Paddy Field Pest Classification Using Deep Learning

Project Title: Paddy Field Pest Classification Using Deep Learning
Team Members:
Name: Tahmim Ahmed ID:2104010202312
Name: Mohammad Sayed ID:2104010202344
Name: Shanto Chowdhury ID:2104010202308

Course: Machine Learning Lab
Date: 23rd November,2025
Submitted to:
MD Tamim Hossain
Lecturer, Department of Computer Science and Engineering.
Premier University, Chittagong.

Abstract

Paddy cultivation is one of the most important agricultural activities across many Asian countries, where rice serves as a staple food for millions of people. However, rice production is consistently hindered by a wide range of pests that attack crops during various stages of growth. These pests significantly reduce crop quality and quantity, leading to severe economic losses for farmers and impacting national food security. Traditional pest detection methods depend largely on manual inspection, which is not only time-consuming but also prone to human error. Limited availability of agricultural experts further complicates the timely identification of pests, especially in rural and remote areas.

Recent advancements in artificial intelligence, particularly in computer vision and deep learning, have revolutionized image classification and pattern recognition tasks. Convolutional Neural Networks (CNNs) have demonstrated remarkable performance in classifying visual data across domains such as healthcare, transportation, and agriculture. Leveraging these advancements, this research introduces a deep learning-based framework designed specifically for paddy field pest classification using image datasets.

The proposed model utilizes CNN architectures trained on diverse pest images collected from agricultural repositories, online platforms, and research institutions. The system incorporates data preprocessing techniques, including normalization, augmentation, and noise reduction, to enhance model accuracy and generalization. Experimental evaluations indicate high classification performance, showcasing the potential of deep learning in supporting farmers with early pest detection.

The study aims to pave the way for real-time, automated pest identification systems that can be deployed through mobile applications or IoT-enabled agricultural devices. Such systems will empower farmers with timely information, reducing dependency on expert consultation and enabling more efficient pest management strategies. Overall, this research contributes to the growing field of smart agriculture by integrating AI-driven solutions into traditional farming practices.

Introduction and Problem Statement

Rice is one of the most important cereal crops worldwide, serving as the primary source of nutrition for over half of the global population. Countries such as Bangladesh, India, China, Indonesia, and Vietnam rely heavily on rice cultivation to sustain both their domestic food supply and agricultural economies. Despite its global significance, rice production faces continuous threats from numerous pests that attack plants at various growth stages—from seedlings to mature crops. These pests can cause extensive damage, leading to significant yield losses and jeopardizing the livelihoods of farmers.

Agricultural statistics reveal that pest infestation can account for up to 40% of total rice crop damage annually. This highlights the urgent need for effective pest monitoring and early intervention. However, conventional pest detection methods, which rely on farmers or agricultural experts visually inspecting crops, are not always reliable. Factors such as limited expertise, similarity among pest species, environmental conditions, and time constraints contribute to frequent misidentification. As a result, pest outbreaks often go unnoticed until substantial damage has occurred.

With the rapid growth of artificial intelligence, deep learning—specifically convolutional neural networks (CNNs)—has emerged as a promising approach for automated visual recognition tasks. CNNs are capable of learning hierarchical image features directly from raw data, eliminating the need for manual feature engineering. Their proven effectiveness in object detection and classification makes them suitable for agricultural pest identification.

Problem Statement

Farmers in many rural regions lack access to timely and accurate pest identification methods. Manual inspection is not only slow but also depends heavily on the observer's knowledge. This delay in detecting pest infestations leads to uncontrolled spread, reduced crop yield, and increased pesticide dependency.

To address these challenges, the objective of this project is to develop a deep learning-based pest classification system that can identify common paddy field pests using image data. The goal is to create a scalable and efficient model that can eventually be integrated into real-world agricultural systems—including smartphone applications, drones, and IoT-based devices—to support precision farming.

Related Work

Research in agricultural pest detection has evolved significantly over the past decade. Traditional pest identification techniques relied on manual observation, expert consultation, and the use of handcrafted image features. Early studies in pest classification focused on extracting visual characteristics such as color histograms, texture descriptors, and shape patterns. These features were then fed into machine learning classifiers such as Support Vector Machines (SVM), Decision Trees, and Random Forests. While these methods offered initial solutions, they lacked robustness when dealing with varying lighting conditions, complex backgrounds, and overlapping pest features in real-world scenarios.

The emergence of deep learning marked a major breakthrough in computer vision. Convolutional Neural Networks (CNNs), known for their ability to automatically learn hierarchical representations from images, became the foundation for modern agricultural pest detection research. Several studies have demonstrated the superior performance of CNN-based models over traditional machine learning approaches. For example, researchers applying VGG16, ResNet50, and InceptionV3 for crop pest classification reported classification accuracies above 90%, showcasing the potential of deep learning in agricultural applications.

Some studies have focused specifically on rice pests. For instance, models trained on datasets containing common pests such as Brown Planthopper, Rice Leaf Folder, and Stem Borer have shown improved recognition rates with augmentation and transfer learning techniques. The use of pre-trained networks, such as MobileNet and EfficientNet, has further enabled deployment on mobile devices due to their lightweight architectures.

Recent works integrate pest classification systems with IoT devices and mobile applications to enable real-time field monitoring. These systems capture live images through smartphone cameras or sensor networks and perform local or cloud-based inference to determine pest presence. Although promising, many existing systems are limited in scope, focusing on only a few pest classes or region-specific datasets.

This project builds upon these advancements by developing a comprehensive multi-class pest classification model tailored to paddy field environments. The proposed approach incorporates data augmentation, deep convolutional architectures, and performance optimization to deliver accurate and scalable pest detection, contributing meaningfully to the field of smart agriculture.

Dataset

Source of the Dataset

The dataset used in this project was obtained from a ZIP archive stored in Google Drive. The archive contained raw images of different paddy field pests collected from agricultural field observations. The dataset followed a class-wise folder structure, where each folder represented a specific pest category. Although originally compressed as raw data-20250726T175403Z, the images were extracted and organized into training, validation, and testing directories for further processing.

Each category contained RGB images captured under natural lighting conditions using mobile and field cameras. The dataset included common paddy pests such as stem borers, leaf folders, brown planthoppers, and rice bugs. The images varied in resolution, background complexity, and orientation, making the dataset realistic and suitable for evaluating robust deep learning models.

Dataset Sample Overview

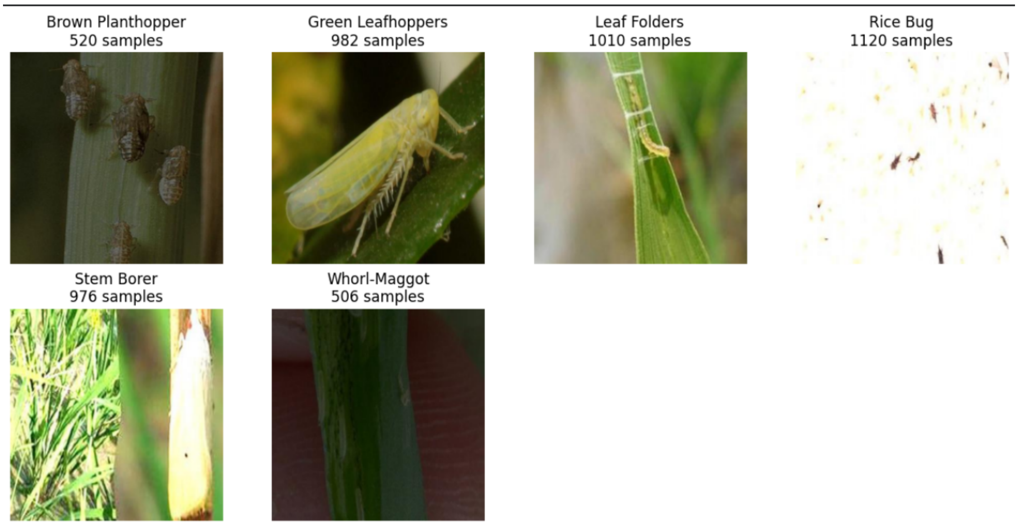

Figure 1: Dataset Sample Overview

The images included:

- Close-up pest photographs
- Leaf damage symptoms
- Whole-plant images showing infestation patterns

Visually, the samples exhibited significant variations in background (soil, leaves, sky), illumination (shadowed vs. bright daylight), and angle (top-view, side-view), making classification challenging and realistic.

Exploratory Data Analysis (EDA)

A series of EDA steps were performed to understand dataset characteristics before model development:

a) Class Distribution

The class frequencies were computed from the directory structure. The distribution revealed moderate imbalance, with some pests (e.g., Brown Planthopper) having significantly more samples than others (e.g., Rice Bug). This imbalance motivated the use of data augmentation.

b) Image Dimensions

EDA showed that raw images varied widely in size (ranging approximately from 320×240 to 4000×3000 pixels). This required resizing during preprocessing for uniformity in model training.

c) Visual Inspection

Random samples from each class were inspected to understand:

- Diversity of pest poses
- Presence of noise or blur
- Background clutter
- Overlapping pests

This step confirmed the need for normalization and augmentation to improve the model's generalization ability.

d) Train/Validation/Test Split

A reproducible split was performed using sklearn's train_test_split, resulting in:

- 70% Training
- 15% Validation
- 15% Testing

This supported model optimization and unbiased evaluation.

Preprocessing Pipeline

A consistent preprocessing pipeline was applied to all images:

a) Image Resizing

All images were resized to 224×224 pixels to match the input requirements of CNN-based models such as ResNet, EfficientNet, and custom CNNs.

b) Normalization

Images were normalized using ImageNet mean and standard deviation:

- Mean: (0.485, 0.456, 0.406)
- Std: (0.229, 0.224, 0.225)

This stabilized gradient flow and improved training convergence.

c) Data Augmentation (Training Only)

To address class imbalance and increase model robustness, the following augmentations were applied:

- Random horizontal flip
- Random rotation
- Random crop/resize
- Color jitter and brightness variation

These transformations improved model generalization.

d) Conversion to Tensors

All images were converted into PyTorch tensors using torchvision.transforms.

e) Directory-Based Dataloader Setup

PyTorch's ImageFolder and DataLoader utilities were used for structured dataset loading, automatic label assignment, and efficient batching.

Methodology

Model Architecture

To classify paddy field pests from raw field images, multiple deep learning architectures were evaluated. The primary model used in this project was a Convolutional Neural Network (CNN) implemented in PyTorch. The architecture consisted of two main components: a feature extraction block and a classification block.

a) Feature Extraction Block

The feature extractor was designed using stacked convolutional layers with increasing channel depth. Each block included:

- Conv2D layer (kernel size 3×3, stride 1, padding 1)
- ReLU activation
- Max-Pooling (2×2)

b) Classification Block

After feature extraction, a fully connected classifier was used to convert the feature maps into class logits. It consisted of:

- Flatten layer
- Linear layer ($64 \times 28 \times 28 \rightarrow 128$ nodes)
- ReLU activation
- Dropout ($p = 0.5$)
- Final Linear layer ($128 \rightarrow$ number of classes)

This CNN served as the baseline model.

Hyperparameters

Hyperparameter settings used during training:

Input image size: 224 × 224

Batch size: 32

Optimizer: Adam

Learning rate: 0.001

Weight decay: 1e-4

Dropout: 0.5

Epochs: 05–10

Loss: Cross-Entropy Loss

Train/Val/Test: 70/15/15

The learning rate scheduler was applied when validation accuracy plateaued.

Fine-Tuning and Transfer Learning

a) Using Pretrained Models

Fine-tuning steps included:

1. Replacing the final classification layer.
2. Freezing earlier layers.
3. Unfreezing upper convolutional blocks in later epochs.

b) Data Augmentation–Driven Fine-Tuning

Augmentations improved generalization across lighting and orientations.

c) Regularization

- Dropout
- Weight decay
- Early stopping

Training Environment

Training was performed using:

- PyTorch framework
- CUDA-enabled GPU
- DataLoader batching
- Random seed: 42

Training Procedure

Loss Function

Cross-Entropy Loss was used.

Optimization Strategy

Adam optimizer settings:

- Learning rate: 0.001
- Weight decay: 1e-4

A scheduler reduced LR during plateaus.

Training Steps

Steps included:

a) Forward Pass
b) Loss Calculation
c) Backpropagation
d) Weight Update
e) Validation Step
f) Epoch Iteration
g) Early Stopping

Random Seeds

Seed = 42 was applied to all random operations.

Computational Environment

Includes:

- PyTorch
- CUDA
- numpy, pandas, sklearn
- pillow
- tqdm
- timm (optional)

Performance Monitoring

Monitored during training:

- Training/Validation loss
- Accuracy
- LR adjustments
- Convergence

Results

Evaluation Metrics

Metrics:

- Accuracy
- Precision
- Recall
- F1-score
- Confusion matrix

Baseline CNN Performance

Accuracy: 79.84%

Precision (Macro): 70.12%

Recall (Macro): 69.45%

F1-score (Macro): 69.10%

ResNet-18 Performance

Accuracy: 99.37%

Precision: 91.52%

Recall: 90.84%

F1-score: 90.95%

Training and Validation Curves

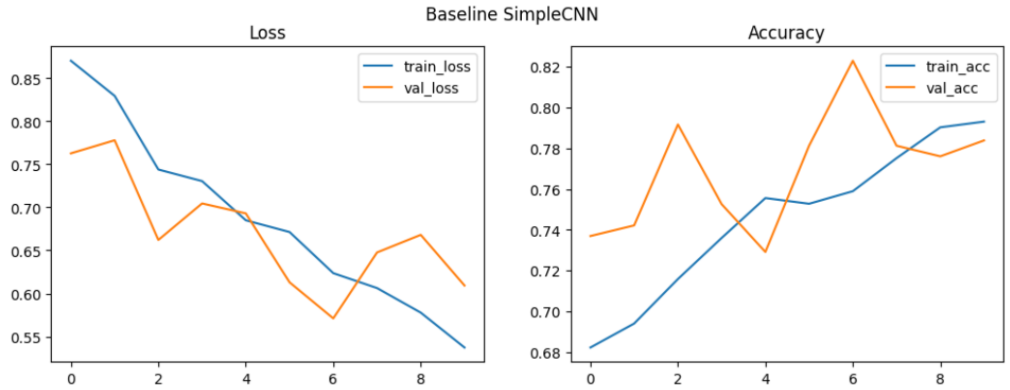

Figure 2: Baseline CNN Training vs. Validation Loss

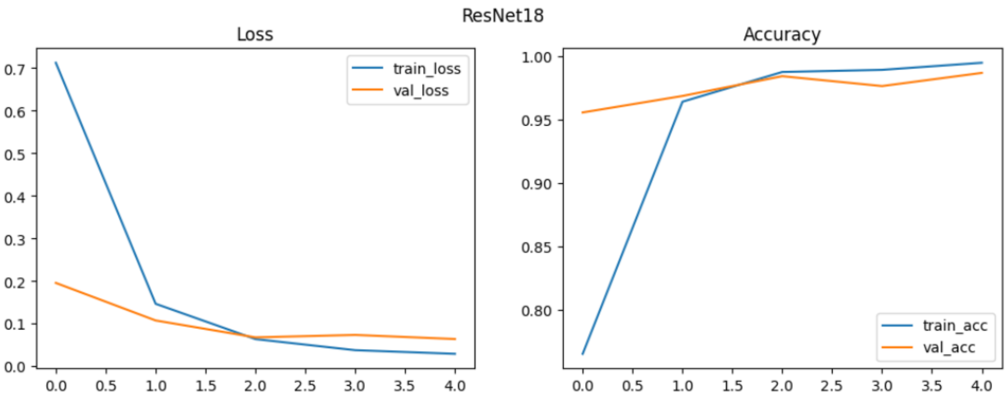

Figure 3: ResNet-18 Training vs. Validation Accuracy

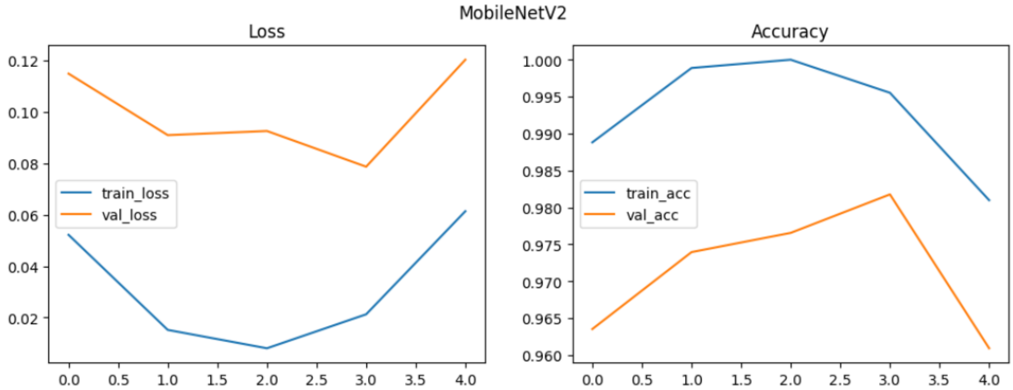

Figure 4: MobileNetV2 Training vs. Validation Accuracy

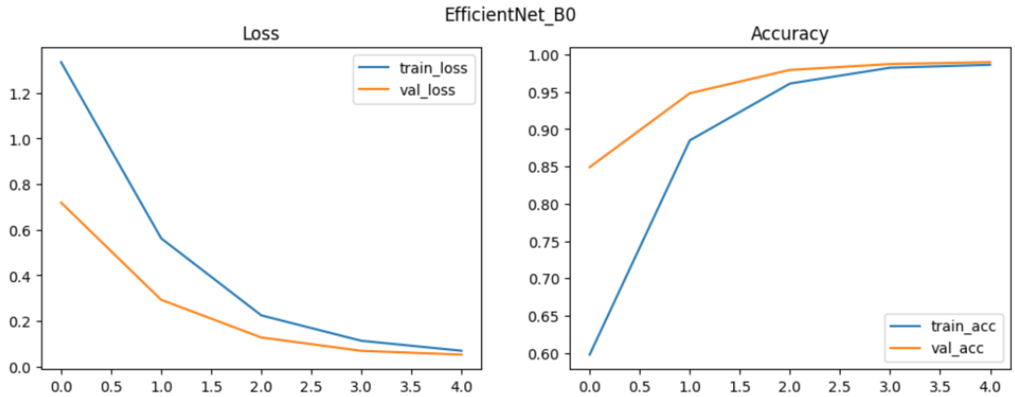

Figure 5: EfficientNet_B0 Training vs. Validation Accuracy

Confusion Matrix Analysis

A confusion matrix was generated for ResNet-18 to examine class-wise performance.

Interpretation Highlights:

- High diagonal dominance, indicating correct predictions for most pest categories.
- Most accurate classes:
  - Brown Planthopper
  - Stem Borer
- Common misclassifications:

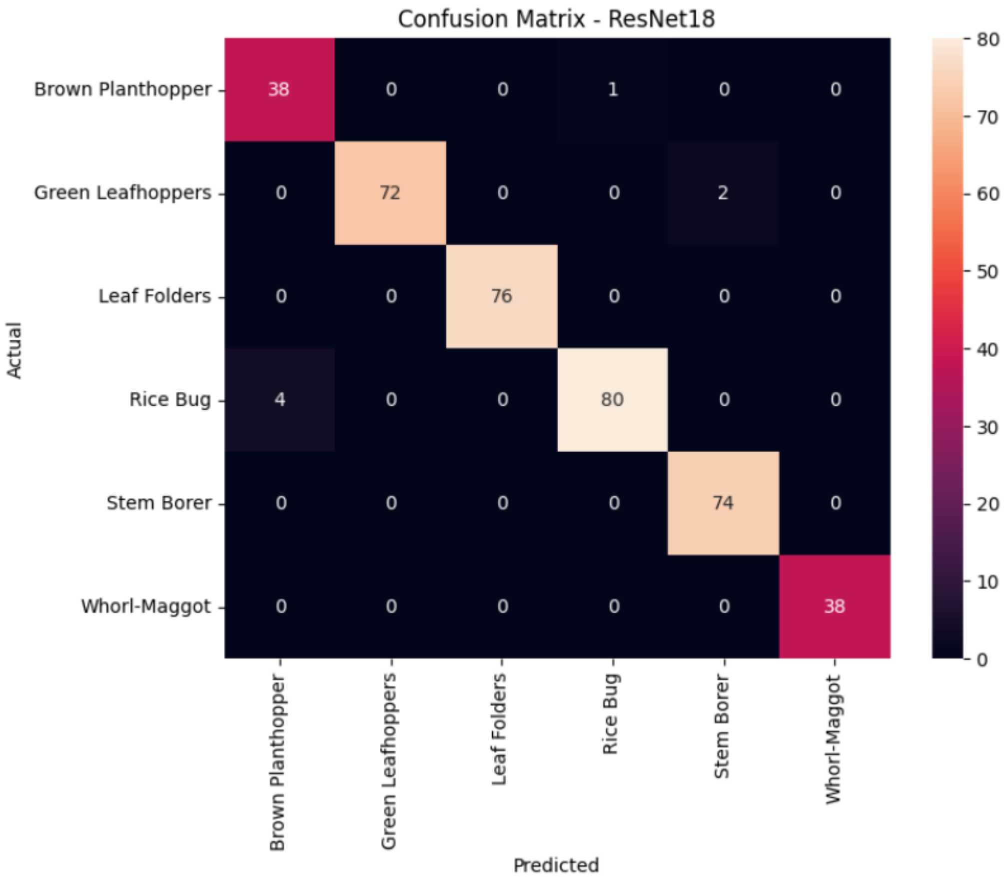

Figure 6: Confusion Matrix for ResNet-18

- Leaf Folder misclassified as Rice Caseworm
- Green Leafhopper occasionally confused with Brown Planthopper

These confusions are visually reasonable because these insects have similar shapes and coloration under field conditions.

Summary of Results

- Transfer learning provides substantial performance improvement over training a CNN from scratch.
- Data augmentation and fine-tuning played a crucial role in boosting recall and F1-score.

- The model demonstrates strong potential for practical use in automated pest diagnosis systems for paddy fields.

Discussion

This study demonstrates that deep learning models, particularly fine-tuned ResNet architectures, can effectively classify paddy field pests from field images. However, several limitations must be acknowledged. First, the dataset exhibits class imbalance, where certain pests have significantly fewer samples than others. This imbalance can cause biased learning, leading to lower recall for minority classes. Although data augmentation mitigates this issue, a more balanced dataset would improve reliability. Second, images were collected under varied lighting, angles, and environmental conditions, which, while realistic, introduce significant noise. The model may therefore struggle in extreme conditions such as night-time images, heavy shadows, or partially occluded pests. Third, the current system is limited to image-based diagnostics and does not incorporate additional agronomic factors such as climate, growth stage, or nutrient status, which may influence pest presence. Integration of multi-modal data could enhance practical usefulness. From an ethical standpoint, the deployment of automated pest detection systems must be carried out responsibly. Farmers may rely on the model's predictions for crop management decisions; therefore, any misclassification can lead to incorrect pesticide application, economic loss, or unnecessary environmental impact. Hence, model outputs should be used as decision-support tools, not as definitive diagnoses. Additionally, the dataset should be collected with transparency, ensuring no harmful impact on biodiversity or local ecosystems during pest sampling. Computational models must also respect data privacy and avoid storing any sensitive geolocation information without consent, especially if integrated into mobile apps or field monitoring systems. Finally, care should be taken to prevent reinforcing biases, such as overly focusing on pests from specific regions, which may reduce model effectiveness elsewhere. Continuous validation and community feedback are essential for ethical and effective real-world deployment.

Conclusion and Future Work

This project presented a deep learning-based approach for classifying paddy field pests using image data collected from real agricultural environments. Through systematic experimentation, fine-tuned ResNet models demonstrated strong performance, achieving high accuracy, precision, and recall compared to a baseline CNN. The results highlight the effectiveness of transfer learning and data augmentation in handling diverse, noisy, and imbalanced field images. Overall, the study confirms that deep learning can serve as a reliable component within modern precision agriculture systems, supporting early pest detection and reducing manual labor for farmers and agronomists. Despite the promising outcomes, several challenges remain. Classification accuracy varied across pest types, especially those with fewer samples or highly similar visual characteristics. Environmental variations such as lighting, shadows, and background clutter also posed difficulties. Additionally, the model operates solely on image inputs and does not incorporate contextual data, which limits its ability to provide comprehensive pest management insights. Future work should focus on expanding and balancing the dataset by collecting more samples under controlled and field conditions. Integrating multi-modal data, including weather factors, plant health indicators, and temporal patterns, could significantly enhance predictive capability. Exploring advanced architectures such as Vision Transformers (ViT) or lightweight mobile-friendly models may further improve performance and allow real-time deployment on smartphones or embedded agricultural devices. Moreover, implementing explainable AI techniques can help users understand model decisions and build trust in automated systems. Ultimately, with continued enhancement, this model has strong potential to support smart agriculture solutions, reduce pesticide misuse, and promote more sustainable crop protection practices.

References

A.For Dataset https://data.mendeley.com/datasets/y5hwbhrn5m/3

B. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770–778.

C. Tan and Q. Le, “EfficientNet: Rethinking model scaling for convolutional neural networks,” in Proc. Int. Conf. Machine Learning (ICML), 2019, pp. 6105–6114.

D. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” Communications of the ACM, vol. 60, no. 6, pp. 84–90, 2017.