University of Hertfordshire

School of Physics, Engineering and Computer Science

Advanced Computer Science Masters Project

7COM1039-0901-2024

Project Title:

Enhancing Rice Crop Health Monitoring: CNN-Based Detection of Rice Leaf Diseases

By

Danish Jalil

Student ID

22016088

Supervisor: *Dr. Muawya eldaw*

Date Submitted: *01/06/2025*

Word Count: *9136*

GitHub address: *Click here to access GitHub*

DECLARATION STATEMENT

This report is submitted in partial fulfilment of the requirement for the degree of Master of Science in Data Science at the University of Hertfordshire. I have read the guidance to students on academic integrity, misconduct and plagiarism information at Assessment Offences and Academic Misconduct and understand the University process of dealing with suspected cases of academic misconduct and the possible penalties, which could include failing the project module or course. I certify that the work submitted is my own and that any material derived or quoted from published or unpublished work of other persons has been duly acknowledged. (Ref. UPR AS/C/6.1, section 7 and UPR AS/C/5, section 3.6). I have not used chatGPT, or any other generative AI tool, to write the report or code (other than where declared or referenced). I did not use human participants or undertake a survey in my MSc Project.

I hereby give permission for the report to be made available on module websites provided the source is acknowledged.

Student SRN number: 22016088

Student Name printed: Danish Jalil

Student signature: Danish Jalil

DEDICATION

It is with great pleasure that I dedicate this research to my family, my colleagues, and my friends, who have continually helped me throughout the whole process of this research with encouragement and unwavering belief. Their support was especially valuable during the times when I felt discouraged and challenged in my work.

ACKNOWLEDGMENTS

It is my sincere pleasure to express my sincere gratitude to my thesis supervisor, Dr. Muawya eldaw, for providing me with steadfast guidance as well as inspiring mentoring throughout the course of my research. In addition, I must acknowledge the incredible support of my family, whose encouragement and motivation have been crucial to bringing this research to fruition. His enthusiasm, patience, and insightful feedback have greatly influenced my work.

Abstract

The rice leaf images were classified using convolutional neural networks (CNNs) and MobileNets based on factors such as bacterial leaf blight, brown spots, and leaf smut. As a result of the classification of 120 high-resolution images into three different types of diseases, we were able to identify the three diseases: Bacterial Leaf Blight, Brown Spots, and Leaf Smut. On the training data, both models achieved 100% accuracy, with CNN exhibiting a significantly lower training loss, suggesting overfitting to the model. As compared to CNN, MobileNet achieved better generalization on the validation set, achieving 91.67% accuracy. While the CNN model excelled at detecting Bacterial Leaf Blight, it struggled to detect Brown Spot and Leaf Smut, exhibiting lower recall and F1-scores for these diseases. For Brown Spot and Leaf Smut, MobileNet showed balanced and higher performance across all disease categories. The results of this comparative analysis demonstrate MobileNet's superior ability to generalize to unseen data and its potential for real-world applications in automating rice disease diagnosis. Agricultural automation, specifically the detection of rice diseases, may benefit from the advancement of deep learning techniques, such as MobileNet.

Keyword : Agricultural automation, image processing, deep learning, rice disease detection, MobileNet, CNN

CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. CHAPTER I: INTRODUCTION

1.1. Problem Overview

Rice is a crucial source of nutrition globally, supporting food security and the economies of many nations(Bandumula 2018). However, agricultural challenges, particularly diseases affecting rice plants, pose significant risks to productivity. In countries like Bangladesh and India, where the majority of the population is dependent on agriculture, rice diseases such as Brown Spot, Tungro, Sheath Blight, Bacterial Leaf Blight, and Bacterial Leaf Blast are major concerns (N. Kumar et al. 2021; Mukherjee et al. 2025; Zhao, Lin, and Chen 2020). The prevalence of these diseases threatens crop yield, quality, and food security, especially in rural areas where farmers struggle to identify and manage these issues.The difficulty in diagnosing rice leaf diseases is compounded by the overlapping symptoms of various diseases, often leading to confusion and ineffective treatments(Strange and Scott 2005). Moreover, climatic factors exacerbate the situation, making timely identification and intervention even more difficult. The need for a precise and reliable diagnostic method for early disease detection in rice crops has become paramount. Agricultural performs can be transformed through AI-driven tools that deliver accurate and effectual disease classification and running, which can convert agricultural performs in this context (Organization et al. 2013).

1.2. Current Issues

The rice segment faces several dangerous challenges connected to crop illnesses, which meaningfully delay yield and excellence. In republics like Bangladesh and India, where rice is important for both food and financial stability, diseases such as bacterial blight, fungal infections, and other leaf diseases lead to substantial losses(Fahad et al. 2019). Early detection and exact analysis are vital to extenuating the blowout of these illnesses and minimalizing the financial impact. Insufficient disease documentation methods, such as graphic reviews by farmers, are time-consuming and often imprecise. This is mainly problematic for large-scale rice agricultural, where physical monitoring is not feasible(Buja et al. 2021). Furthermore, the impact of climate change and increased ecological pressures have made the rice crop more vulnerable to illnesses. Without quick and precise diagnosis, disease occurrences can consequence in abridged crop yields, monetary loss, and food security pressures.

1.3. Project Details

Rice leaf illnesses require deep learning models, such as Convolutional Neural Networks (CNNs) and MobileNet. This study will evaluate CNNs based on images of Brown Spot, Leaf Spot, and Bacterial Blight on rice plants on Kaggle(Hasan et al. 2023). The perfect will contribution in early detection and organisation of disease by sleuthing brown spots, black dots, and water-soaked cuts. As deep learning techniques have increased purchase in agriculture, they are able to deliver faster, more accurate, and automated identifies, which decreases the need for expert information (Wani et al. 2022). For classifying complex designs in rice leaf imageries, CNNs excel at image classification errands.

1.4. Aims and Objectives

Using CNNs and mobilnet for initial disease detection and organisation is the aim of this study. As a consequence of this study, the next objectives will be attained:

1. To provide an impression of the trials associated with rice leaf disease detection and its significance for agriculture and food safety.
2. To design and implement a CNNs-based model for identifying and classifying rice diseases based on Rice leaf images.
3. To review and evaluate current methods of plant disease identification using deep learning methods.
4. To analyze the performance of the CNN model trained on a Kaggle rice leaf disease dataset, comparison it to traditional investigative methods.
5. To assess the potential benefits of AI-driven solutions in improving disease management and mitigating crop losses.

1.5. Research Questions

It is the aim of this study to address the following research questions:

1. Can deep learning models accurately detect sick leaves in rice plants?
2. How can deep learning-based disease detection aid in the diagnosis of rice leaf diseases in agricultural practices?
3. What is the potential impact of this technology on the agricultural sector and the broader economy?

1.6. Motivation of This Study

Farmers have long relied on visual inspection and expert knowledge to identify and manage plant diseases, a time-consuming, error-prone, and inefficient procedure. Plant diseases are increasing due to climate change and environmental stressors, complicating traditional diagnostic methods (John et al. 2023).It is consequently imperative that farmers are helped in detecting and treating crop diseases in a timely manner. Deep learning and computer vision can help resolve these challenges (Amulothu et al. n.d.). Models using deep learning can process large amounts of data and detect disease suggestions accurately, reducing Labor and costs. By means of a rapid, reliable, and available disease detection system, this study aims to improve agricultural does, improve crop health, and decrease financial losses (Bao et al. 2019).

1.7. Novelty of the Study

In this study, progressive deep learning methods, precisely CNNs, are applied to Kaggle datasets of rice leaf images to detect rice leaf disease. Founded on visual symptoms, the model can accurately categorise various rice diseases (Upadhyay and A. Kumar 2022). This method offers farmers a scalable, automatic, and real-time solution that decreases crop losses due to undiagnosed illnesses (Jafar et al. 2024). Furthermore, this investigate donates to the growing body of information on AI applications in agriculture, chiefly in rice farming (Ahmad et al. 2024).

1.8. Final Feasibility, Commercial Context, and Risk

The possibility of this profound learning-based illness detection system is talented, as CNNs have confirmed strong presentation in alike image organisation tasks(Razzak, Naz, and Zaib 2018). The model, once industrialised and tested, can be organized through mobile apps or other stages that allow farmers to take images of their crops and obtain instant disease identifies. This could significantly decrease the time and capitals spent on disease organisation and improve the overall output of rice farming(Singh 2018).In footings of commercialization, this knowledge has the potential to be climbed across various agricultural sectors, counting in republics where rice is a major crop. It could be advertised as part of a exactness agriculture platform, contribution not just disease detection but also references for treatment, pest control, and crop management(Awotide, Karimov, and Diagne 2016). The model, however, expressions risks such as its accuracy under dissimilar environmental circumstances, the need to train on a great and diverse dataset, and its early cost. Ensuring that the model is flexible to numerous climatic conditions and user-friendly for farmers with incomplete technical knowhow will be indispensable for its success in the field.

1.9. Commercial Context

Given that over 80% of the population in countries like India and Bangladesh depend on agriculture, the introduction of AI-driven solutions for disease detection could lead to significant improvements in agricultural productivity and food security(Dhal and Kar 2024; Mohajan 2013). The ability to detect and manage rice leaf diseases quickly can help prevent widespread crop losses and ensure a steady food supply(Balasubramanian et al. 2007).The commercial potential for such AI-based solutions is substantial, not only in rice farming but also in other crops and regions, as the technology could be adapted for various agricultural contexts worldwide.

1.10. Report structure

This report is structured into the following key sections. The **Literature** section reviews existing research, identifying gaps and challenges that the study aims to address. The **Methodology** outlines the approach, techniques, and tools used, providing a detailed explanation of data collection, analysis, and any models or frameworks applied. The **Experimental Results** present the findings of the study, supported by tables, graphs, or figures, and discuss observed patterns and their implications. Finally, the **Conclusion** summarizes the key outcomes, highlights contributions, acknowledges limitations, and suggests directions for future research. This structure ensures a logical flow from existing knowledge to new insights.

2. CHAPTER II: LITERATURE REVIEW

2.1. RELATED WORK

2.2. Introduction to the Field

There are many diseases that affect the color, size, or form of rice around the world, which reduces its productivity. It is estimated that rice diseases contribute to a 37 percent decrease in rice production decline in rice productivity every year, according to Dhaka et al. (Dhaka et al. 2021). These diseases are not only difficult to diagnose, but their detection and management are also poorly understood. Machine learning and deep learning have made it possible to identify and manage rice diseases efficiently. The aim of this section is to review research on deep learning models applied to rice plant disease detection and diagnosis. Even though an appropriate method for detecting and controlling rice illnesses has not yet been developed, there is still a lack of knowledge about the identification and management of these illnesses. Despite this, very few studies have been conducted on diagnosing rice plant diseases. It also discusses new research on deep learning models for rice disease diagnosis (Sethy et al. 2020b).

2.3. Key Studies and Works

The following research highlights various approaches and models used for detecting rice plant diseases:

The Faster R-CNN algorithm was used by Bari et al. (Bari et al. 2021) to distinguish among rice illnesses including blast, brown spot, and Hispa. They were able to determine leaves that were healthy with an accuracy of 99.25%. Both a self-generated dataset and a Kaggle dataset with 2,400 photos in total were used to assess the models. In addition, characteristic maps of diseased leaves were employed for training in a Caffe deep learning technique. An internet database and their own dataset were used to generate a unique Rice Leaf Disease Dataset (RLDD).

Pandian et al.(Pandian J et al. 2022) ResNet19 layers were optimized using evolutionary search. We used data augmentation techniques such as growing, trimming, tumbling, padding, and rotation to add more data to pre-existing photographs. A ResNet197 model was trained using 154,500 photos from 22 different plants, including both normal and sick leaves. The input images were 224 by 224 $\times$ 3 pixels. The model achieved 99.58% classification accuracy after training for up to 1,000 epochs in a GPU environment.

A framework for prompt identification of illnesses was created by Roy and Bhaduri(Roy and Bhaduri 2021), who aimed to maximise the rapidnessand precision of identifying apple illnesses in intricate orchard settings. Their simulation was processed at 56.9 frames per second

(FPS) and had a mean average precision (mAP) of 91.2% and an F1 score of 95.9%. With a mAP gain of 9% and an F1 score improvement of 7.6%, this method demonstrated notable improvements over earlier models.

Saberi Anari (Saberi Anari 2022) diagnosed damaged leaves with Model Engineering. Using deep transfer learning, this method combines radial basis functions, k-NN, decision trees, neural networks, and ensembles to obtain characteristics of images. Specifically, deep transfer learning was used to extract characteristics from pictures of leaves of different fruits.

In order to train and evaluate 5 separate models, Deng et al. (Deng et al. 2021) gathered 33,026 photos of six distinct rice illnesses of the leaves. The methodologies, dataset, camera specs, amount of findings, training rate, frequency of rounds, and model efficacy were all thoroughly described in the paper.

To diagnose rice leaf illnesses, Zhou et al. (Zhou et al. 2023) used Faster R-CNN fusion in conjunction with FCM-KM clustering. 3,010 photos made up the dataset, which had 15,000 rounds and a learning rate of 0.001. For rice explosion, microbial decay, and pollution, the model's accuracy was 96.71%, 97.53%, and 98.26%, respectively. Sethy et al.(Sethy et al. 2020a) used a dataset collected with a mobile phone camera in a farm area for identifying rice fake smuts utilising a Faster R-CNN model. The effectiveness of the model was not disclosed, however the dataset had 50 assessments, 5 rounds, and a learning rate of 0.001.

An improved method was used by Ramesh and Vydeki (Ramesh and Vydeki 2020) to diagnose rice leaf illnesses, and the dataset was also taken from an agricultural area. Rice explosion, microbiological destruction, sheathing rot, and brown spot all had rates of success of 98.9%, 95.78%, and 94%, respectively, for the model they used. A Faster R-CNN model was employed by Deng et al. (2021) and Li et al. (2017) to identify rice illnesses of the leaves using a dataset gathered from crops of rice in China's Anhui and Hunan regions. Both a Sony DSC-QX10 camera and a cell phone camera were used to record the data. 5,320 photos made up the dataset, which had 50,000 rounds and a learning rate of 0.002.

Prajapati et al. (Prajapati, Shah, and Dabhi 2017) Rice leaf disease was diagnosed using NIKON D90 digital SLR camera datasets taken on farmland. The acquisition rate and number of rounds were not specified, but 120 samples were collected. As a result of training and testing, the model gave 93.33 percent, 73.3 percent, 83.80 percent, and 88.57 percent during cross-validation five times. The CNN model was employed by Rahman et al. (2020) for accurate categorisation rice leaf disease diagnosis.

Velusamy et al. (Velusamy et al. 2023) used handmade feature development using segmentation, augmentation, and pre-processing to improve the precision of classification. With this method, the final classification accuracy was 90.63%, an improvement of 3.1%.

In order to lessen dependency on conventional methods for protecting the production of rice, Haridasan et al. (Haridasan, Thomas, and Raj 2023) Machine learning, deep learning, and image processing can be combined to diagnose rice plant diseases automatically. The rice plant's afflicted regions are pinpointed using segmentation methods after picture preprocessing.

Visual traits detect disorders. Rice plant diseases are classified and diagnosed using Support Vector Machine (SVM) and Convolutional Neural Network (CNN) classifiers.

2.4. Relation to Your Research and Hypothesis

The summarized studies provide a foundation for understanding current advancements in rice plant disease diagnosis using machine learning and deep learning models. The high accuracy achieved by models like Faster R-CNN and ResNet19 highlights the potential of leveraging sophisticated algorithms to enhance detection performance. However, the variability in dataset sizes, preprocessing techniques, and performance metrics indicates a need for standardized approaches. Additionally, the reliance on specific datasets and limited exploration of hybrid methods suggests opportunities for further innovation.

2.5. Appropriate Sources and Quality

The studies referenced in this research are sourced from reputable journals, conferences, and datasets, ensuring the credibility and reliability of the information. Each source has been selected based on its relevance to rice plant disease diagnosis and the quality of its methodology and results. High-impact studies, such as those employing advanced models like Faster R-CNN and ResNet19, have been included to highlight the current state-of-the-art approaches in the field. Table 2.1 summarizes the existing literature along with their achievements and limitations.

Table 2.1. Summary of Literature on Rice Plant Disease Diagnosis

Author(s)	**Algorithm/Model Used**	**Dataset Details**	**Techniques**	**Performance**
(Bari et al. 2021)	Faster R-CNN	Self-generated and Kaggle datasets with 2,400 images	Characteristic maps, Caffe deep learning	Accuracy: 99.25%
(Pandian J et al. 2022)	ResNet19 with evolutionary search	Pooled dataset with 154,500 images from 22 plants	Data augmentation (growing, trimming, rotation), pre-trained ResNet19, GPU training	Accuracy: 99.58%
(Roy and Bhaduri 2021)	Deep learning framework	Orchard setting dataset (details unspecified)	High FPS processing (56.9), precision-focused training	mAP: 91.2%, F1: 95.9%

Author(s)	Algorithm/Model Used	Dataset Details	Techniques	Performance
(Saberi Anari 2022)	Deep CNN + SVM models (e.g., RBF, k-NN)	Dataset details unspecified	Deep transfer learning, ensemble SVM models	Performance not specified
(Deng et al. 2021)	Various models	33,026 images of six rice leaf diseases	Thorough methodology description, various camera specs	Performance not specified
(Zhou et al. 2023)	Faster R-CNN + FCM-KM clustering	Dataset of 3,010 images with 15,000 rounds	Hybrid clustering approach	Accuracy: 96.71%-98.26%
(Sethy et al. 2020a)	Faster R-CNN	Mobile camera farm dataset	Minimal rounds (5), learning rate 0.001	Performance not disclosed
(Ramesh and Vydeki 2020)	Improved Faster R-CNN	Agricultural dataset	Comprehensive analysis	Accuracy: 94%-98.9%
(Prajapati, Shah, and Dabhi 2017)	SVM	Nikon D90-based farmland dataset of 120 samples	5/10-fold cross-validation	Accuracy: 93.33%-88.57%
(Velusamy et al. 2023)	CNN-based watershed segmentation	Dataset details unspecified	Handcrafted feature development, segmentation	Accuracy: 90.63%
(Haridasan, Thomas, and Raj 2023)	Automated vision-based approach	Dataset details unspecified	Image preprocessing, segmentation, hybrid SVM-CNN	Performance not specified

2.6. Depth and Breadth of Coverage

The literature reveals that various methodologies, datasets, and technologies have been applied to rice disease diagnosis, showcasing depth and breadth in approaches:

- Studies like that of Zhou et al. (Zhou et al. 2023) incorporated advanced clustering methods (FCM-KM) alongside Faster R-CNN to achieve accuracies exceeding 96%.
- Deng et al. (Deng et al. 2021) explored the use of multiple models trained on a large dataset of 33,026 images, providing a detailed description of the methodologies, camera specifications, and dataset variations.
- The work by Haridasan et al. (Haridasan, Thomas, and Raj 2023) demonstrates the

integration of machine learning and deep learning methods with preprocessing and segmentation to enhance precision in disease detection.

2.7. Comparative Analysis

Comparing the models and methodologies reveals both strengths and limitations:

2.7.1. Strengths

- Models like ResNet19 (Pandian J et al. 2022) and Faster R-CNN (Bari et al. 2021) exhibit high classification accuracy due to robust architectures and comprehensive datasets.
- Techniques such as data augmentation and clustering have significantly enhanced accuracy in disease diagnosis (Zhou et al. 2023).

2.7.2. Weaknesses

- Some studies lack detailed performance metrics or dataset descriptions, which hampers replicability and generalization.

2.8. Identification of Gaps

Despite significant advancements, several gaps remain in the current research:

1. Many studies lack severity-based classification methods, which could aid in better disease management strategies.
2. A comprehensive, standardized dataset for rice disease detection is absent, leading to variations in model performance and generalizability.
3. Few models address real-time deployment in field conditions, as highlighted by the limited FPS in simulations like Roy and Bhaduri's work (Roy and Bhaduri 2021).