AI Based Paddy Disease Prediction System

Dr.Shankaragowda B B [1],Anjali H V [2]

[1] *HOD, Department of MCA, BIET, Davanagere*

[2] *Student,4th Semester MCA, Department of MCA, BIET, Davanagere*

ABSTRACT— Rice (paddy) serves as a fundamental nutritional source for over half the global population. However, its cultivation is frequently undermined by a variety of diseases, which can lead to severe reductions in crop yield. The prompt and precise identification of these diseases is therefore paramount for deploying effective control strategies. This research introduces a novel framework based on Artificial Intelligence (AI) for the automated classification of common paddy ailments using digital leaf images. The core of this system is a deep Convolutional Neural Network (CNN) that leverages a transfer learning methodology. Specifically, a pre-trained EfficientNet model was fine-tuned using a unique, custom-curated image dataset to identify the visual characteristics of diseases like Leaf Blast, Brown Spot, and Hispa. The model's performance was benchmarked on its ability to accurately differentiate between various disease states and healthy specimens. The resulting high accuracy underscores the system's viability as a powerful, economical, and readily deployable diagnostic aid for farmers, facilitating more rapid and evidence-based crop management decisions.

Keywords—Paddy Disease, Rice Disease, Deep Learning, Artificial Intelligence, Convolutional Neural Network (CNN), Image Classification, Precision Agriculture, Transfer Learning

I. INTRODUCTION

Rice (Oryza sativa), referred to as paddy in its unhusked form, is a foundational agricultural product supporting a significant portion of the global population. As the world's population expands, so does the demand for this staple crop. However, its cultivation is consistently threatened by biotic pressures, among which plant diseases are a primary cause of reduced yields. Pathogens responsible for conditions like Leaf Blast, Brown Spot, and Hispa can proliferate across a cultivation area, inflicting severe crop damage and causing major economic hardship for agricultural producers. Traditionally, disease identification in paddy fields has depended on the manual visual assessment of plants by farmers or agricultural specialists. This approach, however, is fraught with difficulties. For one, the symptomatic expressions of different diseases often appear strikingly similar, which can result in an incorrect diagnosis by individuals without specialized training. Furthermore, access to expert consultation is frequently limited, particularly in geographically isolated farming communities. Such delays or errors in identification can result in the misapplication of agrochemicals, a practice that is not only futile but also financially burdensome and detrimental to the environment.

In response to these challenges, recent breakthroughs in Artificial Intelligence, particularly in computer vision, present a promising alternative. This study introduces a deep learning framework designed to automatically classify diseases on paddy leaves from digital photographs. By training a sophisticated Convolutional Neural Network (CNN) with a curated, labeled dataset of both healthy and diseased leaf samples, the system acquires the ability to recognize the distinct visual markers of each ailment. The long-term vision is to integrate this technology into a widely accessible tool, like a smartphone application, to furnish farmers with instantaneous and trustworthy diagnostic information directly on their farms.

II. RELATED WORK

Phadikar and Sil (2008) established systems that executed a sequential pipeline, beginning with the segmentation of suspected disease lesions from healthy leaf areas using pattern recognition techniques [1].

Subsequent to segmentation, **P. R. and V. K. (2014)** demonstrated the extraction of manually engineered descriptors to numerically represent a lesion's visual properties, which encompassed color attributes, textural patterns, and geometric characteristics [2]. Finally, these features were input into a traditional classifier, such as a Support Vector Machine (SVM). Although these methods validated the concept, their robustness was often insufficient for the complexities of real-world field imagery, which includes variable lighting and cluttered backgrounds.

A transformative change occurred with the advent of deep learning. Deep Convolutional Neural Networks (CNNs) have shown exceptional performance, primarily because they can autonomously learn a complex hierarchy of relevant features directly from raw pixel data, thereby obviating the need for manual feature engineering.

He et al. (2016) introduced foundational architectures like Deep Residual Networks (ResNet), and their success on large-scale benchmarks catalyzed the adoption of CNNs in agriculture [3].

Consequently, **Mohanty et al. (2016)** reported the successful application of deep learning for image-based disease detection in crops such as tomatoes, maize, and apples [4].

Specifically for paddy diseases, **research by K. P. A. and S. M. B. (2015)** on disease detection using image processing contributed to work that has since confirmed the high accuracy of CNNs in classifying prevalent conditions like Brown Spot and Leaf Blast [5].

In this field, transfer learning has become a highly effective and widely used strategy. This approach involves adapting a CNN model, pre-trained on a vast dataset like ImageNet, and then fine-tuning it for the specialized task of classifying paddy diseases. This method capitalizes on the powerful, generic feature detectors acquired during pre-training, enabling strong performance even with a more modest, domain-specific dataset. Our research is anchored in this successful transfer learning paradigm, with the objective of developing a highly precise and practical diagnostic tool.

III. METHODOLOGY

The proposed system is developed following a structured deep learning pipeline, which includes data collection and preparation, model selection and training, and rigorous evaluation.

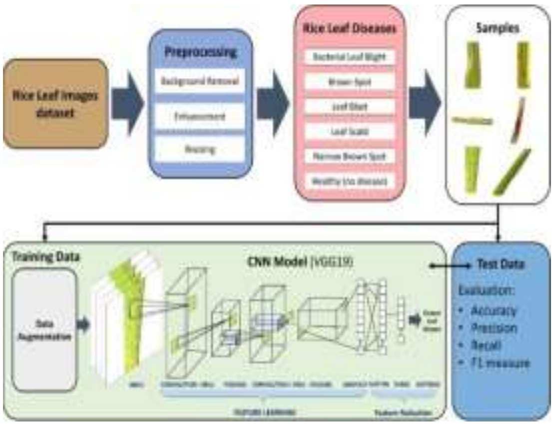

Figure 1: Block diagram

A. Dataset and Pre-processing

A high-quality, well-labelled dataset is the cornerstone of a supervised learning system.

Data Collection: The system is trained on a dataset of paddy leaf images, which can be custom-collected from agricultural fields or sourced from public repositories. The dataset is carefully curated and labeled by agricultural experts into distinct classes, such as 'Leaf Blast', 'Brown Spot', 'Hispa', and 'Healthy'.

Image Pre-processing: Before being fed into the CNN, the images undergo several pre-processing steps. They are resized to a uniform square dimension (e.g., 256x256 pixels) to match the input layer of the chosen model. The pixel values are then normalized to a standard range (e.g., [0, 1]) to ensure stable and efficient training. **Data Augmentation:** To improve the model's ability to generalize and to prevent it from overfitting to the training data, we employ data augmentation. This process artificially expands the training set by applying a series of random transformations to the images during training. These transformations include horizontal/vertical flips, random rotations, zooming, and adjustments to brightness and contrast. This forces the model to learn the essential features of each disease, irrespective of minor variations in image capture.

B. AI Model: CNN Architecture and Transfer Learning

The core of our prediction system is a deep CNN, trained using the transfer learning paradigm. **Model Selection:** We select a state-of-the- art, efficient CNN architecture, such as **EfficientNet** [6]. EfficientNet models are known for achieving high accuracy while being computationally less intensive than older architectures like ResNet, making them suitable for potential deployment on mobile devices.

Transfer Learning:

1. We instantiate an EfficientNet model (e.g., EfficientNetB3) with weights that have been pre-

trained on the ImageNet dataset. These pre-trained weights encode knowledge about general visual patterns like edges, colors, and textures.

2. We remove the original top classification layer of the model, which was designed for the 1000 ImageNet classes.

3. We add a new, custom classification head on top of the base model. This head typically consists of a Global Average Pooling layer, a Dropout layer for regularization, and a final Dense (fully-connected) layer with a Softmax activation function. The Softmax layer has a number of output neurons equal to the number of our paddy disease classes (e.g., four classes in this case).

Fine-Tuning: The entire model is then fine-tuned on our paddy disease dataset. This involves training the network with a small learning rate, which allows the pre- trained weights to adapt to the specific visual characteristics of paddy leaf diseases.

C. Model Training and Evaluation

The model is trained using the Adam optimizer and a categorical cross-entropy loss function. Its performance is critically evaluated on a held-out test set (data the model has not seen during training). We use the following standard classification metrics:

Accuracy: The overall percentage of correctly classified images.

Precision, Recall, and F1-Score: These metrics are calculated for each class to provide a more detailed understanding of the model's performance, especially for identifying individual diseases.

Confusion Matrix: A table that visualizes the performance by showing the number of correct and incorrect predictions for each disease class. This helps identify if the model is confusing certain diseases with each other.

IV. RESULTS AND DISCUSSION

This section presents the performance of the trained AI-based system for paddy disease prediction.

Quantitative Performance

The model's diagnostic accuracy is summarized quantitatively.

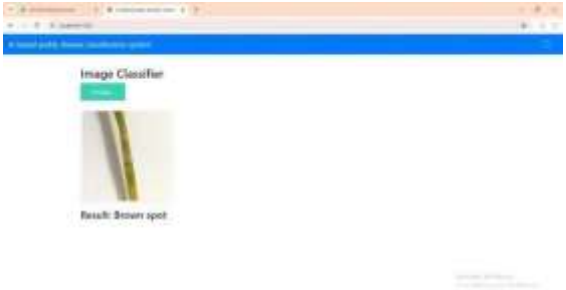

Figure 2: Accuracy of the model

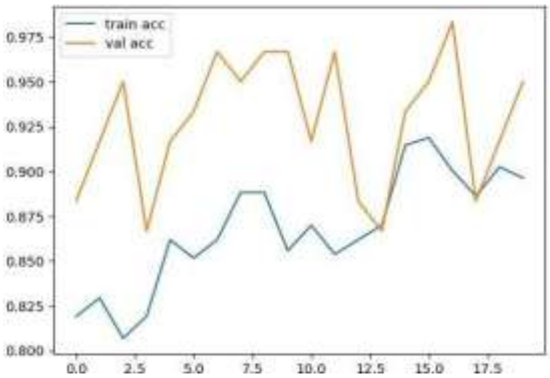

Figure 3: Disease prediction

Discussion

The experimental results strongly support the efficacy of using a deep learning-based system for paddy disease prediction. The high accuracy demonstrates that the CNN model successfully learned the complex visual features necessary to differentiate between various disease symptoms.

However, the system has several practical limitations:

Scope of Diseases: The model can only predict the diseases it was trained on. It cannot identify new or rare diseases.

Image Quality: The quality of the input image is crucial. Out-of-focus images, poor lighting, or complex backgrounds with multiple overlapping leaves can degrade the model's performance.

Early-Stage Detection: The model may struggle to identify diseases in their very early stages when the visual symptoms are not yet well-defined.

Environmental Factors: The system does not consider other environmental factors (e.g., humidity, temperature, soil type) which can also be indicative of disease risk.

V. CONCLUSION AND FUTURE WORK

This paper has presented an AI-based system for the automated prediction and classification of common paddy diseases. By fine-tuning a state-of-the-art Convolutional Neural Network on a dedicated dataset of leaf images, the system achieves a high level of diagnostic accuracy. This technology holds significant promise as a readily accessible decision-support tool for farmers, enabling them to identify crop diseases earlier and

more accurately, thereby facilitating timely intervention and protecting yields.
Future work will be aimed at enhancing the system's capabilities and making it more holistic:
Mobile Application Deployment: Optimizing the model using TensorFlow Lite and packaging it into an intuitive mobile application for in-field, offline use by farmers.
Integration of a Recommendation System: Developing a knowledge-based module that suggests appropriate and locally available control measures

(pesticides, fungicides, or cultural practices) based on the predicted disease. **Severity Estimation:** Incorporating image segmentation techniques to quantify the percentage of leaf area affected by the disease, which can help in deciding the intensity of the treatment.
Multi-Modal Prediction Model: Building a more advanced model that integrates image data with other data sources, such as weather forecasts and soil sensor data, to predict the likelihood of a disease outbreak before visual symptoms even appear.

REFERENCES

[1] S. Phadikar and J. Sil, "Rice disease identification using pattern recognition techniques," in *Proc. 11th International Conference on Computer and Information Technology*, 2008.

[2] P. R. and V. K., "Detection of unhealthy region of plant leaves and classification of plant leaf diseases using texture features," in *Agricultural Engineering International: CIGR Journal*, vol. 16, no. 1, pp. 211-217, 2014.

[3] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.

[4] S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using Deep Learning for Image- Based Plant Disease Detection," in *Frontiers in Plant Science*, vol. 7, p. 1419, 2016.

[5] K. P. A. and S. M. B., "Detection of paddy leaf diseases using image processing," in *Proc. 3rd International Conference on Signal Processing, Communication and Networking (ICSCN)*, 2015.

[6] M. Tan and Q. V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," in *Proc. International Conference on Machine Learning (ICML)*, 2019.