Bangladeshi Rice and Tea leaf disease detection using Machine Learning

BY

Md. Rakib

201-15-13846

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

Supervised By

Mr. Abdus Sattar
Assistant Professor & Coordinator MSc.
Department of CSE
Daffodil International University

Co-Supervised By

Raja Tariqul Hasan Tusher
Assistant Professor
Department of CSE
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

JANUARY 2024

APPROVAL

This Project titled **"Bangladeshi Rice and Tea leaf disease detection using Machine Learning"**, submitted by Md. Rakib, ID No: 201-15-13846 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 *24th January 2024*.

BOARD OF EXAMINERS

Chairman

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

Internal Examiner

Md. Abbas Ali Khan
Assistant Professor
Department of Computer Science and Engineering
Daffodil International University

Internal Examiner

Mohammad Monirul Islam
Assistant Professor
Department of Computer Science and Engineering
Daffodil International University

External Examiner

Dr. Ahmed Wasif Reza
Professor
Department of Computer Science and Engineering
East West University

DECLARATION

We hereby declare that this thesis has been done by us under the supervision of **Abdus Sattar,** Assistant Professor**, Department of CSE** Daffodil International University. We also declare that neither this project nor any part of this project has been submitted elsewhere for the award of any degree or diploma.

Supervised by:

____

Abdus Sattar
Assistant Professor
Department of CSE
Daffodil International University

Submitted by:

Rakib

____

Md. Rakib
ID: 201-15-13846
Department of CSE
Daffodil International University

©Daffodil International University ii

ACKNOWLEDGEMENT

First we express our heartiest thanks and gratefulness to almighty God for His divine blessing makes us possible to complete the final year project/internship successfully.

We really grateful and wish our profound our indebtedness to **Abdus Sattar,** Assistant professor Department of CSE Daffodil International University, Dhaka. Deep Knowledge & keen interest of our supervisor in the field of "Machine Learning & Image Processing" to carry out this project. His endless patience ,scholarly guidance ,continual encouragement , constant and energetic supervision, constructive criticism , valuable advice ,reading many inferior draft and correcting them at all stage have made it possible to complete this project.

We would like to express our heartiest gratitude to **Dr. Sheak Rashed Haider Noori** and Head**,** Department of CSE, for his kind help to finish our project and also to other faculty member and the staff of CSE department of Daffodil International University.

We would like to thank our entire course mate in Daffodil International University, who took part in this discuss while completing the course work.

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

ABSTRACT

Bangladesh's economy is based primarily on agriculture, with the production of tea and rice being essential to maintaining livelihoods and food security. The persistent risk of illnesses in rice and tea plants, however, makes agricultural productivity and financial stability extremely difficult. The goal of this research project is to create reliable and effective disease detection models for Bangladeshi rice and tea leaves by utilizing machine learning, more especially deep learning and image analysis approaches as these two items are consumed in a huge amount on a daily basis from poor to rich. Given Bangladesh's strong agricultural economy, early disease identification in these crops is essential to maintaining both food security and economic stability. The research seeks to offer user-friendly software tools for rapid and precise disease diagnosis, encouraging precision agriculture, lowering pesticide usage, and improving food security. This will be accomplished by training machine learning models on a big datasets of photos of healthy and diseased plants. The project is projected to play a key role in the modernization and resilience of Bangladesh's agriculture by empowering smallholder farmers through technology transfer, promoting sustainable agricultural practices, and supporting government activities. Alongside it will help them to reduce the amount of their loss in farming.

Keywords: EfficientNetB7, MobileNetV3Large, ResNet101, leaf, tea, rice, disease detection, machine learning.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Introduction

The agricultural sector is vital to the country, especially in Bangladesh's fertile soil where the cultivation of rice and tea is a way of life. However, despite the beautiful scenery and growing plantations, there is a silent threat that never goes away which is illnesses that affect the same crops that the country depends on. Bangladesh, a country with a strong agricultural legacy and a sizable farming population, is faced with the ongoing problem of preserving staple crops like rice and the business that drives the country's economy the tea sector. With the goal of using machine learning, the pinnacle of contemporary technological power, to combat the persistent problem of diseases in rice and tea leaves, this thesis sets out on a transformative journey into the heart of Bangladesh's agrarian tapestry. The importance of developing agricultural methods cannot be overstated as the world struggles with the complexities of a rapidly changing climate and the need for sustainable practices.

The research presented here is motivated by two times dedications: first, to protect the customs and means of livelihood that are established in Bangladesh's agricultural past; and second, to welcome the possibility that modern technology can bring these customs into a more robust and technologically advanced future. The focal point is disease detection, which represents a pivotal point at which the combination of traditional methods and modern artificial intelligence can pave the way for increased agricultural yields, financial stability, and food security. This study aims to develop machine learning models that are specifically tailored to the difficulties of disease identification in rice and tea, the two main crops in Bangladesh. A detailed approach is necessary due to the complexity of agricultural ecosystems, where factors include soil conditions and climatic patterns. This research initiatives to create models that can identify early indicators of diseases using a thorough investigation of machine learning algorithms. This will enable farmers to obtain accurate and timely information for efficient action by seeing the leaves.

1.2 Motivation

The motivation for researching "Bangladesh Rice & Tea Leaf Disease Detection Using Machine Learning" began in an understanding of the critical role that agriculture plays in Bangladesh's socioeconomic structure. Being a country with strong roots in agriculture, the productivity and health of secure crops especially rice and tea have a significant impact on the well-being of its citizens. However, the hidden influence of diseases on these vital crops poses a constant threat to this prosperity.

The need to both preserve agricultural traditions and adjust to the demands of a rapidly changing technological landscape highlight how urgent it is to address this challenge. Despite their resilience, traditional farming methods face the challenges of a changing climate and the emergence of novel disease patterns. Within this framework, machine learning offers a revolutionary chance to strengthen traditional methods with the accuracy and insight provided by state-of-the-art technology. The idea behind this research is that Bangladesh's agricultural resilience may undergo a paradigm shift as a result of machine learning being integrated into disease detection procedures. Not only is it an academic project to quickly detect diseases in rice and tea leaves, but it is also a practical solution to the problems that the farming community faces. My goal is to provide farmers with active insights through the use of machine learning algorithms, so they can take specific steps accordingly to protect their crops from the damaging effects of disease that led them to a big loss.

Moreover, the incentive includes a wider social framework. Increased food security and economic stability for farmers have a direct connection with a strong and technologically advanced agricultural sector. I have collected the tea dataset all by myself through visiting few tea states like Finlay tea state, Nurjahan tea state etc. in Sreemangal and rice data is collected from [10] “Dhan Somadhan”. My goal as I set out on this research journey is to make the connection between innovation and tradition as smooth as possible so that the age-old methods of growing tea and rice combine with the revolutionary power of machine learning. By doing this, we hope to create a more robust agricultural industry as well as a more promising and sustainable future for Bangladesh.

1.3 Rationale of the study

The rationale for looking into "Bangladesh Rice & Tea Leaf Disease Detection Using Machine Learning" is deeply rooted in the essential significance of the agricultural landscape, particularly in the context of Bangladesh, where agriculture is the backbone of the economy and a primary source of livelihood for a significant portion of the population. Diseases are a persistent problem for the rice and tea industries, which are essential to the country's food security and economic stability but can have a negative effect on crop yields and farmer incomes.

The crucial requirement for improving Bangladesh's agricultural adaptability in the face of changing environmental and economic dynamics is the driving force behind this study. Farmers who rely on traditional disease detection methods may suffer significant losses because these methods frequently fail to provide timely and accurate insights. The application of machine learning to disease detection procedures shows promise as a novel approach that could completely change the way we protect these essential crops.

Bangladesh also has a wide range of agricultural climates and several types of diseases that present particular difficulties for conventional farming methods. Utilizing machine learning offers the potential for precise and customized disease identification due to its adaptability and capacity to handle complex data patterns. Through the customization of machine learning models to the unique conditions and diseases that are common in Bangladesh, this research seeks to offer a workable and contextually appropriate way to improve the agricultural sector's resilience.

The rationale behind the study is based on the conviction that Bangladesh's commitment to upgrading its agricultural practices is appropriate with the use of modern technology for disease detection. I hope that by equipping farmers with the necessary tools and knowledge to recognize and treat diseases efficiently, crop losses will be minimized and farmers' socioeconomic standing will rise. The study's justification is, in simple terms, strongly connected to the larger goal of Bangladesh's agricultural sector being prosperous, resilient, and empowered by technology.

1.4 Research Questions

- Can machine learning models, specifically deep convolutional neural networks (CNNs), effectively classify and diagnose diseases in rice and tea plants based on visual cues?
- How do machine learning models for rice and tea disease detection compare to traditional, manual methods in terms of accuracy, speed, and reliability?
- In the context of rice and tea leaf disease detection, how does the size and diversity of the dataset influence the performance of deep CNN models?
- What impact do different image enhancement and normalization preprocessing techniques have on the accuracy of machine learning models for rice and tea leaf disease detection?
- Can transfer learning techniques enhance the accuracy of machine learning models for detecting diseases in rice and tea plants in Bangladesh?
- How do environmental factors such as varying light conditions and different stages of disease progression affect the performance of deep CNN models in rice and tea leaf disease detection?
- What are the advantages and disadvantages of employing deep CNNs and computer vision techniques for disease recognition in rice and tea plants in the context of Bangladesh?
- How can the disease detection methods developed for rice and tea plants be practically applied in agricultural practices and conservation efforts?
- Can machine learning models trained for rice and tea leaf disease detection be adapted to recognize diseases in other crops or plant species?
- If efficient disease detection in rice and tea plants is achieved in Bangladesh, how could it positively impact agricultural productivity, environmental conservation, and biodiversity management?

1.5 Expected Output

- ❖ Yes, machine learning models, especially deep convolutional neural networks (CNNs), have shown effectiveness in classifying and diagnosing diseases in rice and tea plants using visual cues.
- ❖ Machine learning models generally outperform traditional manual methods in terms of accuracy, and they can provide faster and more reliable disease detection in rice plants.
- ❖ The size and diversity of the dataset significantly impact the performance of deep CNN models.
- ❖ Normalization and image enhancement by enhancing the caliber and consistency of input data, preprocessing approaches can have a positive effect on the accuracy of machine learning models for illness identification.
- ❖ Yes, transfer learning techniques can enhance the accuracy of machine learning models.
- ❖ Environmental factors like light conditions and disease progression stages can impact the performance of deep CNN models.
- ❖ High precision and automated processing are the advantages for huge datasets and computational resources could be a drawback.
- ❖ Using automated monitoring systems, disease detection techniques can be implemented in real-time, facilitating early intervention and effective conservation measures.
- ❖ Yes, machine learning models can be adapted to recognize diseases in other crops or plant species with appropriate retraining and adjustments.
- ❖ Further crop production, less pesticide use, better environmental conservation techniques, and better biodiversity management in agricultural ecosystems can all result from efficient disease diagnosis.

1.6 Project Management and Finance

- Management of project techniques for organizing, planning, and supervising the identification of rice and tea leaves.
- Managing and optimizing project resources, such as personnel, materials, and financial resources, to ensure the project's success.
- Timely completion of projects is ensured by effective project management.
- Risk assessment and mitigation to guarantee project success within the allocated budget and minimize disruptions.
- Financial planning and budget management to maximize project cost-effectiveness and resource usage.

1.7 Report Layout

Chapter 1: This section discusses the project's significance. A brief description of the project's goals, parameters, and scope is also provided here.

Chapter 2: Reviewing and evaluating the literature on leaf identification, CV, and DL. A Synopsis of Modern Recognition Methodologies.

Chapter 3: A comprehensive description of the process used to collect the data. Actions performed before using the dataset.

Chapter 4: Evaluating the suggested approach in light of the existing situation.

Evaluation of the positive, negative, and ugly in relation to the current situation and potential future directions.

Analysis of the good, the bad, and the ugly in terms of where things stand and where they could go.

Chapter 5: The impacts of the suggested method on sustainability and the environment are carefully taken into account.

Chapter 6: The suggested approach's potential applications and drawbacks are examined.

CHAPTER 2
BACKGROUND

2.1 Terminologies

The use of advanced deep convolutional neural networks (CNNs) in conjunction with other computer vision techniques is used to precisely identify and classify numerous illnesses affecting Bangladesh rice and tea leaf plants based on observable characteristics. As part of this multidisciplinary investigation, novel algorithmic and methodological frameworks are being developed with the goal of improving our capacity to decipher and comprehend images pertaining to diseases of rice and tea leaves. The deep CNN models are trained and tested on an extensive dataset that includes a variety of instances of leaf diseases that affect tea and rice, as well as a range of visual traits related to these crops. The percentage of properly identified cases compared to the total number of samples in the dataset is used to calculate the accuracy of illness identification. The ability of the trained deep CNN models to correctly categorize diseases in rice and tea leaf species that were not included in the training dataset confirms that the models are effective at differentiating between diseases that have not been detected. The training dataset provides evidence of this capacity's validation. Additionally, a methodical approach is used to recognize, evaluate, and put into practice techniques that reduce and manage any risks and uncertainties that can make the rice and tea leaf disease identification project more difficult to complete. In order to successfully complete the project and recognize diseases in the context of rice and tea leaf farming, it is imperative that these risks be addressed.

2.2 Related works

This paper [1] use CNN models for automatically detecting and classifying diseases in tea leaves. They used a dataset of 860 images consisted of 8 classes, which included both healthy plants and plants with various diseases like Algal leaf spot, Gray blight, White spot, Brown blight, Red scab, Bud blight, and Grey blight. The authors trained and tested the model using 80% and 20% of the dataset, respectively. The model achieved an accuracy of 94.45% in identifying the tea leaf diseases. The authors suggest that their proposed model can be further improved by varying the number of layers and applying the transfer learning concept.

In this document [2] involves an integrated approach using IoT and machine learning for the early prediction and broadcasting of tea leaf disease. The process includes the collection of data on temperature and humidity through IoT devices, capturing images of tea leaves using a camera, and analyzing the images using image processing techniques. Then network-based classifier is applied to identify and classify the diseases based on morphological features such as shape, size, and color of the leaves. The data collected from IoT devices but there is no explicit mention about number of data. This approach aims to provide a cost-effective and efficient system for detecting and managing tea leaf diseases.

The authors [3] used a deep neural network model to detect tea leaf diseases. They used a specific architecture for the classifier, which consisted of several convolutional and pooling layers as well as fully connected layers. The authors created a labeled dataset of 5867 tea leaf images and achieved a high accuracy of 96.56% in disease detection. They compared the performance of their model with other existing methods and found that their model outperformed them.

Discussion about the use of deep feature-based identification [4] of rice leaf diseases using SVM. The dataset used in this study consists of 5932 images collected from the agricultural sites of Sambalpur and Bargarh districts in Odisha, India, of four types of rice leaf diseases, namely Bacterial Blight, Blast, Brown Spot, and Tungro. For classification, the authors employed various deep convolutional neural network models, including AlexNet, VGG16, VGG19, GoogleNet, ResNet18, ResNet50, ResNet101, InceptionV3, InceptionResNetV2, DenseNet201, and XceptionNet. They extracted deep features from specific layers of these models, such as fc6, fc7, and fc8, and used them for classification using SVM. They also applied transfer learning to these models. The accuracy of the deep feature plus SVM approach for identifying rice leaf diseases using ResNet50 as the deep feature model is 98.38%.

The article [5] employs an advanced deep convolutional neural network (CNN) known as LeNet to accurately categorize 80 different photos of tea leaf illnesses, such as blister blight, red scab, red leaf spot, and leaf blight. The CNN classifiers were assessed for accuracy using ROC analysis, which revealed good levels of accuracy across several disorders. The accuracy of identifying blister blight illness, for example, was 85%. Despite not achieving 100% accuracy, the LeNet model demonstrates good results, indicating potential for future enhancements. The Table provides the specific classification accuracy of the LeNet algorithm for various illnesses.

The publication explores the application of multi-task deep transfer learning in the identification of illnesses affecting rice and wheat leaves. The experimental findings demonstrate that the enhanced VGG16 model, coupled with the multi-task transfer learning approach, achieves precise identification of illnesses in rice and wheat leaves. The accuracy rate for rice is at 97.22%, while for wheat it reaches 98.75%. The multi-task model has superior performance compared to single-task learning approaches in terms of accuracy, with a 0.43% increase in accuracy rate. The dataset included in the studies comprises photos obtained from both field sources and publicly available photographs from Google. The enhanced VGG16 model surpasses traditional networks such as Resnet50 and Densenet121 in terms of accuracy and generalization skills.

The work explores the application of convolutional neural networks (CNNs) [7] in the prediction of rice leaf diseases. Two models were employed: a basic Convolutional Neural Network (CNN) and the InceptionResNetV2 model with transfer learning. The basic Convolutional Neural Network (CNN) attained an accuracy of 84.75% after 15 epochs, but the InceptionResNetV2 model achieved a higher accuracy of 95.67% after just 10 epochs. The InceptionResNetV2 model, when used with transfer learning, had the highest performance among all models, with an accuracy rate of 95.67%. The dataset utilized was obtained from Kaggle and Google pictures.

Image processing and feature extraction play a vital role [8] in the identification of diseases in tea leaves. The procedure includes the manipulation of pictures by cropping, scaling, and transforming them into threshold values. The collected characteristics are subsequently transmitted to the Artificial Neural Network Ensemble (NNE), which is trained via an ANN model. Negative correlation learning (NCL) is a method employed in neural network ensembles (NNE). The NCL model, which incorporates 10 neural networks with feature extraction, demonstrates the highest performance, obtaining an accuracy rate of 91%. The study utilized a dataset including 50 predetermined pictures.

A deep convolutional neural network (CNN) equipped with a multiscale feature extraction module [9] was employed to detect sick tea leaves. The "CIFAR10-quick model" has been enhanced with the "Alter-second model", resulting in an average identification accuracy of 92.5%. This accuracy surpasses that of typical machine learning methods and classical deep learning approaches. The technology eliminates the need for manual feature extraction and is capable of automatically extracting crucial information from illness photos. A total of 36 photos were processed and enhanced to train the deep Convolutional Neural Network (CNN) model.

2.3 Comparative Analysis and Summary

TABLE I. COMPARATIVE ANALYSIS

Work	**Algorithm**	**Accuracy**
Image analysis and detection of tea leaf disease using deep learning[5]	CNN	85%
Tea leaf diseases recognition using neural network ensemble[8]	NNE	91%
Identification of tea leaf diseases by using an improved deep convolutional neural network[9]	CNN	85%
Bangladeshi Rice and Tea leaf disease detection using Machine Learning	MobileNetV3Large	93.20%

- One common use of deep convolutional neural network models was image classification.
- Several preprocessing techniques were applied, such as image augmentation and normalization.
- It looked into whether implementing previously taught models in transfer learning may improve accuracy.
- Studies have demonstrated that greater variety and size of datasets lead to improved accuracy rates.
- The availability of an extensive dataset of huge photo images is helpful in building a strong model.
- Deep CNN-based algorithms have consistently surpassed more traditional human processes in terms of performance.
- Computerized identification systems have demonstrated significant advancements in productivity and accuracy.
- Various leaf morphological differences, lighting circumstances, and leaf deformations during growth have been recognized as potential barriers.
- One must make use of more sophisticated strategies, including domain adaptation and robust model design, to get beyond these challenges.

2.4 Scope of the Problem

- A deep convolutional neural network (CNN) model specifically designed for rice and tea leaf diseases, with the aim of providing precise detection and classification.
- Using preprocessing methods to improve the CNN model's illness identification accuracy and robustness.
- Analysing how dataset diversity and size affect the rice and tea leaf plant disease detection system's effectiveness.
- Using domain adaptation techniques to enhance the CNN model's capacity to identify and generalise illnesses across various alterations in rice and tea leaf crops.
- Evaluation of the system's functionality in different lighting scenarios and the impact of leaf deformations in rice and tea plants during growth and ripening phases.

- Examining the model's scalability and suitability for identifying illnesses in plant species other than rice and tea leaf crops.
- The implementation of an accreditation scheme aimed at promoting optimal agricultural and environmental practices for disease control in the growing of rice and tea leaves.
- This study examines the wider impacts of disease identification on several domains such as agriculture, ecology, and biodiversity, with particular emphasis on the consequences for Bangladesh's rice and tea leaf crops.

2.5 Challenges

- The various nature of rice and tea leaf plants makes it difficult to diagnose diseases due to variations in their exterior characteristics, including shape.
- Accurate disease classification can be difficult since diverse rice and tea leaf cultivars share many visual traits.
- The quality and appearance of photographs of rice and tea leaf diseases can be affected by changes in illumination during image capture, which complicates the recognition process.
- Consistent disease identification becomes more difficult when rice and tea leaf plants undergo deformations and changes in appearance during the growth stages.
- Training the disease recognition model is hampered by the scarcity of annotated datasets that are specifically designed for the variety of rice and tea leaf species that are unique to Bangladesh.
- Significant computational power is required for the resource-intensive training and fine-tuning of deep convolutional neural networks (CNNs) for the detection of rice and tea leaf diseases.
- A key factor in the actual application of disease recognition is navigating the potential trade-off between model complexity and real-time efficiency.
- For the recognition system to operate reliably, pattern recognition software must be adjusted to different environmental settings and picture capturing hardware.
- Accurate disease identification in rice and tea leaf crops depends on the system's ability to withstand picture noise, occlusions, and other quality issues.

CHAPTER 3

RESEARCH METHODOLOGY

3.1 Research Subject and Instrument

The focus of this research is the detection of diseases in rice and tea leaves in Bangladesh using machine learning techniques. The objective of this project is to develop a model capable of accurately identifying and classifying different types of infected rice or tea leaf based on a given collection of input photographs.

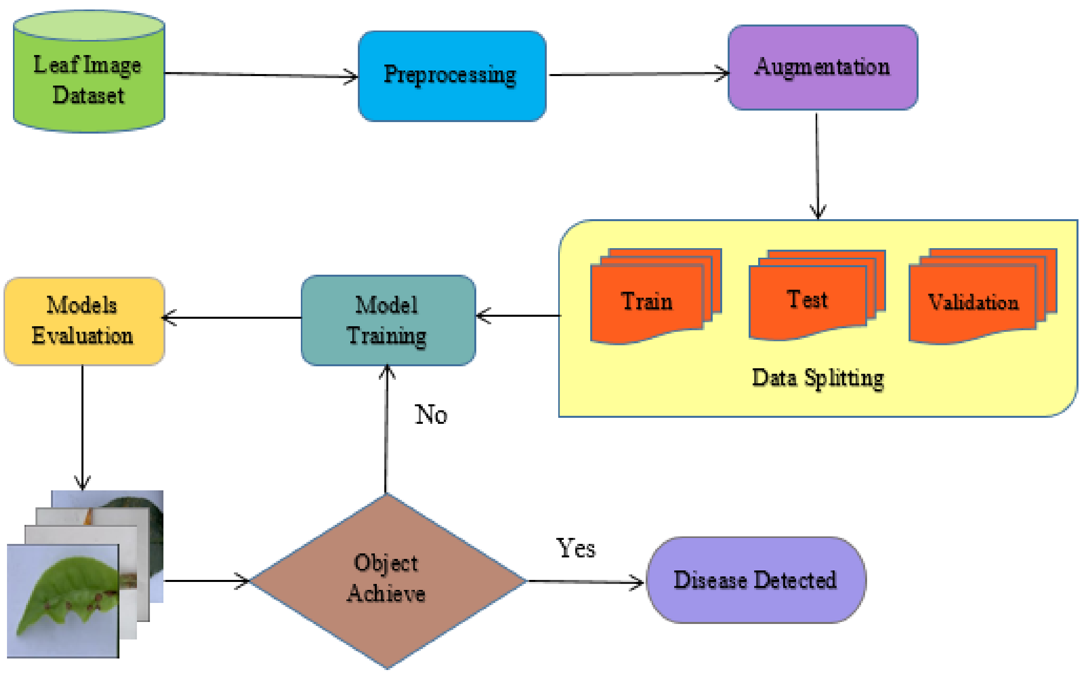


Figure 1: Methodology Diagram

3.2 Data Collection Procedure

Data collecting was carried out at Nurjahan Tea Estate and Finlay Tea Estate in Sreemangal to collect tea diseased leaf. A vast dataset was collected by employing my phone camera, recording intricate information of the plants and their illnesses. In order to achieve consistency and clarity, a standardized method was employed which entailed utilising several angles and viewpoints against a white background. On the other hand, [ ] rice dataset is collected from an online because the dataset is submitted in [10] “Dhan Somadhan”. The original dataset consisted of around 1000 unprocessed photos of rice and tea leaves, which represented a wide range of situations encountered in the field. Afterwards, the dataset experienced a critical period of data augmentation. This procedure entailed the stochastic alteration of preexisting pictures, integrating modifications such as rotations, translations for my dataset. The purpose of the augmentation procedure was to enhance the model's resilience, allowing it to successfully adapt to unanticipated alterations in data. Due to this augmentation process, the dataset significantly increased, reaching a total of 2540 photos. The enhanced dataset encompasses a comprehensive depiction of the diverse range of variances found in real-world situations, hence amplifying the model's ability to effectively recognize and categorize illnesses.

Figure 2 displays a typical portion of the expanded dataset, demonstrating the variety and intricacy of the pictures used to train the model. The careful and thorough process of gathering and enhancing data establishes a strong basis for a resilient and flexible model, capable of addressing the difficulties presented by diseases affecting tea and rice in various agricultural environments.

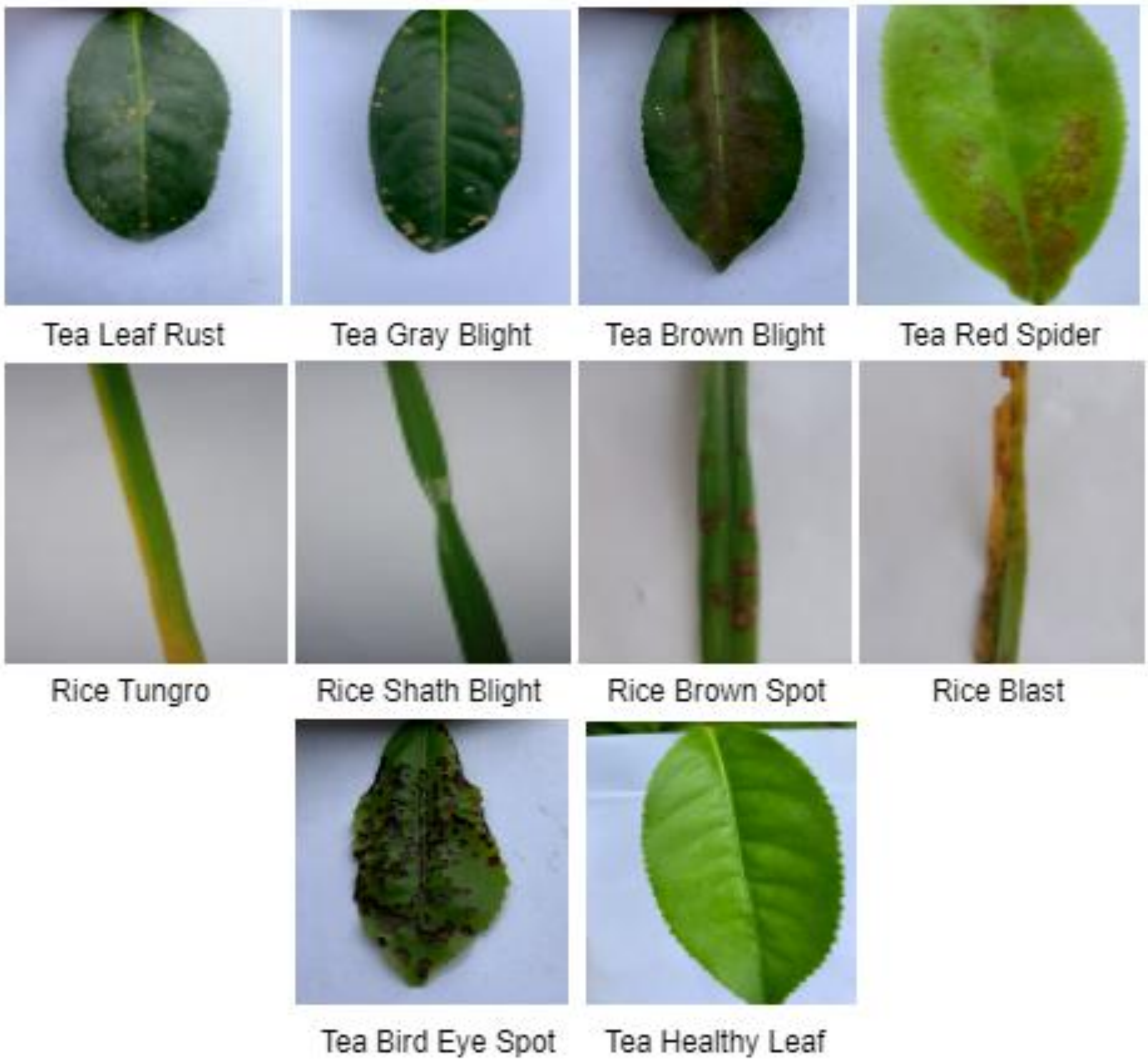


Figure 2: Sample of the dataset.

3.3 Data Augmentation

I improved my train, test, and validation data by applying a rotation of 30 degrees and modifying the width, height shift range to 0.2, and shear range to 0.15. Additionally, I adjusted the zoom range to 0.15.

Data augmentation may enhance the accuracy and resilience of a classifier by transforming the training data. Data augmentation has significantly contributed to achieving state-of-the-art outcomes in many disease detection tasks. Its primary purpose is to improve the generalization capability of the models by including new and remarkable data. The image is rescaled using horizontal flips, brightness shifts, and random RGB color adjustments. Prior to any further processing, it is necessary to perform a multiplication operation on the data, resulting in a rescaled value known as the rescaling value. The model operates in RGB color mode, with RGB coefficients ranging from 0 to 255. However, in this particular circumstance, these values may be excessively high for processing the model. I rescale the image by dividing each pixel value by 255, and then apply a horizontal flip.

3.4 Data Pre-processing

The data has been acquired via an online journal and, with a portion being gathered manually from the field using smartphone. The photographs obtained varied in terms of size and resolution. During the process of training and testing the dataset, I encountered significant challenges and saw several inaccuracies. The current iteration of the dataset has images with a consistent resolution. As per the specifications of my assignment, I have transformed all the photos into a square shape. I standardized the resolution of the photos to 224 × 224. Initially, I performed picture cropping to eliminate any undesired things included in the photographs and lessen the size of the images. Subsequently, I resized it to the specified dimensions for the cropped photos, namely 224 × 224. Furthermore, I conducted training for my models using the RGB color mode.

3.5 Data Separation

Following the completion of the data augmentation step, the pre-processed dataset was segmented into a training set, a validation set, and a testing set. The deep CNN model was produced by employing the training set, which needed the use of transfer learning techniques in order to fine-tune a CNN model that has been trained in the past. Transfer learning may be illustrated by the process of taking a CNN model that has already been trained and modifying the model's final few layers so that it can be utilized for a new job. During the training phase, [11] we kept a close eye on how well the model performed with the validation set and adjusted a number of hyperparameters in order to make it work more effectively. The trained model was then evaluated using data that it had never been exposed to previously throughout the evaluation process. In general, our method consisted of amassing a large number of photographs of leaves, cleaning the data by means of data augmentation, and fine-tuning a pre-trained CNN model by means of transfer learning with the intention of identifying the many kinds of plums that are native to Bangladesh.

3.6 Architecture of the Model

3.6.1 EfficientNetB7:

EfficientNetB7, a constituent of the Efficient Net framework, is built with a distinctive and effective structure. The architecture consists of four convolutional layers, each utilizing common parameters for padding and activation. The model begins with 32 filters for the Conv2D variable and gradually increases to 512 filters in the following layers, which is very noteworthy. The rectified linear unit (ReLU) activation function is constantly utilized to activate every Conv2D layer. The development of EfficientNetB7 incorporates a well-balanced combination of depth, width, and resolution, resulting in an optimized and resource-efficient model that is very effective for a range of computer vision tasks.

3.6.2 MobileNetV3Large:

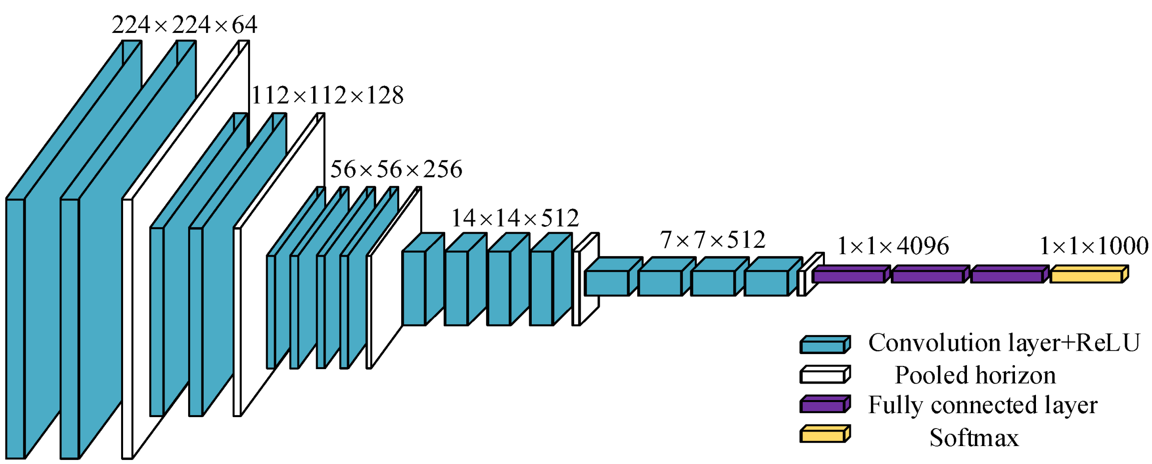
MobileNetV3Large, a member of the MobileNet family, is specifically designed to deliver exceptional performance in situations with limited resources. The model has a unique architecture that includes shared padding and activation parameters in its four convolutional layers. It starts with a Conv2D variable of 32 and ends with 256 filters. The Rectified Linear Unit (ReLU) activation function is consistently utilised for every Conv2D layer. The architecture of MobileNetV3Large focuses on inverted residuals and linear bottlenecks, achieving a trade-off between accuracy and efficiency. This makes it highly suitable for real-time applications on mobile and edge devices.

3.6.3 ResNet101:

ResNet101 is a convolutional neural network that use skip connections to address the issue of disappearing gradients by employing residual learning. By including parallel convolutional layers, this model effectively captures multi-scale properties that are essential for accurate picture categorization. The combination of batch normalization and regularization improves the efficiency of training and the ability to generalize. The novel design of this architecture guarantees strong performance in computer vision tasks.

Dense Architecture:

The base layer of neurons in the brain is referred to as the Dense Layer due to the fact that all the neurons in the layer above transmit information to it. The output generated by the convolutional layers is utilized by the Dense Layer to categories the processed pictures. The intricate structure of each model is meticulously depicted in a highly comprehensive manner in Table I. Every CNN model possesses an identical dense organizational structure.



[Source]

Figure 3: Model Architecture. [12]

TABLE II. ALL MODEL’S DENSE ARCHITECTURE

Layer Name	**Node numbers**	**Activation**
Avg. Pooling	512	
Dense	256	ReLU
Dense	256	ReLU
Dense (Output)	10	Softmax

3.7 Statistical analysis

A Comparative Analysis of CNN and Transfer Learning Models:

Evaluating the Efficiency of the CNN Model and Diverse Transfer Learning Models Statistical tests like t-tests or ANOVA can be employed to compare the performance of the CNN model with different transfer learning methods. One may assess the accuracy, precision, recall, and F1 score of each model to see if there are statistically significant differences between them.

Confusion Matrix:

The confusion matrix is a useful tool for analyzing the misclassifications made by models. This inquiry has the ability to elucidate the categories of jujube spices that provide challenges for the models and aid in identifying areas that require enhancements.

3.8 Applied Mechanism

- Compile a dataset of images of tea and rice leaf from different sources like field collection and online journal, and then separate the dataset into sets for testing, validation, and training.

- Use the proper techniques to add data, resize, and standardize the pictures as part of the first processing.
- Select some CNN architectures (such as EfficientNet, ResNet, or MobileNet) that are suitable for the purpose of identifying diseased rice or tea leaf.
- Using CNN models that have already been pre-trained (e.g. using the pre-trained weights from ImageNet) can facilitate transfer learning.
- Using the gathered dataset, modify the pre-trained model's parameters to make it more appropriate for the task at hand.
- Utilizing the provided training dataset, train the CNN and transfer learning models.
- Perform a performance analysis of the models using metrics like F1 score, accuracy, precision, confusion matrix and recall.
- Perform a confusion matrix-based performance study on the models to identify instances of inaccurate categorization and possible trouble spots.
- Compare and contrast the CNN model's with the transfer learning model's performance using statistical analysis.
- Analyze the information, talk about how effective the technique was, and make recommendations for potential areas for more study or development.

3.9 Implementation Requirements

Initially, a robust computer with a graphics processing unit (GPU) or a Google colab with a T4 GPU and ample storage capacity in disk or Google drive is essential to accommodate both the dataset and the experiment's outcomes. Python, along with other essential libraries and frameworks like TensorFlow, Keras, and PyTorch, is indispensable for software development in the field of deep learning. Image processing software, like OpenCV, is essential for both preprocessing and enhancing data. The research necessitates a thorough and precisely annotated collection of rice and tea leaf images that are classified based on their spice kind. Resizing, normalisation, and data augmentation enhance the quality of the dataset and improve the generalisation of the model. Developing the model requires a comprehensive grasp of CNN architectures, deep learning frameworks, and transfer learning. Utilizing pre-trained models, such as

MobileNetV3Large, ResNet101 or EfficientNetB7 which have been trained on ImageNet, expedites the process. Training and evaluation are necessary when dealing with sets for training, validating, and testing. Optimal model performance may be achieved by the implementation of efficient training techniques, which may involve adjusting hyper parameters. An analysis is conducted on the F1 scores of the models, as well as their accuracy, precision, and recall. Performance analysis include the utilisation of statistics and visualizations. The confusion matrix displays both misclassifications and the opportunity for improvement. Documentation and report writing are essential requirements for doing research. The procedure is recorded in the source documentation, notebooks, and experiment logs. The utilization of visualizations and graphs enhances both the presentation and analytical processes. Researchers might enhance the clarity and comprehensibility of their findings by utilizing word processing or Latex technologies.

CHAPTER 4

EXPERIMENTAL RESULT

4.1 Experimental setup

Collect diverse images of tea and rice leaf in different varieties and situations. Preserve the proportional relationship and adjust the size of the pictures to match the dimensions of the models. Applying rotation, scaling, and flipping operations can enhance the diversity of a dataset and improve the generalization of a model. Given the constraints of model complexity and processing resources, it is recommended to employ a Convolutional Neural Network (CNN) architecture for the purpose of recognizing the leaf. Discover pre-trained convolutional neural network models such as EfficientNetB7, ResNet101 and MobileNetV3Large which have been trained on the ImageNet dataset. Initialize the CNN model by utilizing either randomly generated weights or weights that have been pre-trained. Examine performance metrics such as loss and accuracy on the validation set to track the progress of the training process. Assess the performance of the trained models using the independent testing set for identifying diseased leaf. Assess the performance of models by measuring accuracy, precision, recall, and F1 score. In order to assess the ability of a model to generalize and maintain consistency, it is recommended to employ cross-validation techniques such as k-fold. Evaluate the strengths and limitations of the models by examining the experimental outcomes. Conduct a comparative analysis of CNN and transfer learning models and evaluate their effectiveness.

4.2 Experimental Results & Analysis

TABLE III. MODEL'S ACCURACY REPORT

Model Name	Validation Accuracy (%)	Accuracy (%)	Training Time (sec)
MobileNetV3Large	94.74	93.20	373.18
EfficientNetB7	91.67	88.40	408.86
ResNet101	89.91	86.80	385.45

Table-III displays the performance data of three distinct models—MobileNetV3Large, EfficientNetB7, and ResNet101—according to validation accuracy, test accuracy, and training duration. The MobileNetV3Large model achieves the greatest validation accuracy of 94.74%, followed by the EfficientNetB7 model at 91.67%, and the ResNet101 model at 89.91%. Regarding test accuracy, MobileNetV3Large has the highest score of 93.20%, followed by EfficientNetB7 with 88.40%, and ResNet101 with 86.80%. The duration of training varies across different models, with EfficientNetB7 requiring the most time at 408.86 seconds, followed by ResNet101 at 385.45 seconds and MobileNetV3Large at 373.18 seconds.

4.3 Discussion

According to the accuracy table, model 1 which is MobileNetV3Large outperforms the other models. Consequently, we have determined that MobileNetV3Large will be our optimal choice. Table – IV summarizes the extensive assessment indicators of a classification model, providing insights into its performance across various labels. The precision, recall, and F1-score are painstakingly documented for each class, providing insight into the model's capacity to accurately detect and categorize cases. The model achieves flawless scores in Tea Bird Eye Spot, Tea Brown Blight, Tea Healthy Leaf, and Tea Red Spider, demonstrating remarkably high precision and recall values across these categories.

The model's proficiency in accurately predicting labels is indicated by its total accuracy of 93.20%. The use of macro and weighted averages reinforces a well-rounded and strong performance, highlighting the model's dependability in effectively managing many categories within the dataset.

TABLE IV. BEST MODEL CLASSIFICATION REPORT

Label name	**Precision**	**Recall**	**F1-Score**	**Support**
Rice Blast	0.94	0.91	0.93	68
Rice Brown Spot	0.87	0.87	0.87	30
Rice Shath Blight	0.91	0.98	0.94	49
Rice Tungro	0.94	0.89	0.91	36
Tea Bird Eye Spot	1.00	1.00	1.00	13
Tea Brown Blight	1.00	1.00	1.00	6
Tea Gray Blight	0.87	1.00	0.93	13
Tea Healthy Leaf	1.00	1.00	1.00	13
Tea Leaf Rust	1.00	0.87	0.93	15
Tea Red Spider	1.00	1.00	1.00	7
Accuracy			0.93	250
Macro avg	0.95	0.95	0.95	250
Weighted avg	0.93	0.93	0.93	250

4.4 Model Evaluation

On the completion of 10 epochs, the models employed in this study, including MobileNetV3Large achieved an accuracy of 94.74% for the validation dataset and attained a 93.20% accuracy for the testing dataset. Upon completion of the training and testing sessions, the work achieves a consistent level of correctness. Based on the obtained accuracy and confusion matrix, it can be concluded that the proposed Deep Learning model is suitable for recognizing infected rice and tea leaves.

MobileNetV3Large's training and validation loss are shown in Figure 6.The graph is error-free. This location generates a street line, signifying that the two lines are intersecting. It has been claimed that there is no instance of data loss.

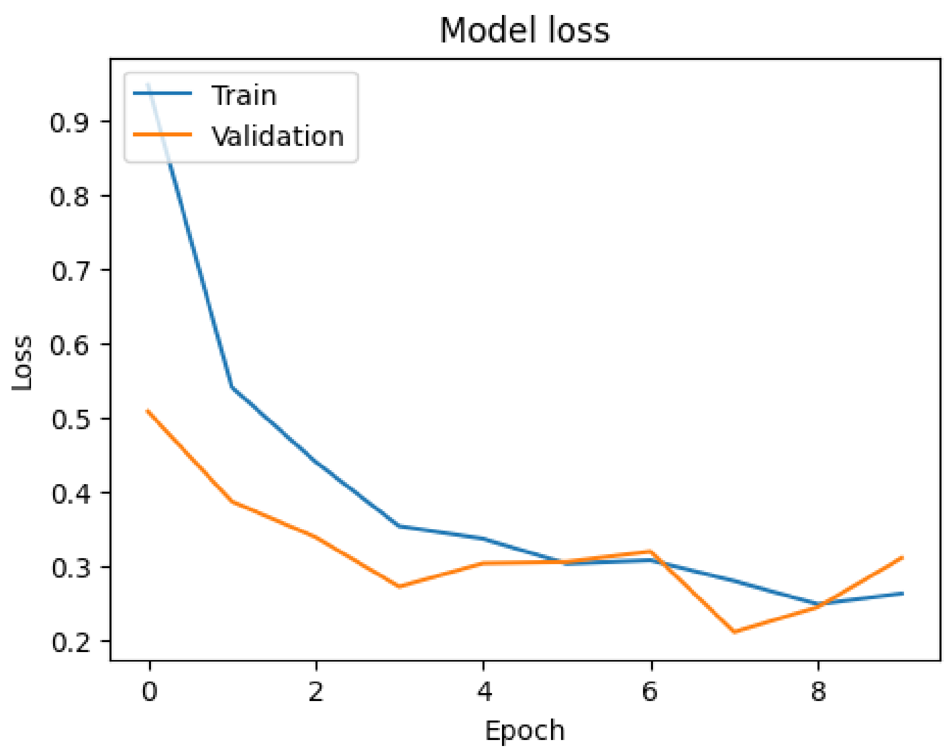


Figure 4: Training vs Validation loss of MobileNetV3Large.

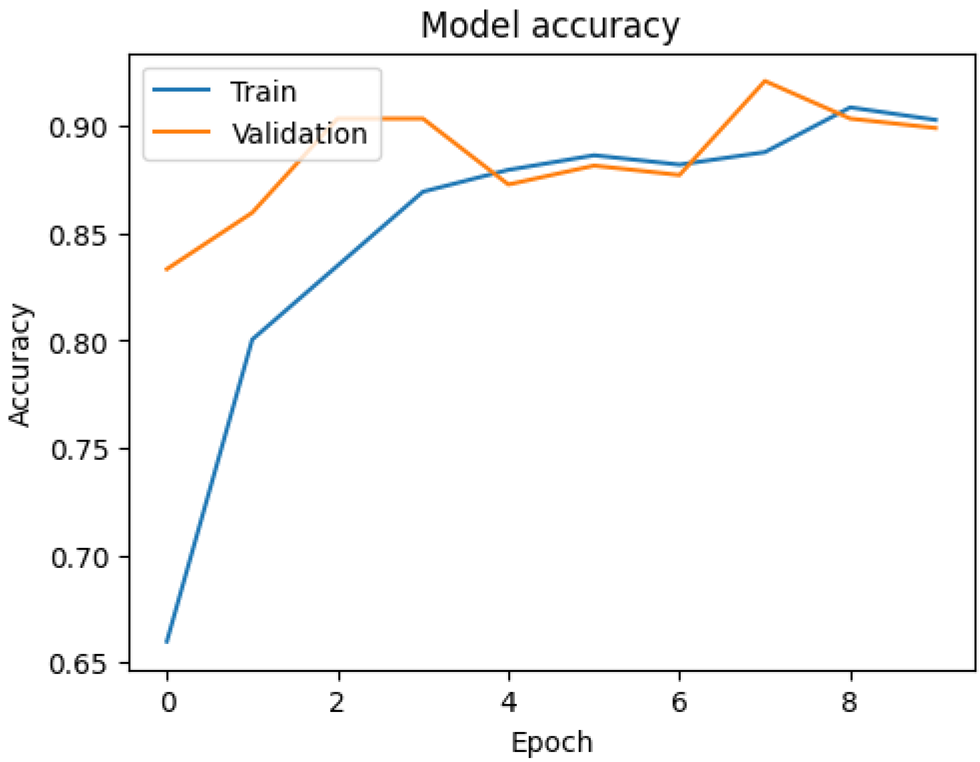


Figure 5: Training vs Validation Accuracy of MobileNetV3Large.

Figure 7 illustrates the disparity between training accuracy and validation accuracy for MobileNetV3Large. This graph exhibits exemplary behavior, properly depicting the absence of mistakes with both lines perfectly overlapping each other. This suggests that this particular model has undergone exceptional training.

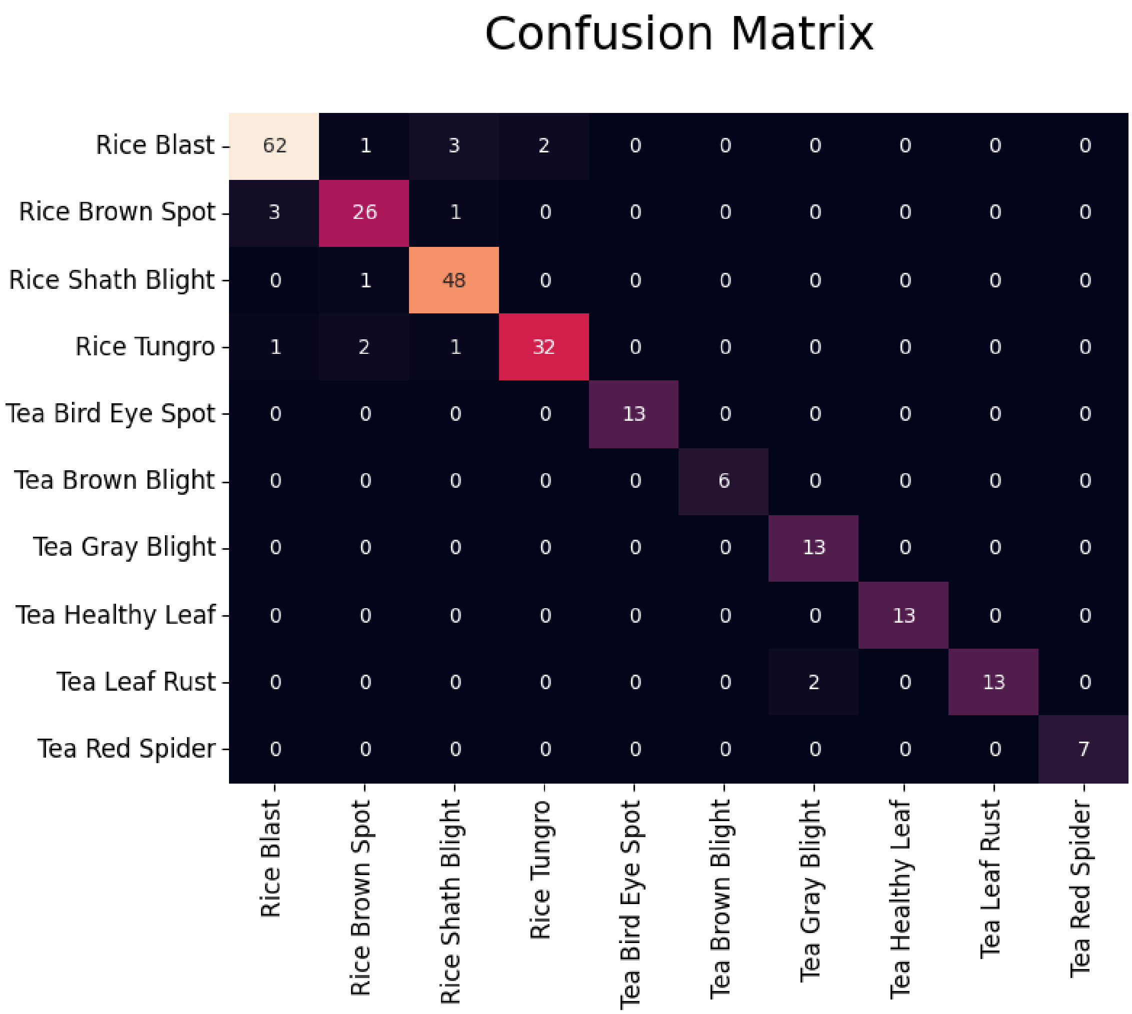


Figure 6: Confusion Matrix of MobileNetV3Large.

CHAPTER 5

Impact on Society, Environment, and Sustainability

5.1 Impact on Society

- Enhanced agricultural productivity due to enhanced agricultural techniques.
- The efficient utilization of existing resources in agricultural production, leading to reduced wastage and increased income levels.
- Ensuring the preservation of plant and animal species, as well as the conservation of genetic diversity.
- Enhancements in ecosystem management and the use of precise data in land-use planning.
- Promoting eco-conscious behavior and sustainable agricultural techniques.
- The adoption of advanced technology in agricultural production and the integration of traditional and modern farming methods.
- The dissemination of knowledge and the establishment of collaborative relationships among agricultural institutions, researchers, and farmers.
- The creation of new business opportunities through targeted consumer markets and goods that offer additional value.
- Promotion of education and dissemination of knowledge on the significance of preserving plant biodiversity and implementing ecologically sustainable agricultural practices.
- Advancements in the domains of deep learning, computer vision, and agricultural research within the scientific community.

5.2 Impact on Environment

- The conservation of biodiversity by accurately categorizing and safeguarding a diverse range of diseased rice and tea leaf.
- The reduction in the use of chemical inputs such as pesticides and fertilizers accomplished by using certain agricultural techniques.

- Enhancements in the administration of ecosystems and protection of natural areas.
- The objective is to minimize environmental pollution and soil degradation by wisely using the resources at hand.
- The promotion of ecologically conscious farming methods that prioritize environmental preservation.
- Providing support for the restoration of habitats and the maintenance of ecological balance.
- Contributing to the conservation of pollinators and other useful insects.
- More efficient irrigation systems lead to reduced water use overall.
- The use of effective management strategies to prevent the spread of infectious illnesses and undesirable organisms.
- The advocacy of agricultural methods that have a reduced impact on the environment and the use of land management techniques that prioritize conservation.

5.3 Ethical Aspects

The effort to identify diseased leaf is driven by ethical reasons. Data utilization, algorithmic equity, and transparency are all constituents of ethical conduct. Strict respect to standards of data privacy and informed authorization is necessary for the ethical collecting and use of leaf photos. To eradicate occurrences of prejudice and discrimination, it is imperative that algorithms and models exhibit objectivity and be subject to thorough examination. The programmer seeks to minimize waste while concurrently advocating for sustainable methods concerning computational power, energy use, and physical resources. The ecological impact of the identification system is comprehensively assessed to guarantee that it will not jeopardize the integrity of any natural habitats or ecosystems. The discussion revolves around three main subjects: technological accessibility, empowerment of farmers, and the promotion of ecologically sustainable agriculture. Prudent implementation and applications ensure avoidance of misuse and breaches of privacy. Decision-making procedures incorporate ethical considerations, and stakeholders are urged to examine the established pros and cons of the system in issue. The diseased rice and tea leaf identification project follows specific ethical principles in both its research and execution. Specific criteria guarantee the project's adherence to principles of impartiality and transparency, its long-term viability, and its commitment to safeguarding both human rights and the environment.

5.4 Sustainability Plan

Sustainable ecological management is the diligent governance of resources with the aim of mitigating adverse effects on the environment. Adoption of environmentally friendly farming methods and less reliance on chemical inputs. Support for efforts to protect biodiversity and maintain ecological equilibrium. Enhancing agricultural productivity and efficiency is crucial for ensuring a nation's economic viability. Optimized resource allocation and enhanced agricultural procedures to achieve maximum efficiency. The establishment of economic opportunities coupled with the facilitation of access to specialist market segments. Empowering farmers and providing assistance to local communities are two components of social equity. The widespread distribution of technology, together with the sharing of knowledge and the enhancement of capacities. Taking into account the varied demands and viewpoints of stakeholders. Sharing research findings, methodology, and best practices is essential for promoting the exchange of technology and knowledge. This work has the capacity to benefit several domains, including deep learning, computer vision, and agricultural research. Support is provided for learning, cooperation, and the promotion of future advancements. Long-term monitoring and fine-tuning: The recognition system will undergo ongoing assessment to improve its performance. Continuous improvement fueled by user feedback and flexibility to accommodate evolving needs. Ensuring the perpetual pertinence and durability of the system. Collaboration and partnership development entail the active engagement of local communities, agricultural institutions, and academic institutes. Utilizing a wide range of knowledge and resources through collaborative efforts. Promoting a sense of accountability and improving the overall efficiency of the project.

CHAPTER 6

SUMMARY, CONCLUSION, IMPLICATION FOR FUTURE RESEARCH

6.1 Summary of the Study

The study primarily employed deep convolutional neural networks (CNNs) and other computer vision techniques to identify several kinds of rice and tea leaf disease from Bangladesh. The research started by conducting a comprehensive examination of the pertinent literature, focusing on the methodology, dataset requirements and collection, performance comparisons and challenges related with this field of study. The study aimed to create a reliable Convolutional Neural Network (CNN) model capable of accurately identifying distinct disease of plants rice and tea by analyzing their leaves. This was achieved by using a raw dataset of various disease. The study's findings indicate that achieving high accuracy rates necessitates considering crucial factors such as dataset size, preprocessing techniques, and illumination conditions. The comparative analysis revealed the superiority of deep Convolutional Neural Network (CNN) models over conventional methods. The focus of the study was primarily on the practical implementation of these models in the domains of agriculture and biodiversity conservation. The research encompassed crucial aspects such as ethics, sustainability, and the effects on society and the environment. The results advocate for ecologically sustainable farming methods, advance the progress of computerized disease identification, and provide practical recommendations for implementation. The research aims to enhance agricultural production, promote ecological conservation, and effectively manage biodiversity in Bangladesh, while also addressing the ethical and sustainable aspects associated with these concerns.

6.2 Conclusion

Three different models were used to identify diseases in rice and tea leaf crops in Bangladesh. Among these models, MobileNetV3Large demonstrated the highest level of accuracy, obtaining an impressive accuracy rate of 93.20%. The effective utilization of deep convolutional neural networks (CNNs) highlights their ability to precisely categorize illnesses. This study emphasizes the potential of deep learning and computer vision in identifying diseases, specifically in the agricultural setting. The impressive accuracy of 93.20% achieved by MobileNetV3Large demonstrates its reliability for practical use, particularly in improving crop management and decision-making procedures. The incorporation of a dataset consisting of 2540 photos underscores the significance of dataset magnitude and variety in effectively training precise models. Preprocessing methods and the careful evaluation of lighting conditions are crucial in improving the illness detection system. High-resolution plant photos obtained using digital cameras, scanners, or drones are preferred more and image preprocessing enhances picture quality and reduces noise, while image segmentation divides the image into segments to isolate the diseased area [11]. In addition to its applications in agriculture, this technology also plays a significant role in ecological research, conservation initiatives and the monitoring of biodiversity. Although the current accuracy is impressive, continuous study and enhancements are necessary to improve the illness detection system's effectiveness and adaptability in various environmental situations. This includes focusing on domain adaptation, constructing strong models, and ensuring scalability.

6.3 Implication for Future Research

- The investigation of more complex strategies for domain adaptation, with the goal of improving the recognition system's generalization capabilities across a wide range of environmental variables and variations.
- Research into ensemble learning methods, which include mixing several different models in order to improve plum species recognition in terms of both its accuracy and its resilience.
- The use of new data modalities, such as spectrum or hyperspectral imaging, in order to extract features that are more complete and hence improve species classification.

- Investigating whether or not the created recognition system can be applied to additional plant species, hence broadening its scope of applicability to include broader agricultural and ecological settings.
- Taking into consideration the possibility of real-time development and deployment of the recognition system in field settings, with an emphasis on overcoming problems relating to computing efficiency and hardware requirements.

REFERENCES

[1]. Latha, R. S., et al. "Automatic detection of tea leaf diseases using deep convolution neural network." 2021 International Conference on Computer Communication and Informatics (ICCCI). IEEE, 2021. [Google Scholar]

[2]. Yashodha, G., and D. Shalini. "An integrated approach for predicting and broadcasting tea leaf disease at early stage using IoT with machine learning–a review." *Materials Today: Proceedings* 37 (2021): 484-488. [Google Scholar]

[3]. Datta, Saikat, and Nitin Gupta. "A novel approach for the detection of tea leaf disease using deep neural network." Procedia Computer Science 218 (2023): 2273-2286. [Google Scholar]

[4]. Sethy, Prabira Kumar, et al. "Deep feature based rice leaf disease identification using support vector machine." Computers and Electronics in Agriculture 175 (2020): 105527. [Google Scholar]

[5]. Gayathri, S., et al. "Image analysis and detection of tea leaf disease using deep learning." 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC). IEEE, 2020. [Google Scholar]

[6]. Jiang, Zhencun, et al. "Recognition of rice leaf diseases and wheat leaf diseases based on multi-task deep transfer learning." Computers and Electronics in Agriculture 186 (2021): 106184. [Google Scholar]

[7]. Krishnamoorthy, N., et al. "Rice leaf diseases prediction using deep neural networks with transfer learning." Environmental Research 198 (2021): 111275. [Google Scholar]

[8]. Karmokar, Bikash Chandra, et al. "Tea leaf diseases recognition using neural network ensemble." International Journal of Computer Applications 114.17 (2015). [Google Scholar]

[9]. Hu Gensheng, Yang Xiaowei, Zhang Yan, Wan Mingzhu "Identification of tea leaf diseases by using an improved deep convolutional neural network" in Elsevier, 2019. [Google Scholar]

[10]. Hossain, Md Fahad; Abujar , Sheikh ; Noori, Sheak Rashed Haider ; Hossain, Syed Akhter (2021), "Dhan-Shomadhan: A Dataset of Rice Leaf Disease Classification for Bangladeshi Local Rice", Mendeley Data, V1, doi: 10.17632/znsxdctwtt.1 [Google]

[11]. Fulari, Utkarsha N., Rajveer K. Shastri, and Anuj N. Fulari. "Leaf disease detection using machine learning." J. Seybold Rep 1533 (2020): 9211. [Google Scholar]

[12]. Wu, Jie, and Xiaoqian Zhang. "Tunnel crack detection method and crack image processing algorithm based on improved retinex and deep learning." Sensors 23.22 (2023): 9140. [Google Scholar]

Plagiarism

Bangladeshi Rice & Tea leaf disease detection

ORIGINALITY REPORT

11%	7%	5%	3%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Daffodil International University Student Paper	2%
2	doctorpenguin.com Internet Source	2%
3	dspace.daffodilvarsity.edu.bd:8080 Internet Source	1%
4	Md. Mamun Sakib, Md. Mehedi Hasan, Rabeya Bibi, Md. Hamidur Rahman, Abdus Sattar. "CNN and Transfer Learning Modeling for Jujube Spices Recognition", 2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT), 2023 Publication	1%
5	Prabira Kumar Sethy, Nalini Kanta Barpanda, Amiya Kumar Rath, Santi Kumari Behera. "Deep feature based rice leaf disease identification using support vector machine", Computers and Electronics in Agriculture, 2020 Publication	1%