CHAPTER 1

INTRODUCTION

Rice is one of the most important staple crops globally, feeding a significant portion of the world's population. However, rice production is often threatened by various diseases that affect the leaves, leading to reduced yield and economic losses. Early detection of these diseases is critical for timely intervention and efficient disease management. Traditional methods of identifying rice leaf diseases involve manual inspection by experts, which is time-consuming, costly, and often prone to errors due to subjective analysis. In recent years, advancements in artificial intelligence (AI) and computer vision have opened new avenues for automating disease detection using digital images, leading to more accurate and rapid diagnosis.

Deep Convolutional Neural Networks (DCNNs) have emerged as a leading technique in image-based disease detection due to their ability to automatically learn and extract features from raw images without the need for manual feature engineering. DCNNs are a type of deep learning model that excel at processing image data by learning hierarchical features, from basic edges and textures to more complex patterns, in a multi-layered fashion. This capability makes DCNNs particularly effective for distinguishing between different types of rice leaf diseases, even when the visual differences are subtle.

In a rice leaf disease detection system using DCNNs, the primary goal is to classify various types of diseases that may affect rice leaves, such as bacterial leaf blight, tungro, bacterial blast and brown spot, among others. The process generally begins with capturing images of infected rice leaves, which are then preprocessed to enhance the quality of the input data. Preprocessing techniques like resizing, normalization, and noise reduction ensure that the DCNN model receives clean and consistent images for training.

The core of the detection system is the DCNN model itself, which is trained on a labeled dataset of rice leaf images. During training, the model learns to associate specific patterns and features in the images with corresponding disease labels. This is achieved through a series of convolutional layers that apply filters to detect features, followed by pooling layers that reduce the dimensionality of the feature maps. As the network progresses through the layers, it captures increasingly complex features that help distinguish between different diseases.

Once trained, the DCNN model can be deployed to classify new, unseen images of rice leaves. The model's ability to generalize from the training data allows it to accurately identify diseases in real-world conditions, even when faced with challenges such as varying lighting, background clutter, or different angles of the leaves. This makes DCNN-based systems highly effective for field use, where environmental conditions may not be ideal.

CHAPTER 2

LITERATURE SURVEY

[1] Mehedi Hasan Bijoy et al., 2023:

The work by Mehedi Hasan Bijoy et al. presents a novel approach for rice leaf disease detection using deep Convolutional Neural Networks (dCNN) and an enhanced dataset. The authors emphasize the need for sustainable agricultural practices by employing a more accurate and efficient method for detecting rice leaf diseases. The research introduces a new dataset that improves model performance by incorporating a wide range of disease patterns, lighting conditions, and background noise. This work demonstrates the potential of using dCNNs for effective disease detection, achieving higher accuracy than conventional methods.

[2] V. K. Shrivastava and M. K. Pradhan, 2021:

This research explores machine learning paradigms based on color feature extraction for rice plant disease classification. The authors focus on using color features as a primary basis for disease classification, employing traditional machine learning algorithms. The study compares various classification models and highlights the potential of color-based features for disease detection. While the model performs well, it lacks the robustness of deep learning models when applied to more complex datasets.

[3] M. A. Azim et al., 2021:

Azim and colleagues propose an effective feature extraction method specifically designed for rice leaf disease classification. The paper focuses on improving the feature extraction process by incorporating advanced techniques that enhance the model's ability to capture crucial disease-related features. Their approach is more computationally efficient compared to full deep learning models but still offers good classification accuracy for rice leaf diseases.

[4] S. Saha and S. M. M. Ahsan, 2021:

This study employs intensity moments and random forest algorithms for rice disease detection. The research highlights how intensity moments—a technique for capturing image patterns—can be combined with the random forest algorithm to classify rice diseases. The study shows that this approach, while simpler than deep learning, can still produce competitive results when applied to structured data.

[5] P. Mekha and N. Teeyasuksaet, 2021:

Mekha and Teeyasuksaet focus on using the random forest algorithm for image classification of rice leaf diseases. The authors explore the effectiveness of traditional machine learning algorithms when applied to large datasets of rice leaf images. Their results demonstrate that random forest can achieve reasonable accuracy, although it lacks the sophistication and precision of deep learning approaches.

CHAPTER 3

PROBLEM STATEMENT

Rice leaf diseases like bacterial leaf blight, brown spot, tungro and leaf blast threaten crop yields and cause economic losses, with traditional detection methods being time-consuming and error prone. The challenge is to develop an automated, CNN-based system that accurately detects and classifies these diseases from images, overcoming environmental challenges to enable early intervention and minimize crop damage.

CHAPTER 4

OBJECTIVES

- Develop a highly accurate deep convolutional neural network (DCNN) model capable of identifying and classifying different types of rice leaf diseases, such as bacterial blight, brown spot, and leaf smut.
- Automate the process of disease detection to reduce the need for manual inspection by experts, minimizing human error and improving the speed of diagnosis.
- Ensure the model can reliably detect diseases in real-world conditions, including variations in lighting, background, and environmental factors.
- Enable early detection of diseases to facilitate timely intervention and prevent severe crop loss, contributing to sustainable agriculture.
- Design a scalable system that can be easily deployed in agricultural fields for real-time monitoring of large-scale rice plantations.

CHAPTER 5

METHODOLOGY

In this study, we aim to develop and evaluate model for the Rice Leaf Disease Detection using Machine Learning, methodology is divided into several stages: dataset preparation, data augmentation, data pre-processing, model architecture, model training and evaluation.

1. **Dataset Preparation:**

   We gathered several pre-existing rice leaf datasets from the internet, but encountered a major challenge in rice leaf disease identification due to the lack of a sufficiently large, publicly available dataset. Additionally, many publicly available datasets are unreliable because they often include duplicate or modified images from the training set in the test set, leading to artificially inflated performance measures. As a result, models trained on these datasets may not perform as well when applied to real-world data. The scarcity of large datasets is further complicated by the difficulty in gathering leaf data that captures subtle disease variations, environmental factors, and the labor-intensive process of accurately annotating images. To address this, we created an expanded dataset of 5,593 images across five disease classes: sheath blight, tungro, brown spot, leaf smut, and bacterial leaf blight. The dataset was divided into training (3,158 images), validation (1,277 images), and test sets (850 images) for model development and evaluation.

Class Name	# Images		
	Training	*Validation*	*Test*
Sheath Blight	371	221	149
Tungro	410	246	163
Brown Spot	936	282	187
Bacterial Blast	410	246	163
Bacterial Leaf Blight	1031	282	188
	Total = 3158	**Total = 1277**	**Total = 850**

Bacterial Leaf Blight Bacterial Blast Brown Spot Sheath Blight Tungro

2. Data Augmentation:

To improve CNN model performance, we addressed the challenge of limited training data by applying data augmentation techniques to expand the dataset. Eight transformations, including rotation, flipping, zooming, cropping, and shifting, were used to increase image diversity. Each image was first resized to 240x240 pixels for consistency, followed by horizontal and vertical shifts (0.2) and flipping with a 50% probability. Zoom-in, zoom-out, and rotation adjustments were applied to simulate real-world scenarios and capture different perspectives. These augmentations significantly enhanced the dataset, enabling the model to better generalize across various visual conditions and backgrounds, ultimately improving its ability to accurately detect rice leaf diseases.

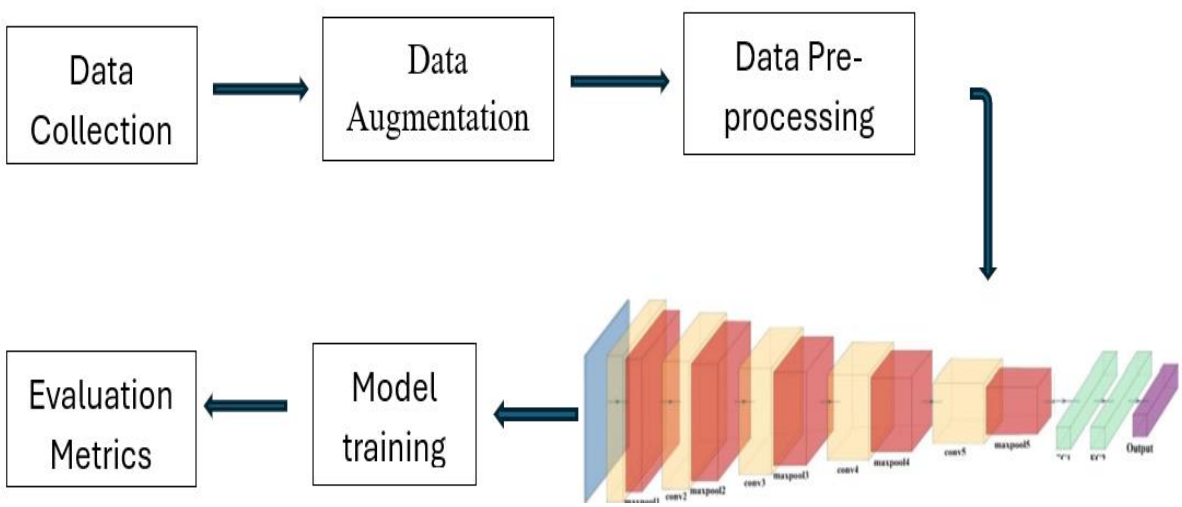
3. Data Pre-processing:

Data preprocessing is crucial for training dCNN models for rice leaf disease detection, helping to standardize and clean the data to overcome challenges like image quality, lighting variations, and background noise. The images are resized to 240x240 pixels to ensure consistency, and pixel values are normalized by scaling them between 0 and 1, which improves model convergence and reduces sensitivity to lighting changes. Noise in agricultural images is addressed using denoising techniques like Gaussian blurring, while brightness and contrast are adjusted using methods like CLAHE to enhance visibility of disease symptoms.

Background distractions such as dirt, machinery, or other plants are minimized using segmentation techniques to isolate the rice leaf, allowing the model to focus on disease-related features. Data augmentation, including flipping, rotation, zooming, and color jittering, is applied to increase dataset diversity, enhancing the model's ability to generalize across varied perspectives and lighting conditions. The dataset is split into training, validation, and test sets, ensuring balanced representation of disease classes. Addressing class imbalance through techniques like oversampling and adjusting the loss function ensures the model performs well across all disease categories, improving detection accuracy.

4. Model Architecture:

Our proposed model for rice leaf disease detection consists of two key components: feature extraction and classification. The feature extraction part uses convolution and pooling layers to capture relevant patterns, while the classification part employs dense (fully connected) layers for accurate predictions. To prevent overfitting, a dropout layer with a 10% ratio is applied. The model uses the Rectified Linear Unit (ReLU) activation function in hidden layers and Softmax in the output layer for probabilistic predictions. With just 0.27M parameters, the model is lightweight compared to larger architectures like AlexNet, ResNet50, and DenseNet121. Despite its small size, the model delivers excellent performance, making it highly efficient for rice leaf disease detection.



5. Model Training:

The model training process involves two key steps: forward propagation and backward propagation. In forward propagation, a batch of images is passed through the model, generating predictions based on the weights and biases. The outputs of hidden and output layer neurons are computed using weight matrices, biases, and activation functions. For backward propagation, the model's weights are updated based on the prediction loss, calculated using the cross-entropy loss function. The softmax function is applied to produce class probabilities, and the loss is computed by taking the negative logarithm of the predicted class probability. The Adam optimizer, which combines momentum and RMSprop, is then used to minimize the loss and update the model's weights and biases with each micro-batch.

6. Evaluation Metrics:

To evaluate the performance of our rice leaf disease detection model, we utilize key performance metrics: Accuracy, Precision, Recall, and F1 Score, which are derived from the confusion matrix.

- Confusion Matrix: This $N \times N$ matrix is used to assess classification models, where N is the number of classes. It compares the model’s predictions against actual labels and provides a detailed view of the model's performance, including the type of errors made. The matrix is structured with actual values on one axis and predicted values on the other. The four key components of a confusion matrix are:
    - True Positive (TP): Correctly predicted positive instances.
    - True Negative (TN): Correctly predicted negative instances.
    - False Positive (FP) or Type-I Error: Incorrectly predicted positive (actual value is negative).
    - False Negative (FN) or Type-II Error: Incorrectly predicted negative (actual value is positive).
- Accuracy: This metric measures the proportion of correctly predicted instances out of the total number of instances in the dataset. It is calculated by dividing the total number of correct predictions by the overall dataset size. A higher accuracy indicates better model performance.

- Precision: Precision quantifies the proportion of true positive predictions among all positive predictions made by the model. It is crucial in scenarios where False Positives (FP) are more concerning than False Negatives (FN), as it reflects the model's reliability.
- Recall: Recall measures the ability of the model to correctly identify actual positive cases. It is particularly important when the cost of False Negatives (FN) outweighs that of False Positives (FP).
- F**1 Score**: The F1 Score, also known as the F-Score or F-Measure, is the harmonic mean of Precision and Recall. It is useful for comparing models, especially when there is a trade-off between low precision and high recall or vice versa.

CHAPTER 6

RESULTS

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 126, 126, 32)	896
max_pooling2d (MaxPooling2D)	(None, 63, 63, 32)	0
conv2d_1 (Conv2D)	(None, 61, 61, 64)	18,496
max_pooling2d_1 (MaxPooling2D)	(None, 30, 30, 64)	0
conv2d_2 (Conv2D)	(None, 28, 28, 128)	73,856
max_pooling2d_2 (MaxPooling2D)	(None, 14, 14, 128)	0
conv2d_3 (Conv2D)	(None, 12, 12, 256)	295,168
max_pooling2d_3 (MaxPooling2D)	(None, 6, 6, 256)	0
conv2d_4 (Conv2D)	(None, 4, 4, 512)	1,180,160
max_pooling2d_4 (MaxPooling2D)	(None, 2, 2, 512)	0
flatten (Flatten)	(None, 2048)	0
dense (Dense)	(None, 512)	1,049,088
dropout (Dropout)	(None, 512)	0
dense_1 (Dense)	(None, 256)	131,328
dropout_1 (Dropout)	(None, 256)	0
dense_2 (Dense)	(None, 5)	1,285

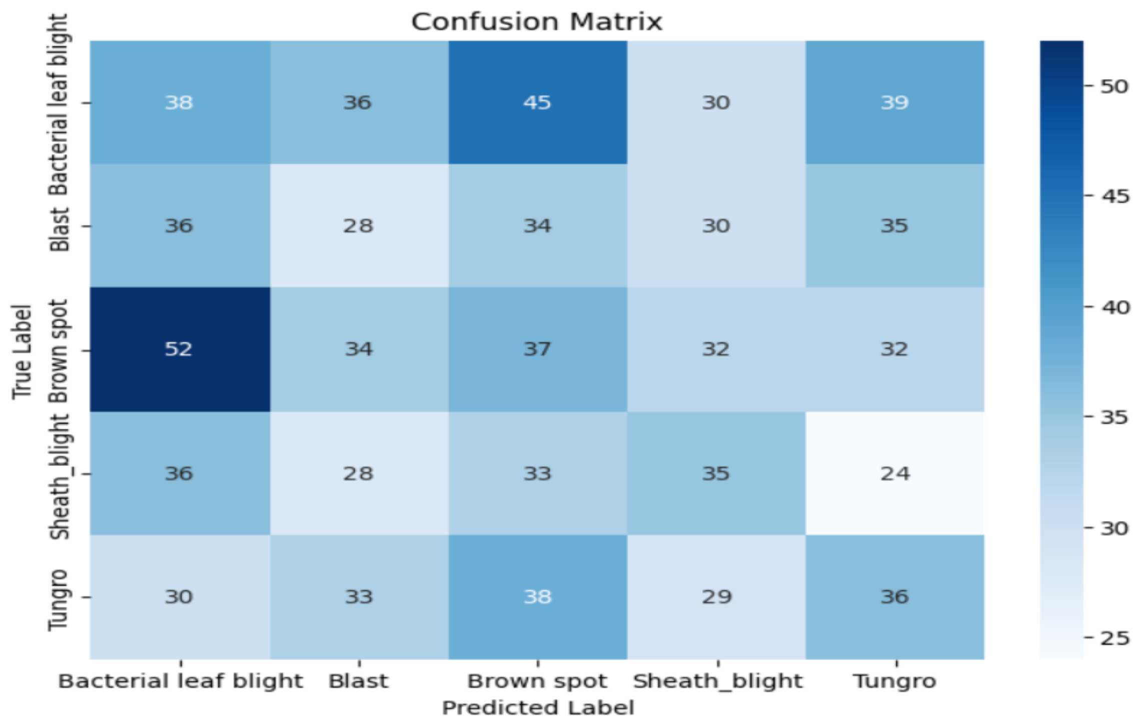
Total params: 2,750,277 (10.49 MB)
Trainable params: 2,750,277 (10.49 MB)

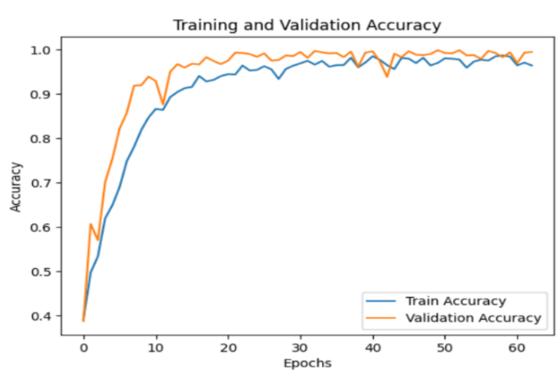
```
# Evaluate on the test data
test_loss, test_acc = model.evaluate(test_generator)
print(f"Test Accuracy: {test_acc * 100:.2f}%")
```

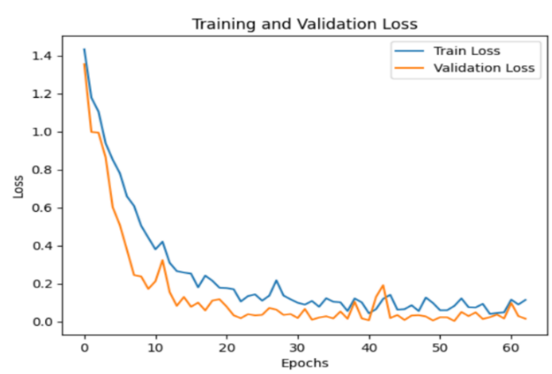
```
27/27 ━━━━━━━━━━━━━━━━━━━━ 413s 16s/step - accuracy: 0.9958 - loss: 0.0077
Test Accuracy: 99.30%
```

```
27/27 ━━━━━━━━━━━━━━━━━━━━ 4s 128ms/step
                       precision    recall  f1-score   support

Bacterial leaf blight       0.20      0.20      0.20       188
                Blast       0.18      0.17      0.17       163
           Brown spot       0.20      0.20      0.20       187
        Sheath_blight       0.22      0.22      0.22       156
               Tungro       0.22      0.22      0.22       166
```







CHAPTER 7

CONCLUSIONS

Detecting rice leaf diseases is vital for improving agricultural productivity, enabling early identification that helps farmers prevent infections from affecting their yields. Traditional detection methods often struggle, especially in resource-limited environments where lightweight models with fewer parameters are necessary. This study introduces a lightweight deep Convolutional Neural Network (dCNN) capable of identifying five prevalent rice leaf diseases: brown spot, tungro, bacterial blight, sheath blight, and bacterial blast.

Extensive trials confirm that our model effectively assists farmers in minimizing output losses, outperforming existing techniques while maintaining lower complexity. We enhance the dataset through expert input, providing new insights into rice leaf diseases and ensuring accurate performance. This research addresses the urgent need for accessible knowledge on disease management, and future efforts will focus on testing the model's effectiveness in real-world agricultural scenarios, equipping farmers with essential guidance for disease treatment and prevention.

REFERENCES

[1] Mehedi Hasan Bijoy1, Nirob Hasan2 , Mithun Biswas3 , Suvodeep Mazumdar4 , Andrea Jimenez4 , Faisal Ahmed5 , Mirza Rasheduzzaman6 , And Sifat Momen2 "Towards Sustainable Agriculture: A Novel Approach for Rice Leaf Disease Detection using dCNN and Enhanced Dataset".IEEE,2023

[2] Pothen and M. L. Pai, ''Detection of rice leaf diseases using image processing,'' in 2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC). IEEE, 2020, pp. 424–430

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

[4] J. Chen, D. Zhang, Y. A. Nanehkaran, and D. Li, ''Detection of rice plant diseases based on deep transfer learning,'' Journal of the Science of Food and Agriculture, vol. 100, no. 7, pp. 3246–3256, 2020.

[5] V. K. Shrivastava and M. K. Pradhan, ''Rice plant disease classification using color features: a machine learning paradigm,'' Journal of Plant Pathology, vol. 103, no. 1, pp. 17–26, 2021.

[6] M. A. Azim, M. K. Islam, M. M. Rahman, F. Jahan et al., ''An effective feature extraction method for rice leaf disease classification,'' Telkomnika, vol. 19, no. 2, pp. 463–470, 2021.

[7] S. Saha and S. M. M. Ahsan, ''Rice disease detection using intensity moments and random forest,'' in 2021 International Conference on Information and Communication Technology for Sustainable Development (ICICT4SD). IEEE, 2021, pp. 166–170.

[8] P. Mekha and N. Teeyasuksaet, ''Image classification of rice leaf diseases using random forest algorithm,'' in 2021 Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunication Engineering. IEEE, 2021, pp. 165–169.