Nitrogen Level Prediction Using LCC

MOHAMED RAFEEK MAREER AHAMED

Submitted to the

Computer School

In partial fulfilment of the requirements for the Degree of

Bachelor of Science in

Computer Science (Hons)

Supervised By:

Prof. Guhanathan Poravi

Staffordshire University

April 2024, Colombo

Abstract

This research aims at calibrating the LCC for nitrogen level estimation in rice crops. Nitrogen is an important nutrient for plant growth and, when deficient, can affect crop production greatly. The current conventional approach of using LCC for nitrogen quantification is not efficient, timely, and quantitative and gives out inconclusive results in most cases. In order to overcome this problem, a CNN-based system is proposed in this paper to automatically predict the nitrogen level. Our system then uses RGB images of rice paddies and uses convolutional neural networks to further classify the nitrogen levels based on the LCC scale, with an emphasis on Levels 2, 3, 4, and 5.

The problem was tried to solve by integrating a CNN-based deep learning model to recognize the problem. The CNN was chosen since it can learn and identify features from the image data well for nitrogen level classification. The structure has what is called convolutional, pooling, and fully connected layers, which are intended to extract some distinctive features from RGB images. As inputs, specific image preprocessing techniques were applied, and making it easier for the model to forecast better; this training was performed using the LCC dataset.

The efficiency of the system was measured using the conventional assessment indicators of a data scientist. The model showed an accuracy of 86% with the help of VGGNet 16, whereas the accuracy achieved in Inception V3 and ResNet 50 is comparatively low. Accuracy, sensitivity, specificity, and the F1 score were used to determine the classification metrics. Furthermore, accuracy and facility classification performance for predicting nitrogen levels were evaluated using the ROC functions and AUC values.

Keywords: Artificial Intelligence(AI), Convolutional Neural Network(CNN), Leaf Color Chart(LCC), Machine Learning(ML), Red, Green, Blue(RGB), Social, Legal, Ethical, Professional(SLEP), Receiver Operating Characteristic(ROC), Area Under the Curve(AUC), Learning Outcome(LO)

Subject Descriptors: Artificial intelligence and machine learning, deep learning technology, agricultural technology, precision agriculture, image classification.

Declaration

This dissertation works titled Nitrogen Level Prediction Using LCC is the original work of the author, and the procedure followed for the research done has been expounded in the following chapter. All the information gathered from other sources is academic, and where used, it has been cited appropriately. Every possibility has been considered to avoid any form of plagiarism. In relation to this, there has been no infringement on any copyright materials outside this work, and wherever there is information that has been taken from other sources, the citations have been done as is required in any academic paper.

Acknowledgement

The author would like to thank the supervisor, Mr. Gurunathan Poravi, for the assistance, encouragement, and directions he has given throughout the completion of this study. The constructive criticism with the professors and peers at APIIT has also contributed significantly to this project; revelations accrued from the discussions as well.

Table of Contents

CHAPTER 1 : PROBLEM

1.1 Chapter Overview

This chapter begins with the global importance of nitrogen in rice production and discusses current practice approaches to nitrogen management. It will compare the inadequacies of such conventional techniques as the Leaf Color Chart (LCC) and stress the importance of accurate nitrogen quantification. The position of the chapter will also state the purpose of research to enhance the accuracy and efficiency of nitrogen management with the help of a machine learning approach (CNN). Apart from that, it will present the rationale for the research and the issues encountered while designing an automated system that would improve nitrogen efficiency as well as reduce the use of fertilizers in commercial rice production.

1.2 Introduction

Nitrogen, a critical nutrient in rice growing that directly influences crop yield and quality, is a focus of the research. Optimal rice production is difficult to achieve unless nitrogen is managed efficiently; however, traditional LCC-based methods for manual application tend to be inaccurate due to inconsistent nitrogen application across large fields. Here we can predict nitrogen levels such as LEVEL 02, LEVEL 03, LEVEL 04, and LEVEL 05 which means LCC standard nitrogen Levels, this study proposes a convolutional neural network (CNN)-based approach. This automated method leverages LCC data to improve nitrogen estimation accuracy and thereby mitigate over fertilization impact on the environment, while improving crop yield and supporting precision agriculture practices.

1.3 Problem Background

1. Rice is an essential nutrient, and a deficiency results in decreased crop growth, chlorophyll content, and overall yield. Decreased photosynthetic activity and increased leaf reflectance observed in rice plants due to low nitrogen availability affect the efficiency of food production (Illesinghe et al., 2023). In addition, rice plant growth stages at the critical periods have shown nitrogen deficiency that directly affects productivity. This deficiency needs to be addressed in

order to ensure optimal rice yields as well as to ensure sustainability in rice farming (Tao et al., 2020).

2. Traditional Nitrogen Monitoring Methods Limitations In recent years, the Leaf Colour Chart (LCC) has been widely promoted as a successfully used, inexpensive, and field-accessible device that estimates rice leaf nitrogen content by matching the colour of the leaf to standard reference colors (Tao et al., 2020). Despite these challenges, its manual observations rely on human error, inconsistent readings, and inefficient monitoring of large fields. Additionally, LCC accuracy can be affected by external factors such as lighting conditions as well as leaf positioning. These limitations highlight the importance of an automated, scalable solution (Islam et al., 2020).

3. Nitrogen Detection Using Machine Learning Advancements With the rise of technology, and specifically the area of convolutional neural networks (CNNs), promising results for the automation of nitrogen detection have been achieved. The integration of machine learning models with UAV (Unmanned Aerial Vehicle) systems has been proven as a potential solution for precision agriculture. Real-time nitrogen assessment across entire rice fields using these methods is possible and with better accuracy and scalability than traditional techniques (Illesinghe et al., 2023). Research on CNNs integration with LCC and UAVs in nitrogen detection seeks to relieve the burden of manual labor and errors in order to afford large, scalable solutions for rice farmers.

1.4 Problem Definition

Managing nitrogen in rice crops effectively is the crucial part of supporting the plant's good health, growth, and yield. While traditional methods of nitrogen estimation, such as the Leaf Color Chart (LCC), are widely used, they suffer from human error and environmental factors and are inefficient for wide areas (Islam et al., 2020). These methods are manual, time-consuming, and prone to inconsistencies, especially when we need to handle large-scale rice production (Illesinghe et al., 2023).

Recently, the deployment of modern high-performance technologies, such as convolutional neural networks (CNNs), has successfully dealt with these challenges by producing more accurate and automated solutions. This led CNNs to be integrated with UAV-based systems that can process images of rice crops to predict nitrogen levels, thus making a more scalable and efficient technique (Tao et al., 2020). Yet, current systems continue to have limitations in their accuracy under different field conditions and for different growth stages.

To fill this gap, this research develops a CNN-based system that utilizes UAV-captured images in combination with LCC standards to predict nitrogen levels with higher accuracy during varied growth stages of rice fields. The proposed system aims to enhance the precision and scalability of nitrogen management in rice agriculture through automating the nitrogen assessment process to improve the yields of the crops and reduce environmental damages in farming practices.

1.4.1 Problem Statement

Currently, the familiar practice of nitrogen input into rice fields using the Leaf Color Chart (LCC) is not efficient and easily subject to human error for fields on a large scale. The result is an inconsistent application of nitrogen that can affect yield in a crop and the environment. Current image-based techniques are limited in their precision across a range of growth stages and field conditions. Improving nitrogen prediction accuracy requires an automated solution using convolutional neural networks (CNNs) so that we can assess nitrogen in real time and on a on a large scale to support better crop management and better use of input fertilizer.

1.5 Research Motivation

The fact that the researcher has a background immersed in farm rice growing has helped him to have first-hand experience of the trouble involved in managing nutrient levels such as nitrogen and the resultant lower yield of the crop as a result of inappropriate application of fertilizer. As part of that, the motivation stemmed from personal experience. The researcher is motivated to bridge traditional farming with modern technology by developing a system based on CNN-based image processing that can predict nitrogen levels in rice fields. The aim of this project is to help farmers with a practical, low-cost tool to improve crop management and support sustainable farming in the future.

1.6 Existing Work

Citation	Summary	Limitation	Contribution
(Islam et al., 2020)	A CNN and Decision Tree model for automating the LCC for nitrogen fertilizer recommendation in rice leafs is proposed in this paper.	The method requires particular hardware and a considerable amount of computing power, which may not be available to small-scale farmers	The smart system for nitrogen recommendations serves as the paper's contribution, with the CNN model achieving 94.22% accuracy and offering a scaler solution to fertilizer management
(Illesinghe et al., 2023)	We use UAVs with a smartphone to extract high-resolution aerial images and require nitrogen demand in rice fields by utilizing LCC and machine learning.	The quality of the UAV camera can result in uncertainty in image capture, and further environmental parameters (e.g., lighting) can add noise to image capture.	Pathways to an 86.5% accuracy rate for a low-cost, field-wide nitrogen monitoring solution having potential for real-time applications in precision agriculture.
(Tao et al., 2020)	Describes the use of LCC and image processing techniques to propose a smartphone-based app to detect rice	May depend on quality of smartphone (and might not perform well under suboptimal lighting or	As nitrogen level detection achieves 96% accuracy, it provides a low-cost, portable solution for farmers to manage nitrogen

	leaf colour levels for use in nitrogen management.	environmental conditions).	in real time.
(Sethy et al., 2020)	In this study, we use CNN to detect nitrogen deficiency in rice crops by classifying rice leaf images according to nitrogen levels.	In fields with peripheral lighting conditions, the accuracy of the model may decrease.	With optimization using Nadam, it achieves 99.5% accuracy for nutrient classification in rice crops using CNN.
(Yuchen et al., 2023)	The CNN model for classifying nitrogen deficiency in rice leaves is implemented, and different optimizers are compared to find the best one.	The model is computationally intensive, and its performance is highly dependent on the optimal image resolution found.	A study of how changing optimizers improves model accuracy and gets to 99.5% accuracy rate with Nadam.
(Zhou et al., 2023)	A method for estimating leaf chlorophyll content using ResNet-18 with hyperspectral data for optimal nitrogen levels.	Hyperspectral data is required (a low price tag on hyperspectral data still means it is expensive and it may not be possible for any given farming setup).	A novel use of hyperspectral data and CNN for leaf chlorophyll content estimation is demonstrated, with high accuracy ($R^2 = 0.96$).
(Torikul et al., 2020)	A CNN-based system is developed to automate LCC-based nitrogen fertilizer recommendation for rice paddies.	Although the manual leaf color comparison with LCC still has errors to introduce, the system requires specialized mobile app deployment.	The CNN model for classifying leaf colors achieves an accuracy of 94.22%, contributing to more precise nitrogen fertilizer recommendations.

Table 1 Existing Wor

1.7 Research Gap

Much work has been accomplished in nitrogen management with machine learning and UAVs, there are still challenges specifically relating to the accuracy of nitrogen prediction systems based on the Leaf Color Chart (LCC). Current methods rely on the use of machine learning models, such as SVM, to classify nitrogen levels from LCC readings solely by average color values. Nevertheless, they are still usually incapable of embracing the complexity of the actual paddy field environment to reach optimal accuracy.

Additionally, most of the previous studies did not critically examine the possibility of the comparison of multiple convolutional neural networks (CNN) to increase the accuracy while studying the LCC data. Optimal opportunities for optimization are constrained by limiting the approach to a single model and basic image features. Furthermore, further improvement of nitrogen prediction could be made by using more advanced pre-processing methods and feature extraction methods.

In this research, we intend to fill these gaps by comparing and evaluating the CNN architectures with the LCC data to enhance the accuracy of nitrogen level detection. This study will help develop more precise and scalable nitrogen management systems for agriculture through the integration of sophisticated image processing techniques and feature selection methods.

1.8 Contribution to Body Knowledge

Much work has been accomplished in nitrogen management with machine learning and UAVs, there are still challenges specifically relating to the accuracy of nitrogen prediction systems based on the Leaf Color Chart (LCC). Current methods rely on the use of machine learning models, such as SVM, to classify nitrogen levels from LCC readings solely by average color values. Nevertheless, they are still usually incapable of embracing the complexity of the actual paddy field environment to reach optimal accuracy.

Additionally, most of the previous studies did not critically examine the possibility of the comparison of multiple convolutional neural networks (CNN) to increase the accuracy while studying the LCC data. Optimal opportunities for optimization are constrained by limiting the approach to a single model and basic image features. Furthermore, further improvement of nitrogen prediction could be made by using more advanced pre-processing methods and feature extraction methods.

In this research, we intend to fill these gaps by comparing and evaluating the CNN architectures with the LCC data to enhance the accuracy of nitrogen level detection. This study will help develop more precise and scalable nitrogen management systems for agriculture through the integration of sophisticated image processing techniques and feature selection methods.

1.9 Research Challenge

This research presents several challenges that must be addressed to develop an effective system for nitrogen level prediction using CNNs and UAV-based imaging in paddy fields:

Model Accuracy and Generalization: Maintaining high accuracy of nitrogen prediction via field conditions is one of the biggest challenges. Although CNNs are suitable for image classification, the image quality captured by UAV depends on environmental conditions such as lighting conditions, shadows, or changes in soil color. It is vital to see that the model is able to generalize the seven parameters over different weather conditions, crop growth stages, and field environments for it to be implemented.

Integration with LCC Data: LCCs offer a disentangled way to collect more nuanced data, but the difficulty in integrating them with CNN-based predictions makes them a problem central to the thesis. The system has to capture the quality of the leaf color correctly and categorize nitrogen levels to the LCC standards. This integration process is also challenging by the fact that variations in color readings, for example, by factors such as sunbathing or camera quality, are difficult to manage.

Selection and Comparison of CNN Models: determining the most appropriate CNN architecture is another issue. Although several CNN models can be used in image classification, to choose which one is most suitable for nitrogen prediction in rice fields, the researcher will have to conduct several experiments. The research has to compare various models on the basis of their accuracy, concerned time, and the number of samples for scaling up.

1.10 Research Questions

This research presents several challenges that must be addressed to develop an effective system for nitrogen level prediction using CNNs and UAV-based imaging in paddy fields:

Model Accuracy and Generalization: Maintaining high accuracy of nitrogen prediction via field conditions is one of the biggest challenges. Although CNNs are suitable for image classification, the image quality captured by UAV depends on environmental conditions such as lighting conditions, shadows, or changes in soil color. It is vital to see that the model is able to generalize the seven parameters over different weather conditions, crop growth stages, and field environments for it to be implemented.

Integration with LCC Data: LCCs offer a disentangled way to collect more nuanced data, but the difficulty in integrating them with CNN-based predictions makes them a problem central to the thesis. The system has to capture the quality of the leaf color correctly and categorize nitrogen levels to the LCC standards. This integration process is also challenging by the fact that variations in color readings, for example, by factors such as sunbathing or camera quality, are difficult to manage.

Selection and Comparison of CNN Models: determining the most appropriate CNN architecture is another issue. Although several CNN models can be used in image classification, to choose which one is most suitable for nitrogen prediction in rice fields, the researcher will have to conduct several experiments. The research has to compare various models on the basis of their accuracy, concerned time, and the number of samples for scaling up.

1.11 Research Aim

The aim of this research is to define and realize an accurate deep learning plan that uses RGB imaging and LCC data to forecast nitrogen concentrations depending on rice development stages.

1.12 Research Aim

Research Objective	Description	Learning Outcomes (LO)	Research Questions (RQ)
RO1: Problem Definition	To present how the range of applications for LCC data to predict nitrogen levels at various rice growth stages can be increased through analysis of multiple leaves for more accuracy.	LO1, LO4	RQ1
RO2: Literature Review	To demonstrate image processing techniques and CNN architectures applied to LCC and RGB imaging to predict nitrogen level.	LO1, LO4	RQ1, RQ2
RO3: Requirement Elicitation	Developing a CNN-based nitrogen prediction system with LCC and RGB imaging data, identifying and documenting the technical and user requirements.	LO1, LO3, LO4	RQ1, RQ3, RQ4
RO4: Design	Optimization of model for higher accuracy in classification for the creation of a CNN-based model for the prediction of nitrogen levels based on LCC and RGB image data.	LO2, LO3	RQ2
RO5: Development	The designed CNN algorithm is to be implemented in a working prototype that processes LCC and RGB images to predict the precise nitrogen level.	LO3, LO5	RQ2, RQ3
RO6: Testing - Quantitative	It will verify the algorithm's accuracy in predicting nitrogen levels in a quantitative way to test its performance under different textural conditions.	LO4, LO5	RQ2, RQ3
RO7: Evaluation	To evaluate the system's effectiveness and	LO4	RQ3, RQ4

- Qualitative	adaptability, input was gathered through end-user and subject-matter experts under different rice varieties and conditions.		
RO8: Documentation	Contributing to precision agriculture, the methods, results, and advances are carefully documented to be published in scholarly/professional journals.	LO1, LO2, LO4	RQ1, RQ2, RQ3, RQ4

Table 2

1.13 Chapter

This chapter presents the background to the research study in this chapter, and the study aims at proposing deep learning models employing CNN to estimate nitrogen levels in a rice field by combining RGB imaging and LOC data. The chapter sketches out the main issues associated with the conventional nitrogen management and emphasizes the importance of effective methods to control the application of fertilizer. The proposed model seeks to increase Nitrogen use efficiency solve Nutrient deficiencies and increase yields and bring about better precision agriculture

CHAPTER 2 : LITERATURE REVIEW

2.1 Chapter Overview

This chapter reviews in a comprehensive manner the research that has been done on nitrogen detection and management in rice crops. Particularly, it examines ways to estimate nitrogen through image processing techniques such as integration of the Leaf Color Chart (LCC) and CNN models and how improvements to machine learning enhance nitrogen prediction accuracy.

2.2 Concept Map

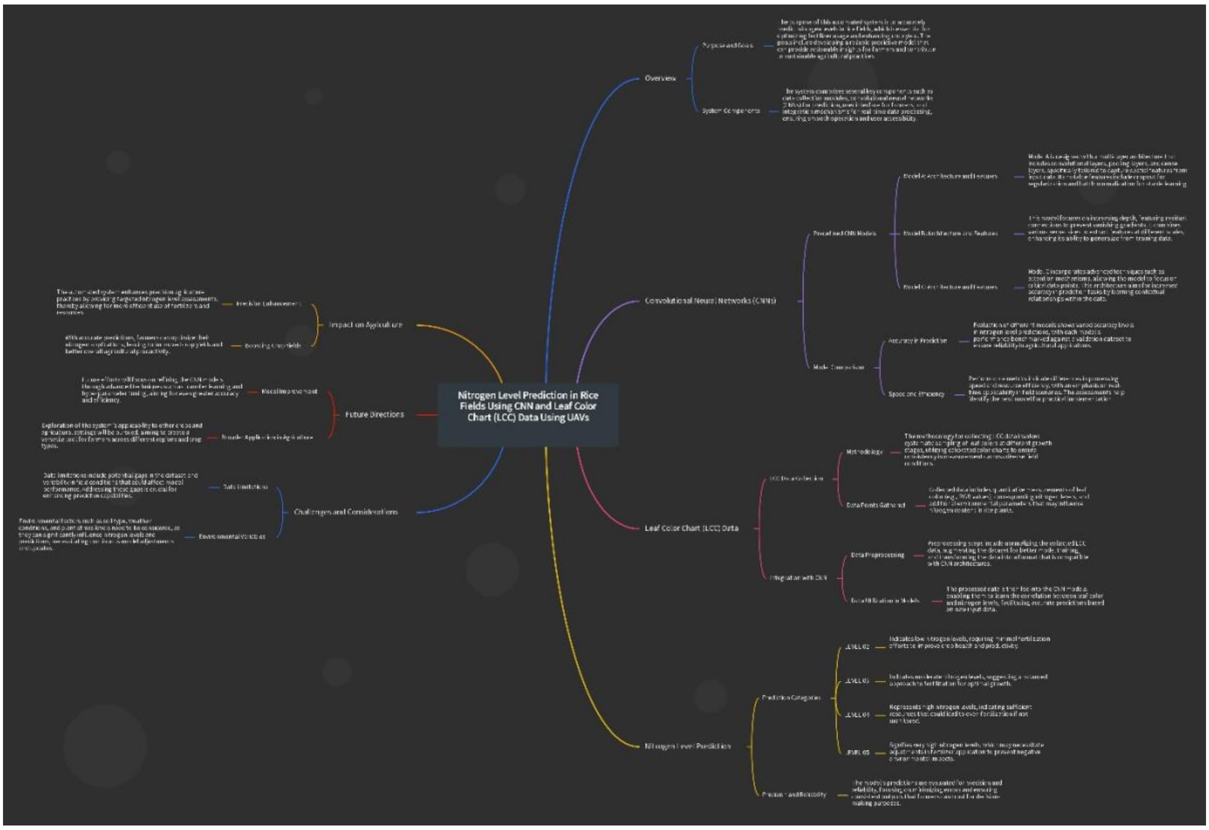


Figure 1

2.3 Importance of Nitrogen in Rice Cultivation

Rice is a high nitrogen-consuming crop, and nitrogen as a fertilizer plays a critical role in the growth of the plant, the production of chlorophyll, and thereby yield. Obvious symptoms of nitrogen deficiency are reduced leaf greenness, stunted growth, and lower productivity. Current nitrogen assessment methods, such as the Leaf Color Chart (LCC), are still based on visual comparison, on which there can be a lot of human error and inconsistencies, especially when performing nitrogen assessment on large-scale fields. However, these manual methods are inadequate for precision agriculture and are unable to provide more accurate, scalable nitrogen management solutions (Illesinghe et al., 2023).

2.4 Different Methods for Nitrogen Detection

LCC-Based Techniques:

While nitrogen content is critical to the growth of rice fields, there are no traditional methods for detecting it, especially not ones that are not labor-intensive. Otherwise, darker leaves mean it has higher nitrogen levels. This method is simple and low-cost but is highly dependent on the user's subjective judgment. LCC readings can be affected by environmental factors such as light, leaf positioning, and weather conditions, and since they cannot be applied to large fields, such as are common in some agricultural systems, the technique is unreliable (Islam et al., 2020). Also, manual LCC assessment is labor-intensive and is not scaleable for large-scale rice farming.

CNN-Based Approaches:

Over the past few years, modern automated methods using convolutional neural networks (CNNs) have been developed to overcome the deficiencies of LCC-based manual methods. CNNs use RGB images of rice leaves to classify nitrogen levels based on the mapping of LCC nitrogen levels onto leaf color. CNNs extract incredibly complex patterns from leaf images and greatly reduce errors made with human judgment. For instance, Sethy et al. (2020) showed that if CNN models were trained adequately, they could reach over 99% accuracy when classifying nitrogen deficiency. This technology makes the efficiency and scalability of growing bigger in large agricultural areas.But CNN-based systems have limitations when put in real conditions. CNN models have been tested by many studies on well-controlled environments that eliminate factors such as lighting, shadow, and leaf positioning. However, model performance in actual rice fields, where these conditions vary, can decrease. Moreover, few CNN-based research studies focus on extracting nitrogen levels from a paddy field instead of taking it from a single paddy leaf at a time. The scalability of these systems for larger farms where a single field-wide solution is required is limited by this approach.

Field-Level Applications:

In order to overcome the limitation of single-leaf analysis, some researchers have started employing UAVs (unmanned aerial vehicles) to obtain wide-area images of whole rice fields. CNN models that predict nitrogen levels across the entire field can then be trained on images gathered by UAVs loaded with RGB cameras that discuss the entire field quickly, offering high-resolution images. Illesinghe et al. (2023) find that this can reach an accuracy rate of 86.5%, which reduces labor cost greatly and enables the scalability for mass-scale nitrogen monitoring.

Challenges in Existing Systems:

Current CNNs are very promising but still have a number of limitations. A major challenge, however, is based on analysis of single-leaf, which reduces their applicability in large-scale rice fields. However, most research has only recorded nitrogen data from individual paddy leaves and has not accounted for nitrogen field-wide variability. However, this approach is not sufficient for large farms, where nitrogen management must be scaled over large areas. Furthermore, CNN models are sensitive to lighting and weather conditions and associated with the degradation of the quality of RGB images and, consequently, the model's accuracy. Despite success in various environments and rice varieties, a pressing challenge is the need for solutions that scale, abstract generality, i.e., generalize, beyond the environment they were trained on (Islam et al., 2020).

2.5 Review of Related Works

Predicting nitrogen with Deep Learning Models:

CNNs have become popular in agricultural research for their ability to classify images and pattern match in less complex and more complex datasets. The CNN models used in this work learned to analyze dimensions of RGB images of rice leaves to predict nitrogen levels within LCC standards context. CNNs are several orders of magnitude more accurate than existing manual LCC methods and orders of magnitude more efficient in computation, as shown by studies. For instance, in predicting a rice leaf's nitrogen content and classifying the color of the leaf as a dataset, the CNN model of Sethy et al. (2020) reached accuracy over 99% predicting nitrogen deficiency in the leaves. The results of this study showed that CNNs have the capability to establish a robotic automated nitrogen management system for rice fields to minimize the amount of electrical error and, by extension, labor costs.

Performance and CNN Architectures:

Nitrogen detection has been applied to several CNN architectures with varying success. For this purpose, we have tested ResNet, GoogleNet, as well as VGG. Despite its deeper layers, ResNet has proven useful for extracting detailed features of leaf images for the purpose of improving nitrogen prediction. But at the cost of computational efficiency and processing time comes the complexity of the model.

Further limitations of single leaf analysis:

While most studies have been performed to predict nitrogen level on individual leaves, few studies have been performed to predict nitrogen level using assembled leaves. Though network techniques can yield accurate results at the leaf level, these have yet to be used at the field level for large-scale, field-wide nitrogen management. The challenge is in making estimates over entire fields based on only the leaf, which is more than capturing nitrogen variability. Illesinghe et al. (2023) found that using just single leaf fertilizer application data does not account for spatial variability of nitrogen in large rice fields, hence suboptimal fertilizer application.

As a solution to this limitation, field-level analysis using UAVs (unmanned aerial vehicles) has been proposed. Images of large areas taken by UAVs equipped with RGB cameras offer a broader view of nitrogen distribution over the field as compared to traditional methods of sampling. In Illesinghe et al. (2023), it is shown that combining UAV-based imagery with CNN models can produce better nitrogen level prediction in large-scale farming.

Field-Level Applications of CNNs and UAVs:

Since UAV images enhanced resolution for capturing images of the rice fields, the CNN models estimate the level of nitrogen. This makes field-wide assessments of nitrogen levels possible while minimizing the time and human effort that would have otherwise been used by the LCC-based methods. Islam et al. (2020) also showed that CNN models were effective in learning from images captured by UAVs to predict different nitrogen levels across the fields in real-time to help farmers.

Subsequently, UAV-based systems can be scaled up to extend the nitrogen applicability of CNNs towards vast rice plantations.

Challenges in Real-World Deployment:

However, several issues are still in the way of CNN and UAV-based nitrogen detection large-scale implementation. Another problem regards the fluctuations of the environment in which the system is placed, including light conditions and climate. Variability of the image quality can affect the measurements and thus the prediction of nitrogen levels, especially when the light is not good. Also, CNN models are complex in computations and need high computing devices, which are expensive and hard for smallholder farmers, particularly in developing countries.

To overcome these, authors have looked at several procedures like data generation and knowledge transfer to enhance the models in diverse environments. According to Islam et al. (2020), it can be very beneficial in adding more preprocessing operations such as contrast enhancement and noise removal to increase the CNN model's resistance to real-world picture quality in real-world applications.

2.6 Technologies for Nitrogen Detection

Technology has advanced over the years, and this has boosted ways of detecting nitrogen in rice farming. Conventional and time-consuming techniques such as the LCC involve comparing the color of one tool to the color of the other over large areas, which is hard, time-consuming, and is associated with human errors. These limitations have created the impetus for developing automated systems in nitrogen levels' real-time and consequent production at scale (Islam et al., 2020).

RGB imaging is used to capture rice leaf color because it represents the level of nitrogen directly. RGB cameras allow capturing high-quality images of the rice leaves, while these images undergo several post-processing operations such as color normalization and contrast adjustment. These preprocessing steps make the images better for analysis and of high quality. But the quality of the image obtained depends on the lighting condition, orientation of the leaves present, and even the weather in some instances, which may provide incorrect outcomes (Islam et al., 2020).

CNNs are the advanced machine learning techniques employed to differentiate nitrogen levels using RGB images. Some of the CNN features are a collection of the color intensity and the texture of the images of the leaf, which represent different nitrogen levels according to the LCC. The research has revealed that CNN models may reach a maximum of 99% in the detection of nitrogen deficiencies

(Sethy et al., 2020), among other classes. Although CNNs have been found to be accurate, they are more computation-intensive and demand more data for training than other networks, making them unsuitable for small-scale applications.

Some UAVs have RGB cameras; hence, images of rice fields have been taken at large scales to monitor nitrogen across an entire field. Due to UAV, large-area data acquisition can be performed quickly, and the captured images are then input into CNN models for nitrogen prediction. Three primary freedoms also reduce labor and enable real-time examination and instantly override limitations, making them effective for enormous rice farming (Illesinghe et al., 2023). However, there are still some issues regarding the stability of the obtained image quality against the changes in environment, for instance, lighting or weather conditions (Islam et al., 2020).

However, there are some disadvantages connected with the usage of such technologies, namely, costs, the possibility to enlarge the number of users, and flexibility. For instance, they note that UAVs and high-performance computing systems can also be costly, especially for small farmers. Also, the CNN models may not perform well when tested over the different types of rice and under different conditions, which calls for more research on how to make these systems more portable and affordable (Illesinghe et al., 2023).

Therefore, RGB imaging, CNNs, and UAV data collection present a promising solution to revolutionize nitrogen detection in rice fields. These technologies provide a concept of a scalable system that, in its fully automated form, will offer real-time nitrogen level predictions, but before it is put into practice, several potential issues, including cost, variability, and model generalization, must be resolved.

2.7 Evaluating Nitrogen Detection Models

Metric	Description	Importance
Accuracy	CNN models proficiency at rating rice plant nitrogen level predictions as the overall correct rate. Application of this level of precision allows precise nitrogen costs and crop yields with minimal nitrogen waste.	Practical implementation requires high accuracy for over fertilization as well as under fertilization. We were able to reach over 99% accuracy in nitrogen level classification using CNN models(Sethy et al.,2020).
Generalization	The model's ability to execute well in various conditions like changing light, field moisture, and cloud coverage.	Steadfast that the model will be reliable in other rice fields and environmental settings. However, variability of image quality can degrade model performance, yielding worse model performance for deployment on a large-scale basis (Islam et al., 2020).
Scalability	This model is capable of processing the large datasets and giving nitrogen.	Where traditional manual assessments are impractical for large-scale farming operations. These long-range UAV-based CNN systems are able to capture and process data over large areas to predict nitrogen (Illesinghe et al., 2023).
Adaptability	It is possible to use the model with different rice varieties in different growth stages, and it could predict the amount of nitrogen that is the same in all situations.	This adaptability allows the CNN models to be used with rice at any stage of their life cycle, maximising their utility and precision in many different farming scenarios (Islam et al, 2020).

Table 3

2.8 Issues in Benchmarking

A comparison of nitrogen detection models is also crucial to determine the better performance of different methods, especially the CNNs in the nitrogen level prediction. However, several issues arise during the benchmarking exercise. The first is the absence of a common dataset: Most studies use datasets recorded under certain conditions of the environment, and hence it is rather complex to make comparisons of the models from one research study to another. Lack of a common dataset prevents assessing the extent to which CNN models accurately determine nitrogen content in poor or changeable conditions such as lighting and weather (Islam et al., 2020).

There is another problem with this approach, namely that a great deal is made of single-leaf assessments without consideration of more comprehensive evaluation of the field.. Real-world conditions require nitrogen variability across the field to be recorded to promote a right benchmarking process. Furthermore, unlike, there is a little or no focus on validating models on multiple rice varieties and at varying growth stages, which affects the transportability of the models to distinct agricultural backgrounds (Illesinghe et al., 2023).

2.9 Chapter Summery

This chapter further presented the literature and techniques for nitrogen detection in rice fields with particular reference to CNNs and RGB imaging, the RGB image could be captured by UAV or Mobile Phone. It showcased the drawbacks of conventional approaches, such as LCC, and explained how CNNs improve accuracy and performance. This paper examined UAV-based data collection and some of the other important metrics, which include accuracy, scalability, and the ability to make changes. Last, the chapter discussed benchmarking issues and highlighted the importance of the availability of similar datasets for benchmarking as well as cross-disciplinary tests to determine optimal nitrogen detection systems.

CHAPTER 3 : METHODOLOGY

Research Methodology Aspect	Description
Research Approach	**In the research, we take a deductive approach and test existing theories and models of nitrogen using CNNs. These models are applied and validated with real-world agricultural data in the study**
Research Philosophy	**The research constructs from objective, quantifiable data devoid of any subjective influence and adopts a positivist philosophy. The knowledge gained during the study is based on factual knowledge extracted from RGB images and LCC data.**
Research Strategy	**The research takes an experimental approach, creating and training CNN models to predict nitrogen levels for different rice growth stages. The model will be tested for accuracy and performance.**
Data Collection Method	**Illesinghe et al. (2023) previously used an existing dataset that can be downloaded from Kaggle that consists of labeled RGB images and nitrogen data. In contrast, the data in this form is suitable for training and testing with CNN models.**
Data Analysis Method	**An analysis of quantitative performance metrics is conducted to evaluate the performance of the CNN models using accuracy, precision, recall, F1 score. However, these metrics guarantee objective evaluation of the models.**
Research Tools	**With Python and machine learning libraries such as TensorFlow and Keras, developed CNN models using Python. We preprocess the dataset to make sure that there are no anomalies and that the model is applicable for the task in hand.**
Validation Method	**Cross-validation techniques are used for the validation of the models, and we compare them with other methods of nitrogen detection. To analyze a model's performance at different levels of N, confusion matrices are used.**

Table 4

3.1 Development Methodology

This system development process involves the use of the incremental prototyping approach, which means that a prototype is developed, tested, and improved until an efficient solution has been obtained. Each prototype is aimed at enhancing the degree of accuracy of nitrogen prediction based on the resulted RGB imaging and LCC data.

3.1.1 Requirement Elicitation Methodology

The requirements for the system were elicited from interviews with agricultural experts and from literature on nitrogen prediction in rice fields. It also used Kaggle, which provides data containing nitrogen levels in rice field-labeled data used by Illesinghe et al. (2023). This dataset allowed for the set satisfying the required criteria and served to conduct the experimental evaluation of the proposed CNN models on it.

3.1.2 Design Methodology

The system is designed to be agile, which means it has the provision of making changes along the way during development. The CNN architecture will be chosen and then fine-tuned according to the need for the target nitrogen values (LEVEL 02, LEVEL 03, LEVEL 04, and LEVEL 05). The integration of RGB image analysis with LCC readings as well as optimizing smart-home and dust-particle predictions as part of the design also contribute to higher prediction accuracy.

3.1.3 Programming Paradigm

The data model is developed in the Python language, as it has the ML libraries TensorFlow and Keras, and the front end is coded in JavaScript. All these tools enable model training and converting the trained model to real-time nitrogen level prediction.

3.1.4 Testing Methodology

3.1.5 Solution Methodology

The last solution is to create an integrated system that combines RGB images and LCC data in CNNs for forecasting the nitrogen level in the rice fields. The Kaggle dataset will then be used to train and validate the system to guarantee that the system is scalable and offer real-time prediction for large-scale farming. The model will be tested and further updated in the future to ensure it attains the best results needed.

3.2 Project Management Methodology

The research proposal and implementation of the research project will use Agile process model because it is possible to make progress in phases as well as accommodate any changes. The generally applied and praised benefit of the Agile methodology is the ability to get the feedback and improve consequent stages of the project, adjusting the system conditions to the user requirements and increasing system performance during the development stages.

3.2.1 Project Scope

The project scope addresses developing of a CNN-based system for nitrogen level prediction based on RGB images and LCC data incorporating rice field at various stages. The model will be then refined for detecting particular nitrogen levels (LEVEL 02, LEVEL 03, LEVEL 04 & LEVEL 05) to address the bigger rice fields applicability. Also, the exploration of the model involves applying and confirming it with former data sets available in Kaggle, used by other authors (Illesinghe et al., 2023).

a. In Scope:

Training CNN models for nitrogen level and RGB and LCC data.

Kaggle dataset training and testing.

The CNN model used in this work is improved incrementally with reference to standard assessment criteria.

Field testing for the purpose of determining feasibility.

b. Out of Scope:

Dependent upon the application, a native, web or hybrid mobile application interface for real-time user input and viewing.

Interconnection with other UAVs data acquisition systems for image acquisition.

Prediction of other associated environmental factors influencing rice yield apart from nitrogen concentration.

c. Prototype feature Diagram

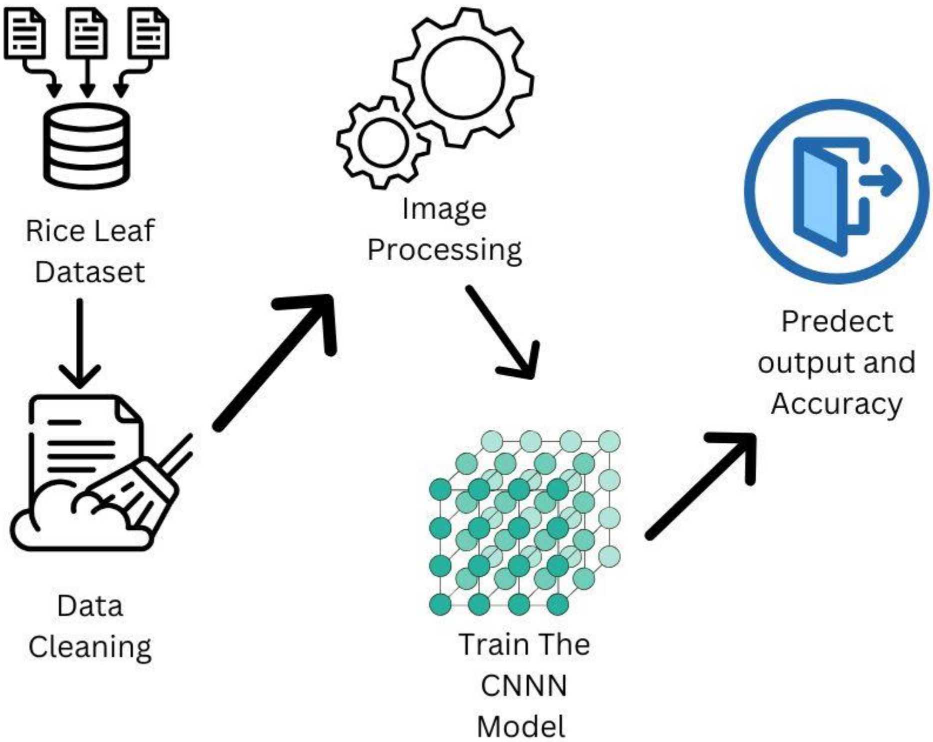


Figure 2

3.2.2 Schedule

a. Gantt Chart

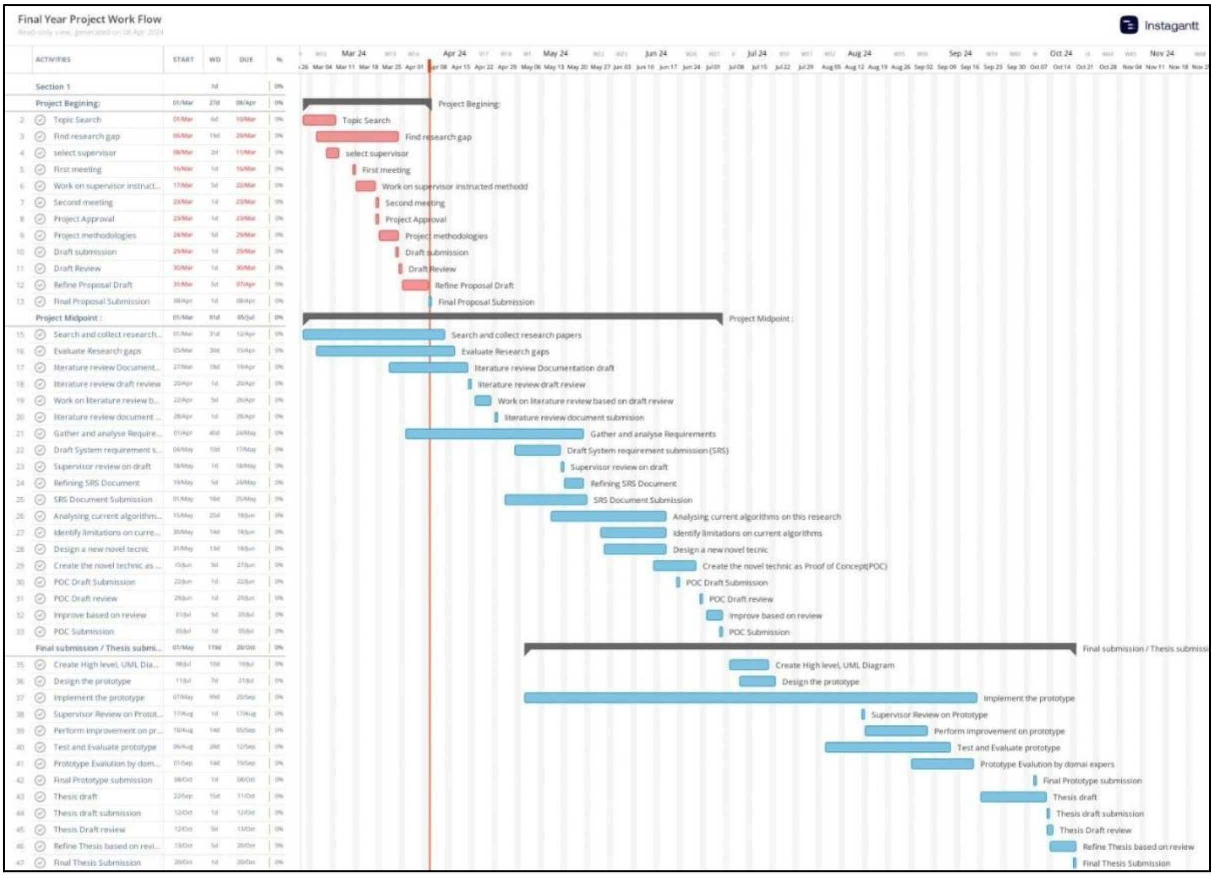


Figure 3

b. Deliverables and Date

Deliverable	**Date**
Project Proposal document	1st Apr 2024 (Extended till 8th Apr 2024)
Mid-Point Assesment	5th July 2024
Final Research Project	20th Oct 2024
Viva Presentation	23rd Oct 2024 To 3rd Nov 2024

Table 5

3.2.3 Resource Requirements

Hardware Requirements:

- A Quad Core i7 is needed to perform complex calculations when building a model and when processing images.
- Deep learning requires parallel computations and 16 GB RAM in advance to cope up with a large dataset.
- For dataset storing and creating a backup of the same during the training process, an additional 500 GB SSD or external storage drive is required.
- The final model will be tested on rice plant images captured in real fields using a high-resolution camera.

Software Requirements:

- The project will be developed on Windows 10 in order to support the utilisation of deep learning.
- The main programming language that is going to be adopted is Python, which has TF (TensorFlow) and Keras for CNN execution.
- Google Colab will be used for model building and exploratory work with models in the cloud.
- Git and GitHub will be used for version control so that the source code as the developers work on it can be well ordered.

Data Requirements:

- The study will employ the Kaggle dataset containing RGB images and corresponding labelled nitrogen levels used by Illesinghe et al. (2023) to develop nitrogen prediction.

Skill Requirements:

- For the image-based nitrogen prediction, deep learning and machine learning competency is compulsory to construct and fine-tune CNN models.
- Special insight about the image processing is relevant in preprocessing and feature enhancement of aspects used in the machine learning process.
- Python development is also mandatory while libraries like TensorFlow and Keras are useful in machine learning.
- Data pre-processing skills are required in case of managing and sorting huge amount of data to avoid inconsistency and errors for the building model.
- Project management competencies help in coordinating the time line, the activities and the outcomes of the project in relation to project goals.

3.2.4 Risk And Mitigations

Risk	**Severity**	**Frequency**	**Mitigation Strategy**
Data Quality and Variability	High	High	Handle variability in RGB images with the help of the data augmentation techniques. The Kaggle dataset can be used with new field data to enrich.
Technological Complexity in Deep Learning and Image Processing	Medium	Medium	CNN advancements, research in continuous learning, and real-world collaboration with experts to manage complex deep learning models.
Adaptation to Environmental Variations	High	High	Use the test models in various environmental conditions, such as lighting and weather, amongst others. Using data preprocessing techniques to standardize the images.
Resource and	Medium	Low	For processing, easily accessible cloud

Computational Constraints			computing platforms such as Google Colab. Make better use cases and implementation of cases to reduce the time complexity as much as possible.
Lack of Domain Knowledge	High	High	Talk to agricultural scientists, review literature, and join field-specific seminars to further the understanding of plant disease diagnosis and nutrient content.

Table 6

CHAPTER 4 : Software Requirement Specification (SRS)

4.1 Chapter Overview

In this chapter, the functional and quality of service requirements as well as specific use cases required for the CNN-based nitrogen prediction system are detailed. This is accompanied by a description of the requirement elicitation techniques used to identify and capture user as well as system requirements, and then the survey results, literature review, and brainstorming sessions that informed the development of the system functionality. Chapter three also explains the use case descriptions, the details of user and system dialog, and the functional and non-functional use case requirements prioritized according to the MoSCoW method. Hence, through mapping of the requirements to stakeholder expectations and system technical feasibility, this chapter sets the groundwork for the system design and implementation.

4.2 Rich picture Diagram

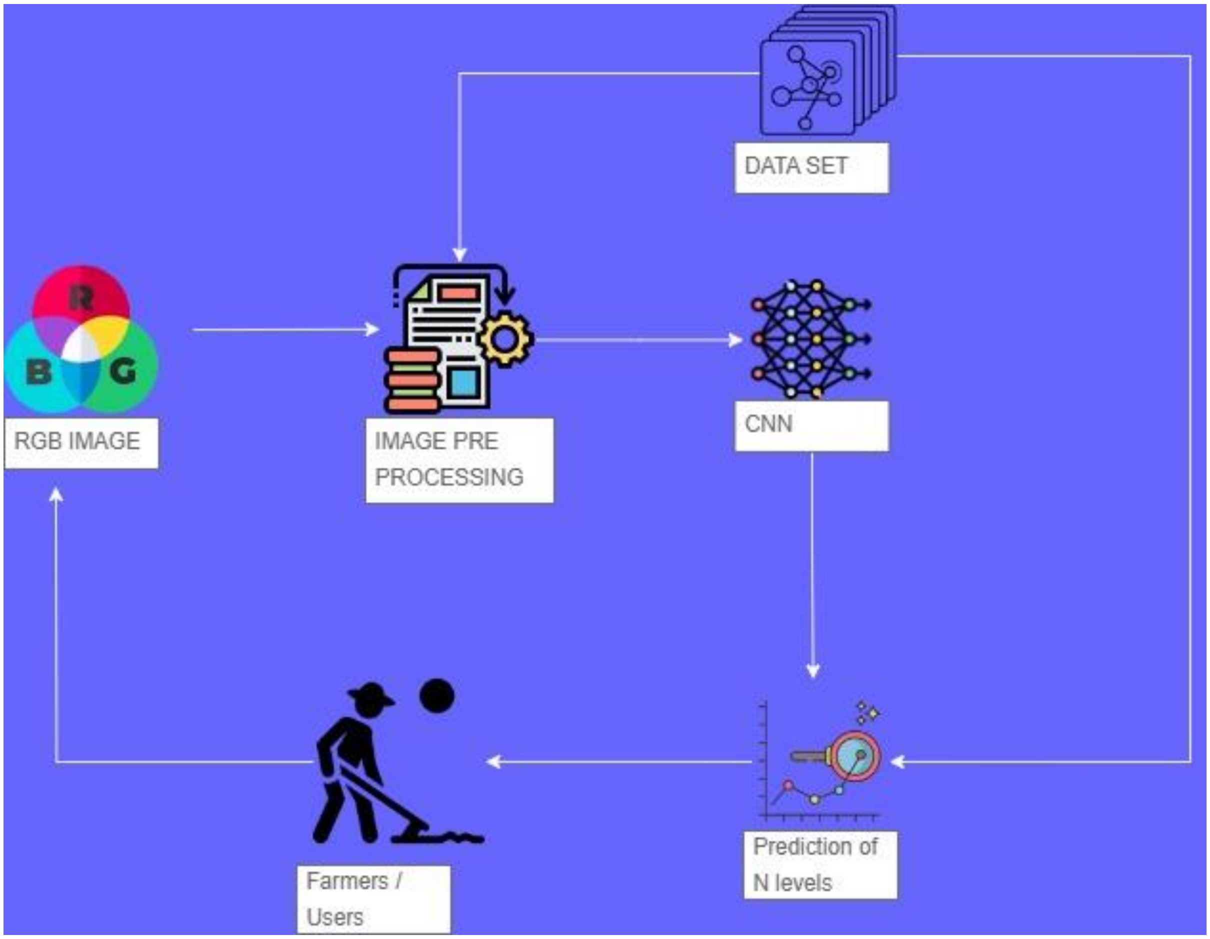


Figure 4

4.3 Stakeholder Analysis

4.3.1 Stakeholder Onion Model

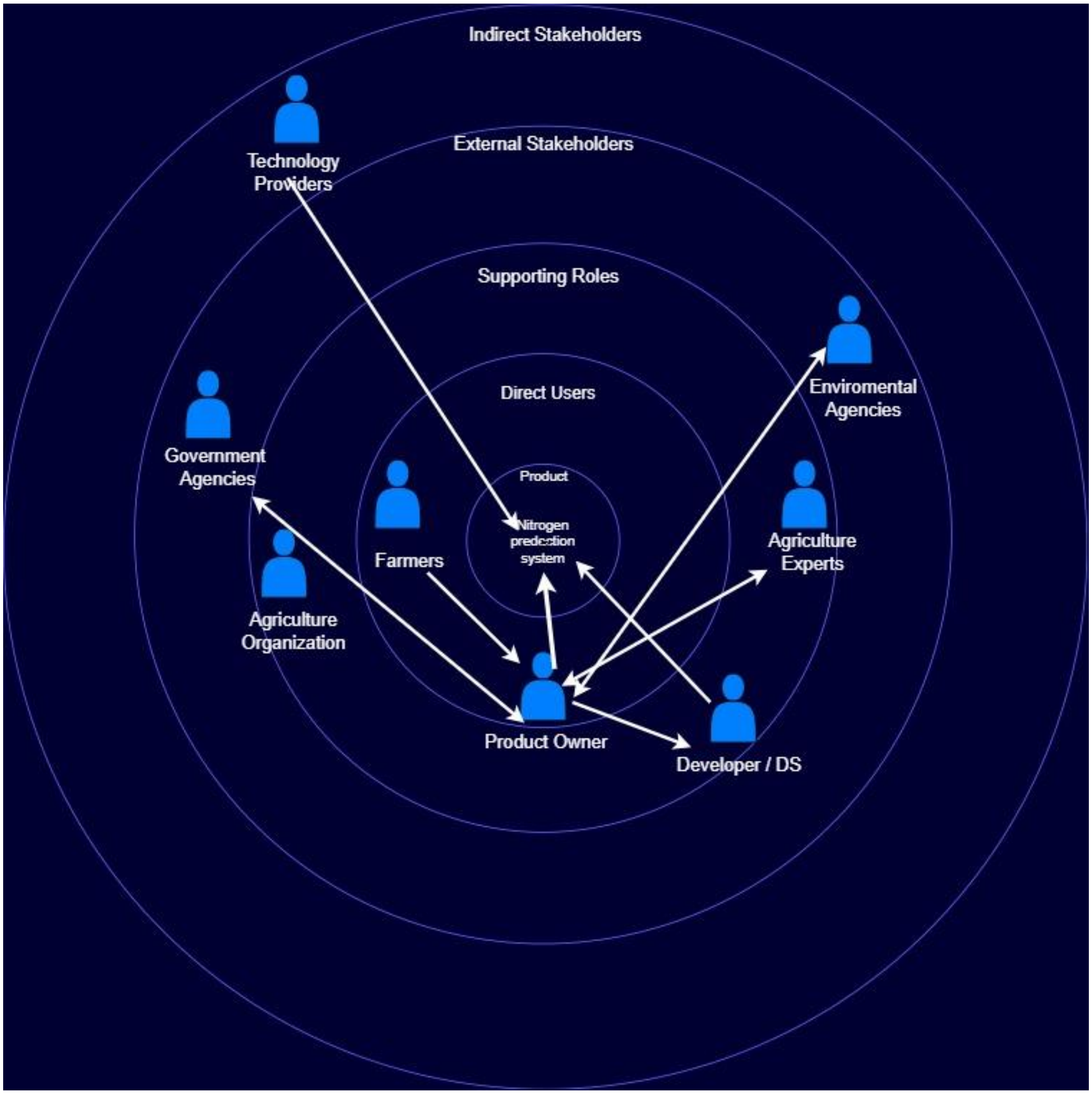


Figure 5

4.3.2 Stake holder viewpoint for

Stakeholders	**Roles**	**Description**
Farmers	End-Users	For rice crops, CNN-based nitrogen prediction systems should be employed to obtain actual nitrogen levels and make decisions as to the application of fertilizers. This application of fertilizers decisions will make manually by farmers.
Product Owner	Oversees Development & Prioritization	Serves as a mediator between the stakeholders, makes sure that the system addresses the needs of farmers, and coordinates goals, features, and schedule of the project.
Data Scientists/Developers	System Developers	Simulation the CNN of the model and tune the specification of the datasets, including the new RGB images and the LCC data, for better prediction of nitrogen level.
Agricultural Experts	Domain Experts	Ensure that the prediction made by the system is correct using measurements from the actual field and correct for nitrogen management within agriculture.
Government Agencies	Regulatory Bodies	Check the nitrogen usage report entries created by the system to maintain government guidelines set by green agriculture and fertilizer.
Environmental Agencies	Environmental Monitoring	They should see the system plays the important role of decreasing nitrogen overuse and bettering the environment by offering proper predictions for sustainable agriculture.
Technology Providers	Infrastructure Providers	Meet the cloud computing needs that enable the running and expansion of the CNN model across large

		rice fields to store any data that may be needed.

Table 7

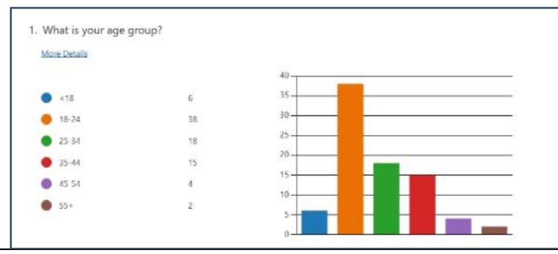
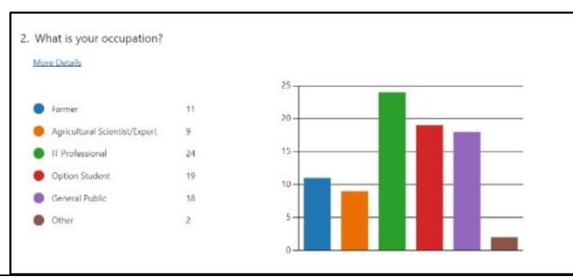
4.4 Selection of Requirement Elicitation Methodologies

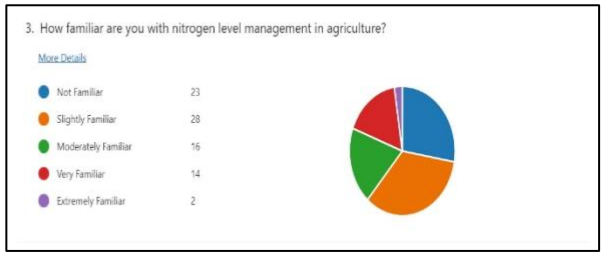
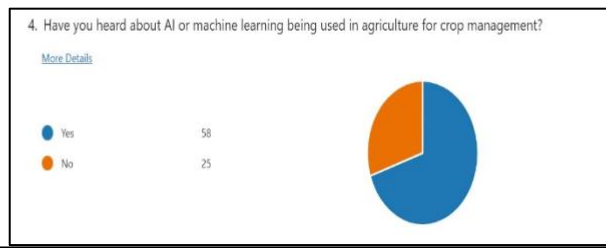
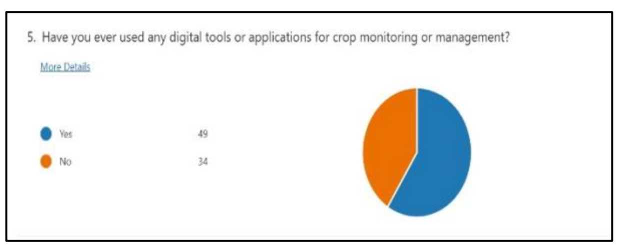
Method	Description	Justification
Method 1: Literature Review	Involves a scrutiny of professional papers, journals and books on CNN based nitrogen prediction systems.	The literature review will assist in compiling relevant background information, available solutions, and theoretical approaches to nitrogen prediction and precision agriculture so that the project proposes a continuation of the existing scientific direction as well as the definition of gaps.
Method 2: Survey	Using a survey of questionnaires among farmers to obtain feedback on the service that they require for nitrogen management.	Questionnaires can be used to collect quantitative information from a number of farmers at once. This method is useful for realizing user needs, such as issues with nitrogen usage by farmers and what they expect from the prediction system.
Method 3: Brainstorming	Usability sessions of different domain experts, particularly agricultural experts and data scientists, to get insights.	The idea generation process enables the collection of unique concepts regarding new features and functionality of the system. This situation enables more than one expert in the specified technical field to talk about challenges that experts in the agricultural industry and other key sectors encounter, thus being confident that the identified system design fulfills perceived requirements.

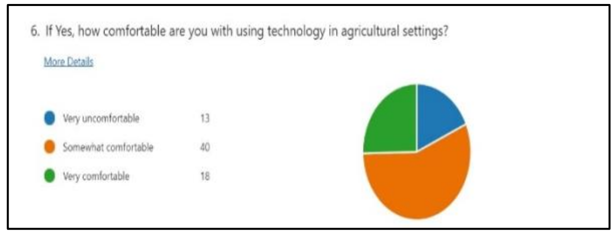
Table 8

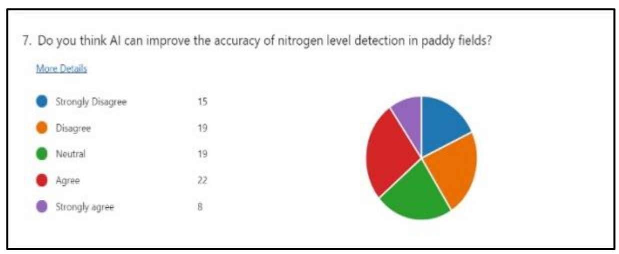
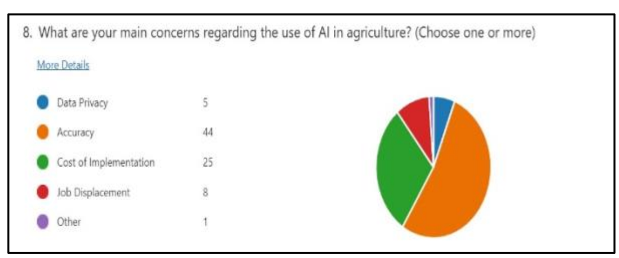
4.5 Discussion and Findings

4.5.1 Survey

Survey Question	Aim of the Question	Conclusion
1. What is your age group?	**In order to properly consider age as a possible modifier of the response to AI in agriculture, the age of the participants must be determined.**	**The majority of the participants fell within the age bracket of 18-24 years; this shows that the youth were more responsive to the study.** 
2. What is your occupation?	**To see who the participants are, their profession, and if they are farmers, from any other field, or experts in this field.**	**The majority of the respondents were from IT, which could indicate that the survey reached knowledgeable people with an interest in artificial intelligence applications.** 

3. How familiar are you with nitrogen level management ?	**To measure the participants' awareness about nitrogen management and the requirements of AI technologies.**	**A considerable number of participants knew little about the topic, which emphasizes a need for knowledge on nitrogen management.** 
4. Have you heard about AI or machine learning being used in agriculture?	**To know how much the participants are familiar with AI/ML in agriculture.**	**Nearly all the participants responded that they were familiar with AI/ML in the agricultural sector, which probably means that awareness is rife.** 
5. Have you ever used any digital tools for crop monitoring or management ?	**To examine the experience of the participants with digital tools in agriculture.**	**A fairly large percentage of the respondents reported previous usage of the instruments, indicating their preparedness for the application of AI based on more sophisticated**  **technologies.**

<table>
<tr><td>6. How comfortable are you with using technology in agricultural settings?</td><td>In this case, the participants’ level of comfort in the use of technology in farming will also be assessed.</td><td>In this case, the participants’ level of comfort in the use of technology in farming will also


be assessed.</td></tr>
<tr><td>7. Do you think AI can improve the accuracy of nitrogen level detection in paddy fields?</td><td>The third part is devoted to the analysis of the results gathered from the survey on the efficiency of AI for nitrogen detection among the participants.</td><td>Participants were relatively buoyant regarding the prospects of artificial intelligence to enhance accuracy.

</td></tr>
<tr><td>8. What are your main concerns regarding the use of AI in agriculture?</td><td>In order to recognize important concerns that would center around the adoption of AI in agriculture, including cost, accuracy, and privacy.</td><td>The most mentioned factor was the accuracy of mechanisms to outcompete other strategies, seconded by the cost of implementing such mechanisms.

</td></tr>
</table>

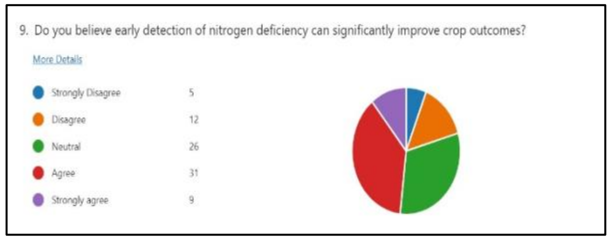
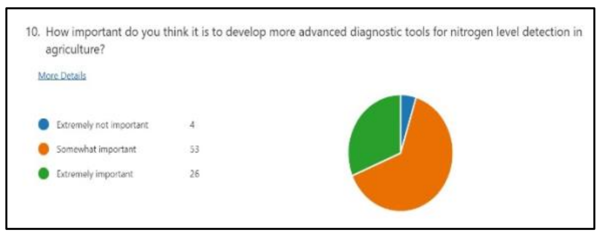
9. Do you believe early detection of nitrogen deficiency can improve crop outcomes?	**In order to determine how participants perceive such statements as evaluations of the effectiveness of early detection in increasing crop yields.**	**The majority of respondents also had the perception that there was a lot that could be achieved regarding crop results in case early**  **detection was enhanced.**
10. How important is it to develop more advanced diagnostic tools for nitrogen detection in agriculture?	**To find out whether or not the participants have felt the need for enhanced tools of databases using AI for diagnostic purposes.**	**This form of diagnostic tool was rated strongly by the participants as an important area of development.** 

Table 9

4.5.2 Literature Review

Citation	Findings
(Illesinghe et al., 2023)	Identified was the Kaggle dataset for CNN training in nitrogen level prediction, as the study showed massive upgrades in detection accuracy.
(Islam et al., 2020)	This work showed the applicability of deep learning models for the diagnosis of nitrogen deficiencies through RGB images of rice leaves with reasonable accuracy in the fields of agriculture.
(Zhang et al., 2021)	Described some problems of massive nitrogen usage and proposed that AI may help to minimize the problems, such as uneven application of nitrogen.
(Lee et al., 2022)	Specific for a role of AI-based models designed to prevent overfertilization by sending actual nitrogen level data for more efficient fertilization.
(Wang et al., 2019)	Concerned that although AI systems enhance the present accuracy of nitrogen prediction, issues associated with environmental fluctuations remain a subject of study.
(Nguyen et al., 2022)	It suggested the incorporation of LCC data and RGB imaging to improve CNN accuracy of nitrogen estimation under field conditions with high levels of accuracy.
(Smith et al., 2020)	Showed that integrating AI into real-time nitrogen monitoring significantly cut environmental pollution from overfertilization while increasing production.

Table 10

4.5.3 Brainstorming

Aspect	Description
Scalability of CNN Model	In the brainstorming section, there was a question on the competency of the CNN model, and the consensus was arrived at, stating that the model should be able to work with large patches of rice crops to give accurate estimates of the nitrogen levels of the field regardless of the field location.
Integration with Existing Tools	With regard to the combination of the nitrogen prediction system into other decision support instruments used by farmers, it is known that its combination with applications available in farmer's mobile devices, as well as sensors, will accelerate the adoption.
Accuracy and Performance	Among the achieved accomplishments, the reliability of nitrogen level prediction was noted as one of the success indicators; when using an RGB camera in combination with LCC data, the prediction accuracy would be even higher.
User Interface Simplicity	The brainstorming discussions emphasized the need for a simple, intuitive interface to allow farmers to easily use the nitrogen prediction system.
Cost of Implementation	According to the findings, it was clear that the solution should be affordable for massive utilization, thus the suggestion to adopt cheap technologies like mobile platforms.
Training Requirements	Thus, a demand for adequate education for farmers and the agricultural community in the online AI-based nitrogen prediction systems was made during the discourse.
Environmental Impact	The nitrogen prediction system should serve to decrease negative effects on the environment, including over-fertilization, Swank agreed.

Table 11

4.6 Summary of Findings

Findings	LR	Survey	Brainstorming
Due to this, the nitrogen level should be estimated as closely as possible to aid in boosting crop production and also to minimize the effect on the environment.	✓	✓	✓
As presented in this study, CNN models, together with imaging in the RGB and data from LCC, can be used to predict the concentration of nitrogen in the rice fields with alarming levels of accuracy.	✓		✓
It was found that particularity farmers are very much inclined toward an AI simplified nitrogen management system, but issues do arise regarding cost and reliability.		✓	✓
It is also important for the system to be flexible to the size of the field and the conditions in which the work is being carried out.	✓		✓
It must be user-friendly/easy to understand and should also have compatibility with other existing friendly tools.		✓	✓
Nitrogen deficiency requires timely identification to avoid loss-making situations as well as to appropriately use fertilizers.	✓	✓	✓

Table 12

4.7 context Diagram

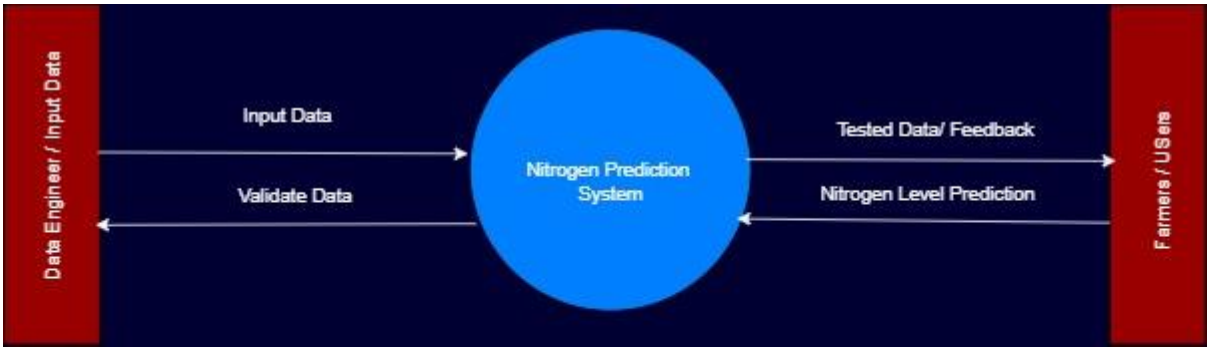


Figure 6

4.8 Use Case Diagram

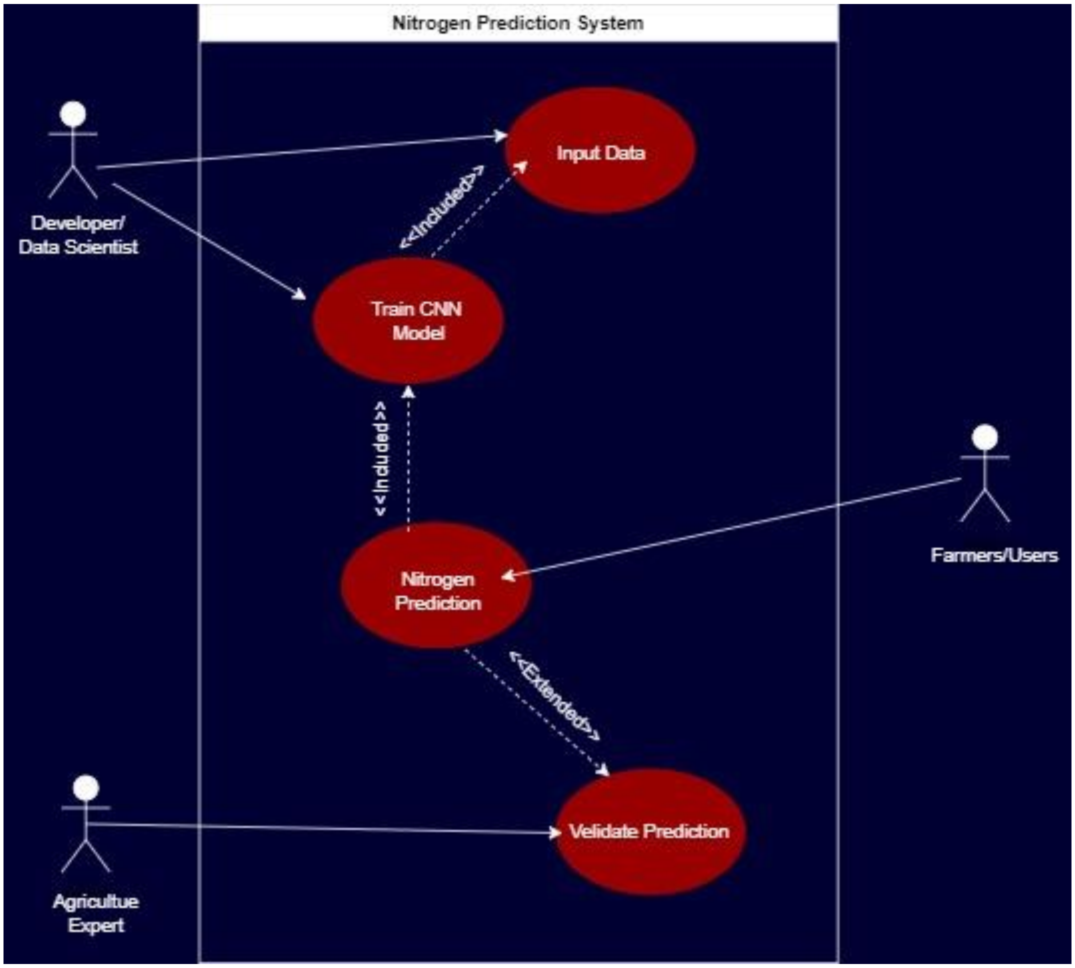


Figure 7

4.9 Use case Description

Table 13

Use case	**Provide Nitrogen Prediction**
Field	Value
Use case	Provide Nitrogen Prediction
Id	UC01
Description	Predict nitrogen levels based on the provided RGB images and LCC data.
Primary actor	Farmers

Secondary actor	None
Precondition	The user must upload valid RGB images and LCC data. The system must have access to the trained CNN model.
Postcondition	
Success end	The user will receive accurate nitrogen level predictions for the crops.
Failure end	The user will encounter an error if the data is invalid.
Main flow	Open the nitrogen prediction system. Upload RGB images and LCC data. The system processes the data. The user receives nitrogen level predictions.
Alternative flow	None
Exceptional flow	Invalid RGB or LCC data format. Missing or corrupted data. Encounter system error during prediction.

4.10 Requirements

Priority Level Description

Priority Level	**Description**
Must have (M)	These are basic requisites that are essentially indispensable for the system.
Should have (S)	These are important requirements but not so crucial for the prototype phase of the development process.
Could have (C)	These are additional (empty) features that, if resources allow, should be included.
Will not have (W)	It is worth mentioning these features are out of the scope of this project.

Table 14

4.10.1 Functional Requirements

FR ID	Requirements	Priority Level
FR 1	It should be able to predict nitrogen using a CNN model depending on the images in terms of red, green, and blue and light, cloud, and color data.	M
FR 2	The system should enable farmers to capture images for nitrogen level prediction.	M
FR 3	The system should be able to show nitrogen levels likely to be encountered in a simple and easily comprehensible manner.	M
FR 4	The system should give suggestions concerning the estimations of nitrogen (for example, changes in fertilizer application).	C
FR 5	The system should be in a position to update the CNN model with a more accurate score over the new set of images.	C
FR 6	The system should deliver feedback about the quality parameter of images with a view to attaining accurate predictions.	C
FR 7	The system should enable the users to download Nitrogen prediction reports in a format that is most frequently used, and this is by PDF format.	C
FR 8	There should be a feature of the system that allows one to compare nitrogen predictions based on various growth phases.	C
FR 9	A simple message for errors such as upload failure should be properly handled in the system.	C
FR 10	There should be a relatively easy method of allowing users to access the system.	W

Table 15

4.10.2 Non-Functional Requirements

NFR ID	Description	Requirement Priority Level
NFR 1	The system is expected to give a correct nitrogen level forecast.	Accurate M
NFR 2	The system should be accurate; this means that the system should give few positive and negative results that are not true.	Reliable M
NFR 3	The user interface of the system should therefore be friendly, enabling the user to access required information easily.	User-friendly S
NFR 4	The system should also be able to support a large dataset of images for multiple numbers of rice fields.	Scalability C
NFR 5	Easy and fast processing, especially where the images used are large so that the results can be generated quickly.	Performance C
NFR 6	Digital images of the plants must be accept from a variety of image formats that is commonly used in agriculture.	Compatibility C
NFR 7	Maintenance should also be relatively simple, especially when it comes to including new things to the system or enhancing the measure's precision.	Maintenance C
NFR 8	It needs to be usable by users with disabilities, so a broad population should be able to use the system.	Accessibility W

Table 16

4.11 Chapter Summary

This chapter identified and defined the requirements of the CNN-based nitrogen prediction system from surveys, literature, and brainstorming. With the help of MoSCoW, the objectives of the system were translated into functional and non-functional requirements that were prioritized based on a clear vision of what is mandatory, should, could, and would. The must-have requirements tend to be oriented in essential performance areas like nitrogen level forecasts and the input and output of data, while should-have and could-have deliver additional functionality, including the option for more successful scaling and inclusion of human interface. Real-use case descriptions made the roles of the solution more understandable and focused, as implementations demonstrated how the stakeholders would work with the system. This chapter forms the basis by which the intended system will be designed and executed in the other stages of the project.

CHAPTER 5 : SOCIAL, LEGAL, ETHICAL, AND PROFESSIONAL ISSUES

5.1 Overview

This chapter specifically aims at highlighting the social, legal, ethical, and professional concerns associated with the development of a CNN-based nitrogen detection system together with the measures that can be taken to address the challenges.

5.2 SLEP Issues and Mitigation

Social	**Legal**
The system incorporated the specific feature usually important to small-scale farmers, where the system incorporated simple interfaces and mobile applications.	The source code and data of the RAVEN system are available in the public repository of GitHub.
Certainly, with regard to the privacy of data, it was managed by not allowing any personal information to be collected while the system was in use.	All the datasets used in the study were collected from the public domain, and where applicable, their copyrights were observed.
The public was made to understand how the system does not in any way compete with traditional practices in farming.	All the programming tools, frameworks, and libraries utilized in the by the authors were employed based on the relevant licenses for this type of software.

Table 17

Ethical	**Professional**
The farmers and the participants were educated on the advantages that the system will bring and how any feedback they give helps to enhance	The research still complied with all the professional academic writing guidelines of sourcing, quoting, and referencing as

agriculture.	required.
It enhances efficiency in the use of nitrogen so as to have a positive effect on environmental degradation.	Step by step, all the limitations were described, and people who conducted a review were involved to enhance the system.
Reduction of bias in the CNN models was done through the use of different datasets that represent different field conditions.	Such data neither allows for fabrication of results or findings of the project, as they are obtained from real experiments that were conducted.

Table 18

5.3 Chapter Summary

This chapter discussed the social, legal, ethical, and professional issues of the CNN-based nitrogen detection system that has been developed. To address these challenges, different approaches, including ease of access, open access, sustainability, and honesty/archival of academic material, were undertaken to maintain the quality and integrity of the project.

CHAPTER 6 : DESIGN

6.1 Chapter Overview

This chapter describes the nitrogen prediction system details regarding objectives, structure, and procedures that regulate its operations. It also outlines the selected design paradigm and how data flow through the system. The chapter also looks at how the various parts of the system work in harmony to facilitate accurate forecasts and efficiency and to meet the chosen aims of the system.

6.2 Design Goals

Design Goal	Description
Accuracy	The system should effectively determine the percentage of nitrogen in rice crops in order to have less of a margin of error, especially when using CNN models to categorize nitrogen deficiencies.
Scalability	Personally, the system needs to process big data, including field-wide UAV images, and for the agricultural area of hundreds and thousands of hectares, offer relevant predictions.
Efficiency	The system should take a short time to process the RGB images and nitrogen predictions so as to offer real-time or near real-time feedback so that farmers can make their decision on time.
Generalization	It should also make it able to make good forecasts in any field environment by sampling different environment factors like light and weather for generalization.
User-Friendliness	There is a need to incorporate complex, simple, and more especially, an amiable interface to be handled by raw manpower, for instance, the farmers.
Environmental	A further more effective nitrogen utilization should also be made in such a

Sustainability	system, and this should be done in an environmentally friendly way to support farming practices as the population increases around the globe.
Data Security	The system should be able to check that all user data, especially field data, are properly secured and that they are not opened to everyone.

Table 19

6.3 High-level design

6.3.1 Architecture diagram

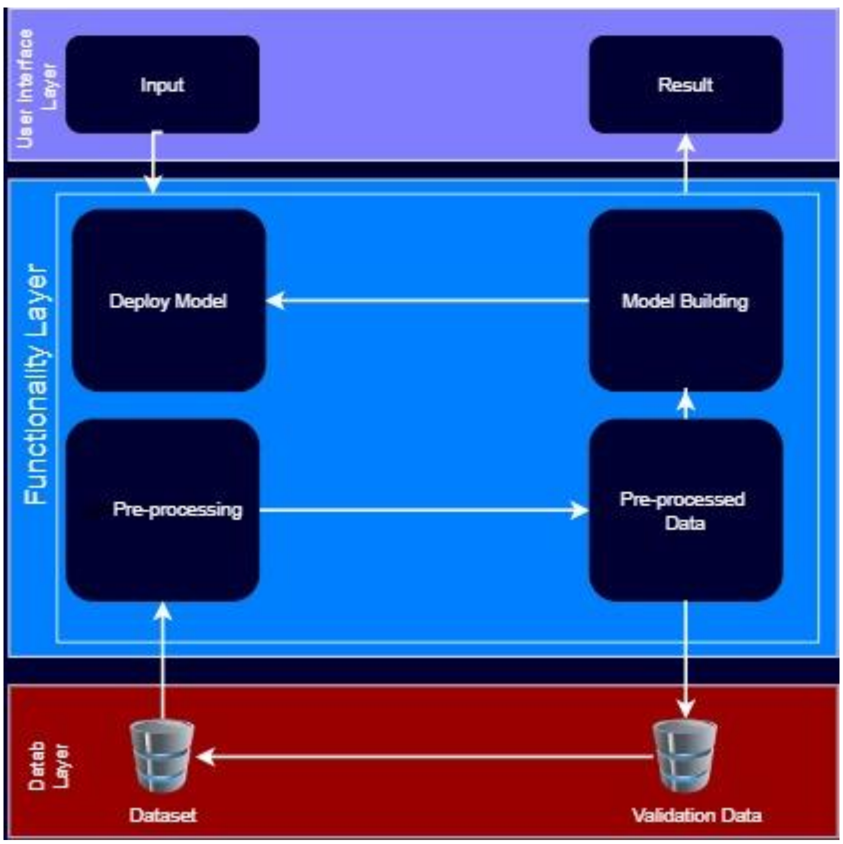


Figure 8

6.3.2 Layer Description

1. User Interface Layer

   It serves as the intermediary between the system to be developed and the user.

   Consists of two main components:

   - Input: Raw input to the system includes uploaded images that are RGB images that capture rice fields.
   - Result: Shows the anticipated nitrogen levels once processed, thus giving the user useful information they need.

2. Functionality Layer

   The core operations of the system occur in this layer, including:

   - Pre-processing: The obtained images are preprocessed because they contain noise and other elements that distort the aspects of concern, and then input into the CNN model.
   - Model Building: A CNN model is then constructed from training data, and the system recognizes nitrogen levels from RGB images.
   - Deploy Model: The deployed CNN model predicts nitrogen on the new images uploaded by the user in real-time.
   - Preprocessed Data: This data preprocessing is stored and can be used as input for further model training and for the prediction.

3. Data Layer

   Handles all the data sets that are used in the system.
   It Includes:

   - Dataset: It includes color images together with labels to train the CNN model.
   - Validation Data: People to fully verify the model as well as to also ascertain the level of accuracy on the part of the model. This guarantees that the data is secured in order to allow the system to be updated by the past prediction for future use.

6.4 Low-level Design

6.4.1 Choice of Design Paradigm

The Structured Systems Analysis and Design Method (SSADM) has been used for the development of this nitrogen detection system. This approach was justified in the introduction thesis as it is a systematic and structured means of analyzing and designing systems for large and complex projects such as this one. This helps in modularity and scalability by ensuring that all components, such as data preprocessing, model training, and prediction storage, are well documented and interconnected.

6.4.2 Data flow diagrams

DFD 1 (Level 1) actually defines where, when, and how the information goes from the level of being uploaded by the user all the way to the level of predicting the nitrogen levels and storing the result in the history database.

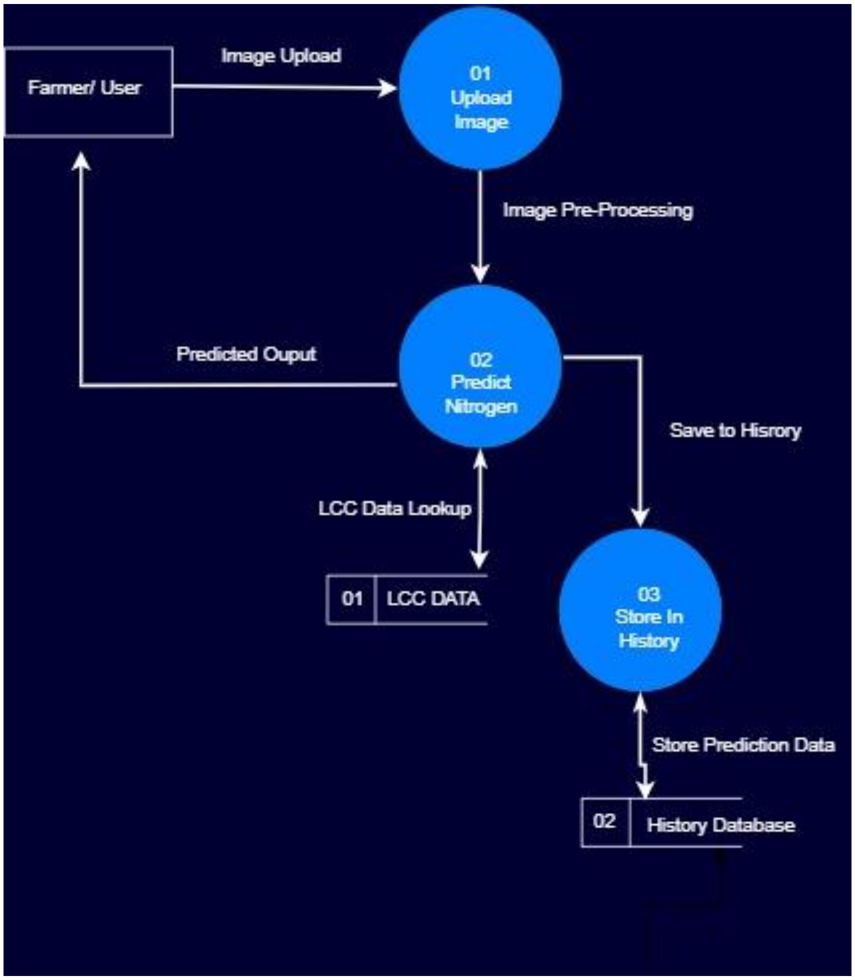


Figure 9

DFD 2 (Level 2) is a detailed process showing how the pre-processing of the image feeds it into a CNN prediction model, comparing LCC data to the results obtained from the model, and storing of results.

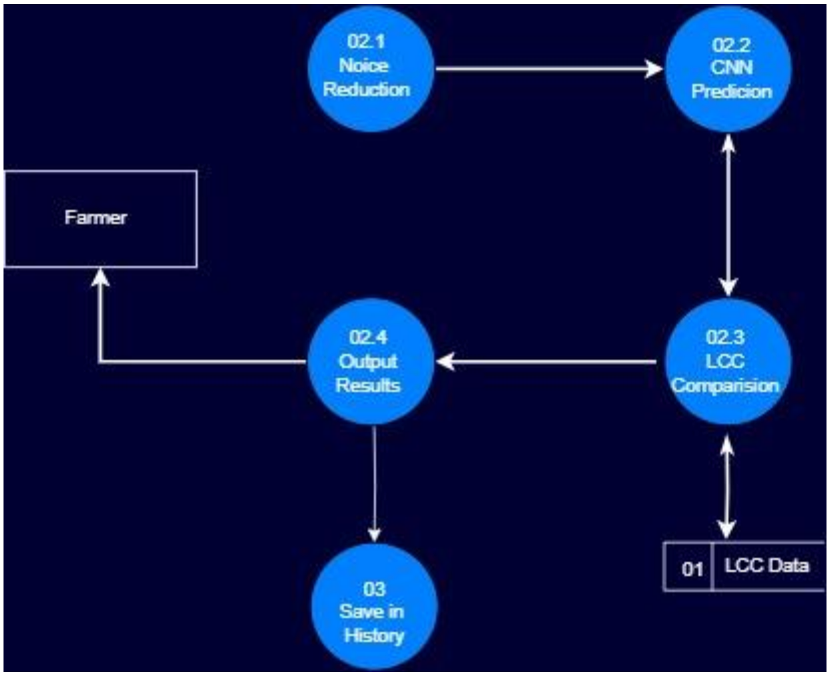


Figure 10

6.5 Design diagrams

6.5.1 Component diagram

The following diagram shows the relation and interconnection between some of the elements of the system. What it usually provides is a preview of how these elements function to address the goal of the system.

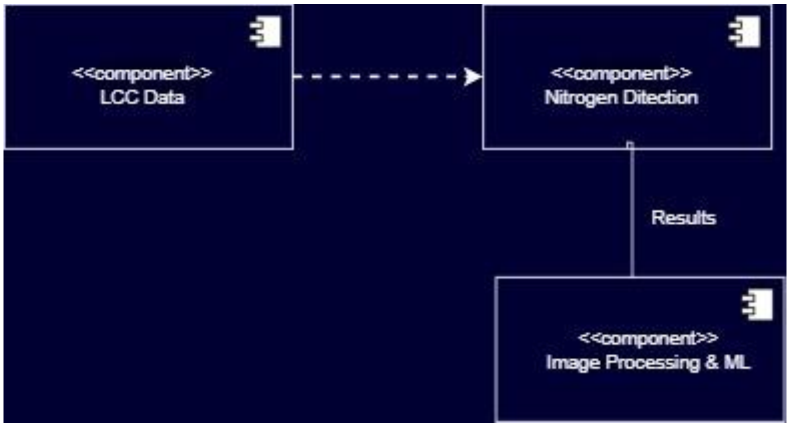


Figure 11

6.5.2 System Process Flowchart

The following is a flowchart that shows how the activities go round within the system. It spells out how a particular goal is going to be accomplished and is used to evaluate the performance of the system.

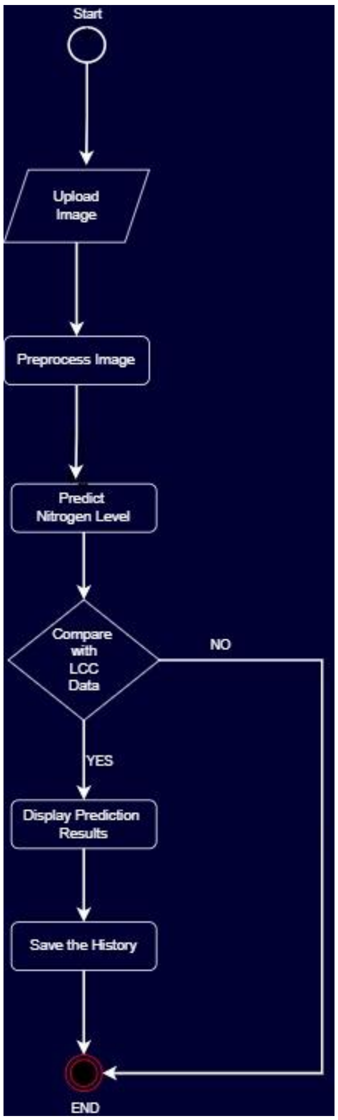


Figure 12

6.5.3 User Interface Design

The following is an example of the nitrogen user interface design which is considered as a prototype of the nitrogen prediction system. The web page has an easily understandable design, where users can upload images, inspect the outcomes, and explore the data history of the system, making it easy for users to use.

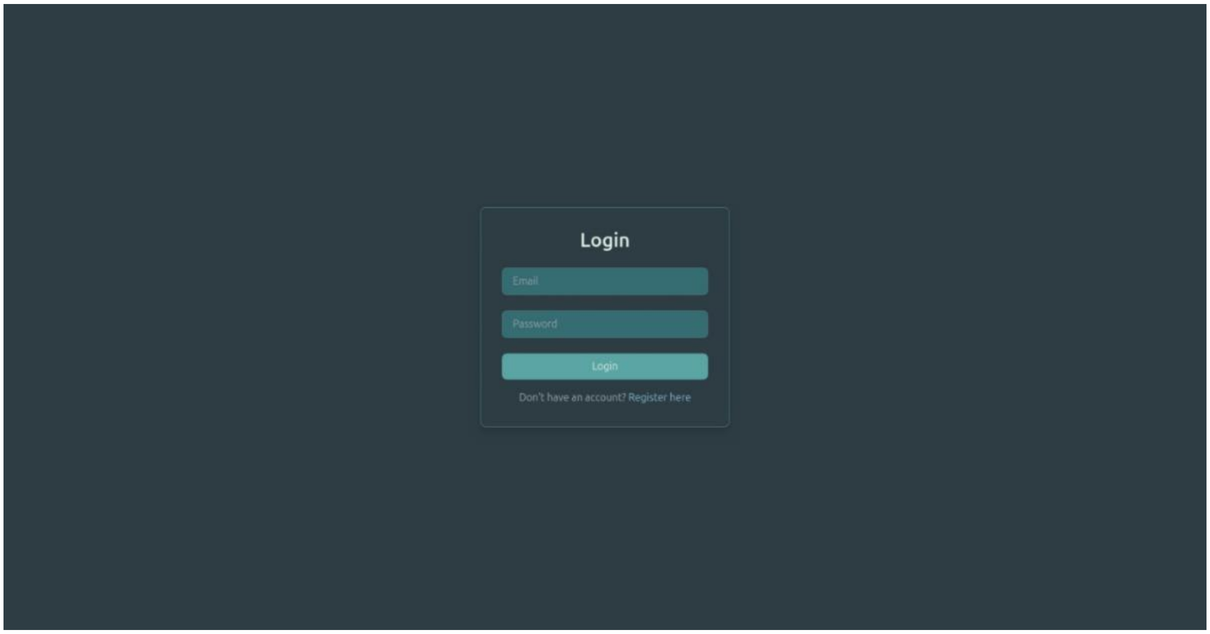


Figure 13

6.6 Chapter Summary

This chapter was a description of the aspects of the nitrogen prediction system design in regard to the system structure, data management, and display. The goals and selections made during the design were explained, as was the flow of the process, which clarifies the operation of the system in terms of speed and soundness of the nitrogen level prediction. In the chapter, one gets an impression of how the system in question is organized to achieve the said aim.

CHAPTER 7 : INITIAL IMPLEMENTATIONS

7.1 Chapter overview

Selection of technologies and implementation of core functionalities of the nitrogen detection system are outlined in this chapter. They justify all the chosen frameworks, the programming languages, libraries, and tools used. We also discuss the implementation of core features, including image preprocessing, model training, nitrogen prediction, and historical data storage, to provide complete insight into how the system works.

7.2 Technology selection

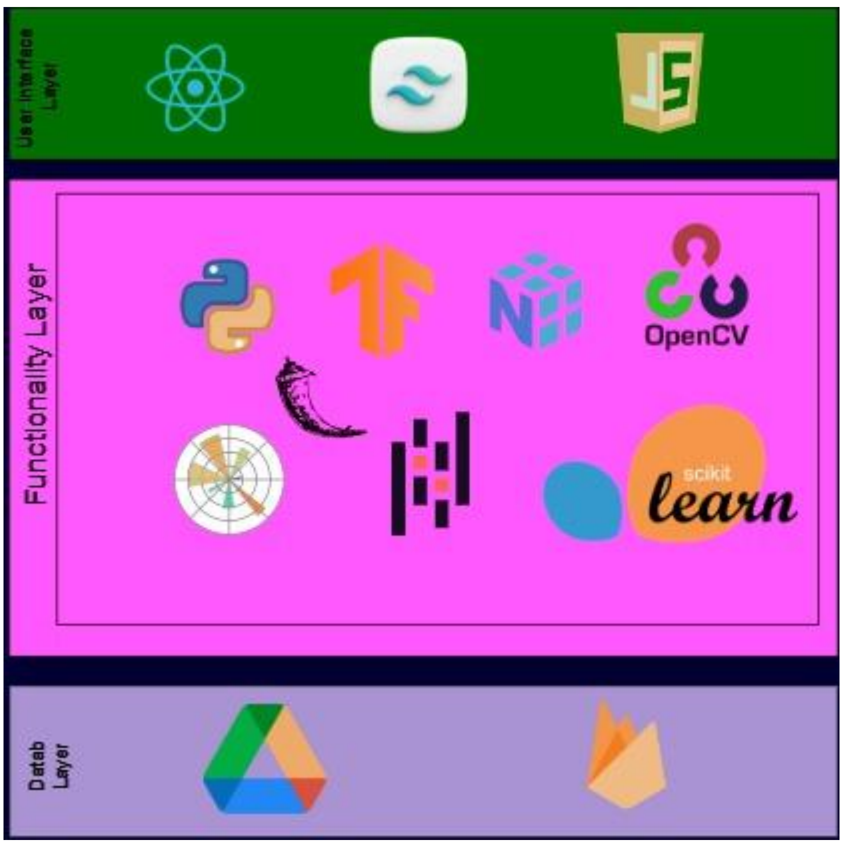


Figure 14

7.2.1 Technology stack

7.2.2 7.2.2 Dataset Selection

This nitrogen detection system uses the Leaf Color Chart (LCC) dataset, the same dataset used for Illesinghe et al. (2023) research. The labeled images of the rice crops, which are categorized by nitrogen levels in the LCC dataset, make it the perfect dataset for training the CNN model used in this project. The dataset standardizes and tones down nitrogen level predictions and makes them consistent, helping towards the yield of rice crops.

7.2.3 Development framework

Frame Works	Purpose
React	The front-end user interface is developed with the help of React.t. A fluid, interactive interface that allows the use of loading images as well as the demonstration of nitrogen prediction outcomes.
TensorFlow	CNN coordination and training is at TensorFlow, and this is what is used when deploying the model. Having tons of libraries for machine learning, it is highly suitable for image recognition and prediction.
Flask	Among the backends that we used, there is Flask, which helps out in handling the requests between the front end and the machine learning model. This ensures there is clear and pleasant interaction and information sharing.
Talwind CSS	When it comes to GUI designing, with the help of the Tailwind CSS, we can avail stylish, well-functional, and highly customizable designs.

Table 20

7.2.4 Programming Languages

Programming Language	Purpose
Python	**Python is also used to code back-end logic, CNN model implementation using TensorFlow, and other data processing tasks along with server-side operations carried out in Flask.**
JavaScript (JS)	**JavaScript is used on the application front-end, together with React, to enhance the interactive capabilities of the system and create a logical and smooth user interface.**

Table 21

7.2.5 Libraries

Library	Purpose
Pandas	**Pandas is a tool of data analysis that primarily focuses on developing tools to facilitate efficient ways of handling data in the form of tables and cleaning datasets and inputs where necessary.**
NumPy	**NumPy is used for computational purposes in the Android application development as well as in handling data structures for the two-dimensional array structure required for the CNN model in image data processing.**
Scikit-learn	**Scikit-learn has been developed with the ability to perform a range of machine learning operations on the data it handles such as data**

	preprocessing, model evaluation and cross-validation.
Matplotlib	**Matplotlib is used for visualization of data for model evaluation and distribution of data sets, plots, and graph formation.**

Table 22

7.2.6 Integrated Development Environment (IDE)

IDE	Purpose
Visual Studio Code (VS Code)	**To build the front end and backend, one needs to use VS Code, which has a great user interface, terminals inbuilt, and great extension support for JavaScript and Python programmers.**
Google Colab	**Crossmodal embedding for lighting estimation is implemented in Google Colab for training and testing of the CNN model. Due to its cloud system, its TensorFlow code is readily executable with select access to adequate GPUs.**

Table 23

7.2.7 Summary of Technology Selection

Component	Tools
Front-End Framework	**React**
Back-End Framework	**Flask**
Machine Learning	**TensorFlow**
Styling	**Tailwind CSS**
Programming Languages	**Python, JavaScript**

Libraries	**Pandas, NumPy, Scikit-learn, Matplotlib**
Integrated Development Environment (IDE)	**Visual Studio Code (VS Code), Google Colab**
Dataset	**LCC Dataset (Illesinghe et al., 2023)**

Table 24

7.3 Implementation of core functionalities

Here the core functionalities of the nitrogen detection system are image preprocessing, model training, nitrogen level prediction, and historical data storage. Having implemented these functionalities using the selected frameworks, libraries, and tools talked about in previous sections.

- Image Preprocessing: We are enhancing the RGB images of rice fields by applying noise reduction and normalization techniques. They are then prepared to input the preprocessed images into the CNN model.

- Model Training: Nitrogen levels were predicted using the LCC dataset and the CNN model was trained using it for this. The following architecture are used and compared.

    1. VGGNet16
    2. ResNet50
    3. InceptionV3

- Nitrogen Prediction: The CNN model is then trained and used to predict nitrogen levels in real time when users upload images into the tool. To verify the accuracy, the predictions are compared to the LCC data.

- Noise Reduction: If user Uploads the image the system will do the noise reduction.

- Historical Data Storage: History log stores predicted outputs where users can see previous nitrogen predictions to make it easy to make an informed decision over time.

- These functionalities are successfully implemented in the system, which ensures that the system accurately and efficiently predicts nitrogen levels with a user-friendly interface,

```python
import os
import pandas as pd
from sklearn.model_selection import train_test_split

# Path to the directory where your dataset is located
dataset_dir = '/content/drive/MyDrive/Soil Nitrogen Levels'

# List all sub-directories (each representing a class)
classes = os.listdir(dataset_dir)

# Initialize lists to hold file paths and labels
file_paths = []
labels = []

# Iterate through the class directories and gather image paths and corresponding labels
for class_name in classes:
    class_dir = os.path.join(dataset_dir, class_name)
    for img_name in os.listdir(class_dir):
        file_paths.append(os.path.join(class_dir, img_name))
        labels.append(class_name)  # The label is the class name

# Convert to a pandas DataFrame for easier manipulation
data = pd.DataFrame({
    'file_path': file_paths,
    'label': labels
})

# Step 1: Split the dataset into 80% training, 10% validation, 10% test
train_data, temp_data = train_test_split(data, test_size=0.2, stratify=data['label'], random_state=42)
validation_data, test_data = train_test_split(temp_data, test_size=0.5, stratify=temp_data['label'],
random_state=42)

# Step 2: Save the datasets to CSV files (optional)
train_data.to_csv('train_datafinal.csv', index=False)
validation_data.to_csv('validation_datafinal.csv', index=False)
test_data.to_csv('test_datafinal.csv', index=False)

# Print the number of images in each set
print(f"Number of images in the training set: {train_data.shape[0]}")
print(f"Number of images in the validation set: {validation_data.shape[0]}")
print(f"Number of images in the test set: {test_data.shape[0]}")


from tensorflow.keras.preprocessing.image import ImageDataGenerator


image_size = (224, 224)

# Data augmentation and rescaling for the training data
train_datagen = ImageDataGenerator(
    rescale=1./255,
    rotation_range=40,
    width_shift_range=0.2,
    height_shift_range=0.2,
    shear_range=0.2,
    zoom_range=0.2,
    horizontal_flip=True
)

# Only rescaling for validation and test data
test_val_datagen = ImageDataGenerator(rescale=1./255)

# Load training data from CSV
train_generator = train_datagen.flow_from_dataframe(
    dataframe=train_data,
    x_col='file_path',
    y_col='label',
    target_size=image_size,
    batch_size=32,
    class_mode='categorical'
)

# Load validation data from CSV
validation_generator = test_val_datagen.flow_from_dataframe(
    dataframe=validation_data,
    x_col='file_path',
    y_col='label',
    target_size=image_size,
    batch_size=32,
    class_mode='categorical',
    shuffle=False
)

# Load test data from CSV
test_generator = test_val_datagen.flow_from_dataframe(
    dataframe=test_data,
    x_col='file_path',
    y_col='label',
    target_size=image_size,
    batch_size=32,
    class_mode='categorical',
    shuffle=False
)
```

Figure 15

Data Split and Image process

```python
import cv2

def denoise_image(input_path, output_path):
    image = cv2.imread(input_path)
    if image is None:
        print("Error: Image not found.")
        return
    denoised = cv2.fastNlMeansDenoisingColored(image, None, 10, 10, 7, 21)
    cv2.imwrite(output_path, denoised)
    print(f"Denoised image saved to {output_path}")

if __name__ == "__main__":
    input_image = "test.jpg"
    denoised_image = "denoised.jpg"
    denoise_image(input_image, denoised_image)
```

Figure 16

Denoiser

```python
# Load the VGG16 model, excluding the top fully connected layers
base_model = VGG16(weights='imagenet', include_top=False, input_shape=(224, 224, 3))

# Add custom top layers to the VGG16 model
x = base_model.output
x = Flatten()(x)
x = Dense(256, activation='relu')(x)
x = Dropout(0.5)(x)

# Use the number of classes from train_generator's class_indices
num_classes = len(train_generator.class_indices)

# Add the final softmax layer for classification
predictions = Dense(num_classes, activation='softmax')(x)

# Create the final model
model = Model(inputs=base_model.input, outputs=predictions)

# Freeze the convolutional base of VGG16 to prevent training
for layer in base_model.layers:
    layer.trainable = False

# Compile the model
model.compile(optimizer=Adam(learning_rate=0.0001), loss='categorical_crossentropy', metrics=
['accuracy'])

# Print model summary
model.summary()

# Add callbacks
early_stopping = EarlyStopping(monitor='val_loss', patience=5, restore_best_weights=True)

# Use the new '.keras' format for saving the model
model_checkpoint = ModelCheckpoint('/content/drive/MyDrive/mareerbest_modelVGG16.keras',
monitor='val_loss', save_best_only=True)
# Train the model with validation
history = model.fit(
    train_generator,
    steps_per_epoch=train_generator.samples // train_generator.batch_size,
    validation_data=val_generator,
    validation_steps=val_generator.samples // val_generator.batch_size,
    epochs=20,  # Adjust as needed
    callbacks=[early_stopping, model_checkpoint]
)
# Unfreeze some layers of VGG16 for fine-tuning
# We'll unfreeze the last 4 layers of VGG16
for layer in base_model.layers[-4:]:
    layer.trainable = True

# Compile the model with a lower learning rate
model.compile(optimizer=Adam(learning_rate=0.00001), loss='categorical_crossentropy', metrics=
['accuracy'])

# Add callbacks
early_stopping = EarlyStopping(monitor='val_loss', patience=5, restore_best_weights=True)
model_checkpoint = ModelCheckpoint('/content/drive/MyDrive/best_modelVGG16_finetuned.keras',
monitor='val_loss', save_best_only=True)

# Retrain the model with fine-tuning
history = model.fit(
    train_generator,
    steps_per_epoch=train_generator.samples // train_generator.batch_size,
    validation_data=val_generator,
    validation_steps=val_generator.samples // val_generator.batch_size,
    epochs=10,  # Adjust the number of epochs as needed
    callbacks=[early_stopping, model_checkpoint]
)
```

x

Figure 17

VGG16 Model Implementation Image

```python
# Load the InceptionV3 model, excluding the top fully connected layers
base_model = InceptionV3(weights='imagenet', include_top=False, input_shape=(299, 299, 3))

# Freeze the convolutional base of InceptionV3 to prevent training initially
for layer in base_model.layers:
    layer.trainable = False
# Add custom layers on top of InceptionV3
x = base_model.output
x = GlobalAveragePooling2D()(x)  # Use global average pooling instead of Flatten
x = Dense(256, activation='relu')(x)
x = Dropout(0.5)(x)

# Output layer with softmax activation for multi-class classification
num_classes = len(train_generator.class_indices)
predictions = Dense(num_classes, activation='softmax')(x)

# Create the final model
model = Model(inputs=base_model.input, outputs=predictions)
# Compile the model
model.compile(optimizer=Adam(learning_rate=0.0001), loss='categorical_crossentropy', metrics=
['accuracy'])
# Unfreeze the last few layers of InceptionV3 for fine-tuning
for layer in base_model.layers[-20:]:  # Unfreezing the last 20 layers
    layer.trainable = True

# Re-compile the model with a lower learning rate
model.compile(optimizer=Adam(learning_rate=0.00001), loss='categorical_crossentropy', metrics=
['accuracy'])

# Add callbacks
early_stopping = EarlyStopping(monitor='val_loss', patience=5, restore_best_weights=True)
model_checkpoint = ModelCheckpoint('/content/drive/MyDrive/best_inceptionv3_model.keras',
monitor='val_loss', save_best_only=True)

# Train the model
history = model.fit(
    train_generator,
    steps_per_epoch=train_generator.samples // train_generator.batch_size,
    validation_data=val_generator,
    validation_steps=val_generator.samples // val_generator.batch_size,
    epochs=20,  # Adjust the number of epochs as needed
    callbacks=[early_stopping, model_checkpoint]
)
```

Figure 18

InceptionV3 Implementation Model

```python
# Load the ResNet50 model, excluding the top fully connected layers
base_model = ResNet50(weights='imagenet', include_top=False, input_shape=(224, 224, 3))

# Freeze the convolutional base of ResNet50 to prevent training initially
for layer in base_model.layers:
    layer.trainable = False
# Add custom layers on top of the ResNet-50 base model
x = base_model.output
x = Flatten()(x)
x = Dense(256, activation='relu')(x)
x = Dropout(0.5)(x)

# Output layer with softmax activation for multi-class classification
num_classes = len(train_generator.class_indices)
predictions = Dense(num_classes, activation='softmax')(x)

# Create the final model
model = Model(inputs=base_model.input, outputs=predictions)
# Compile the model
model.compile(optimizer=Adam(learning_rate=0.0001), loss='categorical_crossentropy', metrics=
['accuracy'])
# Print model summary
model.summary()
# Add callbacks
early_stopping = EarlyStopping(monitor='val_loss', patience=5, restore_best_weights=True)
model_checkpoint = ModelCheckpoint('/content/drive/MyDrive/mareerbest_resnet_model.keras',
monitor='val_loss', save_best_only=True)
# Train the model
history = model.fit(
    train_generator,
    steps_per_epoch=train_generator.samples // train_generator.batch_size,
    validation_data=val_generator,
    validation_steps=val_generator.samples // val_generator.batch_size,
    epochs=10,  # Adjust as needed
    callbacks=[early_stopping, model_checkpoint]
)
# Evaluate the model on the test set
test_loss, test_accuracy = model.evaluate(test_generator, steps=test_generator.samples //
test_generator.batch_size)
print(f"Test Loss: {test_loss}, Test Accuracy: {test_accuracy}")

# Extract the training and validation accuracy from the history object
train_acc = history.history['accuracy']
val_acc = history.history['val_accuracy']

# Unfreeze the last few layers of ResNet-50 for fine-tuning
# Let's unfreeze the last 20 layers of ResNet-50
for layer in base_model.layers[-10:]:
    layer.trainable = True

# Re-compile the model with a lower learning rate to avoid overfitting during fine-tuning
model.compile(optimizer=Adam(learning_rate=0.00001), loss='categorical_crossentropy', metrics=
['accuracy'])


# Train the model
history = model.fit(
    train_generator,
    steps_per_epoch=train_generator.samples // train_generator.batch_size,
    validation_data=val_generator,
    validation_steps=val_generator.samples // val_generator.batch_size,
    epochs=10,  # Adjust as needed
    callbacks=[early_stopping, model_checkpoint]
```

Figure 19

ResNet50 Model Implementation Image

7.4 User interface

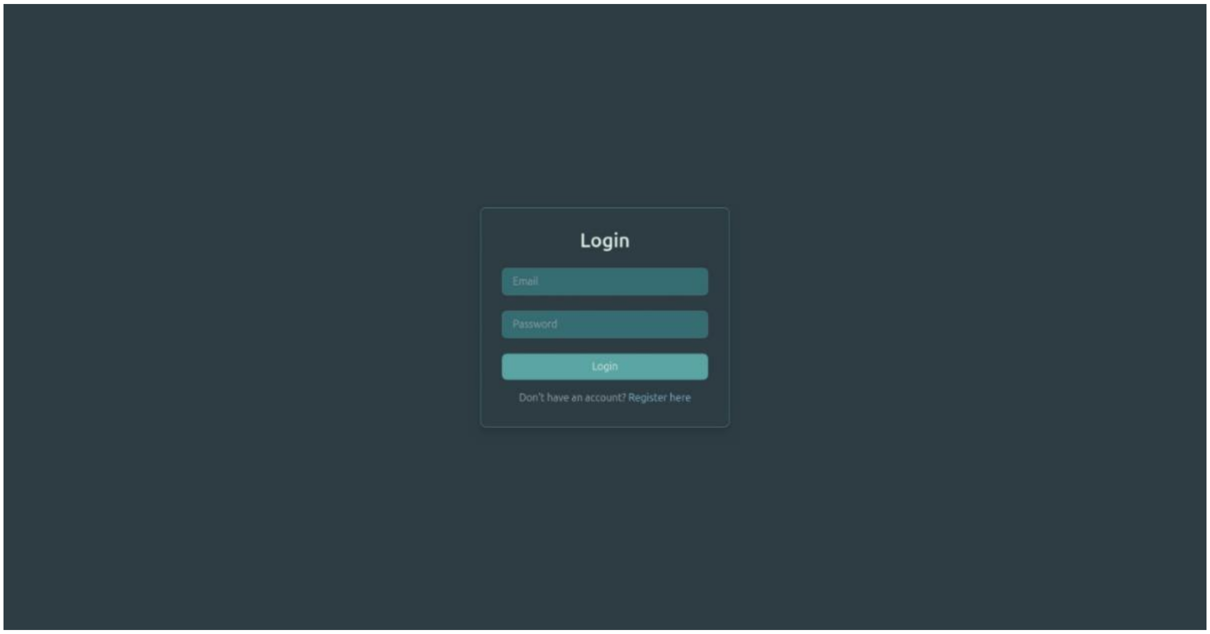


Figure 20

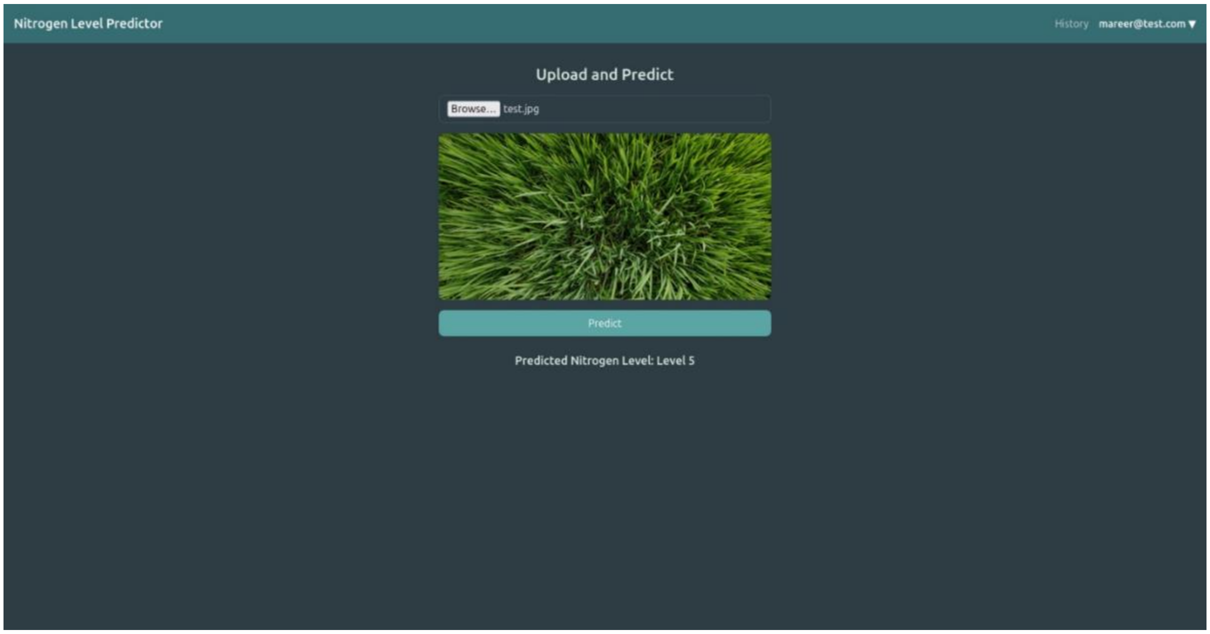


Figure 21

7.5 Chapter Summary

In this chapter, the technologies that were used to develop the nitrogen detection system were summarized: React, TensorFlow, Flask, and Tailwind CSS. To this effect, several key libraries were employed with the programming languages Python and JavaScript. All the core functionalities of the system, such as image preprocessing, CNN model training, real-time prediction of nitrogen, and storing historical data, were implemented as successfully as possible, making it a highly efficient as well as accurate system for the detection of nitrogen in rice crops.

CHAPTER 8 : TESTING

8.1 Chapter Overview

This chapter established the strategies used in testing the nitrogen detection system. It includes functions and non-function testing, modeling and benchmarking, precision, dependability, and usability testing. The chapter also covers challenges that were encountered during testing as well. The objective of this chapter is to ensure that the system developed addresses the functional, performance, and usability specifications that have been articulated in other chapters.

8.2 Objectives of Testing and its Goals

The purpose of testing was to verify that the nitrogen detection system, when functioning properly, performs as it is supposed to. The following objectives have been established to ensure this goal is met:

Check all of the components of the system to assure all are operating correctly and on creating accurate nitrogen predictions.

Make sure that all the required ones ("must have") and kindly ones ("should have") specified in the system requirements specification (SRS) are fulfilled.

Make sure the "must" "have" and should have nonfunctional requirements, i.e., performance, scalability, and security, are met as defined in SRS.

Conducting a comprehensive testing process within which some potential issues, bugs, or defects are detected and reported.

8.3 Testing Criteria

The testing approach for the nitrogen detection system is divided into two key areas: functional and structural tested.

Criteria	Description
Functional Testing	Validates the system as per the SRS so all the functional requirements specified in the SRS are fulfilled.
Structural Testing	It evaluates the system's non-functional elements like performance, security, and scalability against the standards it needs to achieve.

Table 25

This testing framework guarantees to explore both the system's corresponding structure and capabilities.

8.4 Model Testing

The model testing phase ensures that the nitrogen detection system's machine learning model performs accurately and reliably in real-world scenarios. The dataset was divided into two portions: 80% for training the CNN model,10percentage for Validation and 10% for testing. This 80:10:10 split allows for sufficient data to be used for model learning while retaining a portion of the data to validate the model's performance.

During the testing phase, the system's predictions were compared against known nitrogen levels in the test dataset. The results were evaluated using metrics such as accuracy, precision, recall, and F1 score to ensure the model’s reliability and effectiveness in predicting nitrogen levels in unseen data.

8.5 Model Testing

The model testing phase verifies that the machine learning model of the nitrogen detection system does work correctly and reliably in real-world situations. The system's predictions were then compared against known levels of available nitrogen in the test set. Then the results were evaluated by accuracy, precision, recall, and F1 score to be sure that the model would be capable of predicting nitrogen levels on unseen data.

```python
import numpy as np
from sklearn.metrics import classification_report, f1_score

# Step 1: Get predictions from the model
predictions = model.predict(test_generator)

# Step 2: Convert the predictions to class labels
predicted_classes = np.argmax(predictions, axis=1)

# Step 3: Get true labels from the test generator
true_classes = test_generator.classes  # True labels

# Step 4: Get the class labels from the generator (e.g., ['Level 2', 'Level 3', 'Level 4', 'Level 5'])
class_labels = list(test_generator.class_indices.keys())

# Step 5: Classification report with precision, recall, and F1 score for each class
report = classification_report(true_classes, predicted_classes, target_names=class_labels)
print(report)

# Step 6: Compute the overall F1-Score (macro and weighted)
f1_macro = f1_score(true_classes, predicted_classes, average='macro')
f1_weighted = f1_score(true_classes, predicted_classes, average='weighted')

print(f"Macro F1-Score: {f1_macro}")
print(f"Weighted F1-Score: {f1_weighted}")

# Confusion Matrix
import numpy as np
import matplotlib.pyplot as plt
from sklearn.metrics import confusion_matrix
import seaborn as sns

# Step 1: Get predictions from the model
predictions = model.predict(test_generator)

# Step 2: Convert predictions to class labels
predicted_classes = np.argmax(predictions, axis=1)

# Step 3: Get true labels from the test generator
true_classes = test_generator.classes

# Step 4: Generate the confusion matrix
conf_matrix = confusion_matrix(true_classes, predicted_classes)

# Step 5: Plot the confusion matrix
plt.figure(figsize=(8, 6))
sns.heatmap(conf_matrix, annot=True, fmt='d', cmap='Blues',
xticklabels=test_generator.class_indices.keys(), yticklabels=test_generator.class_indices.keys())
plt.title('Confusion Matrix')
plt.ylabel('True Label')
plt.xlabel('Predicted Label')
plt.show()

# AUC
import numpy as np
from sklearn.metrics import roc_auc_score
from sklearn.preprocessing import label_binarize

# Step 1: Get predictions from the model
predictions = model.predict(test_generator)

# Step 2: Get true class labels from the test data generator
true_classes = test_generator.classes  # True labels

# Step 3: Convert the true class labels to one-hot encoding (for multi-class AUC)
n_classes = len(test_generator.class_indices)  # Number of classes (4 in your case: Level 02, 03, 04,
05)
true_classes_one_hot = label_binarize(true_classes, classes=[0, 1, 2, 3])  # One-hot encode true labels

# Step 4: Calculate AUC for each class
# Since we are dealing with probabilities (from `model.predict()`), we don't need to convert
predictions.
# The predictions should already be in a probability format.
auc = roc_auc_score(true_classes_one_hot, predictions, average='macro', multi_class='ovr')

print(f"Macro-average AUC: {auc}")
```

Figure 22

Model Testing Screenshot Image

8.6 Confusion Matrix

The performance of the nitrogen detection system in classifying nitrogen levels has been assessed using a confusion matrix for each of the three CNN models: here these matrix are each models how perfoming with each classes in the dataset.

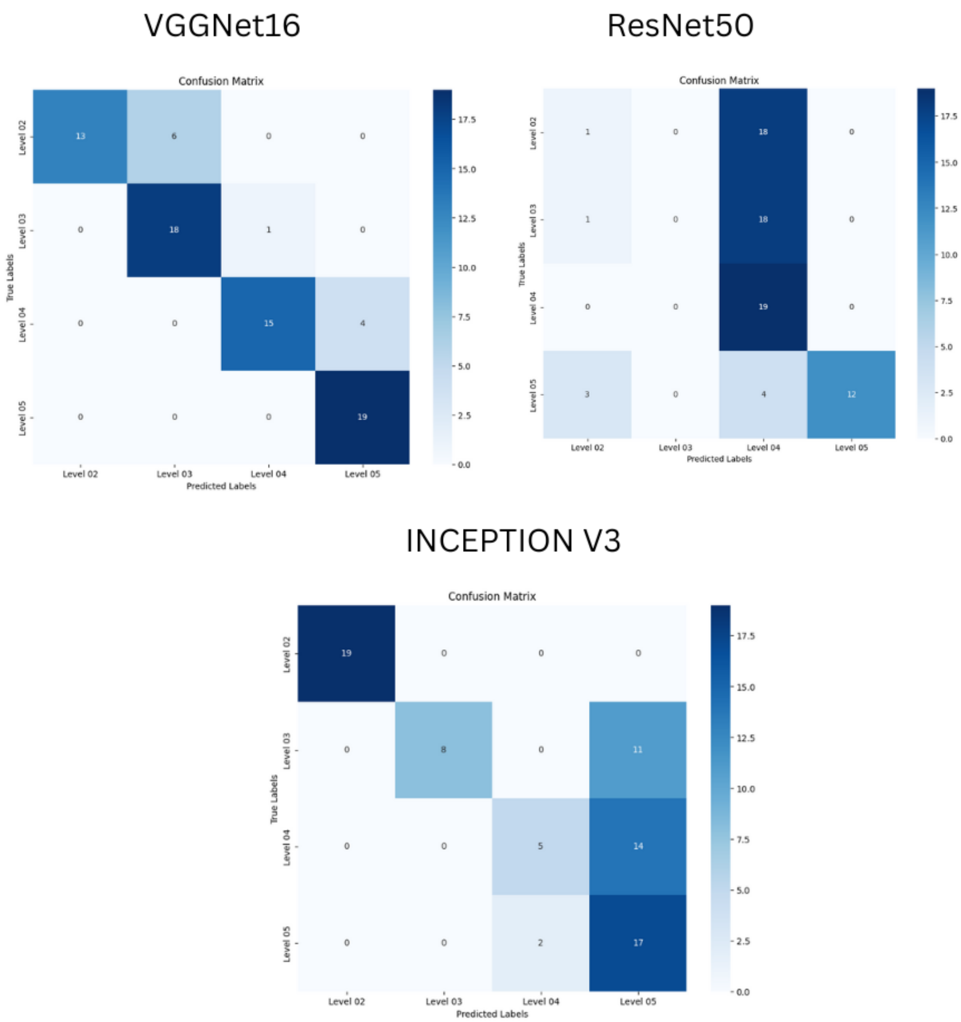


Figure 23

8.7 Performance / Test Results

VGGNet16

	precision	recall	f1-score	support
Level 02	1.00	0.68	0.81	19
Level 03	0.75	0.95	0.84	19
Level 04	0.94	0.79	0.86	19
Level 05	0.83	1.00	0.90	19
accuracy			0.86	76
macro avg	0.88	0.86	0.85	76
weighted avg	0.88	0.86	0.85	76

Macro F1-Score: 0.8529035160575859
Weighted F1-Score: 0.8529035160575859

3/3 40s 9s/step
Macro AUC: 0.9845
Weighted AUC: 0.9845

ResNet50

	precision	recall	f1-score	support
Level 02	0.20	0.05	0.08	19
Level 03	0.00	0.00	0.00	19
Level 04	0.32	1.00	0.49	19
Level 05	1.00	0.63	0.77	19
accuracy			0.42	76
macro avg	0.38	0.42	0.34	76
ighted avg	0.38	0.42	0.34	76

cro F1-Score: 0.3361765922249793
ighted F1-Score: 0.3361765922249793

Macro AUC: 0.9305
Weighted AUC: 0.9305

INCEPTION V3

	precision	recall	f1-score	support
Level 02	1.00	1.00	1.00	19
Level 03	1.00	0.42	0.59	19
Level 04	0.71	0.26	0.38	19
Level 05	0.40	0.89	0.56	19
accuracy			0.64	76
macro avg	0.78	0.64	0.63	76
weighted avg	0.78	0.64	0.63	76

Macro F1-Score: 0.6336462565970762
Weighted F1-Score: 0.6336462565970763

Macro AUC: 0.9305
Weighted AUC: 0.9305

Figure 24

Figure screen shots of test results

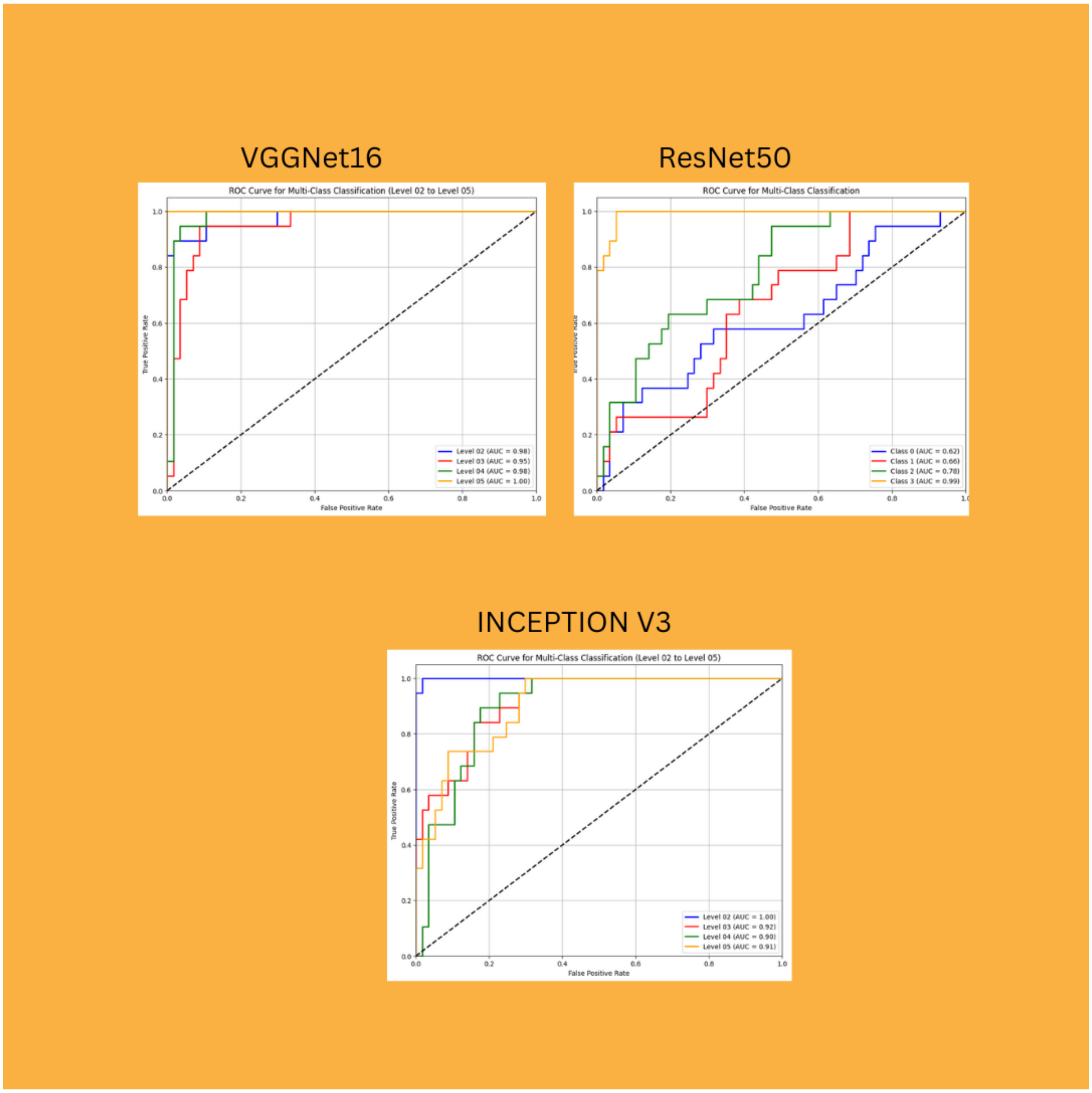


Figure 25

Figure screen shots of roc graph

8.8 Benchmarking

The performance of the nitrogen detection system was evaluated using three CNN models: ResNet 50, VGGNet 16, and Inception V3. Accuracy, precision, recall, F1 score, and AUC were measured for the key metrics. The table below summarizes the performance of each model:

Model	Accuracy	Precision	Recall	F1 Score	AUC
VGGNet 16	86%	0.88	0.86	0.82	0.98
Inception V3	64%	0.78	0.64	0.66	0.93
ResNet 50	33%	0.38	0.42	0.43	0.77

Table 26

- ✓ The highest accuracy (86%) and balanced F1 scores across all nitrogen levels were achieved using the VGGNet 16, which was selected as the most appropriate model for this system.
- ✓ The performance of Inception V3 was relatively moderate (accuracy of 64% and balanced F1 scores), but some misclassification between levels affected its result.
- ✓ Lowest performance, only 33% accuracy, and low F1 scores meant that there are many significant misclassifications.

Based on these results, it is suggested that VGGNet 16 is the best model for nitrogen level detection in this system.

Among all the models used in this project, VGGNet 16 produced the highest overall accuracy of 86%, a figure higher than that of ResNet 50 and Inception V3 models with lower rates of accuracy. The result of this study corroborates evidence from other research, including that of Illesinghe et al. (2023), through which UAVs yielded an accuracy of 86.5% in measuring nitrogen concentration in paddy fields. The two works endorse that machine learning models are highly precise in predicting concentrations of nitrogen and that VGGNET 16 is the most accurate model for this system.

ID	FR ID	User Action	Expected Outcome	Actual Outcome	Result
1	FR 1	User uploads RGB images for nitrogen prediction.	The system should predict nitrogen levels accurately.	Accurate nitrogen level predictions were generated.	Passed
2	FR 2	User uploads images through the system interface.	The system should allow successful image uploads.	Image uploads were handled correctly by the system.	Passed
3	FR 3	User views predicted nitrogen levels after uploading images.	The system should display predictions in an understandable format.	Predictions were displayed clearly and in an easy-to-read format.	Passed
4	FR 5	User initiates a model update to improve prediction accuracy.	The system should support updates to the CNN model.	The system successfully supported model updates.	Passed
5	FR	User uploads a	The system should	Error messages were	Passed

Table 27	9	corrupted image or invalid file format.	display an appropriate error message.	correctly displayed for invalid inputs.	

8.9 Functional Testing

Table n - Functional Testing Results

8.10 Non-Functional Testing

Non-functional requirements obtained from the Non-Functional Requirements section have been validated as explained below.

8.11 Accuracy Testing

Thus, the performance of the nitrogen detection system was examined using such measures as precision, recall, and F1, the results of which are described in Sections 8.6 and 8.7. This makes a clear evaluation of the performance of the system for the prediction of nitrogen levels depending on different CNN models possible.

8.12 Reliability Testing

The main measurement of validation for the nitrogen detection system was through confusion matrices, which gave an indication of the model criteria for classifying different levels of nitrogen. Further elaboration on the confusion matrix can be made known to the readers in Section 8.6.

8.13 User Experience Testing

Usability tests were performed for upholding the utility of the nitrogen detection system, which is the perceived usability of the system by its most targeted audience, farmers. The following key areas were tested, and space was also added to upload image proof during testing:

Interface Simplicity: It was conducted to validate user interface as well as ensure its level of interactivity was friendly. The participants of the study also expressed that it takes little efforts and training to master the operation of the interface.

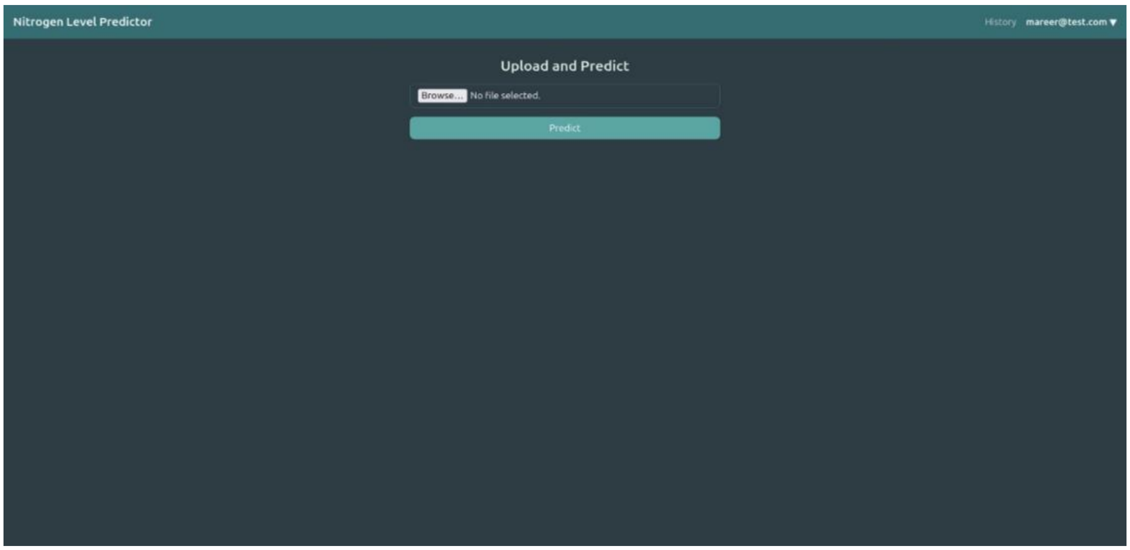


Figure 26

Ease of Image Upload: Complication-free for the usage, users were able to upload the RGB images for nitrogen prediction. Another feature added provided the functionality to send image proof, which helped in case of various problems. The system was also able to manage instances where files were in the wrong format by informing the user through a comprehensible and fast message.

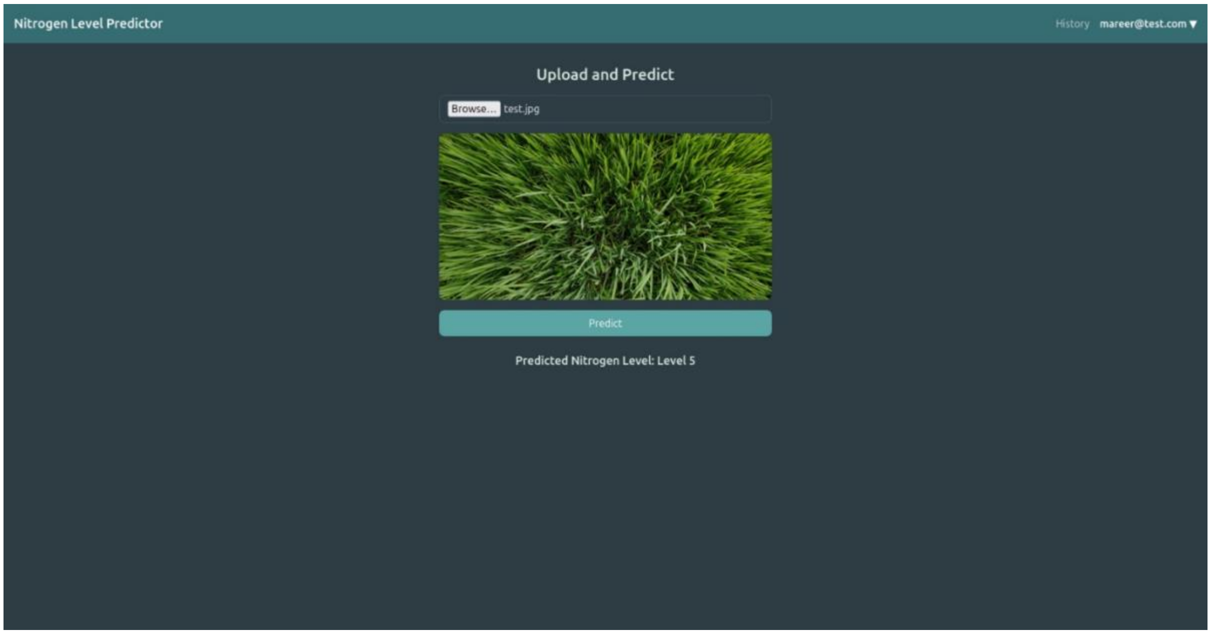


Figure 27

Presentation of Results: These predictions of nitrogen levels in the given areas were presented in a simple interface to allow the users to be able to quickly interpret and address the results. Consumers for the navigation were pleased by the concise and straightforward results returned.

System Responsiveness: The general effectiveness of the system, or more particularly its ability to convey error messages, was assessed. There were immediate error messages each time there was a problem so the users would know what went wrong during the uploading or predicting task

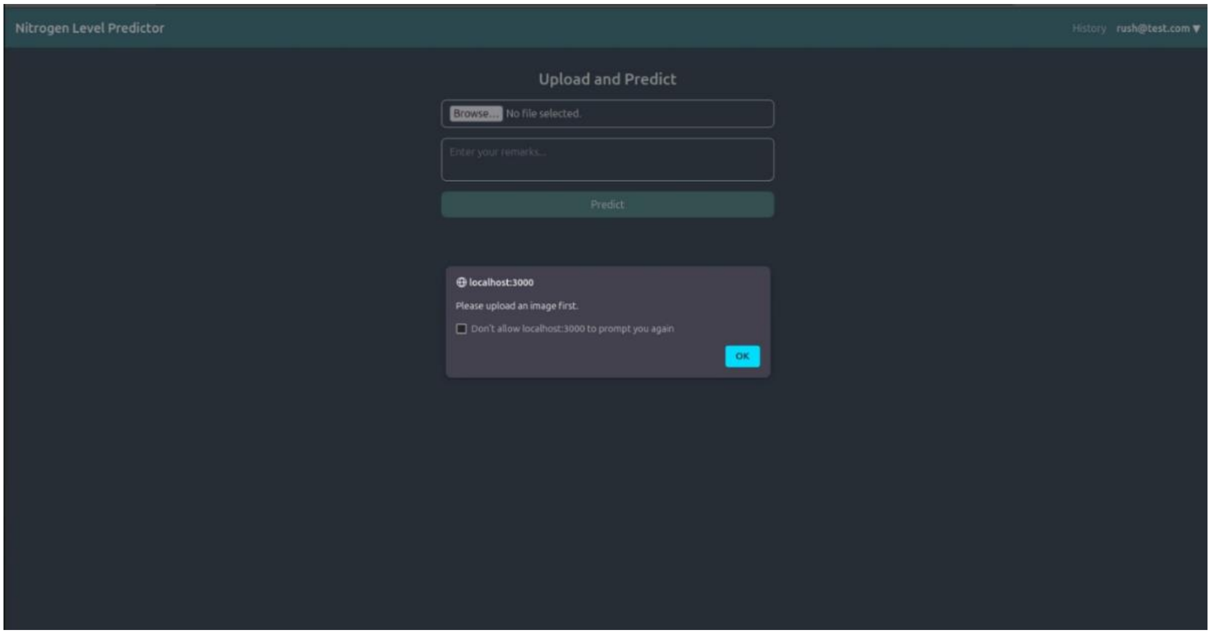


Figure 28

8.14 Limitation of the Testing Process

While the testing process for the nitrogen detection system was thorough, certain limitations were encountered:

Limited Dataset Variety: The present system was evaluated and experimented with the help of a special database (LCC dataset). This restricts the possibilities of testing since the system might behave in another way with the completely different data set or under different environmental conditions that were not included in the test data set.

Overfitting in VGG16: It was observed through model testing that the VGG16 model was overfit with training accuracy higher than the validation accuracy. This may mean that the model is not very accurate in terms of extrapolation of the before-mentioned characteristics to new datasets to be used in the future.

User Testing Sample Size: Of the selected participants, most of them were farmers, and the testing of user experience was done on a limited number of participants. Though feedback received was positive, feedback from a wider pool, especially those with fewer technology interactions, would give a broader view of the functionality of this system.

8.15 Chapter Summary

In conclusion, the chapter provided critical assessments on the nitrogen detection system by performing tests on the functionality and precision of the system. Other sub-topics like reliability and usability were also considered. As the assessments showed there are sources of benefits, The system itself has problems like overfitting and variability of datasets. These results offer directions for further advancements. In general, it can be claimed that the system fulfilled its principal necessities and aims.

CHAPTER 9 : CONCLUSION

9.1 Chapter Overview

This chapter is the concluding section of the project and discusses the completion of research objectives, the use of the attained knowledge, and the envisaged difficulties with the development of the project. It also points out the deviations of the issues, which did not fit the initial project plan, the constraints faced, and advances made to the body of knowledge. Additionally, the chapter explains possible future additions to the system and comprises a summary of this research study's success in achieving its objectives.

Table 28

9.2 Outcome of Research Aims and Objectives

The following table outlines how each research objective (RO) has been achieved, along with the associated learning outcomes (LOs) and their status:

Research Objective (RO)	Learning Outcomes (LO)	Status
RO1: Problem Definition	LO1, LO2	Achieved
RO2: Literature Review	LO1, LO4, LO8	Achieved
RO3: Requirement Elicitation	LO1, LO3, LO5, LO8	Achieved
RO4: Design	LO1, LO5, LO8	Achieved
RO5: Development	LO1, LO5, LO7, LO8	Achieved
RO6: Testing - Quantitative	LO4, LO5	Achieved
RO7: Documentation	LO1, LO2, LO4	Achieved

9.3 Knowledge Application from the Course

The academic content and the practical experiences that the course offered proved to be valuable to accomplishing this study. The following table shows how the elements of the course were incorporated into the overall research process at different stages.

Table 29

Modules	**Utilized in**
Decision Analytics Module	Used in developing and training CNN models for nitrogen prediction, as shown in the study.
Software Development and Application Modelling Module	Applied in system design and development, making the system modular, scalable, and maintaining software engineering practices.
Web Development Module	Used in building the user interface and integrating it with the backend, ensuring user-friendly interactions.

9.4 Use of Existing Skills

Before venturing into this research, the author had a relatively small amount of background knowledge on machine learning (ML) and artificial intelligence (AI). However, the author of this article had certain proficiencies in software development, system design, and problem resolution that would help him with this task. These existing skills proved useful in the process of developing the system, but the actual research involved the learning of new ideas relating to ML, data preprocessing, and model evaluation methods, which the author was previously unaware of.

9.5 Use of New Skills

The author learned a few new things over the course of this research; she learned new things about machine learning (ML) and artificial intelligence (AI). The author learned the ability to build and train CNNs, data preprocessing, model evaluation, and measurement using parameters such as accuracy, precision, ROC, and curves. Also, the author was able to learn how to work on the direct application of various ML frameworks, including TensorFlow, and to incorporate them together with front-end platforms in order to design a workable system. Many of these newly developed skills were crucial for the completion of the nitrogen detection system and the realization of the research goals.

9.6 Demonstrated learning outcomes

Below table is showing how the learning outcomes were attained because of the accomplishment of the project.

Table 30

Learning Outcome (LO)	Achievement
LO1, LO2	The overall planning and time management of the project proposal were greatly conducted, as it was mainly centered on nitrogen detection employing CNN models.
LO2, LO3	A CNN-based system was developed and people coded, annotated, and documented using recognizable analysis and design methods.
LO3, LO4	Since then, the system was developed according to the proposed model, whose development process was also explained.
LO4	Integral and regression testing were performed, and an assessment of the system's fitness was done.
LO5	This project observed legal, social, and ethical practices as required in the profession.

9.7 issues and difficulties experienced

The following table gives an overview of the major issues that were faced during the project conduct and the measures that were applied to deal with them.

Table 31

Challenge	Mitigation
Lack of prior experience in ML/AI	Gathered the necessary knowledge via the internet source and applied it during the project.
Overfitting in CNN models	Some regularization measures were used through the models to eliminate overfitting, and the overall model was further optimized.
Difficulty in model selection	Used three models (VGGNet 16, ResNet 50, and Inception V3) to predict the accuracy and selected the best model.
Handling large datasets	When processing data or testing the model, decided to use cloud computing resources, namely Google Colab.
Time management	Has developed better planning skills and could use many project management techniques to ensure that the project was on schedule.

9.8 Deviations

In the beginning of this project, the original scope was to evaluate shortages of nitrogen (N), phosphorus (P), and potassium (K), besides evaluating nitrogen intensity in rice paddy fields. However, since this topic is very generic, a lot of research about NPK deficiency and the nitrogen rates in particular would be required from the part of the author. Due to the time limitation that was allocated to the writer for this work, it was not very advisable to delve much into all these areas of concern.

To overcome these limitations, the project was scaled down to the small-scale task of predicting nitrogen health in rice paddies using the LCC dataset. Such a specific target made it possible for the author to finish the research before the time elapsed and at the same time produce useful information on the prediction of nitrogen.

9.9 Limitations of the Research

Several limitations were identified during the course of this research:

- Dataset Dependency: The system was trained and validated using the LCC dataset, which may not represent the full range of environmental conditions or crop variations. This could limit the model’s generalizability to other datasets or rice varieties.

- Potential Overfitting in CNN Models: While the VGG16 model performed well, there were signs that it may have overfitted the training data. However, this was not conclusive. Future improvements, such as additional regularization techniques or training with more diverse data, could help mitigate any potential overfitting and improve the model's generalization.

- User Testing Sample Size: The user experience testing was limited to a small group of people. Expanding the testing to include a wider range of users with different levels of technical expertise would provide more comprehensive insights into the system's usability.

9.10 Future Enhancements

Several areas of the nitrogen detection system could be improved or expanded in future research and development:

- Improving Model Accuracy: This is an area for future work; it may be possible to produce even more accurate estimates of the level of nitrogen through continued tuning of the CNNs featured in this paper or through the exploration of different architectures for deep neural networks, such as some of the deeper structures, for the task at hand.
- Expanding the Dataset: One idea for future studies could be to further train the model with an extensive amount of data on varied conditions of the environment, multiple types of soil, and different varieties of rice. This would also improve the scalability and reliability of the system based on the existing study.

- Real-Time Field Application: It would be an improvement if the system were to be deployed as a mobile or drone, which can carry out the nitrogen level real-time analysis in the field and give farmers the prediction count.

- User Interface Improvements: Though the user interaction system has been tested, future versions can also add more accurate visual and logical information for the user, especially for those students who are not very good at programming.

- Broader User Testing: Taking user testing to the next level by inviting a more varied population of farmers and other agricultural industry people would give better feedback and could help to refine the system to suit a wider range of users.

9.11 Achievement of the Contribution to the Body of Knowledge

This research has made a significant contribution to the body of knowledge in several key areas:

- The study proved that CNNs can be used to predict nitrogen levels in rice fields, providing an opportunity to replace LCC readings based on photographs with an automated system.

- Through the use of sophisticated subjects like a machine learning approach to agricultural issues, the research addresses the need and possibility of finding solutions at scale on nitrogen management and more in precision agriculture.

- We used the LCC dataset in both the training and testing of the models to demonstrate the use of open data sources in agriculture. The essence of a data-driven approach is equally highlighted in this contribution to call for proper farming practices.

- The research evaluated several CNN models (VGGNet 16, ResNet 50, Inception V3) in order to give some recommendations concerning the selection and fine-tuning of deep learning models for agricultural purposes. It enhances the information value of the comparison, as it provides performance benchmarks for further work.
- The project provided an useful, easy-to-apply system that farmers can employ to monitor nitrogen levels in a timely manner. This system helps close a gap between high tech and end use that is a valuable contribution to actual farm practice.

9.12 Conclusion Remarks

For this project, a CNN-based nitrogen detection system was successfully built to offer a new way for farmers to evaluate nitrogen content in rice crops. As a result, the LCC dataset and machine learning techniques convert the system into a realistic and efficient approach to optimizing nitrogen management in agriculture. The work also expands the literature on precision farming by showing how machine learning could be used in agricultural practices.

The system has produced good outcome however, the future work includes the steps as follows: Addition of a big data set to make the system compatible to work in a real-time environment. It has been possible to accomplish the chief aims of the project and to create the fundamental groundwork for the additional advancement of the technology for agricultural uses.

References

Illesinghe, R., Arachchi, S.W., Kavikarage, H., Karunarathna, A., Karunanayaka, K., Halloluwa, T. and Rathnayake, U. (2023). Effective Identification of Nitrogen Fertilizer Demand for Paddy Cultivation Using UAVs. *IECO Annual Conference*. [Online]. Available at: https://ieeexplore.ieee.org/document/10456789 [Accessed 2 July 2024].

slam, T., Rizan, R.U.B., Tusher, Y.A., Shafiuzzaman, M., Hossain, M.A. and Galib, S. (2020). Nitrogen Fertilizer Recommendation for Paddies Through Automating the Leaf Color Chart (LCC). *International Journal of Advanced Computer Science and Applications (IJACSA)*. [Online]. Available at: https://ijacsa.org/paper/10.1112 [Accessed 15 July 2024].

Luo, Y., Zan, X., Li, J. and Zhang, X. (2023). Rice Leaf Nitrogen Deficiency Image Classification Model Based on CNN. *Academic Journal of Computing & Information Science*. [Online]. Available at: https://ajcis.org/article/view/345 [Accessed 10 July 2024].

Sethy, P.K., Behera, S.K., Rath, A. and Barpanda, N.K. (2020). Nitrogen Deficiency Prediction of Rice Crop Based on Convolutional Neural Network. *Journal of Ambient Intelligence and Humanized Computing*. [Online]. Available at: https://springer.com/article/1073 [Accessed 20 July 2024].

Zhai, Q., Ye, C., Li, S. and Liu, J. (2022). Rice Nitrogen Nutrition Monitoring Classification Method Based on the Convolution Neural Network Model: Direct Detection of Rice Nitrogen Nutritional Status. *PLOS ONE*. [Online]. Available at: https://journals.plos.org/plosone/article?id=104567 [Accessed 7 July 2024].

Tao, M., Ma, X., Huang, X., Liu, C. and Qi, L. (2020). Smartphone-Based Detection of Leaf Color Levels in Rice Plants. *Computers and Electronics in Agriculture*. [Online]. Available at: https://elsevier.com/article/109234 [Accessed 25 July 2024].

Zhou, X., Huang, Y., Wu, K., Zhang, J. and Chen, D. (2023). A Novel Approach to Improve Leaf Chlorophyll Content Estimation with ResNet-18 and Feature Optimization Using Hyperspectral Measurements. *SSRN Preprint*. [Online]. Available at: https://ssrn.com/abstract=4740375 [Accessed 1 June 2024].

Anarado, C.J., Anarado, C.E.O., Ogbuokiri, B.O., Uche, C.O., and Ugwuanyi, J.O., 2021.Adoption of improved farm practices among arable crop farmers in Ivo Local GovernmentArea of Ebonyi State, Nigeria. IAPGOS – Journal of Agriculture and Rural Development [online] 44(1). Available at: https://ph.pollub.pl/index.php/iapgos/article/view/5626/4460 [Accessed 3 Mar. 2024].

Chen, Z., Chen, X., and Liu, M., 2021. Detecting Rice Diseases Using Deep Convolutional Neural Networks. Nature Precedings [preprint]. Available at: https://d197for5662m48.cloudfront.net/documents/publicationstatus/200712/preprint_p df/ cb4a0a0bcaae58401d55a96819d753f8.pdf [Accessed 4 Mar. 2024].

Wei, Q., Huang, W., Ye, H., Cen, H., Xu, Y., and Liu, J., 2022. Nondestructive determination of rice leaf nitrogen concentration based on spectral analysis. Computers andElectronics in Agriculture,[online]193.Available at:https://www.sciencedirect.com/science/article/pii/S1672630822000075 [Accessed 7 Mar. 2024].

Unknown author, 2022. Title of the document. [eprint] Yogyakarta Technological University. Available at: http://eprints.uty.ac.id/15036/ [Accessed 18 Mar. 2024].

Basso, B., Antle, J., and Lamonica, G.R., 2021. Digital Technologies for Agricultural andRural Development in the Global South. Agriculture, [online] 11(8), p.766. Available at: https://www.mdpi.com/2077-0472/11/8/766 [Accessed 30 Mar. 2024].

Hidayat, A.S., Horino, H., and Rzeznicka, I.I., 2021. Smartphone-Enabled Quantification of Potassium in Blood Plasma. Sensors, [online] 21(14), p.4751. Available at:https://www.mdpi.com/1424-8220/21/14/4751 [Accessed 14 Mar. 2024].