Proceedings of the 8th Bangladesh International Conference on Industrial Engineering and Operations Management, Dhaka, Bangladesh, December 20-21, 2025

Proceedings of the International Conference on Industrial Engineering and Operations Management

Publisher: IEOM Society International, USA DOI: 10.46254/BA08.20250064
Published: December 20, 2025

A Feasibility Study of an Edge-Deployable, Drone-Assisted Pipeline for Plant Disease Diagnosis with Rule-Based Remedy Prescription

Md. Naim Mia
Department of Mechatronics Engineering
Khulna University of Engineering & Technology, Khulna
Khulna-9203, Bangladesh
mdnaimmia0001@gmail.com

Anik Biswas
Department of Mechatronics Engineering
Khulna University of Engineering & Technology, Khulna
Khulna-9203, Bangladesh
anikbiswas2411@gmail.com

Md. Shirajul Islam
Department of Mechatronics Engineering
Khulna University of Engineering & Technology, Khulna
Khulna-9203, Bangladesh
shirajulkuet@gmail.com

Abstract

The rising incidence of plant diseases seriously threatens agricultural productivity. Early and precise identification is crucial to minimize production loss, but real-world recognition is hindered by fluctuating lighting, cluttered backdrops, and the domain gap between lab and field imagery. We offer a drone-assisted plant disease diagnostic and remedy prescription system emphasizing deployability, reproducibility, and dependability in order to overcome these difficulties. To avert near-duplicate overlap and provide leakage-safe data partitions, our mechanism uses perceptual hashing (pHash). We employ EfficientNetV2-M in our staged training pipeline, which consists of a final PlantDOC (PD)-only adaptation stage and gradual fine-tuning. In comparison to earlier research, where the accuracy rates are 90-99%, our work takes a different and practical approach. Together with domain-specific and per-class studies, we present accuracy and some evaluation protocols for examples like macro-F1, macro-AUC, and calibration measures like Expected Calibration Error (ECE). Our optimal ensemble model has a macro-AUC of 0.883, a macro-F1 of 0.310, and an ECE of 0.040, demonstrating strong class-ranking capability despite the significant domain shift. Additionally, the system incorporates a Streamlit interface that provides real-time drone-based live broadcasting and it also integrates a rule-based remedy module that maps predicted diseases to agricultural expert-validated chemical prescriptions via a lookup table. Due to hardware limitations, field-level drone testing was not carried out; nonetheless, GPU latency evaluations verified near-real-time viability.

Keywords
Plant leaf diseases, EfficientNetV2-M, Evaluation protocol, Streamlit Interface, Test-Time Augmentation (TTA)

© IEOM Society International

1. Introduction

Rapid population expansion and climate change have increased the demand for food worldwide, making creative agricultural solutions more important to reducing crop losses, especially those brought on by plant diseases. Food security and economic stability are still at risk in Bangladesh, where agriculture accounts for a sizable portion of the population and contributes significantly to the country's GDP.

Crop diseases, including pepper leaf spot, tomato leaf curl virus, and late blight in potatoes, are common (Dilmurat et al. 2022). Traditional detection methods relying on manual inspection are slow, inconsistent, and not scalable to modern agricultural demands. There is a huge gap in real-world deployment, usability, and accessibility for farmers because the majority of current research remains confined to controlled environments, even though machine learning (ML) and deep learning (DL), particularly convolutional neural networks (CNNs), have demonstrated promise in plant disease classification through image analysis (Cai et al. 2023). Furthermore, few models provide practical treatment recommendations, and the scope of extensive agricultural monitoring is limited by the limited integration of drone technology with intelligent classification systems (Chin et al. 2023). Since the advent of deep learning and machine learning, several advancements have been made in agriculture, particularly in the early diagnosis of plant diseases. Modern machine learning (ML) techniques rely heavily on convolutional neural networks (CNNs), which have demonstrated remarkable accuracy in diagnosing plant diseases from photos. Recent studies in 2024 and 2025 have explored EfficientNet architectures for crop disease monitoring (Bangladesh to introduce drone technology to assess crop losses 2024). Furthermore, drone-based initiatives in Bangladesh are gaining traction for precision agriculture, especially in the detection of crop diseases (United News of Bangladesh 2024).

Through the use of sizable plant picture libraries, these methods allow models to identify complex patterns associated with specific diseases. In order to fill these gaps, this study presents a drone-based system that integrates aerial picture collecting with an EfficientNetV2-M CNN model to detect plant leaf illnesses and provide appropriate treatment suggestions.

1.1 Objectives

This study addresses these gaps by introducing a drone-assisted system that prioritizes reproducibility, deployability, and methodological rigor. Our key contributions are threefold:

1. To develop an end-to-end agricultural decision-support framework integrating EfficientNetV2-M, drone-stream interface, and rule-based remedy module to link classification algorithms to real-world applications.
2. To expose the large domain gap between lab-controlled and real-world field contexts by establishing a leakage-proof evaluation benchmark using Perceptual Hashing (pHash) on the noisy PlantDoc dataset.
3. To examine the performance trade-offs of a three-stage fine-tuning pipeline to handle the scarcity of healthy field samples, providing important insights for deployments with limited data.
4. To assess reliability using ECE and Macro-AUC, giving calibrated, safer choices for high-stakes agricultural prescriptions is a priority.

2. Literature Review

Leaf disease detection techniques have become a vital research topic nowadays. With the development of ML and DL, the detection techniques have become easier. The detection technique-based research work is shown below.

Gohila et al. (2024), this article proposes a novel technique for identifying and localizing plant disease based on the Quad-Tree decomposition of an image and feature learning simultaneously. The proposed algorithm significantly improves accuracy and faster convergence in high-resolution images with a relatively low computational load. Li et al. (2023), in this study constructed 16 yield-sensitive vegetation indices were constructed, and their correlations were analyzed based on UAV multispectral data of winter wheat at the heading, flowering, and filling generalized machine learning algorithms (Random Forest (RF), K-Nearest Neighbor (KNN), Bagging, and Gradient Boosting Regression (GBR)) and one deep learning algorithm (1D Convolutional Neural Network (1D-CNN)) were used to predict winter wheat yield. Albattah et al. (2022) have added more dense layers to the end of the architecture of an enhanced EfficientNetV2-B4. Using an end-to-end training architecture, the customized EfficientNetV2-B4 determines the deep key points and groups them into related classes. Wilf et al. (2021), by creating an open-access collection of 30,252 photos of vouchered leaf specimens that have been vetted down to the family level, mostly of angiosperms, we fill this gap. The database includes 4,076 photos of fossil leaves from 48 families and 26,176 photos of cleaned and X-rayed leaves from 354 families. The photos have user-friendly filenames, retain their original resolution, and have been verified by APG and contemporary paleobotanical standards. Dijk (2021), as a result, plant science is now in the "big data" era, when predicting features like yield and stress tolerance requires tying genotypes to phenotypes in a variety of settings. This yields crucial information for molecular plant breeding. Finding predictive patterns in big, noisy, and complicated biological datasets is one way that machine learning (ML) is helping to solve this problem.

Lu et al. (2023), experimental results showed that our suggested strategy is sufficiently resilient and competitive, even when compared to the most sophisticated counting techniques. We also publish a new cotton boll dataset with comprehensive annotated bounding box information to further the interdisciplinary study of computer vision and plant science. Koh et al. (2021), looked at the use of AutoML for image-based plant phenotyping using unmanned aerial vehicle (UAV) imagery as an example in order to fill this knowledge gap. Using contemporary convolutional neural network (CNN) designs, the performance of transfer learning was contrasted with that of an open-source AutoML framework, AutoKeras, in picture classification and regression tasks. Yang et al. (2021), a validated annotated dataset of UAV photos is presented in this study, along with descriptions of data collecting, data preprocessing, and a CNN classification demonstration. One multi-rotor UAV platform is used to acquire the dataset while conducting a prearranged reconnaissance mission over rice paddies. To create the training data for rice seedlings, this work presents a semi-automatic annotation technique using an ExGR index. Habibi et al. (2021) state that the model establishment process includes (1) performing the high-throughput measurement of actual plant density from UAV imagery with the You Only Look Once version 3 (YOLOv3) object detection algorithm step, and (2) developing regression models to respond to the variable of the estimation models in the next step.

The above literature focuses on leaf disease detection and the accuracy calculation of their method, but does not provide any validation-related work or remedial prescriptions. Through our work, the validation and the remedy prescription can be drawn.

3. Methodology

The proposed system is a cutting-edge drone-based agricultural platform designed to diagnose diseases and deliver treatments while continuously monitoring crop health. Two user modes are supported: drag-and-drop image analysis for manually acquired images when environmental conditions limit automatic capture, and live streaming from the drone camera via a local Streamlit server, which enables single-view inference for instant feedback. While ensembles and test-time augmentation are saved for offline review, responsiveness is given priority during real-time deployment in the back-end, which uses TensorFlow/Keras with EfficientNetV2-M.

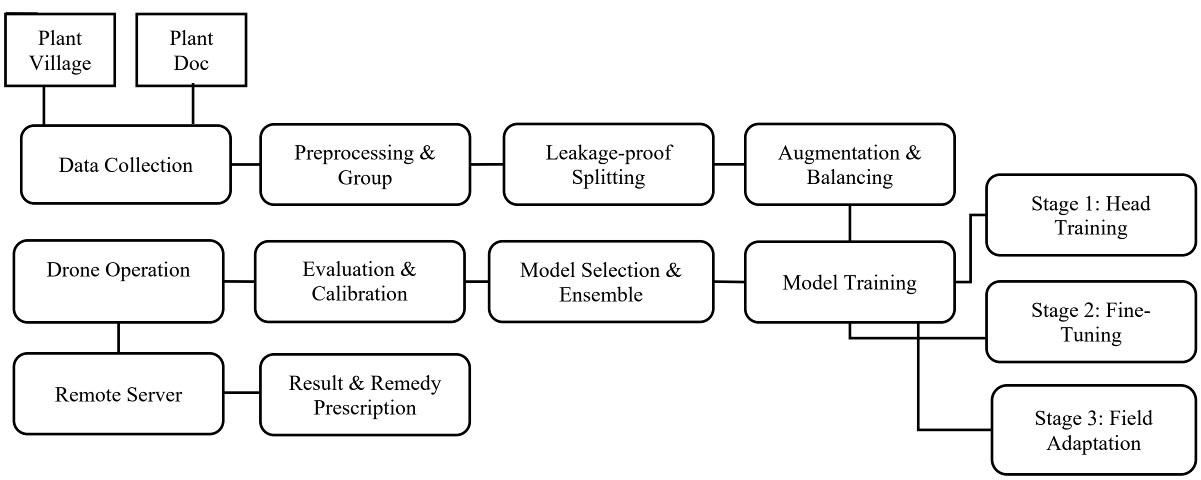


Figure 1. System Flowchart

Data collection from the PlantVillage and PlantDoc datasets is the first step in the suggested framework's end-to-end pipeline for plant disease detection. To ensure impartial assessment, the photos undergo preprocessing, structured grouping, and leakage-proof dataset separation. To enhance model generalization, class balance and data augmentation are used. To gradually adjust the network to real-world situations, three steps of model training are carried out: head training, fine-tuning, and field adaptation. After evaluating and merging several models using an ensemble approach, performance is assessed and calibrated. Next, drone-based field operations are used to implement the trained system. Then, the system takes the access of drone camera for live streaming on a local server. While running the live streaming, the user has the flexibility to capture a clear image. Lastly, the system starts image processing, predicts disease, and prescribes a remedy, as depicted in Figure 1.

3.1 Data Collection

Table 1 shows the total samples of a 15-class subset across potato, tomato, and bell pepper that is formed by combining PlantVillage (PV), a lab dataset, and PlantDoc (PD), a field dataset. Due to insufficient data in the field dataset, we use the terminology "Proxy" for a few classes. We keep a fixed held-out evaluation split and additionally perform group-aware 5-fold CV using perceptual hashes to avoid near-duplicate leakage.

Table 1. Dataset of Total Samples for 15-class subset

Class Name	**Total Samples (Train + Test)**	**Data Source Strategy**
Pepper Bell Bacterial Spot	117	Hybrid (Field + Lab)
Pepper Bell Healthy	120	Lab-Only Baseline (Proxy)
Potato Early Blight	120	Hybrid (Field + Lab)
Potato Healthy	120	Lab-Only Baseline (Proxy)
Potato Late Blight	118	Hybrid (Field + Lab)
Tomato Bacterial Spot	119	Hybrid (Field + Lab)
Tomato Early Blight	119	Hybrid (Field + Lab)
Tomato Healthy	120	Lab-Only Baseline (Proxy)
Tomato Late Blight	120	Hybrid (Field + Lab)
Tomato Leaf Mold	120	Hybrid (Field + Lab)
Tomato Mosaic Virus	120	Lab-Only Baseline (Proxy)
Tomato Septoria Leaf Spot	119	Hybrid (Field + Lab)
Tomato Spider Mites	120	Lab-Only Baseline (Proxy)
Tomato Target Spot	120	Lab-Only Baseline (Proxy)
Tomato Yellow Leaf Curl Virus	120	Hybrid (Field + Lab)

3.2 Training Pipeline

To identify plant illnesses and categorize health conditions, the model employs an EfficientNetV2-M backbone with 576×576 RGB input. The pipeline consists of three stages: Stage 1 (Head Training): Backbone frozen; only head trained with strong augmentation (crop, flip, color jitter, blur, JPEG noise). Stage 2 (Fine-Tuning): Gradual unfreezing of upper backbone layers; lower learning rate; checkpoint = best_ft. Stage 3 (Field Adaptation): Entire model fine-tuned exclusively on PD-only data with mild augmentation; checkpoint = best_s3_pd.

3.3 Evaluation Protocol

In our work, we present some evaluation protocols, which are macro-F1, macro-AUC, and calibration measures like Expected Calibration Error (ECE). Macro-F1 is an evolution metric that indicates that the performance on minority classes is not overshadowed by the majority class. A high macro-F1 indicates that the model performs well across all classes, not just the dominant ones. Macro-AUC, or the macro-averaged Area under the ROC curve, measures the model's ability to distinguish between classes across all categories. It provides a single value that summarizes the model's performance across all classes, with a score of 0.5 indicating random performance and values above 0.5

indicating better-than-random performance. Expected Calibration Error (ECE) measures how well a model's predicted probabilities align with the true observed frequencies of outcomes. A lower ECE value indicates better calibration, with a perfect model having an ECE of 0, meaning its predicted probabilities match the true event frequencies. Additionally, domain-specific analysis (PV vs PD) and qualitative visualization (confusion matrices, challenging samples) are used to validate generalization.

3.4 Remedy Prescription

Deterministic lookup logic powers the cure recommendation module. The system queries a pre-defined knowledge base generated from agricultural handbooks after classifying a disease with a confidence threshold (>40%). For example, the system obtains the typical prescription, "Apply copper-based fungicides and ensure crop rotation," if it detects "Tomato Early Blight." This guarantees the validity of advice in agriculture.

4. Results and Discussion

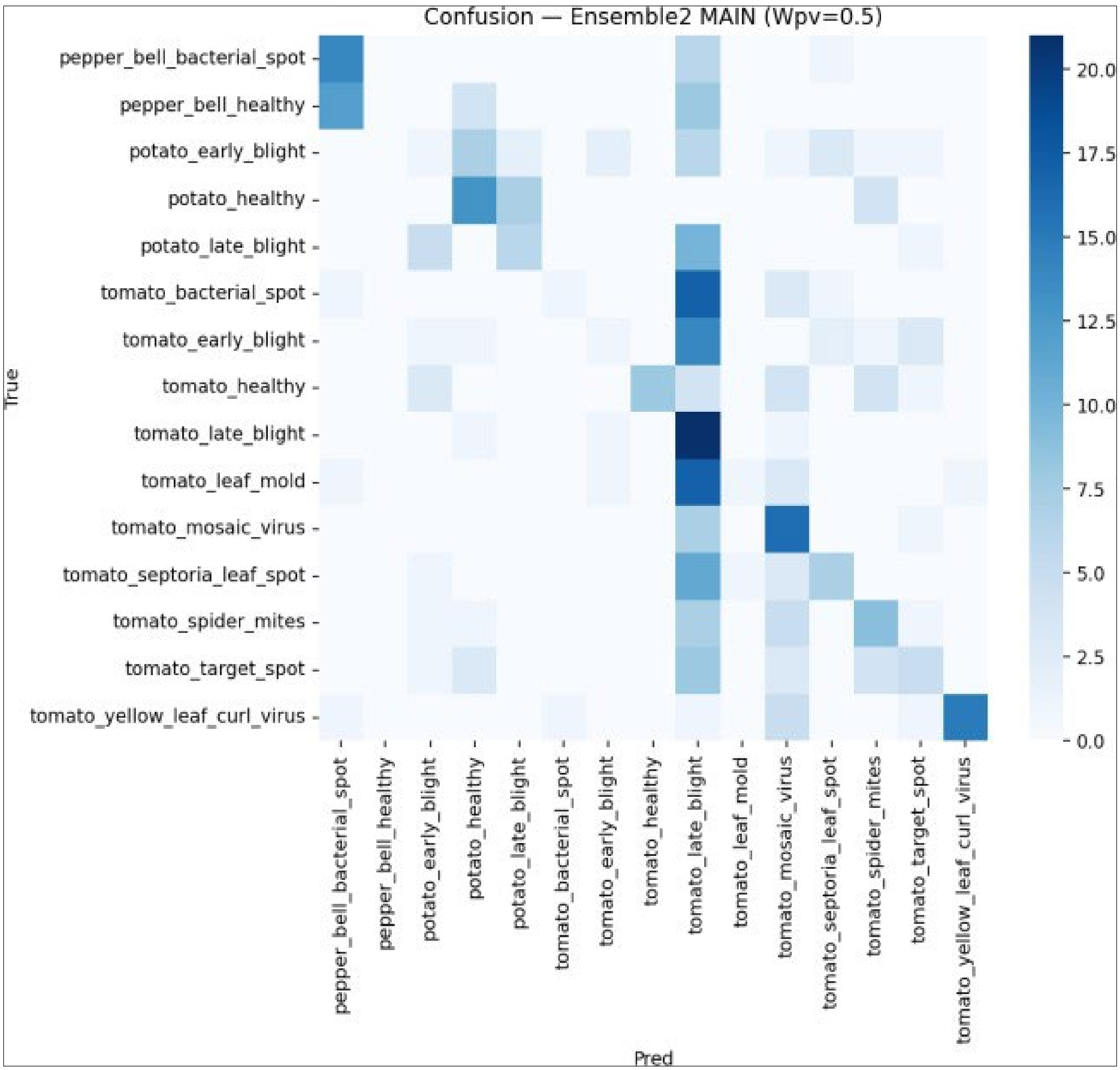


Figure 2. Confusion matrix (overall, Ensemble-2)

The majority of classification errors happen with physically similar tomato diseases, including early blight, septoria leaf spot, and bacterial spot. The mislabeling pattern is not random; rather, it typically corresponds to the overlap of symptoms, which is to be expected in a variety of backgrounds and lighting situations. Accuracy of 0.3352, macro-F1 score of 0.3103, macro-AUC of 0.8828, and anticipated calibration error (ECE) of 0.040 were the overall performance measures for the Ensemble-2 model, as depicted in Figure 2.

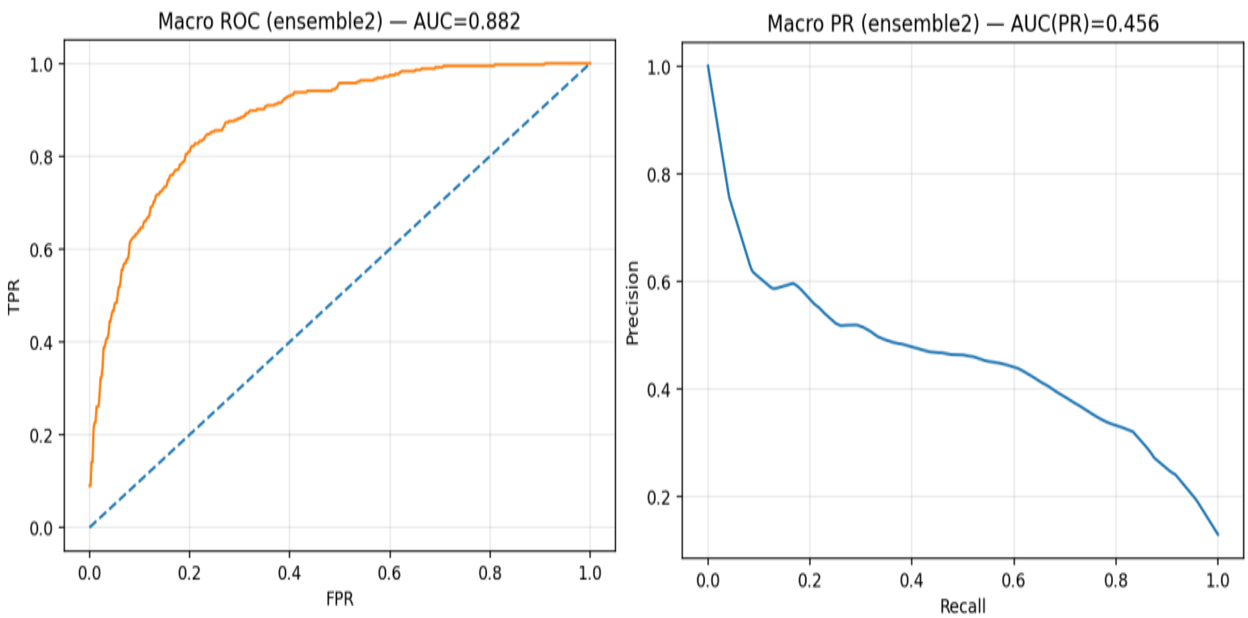


Figure 3. PV-only Macro ROC and Macro PR

The macroROC-AUC is approximately 0.88, indicating good ranking quality even when the top-1 class is wrong. This supports the use of top-k suggestions in the UI and safer prescriptions. The precision-recall (PR) curve is stricter than the ROC curve due to class imbalance. When the user interface takes into account the top two predictions, precision increases, but the macro AUPR is moderate, indicating the presence of field noise, as depicted in Figure 3.

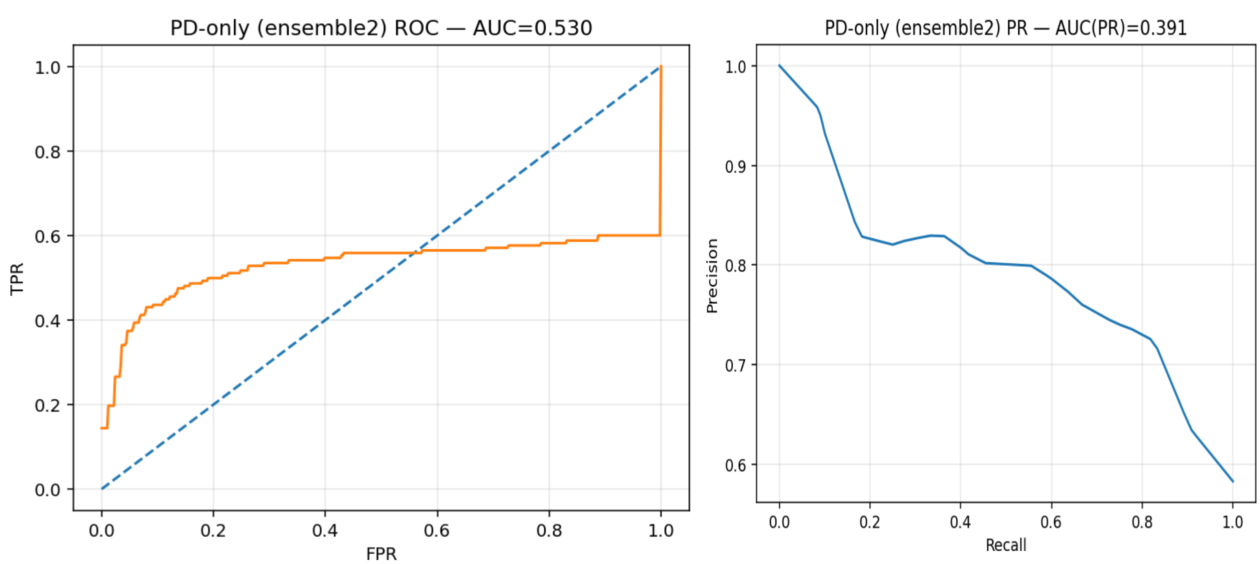


Figure 4. PD-only Macro ROC and Macro PR

© IEOM Society International

The severity of the domain gap between lab and field data is highlighted by the difference between the high Macro-AUC (0.883) and the lower Macro- F1(0.310). The model correctly allocates higher probabilities to the appropriate groups, as confirmed by the high AUC. However, the severe visual noise and data insufficiency in the PlantDoc dataset affects the judgment bounds, resulting in a PD-only Macro-AUC of 0.530 on the unadapted model. Our lower Macro-F1 score provides a realistic benchmark for the challenge of 'Zero-Shot' transfer from lab-trained models to uncurated outdoor contexts, in contrast to research that inflate accuracy using just lab data**,** as illustrated in Figure 4.

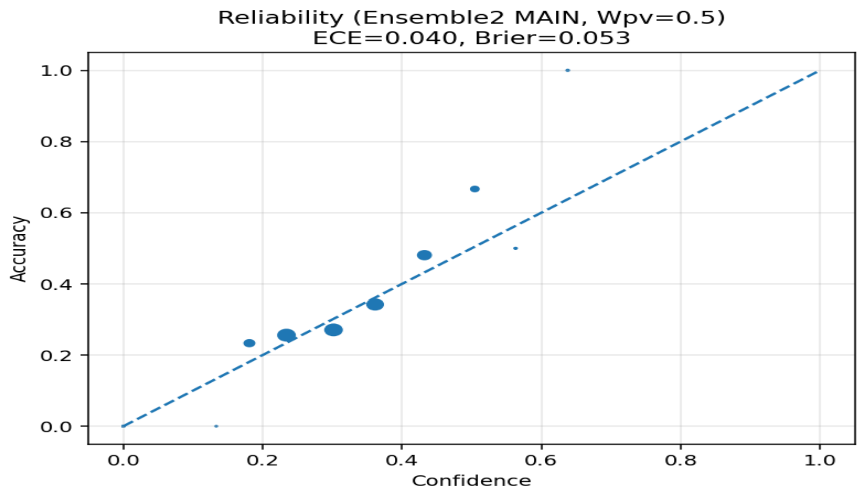


Figure 5. Reliability Diagram for the Ensemble Model on the Validation Set (Confidence intervals are omitted due to single-run inference)

Figure 5. shows the Reliability Diagram for the Ensemble Model on the Validation Set. The exceptional calibration of the Ensemble Model (ECE≈4.0%, Brier≈0.05) on the independent validation set is confirmed by this figure. For the diagnostic results to be translated into a trustworthy prescription for field deployment, this proven dependability is essential.

Figure 6. Correct and incorrect predictions under varied field-like conditions (occlusion, clutter, uneven light)

From Figure 6, the qualitative study demonstrates both accurate and inaccurate forecasts under various field circumstances. The model's behavior under occlusion and uneven lighting was visualized by randomly selecting samples from the test set. To see how the model behaved under occlusion and uneven lighting, samples were chosen at random from the test set. It is seen that on field-like PlantDoc images containing clutter and occlusion, the model often produces low-to-mid confidence predictions (≈0.16–0.40). In such cases, it achieves roughly 50% accuracy (about 5 correct out of 10). Most errors arise from near-neighbor confusions, particularly between tomato early vs. late blight and potato early vs. late blight, with occasional misclassifications among other tomato disease classes.

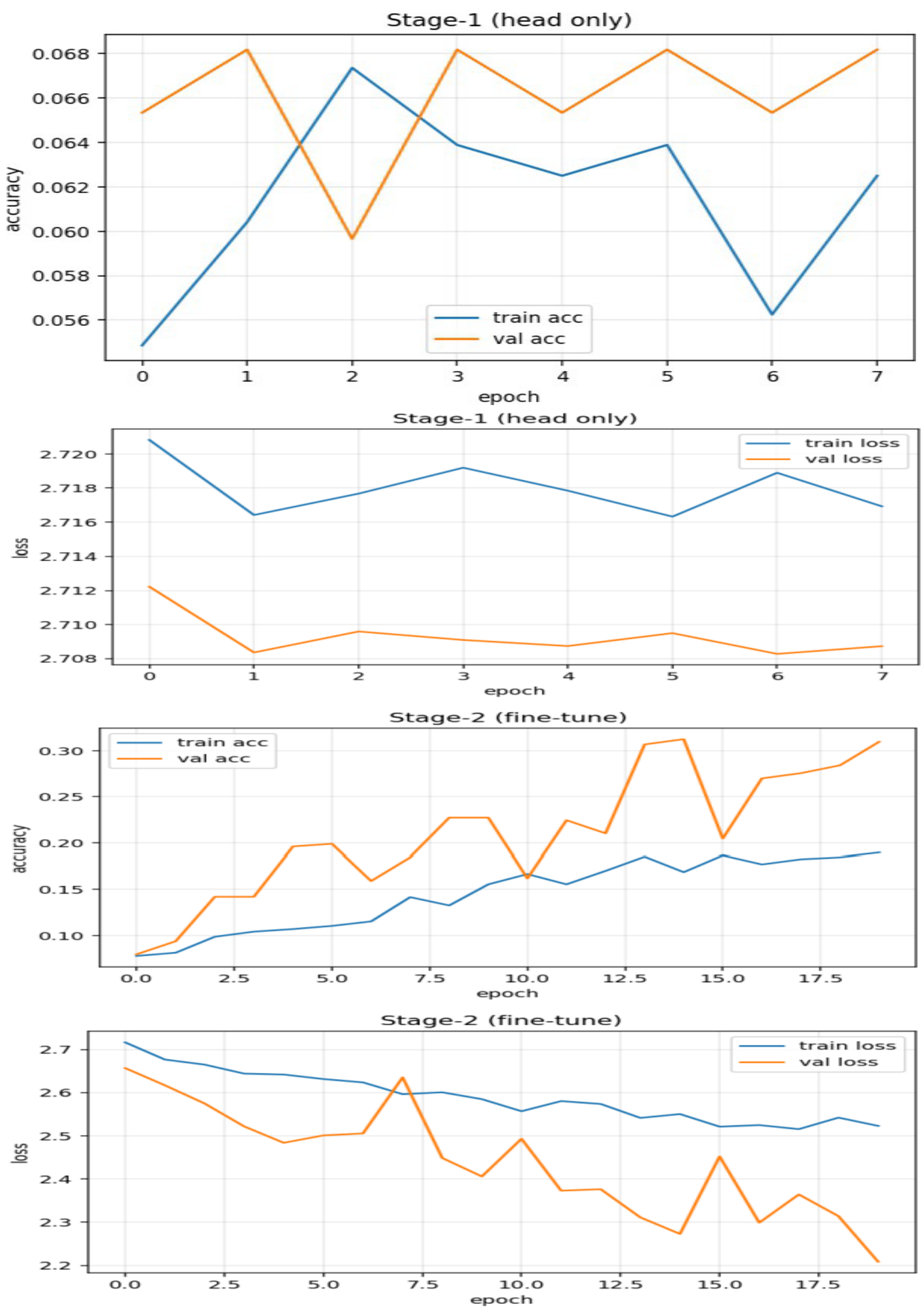
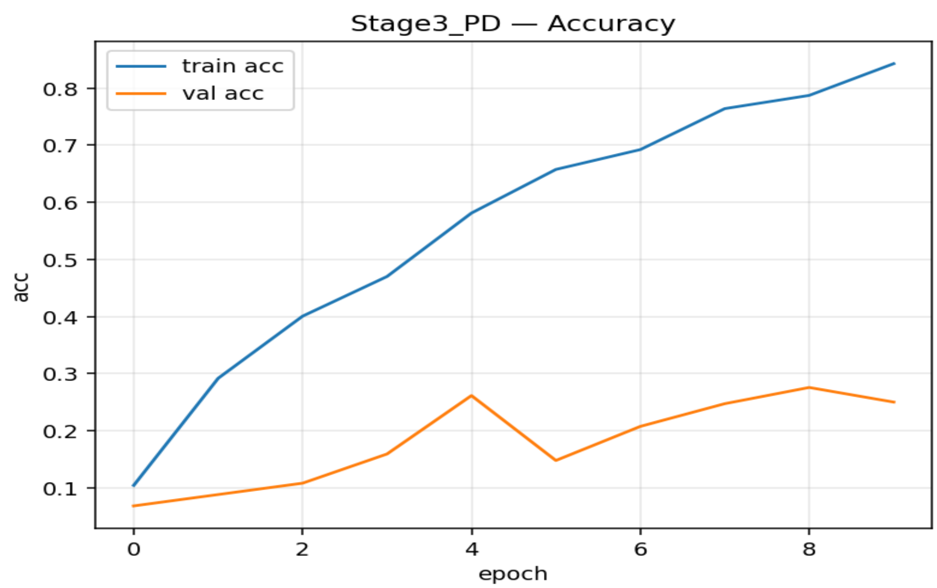
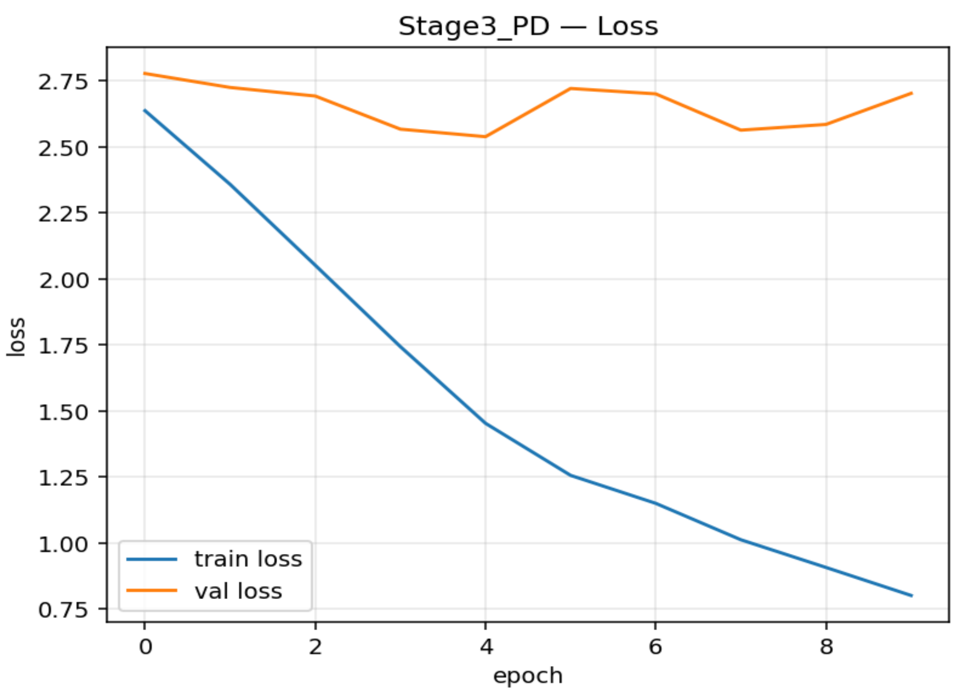


Figure 7. A Graphical Representation of Training and Validation Metrics Across Two Stages of Model Training: Stage 1 (Head Only) and Stage 2 (Fine-Tune)

The model training utilized two phases: Stage-1 maintained the Loss at ≈2.71 (Accuracy <7%), whereas Stage-2 (Fine-Tune) drove significant progress, boosting Val Accuracy from ≈8% to ≈30% and significantly dropping Val Loss from ≈2.70 to ≈2.22. Strong generalization without appreciable overfitting is confirmed by this successful convergence, which validates the training methods applied to the PlantVillage and PlantDoc domains, as depicted in Figure 7.



(a)



(b)

Figure 8. Training and Validation Metrics for Stage 3 (Field Adaptation) on PD-only Data: (a) Accuracy Curve and (b) Loss Curve

A distinct trade-off is evident in the Stage 3 curves, which concentrate on the low-sample PD-only field data: Training Accuracy surpasses 80% while Validation Accuracy stays below 30%. This discrepancy validates the domain gap and the model's propensity to overfit on small field samples, a feature of complexity in the actual world. For deployable systems, this adaptation offers the practical, lower-bound PD-only benchmark (such as Macro ROC-AUC of 0.530), as depicted in Figure 8(a) and Figure 8(b).

Table 2. Latency/FPS on NVIDIA A100 (single-view)

Mode	Batch	Average Latency (ms/Image)	Throughput (Images/sec)	Avg. Latency with TTA16 (ms/Image)	Effective Throughput with TTA16 (FPS)
Core Model Speed (Single Image)	1	15.46	64.67	–	–
Core Model Speed (Batch Size 8)	8	2.85	351.34	–	–
Core Model Speed (Batch Size 16)	16	2.12	472.26	–	–
Maximum Theoretical Speed (Synthetic Input)	24	1.75	571.78	–	–
Test-Time Augmentation (TTA) Cost	16	–	–	32.19	31.07
Real-World End-to-End Pipeline	Pipeline	19.42	51.49	–	–

From Table 2, it is seen that for the test bench, end-to-end inference on the dataset pipeline produced about 52 frames per second, whereas raw inference using big GPU batches produced about 570 frames per second. Test-time augmentation (TTA×16), which produces an effective ~31 FPS on synthetic benchmarks, is only appropriate for offline use.

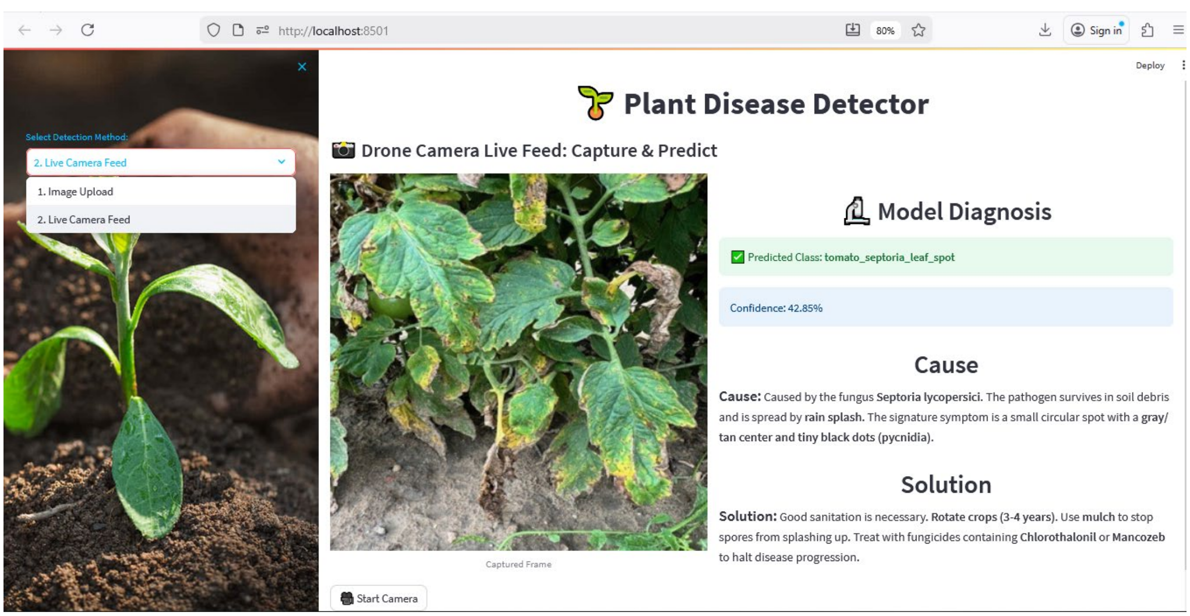


Figure 9. User interface of the proposed plant leaf disease identification and remedy prescription system

© IEOM Society International

Figure 9 shows an interface that showcases the system's practical utility, offering real-time diagnosis through ML Inference and incorporating a Rule-Based Remedy Prescription Module for actionable field recommendations.

6. Conclusion

This work integrates automated image collecting, near-real-time inference, and rigorous evaluation to offer a deployable drone-assisted system for plant disease detection and treatment suggestions. It creates a realistic zero-shot baseline using perceptual hashing to strictly segregate training and testing data (using uncurated PlantDoc photos) by revealing the domain gap between lab and field photographs. Hash-based dataset merging, phased training, and ensemble approaches are used to address cross-domain resilience. Evaluation avoids too optimistic previous approaches by emphasizing practical performance with accuracy, macro-F1, macro-AUC, and calibration metrics. Benchmarks for GPU latency verify viability, and a Streamlit interface facilitates both live drone-stream inference and offline analysis. The system exhibits great potential for precision agriculture, laying the foundation for upcoming advancements in domain adaptation, stabilization, and edge deployment despite issues with drone instability affecting image quality.

The lack of healthy field samples was found to be a major drawback, requiring a mixed training strategy. The study effectively evaluates the difficulties of real-world deployment, despite the fact that this affected the F1 score. To close this domain gap, future research will concentrate on gathering a balanced, localized dataset.

References

Dilmurat., Estimating Crop Seed Composition Using Machine Learning from Multisensory UAV Data, *Remote Sens,* vol.14, no.19, pp. 1-29, ,25, 2022, doi: 10.3390/rs14194786.

Cai, P., Detecting Individual Plants Infected with Pine Wilt Disease Using Drones and Satellite Imagery: A Case Study in Xianning, China, *Remote Sens,* vol.15, no.10, pp. 1-17, May 20, 2023, doi: 10.3390/rs15102671.

Chin, R., Plant disease detection using drones in precision agriculture, *Precision Agriculture,* vol.24, pp. 1663–1682, March 28, 2023, doi: 10.1007/s111119-023-10014-y.

Bangladesh to introduce drone technology to assess crop losses: ADB, ICRISAT Pressroom, Apr. 18, 2024. Available: https://pressroom.icrisat.org/bangladesh-to-introduce-drone-technology-to-assess-crop-losses-adb.

United News of Bangladesh, "Bangladesh to introduce drone tech to assess crop losses: ADB," New Age, Apr. 17, 2024. Available: https://www.newagebd.net/post/economy/232823/bangladesh-to-introduce-drone-tech-to-assess-crop-losses-adb.

Gohila, M. K., A Hybrid Technique For Plant Disease Identification And Localisation In Real-Time, pp. 1-17, December 27, 2024, doi:10.48550/arXiv.. 2412.19682.

Li, Z., Deep Learning Models Outperform Generalized Machine Learning Models in Predicting Winter Wheat Yield Based on Multispectral Data from Drones, *Drones,* vol. 7, no. 8, pp. 1-25, August 2, 2023, doi: 10.3390/drones7080505.

Albattah, W., Artificial Intelligence-Based Drone System for Multiclass Plant Disease Detection Using an Improved Efficient Convolutional Neural Network, *Plant Science,* vol. 7, pp. 1-16, June 09, 2022, doi: 10.3389 /fpls.2022.808380.

Wilf1, P., An image dataset of cleared, x-rayed, and fossil leaves vetted to plant family for human and machine learning, *Phytokeys,* pp. 93-128, December 5, 2021, doi: 10.3897/phytokeys 187.72350.

Dijk, A. D. J. V., Machine learning in plant science and plant breeding, *iScience,* vol. 7, pp. 1-112, January 22, 2021, doi: 1016/j.isci.2020.101890.

Lu, D. et al., Plant Detection and Counting: Enhancing Precision Agriculture in UAV and General Scenes, *IEEE,* vol. 11, pp. 116196 – 116205, October 18, 2023, doi: 10.1109/ACCESS.2023.3325747.

Koh, J.C.O., Automated Machine Learning for High-Throughput Image-Based Plant Phenotyping, *Remote Sens,* vol. 13, no. 5, February 25, 2021, doi: 10.3390/RS13050858.

Yang, M. D., AUAV Open Dataset of Rice Paddies for Deep Learning Practice, *Remote Sens,* vol. 13, no. 7, April 01, 2021, doi: 10.3390/rs13071358.

Habibi, L.N., Machine Learning Techniques to Predict Soybean Plant Density Using UAV and Satellite-Based Remote Sensing, *Remote Sens,* vol. 13, no. 13, June 29, 2021, doi: 10.3390/rs13132548.

© IEOM Society International

Biographies

Md. Naim Mia graduated from Khulna University of Engineering & Technology (KUET) with a Bachelor of Mechatronics Engineering degree. His areas of interest are mechatronic systems and artificial intelligence, with an emphasis on developing intelligent solutions for practical applications. His goal is to create scalable, socially significant technology in fields including security, agriculture, and the environment.

Anik Biswas earned his bachelor's degree in Mechatronics Engineering from Khulna University of Engineering & Technology (KUET), Khulna. He is currently pursuing his master's degree in the same department at KUET, where he also serves as a Teaching Assistant. He also worked as a Calibration Engineer at 'Quality Calibration Solution Private Limited'. His academic and research interests span mechatronics systems, automation, and advanced engineering technologies.

Md. Shirajul Islam is currently pursuing an MS in Mechatronics Engineering from Khulna University of Engineering & Technology, Khulna. He received his BSc in the same department from the same institution. Now he serves as a Teaching Assistant in the same place where he is pursuing his post-graduation degree, and also serves as adjunct faculty in the Mechanical Engineering Department at Imperial College of Engineering (Affiliated with University of Rajshahi), Khulna. His interest lies in Robotics, Control systems, Automotive systems, Bio-embedded systems, and Image Processing.

© IEOM Society International