseases. We will examine previous studies on the utilization of deep learning models for the detection of diverse agricultural diseases, emphasizing the possibilities and constraints of these methodologies. This paper aims to build a comprehensive grasp of the latest advancements in disease identification using deep learning techniques. It will also set the stage for introducing the suggested ViT model, which is designed for classifying guava leaf diseases.

The paper by Assad Doutoum, Recep Eryigit and Bulent Tugrul mentioned classification of Guava Leaf Disease using Deep Learning involves training Convolutional Neural Network (CNN) models to detect and classify various guava diseases. Studies have shown that deep learning algorithms, such as ResNet-101, Bagged Tree, and SVM, have achieved high accuracy rates in identifying diseases like Canker, Mummification, Rust, and Dot in guava plants [1]. By utilizing deep learning methods like EfficientNet-B3, researchers have achieved a 94.93% accuracy in classifying guava leaf diseases, outperforming traditional methods and providing a more reliable approach for early disease detection and prevention in guava trees. This automated system aids farmers in diagnosing diseases promptly, implementing effective control strategies, and ultimately improving guava production by providing valuable insights into guava leaf diseases.

GLD-Det is an advanced framework that uses a lightweight deep learning algorithm based on MobileNet to detect guava leaf disease in real-time. The framework demonstrates exceptional performance in terms of speed, reliability, accuracy, efficiency, and cost-effectiveness [2]. The model includes several components such as pooling layers, dropout layers, batch normalization layers, dense layers with ReLU activation, and a smaller dense layer with SoftMax activation to improve efficiency and mitigate overfitting. The model's efficacy has been verified in diverse geographical regions such as Pakistan and Bangladesh, guaranteeing its strength and precision using preprocessing approaches and assessment metrics. The integration of the Grad-CAM technique into cellphones enhances confidence and transparency in guava leaf disease detection, making it an optimal and enduring solution.

Another academic study introduces a hybrid deep learning model designed to identify numerous diseases in real-time on a single guava leaf. The suggested approach has three primary elements: the detection of contaminated guava patches using GIP-MU-NET, the segmentation of guava leaves using the Guava Leaf Segmentation Model (GLSM), and the identification of numerous illnesses on a guava leaf using the GMLDD model. This method facilitates the effective and precise identification and pinpointing of illnesses on guava leaves, hence improving disease control measures in agriculture [3].

By integrating these advanced deep learning techniques, the model can proficiently examine guava leaf photos to detect different diseases, offering a comprehensive solution for real-time monitoring and management of guava plant health.

The study paper, entitled "A Hybrid Framework for Detection and Analysis of Leaf Blight Using Guava Leaves Imaging," introduces a new method devised by Mumtaz et al. to identify and analyze leaf blight on guava leaves [4]. This work exemplifies a collaborative endeavor among academics hailing from Pakistan, Nigeria, Saudi Arabia, and the UK, demonstrating a multidisciplinary approach to addressing agricultural difficulties. The framework combines computer science, engineering, and industrial engineering to successfully tackle the challenge. The framework seeks to improve the precision and effectiveness of leaf blight detection by integrating knowledge from several disciplines. This offers a hopeful resolution for farmers and agricultural professionals. This novel method highlights the significance of interdisciplinary collaboration in the development of cutting-edge agricultural technologies.

The study paper, titled "A Deep Features Extraction Model Based on the Transfer Learning Model and Vision Transformer," introduces a new method for classifying plant diseases. This method utilizes a hybrid model called TLMViT. This algorithm utilizes transfer learning and a vision transformer to extract profound characteristics from plant leaves, resulting in precise disease classification [5]. The TLMViT framework comprises four stages: data collecting, image augmentation, initial feature extraction utilizing pre-trained models, and deep feature extraction with ViT, culminating in classification using an MLP classifier. The study emphasizes the efficacy of utilizing ViT for extracting comprehensive and relevant information from plant leaves in order to enhance the accuracy of disease categorization.

Visual intelligence is essential in precision agriculture, specifically in identifying plant diseases utilizing efficient vision transformers [6]. The field of agriculture encounters major obstacles as a result of plant infections, which in turn result in impaired food security and large crop losses. In order to tackle these problems, the incorporation of advanced deep learning methods, such as vision transformers, has demonstrated potential in automating the process of disease identification through the study of images. Machine learning techniques, such as Support Vector Machine (SVM) in conjunction with feature extraction methods like Scale Invariant Feature Transform

(SIFT), have been employed to categorize plant diseases using picture data. Researchers seek to optimize disease detection processes in agriculture by utilizing advanced technologies such as vision transformers. This approach aims to improve the efficiency and accuracy of disease identification, leading to enhanced crop management and increased food production.

Image classification of Guava leaves for disease diagnosis is crucial for maintaining plant health and fruit yield. In the study by M.Thilagavathi S.Abirami, image processing techniques like region growing segmentation, color transformation (YCbCr, CIELAB), and Scale Invariant Feature Transform (SIFT) were applied to resized leaf images with improved contrasts to classify diseases accurately. Support Vector Machine (SVM) and k-Nearest Neighbor (k-NN) classifiers were evaluated for disease-wise classification accuracy using 125 leaf samples with 128 texture features per sample. While both SVM and k-NN performed well, SVM showed slightly superior accuracy, highlighting its potential for precise disease classification in Guava leaves [7]. This research demonstrates the effectiveness of utilizing image processing and machine learning algorithms for automated and accurate disease diagnosis in Guava plants, aiding in timely and targeted treatment strategies.

PlantKViT is a novel model that combines the Vision Transformer architecture with the KNN algorithm for forest plant classification, offering high efficiency and ease in adding new plant species [8]. By utilizing the Vision Transformer as an embedding network to capture feature vectors of plant images and storing them in a database, the model simplifies the identification process by comparing these vectors using KNN. Tested on the DanangForestPlant dataset, PlantKViT achieved an impressive accuracy of 93%, outperforming other models like ConvNeXt and Resnet-152 significantly. The successful development of applications like 'plant id' and 'Danangplant' for iOS and Android platforms showcases the model's potential for widespread forest plant identification, emphasizing the need for future work to enhance accuracy and performance by expanding the dataset and refining the model further.

The research paper proposes an automatic pest identification method utilizing the Vision Transformer (ViT) to enhance the accuracy of computer classification of plant diseases and insect pests in precision agriculture and smart agriculture applications [9]. By augmenting plant diseases and insect pests datasets with techniques like Histogram

Equalization, Laplacian, Gamma Transformation, CLAHE, Retinex-SSR, and Retinex-MSR, the ViT neural network achieved a test recognition accuracy rate of 96.71% on the Plant_Village dataset, surpassing traditional convolutional neural networks like GoogleNet and EfficientNetV2 by approximately 1.00% in identifying plant diseases and insect pests. This innovative approach showcases the potential of Vision Transformers in automating the identification and detection of plant pests and diseases, contributing to more precise classification and improved system performance in agricultural settings.

The categorization of olive diseases can be efficiently accomplished by employing a hybrid feature extraction technique that integrates convolutional neural networks (CNN) and vision transformers (ViT) [10]. ViT models, which draw inspiration from transformer models in natural language processing, have exceptional proficiency in comprehending global interdependencies within input images. This is achieved by dividing the images into subimages and employing self-attention modules. ViTs, unlike CNNs, preserve position information throughout the processing stages, which makes them well-suited for picture classification tasks. The ViT architecture, equipped with a self-awareness mechanism, presents a potential method for precise categorization of olive diseases. It combines the advantages of both CNNs and transformers in extracting visual features and performing classification. This novel approach demonstrates the capabilities of ViT models in improving picture classification tasks, such as the detection of olive diseases.

The proposed method for real-time automated plant disease categorization involves using the Vision Transformer (ViT), which is a lightweight deep learning approach [11]. The objective of this technology is to offer farmers visual information in order to take prompt preventive actions against plant diseases. The ViT architecture emulates the human process of picture categorization by directing attention to specific regions of interest, hence improving the accuracy of classification. Although pre-designed architectures such as ResNet and ViT use extensive datasets to achieve optimal parameter values, lightweight hand-designed architectures can deliver similar performance without the computational load, which is essential for real-time applications. By integrating attention blocks with CNN blocks, it is possible to reduce the trade-off between speed and accuracy, hence enhancing the efficiency of the classification process. The study contrasts ViT-based models with CNN-based architectures, emphasizing the significance of model intricacy and prediction velocity in crop disease categorization.

The authors of the research article "Evaluating Deep CNNs and Vision Transformers for Plant Leaf Disease Classification" did a comprehensive study to compare Convolutional Neural Networks (CNNs) and Vision Transformers (ViT) in their ability to classify plant leaf diseases. The study concentrated on datasets pertaining to Rice leaf, Tea leaf, and Maize leaf, assessing the efficacy of both CNN and ViT models on these datasets. The objective of the research was to gain useful insights into the most effective setup of ViT for achieving exceptional performance in the categorization of plant leaf diseases. The research also emphasized the difficulties connected with this task [12]. Comparing Convolutional Neural Networks (CNNs) and Vision Transformers (ViTs) in this specific context provides a thorough evaluation of their skills, which helps in the progress of artificial intelligence (AI) applications in smart agriculture for the effective diagnosis and control of diseases.

2.3 Comparative Analysis

An essential component of this study entails the comparative review of various deep learning models and papers for the categorization of guava leaf diseases. This research provides insight into the advantages and disadvantages of several approaches, enabling us to determine the most efficient model for this work. The work rigorously assesses the performance of a Vision Transformer (ViT) model compared to well-established transfer learning models such as ResNet50, VGG19, and EfficientNetB5. We evaluate the accuracy, precision, recall, and other important parameters for each model by subjecting them to rigorous training, validation, and testing procedures.

This comparative inspection extends beyond the mere identification of the most precise model. It offers significant perspectives on the appropriateness of various architectures for classifying guava leaf disease. Through an analysis of the performance differences among models, we may enhance our comprehension of the particular characteristics that each model excels at capturing in the leaf images. Furthermore, the evaluation takes into account the computing efficiency of each model, as the constraints of available resources can impact the practical use of these models in agricultural applications.

The results obtained from this comparative analysis will provide valuable insights for the discussion in the next chapters. By emphasizing the efficacy of the ViT model and its superiority over conventional transfer learning methods, we bolster the argument for its potential use in agricultural disease management. This investigation not only contributes to the progress of deep learning in the identification of plant diseases, but also offers practical recommendations for choosing suitable models in different agricultural settings. In conclusion, this comparative analysis establishes the foundation for future research that investigates the wider implementation of ViT and other deep learning models in sustainable farming practices.

Table 2.1: Analysis of comparison using previous research

Author	**Image type**	**Model**	**Accuracy**
Assad Doutoum Recep Eryigit Bulent Tugrul	Guava Leaf	VGG16, ResNet50, InceptionV3, EfficientNet-B3	EfficientNet-B3 achieved 94.93% accuracy on the test data.
Md Mustak Un Nobi Md Rifat M Mridha Mejdl Safran Dunren Che	Guava Leaf	EfficientNetV2 B2, EfficientNetB0, EfficientNetB2, EfficientNetB1, MobileNetV2, GLD-Det	GLD-Det model accuracy: 0.9714 on Dataset D1, 0.9712 on Dataset D2.
Javed Rashid Imran Khan Ghulam Ali Shafiq Ur Rehman Fahad Alturise Tamim Alkhalifah	Guava Leaf	YOLOv3 YOLOv4 YOLOv4tiny GMLDDmodel	Sensitivity = 90.07% and Specificity = 92.16%
Maciej Zaborowicz Jakub Frankowski Sidrah Mumtaz Mudass Raza Ofonime Okon Saeed Rehman Adham Ragab Hafiz Tayyab Rauf	Guava Leaf	SidNet, DarkNet-53, AlexNet	Achieved 98.9% accuracy with 5-fold and 99.2% with 10-fold cross validation.

Hamoud Alshammari Karim Gasmi Ibtihel Ltaifa Moez Krichen Lassaad Ammar Mahmood Mahmood	Olive Diseases	ViT+VGG-16, ViT+VGG-19, Transformer (ViT)	Achieved 96% accuracy for multiclass and 97% for binary classification.
Yasamin Borhani Javad Khoramdel Esmaeil Najafi	Infected rice leaves and wheat rust	Vision Transformer (ViT), convolutional neural network (CNN), AlexNet, SqueezeNet and GoogleNet	GoogleNet = 99%
Parag Bhuyan Kumar Pranav Singh Pranav Kuma	Plant Leaf	Resnet50, VGG16, DenseNet121, MobileNetV2, InceptionV3, ViT-10, ViT-20, ViT-30	ViT-30 achieved 98.41% accuracy on Rice Leaf Dataset. ViT-20 achieved 67.75% accuracy on Tea Leaf Dataset. ViT-30 achieved 96.95% accuracy on Maize Leaf Dataset.
P Perumal Kandasamy Sellamuthu K Vanitha V Manavalasundaram [13]	Guava Leaf	Support Vector Machine (SVM)	Overall accuracy of the proposed system: 98.17%
AhmadAlmadhor, HafizTayyabRauf, Muhammad Ikram Ullah Lali, Bader Alouffi and AbdullahAlharbi [14]	Guava Plant	Fine KNN, Cubic SVM, Complex tree, Boosted tree, and Bagged tree	Bagged tree accuracy is 99%
Lama Aldakhil Aeshah Almutairi [15]	6 types of Fruits	EfficientNetV2	EfficientNetV2 = 99.49%
My Proposed Model	Guava Leaf	ViT	ViT = 97.54%

2.4 Scope of The Problems

This study addresses the difficulty of precisely and automatically classifying diseases in guava leaves. Guava, an essential fruit crop in numerous tropical and subtropical areas, has substantial decreases in production as a result of a variety of leaf diseases. Timely and accurate disease diagnosis is essential for adopting efficient disease management methods. Conventional techniques used to diagnose guava leaf diseases mostly depend on visual examination conducted by specialists, which can be subjective, time-consuming, and potentially susceptible to inaccuracies.

This paper introduces an innovative framework that utilizes a Vision Transformer (ViT) model to overcome these limitations. Visual Transformer (ViT) are highly proficient in catching intricate relationships between distant elements in photos, which makes them exceptionally well-suited for extracting nuanced disease patterns from guava leaf images. Moreover, the study investigates the capacity of transfer learning, wherein a pre-trained ViT model utilizes knowledge acquired from other image classification tasks to enhance its accuracy in classifying diseases in the guava leaf disease dataset. This research endeavors to enhance the objectivity, efficiency, and reliability of guava leaf disease classification by employing a ViT model with transfer learning. The ultimate goal is to empower farmers and agricultural researchers in their efforts to combat these diseases and achieve optimal crop production.

CHAPTER 3

RESEARCH METHODOLOGY

3.1 Research Subject and Instrumentation

The main focus of the study is " Guava Leaf Disease Detection: A Powerful Approach with Vision Transformers." The goal is to develop and implement a novel approach that uses transformer methods—more especially, the Vision Transformer (ViT)—to diagnose diseases in guava leaf. There are some steps to complete the study. The complete procedural strategy is depicted in Figure 3.1.

1. Data Collection: The first step is to collect data about guava leaf disease from an online source, mainly Kaggle.
2. Data Preprocessing: Preprocessing will be applied to the images after data collection to guarantee a uniform and standardized dataset. Several image processing techniques were used, including sharpening, contrast correction, and noise reduction.

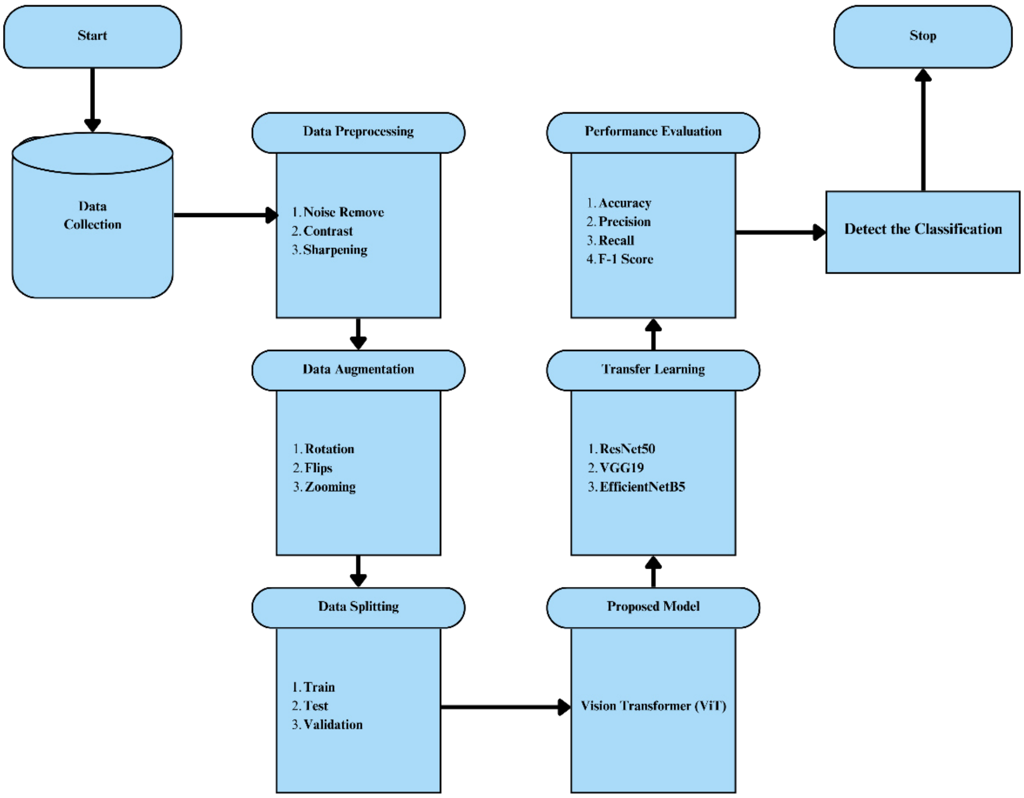

Figure 3.1: The General Approach to Categorizing Guava Leaf Disease Classes

3. Data Augmentation: After preprocessing the image, apply some data augmented techniques such as rotation, flip's and zooming.
4. Data splitting: Split the data into training, test, and validation.
5. Proposed model: Applied proposed model into the dataset.
6. Other transfer learning models: Applied some transfer learning models into the dataset.
7. Performance Evaluation: Compare the proposed model to the transfer learning model and evaluate the performance of those models.

3.2 Data Collection Procedure:

In this research, the guava leaf dataset is publicly available on Kaggle. This dataset contained five major classes: Canker, Mummification, Dot, Rust, and Healthy. This guava leaf dataset is divided into three major categories: training, validation, and testing directories. The dataset has 4749 images of guava leaves. There are 3099 images in the training set, 795 images in the validation set, and 855 images in the testing set. The resized images have a resolution of 224 × 224 pixels.

Canker

Dot

Healthy

Mummification

Rust

Figure 3.2: Sample of Guava leaf diseases

3.2.1 Image preprocessing

The process of transforming raw image data into a format that is both useful and meaningful is known as image preprocessing. It helps you improve some aspects necessary for computer vision applications and eliminate unnecessary distortions. Preprocessing is an essential initial step before putting image data into deep learning models. To reduce the noise firstly apply a Gaussian filter then apply contrast and brightness adjustment techniques to enhance the quality of the images and finally, sharping techniques are applied.

3.2.2 Removing noise using Gaussian filter:

Removing or minimizing undesirable noise or artifacts from an image is known as noise reduction. It is significant because it may enhance the image's clarity and visual quality and make it simpler to interpret or analyze using computer algorithms. In this study, I applied a Gaussian filter to reduce the noise it usually produces the smoothest image. In this process, I implemented a kernel size of 3x3.

3.2.3 Contrast and Brightness Adjustment:

An image's visual attractiveness and efficacy can be greatly impacted by adjusting its brightness and contrast. It can be useful in enhancing detailed visibility and fixing any imperfections or faults in the image. For an image to be both visually appealing and functional, the brightness and contrast levels must be carefully balanced. In this study, I used the brightness value is 1.5 and the contrast value is 1.

3.2.3 Sharping techniques:

To make images look sharper and more defined, sharpening is the process of boosting the edges and small characteristics in the image. It is significant because it is useful in highlighting the elements and details of an image, improving its visual appeal and simple comprehension. The sharping method was implemented with a kernel size of 3x3.

3.2.4 Data Augmentation:

Data augmentation is the process of creating new images from preexisting ones to expand the amount and diversity of a dataset. This is accomplished by giving the images several adjustments, where parameters were set as “width_shift_range” = 0.2, “height_shift_range” = 0.2 “rotation_range” = 0.2, “vertical_flip” = True, and “horizontal_flip” = False. To avoid overfitting and enhance the model's capacity for generalization, the training dataset's size and diversity should be increased.

3.3 Statistical Analysis

This study utilized rigorous statistical analysis to confirm the strength and applicability of our findings on the classification of guava leaf disease using a Vision Transformer (ViT) model. The performance parameters (accuracy, precision, recall) achieved by the ViT model throughout training, validation, and testing datasets were characterized using descriptive statistics, such as mean and standard deviation. This facilitated a lucid comprehension of the central tendencies and variability in the model's performance.

In addition, inferential statistics, such as hypothesis testing and confidence intervals, were used to evaluate the statistical significance of the observed variations between the ViT model and the other transfer learning models (ResNet50, VGG19, EfficientNetB5). By employing statistical rigor, we were able to assert with confidence that the improved performance of ViT is not a result of random chance.

By confirming the statistical significance of the results, we may confidently extrapolate our findings to a wider population of guava leaf photos. This enhances the overall credibility of the suggested method and showcases its potential for practical use in identifying guava diseases.

Ultimately, the statistical analysis was crucial in deriving significant insights from the experimental data. The study not only confirmed the efficacy of the ViT model but also established a basis for its wider implementation in the domain of agricultural disease identification.

3.4 Proposed Methodology

Vision Transformer (ViT): Transformers have achieved remarkable success in natural language processing tasks, largely attributed to their attention mechanisms. Building upon this concept, the Vision Transformer (ViT) has emerged as a powerful architecture for image classification. There are three fundamental parts in Vision Transformer (Vit) which are Patch Embedding, Transformer Encoder, and Classification. In ViT, it breaks down the images into small pieces called patches. After that, Flatten the patches from 2D to 1D vectors where each patch is a separate input token and also linear projection flattened patches transforming every 1D vector into a lower dimension, preserving the relationship of important features.

ViT relies heavily on positional embeddings. ViT, in contrast to traditional feed-forward networks, represents each picture patch as a sequentially indexed vector and maintains context using positional embeddings. The transformer framework receives these vectors after they have been expanded and joined with an embedding matrix.

Multiple multi-head self-attention mechanisms are included in the transformer encoder block. These mechanisms provide an entire input image representation by using its position embeddings and embedded patch sequence. In the end, classifying the input image into various categories is the responsibility of the classification head, which is usually implemented as a fully connected layer.

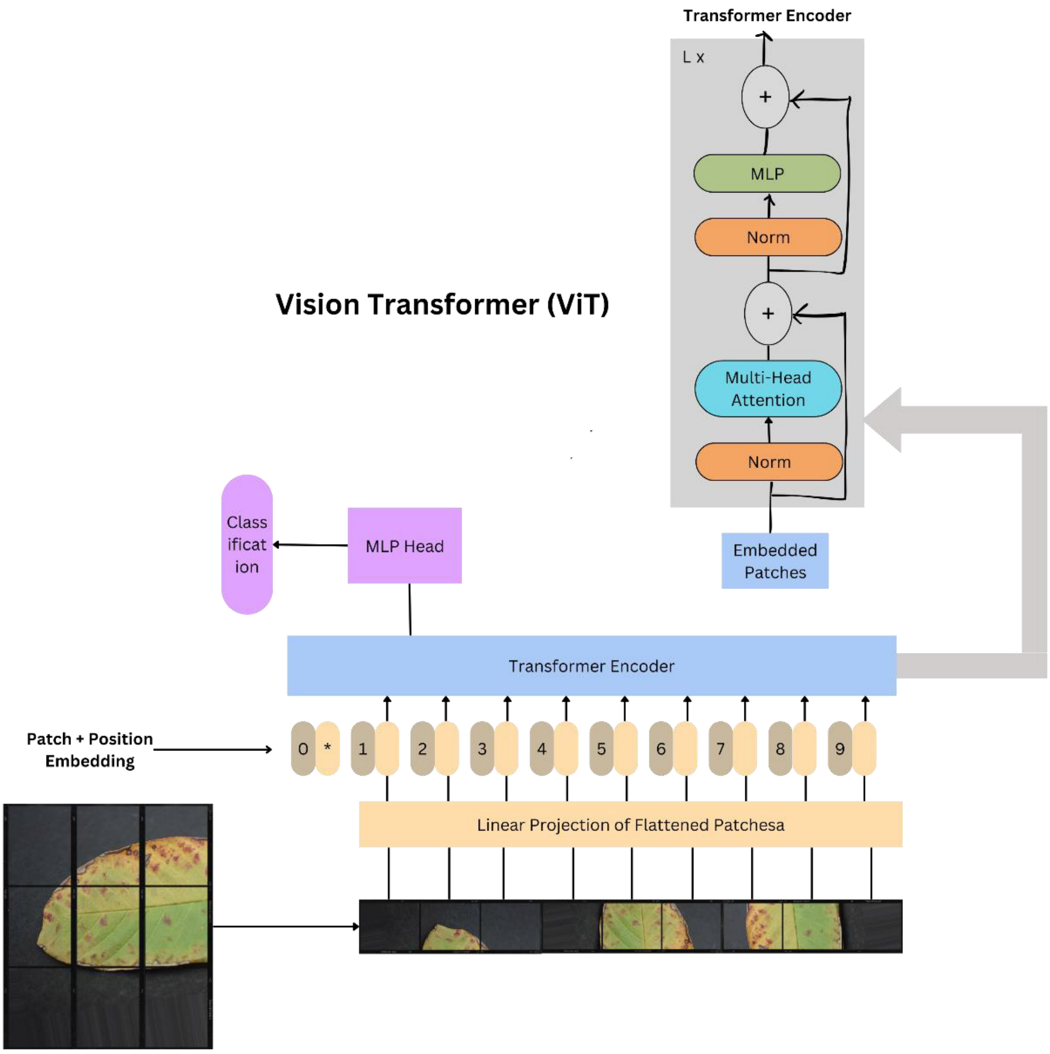

Figure 3.3: Architecture of Proposed ViT Model

In my proposed model, I used ViT-B16 where the image size is 224x224, the patch size is 16x16, the stride is 16, and applied 20 epochs. Implemented L2 regularization technique and dropout 0.5 to mitigate the issue of overfitting. Used relu for the activation function. Finally, add two fully connected layers with 512 and 256 neurons respectively which subsequently improves the performance.

3.4.1 Transfer Learning Algorithm

ResNet50: ResNet50 is a significant image classification model that produces novel results when trained on huge datasets.

A significant advancement in it is the utilization of residual connections, which enable the network to make up a set of residual functions that translate the input into the intended output. With these remaining connections, the network is able to learn much deeper structures without encountering the vanishing gradients issue. It is a 50-layer convolutional neural network where 48 convolutional layers, one MaxPool layer, and one average pool layer. It is divided into four main parts which are convolutional layers, identity block, convolutional block, and fully connected layers. From the input image, the convolutional layers extract features, which are then processed and transformed by the identity block and convolutional block. The final classification is then determined using the fully connected layers. In this study, I used dense 128 neurons in fully connected layers for the classification.

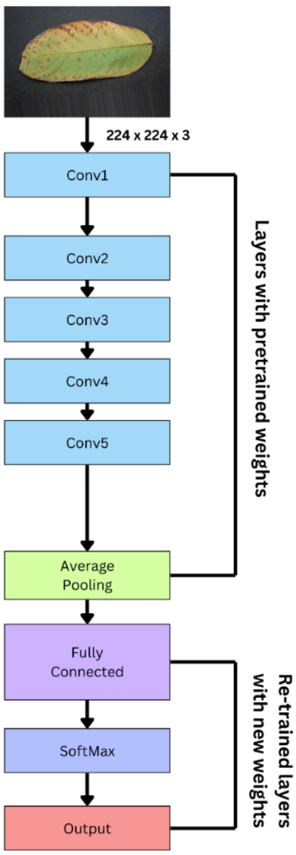

Figure 3.4: Architecture of RestNet50

VGG19: The VGG architecture is a straightforward and efficient CNN design, which has been applied to a variety of image recognition tasks, including object detection, segmentation, and object classification. Deep learning neural network VGG-19 has 19 connection layers, comprising 3 fully connected levels and 16 convolution layers. The fully connected layers will categorize the disease images based on the attributes that the convolution layers have extracted from the input images. A consistent receptive field is maintained by the convolutional layers by the use of small 3x3 filters, while computational efficiency is increased by the max-pooling layers through downsampling the spatial dimensions. To aggregate the learned features and provide predictions, the fully connected layers at the conclusion of the design function as classifiers. In this study, I used dense 1024, 512 neurons and dropout 0.8 in fully connected layers for the classification.

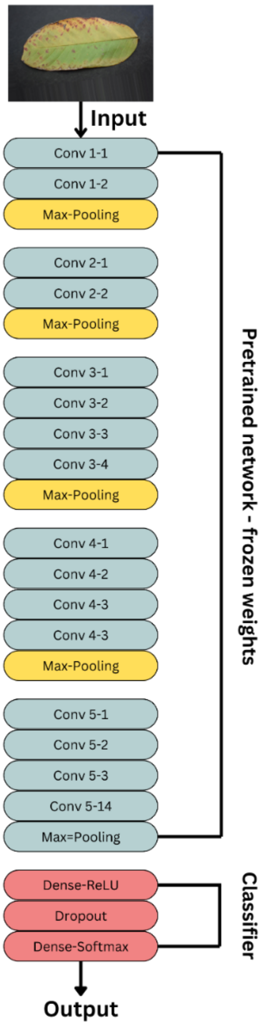

Figure 3.5: Architecture of VGG19

EfficientNet-B5: EfficientNet-B5 is a convolutional neural network (CNN) architectural type that is a member of the EfficientNet family. The depth, width, and resolution of the network are adaptively modified in the EfficientNet models to maximize accuracy and efficiency. There are several iterations of the network, from B0 to B7. EfficientNetB5 was the study's choice. In this network, the input image dimensions are defined as 224 x 224. In this study, I used dense 1024, 512 and dropout 0.5 in fully connected layers for the classification.

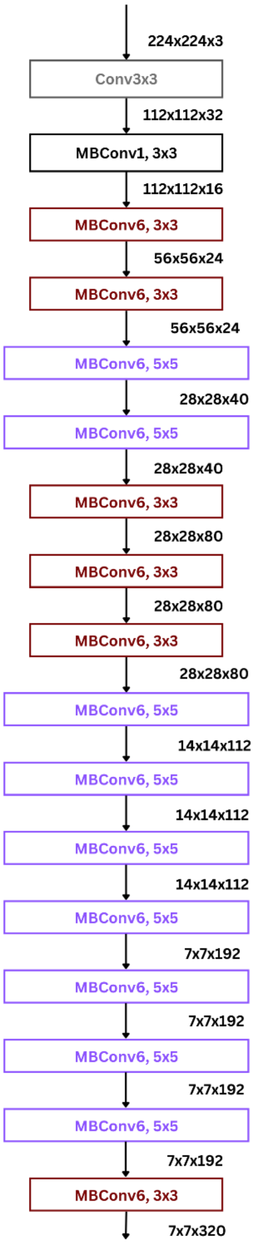

Figure 3.6: Architecture of EfficientNetB5

3.5 Implementation Requirements

To develop the ViT model there are some requirements.

- Computer with sufficient processing power and graphics processing unit (GPU) for deep learning tasks.
- Deep learning framework: Software libraries like TensorFlow, PyTorch, or Keras used to develop and train the ViT model.
- Image processing libraries: Libraries like OpenCV used for image manipulation tasks during data preprocessing.
- Data analysis libraries: Libraries like pandas and NumPy used for data organization and manipulation during the analysis phase.
- Data Visualization Library: Libraries like matplotlib and Sklearn used for data visualization during this phase.
- Dataset: Online available guava leaf disease image datasets on Kaggle website.
- Online resources and communities: Access to online tutorials, documentation, and forums related to deep learning, ViT applications, and the chosen deep learning framework is valuable for troubleshooting and learning best practices.

CHAPTER 04

EXPERIMENTAL RESULTS AND DISCUSSION

4.1 Experimental Setup

All the information I gathered was thoroughly analyzed. Kaggle is an online data collection tool that is the source of my data set. I applied to three transfer learning models, which are ResNet50, VGG19, and EfficientNet-B5. My proposed model is a Vision Transformer (ViT). I compared my proposed model to those transfer learning models, and I saw that my proposed model works better than those transfer learning models. The goal of this study is to find a state-of-the-art model to classify the images.

4.2 Experimental Results & Analysis

Evaluation Metrics:

Accuracy: Accuracy quantifies how frequently the classifier makes accurate predictions. Accuracy can be defined as the ratio of the total number of predictions to the number of right predictions.

$$Accuracy = \frac{TruePositive + TrueNegative}{TruePositive + FalsePositive + TrueNegative + FalseNegative}$$

Precision: The percentage of correctly predicted sightings among all positively predicted observations is known as precision. The statistic, which shows the precise proportion of accurate positive predictions the model produced, reflects the sensitivity of the model. It's crucial to remember that while many models target a high recall, this can be deceptive since, when accuracy and other performance metrics are taken into account, recall is not always a reliable indicator of success.

$$Precision = \frac{TruePositive}{TruePositive + FalsePositive}$$

Recall: Recall, which is the percentage of correctly predicted positive data points to the total number of actual positive data points. Assessing the model's ability to identify favorable results is crucial. It calculates the percentage of all positive labels that result in the expected positives. Although a high accuracy rate is generally desired, if it is not examined in conjunction with other important matrices, it may occasionally be deceptive.

$$Recall = \frac{TruePositive}{TruePositive + FalseNegative}$$

F-1 Score: The F1 score, which is computed as the harmonic mean, is a statistic that evaluates a classifier's performance by accounting for both accuracy and recall. Its primary goal is to evaluate and compare the performance of two classifiers. It is essential to consider that a model's inadequate efficiency can necessitate more adjustments if the harmonic mean of these measures has a particularly low mean.

$$F1 = 2 * \frac{Recall * Precision}{Recall + Precision}$$

Results & Analysis

In this phase, I evaluated the effectiveness of my proposed Vit model for classifying guava leaf diseases using a curated dataset. I compared its performance against established models, ResNet50, VGG19, and EfficientNetB5, all fine-tuned with appropriate settings.

Table 4.1: Accuracy for Classification of Individual Models

Model Name	Training Accuracy (%)	Validation Accuracy (%)	Testing Accuracy (%)
EfficientNetB5	96.59	93.33	94.97
VGG19	90.39	88.77	88.77
ResNet50	98.76	96.10	95.00
Vision Transformer (ViT)	97.42	95.60	97.54

To thoroughly assess their abilities, I employed confusion matrices and analyzed essential metrics like accuracy, precision, recall, and F1-score. This evaluation provided a comprehensive understanding of each model's capacity to identify guava leaf diseases accurately. By assigning distinct challenges during training and testing, I aimed to identify the best-performing model for guava leaf disease classification.

EfficientNetB5:

After training the model, I evaluated accuracy, recall, and F1 score by applying the learned model to my testing dataset.

These are my model results which are displayed in the table below.

Table 4.2: The Table of EfficientNetB5 That Shows Accuracy

Model Name	Training Accuracy (%)	Validation Accuracy (%)	Testing Accuracy (%)
EfficientNetB5	96.59	93.33	94.97

Table 4.3:The Table of EfficientNetB5 That Shows the Precision, Recall and F1 score

Class Name	Precision	Recall	F-1 Score
Canker	0.88	0.82	0.85
Dot	0.89	0.60	0.72
Healthy	0.97	1.00	0.98
Mummification	0.96	0.91	0.94
Rust	0.94	0.97	0.96

Confusion Matrix of EfficientNetB5:

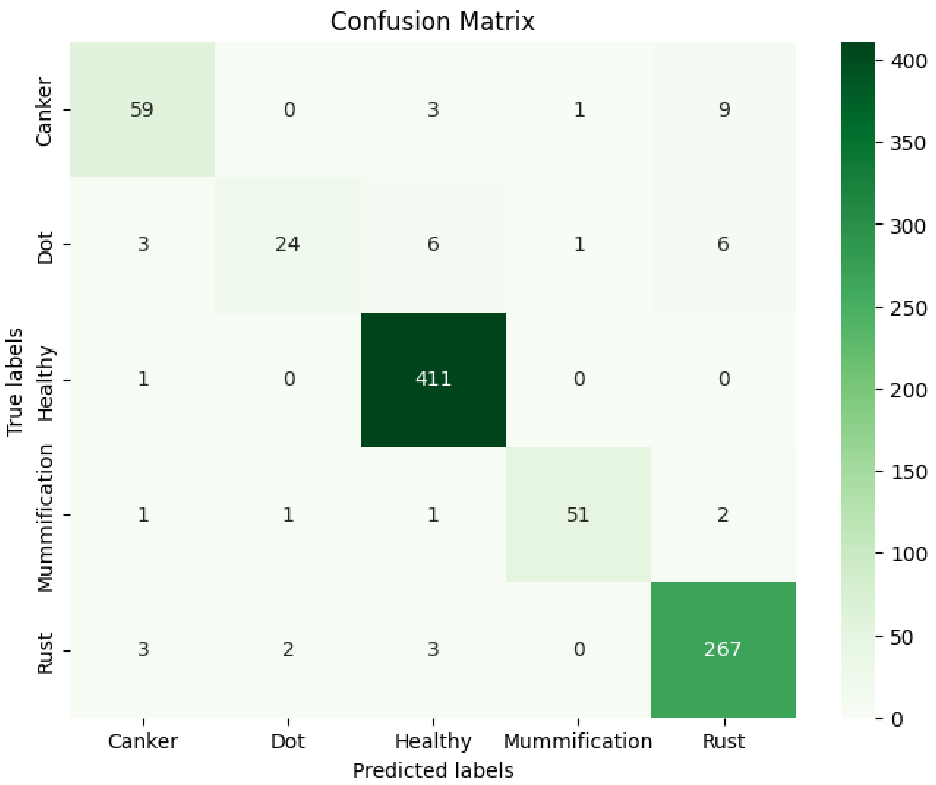

Figure 4.1: The Confusion Matrix of EfficientNetB5

Training, Validation Accuracy, and Loss Graph of EfficientNetB5:

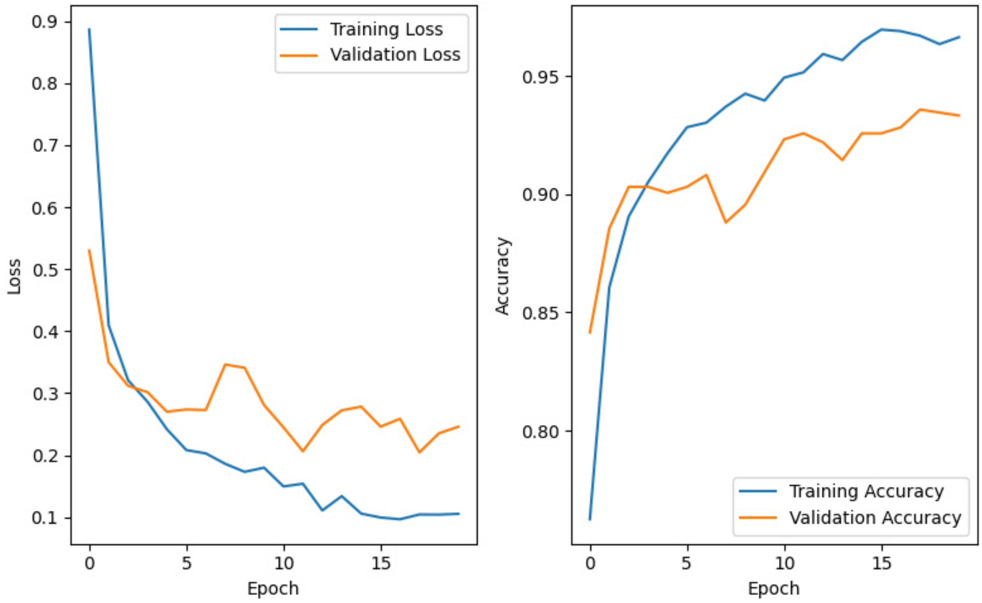

Figure 4.2: Training, Validation Accuracy, And Loss Graph of EfficientNetB5

VGG19:

After training the model, I evaluated accuracy, recall, and F1 score by applying the learned model to my testing dataset.

These are my model results which are displayed in the table below.

Table 4.4: The Table of VGG19 That Shows Accuracy

Model Name	Training Accuracy (%)	Validation Accuracy (%)	Testing Accuracy (%)
VGG19	90.39	88.77	88.77

Table 4.5: The Table of VGG19 That Shows the Precision, Recall and F1 score

Class Name	Precision	Recall	F-1 Score
Canker	0.74	0.58	0.65
Dot	0.50	0.05	0.09
Healthy	0.95	0.97	0.96
Mummification	0.80	0.84	0.82
Rust	0.85	0.98	0.91

Confusion Matrix of VGG19:

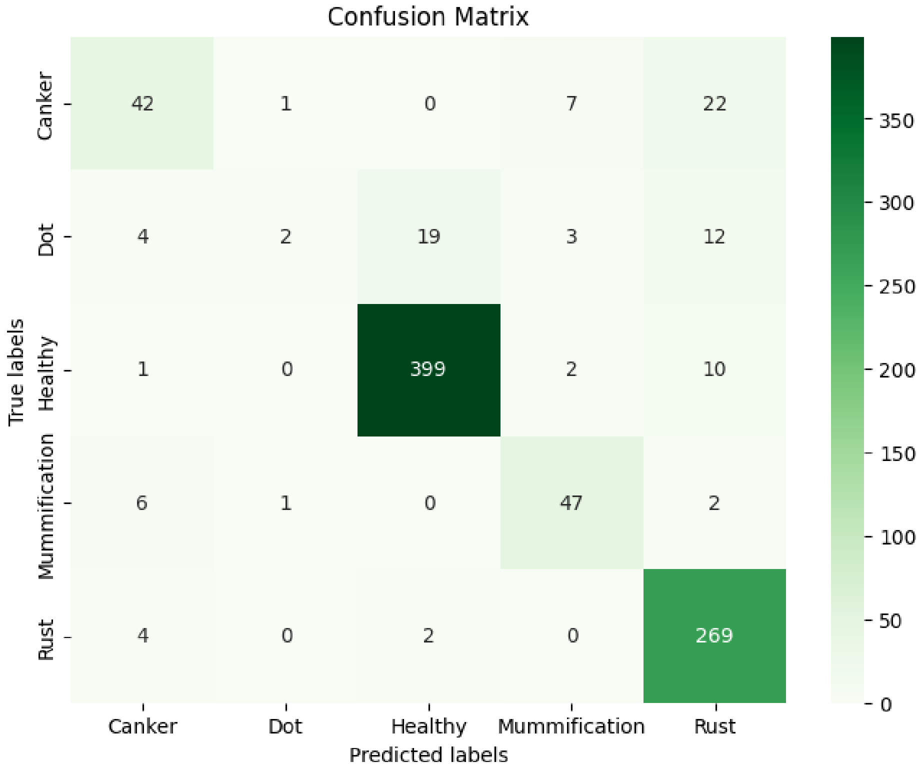

Figure 4.3: The Confusion Matrix of VGG19

Training, Validation Accuracy, and Loss Graph of VGG19:

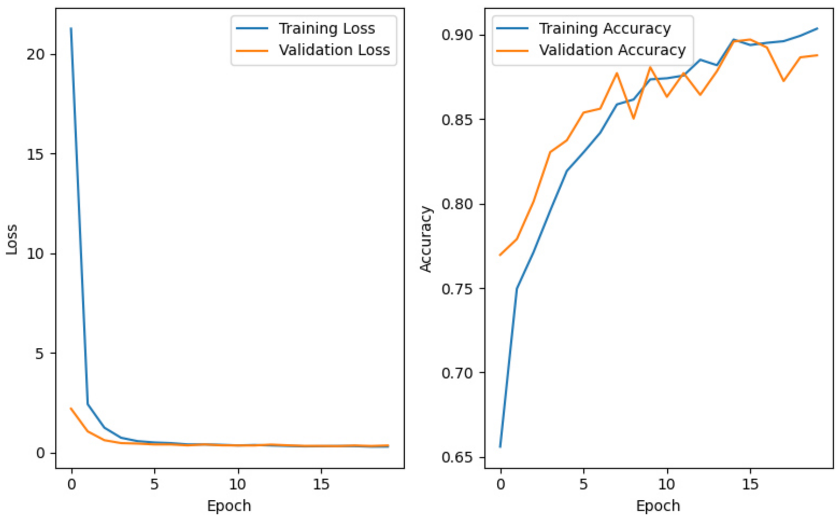

Figure 4.4: Training, Validation Accuracy, And Loss Graph of VGG19

ResNet50:

After training the model, I evaluated accuracy, recall, and F1 score by applying the learned model to my testing dataset.

These are my model results which are displayed in the table below.

Table 4.6: The Table Of ResNet50 That Shows Accuracy

Model Name	**Training Accuracy (%)**	**Validation Accuracy (%)**	**Testing Accuracy (%)**
ResNet50	98.76	96.10	95.00

Table 4.7: The Table of ResNet50 That Shows the Precision, Recall and F1 score

Class Name	**Precision**	**Recall**	**F-1 Score**
Canker	0.92	0.68	0.78
Dot	0.94	0.85	0.89
Healthy	1.00	0.99	0.99
Mummification	0.72	0.96	0.82
Rust	0.97	0.99	0.98

Confusion Matrix of ResNet50:

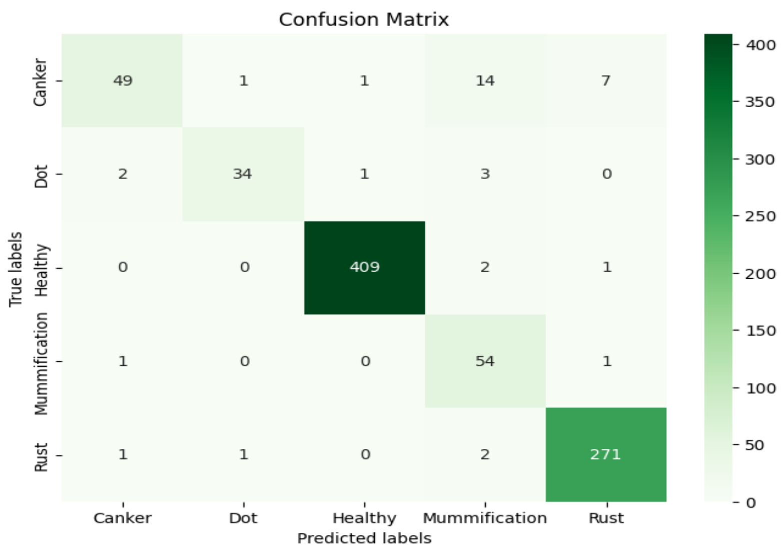

Figure 4.5: The Confusion Matrix of ResNet50

Training, Validation Accuracy, and Loss Graph of ResNet50:

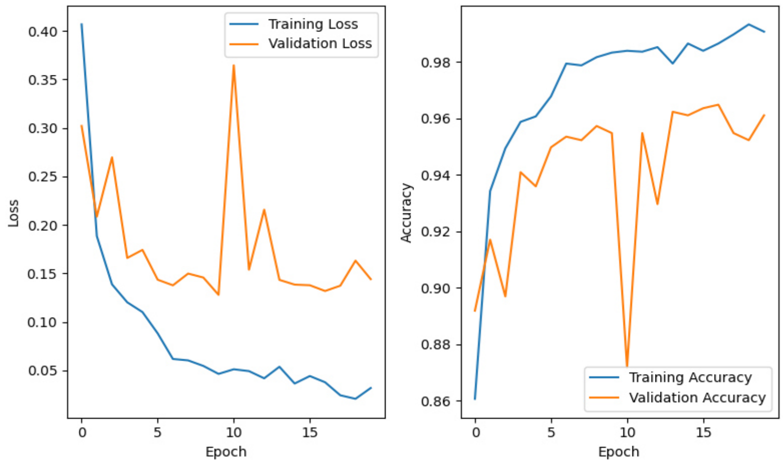

Figure 4.6: Training, Validation Accuracy, And Loss Graph of ResNet50

Proposed Vision Transformer (ViT):

After training the model, I evaluated accuracy, recall, and F1 score by applying the learned model to my testing dataset.

These are my model results which are displayed in the table below.

Table 4.8: The Table of Proposed ViT That Shows Accuracy

Model Name	Training Accuracy (%)	Validation Accuracy (%)	Testing Accuracy (%)
Vision Transformer (ViT)	97.42	95.60	97.54

Table 4.9: The Table of Proposed ViT That Shows the Precision, Recall And F1 score

Class Name	Precision	Recall	F-1 Score
Canker	0.91	0.88	0.89
Dot	0.94	0.85	0.89
Healthy	0.99	1.00	0.99
Mummification	1.00	0.98	0.99
Rust	0.97	0.99	0.98

Confusion Matrix of Proposed Vision Transformer (ViT):

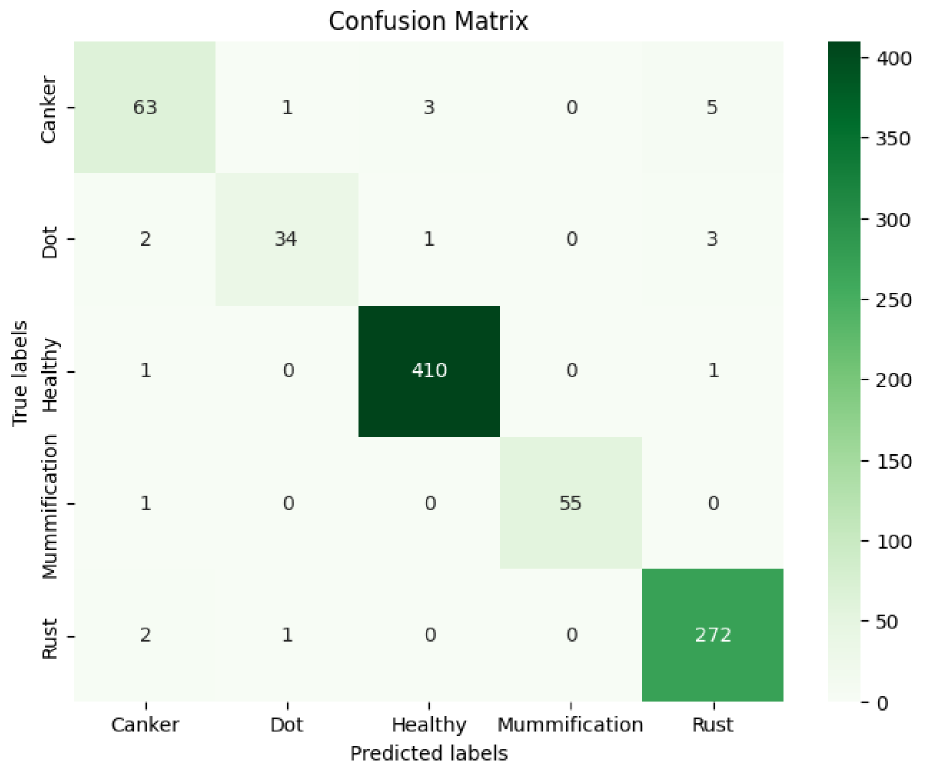

Figure 4.7: The Confusion Matrix of Proposed Vision Transformer (ViT)

Training, Validation Accuracy, and Loss Graph of Proposed Vision Transformer (ViT):

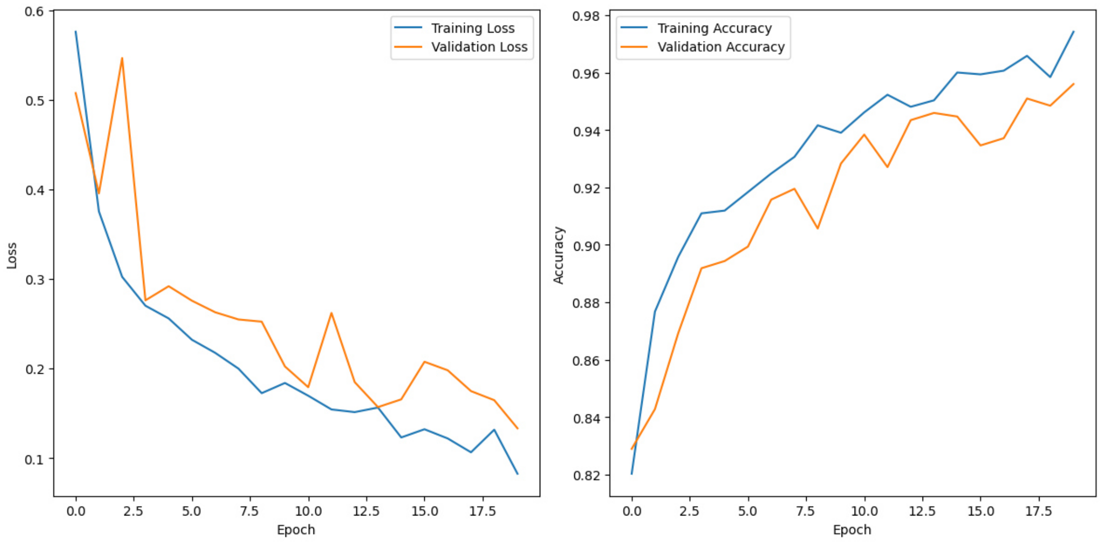

Figure 4.8: Training, Validation Accuracy, And Loss Graph of Proposed Vision Transformer (ViT)

Table 4.10: The performance comparison of the proposed study with other state-of-the-art studies in classification Guava Leaf Disease.

Paper Name	Applied Model	Accuracy
A. S. Doutoum, R. Eryigit, and B. Tugrul, "Classification of Guava Leaf Disease using Deep Learning [1]	VGG16, ResNet50, InceptionV3, EfficientNet-B3	EfficientNet-B3 achieved 94.93% accuracy on the test data.
J. Rashid, I. Khan, G. Ali, S. ur Rehman, F. Alturise, and T. Alkhalifah, "Real-Time Multiple Guava Leaf Disease Detection from a Single Leaf Using Hybrid Deep Learning Technique [3]	YOLOv3, YOLOv4, YOLOv4tiny, GIP-MU-Net, GLSM ,GMLDDmodel	GIP-MU-Net model achieved 92.41%, GLSM gained 83.40%, GMLDD technique achieved 73.3% precision, 73.1% recall.
H. Alshammari, K. Gasmi, I. Ben Ltaifa, M. Krichen, L. Ben Ammar, and M. A. Mahmood, "Olive Disease Classification Based on Vision Transformer and CNN Models [10]	ViT+VGG-16, ViT+VGG-19, Transformer (ViT)	Achieved 96% accuracy for multiclass and 97% for binary classification.
Md. Mustak Un Nobi, Md. Rifat, M. F. Mridha, S. Alfarhood, M. Safran, and D. Che, "GLD-Det: Guava Leaf Disease Detection in Real-Time	EfficientNetV2B2, EfficientNetB0, EfficientNetB2, EfficientNetB1, MobileNetV2, GLD-Det	GLD-Det model accuracy: 0.9714 on Dataset D1, 0.9712 on Dataset D2.

Using Lightweight Deep Learning Approach Based on MobileNet [2]		
My Proposed Model	Vision Transformers (ViT)	ViT = 97.54%

Table 4.10 presents a performance comparison of our suggested study with other recent studies. To ensure fair comparison, we have used the most current published research from 2021 to 2024. The research demonstrates that there is currently no existing study focused on guava leaves using the proposed ViT model. Furthermore, our model achieved an impressive accuracy score of 97.54% in accurately classifying guava leaf diseases.

4.3 Discussion

The proposed ViT model, after training on extensive and diverse Guava leaf diseases, could provide accurate predictions for early identification and monitoring. Because transformer techniques are effective at reducing the model complexity and performance, they provide a promising contender to improve the effectiveness of guava leaf disease prediction. My proposed model accuracy is 97.54% where the loss is 0.0911.

CHAPTER 5

IMPACT ON SOCIETY, ENVIRONMENT AND SUSTAINABILITY

5.1 Impact on Society

The proposed approach for detecting guava leaf disease offers significant advantages to both the agriculture industry and society as a whole. My model, created through careful examination of actual datasets, is precisely crafted to analyze and detect key components and traits that indicate guava leaf illnesses. This study aims to provide multiple social benefits, specifically in the field of education and raising awareness about the incidence of guava leaf diseases and the preventive methods that are available.

The high level of accuracy demonstrated by my model in detecting diseases and continuously monitoring them has the potential to bring about significant changes. It enables early intervention and empowers guava producers to make well-informed decisions in managing potential diseases in their orchards. My method is highly efficient and decreases both time and computing needs, resulting in accurate disease prediction.

This study attempts to uncover the underlying causes that contribute to guava leaf illnesses by conducting a thorough analysis of data using advanced diagnostic techniques. At a societal level, the goal is to see the recommended technique being widely accepted and implemented by the guava farming community. The main goal is to promote the spread of knowledge and encourage proactive health practices in guava farming, to create a better knowledgeable and health-conscious guava agricultural community.

The primary objective is to provide guava farmers with the required tools and expertise to efficiently oversee the well-being of their orchards. This will help reduce the effects of guava leaf diseases and promote a more robust and sustainable guava farming sector. To summarize, my model provides a potential approach for accurately detecting diseases in guava leaves, making a substantial contribution to improving guava health practices and raising awareness on a larger scale.

5.2 Impact on Environment

The suggested approach for detecting guava leaf disease shows great potential for use in farmlands and underprivileged regions. It offers simplified diagnostic approaches that can reduce complexity and save time efficiently. The method's direct and non-intrusive characteristics guarantee a beneficial environmental outcome without any negative consequences. Healthcare has become more accessible and convenient in distant places as individuals no longer have to go to urban hubs for guava leaf disease detection. This prediction model enhances orchard health monitoring by offering accurate projections of probable outcomes. It effectively addresses concerns regarding the expenses associated with local treatment and the affordability of guava leaf disease identification. The design of the product is user-friendly, allowing users with any degree of competence to easily utilize it.

By implementing my suggested strategy, the ability to identify and control guava leaf diseases becomes a tangible outcome, greatly improving the agricultural and social environment. I am confident that implementing my proposed model will accelerate significant progress in the field of agricultural scientific technology, ultimately enhancing the quality of guava health care and agricultural technology universally.

5.3 Ethical Aspects

Before implementing my guava leaf disease detection system, it is imperative to prioritize ethical considerations to prevent the unintentional disclosure of sensitive information, diagnostic outcomes, or unforeseen consequences. The detection and management of guava leaf diseases, both in practical applications and future research endeavors, stand to benefit significantly from my recommended approach. Guava leaf disease concerns transcend geographical boundaries and impact a global population. My method empowers individuals, whether they are guava farmers or well-informed stakeholders, to anticipate the onset and progression of guava leaf diseases, providing a valuable tool for proactive healthcare management within the guava farming community on a global scale.

5.4 Sustainability Plan

Below are a few possible elements that could be included in a sustainability strategy for utilizing deep learning in the early identification of guava leaf diseases:

- Efficiency for Energy: Reducing the environmental impact of diagnostic procedures can be accomplished significantly by improving the energy consumption of computer systems and other diagnostic equipment utilized in guava leaf disease detection.
- Sustainable energy utilization: Mitigating the environmental effect of the diagnostic procedure might be achieved by integrating sustainable sources of energy, such as solar or wind power, to power computers and other devices involved in guava leaf disease detection.
- Responsible disposal of equipment: Ensuring proper disposal of computers and other electronic devices used in guava leaf disease detection when they are no longer useful can aid in reducing waste and preventing the introduction of hazardous compounds into the environment.
- Healthcare provider collaboration: Collaborating with agricultural experts and stakeholders to integrate the guava leaf disease detection procedure into regular agricultural services can effectively diminish the overall environmental impact by minimizing the need for additional transportation and resources.
- Education and outreach: Increasing awareness and acceptance of guava leaf disease detection through education and outreach initiatives targeted at farmers, agricultural professionals, and stakeholders can contribute to more sustainable practices and improved orchard health management.
- Continuous improvement: Consistently evaluating and enhancing the sustainability of guava leaf disease detection processes can optimize their long-term environmental friendliness and contribute to the overall sustainability of agricultural practices.

CHAPTER 6

SUMMARY, CONCLUSION, RECOMMENDATION, AND IMPLICATION FOR FUTURE RESEARCH

6.1 Summary of the Study

Several recent research has investigated the utilization of deep learning techniques for the detection of plant diseases. Our current research expands on this previous research by introducing an innovative method for classifying guava leaf diseases. We utilize a Vision Transformer (ViT) model for this purpose. The results of our study outperform conventional transfer learning models such as ResNet50, VGG19, and EfficientNetB5, attaining higher accuracy (97.54%) while still preserving computational efficiency.

This research makes a substantial contribution to the current endeavors of utilizing artificial intelligence (AI) for sustainable agriculture. ViT has the potential to enhance crop productivity and food security by enabling the early diagnosis of plant diseases. Subsequent research endeavors may analyze the utilization of ViT for categorizing diseases in different crops and assess its appropriateness for promptly detecting diseases in field environments. Through the utilization of deep learning and artificial intelligence, we have the ability to transform agricultural methods, granting farmers greater control and guaranteeing worldwide food stability.

6.2 Conclusion

This study aimed to assess the effectiveness of a Vision Transformer (ViT) model in accurately categorizing guava leaf diseases. The approach we provided yielded impressive outcomes, with a training accuracy of 97.42%, a validation accuracy of 95.60%, and a testing accuracy of 97.54%. The performance of ViT in guava leaf disease classification outperformed classic transfer learning models such as ResNet50, VGG19, and EfficientNetB5, demonstrating its promise.

The results of our study add to the increasing amount of research investigating the use of deep learning methods in the field of agriculture. The capacity of ViT to capture extensive interdependencies in images demonstrates its value in precise disease identification, perhaps resulting in prompt intervention and enhanced agricultural productivity.

This finding sets the stage for additional investigation of ViT in many agricultural applications. Subsequent research endeavors could examine its efficacy in categorizing ailments of different crops and delve into the possibility of detecting diseases in real-time inside field environments. Through harnessing the capabilities of deep learning and artificial intelligence, we can make substantial contributions to the advancement of sustainable agriculture methods. This will help guarantee food security on a worldwide scale and empower farmers to make well-informed choices on the most effective crop management strategies.

6.3 Implication for Further Study

The study effectively showcased the capabilities of Vision Transformers (ViT) in accurately classifying guava leaf diseases. However, there exist intriguing prospects for additional investigation in this field.

An area for future research entails examining the efficacy of data augmentation approaches. By intentionally augmenting the dataset with differences in illumination, rotation, and scale, we have the potential to enhance the model's capacity to generalize and handle unseen data more effectively. In addition, investigating the incorporation of methods such as saliency map visualization could offer important understanding of the decision-making mechanism of the ViT model, assisting in the interpretation of its predictions and potentially pinpointing areas for further enhancement.

Moreover, it is necessary to investigate the feasibility of implementing the proposed ViT model in practical agricultural environments. Creating a mobile app that utilizes smartphone cameras to identify diseases on-site could enable farmers to take prompt actions and enhance crop management techniques. Further investigation could focus on optimizing the ViT model for mobile implementation, guaranteeing both efficiency and accuracy in disease categorization, even when computational resources are constrained. By exploring these paths for more research, we can improve and optimize the practical implementations of ViT technology in the detection of agricultural diseases. This will make a substantial contribution to the advancement of sustainable and efficient agricultural methods, ultimately resulting in enhanced crop productivity and global food security.

REFERENCE

[1] A. S. Doutoum, R. Eryigit, and B. Tugrul, "Classification of Guava Leaf Disease using Deep Learning," WSEAS TRANSACTIONS ON INFORMATION SCIENCE AND APPLICATIONS, vol. 20, pp. 356–363, Oct. 2023, doi: 10.37394/23209.2023.20.38.

[2] Md. Mustak Un Nobi, Md. Rifat, M. F. Mridha, S. Alfarhood, M. Safran, and D. Che, "GLD-Det: Guava Leaf Disease Detection in Real-Time Using Lightweight Deep Learning Approach Based on MobileNet," Agronomy, vol. 13, no. 9, p. 2240, Aug. 2023, doi: 10.3390/agronomy13092240.

[3] J. Rashid, I. Khan, G. Ali, S. ur Rehman, F. Alturise, and T. Alkhalifah, "Real-Time Multiple Guava Leaf Disease Detection from a Single Leaf Using Hybrid Deep Learning Technique," Computers, Materials & Continua, vol. 74, no. 1, pp. 1235–1257, 2023, doi: 10.32604/cmc.2023.032005.

[4] S. Mumtaz, M. Raza, O. D. Okon, S. U. Rehman, A. E. Ragab, and H. T. Rauf, "A Hybrid Framework for Detection and Analysis of Leaf Blight Using Guava Leaves Imaging," Agriculture, vol. 13, no. 3, p. 667, Mar. 2023, doi: 10.3390/agriculture13030667.

[5] A. Tabbakh and S. S. Barpanda, "A Deep Features Extraction Model Based on the Transfer Learning Model and Vision Transformer 'TLMViT' for Plant Disease Classification," IEEE Access, vol. 11, pp. 45377–45392, 2023, doi: 10.1109/access.2023.3273317.

[6] S. Parez, N. Dilshad, N. S. Alghamdi, T. M. Alanazi, and J. W. Lee, "Visual Intelligence in Precision Agriculture: Exploring Plant Disease Detection via Efficient Vision Transformers," Sensors, vol. 23, no. 15, p. 6949, Aug. 2023, doi: 10.3390/s23156949.

[7] M. T. SAbirami, "Application of Image Processing in Diagnosing Guava Leaf Diseases," International Journal of Scientific Research and Management, vol. 5, no. 7, pp. 5927–5933, Jul. 2017, doi: 10.18535/ijsrm/v5i7.19.

[8] N. Van Hieu, N. L. H. Hien, L. Van Huy, N. H. Tuong, and P. T. K. Thoa, "PlantKViT: A Combination Model of Vision Transformer and KNN for Forest Plants Classification," Journal of Universal Computer Science, vol. 29, no. 9, pp. 1069–1089, Sep. 2023, doi: 10.3897/jucs.94657.

[9] H. Li, S. Li, J. Yu, Y. Han, and A. Dong, "Plant disease and insect pest identification based on vision transformer," Apr. 2022, doi: 10.1117/12.2628467.

[10] H. Alshammari, K. Gasmi, I. Ben Ltaifa, M. Krichen, L. Ben Ammar, and M. A. Mahmood, "Olive Disease Classification Based on Vision Transformer and CNN Models," Computational Intelligence and Neuroscience, vol. 2022, pp. 1–10, Jul. 2022, doi: 10.1155/2022/3998193.

[11] Y. Borhani, J. Khoramdel, and E. Najafi, "A deep learning-based approach for automated plant disease classification using vision transformer," Scientific Reports, vol. 12, no. 1, Jul. 2022, doi: 10.1038/s41598-022-15163-0.

[12] P. Bhuyan and P. K. Singh, "Evaluating Deep CNNs and Vision Transformers for Plant Leaf Disease Classification," Lecture Notes in Computer Science, pp. 293–306, 2024, doi: 10.1007/978-3-031-50583-6_20.

[13] P. P. . et. al., "Guava Leaf Disease Classification Using Support Vector Machine", TURCOMAT, vol. 12, no. 7, pp. 1177–1183, Apr. 2021.

[14] A. Almadhor, H. T. Rauf, M. I. U. Lali, R. Damaševičius, B. Alouffi, and A. Alharbi, "AI-Driven Framework for Recognition of Guava Plant Diseases through Machine Learning from DSLR Camera

Sensor Based High Resolution Imagery," Sensors, vol. 21, no. 11, p. 3830, Jun. 2021, doi: 10.3390/s21113830.

[15] L. A. Aldakhil and A. A. Almutairi, "Multi-Fruit Classification and Grading Using a Same-Domain Transfer Learning Approach," IEEE Access, vol. 12, pp. 44960–44971, 2024, doi: 10.1109/access.2024.3379276.

APPENDIX

This thesis is complemented by multiple appendices that offer additional information to improve comprehension and replicability of the research.

- **Guava Leaf Disease Dataset:** It provides a comprehensive description of the dataset containing information about guava leaf diseases. The dataset was utilized for both training and evaluating the ViT model. The report will provide details regarding the overall quantity of photos, the allocation of images among various illness categories, and any preprocessing methods employed to verify the data's appropriateness for the model.
- **Vision Transformer Model Architecture:** This contains a comprehensive graphic depiction of the architecture of the ViT model. The diagram will depict the different constituents of the model, encompassing the encoder and decoder blocks, along with their distinct operations. This visual assistance will offer readers a precise comprehension of the internal framework of the ViT.
- **Code Implementation:** The appendix includes chosen elements of the source code used to design and train the ViT model, to enhance transparency and repeatability. This may include essential tasks associated with data pre-processing, defining the model architecture, conducting the training process, and evaluating the metrics.
- **Experimental Results:** This appendix contains detailed tables describing the numerical outcomes of the trials. The report will provide information on the training and validation loss curves, accuracy metrics such as accuracy, precision, recall, and F1-score, and maybe confusion matrices to visually represent the model's performance across several disease categories.
- **Supplementary Visualizations:** It includes any extra visual components that improve the comprehension of the research. Examples of possible visualizations include sample photos from the guava leaf disease dataset, illustrating both healthy and sick leaves. Additionally, attention maps generated by the ViT model can be used to analyze its focus during classification. Other relevant visualizations that assist in analyzing the results may also be included.

CO Description for FYDP

CO	CO Descriptions	PO
	Phase -I	
CO1	To solve a **guava leaf disease detection** problem for the Final Year Design Project (FYDP), use information that has been learned recently and in previous years.	**PO1**
CO2	To create an outcome for this FYDP, analyze several parts of the objectives.	**PO2**
CO3	Identify the problems, define the topics of study, and set the objectives for the FYDP by doing a literature review.	**PO4**
CO4	Throughout the FYDP's development life cycle, conduct cost and economic analysis and make appropriate use of project management techniques.	**PO11**
	Phase -II	
CO5	In this FYDP, take into account cultural, socioeconomic, and environmental aspects while designing and developing technological solutions, system components, or processes that satisfy requirements and adhere to public health and safety regulations.	**PO3**
CO6	Adhere to relevant limitations in this FYDP while selecting and utilizing suitable approaches, resources, and modern engineering and IT technologies to handle complex engineering processes, including modeling and prediction.	**PO5**
CO7	Applying logical reasoning informed by contextual awareness, analyze societal, health, safety, legal, and cultural factors, as well as related obligations, in the context of professional engineering practice and the solution to this challenge.	**PO6**

CO8	Recognize and assess the long-term viability and influence of professional engineering initiatives in resolving complex engineering problems in social and environmental contexts.	**PO7**
CO9	Adhere to professional standards and norms and apply ethical concepts in this FYDP.	**PO8**
CO10	Ability to function well in a variety of teams and multidisciplinary environments as a leader or as a team member in addition to working alone during this FYDP.	**PO9**
CO11	Capable of understanding and producing detailed reports and design documentation, as well as giving and receiving clear instructions throughout this FYDP, effectively communicating with the technical community and the general public about complicated engineering projects.	**PO10**
CO12	Recognize the value of self-directed learning and lifetime learning in the context of the rapidly changing technological world and be prepared and able to participate in lifelong learning activities.	**PO12**

GUAVA LEAF DISEASE DETECTION: A POWERFUL APPROACH WITH VISION TRANSFORMERS

ORIGINALITY REPORT

25%	19%	15%	13%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	dspace.daffodilvarsity.edu.bd:8080 Internet Source	6%
2	Submitted to Daffodil International University Student Paper	2%
3	wseas.com Internet Source	1%
4	www.mdpi.com Internet Source	1%
5	"Human-Centric Smart Computing", Springer Science and Business Media LLC, 2024 Publication	1%
6	ijece.iaescore.com Internet Source	1%
7	123dok.com Internet Source	1%
8	Submitted to Midlands State University Student Paper	1%

www.ijisae.org