International Journal of Business, Arts and Scientific Study (IJBASS)
ISSN: 2709-0825 (Print), ISSN: 3078-2775 (Online), Volume-07, Issue-02, August, 2026
A multidisciplinary journal of Bangladesh Center for Research and Planning (BCRP)

DOI: https://doi.org/10.70818/ijbass.2026.v07i02.2621

Automated Leaf Color Chart (LCC)-Based Nitrogen Fertilizer Recommendation Using Paddy Leaf Image Color Classification

Rusanara Akhter[1]
Undergraduate Fellow, Dept. of CSE, Dhaka International University,
Email: so-wa.mitu@gmail.com

Md. Zahidul Islam[2]
Asst. Prof. Dept. Of CSE, Dhaka International University,
Email: zahidul.cse.bd@gmail.com

Keywords:
Paddy leaf color classification, Leaf Color Chart (LCC), nitrogen management, deep learning, Custom CNN, EfficientNetB7, ResNet50, VGG16, Lla MA 3.2-11B Vision

ABSTRACT: The demand for accurate and timely assessment of crop health is high in precision agriculture to support proper nutrient management. In this work, we present an intelligent deep learning-based solution that evaluates the colour of paddy rice leaves to provide real-time, context-sensitive recommendations regarding the addition of nitrogen fertilizers to grow rice sustainably. For optimum nitrogen management and sustainability in the production of rice crops, an accurate assessment of the color of paddy leaves is required. Herein, we present a deep learning-based framework for paddy leaf color classification with the capability of recommending an appropriate amount of nitrogen fertilizer as needed in real-time. In this research, the Paddy Net dataset containing 12,000 augmented rice leaf images will be used for training the neural network used for classifying the rice leaf's color based on the Leaf Color Chart (LCC) system, which consists of LCC levels from 2 to 5. Additionally, several deep learning architectures were trained and tested, specifically Custom CNNs, VGG16, ResNet50 and EfficientNetB7, to identify which architecture provided the best classification performance of rice leaf color. In terms of the experimental results, it can be concluded that the Custom CNN with a classification accuracy of 97.55% significantly outperforms the transfer learning models, VGG16 (78.89%) and EfficientNetB7 (78.67%). To also improve practicality, multimodal vision-language model LLa MA 3.2-11B Vision is integrated into the system, allowing it to provide real-time, context-based nitrogen fertilizer recommendations based on visual input images. Additionally, the proposed framework takes into account other agronomic factors such as the growth stage of the crop, type of soil, and fertilizer application history. This allows for more accurate and relevant recommendations. The proposed framework allows for precise and objective nitrogen management without the subjective use of visual assessment methods typically used by farmers. Overall, the proposed system for fertilizer application is significant for the development of precision agriculture because it increases the efficiency of the decision-making process, optimizes the use of fertilizer, and increases the practice of sustainable rice cultivation.

1. Introduction

On Chlorophyll production, photosynthetic rates, and overall yield from rice are all affected by the availability of nitrogen to rice plants. Consequently, efficient nitrogen management is an essential aspect of maximizing productivity and achieving resource sustainability to produce rice. Rice (Oryza sativa) is a major food source for over half the world's population. It constitutes at least 70% of average daily calorie intake in Bangladesh, India, China, and Indonesia (in most Asian countries). Rice contributes over 750 million metric tons of rice per year to the world; due to its importance for food security and agricultural income, it also plays an important role in rural development, employment, and economic development[1].

Determining how much nitrogen fertilizer to use, which is very important, can be very difficult for farmers. Farmers who use too much or not enough nitrogen can have problems with how well crops yield, how easily crops become infested with insects or diseases, and serious environmental problems such as soil degradation and water pollution caused by nutrient runoff[2]. There is a very good solution for this problem; the Leaf Color Chart (LCC) is a cheap, easy way to find out what kind of nitrogen fertilizer (if any) is needed based on the color of leaves on the plant[3].

LCC (Leaf Color Chart) is very dependent on how much an individual can visually interpret colors. This creates a lack of consistency and subjectivity in the application of this method in real-world situations. Changes in light (time of day or weather), people's ability to see (some people are color blind), and environmental changes all lead to inaccuracies when determining the color of the leaves of crops to determine nitrogen needs. This causes farmers to miscalculate how much nitrogen to apply, resulting in waste of fertilizer and money[4]. Also, because deficiencies in phosphorus and potassium can affect the color of leaves, determining if the color difference is due to nitrogen deficiency is difficult, causing farmers to have more trouble making decisions[5]. Additionally, because LCC is manual, it cannot be used in a large-scale manner by users who may not have much technical knowledge.

Recent discoveries in deep learning and computer vision show that these technologies could tremendously impact or provide solutions to problems faced in agriculture through automated and objective analyses of agricultural data. CNNs (convolutional neural networks), along with advanced CNN architectures like VGG16, ResNet 50, and Efficient Net, demonstrate exceptional performance when used for image classification tasks like identifying healthy vs. diseased plants or monitoring crops[6][7]. These technologies also provide exceptional feature extraction, resulting in the ability to accurately identify differences in the color and texture of leaves as a result of the change in environmental conditions.

A deep learning technology-based comprehensive framework is proposed to classify paddy leaves color and recommend nitrogen fertilizer. This framework consists of several models, including a custom CNN, EfficientNetB7, ResNet50, and VGG16, for classifying images of the leaves at various LCC levels (2–5). Furthermore, an additional multimodal vision-language model (LLaMA 3.2–11B Vision) will be used to generate fertilizer recommendations based on a vision analysis of the leaf images combined with agronomic factors, such as growth stage, soil type, and prior fertilizer applications, and will be presented in a human-interpretable manner[8].

This research's main goal is developing a scalable, accurate, and simple-to-use tool for estimating the color of a paddy field leaf from real-world photographs, while at the same time overcoming the shortcomings of traditional LCC (leaf color chart) techniques. The proposed approach combines deep learning classification with intelligent recommendation generation to improve decision-making efficiency and nitrogen usage and to encourage sustainable rice growing in the context of precision agriculture[9].

2. Literature Review

Automated systems for analyzing paddy leaves through precision agriculture have been evolving rapidly. In particular, automated systems that detect nitrogen deficiencies and classify diseases utilizing image processing and deep learning methods have been developed recently. An early effort in this area by Islam et al. created a CNN-based model that classified paddy leaf images based on Leaf Color Chart (LCC), achieving a classification accuracy of 94.22% when applied to a controlled dataset of 5,000 images. The study demonstrated the potential for automating the recommendation of nitrogen, but the need for controlled lighting conditions limited the ability to apply the results to real-world applications due to the variability in illumination and background encountered in practice.

Various traditional machine learning methods have been attempted, including Umam & Susanto, who achieved a 91.67% success rate in LCC classification by using a color space transformation and support vector machines (SVMs)[11]. However, traditional methods are highly reliant on hand-crafted features and will struggle with varying illumination levels; therefore, they are less robust than deep learning approaches. Likewise, classical image processing techniques such as segmentation, clustering, and feature extraction using GLCM and Gabor filters have been previously used for disease detection[15][16]; however, while they can work well in controlled environments, they generally cannot transfer to other environments due to being sensitive to noise, changes in the environment, and requiring careful tuning of parameters.

ResNet 9 (and similar lightweight deep learning models) have been developed to meet hardware limitations so they can be deployed on mobile devices[12][19]. These models are small and require fewer computing resources; however, this creates additional challenges in accuracy and difficulty extracting complex features, depending on environmental considerations. While the hybrid model, VGG-ICNN, has had high accuracy with controlled data sets, it has not yet been adequate when put to use in the real-world agricultural environment[20]. The Leaf Color Chart (LCC) has had extensive investigation as a low-cost method of nitrogen management[5][13][14]. Despite improving overall nitrogen use efficiency and increasing yield, it is an inherently subjective method of evaluation requiring human interpretation; as such, results from LCC can vary in accuracy due to variables including lighting, human perception, and environmental conditions. In addition, readings from the LCC may reflect other deficiencies in nutrients, which could further reduce the accuracy of the LCC as a source of total nutrient diagnostics.

All the previous studies published show good advancement in the automatic analysis of paddy leaves. Nevertheless, the following challenges still exist: dataset diversity, model generalization, interpretability, and real-time deployment of models. As these limitations suggest, there is a need for more robust, scalable, and context-dependent systems that can produce accurate results across a large variety of agricultural environments[17][18][21].

3. Methodology

A structured pipeline was used to develop and evaluate an automated paddy leaf color classification (LCC 2-5) and nitrogen recommendation system. Research began with standardizing field protocol to collect paddy leaf images, followed by preprocessing images through resizing, illumination normalization, color-space transformation (using CIE Lab), and data augmentation. Relevant color and texture features were extracted from the image and utilized to create multiple deep-learning models, including Custom CNN, EfficientNetB7, ResNet50, VGG16, and LLaMA 3.2-11B Vision, using stratified cross-validation methods. Model performance was evaluated using accuracy, precision, recall and F1-score, while comparative evaluation allowed identification of the most robust and efficient architecture. Validation of models on previously unseen data confirmed real-world applicability and reproducibility. The developed methodology provides a robust and reliable framework for automated nitrogen management in rice cultivation.

3.1 Data Collection

This study used the PaddyNet dataset, which is publicly available through the Mendeley Data repository, to enable nitrogen fertilizer recommendations using paddy leaf color analysis. The data were collected during the Aman growing season in paddy fields in Jashore, Bangladesh, with the cooperation of the Bangladesh Agricultural Research Institute (BARI) and local agricultural extension officers. To produce high-quality images that accurately represent field conditions, a custom-built mobile application was used to acquire images of the top leaves on paddy plants, using a Nokia 3 (8 MP camera) or Samsung S8 (12 MP camera) in a real agricultural environment. As the images were collected, they were annotated according to the Leaf Color Chart (LCC) with the assistance of BARI's principal scientific officers, who validated the annotations to ensure proper labeling and reliability. At the time of image collection, there were 560 good-quality images of the uppermost leaves of paddy plants, all of which had been standardized to a resolution of 500 × 500 pixels and were categorized by four LCC levels between 2–5, with the number of images in each category being 158 for LCC level 2; 162 for LCC level 3; 124 for LCC level 4; and 116 for LCC level 5.Augmentation of data was utilized to add different types of "noise" (e.g., rotations and flips) and create 12,000 new images to make the dataset larger, thus improving both model performance and generalization. All augmentations were added so that they created images that were similar to different orientations of leaves under different environmental conditions. All images were saved as JPEGs, so they can be used with all popular deep learning frameworks. With its structure and augmentations, this dataset will form the basis for training, validating, and evaluating the models proposed here. Table 1 shows an example of a subset of the images in this dataset, which were used in this study.

Table 1. Performance comparison of deep learning models Utilizing automated LCC to recommend nitrogen fertilizer for paddy with a color classification approach of paddy leaf images.

Model	Accuracy	Precision	Recall	F1-Score
Hybrid Deep Learning	96.5%	95.8%	96.0%	95.9%
Visual feature analysis + ML	78.88%	79.37%	78.88%	78.64%
Leaf shape, texture, color + ML	92.1%	90.5%	91.0%	90.8
KNN + K-Means	89.4%	87.0%	88.2%	87.6%
Deep Learning CNN	95.7%	94.8%	95.2%	95.0%
Lightweight Deep Learning	94.3%	93.7%	93.9%	93.8%
Custom CNN,VGG16, ResNet50, Efficient NetB7 (proposed)	97.55%	97.57%	97.55%	97.54%

A table (Table 1) summarizing the comparative performance analysis of different deep learning and machine learning models used on paddy leaf images to predict the recommended nitrogen amount via an automated Leaf Color Chart (LCC). Key evaluation metrics include accuracy, precision, recall and F1-score. Out of all the models tested for use with nitrogen determination, the custom hybrid model of CNN and VGG16 + ResNet50 + EfficientNetB7 produced the best performance results, as evidenced by the highest accuracy= 97.55%, precision= 97.57%, recall= 97.55%, F1- score= 97.54%, and therefore established itself as the superior model than that of the traditional approaches. Traditional models that produce visual features and utilized ML, together with the use of KNN + Kmeans produced lower overall performance with accuracies of between 78.88% to 89.4%. While deep learning (deep neural networks) together with lightweight deep learning neural networks have produced overall accuracy results of > 94%, thus further establishing the capability and efficacy of deep learning models for use in the automated management of nitrogen usage in rice production. These comparisons illustrate that hybrid/ and or custom-built deep learning models have the greatest potential for improving classification reliability, thus providing agriculture professionals with a powerful tool for precision or site-specific agriculture.

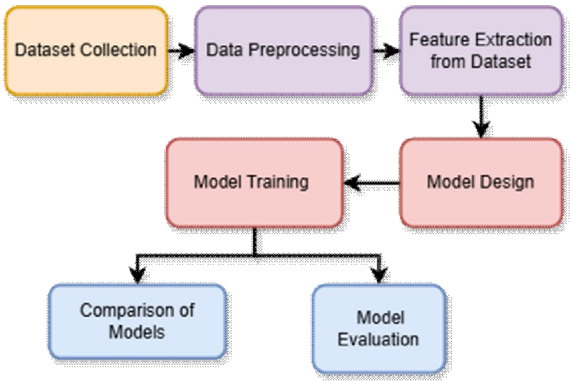

Fig. 1. End-to-End Methodological Pipeline for Utilizing automated LCC to recommend nitrogen fertilizer for paddy with a color classification approach of paddy leaf images.

This research utilizes an organized methodology to create a computer-assisted system for classifying color of paddy (rice) leaves and providing recommendations on the amount of nitrogen fertilizer to use. The research starts with data collection, where images of paddy (rice) leaves taken under multiple environmental conditions at differing levels of color shade are collected according to established protocols for taking field photographs. After data collection of the images, the images undergo pre-processing before analysis to ensure quality and consistency in appearance. Pre-processing includes resizing, normalizing illumination and transforming image data from one color space to another (i.e., RGB to HSV or LAB) to reduce variations in appearance due to lighting and background conditions. The next step in the process is to extract features from images that will provide a basis for automated classification by extracting the relevant color and texture features. The following step is designing and evaluating multiple custom-designed CNN deep learning models and pre trained models (e.g., EfficientNet B7, ResNet50, VGG16, and LLaMA 3.2-11B Vision). The models are trained with stratified cross-validation to ensure optimal performance and prevent over fitting. The models are also evaluated using standard performance metrics of accuracy, precision, recall, and F1-score by using both hold-out validations against unseen data and through pair-wise comparisons of models. The results of this comparative evaluation allow for the identification of the best-performing model based on predictive ability, computational efficiency, and scalability. As a whole, this approach to study design and implementation creates a valid and replicable method of developing a visual-based system that will classify paddy leaf color accurately and provide information that can lead to improved nitrogen fertilizer-use decisions in rice farming. The methodology diagram provides a visual way to display the workflow of the project, starting with dataset collection and ending with final model assessment and the sequential order in which each part of this research process occurred.

3.2 Data Preprocessing

When collecting paddy leaf images, a set of sequential preprocessing methods are done to adjust and improve the data used to train the models. This is done to ensure that the classification of the data will not be negatively impacted by differences in quality of images such as; lighting, background, and size of the image.

Before proceeding with training the model, the paddy leaf image dataset must go through a preprocessing phase to standardize, refine, and expand its diversity. The first step in data preprocessing is to resize all images (to fixed dimensions such as 224×224 pixels and 256×256 pixels) and normalize the pixel intensities (scale the pixel intensities to be within [0, 1]), thereby reducing the amount of variability due to lighting conditions and increasing the stability of the model. To increase the variability in the dataset, which will help to improve generalization, data augmentation techniques (randomly rotating, flipping, and making small geometric changes to the leaves) will be applied to images during training and provide simulated examples of how leaves may look given the variances associated with leaf orientation and the conditions in which they were acquired. This preprocessed and augmented dataset will then be efficiently passed through tf.data pipelines to allow real-time batching, caching and pre fetching during the model training process. The augmentations will only be applied to the training dataset to prevent information leakage and ensure a fair evaluation of the model during the evaluation phase. The combination of these updates to the original processed image set will improve feature learning while limiting the amount of overdosing and develop a strong model performance when conducting paddy leaf classification.

3.3 Model Architecture for Paddy Leaf Analysis

Several AI models are described in this part of the research that will help classify different types of paddy leaves. Depending on how they were built, each model has its own way of finding the most helpful color and texture patterns; this is very important for estimating how much nitrogen is in the plant and detecting diseases in plants. The types of models used in this research ranged from custom-built lightweight CNNs to state-of-the-art deep learning, multimodal vision, and language-based models.

Custom CNN. Custom Convolutional Neural Networks (CNNs) are tailored to the specific needs of a dataset and its intended analytical task so that domain-based features can be extracted efficiently. For example, in paddy leaf analysis, custom CNN architectures usually consist of multiple convolutional layers, subsequently followed by max pooling layers after each convolution layer, and a ReLU activation function that produces a non-linear response at each step of the processing. Finally, a Softmax output layer computes an aggregated output for multi-class classification of relevant image objects, representing the distinct classes of analyzed rice plant diseases. As an example, Pal (2022) used a custom CNN with three convolution layers plus two max pooling layers to classify images of rice plant diseases, yielding an accuracy rate of approximately 98.6% (for field-based images). This simple design reduces both model complexity and over fitting, allowing the CNN to focus on important visual characteristics such as leaf color and texture differences; in addition, by augmenting the dataset, the resulting CNN is robust to variations caused by the imaging conditions (e.g., lighting conditions, leaf orientation, and size) typically found in real-world settings. Compared to more common architectures, custom CNNs have a number of advantages: less resources needed resulting from lower total parameter counts, easier interpretability, rapid prototyping or model deployment capabilities (particularly in mobile environments), and more effective for resource-limited use. However, custom CNNs' design simplicity may decrease their performance compared to larger, more complex, pre-trained models due to their use of large datasets and generalization of feature representation.

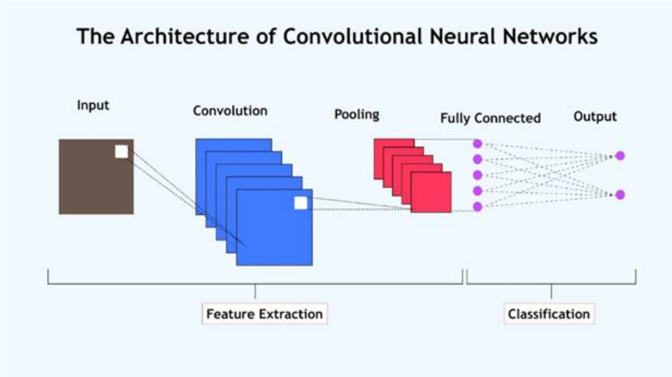

Fig. 2. The Architecture of CNN

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 254, 254, 32)	896
max_pooling2d (MaxPooling2D)	(None, 127, 127, 32)	0
conv2d_1 (Conv2D)	(None, 125, 125, 64)	18,496
max_pooling2d_1 (MaxPooling2D)	(None, 62, 62, 64)	0
conv2d_2 (Conv2D)	(None, 60, 60, 128)	73,856
max_pooling2d_2 (MaxPooling2D)	(None, 30, 30, 128)	0
flatten (Flatten)	(None, 115200)	0
dropout (Dropout)	(None, 115200)	0
dense (Dense)	(None, 128)	14,745,728
dropout_1 (Dropout)	(None, 128)	0
dense_1 (Dense)	(None, 4)	516

Fig. 3. Custom CNN Model Summary

A custom CNN was created to perform multi-class image classification (such as classifying the color of paddy leaves and identifying leaf diseases). The CNN is made up of three convolutional layers of 32, 64, and 128 convolutional filters (3 x 3 kernel) to create progressively more complex spatial representations. After a convolution, a max pool layer is added for each convolution and progressively reduces (down-samples) each subsequent feature map, keeping only the most salient features. Once the output feature maps from the convolutions and pooling layers are completed, they are all flattened into one-dimensional vectors and pass through two dropout layers (each with a 20% dropout) to reduce the likelihood of over fitting and to improve generalization. After that, there is one fully connected (dense) hidden layer with 128 neurons and a ReLU activation function used to further process the features extracted by the convolutional and pooling layers so that the CNN can learn non-linear combinations of patterns present in the data. The last layer of the CNN is a dense output layer with a softmax activation function producing the probabilities of each target class (n_classes), allowing for accurate multi-class classification of input image data. This architecture has effectively combined hierarchical feature extraction with regularization techniques, providing appropriate trade-offs between the model capacity, robustness and computational efficiency. Thus, this architecture will work well with real-world images that often display large variations in lighting, positional orientation and leaf size, providing reliable classification results.

Efficient Net B7. The architecture of Efficient Net B7 is a highly optimized convolutional neural network architecture that offers a balanced trade-off between the depth, width, and input resolution of the network by compound scaling. An optimised architecture will allow the model to efficiently use computational resources and parameters while simultaneously providing hierarchical feature extraction. Efficient Net B7 has achieved state-of-the-art performance on many image classification tasks consistently, which makes it a strong candidate for high-accuracy applications. Efficient Net B7 has also shown exceptional capability to capture very fine-grained details in high-resolution agricultural images, which is critical for identifying subtle differences in color that represent nitrogen status or disease on paddy leaves. By initially using transfer learning from weights pre-trained on Image Net, Efficient Net B7 can be fine-tuned using paddy leaf datasets and converge on a solution quickly and accurately, even when there is little domain-specific data available for training. While Efficient Net B7 is extremely effective, due to its increased computational requirements compared to lighter-weight models, deploying it in edge devices or environments that have limited computing resources may not be possible. However, Efficient Net B7's fast convergence and ability to perform at a high level of accuracy make it an excellent choice for applications that require accuracy, such as crop monitoring and determining the nitrogen level of crops.

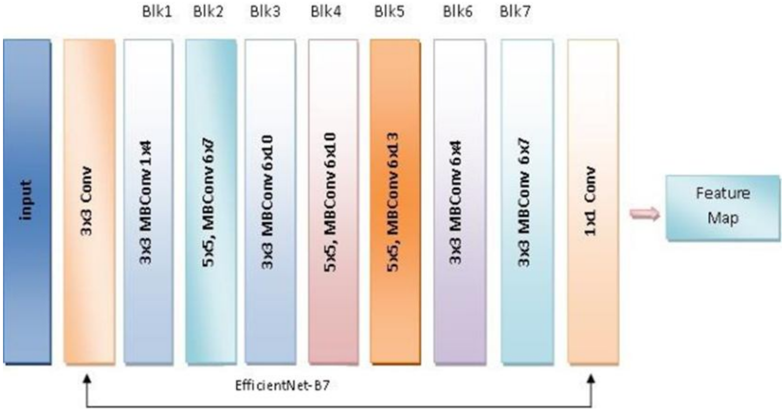

Fig. 4. The Architecture of EfficientNet-B7

ResNet50.ResNet50 has been widely adopted in classifying paddy leaf disease and nutrient deficiencies via agricultural image analysis because it can capture detailed textural and color differences that distinguish between closely related levels of the Leaf Color Chart or subtle symptoms of different leaf diseases, as the residual blocks of the network enable efficient training even on moderately sized datasets through application of transfer learning from pre-trained weights. While ResNet50 requires more computational power than smaller custom CNNs, it demonstrates a great balance of depth, accuracy and robustness of performance and has been proven effective enough to be used frequently for practical operations in smart farming where reliable, detailed agricultural image analysis will be performed.

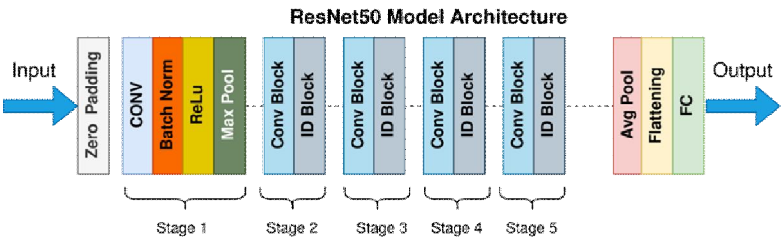

Fig. 5. The Architecture of ResNet50

VGG-16.While VGG16 is an early and very successful CNN architecture with a simple and uniform design with only 16 weight layers, it has proven to be an effective way for CNNs to learn hierarchical feature representations and has a very simple-to-understand and implement concept. This architecture primarily uses small

(3×3) convolutional filters stacked together in multiple layers, representing the first half of the network, where the front half consists of convolutional layers, and the back half consists of fully connected layers. Due to its simple architecture, VGG16 provides an easy way to learn hierarchical feature representations in a manner that is not only conceptually clear but also enables CNNs to learn better at their respective types of input data. While VGG16 was one of the earliest CNN architectures developed, it continues to be one of the most widely used CNN architectures for agriculture image classification due to its ability to extract features effectively. VGG16 has demonstrated competitive accuracy for the identification of paddy leaf disease and classification of leaf color. According to Pal (2022), VGG16 provides excellent feature extraction capabilities and is very easy to fine-tune for specific domains because of the same design features. However, the VGG16 CNN architecture has limitations as well. VGG16 has a very large number of parameters, making it difficult and computationally expensive to train. Additionally, the architecture is prone to over fitting and can require significant amounts of regularization and data augmentation techniques to prevent over fitting. Therefore, it is important to keep this in mind when implementing VGG16. Training time with the VGG16 architecture can be longer than some of the newer or more compact architectures.

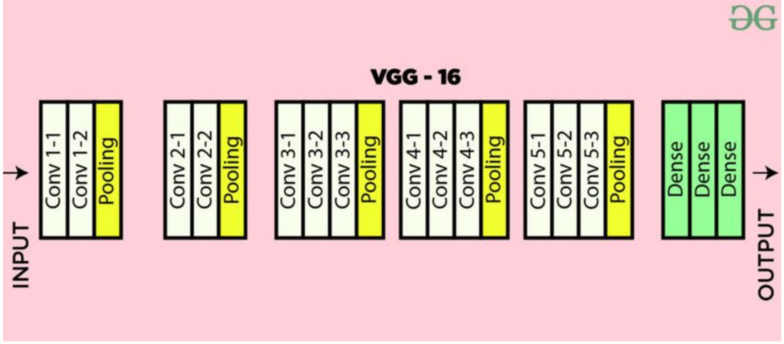

Fig. 6. The Architecture of VGG-16

LLaMA 11B Vision. The LLaMA 11B Vision is an advanced large language model with multimodal capabilities and new vision capabilities that allow it to analyze high-resolution images and text together. Unlike CNNs which typically classify data, LLaMA Vision's transformer architecture utilizes attention-based mechanisms for holistic visual reasoning as well as generating human-readable explanations and predictions. The 11B model contains a combined global understanding of context with detailed local characteristics, making it particularly capable at completing complicated agriculture diagnosis tasks. For instance, its ability to produce interpretable solutions (e.g., actionable information for farmers on how to manage nitrogen or minimize disease) when interpretability is most important is a substantial benefit. With its ability to do visual reasoning, followed by text-based reasoning, LLaMA Vision transcends standard classification and generates explainable recommendations that can improve the quality of decision-making in precision agriculture. However, LLaMA 11B Vision's extreme size and computational demands require high-performance hardware (or cloud infrastructure), making it difficult to deploy on-device. Nevertheless, the model is a major innovation in the agriculture multimodal artificial intelligence space, as it combines state-of-the-art image understanding with contextual reasoning for generating actionable, interpretable information insights.

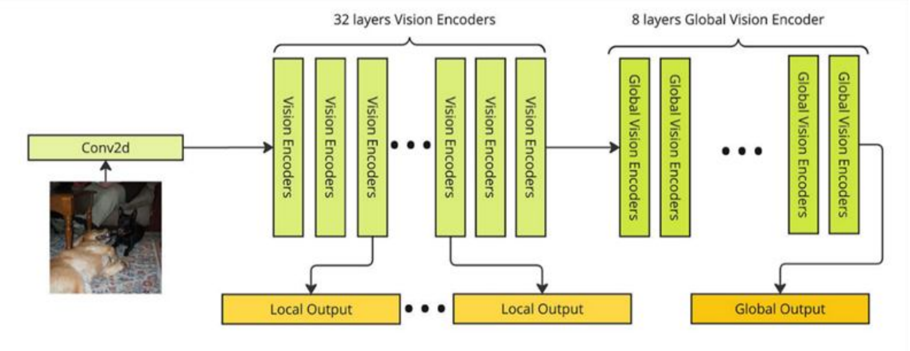

Fig. 7. The Architecture of LLaMA 11B Vision

4. Experiments and Results

A set of tests was performed to assess the performance of the deep learning models already developed based on the Paddy Net dataset to determine their ability to properly classify paddy leaves by color level (LCC 2-5) for providing nitrogen recommendations. The following standard metrics were used to evaluate performance: accuracy, precision, recall, F1, and confusion matrices, which allow for a well-rounded cross-comparison of model robustness and generalization for use in a real-world environment.

4.1 Experimental-Setup

Carefully designed to evaluate the performance of five different deep learning models (Custom CNN, Efficient Net B7, ResNet50, VGG16 & LLaMA 3.2-11B Vision), the experimental setup allows for sufficient understanding of these models' abilities to classify paddy leaf color levels (LCC 2 – 5) to aid in the recommendation of nitrogen fertilization. Experiments utilized the PaddyNet dataset, which was made up of 12,000 expertly diagnosed images of paddy leaves and was normalized to a size of 500x500 pixels and then split into an 80% training set and a 20% test set. All models were trained under the same preprocessing conditions; specifically, all images were resized to a size of 256x256 pixels, normalized (where pixel values fall within the range of 0-1), and had undergone random rotation (with a random rotation lasting +/- 20 degrees) to increase generalisability of the models trained with each model using a batch size of 64 while being trained by the Adam optimizer and with a learning rate of 0.0001 for a total of 200 epochs while also employing early stopping and model check pointing to minimize potential for over fitting. The base layers of transfer learning models (EfficientNet B7, ResNet50, VGG16) were frozen during training of the top classification layers; then fine-tuning was done to adapt the networks to the characteristics of paddy leaves. The Custom CNN was constructed from scratch, and had multiple convolutional and pooling layers with a ReLU activation function, dropout regularization, and softmax outputs. A multimodal inference model was used with the LLaMA 3.2-11B Vision to generate interpretable predictions based on the use of leaf images and prompts. High-performance hardware (NVIDIA RTX 3090 or A100 GPUs) was used for the experiments, with optimized Tensor Flow 2.x pipelines using tf.data prefetching and parallel processing to reduce I/O bottlenecks. Evaluation of the models included an extensive array of metrics: accuracy, precision, recall, F1-score, and confusion matrices, allowing for a thorough assessment of model performance on the subtle differences present between the colors on the leaves tested.

4.2 Performance Evaluation Metrics

Performance Evaluation Metrics are discussed in details in below. The performance of my research work is shown in the table.2

Table 2. Algorithm performance.

Model	**Accuracy**	**F1 Score**	**Recall**	**Precision**
Custom CNN	97.55%	97.54%	97.55%	97.57%
VGG16	78.88%	78.64%	78.88%	79.37%
ResNet50	83.91%	83.88%	83.91%	84.14%
EfficientNetB7	78.67%	78.17%	78.67%	81.45%

An evaluation was conducted to measure the performance of the different evaluation techniques by assessing both deep-learning model accuracy, precision, recall and F1-score. The number of correct predictions made by each model is referred to as the model's accuracy. Precision & recall provide insight as to how effectively a model has minimized the number of classifications that were incorrectly classified as positive (precision) and the number of classifications that were correctly classified as positive (recall). When doing multi-classification tasks, such as paddy leaf color level classification, the importance of using the F1-Score is to determine how balanced the precision and recall scores are among the models. The Custom CNN model had the highest scores among all the models, with 97.55% Accuracy, 97.57% Precision, 97.55% Recall and a 97.54% F1-Score. The consistently high scores suggest that this model has excellent capabilities for classifying images in the dataset, has also learned features effectively and has generalized well on the Paddy Net dataset. VGG16 and Efficient Net B7 both have low scores in comparison with the Custom CNN, with the models scoring 78.89% & 78.67% Accuracy, respectively. Although the Precision scores for both models are still reasonably high, their Recall and F1-Scores are lower than they should be, causing increased misclassification of images, especially between similar classes. ResNet50 performed moderately better with a score of 83.91% Accuracy, but was still not as good as the Custom CNN. However, ResNet50 had balanced Precision & Recall scores. The overall results of this evaluation demonstrate that well-designed custom models can outperform pre-trained models with respect to image classification tasks in a domain such as agriculture.

4.3 Comparative Analysis of Model Accuracy

The confusion matrix analysis for implemented models is shown in Figures 8, 9, 10, and 11.

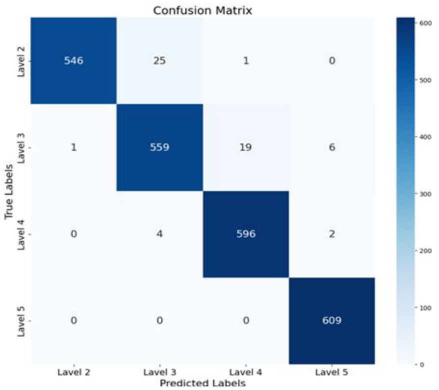

Fig. 8. Confusion matrix of CNN

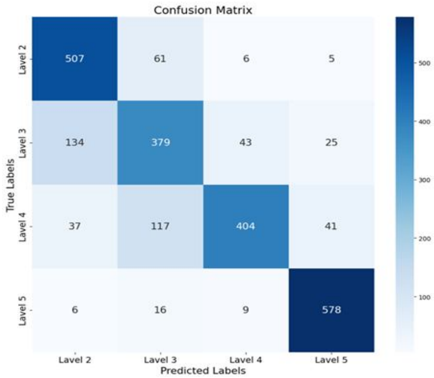

Fig. 9. Confusion matrix of VGG16

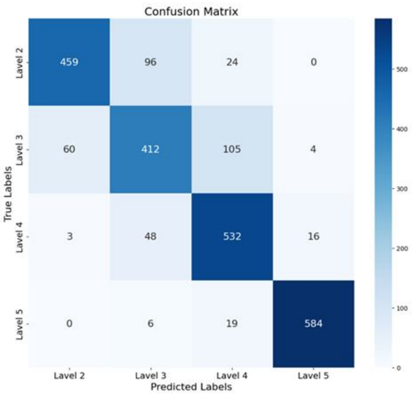

Fig. 10. Confusion matrix of Res Net 50

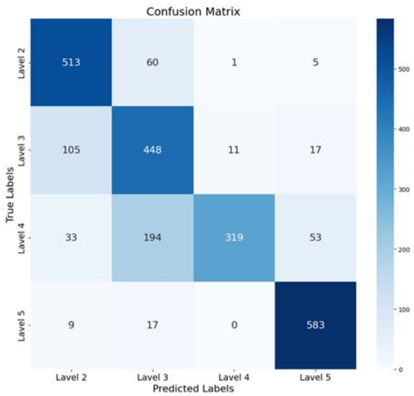

Fig. 11. Confusion matrix of Efficient Net B7

The analysis of confusion matrices reveals significant variation between models in their classification performance, with Custom CNN showing an exceptional balance of accuracy, precision, recall, and F1-score, all approximately at 97.55% as well as very few misclassifications indicating reliable identification of all Misclassifications are few in the Custom CNN model, meaning that the model is capable of successfully identifying LCC levels (2-5) with very little chance of producing either false-positive or false-negative prediction. The Custom CNN model's high precision will mean that nitrogen-deficient leaves are seldom overestimated, helping to avoid over-application of fertilizer. High recall will result in almost all nitrogen-deficient leaves being correctly identified, reducing risk of under-application of fertilizer. These features support the conclusion that Custom CNN has learned robust and discriminating feature representations relative to fine-grained color classification. In contrast to the Custom CNN model, VGG16, Efficient Net B7 and ResNet50 all display much lower overall performance. Their corresponding confusion matrices show the presence of relatively higher off-diagonal values representing a higher number of misclassifications occurring with neighboring LCC levels. The corresponding lower recall (78.67-83.91%) produced by these models suggests that nitrogen-deficient leaves are more likely to be misclassified as healthy plants than by the Custom CNN model, resulting in a much greater risk of under-applying nitrogen fertilizers. Efficient Net B7 achieves relatively high precision (81.46%) but the cost of reduced recall, suggesting that the model is conservative in avoiding false-positive samples and thus leads to more false-negative samples. The overall moderate precisions and recalls of VGG16 and ResNet50 suggest that they have weaker overall capabilities in classifying colored samples.

4.4 ROC Curve Analysis

A Receiver Operating Characteristic (ROC) curve assesses the effectiveness of a classification model, including those that classify binary and multi-class items like the paddy leaf color level (LCC). An ROC curve plots the True Positive Rate (TPR = Sensitivity or Recall) on the Y-axis and the False Positive Rate (1 - Specificity) on the X-axis. By plotting these two rates at varying classification thresholds, you can see how well your classification model distinguishes between classes based on given decision boundaries. If the ROC curve of your model is close to the top-left corner, it means that it correctly identifies many of the positive cases and minimizes the false-positive cases. In contrast, a curve that is vertically or horizontally centered in relation to both axes means that the model performs at least as well as random chance. The Area under the Curve (AUC) expresses the ROC curve on an absolute scale. An AUC close to 1 represents an excellent model with the predictive capability to correctly assign a classification to an item across all classification thresholds. An AUC of 0.50 represents a model that does not provide valid discrimination between classes. For multiple classes, individual ROC curves are generated for each class using a one-versus-all methodology and then summarized to provide an overall Macro Average of AUC for the entire model's performance.

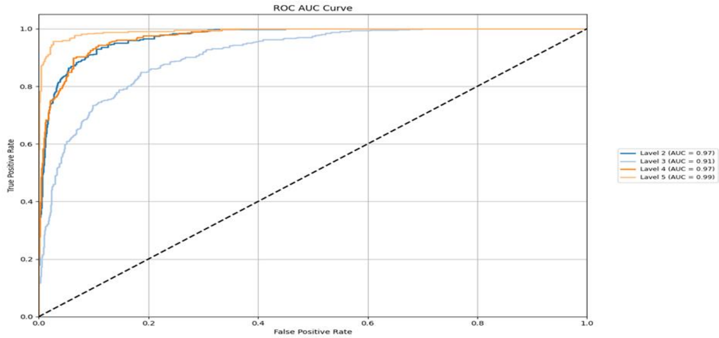

Fig. 12. ROC Curve of Custom CNN

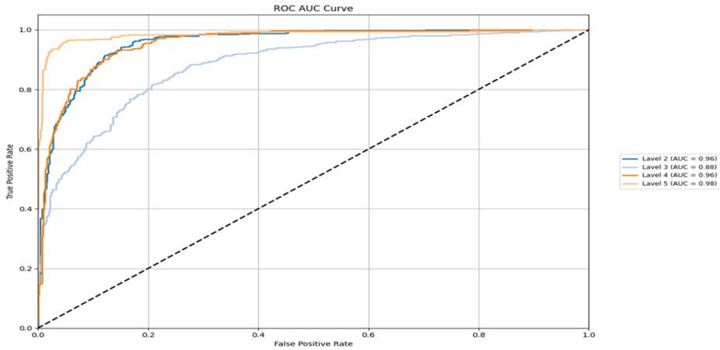

Fig. 13. ROC Curve of VGG16

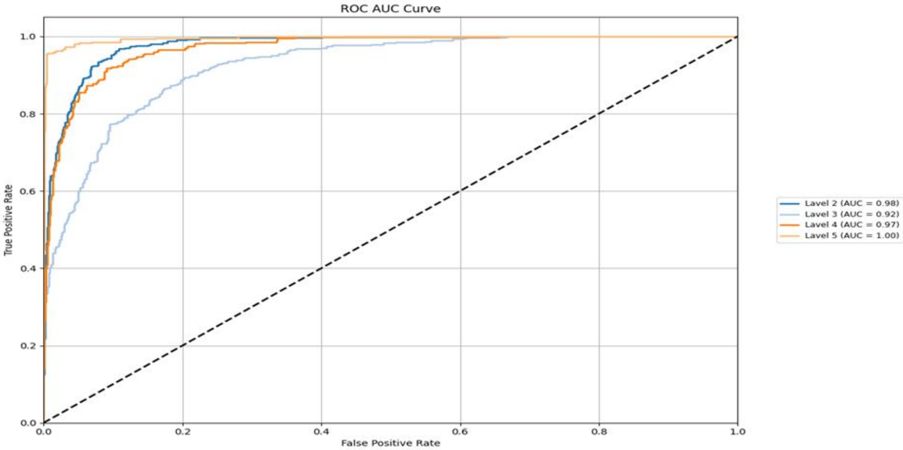

Fig. 14.ROC Curve of ResNet50

4.5 ROC and Training/Validation Performance

The Receiver Operating Characteristic (ROC) Curve is a major evaluation metric for classification models used for binary and multi-class tasks, such as paddy leaf color classification. It is created by plotting the true positive rate (also known as sensitivity or recall) against the (1 – Specificity) or false positive rate across a range of threshold levels. Thus, the ROC Curve allows an analyst to see how well a classification model can discriminate between classes. If the ROC curve is close to the top left corner of the graph, it means that the model is performing well at classifying data with a high true positive rate and few false positives. The Area Under the Curve (AUC) provides a numerical representation of the ROC Curve's performance. Specifically, when the AUC value is close to 1, it means that the predictive power of the model should be very strong, while an AUC value of 0.5 means that the predicted classifier is effectively guesswork. For multi-classification tasks, it is possible to derive a ROC curve for each of the classes by using a one-vs-rest approach, and then take the macro-average of the AUC values from the individual ROC curves to summarize the performance of the classification model. In addition to evaluating classification models using ROC analysis, training and validation accuracy, and loss curves provide additional evidence of how the model is learning the underlying relationships between the data. The accuracy curves reveal how quickly the model converges, and whether it generalizes well to new/unseen data, while the loss curves provide information about the rate of decline in the number of errors found throughout the training process. In an ideal scenario, both the training and validation curves converge at the same time without significant divergence, indicating that there is little to no over fitting or under fitting occurring.

Observing these trends helps verify that the model has learned stable and discriminative features from the paddy leaf images, and that it maintains reliable performance across both seen and unseen samples.

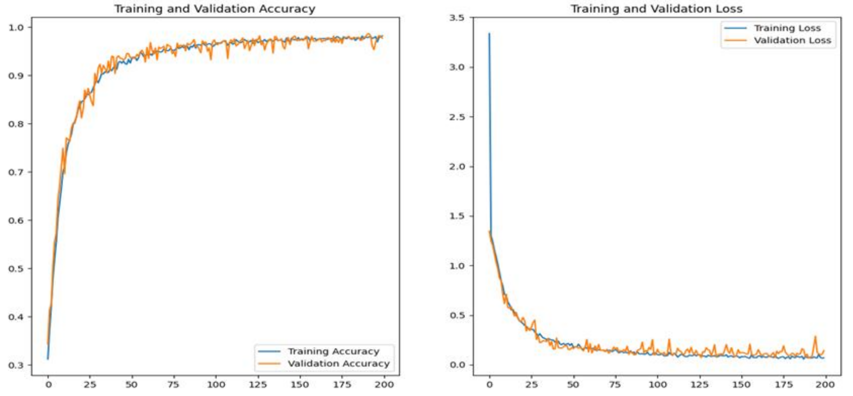

Fig. 15. Training & Validation Accuracy and Loss (Custom CNN)

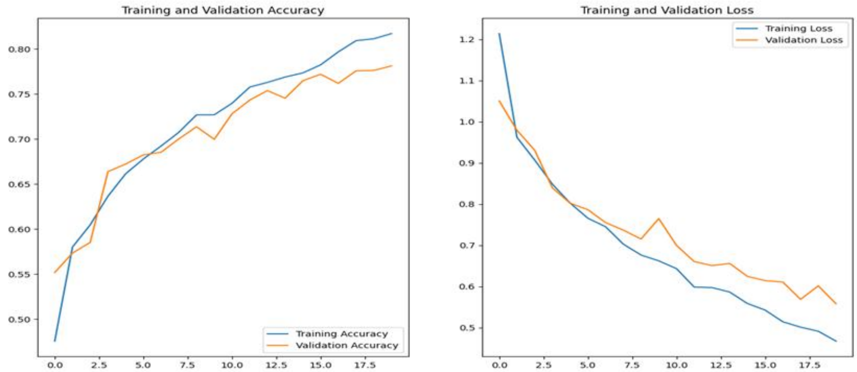

Fig. 16. Training & Validation Accuracy and Loss (VGG16)

Fig. 17. Training & Validation Accuracy and Loss (EfficientNetB7)

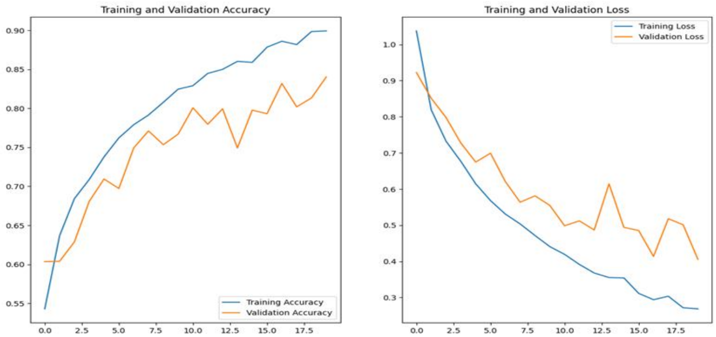

Fig. 18. Training & Validation Accuracy and Loss (Resnet50)

4.6 Real-time Nitrogen Recommendation Assessment Using Llama (11B)

The integration of LLaMA 3.2-11B Vision enables real-time, intelligent assessment of paddy leaf color levels (LCC 2–5) and provides actionable nitrogen fertilizer recommendations. By analyzing high-resolution leaf images and considering contextual field data such as growth stage, soil type, and prior fertilizer use—the model accurately identifies nitrogen status, from severe deficiency to optimal levels. Leveraging its multimodal vision-language capabilities, LLaMA generates human-readable advisory reports detailing precise urea dosages, split application timing, and method-specific guidance tailored for smallholder farmers. The system dynamically adjusts recommendations based on crop phonology; for example, LCC Level 2 at the tailoring stage triggers urgent split-dose urea application, while LCC Level 4 suggests maintenance dosing. By combining visual diagnosis with agronomic expertise, LLaMA 3.2 transforms Smartphone-based leaf imaging into a comprehensive digital advisory tool, reducing subjectivity, enhancing nitrogen use efficiency, preventing under- and over-fertilization, and supporting sustainable rice production with improved yields.

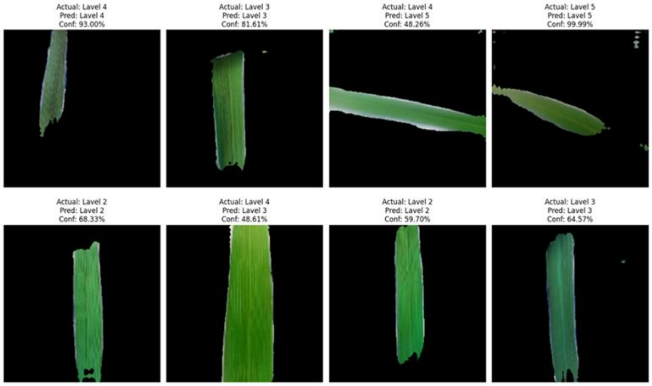

Fig. 19. Real-time inference

5. Discussion and Conclusion

The study concludes that deep learning based automated classification of paddy leaf color (level 2–5) is an effective tool for nitrogen management. The Custom CNN model clearly outperformed the models that utilized transfer learning technologies (VGG16, ResNet50, and Efficient Net B7) in terms of accurate performance (97.55%) and almost all other performance characteristics including precision, recall, and F1-score. The results indicate that purpose-built architectures are able to better represent the subtle, variable characteristics of leaves (e.g., degree of greenness) associated with nitrogen deficiency than can a general-purpose, pre-trained architecture. The proposed system has practical implications for sustainable rice agriculture through the ability to provide real-time, unbiased assessments of nitrogen need using Smartphone-based leaf images. This allows farmers to apply fertilizers optimally, improve yield, reduce costs associated with fertilizer, and limit environmental harm created by nitrogen leaching and emissions of greenhouse gases. The light weight of the Custom CNN also lends itself to deployment on the device, making it affordable and accessible to smallholder farmers working in resource poor environments. Some limitations of the research include: the study was limited to a geographically specific dataset (Jashore, Bangladesh), and there is no consideration of color-based classification due to disease or water stress and/or other nutrient deficiencies. Future studies should not only expand the study's datasets to cover more agro-ecological regions, but also consider additional data collection methods (e.g., weather, soil, and multi-spectral imaging) to improve robustness and diagnostic capability for each model. The research has shown that personalized models of deep learning can consistently provide the capability to automate leaf color assessments and therefore serve as scalable, cost-efficient, and environmentally sustainable tools for managing nitrogen in precision agriculture when applied to growing rice.

6. References

1. R. Hu, D. Xiao, Z. Ding, Y. Cao, J. Hou, and X. Wang, “Effects of local nitrogen supply and nitrogen fertilizer variety coupling on rice nitrogen transport and soil nitrogen balance in paddy fields,” Frontiers in Sustainable Food Systems, vol. 7, Aug. 2023, doi: https://doi.org/10.3389/fsufs.2023.1252324.
2. Yoseftabar, Y. Researchers, E. Club, and S. Branch, “Role of biological nitrogen fixation in rice.” Accessed: Oct. 06, 2025. [Online]. Available: https://www.woarjournals.org/admin/vol_issue1/upload%20Image/IJGAES011104.pdf

3. X. Wang, H. Zhu, S. Hou, H. Cui, and B. Yan, "Environmental impacts of fertilization during rice production in saline-alkali paddy fields based on life cycle assessment," Journal of Cleaner Production, vol. 467, p. 142947, Jun. 2024, doi: https://doi.org/10.1016/j.jclepro.2024.142947.
4. "Smart Fertilizer Recommendation System through Leaf Color Chart (LCC) Automation," Jashore University of Science & Technology, 2023. https://just.edu.bd/research/projects/62ada463d5313009ceae039d (accessed Oct. 06, 2025).
5. K. Gokulnandan, R. Agronomy, N. Das, and K. Rao, "Leaf color chart (LCC) a reliable tool for nitrogen management in (Dry seeded rice and transplanted rice): A review," ~ 3974 ~ The Pharma Innovation Journal, vol. 12, no. 6, pp. 3974–3979, 2023, Available: https://www.thepharmajournal.com/ archives/2023/ vol12issue6/PartAT/12-6-215-860.pdf
6. G. Saleem, M. Akhtar, N. Ahmed, and W. S. Qureshi, "Automated analysis of visual leaf shape features for plant classification," Computers and Electronics in Agriculture, vol. 157, pp. 270–280, Feb. 2019, doi: https://doi.org/10.1016/ j.compag.2018.12.038.
7. C. Subbarayudu and M. Kubendiran, "An automated hybrid deep learning framework for paddy leaf disease identification and classification," Scientific Reports, vol. 15, no. 1, Jul. 2025, doi: https://doi.org/10.1038/s41598-025-08071-6.
8. M. Keivani, J. Mazloum, E. Sedaghatfar, and M. Tavakoli, "Automated Analysis of Leaf Shape, Texture, and Color Features for Plant Classification," Traitement du Signal, vol. 37, no. 1, pp. 17–28, Feb. 2020, doi: https://doi.org/10.18280 /ts.370103.
9. T. Islam, R. Uddin, Y. Arefin, M. Shafiuzzaman, Md. Alam, and Syed, "Nitrogen Fertilizer Recommendation for Paddies through Automating the Leaf Color Chart (LCC)," International Journal of Advanced Computer Science and Applications, vol. 11, no. 8, 2020, doi: https://doi.org/10.14569/ ijacsa.2020.0110891.
10. K. S. Rahman, Md. R. I. Rakib, M. M. Salehin, Md. R. Ali, and A. Rahman, "Assessment of paddy leaves disease severity level using image processing technique," Smart Agricultural Technology, vol. 7, p. 100410, Mar. 2024, doi: https://doi.org/10.1016/j.atech.2024.100410.
11. K. Umam and E. H. Susanto, "Classification Of Rice Leaf Color Into Leaf Color Chart Using LAB Color Space," CCIT Journal, vol. 13, no. 2, pp. 168–174, Aug. 2020, doi: https://doi.org/10.33050/ccit.v13i2.1008.

12. M. Ajith et al., “Issue 4 www,” IJRAR1DUP010 International Journal of Research and Analytical Reviews, vol. 11, p. 41, 2024, Accessed: Oct. 06, 2025. [Online]. Available: https://www.ijrar.org/papers/IJRAR1DUP010.pdf
13. M. M. Alam et al., “Leaf Color Chart for Managing Nitrogen Fertilizer in Lowland Rice in Bangladesh,” Agronomy Journal, vol. 97, no. 3, pp. 949–959, May 2005, doi: https://doi.org/10.2134/agronj2004.0206.
14. Dinesh Timilsina and Santosh Marahatta, “Evaluating Leaf Color Chart-Based Nitrogen Management and Tillage Methods for Improved Maize Yield, Nitrogen Use Efficiency and Economic Returns in Chitwan, Nepal,” Turkish Journal of Agriculture - Food Science and Technology, vol. 13, no. 8, pp. 2048–2055, Aug. 2025, doi: https://doi.org/10.24925/turjaf.v13i8.2048-2055.7625.
15. M. R. Larijani, E. A. Asli□Ardeh, E. Kozegar, and R. Loni, “Evaluation of image processing technique in identifying rice blast disease in field conditions based on KNN algorithm improvement by K□means,” Food Science & Nutrition, vol. 7, no. 12, pp. 3922–3930, Nov. 2019, doi: https://doi.org/10.1002/fsn3.1251.
16. R. Jayasooriya and S. Arachchi, “Issue 1 IJRARJFM1392,” International Journal of Research and Analytical Reviews, vol. 8, 2021, Accessed: Oct. 06, 2025. [Online]. Available: https://www.ijrar.org/papers/IJRARJFM1392.pdf
17. K. S. Rahman, Md. R. I. Rakib, M. M. Salehin, Md. R. Ali, and A. Rahman, “Assessment of paddy leaves disease severity level using image processing technique,” Smart Agricultural Technology, vol. 7, p. 100410, Mar. 2024, doi: https://doi.org/10.1016/j.atech.2024.100410.
18. A. Ahmad, D. Saraswat, V. Aggarwal, A. Etienne, and B. Hancock, “Performance of deep learning models for classifying and detecting common weeds in corn and soybean production systems,” Computers and Electronics in Agriculture, vol. 184, p. 106081, May 2021, doi: https://doi.org/10.1016/j.compag.2021.106081.
19. M. Ajith et al., “Issue 4 www,” IJRAR1DUP010 International Journal of Research and Analytical Reviews, vol. 11, p. 41, 2024, Available: https://www.ijrar.org/papers/IJRAR1DUP010.pdf
20. “AI in Agriculture: A Survey of Deep Learning Techniques for Crops, Fisheries and Livestock,” Arxiv.org, 2021. https://arxiv.org/html/2507.22101v1
21. W. Shafik, A. Tufail, L. C. De Silva, and R. A. A. H. M. Apong, “A lightweight deep learning model for multi-plant biotic stress classification and detection for sustainable agriculture,” Scientific Reports, vol. 15, no. 1, Apr. 2025, doi: https://doi.org/10.1038/s41598-025-90487-1.