Heaven's Light is Our Guide

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

Rajshahi University of Engineering and Technology, Bangladesh

Rice Leaf Diseases Classification Using CNN and Swin Transformer with Transfer Learning

Author

Sajidul Islam Ronju
Roll No. 1803161
Department of Computer Science and Engineering
Rajshahi University of Engineering and Technology

Supervised by

Prof. Dr. Md. Rabiul Islam
Professor
Department of Computer Science and Engineering
Rajshahi University of Engineering and Technology

ACKNOWLEDGEMENT

First and foremost, I want to thank the almighty Allah for giving me the knowledge and ability to research and complete this thesis successfully. Then, I offer my sincere gratitude and indebtness to my thesis supervisor, **Prof. Dr. Md. Rabiul Islam** Professor, Department of Computer Science and Engineering, Rajshahi University of Engineering and Technology, Rajshahi. who has supported me throughout my thesis with his patience and knowledge. I shall ever remain grateful to him for his valuable guidance, advice, encouragement, cordial and amiable contribution to my thesis.

I am also grateful to respected **Prof. Dr. Md. Ali Hossain**, Head of the Department of Computer Science and Engineering and all the respective teachers of Department of Computer Science and Engineering, Rajshahi University of Engineering and Technology, Rajshahi for their valuable suggestions and inspirations from time to time.

Finally, I would like to convey my thanks to my parents, friends, and well-wishers for their true motivations and many helpful aids throughout this work.

January 20, 2025 Sajidul Islam Ronju

RUET, Rajshahi

Heaven's Light is Our Guide

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

Rajshahi University of Engineering and Technology, Bangladesh

CERTIFICATE

*This is to certify that this thesis report entitled **"Rice Leaf Diseases Classification Using CNN and Swin Transformer with Transfer Learning"** submitted by **Sajidul Islam Ronju, Roll:1803161** in partial fulfillment of the requirement for the award of the degree of Bachelor of Science in Department of Computer Science and Engineering of Rajshahi University of Engineering and Technology, Bangladesh is a record of the candidate own work carried out by him under my supervision. This thesis has not been submitted for the award of any other degree.*

Supervisor

____

Prof. Dr. Md. Rabiul Islam
Professor
Department of Computer Science and Engineering
Rajshahi University of Engineering and Technology
Rajshahi-6204

External Examiner

____

Md Rakibul Haque
Lecturer
Department of Computer Science and Engineering
Rajshahi University of Engineering and Technology
Rajshahi-6204

ABSTRACT

Rice is one of the major cultivated crops in Bangladesh which is affected by various diseases at various stages of it's cultivation. However, it is challenging for farmers to accurately identify these diseases due to their limited knowledge. Recent advancements in deep learning have demonstrated the potential of Convolutional Neural Networks (CNN) for automatic image recognition, which can greatly benefit in addressing such issues. As the availability of rice leaf disease image dataset is limited, We have collected a medium-sized dataset which contains 18445 images from the internet and we have conducted our experiment on ten major diseases namely bacterial_leaf_blight, brown_spot, healthy_leaves, leaf_blast, leaf_scald, narrow_brown_spot, neck_blast, rice_hispa, sheath_blight, tungro. We have utilized pre-trained and fine tuning deep learning model. Through conducting experiments with a total of 18,445 images, it was observed that the fine tuning CNN with swin transformer architecture exhibited the highest accuracy in detecting and classifying rice leaf diseases and achieved low training and validation loss compared to other papers . The classification accuracy achieved by the pre-trained VGG 16 and fine tuning CNN with Swin Transformer models were 91.83% and 96.40% respectively.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction

1.1 Introduction

Rice is the staple source of food in Bangladesh as well as across the world. It is attacked by a variety of diseases in various stages of its cultivation. Therefore, early detection and remedy of such diseases are beneficial to ensure high quantity and best quality but this is very difficult due to the huge expanse of land under individual farmers and the huge variety of diseases as well as the occurrence of more than one disease in the same plant. Agricultural expert knowledge is not accessible in remote areas and it is a time taking process [1]. It is important that disease symptoms are identified early and accurately for undertaking disease control actions or remedial measures. The symptoms of rice diseases, such as rice bacterial leaf blight, rice sheath blight and rice blast, appear initially as spots around the infected areas. So the detection of these diseases mainly relies on their spots. At present, farmers or technicians mainly rely on their experiences, the guide books and experts to identify these diseases. It is suitable for detecting some familiar and typical diseases. However, some different diseases can cause similar spots and the same diseases can cause different spots because of different rice varieties and local conditions. It increases the complexity of identifying rice diseases [4]. So with the high amount of production and consumption, it is also very much prone to the various types of plant diseases which will have negative consequences on the quality and quantity of rice production [5]. Therefore, we have come up with an automated system, which will contribute to the development of Bangladesh's agriculture sector. Our proposed system will automatically detect if a leaf is healthy or infected from the leaf's image[6].

1.2 Significance of Disease Detection in Rice Plants

Rice does be a mat staple eat for a big portion of worldy's folks, and bug can seriously impact it yields. The here is reason why spotting disease in rice plant early do be crucial:

1.2.1 Reduced Crop Loss

Illnesses may lead to tremendous harvest failures, varying between 20-40%! based on few researches. Timely recognition permits farmers to take responding actions to halt the extension of the illness and decrease harm.

1.2.2 Improved Food Security

Rice be a crucial source of calorie for billions of individuals. Through decrease crop loss, detect disease can assist in guaranteeing food safety and decrease hunger risk.

1.2.3 More Efficient Use of Resources

When sicknesses avoid being detected, farmers might end up using unnecessary toxicants. Spotting the issue early and precisely permits focused administration of suitable remedies, lessening excess and environmental repercussions.

1.2.4 Increased Farmer Income

Crops that are in good health end up with more harvests and better quality goods that results in more money for farmers at the end of the day [7].

1.3 Overview

Agriculture, being an essential source of earnings and livelihood in many regions, with a lot of croppy things, rice being a kind of prominent choice, particularly in countries of the Asian persuasion, tends to have challenges with various sick issues during the growth stage. These sicknesses can really hit the crop quality and how much you get. Spotting these sick things can get you kinda confused because some of them seem same. Anyway, these machines do a bit of the work spotting which is helpful to stop issues early. Science folks have been using the smarts of machines in some cool ways to make these spotting machines. Farmers kinda struggle with telling one disease from another because they're not super experts. The smart technology has been helping out with fancy image-knowing systems using the CNN model to tackle these issues.

1.4 Overview

The rise and fall of rice in our nation has taken a tumble, no thanks to the rise in leafy rice diseases. Stated in [8], these pesky diseases, often brought on by bacterium, viruses, or funguses, really put a damper on productivity. Utilizing the feisty technology of computer vision for keeping these diseases at bay is usually the easiest and cheapest way to go about it. The global distribution of rice diseases is illustrated. Some of the diseases that can impact rice leaves include Bacterial leaf blight, Brown Spot, Leaf Blast, Leaf Scald, Leaf Tungro, Leaf Ufra, Narrow Brown Leaf Spot, and Sheath Blight. Our research has identified four main rice leaf diseases: bacterial leaf blight, brown spot, leaf blast, and sheath blight, as shown . In Bangladesh, them diseases are most common kind of rice leaf diseases. Bacterium blight be caused by Xanthomonas oryzae pv. Lead to seedlings wiltin', leaves turnin' yellow, and dryin' out [9]. Brown spot, caused by the mushroom, impacts different parts of rice plant like coleoptile, leaves, leaf sheath, panicle branches, glumes, and spikelets, creatin' big spots on leaves and possible death of whole leaf. When infected, the seed may develop unfilled grains, speckled, or discolored seeds [10]. Blasting, caused by the fungus Magnaporthe oryzae, affects aboveground parts of the rice plant including the leaf, collar, node, neck, portions of the panicle, and sometimes the leaf sheath [11]. Rhizoctonia solani is a fungus responsible for causing sheath blight disease. Infected leaves tend to senesce or dry out and perish at a faster rate. Young tillers can be eliminated as a result. The accurate diagnosis of rice leaf disease is currently achievable through convolutional neural network (CNN) technology, which is a component of artificial neural networks. CNN enables us to work with vast amounts of data. Transfer learning primarily focuses on handling large data volumes. Machine learning (ML) research on transfer learning (TL) involves storing information obtained while solving one problem and applying it to another similar yet unrelated problem. Utilizing the newest finetic CNN structure rooted from the Swin Transformer model with our unique dataset was a main part of our research. We also applied a pre-used CNN structure from VGG-16 with transfer learning with our dataset. In conclusion, the Swin Transformer model showed higher accuracy and precision than VGG-16.

1.5 Motivation

Bangladesh is the fourth-largest rice producer. Rice is grown on about 10.5 million hectares. Almost all of the 13 million farm families of Bangladesh grow rice [12]. So, it is essential to produce good quality and quantity of rice. But the main problem is the disease of rice crops. Detection of rice plant disease has always been challenging. It requires constant human observations. But it is much more time consuming and requires more labor. This research presents an efficient model that detects if any of the crops are infected by disease or not. This will make the life of farmers easier, reduce the cost of labor and save time. Our proposed system is less time consuming and efficient.

1.6 Objectives of the Thesis

The agriculture industry now support the majority of countries. The agriculture industry is equally important for a country's economic growth. However as the population grows, the environment is impacted more and more. The main Objectives of this work are as follows:

- Developing a deep learning model for identifying and classifying large number of rice leaf disease images.
- Utilizing deep learning techniques such as convolutional neural network, vision transformer to enhance detection accuracy and handle the variability of disease symptom.
- Improve the efficiency and precision of existing paper model and reduce the computational parameters.

1.7 Challenges of Deep Learning Techniques Without Transfer Learning

Deep learning holds enormous promise, but it is still plagued with significant problems when it comes to preparing models from scratch. These are some of the difficulties:

- hunger: deep learning models need large amounts of labeled data to be effective. This data can be costly and time-consuming to collect, label, and clean. The model may have difficulty generalizing and performing well on unseen samples if there is insufficient data.
- overfitting: it may rely too much on certain features of the underlying data if it is trained solely on the training data. It overfits when the model performs well on the training data but poorly on the new data.
- computing Cost: Significant computing resources, such as strong GPUs and large quantities of memory, are needed to train deep learning models from start. Smaller businesses and individuals may find this to be a hurdle.
- Time Consumption: Deep learning model training may be somewhat time-consuming, particularly for complicated models. This could become a snag for projects with short deadlines.
- Interpretability: Deep learning models are sometimes like ”black boxes,” with little to no explanation for the predictions they provide. This lack of interpretability may be a challenge in situations where it is essential to comprehend the logic behind a choice.

By utilizing pre-trained models and their acquired characteristics, transfer learning aids in addressing a number of these difficulties. This lowers the possibility of overfitting, speeds up training, and enables us to train models with less data [13].

1.8 Thesis Organization

The report is organized into 7 chapters including this chapter: ***Introduction*** where all the related topics are discussed which are needed for understanding the research work. The outline of rest of the works are organized as follows:

Chapter 2

Topic - Background Study

This chapter explores the applications of deep learning methods in detecting and classifying rice leaf diseases. It also covers the advantages of such techniques for reducing feature sizes.

Chapter 3

Topic - Literature Review

This chapter covers some works related to rice leaf disease detection and classification with their contributions and limitations.

Chapter 4

Topic - Proposed Methodology

The dataset and suggested methods are covered in this chapter. It also provides a thorough explanation of the suggested architecture, model training, and data pre-processing.

Chapter 5

Topic -Implementation

The implementation of base paper model and proposed model are covered in this chapter. It also provides a thorough explanation of the implemention.

Chapter 6

Topic - Result and Performance Analysis

This chapter compares the suggested architecture with related works and examines the experimental results and performance of the design. This chapter also includes a description of the measures that were utilized to assess my model.

Chapter 7

Topic - Conclusion and future work

The research has come to an end with this chapter. The results of my research are summarized and the future work in this chapter. We have also attempted to draw attention to our work's shortcomings and possible areas for improvement in the future.

1.9 Conclusion

This chapter lays the groundwork for the next investigation into the use of deep learning models for rice leaf disease detection and classification. It recognizes the difficulties farmers encounter in correctly diagnosing these illnesses and emphasizes how well-suited deep learning is for this purpose. This chapter explores the reasons for doing this research, the particular goals that the study seeks to accomplish, and the possible roadblocks that might appear along the way. In the chapters that follow, these difficulties and goals will be covered in more detail, giving the reader a road map for the research process.

Chapter 2

Background Study

2.1 Introduction

Globally speaking, rice is a staple meal for billions of people, and environmental factors have a big influence on both its output and quality. Accurately identifying these circumstances in advance is essential for efficient operations and reducing crop loss. Farmers visually check samples to detect diseases using traditional methods, which may be laborious, subjective, and error-prone.

2.2 Image Processing and Machine Learning

Prior to deep learning, scientists investigated feature extraction from rice leaf photos using image processing techniques, followed by classification using machine learning algorithms such as Support Vector Machines (SVMs) or decision trees.This strategy had some success, but it needed a lot of feature engineering knowledge. [14].

2.2.1 Image Processing for Rice Leaf Disease Detection

- **Prepares the Data:** methods for processing images tidy, improve, and pre-process the unprocessed photos of rice leaves. This might include: deleting from the picture any undesired components, such as dust or grain. changing the image's color space (from RGB to another, maybe more effective) to better emphasize illness signs. separating the leaf, or area of interest, from the backdrop to allow for more in-depth examination. To ensure uniformity throughout the dataset, resize and modify picture characteristics [13].

- **Feature Extraction:** Following pre-processing, image processing can identify pertinent characteristics from the pictures that are useful for categorizing different diseases. Analyzing color distribution or the existence of particular color patterns linked to illnesses are a few examples of these properties. analyzing the differences in texture on the leaf surface, which might be caused by illness. examining the general form and the existence of disease-related deformities [13].

2.2.2 Machine learning for Rice Leaf Disease Detection

- **Classification:** Machine learning techniques may be used to categorize the leaves into groups based on health once characteristics have been retrieved. Conventional algorithms such as:
  - **Support Vector Machines (SVMs):** These algorithms learn to create a separation hyperplane between healthy and diseased feature sets.
  - **Decision Trees:** In order to arrive at a final classification (healthy or ill), these algorithms construct a tree-like structure, where each node reflects a choice based on features [14].

2.3 Limitations of Traditional Machine Learning

- **Feature Engineering:** takes a great deal of experience to choose and create the best characteristics for categorization.
- **Limited Accuracy:** For difficult illness detection tasks, deep learning models may outperform traditional methods in terms of accuracy [14].

2.4 The Rise of Deep Learning

These restrictions are removed by deep learning, a kind of machine learning, which automatically extracts information from the photos. at particular, Convolutional Neural Networks (CNNs) excel at this task.

- **CNNs for Feature Learning:** Multiple convolutional layers make up CNNs, which automatically extract information at various degrees of abstraction from the pictures, ranging from low-level edges to high-level disease patterns.
- **Superior Performance:** When compared to conventional machine learning techniques, deep learning models detect rice leaf disease with noticeably greater accuracy [15].

2.5 Transfer learning for Rice Leaf Disease Detection and Classification

a pre-trained model that can distinguish between diverse things, such as automobiles, pets, and aircraft, using ImageNet training. We can take advantage of this model's preexisting capacity to identify forms, textures, and colors by transferring it and refining it on a dataset related to rice leaf disease. Building on the prior information, the fine-tuning procedure then specialized the model to recognize disease patterns on rice leaves. When utilizing deep learning models for rice leaf disease detection and classification, transfer learning has a number of benefits.

- **Reduced Training Time:** Transfer learning makes use of large, general picture datasets such as ImageNet to apply pre-trained models. Low- and mid-level characteristics that are often helpful for image identification tasks have previously been taught to these models. Our rice leaf disease detection model may be trained much faster by utilizing these pre-trained weights.
- **Improved Performance:** A solid basis for acquiring disease-specific properties is offered by pre-trained models. When the final layers of the pre-trained model are fine-tuned using our rice leaf disease dataset, the model becomes more adept at identifying these particular patterns, which frequently results in higher accuracy than when the model is trained from scratch using a smaller dataset.
- **Reduced Computational Resources:** Transfer learning uses less processing power to train our model because a large amount of the learning is accomplished using the pre-trained weights, making it more accessible to users with modest resources.

Overall, transfer learning is a valuable technique for rice leaf disease detection as it allows us to:

- **Achieve good results with smaller datasets**
- **Train models faster**
- **Reduce computational demands**

This makes it a useful and effective method, particularly for users with limited resources or data, for creating accurate models for rice leaf disease detection [3].

2.6 Hybrid CNN Model

To overcome certain constraints or improve job performance, a hybrid CNN model combines the advantages of Convolutional Neural Networks (CNNs) with other machine learning architectures [16]. This is an explanation:

- **Convolutional Neural Networks (CNNs):** Deep learning models are quite good at problems involving picture identification. They excel in taking spatial elements out of data, such as photographs.
- **Hybrid Approach:** A CNN and another machine learning model are combined to take use of each other's complimentary advantages in a hybrid CNN model. Here are a few typical pairings: When working with sequential data, such as in video analysis, this combination works well since the CNN can extract features from individual frames while the RNN can record the temporal correlations between them.This might help increase accuracy in classification jobs when the SVM does the final classification after the CNN extracts the information.
- **Benefits of Hybrid CNN Models** CNNs perform well in image recognition tasks, however they may not perform well in jobs involving sequential data or global context. Hybrid models, which combine many designs, can fill in these gaps. In some jobs, a hybrid method might result in increased accuracy, efficiency, or robustness by fusing the capabilities of many models.

2.7 CNN With Swin Transformer Architecture

Convolutional Neural Networks (CNNs) and Swin Transformers are being studied together in an attempt to enhance photo recognition. CNNs are known for their ability to capture local image details, and they are coupled with Swin Transformers, who are experts at interpreting the bigger picture within an image. This combination aims to use the advantages of each, with CNNs providing fine-grained local features and Swin Transformers capturing the connections between these elements for a more comprehensive image interpretation. By combining the benefits of both approaches, this powerful duo may boost productivity and enhance accuracy in image classification and application recognition. In an effort to fully explore the possibilities of this fascinating hybrid architecture, research is being done on dual-branch models and feature fusion, among other tactics. [17].

2.8 CNN Based on VGG-16 Architecture With Transfer Learning

Another approach, relying on transfer learning, uses the pre-trained VGG-16, which is trained in advance in networks with a large dataset and stands out with deep convolutional layers that help to extract a lot of visual information. VGG-16's few final classification levels are fully replaced by new layers suitable directly for classifying rice leaf diseases. If we refine these new layers with a collection of rice disorder images, the computer will be able to distinguish between normal and diseased leaves, even with diseases that have very similar characteristics. Although the trained VGG-16 has a wide repertoire of pattern recognition, the recognition program focuses exclusively on the specific task of classifying rice leaf diseases. As a result, such a learning framework provides quick recognition and at the same time achieves high accuracy without elaboration from scratch [18].

2.9 Conclusion

Rice growers have always been troubled with identifying the diseases that can substantially impair their harvest. However, the application of machine learning and image processing methods makes this task solvable. Traditional machine learning approaches provide a background for the methods applied, but their use is limited. Deep learning solves this issue. It autonomizes feature extraction from photos, thanks to the strengths of image processing , which, eventually, increases the classification of accuracy of rice leaf diseases. Eventually, it allows for actually applied implementations in agriculture. Due to the early and accurate identification of the disease, the farmer has an opportunity to react properly and on time. This improves the crop and implicitly increases the core management and yields at the end.

Chapter 3

Literature Review

3.1 Introduction

This review of the literature explores the state-of-the-art in rice leaf disease classification and detection at the moment. I will examine many image processing methods used for feature extraction and segmentation of disease regions. I will also look at how various machine learning algorithms and deep learning techniques are applied to the classification of diseases and talk about the results these techniques provide. The review will point out the drawbacks and restrictions of the current methods and suggest possible directions for further investigation. The overall goal of this review is to present a thorough overview of the developments and prospects in automated machine learning, deep learning and image analysis-based rice leaf disease detection and classification.

3.2 Related Works

The CNN architecture VGG16 has been used for training and testing on datasets gathered from the internet as well as paddy fields. Applying this model resulted in 92.42% accuracy. The three main diseases—bacterial leaf blight, brown spot, and rice leaf blast—were the focus of the investigation, which examined 1649 images of rice leaves. [1]. TThis work developed the SVM approach, which successfully identified data based on form, color, and texture aspects using 108 training examples. With an accuracy of 97.2%, the results showed that SVM was successful in identifying and classifying illness areas. [4].

In order to classify rice plant illnesses based on photographs, the study evaluated a variety of pre-trained deep CNN models, such as AlexNet, Vgg16, ResNet152V2, InceptionV3, InceptionResNetV2, Xception, MobileNet, DenseNet169, NasNetMobile, and NasNetLarge. The research used a dataset consisting of 1216 photos of damaged rice plants from actual fields; the photos were divided into seven classes: brown spot, rice blast, bacterial leaf blight, sheath rot, fake smut, and healthy leaves. Of all the models that were evaluated, Vgg16 had the best classification accuracy, coming in at 93.11%. [19].

The authors' work employed CNNs approaches to analyze eight important rice diseases, including bacterial leaf blight, false smut, rice hispa, blast, stemborer, sheath blight, brown spot, and brown planthopper. The report is divided into two main sections: the first part describes the research's survey methodology, and the second part explores the most advanced method for detecting rice disease (RDD) using CNNs. [5].

Using the YOLOv5 deep learning architecture, the authors of this paper presented a technique for identifying and detecting illnesses of rice leaves. A dataset with 1500 annotations was used to test, train, and assess the model. The following performance metrics were set: 90%, 67%, 76%, and 81% for recognition precision, recall, map values, and F1 score, in that order.[12].

The system created by Mobeen Muhammad and colleagues [20] was evaluated to tackle the problems of negative transfer learning and overfitting in knowledge transfer across domains. Two plant disease datasets, including the publicly accessible PlantVillage dataset, were used to assess the system. With an accuracy of 99% on the Pepper dataset and 99.69% on the PlantVillage dataset, the suggested approach beat earlier efforts on the PlantVillage dataset. Sukhvir Kaur and her colleagues (Kaur et al., 2014) [21] carried out an investigation on many commonly studied infections, providing an example of a research scenario at various stages of a disease detection system. They evaluated the effectiveness of cutting-edge techniques in order to pinpoint those that have promising outcomes across a range of crops or categories. The report outlines prospective study topics and highlights essential issues by showcasing a number of workable methodologies. The purpose of this survey is to help researchers better understand how computer vision is used to diagnose plant diseases. In their work, Vimal Shrivastava and associates (Shrivastava et al., 2015) created a machine learning method based on pictures to recognize and categorize plant illnesses. Their study concentrated on illnesses that affect Oryza sativa, or rice plants. The photos that showed signs of illness on the stems and leaves were obtained from real rice fields. The four categories that comprised the 619 photos of sick rice plants were Rice Blast (RB), Bacterial Leaf Blight (BLB), Sheath Blight (SB), and Healthy

Leaves (HL). They obtained promising results by using a pre-trained deep convolutional neural network (CNN) as a feature extractor and Support Vector Machine (SVM) as a classifier. This early warning system and preventative action for rice illnesses might be provided by this early detection system. Additionally, It might develop into a useful instrument for diagnosing rice plant problems in actual farming settings.

A novel lightweight custom convolutional neural network (CCNN) model intended for efficient crop leaf disease detection was reported in the study. Three distinct datasets were used to assess the model's performance, and it showed good accuracy with fewer parameters. The accuracy, number of parameters, and training duration of the models were thoroughly compared with those of transfer learning (TL). Furthermore, heatmaps showing important areas during detections were produced using Gradient Activation Map (Grad-CAM) approaches. The results of this study enhance crop leaf disease detection techniques and have applications in improving agriculture and lessening the effects of crop diseases. [22].

Food production and quality are significantly impacted by diseases that harm rice plants, as demonstrated by the study carried out by [13]. Improving the quality of output can result from the detection of certain disorders. This study investigates a range of methods, including image processing, machine learning, and deep learning, that are employed in the identification of detrimental illnesses in rice plants. Many attempts have been made to use leaf photos for automated rice plant disease detection. The study examines many techniques for diagnosing problems in rice plants and concludes that deep learning methods have greater potential than the other two.

Using machine learning techniques, this study offers a unique way to diagnose illnesses of rice leaves. The study focuses on identifying three common diseases that affect rice crops: brown spot infections, bacterial leaf blight, and leaf smut. The input for the system needs crisp pictures of infected rice leaves on a white background. The dataset undergoes necessary pre-processing before training, and a range of machine learning methods, such as KNN (K-Nearest Neighbor), J48 (Decision Tree), Naive Bayes, and Logistic Regression, are used to train the dataset. When evaluated on the dataset, the decision tree method shows an amazing accuracy of more than 97% after 10-fold cross validation. [14]. tK Kiratiratanapruk et al. [15] carried out experiments in their research to train and evaluate each model using a dataset of 6,330 photos. Based on the results of the studies, it was found that YOLOv3 performed the best in terms of mean average precision (mAP), scoring 79.19% for rice leaf disease detection and classification. On the other hand, Mask R-CNN, Faster R-CNN, and RetinaNet achieved 75.92%, 70.96%, and 36.11% of precision, respectively. The photographs that show indications of rice illnesses on the leaves and stems of rice plants in a rice field are shown in this article. The six hundred and ninety-nine photos that showed sick rice plants were categorized into four groups: Rice Blast (RB), Bacterial Leaf Blight (BLB), Sheat Blight (SB), and Healthy Leave (HL). A pre-trained deep convolutional neural network (CNN) was used to extract features, and a Support Vector Machine (SVM) was used as the classifier. The results showed promise. This early detection technique for rice illnesses has the potential to serve as both an early warning system and a preventative tool. Additionally, it may be improved upon to provide a useful method for determining rice plant illnesses in the agricultural sector.citer11 In a research, Ahmad et al. [20] introduced a stepwise transfer learning technique to address problems with quick convergence, decreased overfitting, and avoided negative transfer learning when transferring information between domains. Two plant disease datasets were used to test the method: one from the National Institute of Horticultural and Herbal Science in the Republic of Korea, which is a pepper disease dataset, and the other from PlantVillage, a publicly available dataset. The pepper dataset's variety of photos depicting various plant sections, such as the leaf, pulp, and stem, made it a special challenge. With accuracy rates of 99% on the Pepper dataset and 99.69% on the PlantVillage dataset, respectively, the suggested approach outperformed earlier research on the PlantVillage dataset.

3.3 Conclusion

Out of the aforementioned papers, we have selected the research conducted by Shreya et. al (2020) [1] in order to enhance their research outcomes. The utilization of performance metrics was not comprehensive, and the accuracy of their proposed VGG-16 model based on CNN was relatively low at 92%. Additionally, they did not mention the computational parameters and hyperparameters employed. Furthermore, their model was trained on a small dataset consisting of only 1649 images of rice leaf diseases. In order to overcome these limitations, we have devised and implemented our proposed methodology and dataset to address this issue.

Chapter 4

Proposed Methodology

4.1 Introduction

This chapter presented a comprehensive methodology for automatic rice leaf diseases classification using convolutional neural network. By implementing this methodology, a robust system can be developed to assist farmers in early and accurate disease identification, ultimately leading to improved crop management and increased rice yield.

4.2 Overview of the proposed Methodology

The technique described in Shreya et al. (2020) [1] has been applied to the proposed framework, with the inclusion of stages as shown in Figure 4.1. To improve data efficacy and consistency, a larger dataset was added, and picture normalization was used. We used the suggested dataset to create our proposed CNN with Swin Transformer architecture. We evaluated our model's performance by adjusting its hyperparameters. Figure 4.1 presents an excellent illustration of the complete process that demonstrates the seamless succession of implementation phases.

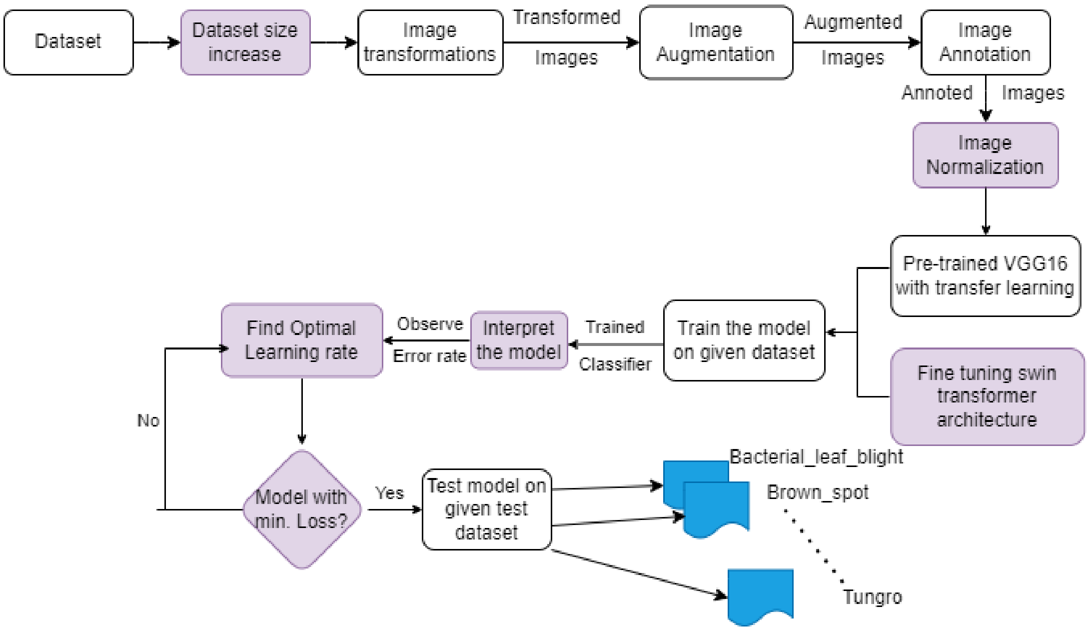


Figure 4.1: Proposed CNN based architecture

4.2.1 Dataset Descriptions

This experiment utilized a rice leaf disease dataset containing 18,445 images obtained from the reliable source Kaggle. The dataset encompassed ten classes: bacterial leaf blight, brown spot, healthy leaves, leaf blast, leaf scald, narrow brown spot, neck blast, rice hispa, sheath blight, and tungro. A CNN technique was applied to this dataset and achieved the highest precision. The distribution of the data is visualized in a seaborn chart.

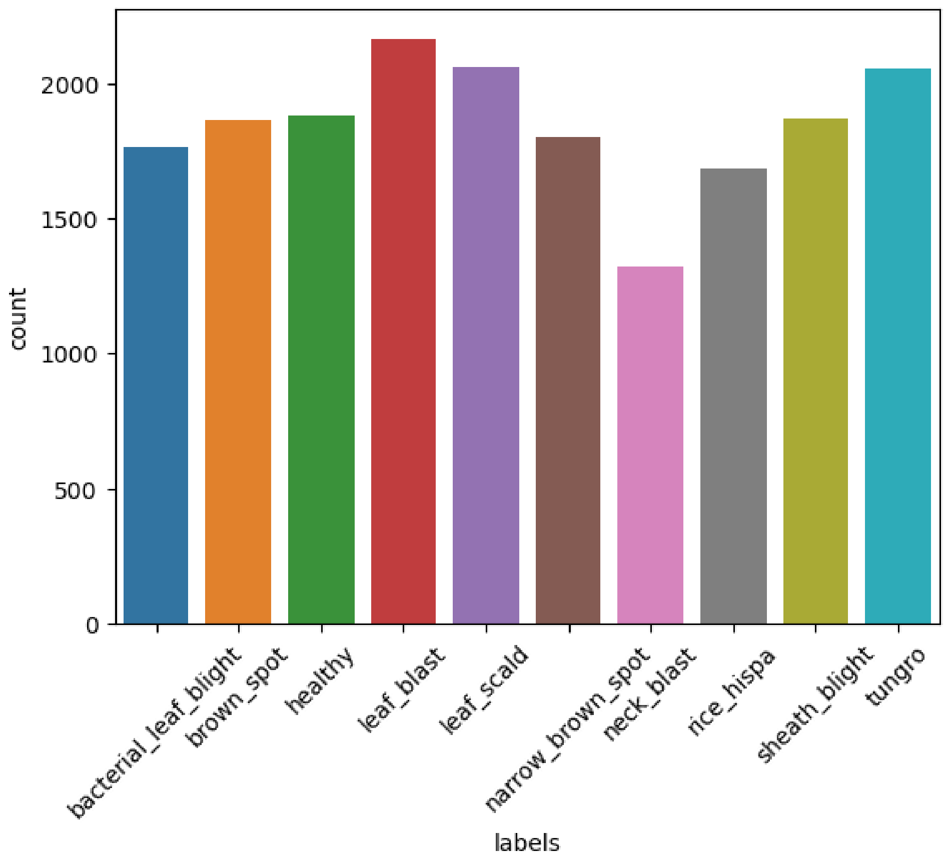


Figure 4.2: Seaborn chart representation of rice leaf diseases dataset

Source: https://www.kaggle.com/datasets/loki4514/rice-leaf-diseases-detection
The below sections describe the classes of Rice Leaf diseases on which we have worked.

A. Bacterial_Leaf_Blight:

Xanthomonas oryzae is the causative agent of leaf blight, a bacterial disease. The diseased leaves curl up after being discolored, turning from green to a grayish tinge. Subsequently, the leaves start to turn yellow, afterwards becoming straw-colored, and ultimately withering and dying. The leaf lesions with wavy edges gradually spread toward the root. Bacterial slime, which resembles drops of morning dew, is observed early in the formation of lesions. For an illustration of the effects of Leaf Blight on leaves, please see Figure 4.3(a).

B. Brown_Spot:

That's a fungal illness. The affected leaves may finally completely die as a result of several large spots. The leaves initially exhibit small, circular lesions that range in hue from dark brown to purple-brown. The fungus produces a toxin that causes the lesions to take on an oval or circular

Figure 4.3: (a)-(j) From Left to right. (a)bacterial_leaf_blight (b) brown_spot (c) healthy (d) leaf_blast (e)leaf_scald (f) narrow_brown_spot (g) neck_blast (h) rice_hispa (i) sheath_blight (j) tungro

form with a light brown to gray center. It has a reddish brown rim. The small dark brown lesions on the leaves affected by Brown Spot are seen in Figure 4.3(b).

C. Healthy_Leaves:

Healthy leaves are free from all types of diseases. In contrast to sick leaves, which have an uneven color distribution, their color distribution is consistent and homogeneous. We tested our model's ability to detect pictures free of illness using leaves from healthy rice plants. The corresponding image is displayed in figure 4.3 (c).

D. Leaf_Blast:

Magnaporthe oryzae is the fungus that is causing the illness. Spindle-shaped or elliptical white to grey-green dots with dark red to brownish edges may be seen at first. Like a diamond, certain lesions may have wide centers and pointed ends. Figure 4.3(d) shows spindle-shaped lesions with dark brown edges and white spots.

E. Leaf_Scald:

epidemic struck Burma in 1935, and in the Philippines during rainy season, a noteworthy outbreak was reported [18]. The symptoms, which resemble greenish spore-like balls, appear on the panicle grains. These spore balls, which have a diameter of at least one centimeter, include whole floral components. They are flat, golden, and smooth when they initially appear. They are covered with a membrane, but as they develop, it breaks. The color eventually turns orange, then yellowish-green. Figure 4.3(e)

G. Neck_Blast:

Neck blast disease on rice leaves can result in small, grayish-white spots with dark green borders, albeit it is less common than on the neck itself. These lesions are smaller than neck lesions, but they have the potential to enlarge over time. They can be spindle- or elliptical-shaped, with white or gray centers and reddish-brown or necrotic margins [23].

H. Rice_Hispa:

The rice hispa is one type of larval insect that defoliates plants. CNNs have not been used extensively in research on rice hispa illness; just one publication [24] has addressed this issue. However, the rice hispa disease has not been properly researched in CNNs. In a recent paper, researcher [25] provided a method for clustering and classification using k-means, SVM, and CNN approaches. The study employed a dataset of 50 images of rice hispa with a 95% accuracy rate.

I. Sheath_Blight:

Rhizoctonia solani is the etiology of SB disease (Nandakumar et al., 2001 [26]). The disease initially affects the bottom sheath of the plant before moving on to the upper sheath and leaves. It can spread from plant to plant in the wild (Zhou et al., 2015 [27]). SB sickness has developed into a very dangerous illness in the world's major rice-producing nations (Zhou et al., 2015 [27]). For this study, one hundred images depicting SB illness have been gathered.

.,

J. Tungo_Disease_Leaf:

Asian cultivated rice, scientifically known as Oryza sativa L., is a vital crop that belongs to the Poaceae (Gramineae) family. Being a staple food for a sizable portion of the global population, it is one of the most significant crops in the world. Unfortunately, a disease called Rice Tungro condition (RTD) affects rice production greatly and costs the world's rice industry about US $1.5 billion annually. This disease lowers rice production in South and Southeast Asia by 5% to 10%. The rice tungro bacilliform virus (RTBV) and the rice tungro spherical virus (RTSV) are the viruses that cause rice tungro. The RTBV virus, a double-stranded (ds) DNA genome virus, belongs to the Tungrovirus genus within the Caulimoviridae family. The length of its particles varies from 100 to 300 nm. Nevertheless, RTSV and other single-stranded (ss) RNA viruses are members of the same family. Helper viruses, such as RTBV and RTSV, are known to facilitate the spread of RTBV and are present in several countries that produce rice. [28].

4.3 Dataset Preparation

The pre-processing and data transformation procedures used in this research project are described in this section. The two subsections that follow are separated inside this section:

4.3.1 Dataset Preprocessing and Augmentation

The collected images have been resized to 224 by 224 pixels. To train an effective classifier for rice leaf disease, images need to be augmented (rotated, flipped), and preprocessed (resized, noise reduced, normalized, and color space translated). This increases the dataset's size and diversity, which improves the model's capacity to generalize and make it more resilient to changes in the real world.

4.3.2 Image Transformation

Image augmentation is done during the complete data translation process. Magnified versions of the photos included in the dataset are used in picture augmentation, a frequently used technique for data augmentation that increases the size of the training set. The process of altering the images in the training set is known as image augmentation. An image data generator is utilized in this project to improve photographs for CNN. Numerous transformation operations, including rotate, shift, zoom, flip, and many more, are used in this approach. While rescaling photographs is a part of data pre-processing, image augmentation is a separate process. Rather, this technique modifies specific images to enable more modeling. are carried out in this research project using CNN model training data. At this stage, the batch size for the complete dataset is first determined. Batch size is one of the most important hyperparameters for optimizing deep learning models. Each batch consists of 32 photographs when the batch size is set to 32. The batch size should be accurately defined for each model. No matter the size, poor generalization results from size. After the batch size is established, each image is turned around. It is possible to rotate images thirty degrees in a clockwise direction. Image rotation is important because every pixel is fetched throughout the rotation process, even if it's possible that some pixels may be out of range. The picture flip is an additional important element. This is done in a vertical flip.

4.3.3 Image Normalization

Picture normalization is the first step in the use of deep learning models for rice leaf disease diagnosis. The following are some strategies we might employ and why they are important: Pictures can differ in terms of size, color balance, and lighting. By reducing these variations, normalization enables the model to concentrate on characteristics unique to a particular situation. The model trains more effectively and converges more quickly when all of the pictures have the same scale.The process of gathering pertinent information, such as color, texture, and shape, for the categorization of diseases is improved by normalization. You may greatly increase the precision and effectiveness of your rice leaf disease detection system by utilizing image normalization.

4.4 Data Modeling

In this section, six different machine learning models and their functionalities are discussed. Each model serves a specific purpose.

4.4.1 Overview

Significant variations in superstructures and equipment were revealed by a study of rice leaf disease photos, underscoring their unique characteristics. A CNN-Swin Transformer model was created to combine the advantages of both architectures for rice leaf disease image analysis since CNNs are good at extracting local features and transformers with self-attention are good at capturing particular characteristics.Figure 4.4 displays the structure of the model network.
The three parts of the model are the CNN backbone, transformer backbone, and multi-layer perceptron (MLP) neural network-based classification module. In order to reliably and effectively extract image features, the model makes use of the CNN structure for feature extraction and a CNN backbone. The CNN backbone generates a feature vector by continually stacking CNN-Blocks, which enriches the lowest layer of the feature map with more semantic information. After employing a CNN-Block in the CNN backbone to reduce the resolution of the feature map, four layers are used to extract data. Each layer has 2, 2, 6, 2 CNN-Blocks; the

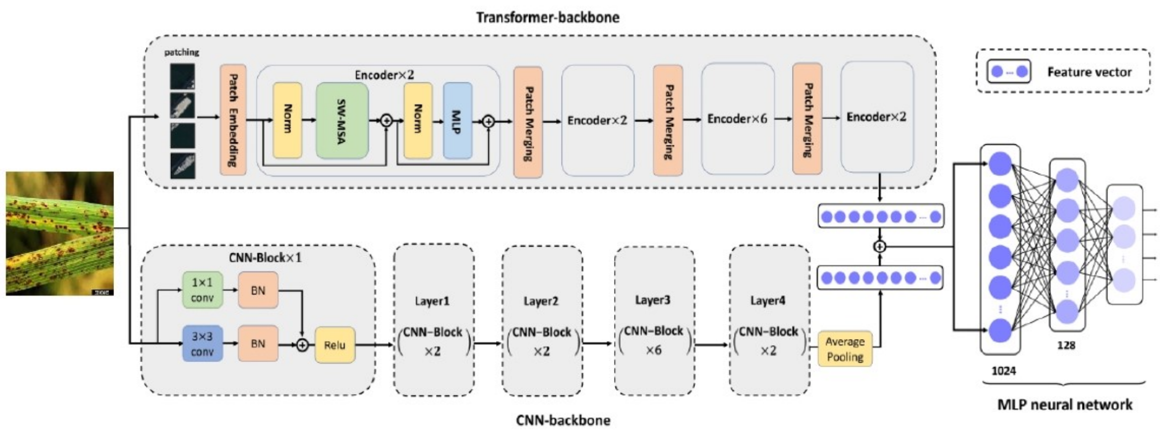


Figure 4.4: Network architecture of the proposed model [2]

first CNN-Block in each layer employs a convolutional kernel with a two-pixel stride in order to lower the resolution. This method was created using the research [29] as a basis.

It appears from empirical data that replacing pooling layers with convolutional layers that have strides longer than one pixel does not reduce classification accuracy—in fact, it may even somewhat increase it. As a result, the feature map size is condensed using a stride of two pixels, guaranteeing improved accuracy, a smaller spatial dimension, and the removal of duplication.

The transformer backbone is supported by the Swin transformer. Initially, each visual channel is split up into 48 patches, each measuring four by four. These patches are called ”tokens” because they are the smallest processing units in the model, making up the full image without any overlap. The transformer backbone then uses encoders to obtain picture attributes after embedding each patch. Local information is included in the feature vector that is produced. Moreover, the transformer backbone extracts features via patch merging before to each encoder, producing feature maps that include high-level semantic information. During each patch merging process, tokens undergo a downsampling by a factor of two in order to concatenate the properties of adjacent 2×2 tokens.

Three fully connected layers make up the MLP neural network: an output layer whose size is dependent on the number of classes to be detected, a hidden layer with 128 dimensions, and an input layer with 1024 dimensions. The CNN-Swin model maps the fused feature vectors to ship classification using an MLP neural network.

4.4.2 The Backbone of a Convolutional Neural Network (CNN)

In order to minimize overfitting and provide better results, the CNN backbone of the model should be a condensed CNN structure like the VGG model [30], which has minimal parameters. On ImageNet, however, VGG-19's Top-1 accuracy is 72%, and on the military ship dataset utilized for this investigation, it was 84.5%. Both have less classification accuracy when compared to the sophisticated CNN models of today. The CNN backbone utilized in this study created Resnet-based models by taking cues from Resnet [31]. Better results are obtained with the multi-branch CNN-Block module, which is based on CNN structures with fewer parameters (like the VGG model).

4.4.3 Layers

Cascaded CNN-Blocks are used by the CNN backbone to extract features from the picture. It is split into two primary sections. The initial segment comprises distinct CNN-Blocks designed to reduce picture resolution, a 3×3 convolutional layer, a BN (batch normalization) layer, a parallel structure made up of a BN layer and a 1×1 convolutional layer, and so on. To create the feature map, a rectified linear (ReLU) activation function is then employed. The final result is the output of the second portion, which consists of a global pooling layer and four layers with different numbers of Resnet blocks in each layer. In Figure 4.5, the Layer architecture is displayed. The convolution channel counts of 64, 128, 256, and 512, respectively, are the only differences between Layer 1 and Layers 2, 3, and 4. Furthermore, Layer 3's volume is three times that of Layer 1's. Several CNN-Blocks with comparable structural patterns comprise each layer. The main distinction is shown in the first CNN-Block, where variations in the resolution and channel numbers relative to the input X are caused by the 2 pixel strides in the 3×3 convolution layer and the 1×1 convolutional layer. In the ensuing CNN-Blocks, the 1×1 and 3×3 convolutional layers have one pixel strides, guaranteeing that the feature map's dimensions and channel counts don't change.

4.4.4 CNN Block

The CNN-Block architecture comes from the Resnet architecture, which employs residual connections to get feature information via $Y = \mathrm{ReLU}(F(X) + X)$, a fundamental residual module equation. In this scenario, the input is X and the output is Y, and the network has to learn the

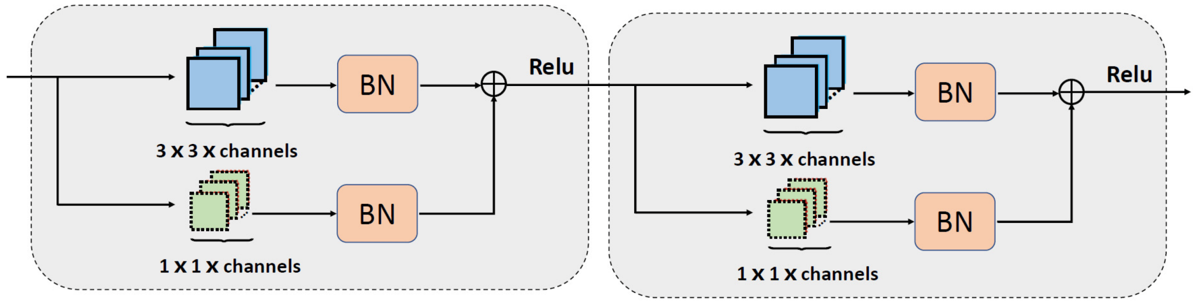


Figure 4.5: Network architecture of layer1 [2]

residual mapping F(X). To modify the size of X, one must use a 1 x 1 convolutional layer to match F(X) when the feature map F(X) decreases. Y = ReLU(F(X) + G(X)) is the convolutional shortcut that is used in the 1 x 1 convolutional layer. Because of its multi-branch architecture, the Resnet structure benefits from an implicit ensemble of numerous shallower models [32].

4.4.5 Transformer Backbone

The CNN-Swin model, which is based on the Swin transformer model, is used to extract visual characteristics. Weight sharing and local perception are made possible by the CNN structure's unique induction bias, but it also somewhat reduces the amount of picture characteristics that can be retrieved. However, because of its lower induction bias, the transformer structure is more akin to the focus that a human sees when identifying an image. As a result, it can extract features more effectively. The Resnet-like hierarchical transformer structure is very useful for extracting characteristics from ship photos. The transformer model used in this study consists of an encoder and a patch embedding. After patch embedding, the input ship picture is transmitted via a multi-layer cascaded encoder module to extract its features. A norm operation, shifted, window-based, multi-head, self-attention (SW-MSA), and an MLP make up the encoder.

4.5 Conclusion

The following section delves into the basic components required for the proposed approach within the deep learning framework. These components include dataset organization, data pre-processing and enhancement, image conversion, image standardization, and a model summary that includes layers, CNN backbone, transformer backbone, and CNN backbone. The essential elements required for the suggested strategy in the deep learning-based framework are examined in the next chapter.

Chapter 5

Implementation

5.1 Introduction

This section provides an explanation of the models developed for the CNN and deep learning model development project. These algorithms attempt to detect and categorize rice leaf illnesses using pictures of both healthy and diseased rice leaves. The ailments are divided into three categories: healthy, bacterial leaf blight, and brown spot. In addition, the following categories are included: narrow brown spot, neck blast, leaf blast, leaf scald, tungro, rice hispa, and sheath blight. To complete the project, Jupyter is used as the Integrated Development Environment (IDE) and Python version 3.8 is used. It is important to keep in mind that training the CNN model takes a substantial amount of time. Therefore, a good system setup, a GPU, and the proper installation of all necessary libraries are required before training and executing the model. Many libraries, including Matplotlib, Pillow, Scikit-learn, PyTorch, TensorFlow, and OpenCV, were installed in order to do this. A popular framework for working with images and creating effective models is TensorFlow. Conversely, Keras is an interface for TensorFlow and a potent framework for creating deep learning models. Regression and classification applications require machine learning techniques, which are implemented using Scikit-learn. Deep learning model development and training are the main uses for PyTorch, an open-source machine learning framework. especially for applications like recommender systems, natural language processing, and picture recognition. Matplotlib from the Python programming language is used for charting, and OpenCV and Pillow are utilized to manage picture data. To set up the environment for the project's real implementation phase, all of these libraries were installed. The initial model, the CNN model, eliminates some elements from the pictures.

5.2 CNN With Swin Transformer Model Training With Proposed Dataset

The model architecture used to categorize rice leaf infections is described in the table. Two distinct processing units receive the 224 × 224 pixel images: a Transformer and a Convolutional Neural Network (CNN). They all take useful information from the picture at the same time. The model then applies a straightforward addition operation to each linked element in order to combine the information from the CNN and Transformer. After that, a three-layer Multi-Layer Perceptron (MLP) neural network receives this aggregated feature vector. The merged feature vector is sent to the input layer. In the buried layer, 128 neurons process the information through calculations. The output layer generates the final output, which includes likelihood probabilities related to various diseases, using a softmax.

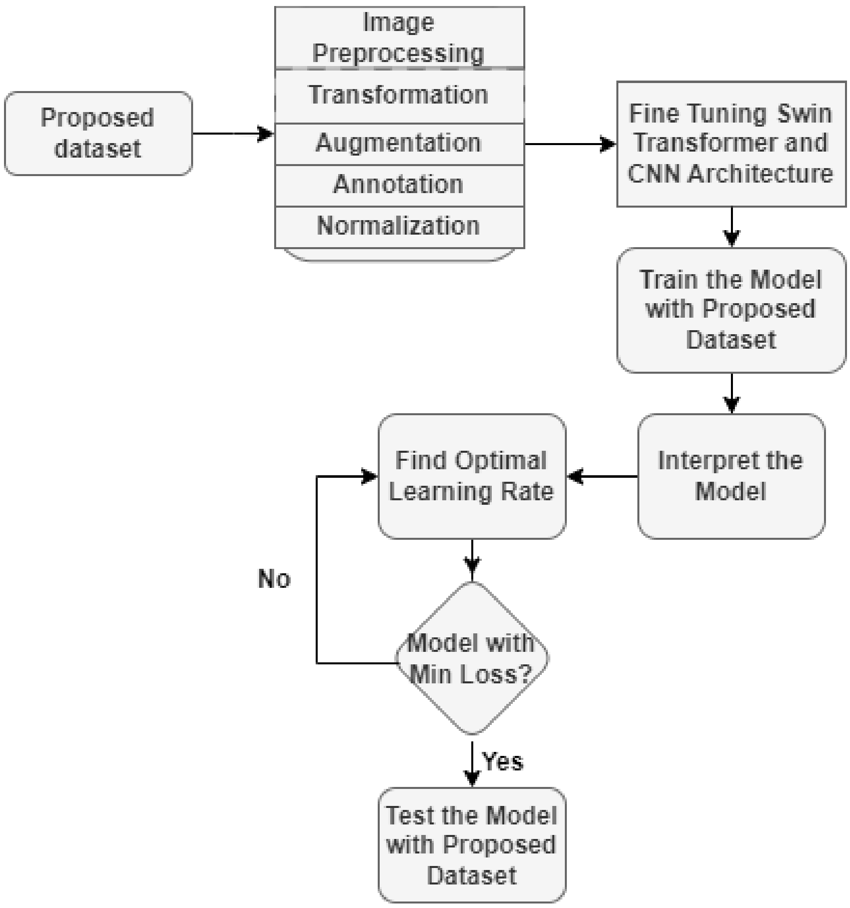


Figure 5.1: Overview of the steps of Swin Transformer and CNN model.

5.3 VGG-16 CNN Model Training With Proposed Dataset

The training and testing process starts with the loading of the image data set. The relevant images and class labels are stored in separate arrays for training. To ensure a fair distribution, twenty percent of the data is utilized for training and the remaining eighty percent is reserved for testing. The train_test_split function is used for this.
Moreover, thirty percent of the eighty percent training data is reserved for validation when it is divided again. The class labels are encoded as numbers to speed up processing as much as feasible. After that, one-hot encoding is applied to these labels, converting each label's representation from an integer to a vector. Next, the keras library is used to import the VGG-16 model. The residual layers of the model are the ones that remain after the final completely linked layers are eliminated. After then, the remaining layers are designated as untrainable. Following the flattening of the feature extractor component's output, an output layer with softmax activation and a fully linked layer emerge. The model is constructed using the Adam optimizer with categorical_crossentropy as the classification loss function in order to get it ready for training. Fifty epochs of training are conducted, and it is concluded that the outcomes have stabilized at that point. The steps involved in the categorization process are shown in Figure 5.1.

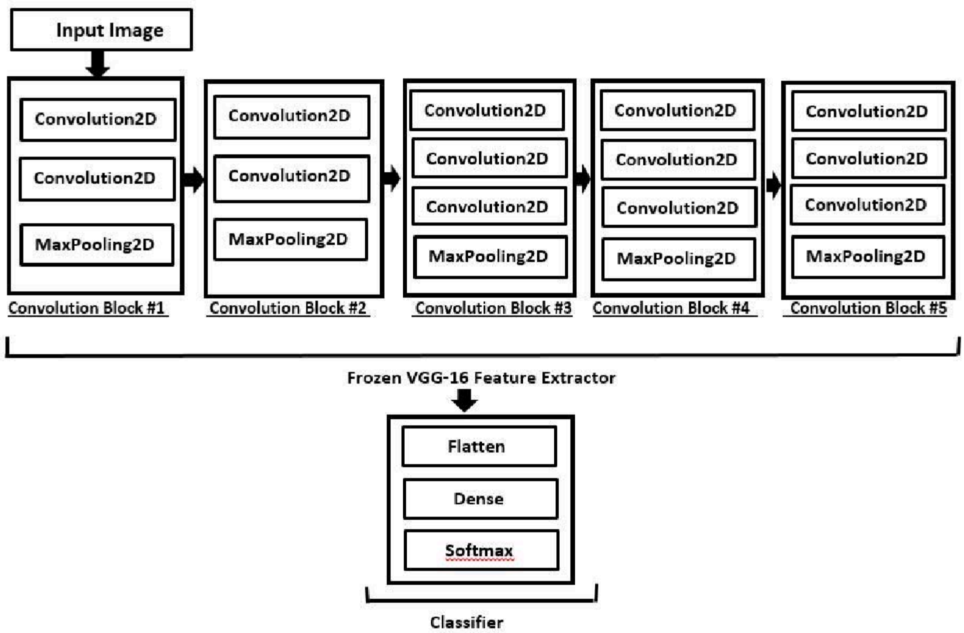


Figure 5.2: VGG-16 Architecture fine–tuned with the last two layers with 128 Dense FC Layer and 4 Dense Softmax Layer as the output [1].

5.3.1 Justification For The Chosen Model

The practice of using knowledge from one environment to enhance performance in another is known as transfer learning. Because there are many real-world instances where there is insufficient labeled data to support complicated models, this strategy is beneficial because it shortens the time required to train a neural network model. To train a neural network from scratch, a sizable dataset is frequently needed, albeit this data may not always be accessible. Because the model has previously been pre-trained, transfer learning enables the creation of trustworthy machine learning models with fewer training data. Consequently, we trained the pre-trained VGGNet on our little dataset and then adjusted it for classification [33].

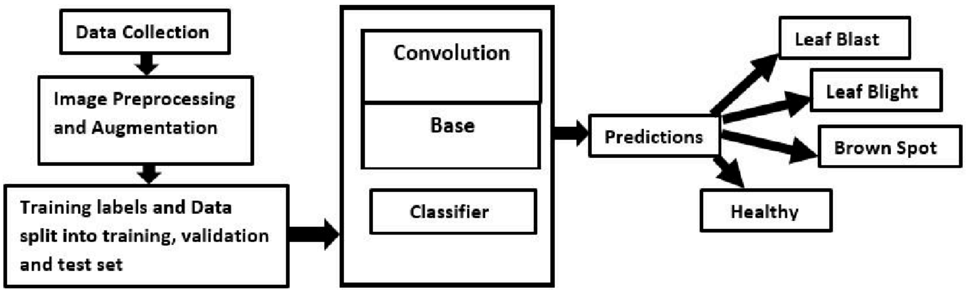


Figure 5.3: Overview of the steps of the base paper model [1].

5.4 Conclusion

This section offers a comparison between our proposed system and the methodical way it is implemented. Our dataset collection technique, which has been enlarged from the original study, has been shown. Additionally, we have effectively implemented the CNN and Swin architecture that we proposed. Furthermore, the suggested dataset was used to train the VGG-16 CNN model that we developed. Finally, we have compared the performance of the architecture in the original study with the design in our model architecture.

Chapter 6

Results and Performance Analysis

6.1 Introduction

This chapter serves as an analog for the systematic execution of our suggested system. We have shown how we collected data, as well as identified the flaws that came from our exploratory data analysis and offered our proposed solutions. In addition, we implemented the base paper model with our proposed dataset. We also implemented our prosed model architecture with proposed dataset. Finally, we evaluated the proposed model and base paper model by using some measuring tools like accuracy, precision, recall, F1 score and confusion matrix.

6.2 Evaluation Metrics

Metrics for evaluating performance are indispensable for determining the efficacy of models, algorithms, or systems in numerous disciplines, such as machine learning, data analysis, and decision-making. These metrics provide quantitative measures for evaluating the quality, precision, and efficacy of a model's predictions or classifications.

True Positives (TP): In these cases, the model predicts the positive class (1) accurately, and the data point's actual value is also positive (1).

True Negatives (TN): In these cases, the model predicts the negative class (0) accurately, and the data point's actual value is also negative (0).

False Positives (FP): while the model predicts the positive class (1) while the actual value of the data point is negative (0), this is known as a false positive. The model predicted the positive class, which is why the term ”positive” is used, whereas the term ”false” is used when the model produced an inaccurate prediction.

False Negatives (FN): while the model predicts the negative class (0) while the actual value of the data point is positive (1), this is known as a false negative. The word ”false” denotes an inaccurate prediction produced by the model, akin to false positives, while ”negative” denotes the prediction of a negative class by the model.

In conclusion, by contrasting the predicted labels of the data points with the actual labels, these metrics are used to measure the effectiveness of classification algorithms. Now that we are aware of these metrics, we can go on to discuss each assessment indicator and how it helps determine the efficacy of the model.

6.2.1 Confusion Matrix

Positive, and False Negative forecasts allows it to do this. This research provides insightful information about the accuracy and kinds of mistakes the model makes while categorizing different objects. An examination of the components that make up a confusion matrix is provided below:

		Predicted	
		Negative **(N)** -	Positive **(P)** +
Actual	Negative -	True Negative **(TN)**	**False Positive (FP)** **Type I Error**
	Positive +	**False Negative (FN)** **Type II Error**	True Positive **(TP)**

Figure 6.1: Confusion Matrix [3]

6.2.2 Accuracy

Especially in classification tasks, accuracy is a critical performance indicator for machine learning systems. It calculates the ratio of the algorithm's accurate predictions to all of the predictions made. It's critical to remember that accuracy by itself could not give a whole view of the model's performance, particularly in cases where datasets are unbalanced or if the costs associated with false positives and false negatives fluctuate greatly. Furthermore, while accuracy is helpful in assessing the model's performance using the available dataset. In order to evaluate the model's capacity for generalization and prevent overfitting to the training set, methods such as hold-out testing and cross-validation must be used.

The ability of the model to accurately identify positive examples (true positives) among all instances that the model predicted as positive (including true positives and false positives) is measured by a performance indicator called precision in binary classification tasks. To put it another way, precision gauges how well the model predicts favorable outcomes.

$$Precision = \frac{TP}{TP + FP} \tag{6.1}$$

When a model has a high precision score, it is more accurate in recognizing positive instances and produces fewer false positive predictions. It implies that a data point is more likely to be accurate when the model predicts it to be positive. A low accuracy score, on the other hand, indicates that the model may be less accurate in differentiating between positive and negative occurrences and has a tendency to provide more false positive predictions.

6.2.3 Recall

Recall is another crucial performance parameter used in binary classification tasks; it is sometimes referred to as sensitivity or true positive rate. The assessment gauges the model's capacity to accurately detect positive examples, or true positives, among all occurrences that are genuinely positive, including false negatives..

$$Recall = \frac{TP}{TP + FN} \tag{6.2}$$

A high recall score means that the model can accurately identify positive examples in the dataset. Stated differently, the model has a higher probability of accurately classifying a positive data point when it comes across one. Conversely, a low recall score indicates that the model is less successful in correctly detecting every positive event and is more likely to overlook positive examples.

6.2.4 F1-Score

In binary classification problems, the F1-score is a performance indicator that helps balance recall and precision. When there is an imbalance in the dataset—that is, when one class is significantly more common than the other—it is very helpful. An even more thorough assessment of the model's performance is provided by the F1-score, which offers a single figure that accounts for both precision and recall.

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{6.3}$$

The F1-score will be near to 1, suggesting a balanced model performance in terms of both the accuracy of positive predictions and the capacity to capture all positive cases, when precision and recall have similar values. The F1-score, on the other hand, will be lower if accuracy and recall diverge much, indicating that the model was unable to strike a compromise between the two measures.

6.3 Shreya Ghosal et al. (2020) [1] Classification Findings

The authors advise using transfer learning in conjunction with a pre-trained VGG-16 model. After being tested on 647 test data points and trained for 25 epochs on 1509 training data points, the model had a test accuracy of 92.4% and a training set accuracy of 97%. They also divided the data into test, validation, and train sets and performed trials on the same data set with an alternative CNN model without transfer learning. By changing the optimizer, epoch count, and batch size to 16, 30, and rmsprop, respectively, and by employing a dropout rate of 0.4, they were able to achieve the highest accuracy of 74%. This CNN model, which does not use transfer learning, consists of four convolution layers, each of which is followed by ReLU, Maxpooling, dropout, and two fully connected layers. in addition to SoftMax. Table 6.1 displays a comparison of the accuracy of the proposed CNN model with transfer learning vs the CNN without transfer learning. [1].

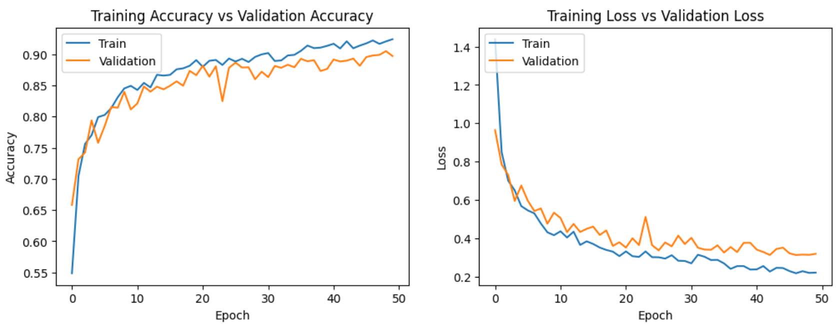
Model	*Test Accuracy*
CNN With Transfer Learning	**92.46%**
CNN Without Transfer Learning	74%

Table 6.1: Performance Comparison of CNN model with and without Transfer Learning [1].

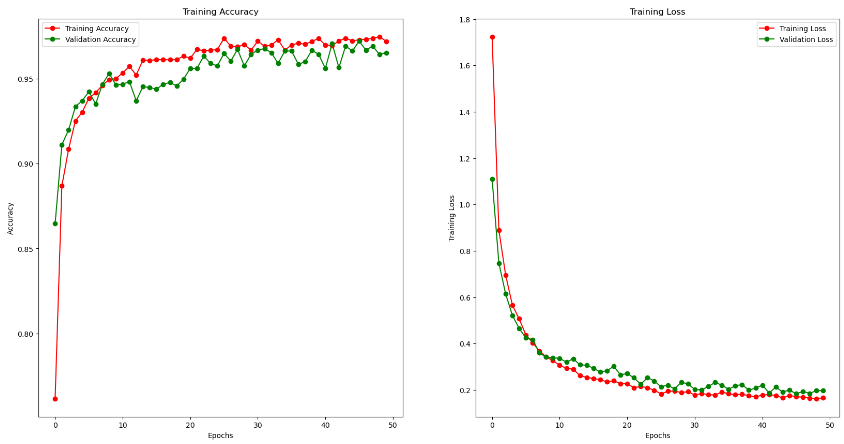
6.4 Classification Results With Proposed Dataset

This portion shows and addresses the result derived by the proposed methodology and proposed dataset. A pre-trained VGG-16 with transfer learning and fine tuning CNN with swin transformer architecture were experimented with training and test sets containing 70% and 30% respectively (50 epochs). The existing paper architecture and proposed architecture with proposed datasets results are shown in the following tables and figures.

6.4.1 Training vs Validation Accuracy and Loss Graph



(a)



(b)

Figure 6.2: Accuracy and Loss graph of training and validation for CNN model i.e, (a) Pre-trained VGG-16, (b) Fine tuning CNN with swin transformer architecture.

6.4.2 Model Performance with Confusion Matrix

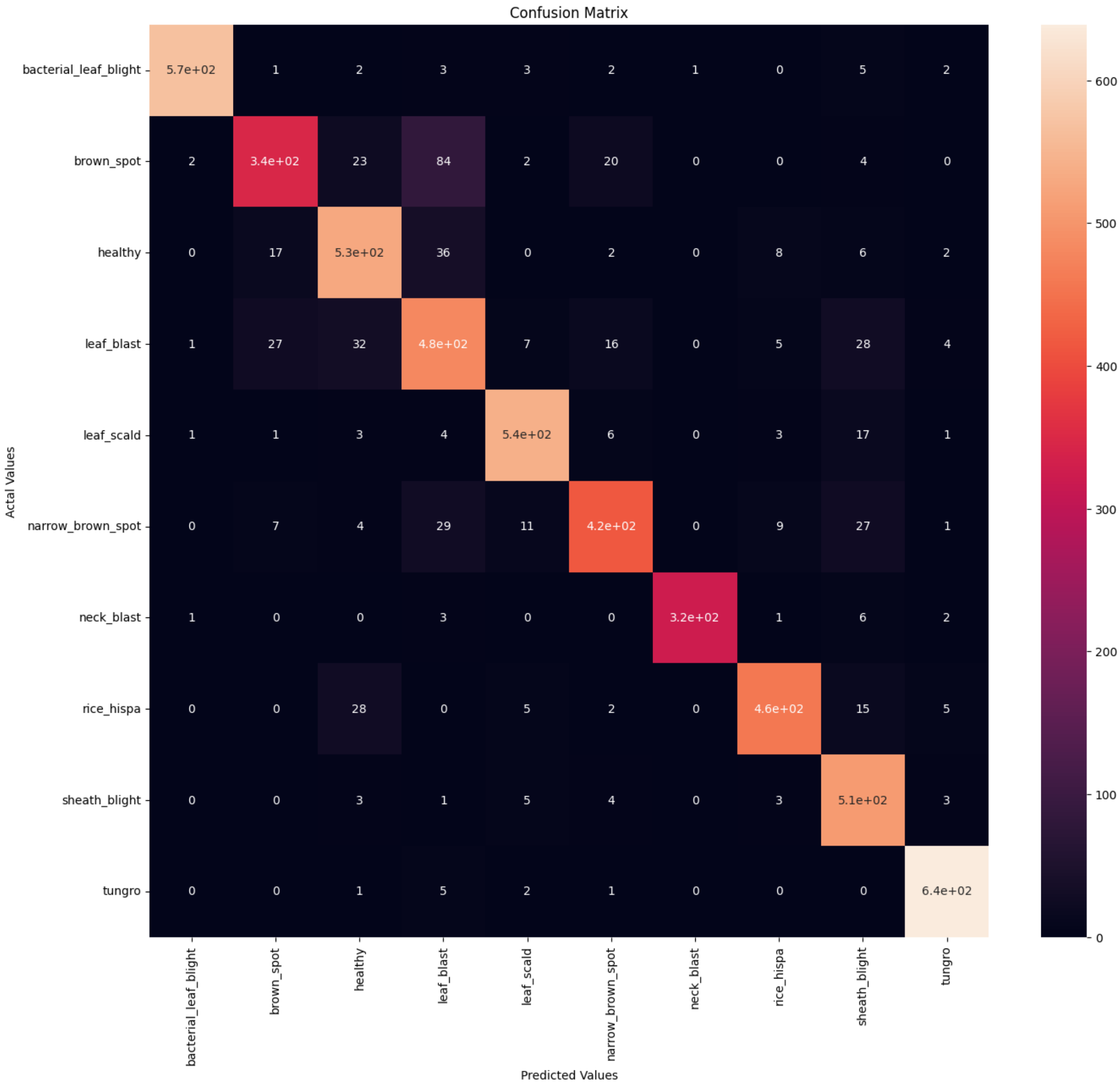


Figure 6.3: Confusion Matrix of Pre-trained VGG-16 with Proposed Dataset.

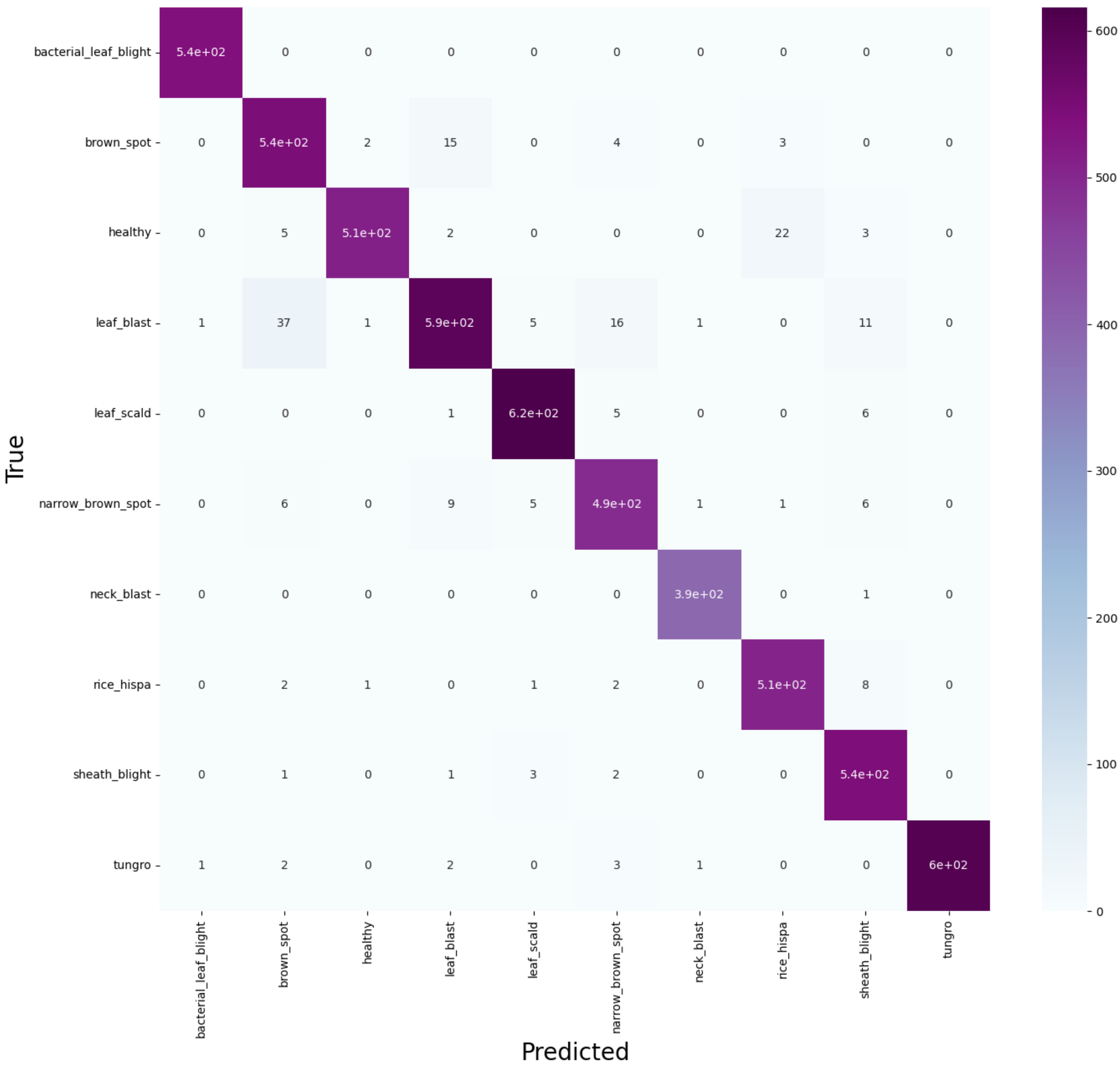


Figure 6.4: Confusion Matrix of Fine Tuning CNN with Swin Transformer Architecture.

6.4.3 Model Performance with Precision, Recall and F1-Score

Disease Leaves	Precision	Recall	F1-score	Support
bacterial_leaf_blight	0.99	0.97	0.98	586
brown_spot	0.87	0.72	0.79	480
healthy	0.85	0.88	0.86	600
leaf_blast	0.74	0.80	0.77	600
leaf_scald	0.94	0.94	0.94	576
narrow_brown_spot	0.89	0.83	0.86	504
neck_blast	1.00	0.96	0.98	336
rice_hispa	0.94	0.89	0.92	512
sheath_blight	0.82	0.96	0.89	528
tungro	0.97	0.99	0.98	648
accuracy			0.9183	5370
macro avg	0.90	0.89	0.90	5370
weighted avg	0.90	0.89	0.89	5370

Table 6.2: Precision, Racall and F1-score of Pre-trained VGG-16 with Proposed Dataset.

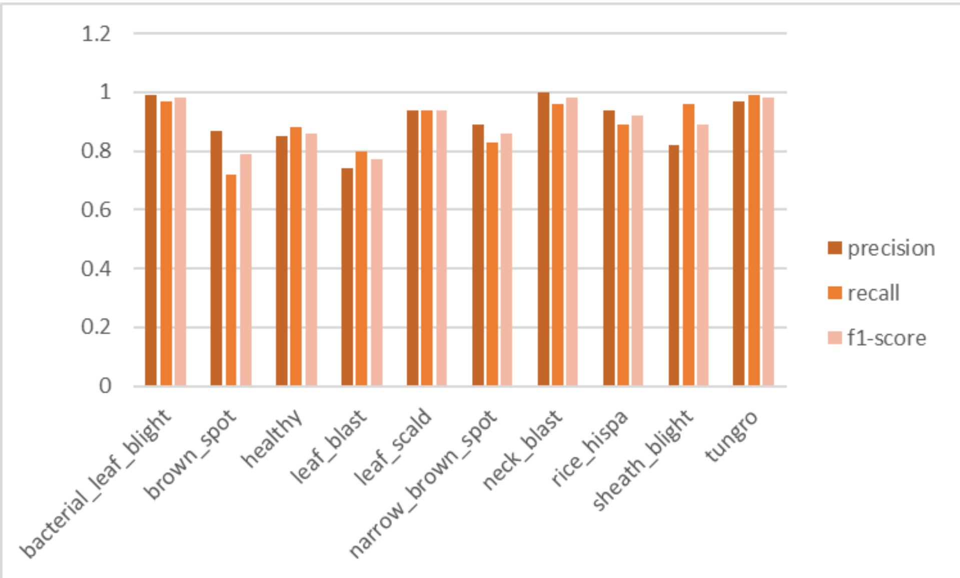


Figure 6.5: Precision, Racall and F1-score bar graph of Pre-trained VGG-16 with Proposed Dataset.

Disease Leaves	Precision	Recall	F1-score	Support
bacterial_leaf_blight	1.00	1.00	1.00	537
brown_spot	0.91	0.96	0.93	568
healthy	0.99	0.94	0.97	543
leaf_blast	0.95	0.89	0.92	664
leaf_scald	0.98	0.98	0.98	628
narrow_brown_spot	0.94	0.95	0.94	522
neck_blast	0.99	1.00	0.99	391
rice_hispa	0.95	0.97	0.96	520
sheath_blight	0.94	0.99	0.96	550
tungro	1.00	0.99	0.99	611
accuracy			0.96	5534
macro avg	0.97	0.97	0.97	5534
weighted avg	0.96	0.96	0.96	5534

Table 6.3: Precision, Racall and F1-score of Fine Tuning CNN with Swin Transformer on Proposed Dataset.

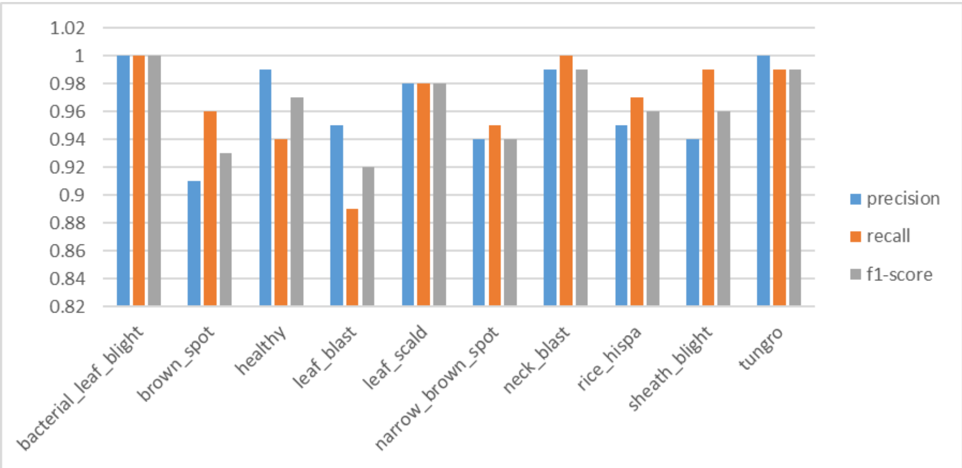


Figure 6.6: Precision, Racall and F1-score bar graph of Fine Tuning CNN with Swin Transformer architecture on Proposed Dataset.

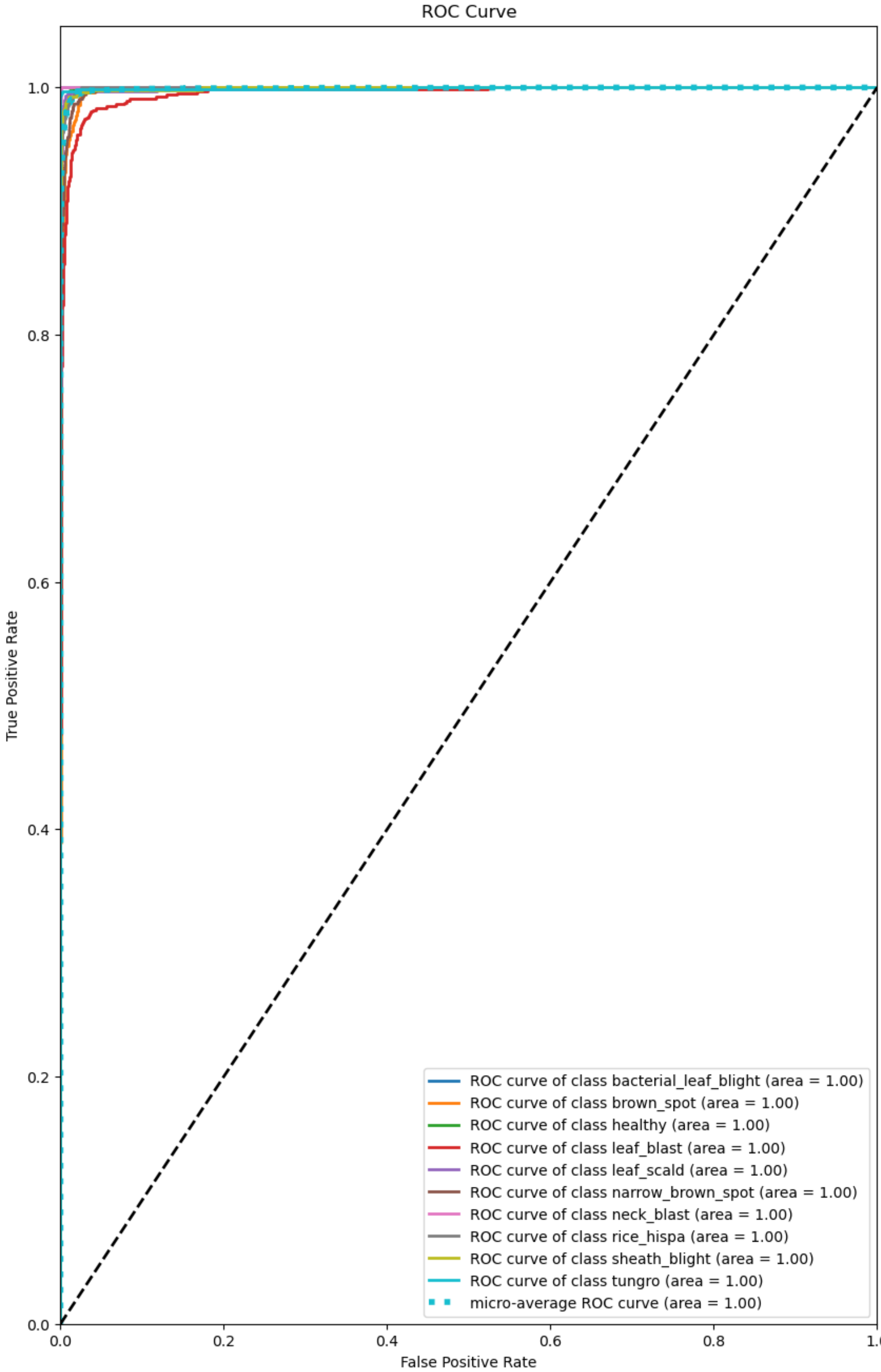


Figure 6.7: ROC curve of proposed method.

6.4.4 Model Comparison

Criteria	Shreya et.al (2020)[1] VGG-16	VGG-16	swin transformer
Used Dataset(images)	1649	18445	18445
Train Accuracy	97%	92.42%	97.21%
Test Accuracy	92.40%	91.83%	96.40%
Train loss	–	0.2208	0.1660
Validation loss	–	0.3191	0.1970
Total parameters	–	211 M	1955 M

Table 6.4: Performance Comparison between existing and proposed model.

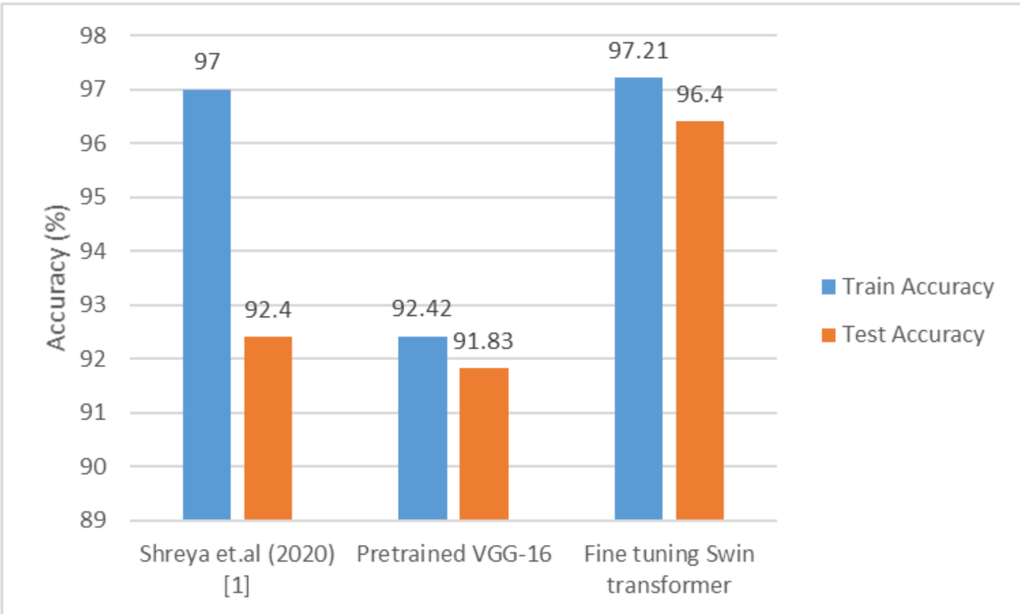


Figure 6.8: Bar graph representation of existing model and proposed model.

6.5 Performance Analysis

Our study investigated the use of a pre-trained VGG-16 convolutional neural network with transfer learning and fine tuning CNN with Swin Transformer architecture for classifying rice leaf diseases. We employed a large dataset exceeding the size used in the base paper. Ten major rice plant diseases were included in the analysis. The results, visualized in Figures 6.5, 6.6, 6.8, and table 6.4 demonstrate that the Swin Transformer achieved better performance. It reached a peak training accuracy of 97.21% and testing accuracy of 96.40% within 50 epochs. Swin transformer gave better result because their hierarchical processing approach allows them to effectively capture long-range dependencies within the images, which can be crucial for achieving better performance. However, this improvement comes at the cost of increased computational demands, as the Swin Transformer is a more intricate and parameter-intensive model compared to the pre-trained VGG-16 with transfer learning.

6.6 Conclusion

This chapter incorporates the utilization of the confusion matrix and various evaluation metrics to measure the model performance for all the models. By comparing the performance of the suggested deep convolutional approach with relevant architectures, the findings demonstrate that the proposed approach outperformed from the base paper method.

Chapter 7

Conclusion and Future Works

7.1 Introduction

This chapter provides a comprehensive summary of the research. It dives into the problem we addressed, what others have done before us, and the unique contributions we bring. We then detail our experiments and the conclusions we reached. Finally, we offer a glimpse into exciting possibilities for future exploration.

7.2 Thesis Summary

This research investigates how dataset size and model architecture influence CNN performance for computer vision tasks. We emphasize data quality using Seaborn to ensure balanced classes. Data preprocessing guarantees consistent formatting and avoids compatibility issues, while augmentation expands the training data for robustness. Normalization prevents bias towards features with large initial values. Shreya Kamal et al. [1] achieved 92% accuracy on a small dataset (1649 images) using VGG-16, highlighting its limitations with restricted data. We implemented their VGG-16 architecture on a medium-sized dataset (18445 images) but saw minimal improvement, suggesting it struggles with larger datasets. To address this, we implemented a novel CNN-Swin Transformer hybrid model. This model achieved significantly better results (97.21% training accuracy, 96.40% testing accuracy) even with the larger dataset. The hybrid model integrates Swin Transformer blocks into the CNN backbone, enabling local feature extraction via CNN layers and global context capture through the attention mechanism of the Swin Transformer, potentially leading to superior performance. However, this increased accuracy comes at the cost of using more parameters compared to VGG-16, as Swin Transformers rely on an attention mechanism for long-range dependencies.

7.3 Limitations

Though the proposed model outperformed, it also possesses some limitations. For hyperparameter tuning, this paper used hybrid swin transformer architecture, which has several limitations:

- While the hybrid CNN-Swin Transformer architecture offers significant accuracy improvements, its performance might suffer when trained and tested directly on unprocessed field images.
- Training and running the model requires more computational resources due to the additional parameters and operations involved in the Swin Transformer blocks.
- With a more complex model, there's a higher risk of overfitting, especially with small size datasets. Careful regularization techniques are crucial to mitigate this.
- The larger number of parameters in the hybrid model translates to a larger model footprint. This can be a concern for devices with limited memory resources.
- raining a complex model like a hybrid architecture can be more time-consuming and require careful hyperparameter tuning to achieve optimal performance.

7.4 Future Works

There are numerous possibilities to expand the scope of this study in the future. Here are a few examples of potential tasks:

- The main focus will be to evaluate the model performance by using dataset that contains large number of direct field images.
- We will try to evaluate the performance of proposed method on images of rice fields that are affected by multiple diseases simultaneously.

- The goal is to improve the accuracy and precision by exploring different configurations for the swin transformer blocks within the hybrid model. Experiment with the number of stages, transformer heads and window sizes to potentially improve performance.
- Can be considered other approaches for better performance.

If circumstances allow, we aspire to make a contribution to these challenges in the coming time, with the goal of enhancing the model to be more versatile and widely applicable than ever before.

7.5 Conclusion

This study carefully considered difficulties to show that fine-tuning how data is prepared (pre-processing) can really make a difference in how well deep learning models work. We found that the way data is set up is linked to how well the model does. Because of this, We recommend a more complete approach that considers all the steps - getting the data ready, using it to build the model, and checking how well it works. This way of doing things helps get the most out of data for making predictions

REFERENCES

[1] S. Ghosal and K. Sarkar, "Rice leaf diseases classification using cnn with transfer learning," in *2020 IEEE Calcutta Conference (Calcon)*, pp. 230–236, IEEE, 2020.

[2] L. Huang, F. Wang, Y. Zhang, and Q. Xu, "Fine-grained ship classification by combining cnn and swin transformer," *Remote Sensing*, vol. 14, no. 13, p. 3087, 2022.

[3] V. K. Shrivastava, M. K. Pradhan, S. Minz, and M. P. Thakur, "Rice plant disease classification using transfer learning of deep convolution neural network," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 42, pp. 631–635, 2019.

[4] Q. Yao, Z. Guan, Y. Zhou, J. Tang, Y. Hu, and B. Yang, "Application of support vector machine for detecting rice diseases using shape and color texture features," in *2009 international conference on engineering computation*, pp. 79–83, IEEE, 2009.

[5] R. Sharma, V. Kukreja, *et al.*, "Rice diseases detection using convolutional neural networks: a survey," in *2021 International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)*, pp. 995–1001, IEEE, 2021.

[6] S. S. H. Rumy, M. I. A. Hossain, F. Jahan, and T. Tanvin, "An iot based system with edge intelligence for rice leaf disease detection using machine learning," in *2021 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS)*, pp. 1–6, IEEE, 2021.

[7] K. Ahmed, T. R. Shahidi, S. M. I. Alam, and S. Momen, "Rice leaf disease detection using machine learning techniques," in *2019 International Conference on Sustainable Technologies for Industry 4.0 (STI)*, pp. 1–5, IEEE, 2019.

[8] S. Lamba, A. Baliyan, and V. Kukreja, "A novel gcl hybrid classification model for paddy diseases," *International Journal of Information Technology*, vol. 15, no. 2, pp. 1127–1136, 2023.

[9] S. Lamba, A. Baliyan, V. Kukreja, and R. Tripathy, “An ensemble (cnn-lstm) model for severity detection of bacterial blight rice disease,” in *Mobile Radio Communications and 5G Networks: Proceedings of Third MRCN 2022*, pp. 159–171, Springer, 2023.

[10] F. Zhou, H. Zhao, and Z. Nie, “Safety helmet detection based on yolov5,” in *2021 IEEE International conference on power electronics, computer applications (ICPECA)*, pp. 6–11, IEEE, 2021.

[11] F. Jubayer, J. A. Soeb, A. N. Mojumder, M. K. Paul, P. Barua, S. Kayshar, S. S. Akter, M. Rahman, and A. Islam, “Detection of mold on the food surface using yolov5,” *Current Research in Food Science*, vol. 4, pp. 724–728, 2021.

[12] M. E. Haque, A. Rahman, I. Junaeid, S. U. Hoque, and M. Paul, “Rice leaf disease classification and detection using yolov5,” *arXiv preprint arXiv:2209.01579*, 2022.

[13] M. Bala and V. Mehan, “Identification of rice plant diseases using image processing, machine learning & deep learning: a review,” in *CEUR workshop proceedings*, vol. 3058, pp. 0–3, 2021.

[14] K. Ahmed, T. R. Shahidi, S. M. I. Alam, and S. Momen, “Rice leaf disease detection using machine learning techniques,” in *2019 International Conference on Sustainable Technologies for Industry 4.0 (STI)*, pp. 1–5, IEEE, 2019.

[15] K. Kiratiratanapruk, P. Temniranrat, A. Kitvimonrat, W. Sinthupinyo, and S. Patarapuwadol, “Using deep learning techniques to detect rice diseases from images of rice fields,” in *International conference on industrial, engineering and other applications of applied intelligent systems*, pp. 225–237, Springer, 2020.

[16] M. H. Tunio, L. Jianping, M. H. F. Butt, and I. Memon, “Identification and classification of rice plant disease using hybrid transfer learning,” in *2021 18th International computer conference on wavelet active media technology and information processing (ICCWAMTIP)*, pp. 525–529, IEEE, 2021.

[17] Z. Zhang, Z. Gong, Q. Hong, and L. Jiang, “Swin-transformer based classification for rice diseases recognition,” in *2021 International Conference on Computer Information Science and Artificial Intelligence (CISAI)*, pp. 153–156, IEEE, 2021.

[18] L. Seth *et al.*, “Studies on the false-smut disease of paddy caused by ustilaginoidea virens (cke.) tak.,” *Indian Journal of Agricultural Science*, vol. 15, no. 1, 1945.

[19] V. K. Shrivastava, M. K. Pradhan, and M. P. Thakur, “Application of pre-trained deep convolutional neural networks for rice plant disease classification,” in *2021 international conference on artificial intelligence and smart systems (ICAIS)*, pp. 1023–1030, IEEE, 2021.

[20] M. Ahmad, M. Abdullah, H. Moon, and D. Han, “Plant disease detection in imbalanced datasets using efficient convolutional neural networks with stepwise transfer learning,” *IEEE Access*, vol. 9, pp. 140565–140580, 2021.

[21] S. Kaur, S. Pandey, and S. Goel, “Plants disease identification and classification through leaf images: A survey,” *Archives of Computational Methods in Engineering*, vol. 26, pp. 507–530, 2019.

[22] N. Jannat, S. M. Hasan, A. Y. Srizon, M. Oishe, A. H. Efat, M. F. Taraque, M. Mitu, and M. F. Faruk, “Efficient detection of crop leaf diseases: A lightweight convolutional neural network approach for enhanced agricultural productivity,” in *2023 International Conference on Information and Communication Technology for Sustainable Development (ICICT4SD)*, pp. 99–103, IEEE, 2023.

[23] K. Seebold Jr, L. Datnoff, F. J. Correa-Victoria, T. Kucharek, and G. Snyder, “Effects of silicon and fungicides on the control of leaf and neck blast in upland rice,” *Plant Disease*, vol. 88, no. 3, pp. 253–258, 2004.

[24] C. R. Rahman, P. S. Arko, M. E. Ali, M. A. I. Khan, S. H. Apon, F. Nowrin, and A. Wasif, “Identification and recognition of rice diseases and pests using convolutional neural networks,” *Biosystems Engineering*, vol. 194, pp. 112–120, 2020.

[25] C. R. Rahman, P. S. Arko, M. E. Ali, M. A. I. Khan, S. H. Apon, F. Nowrin, and A. Wasif, “Identification and recognition of rice diseases and pests using convolutional neural networks,” *Biosystems Engineering*, vol. 194, pp. 112–120, 2020.

[26] R. Nandakumar, S. Babu, R. Viswanathan, T. Raguchander, and R. Samiyappan, “Induction of systemic resistance in rice against sheath blight disease by pseudomonas fluorescens,” *Soil Biology and Biochemistry*, vol. 33, no. 4-5, pp. 603–612, 2001.

[27] V. K. Shrivastava, M. K. Pradhan, S. Minz, and M. P. Thakur, “Rice plant disease classification using transfer learning of deep convolution neural network,” *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 42, pp. 631–635, 2019.

[28] H. Bunawan, L. Dusik, S. N. Bunawan, and N. M. Amin, “Rice tungro disease: From identification to disease control,” *World Appl. Sci. J*, vol. 31, no. 6, pp. 1221–1226, 2014.

[29] J. T. Springenberg, A. Dosovitskiy, T. Brox, and M. Riedmiller, “Striving for simplicity: The all convolutional net,” *arXiv preprint arXiv:1412.6806*, 2014.

[30] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” *arXiv preprint arXiv:1409.1556*, 2014.

[31] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016.

[32] A. Veit, M. J. Wilber, and S. Belongie, “Residual networks behave like ensembles of relatively shallow networks,” *Advances in neural information processing systems*, vol. 29, 2016.

[33] S. Bhujel and S. Shakya, “Rice leaf diseases classification using discriminative fine tuning and clr on efficientnet,” *Journal of Soft Computing Paradigm*, vol. 4, no. 3, pp. 172–187, 2022.