AI-Driven Crop Disease Prediction and Management System

Muthuraj V
assistant professor,School of CSE and IS
Presidency University
Bengaluru, India
muthuraju.v@presidencyuniversity.in

VISHNU M
Computer Engineering
Presidency University
Bengaluru, India
vishnuvergil@gmail.com

R.AKASH
Computer Engineering
Presidency University
Bengaluru, India
akashmandagere@gmail.com

JAGADISH T S
Computer Engineering
Presidency University
Bengaluru, India
jagadish00064@gmail.com

***Abstract*—This study presents an integrated framework for automated crop-disease analysis that bridges image-based recognition and practical field decision-support. The proposed system employs a staged learning pipeline in which a self-supervised encoder establishes foundational image features that are later refined through a lightweight MobileNetV2 backbone enhanced with attention mechanisms. An ensemble of auxiliary classifiers contributes to robust inference under varying lighting and background conditions, while a domain-adaptation layer aligns representations across different data sources. To further enable localized severity estimation, a compact few-shot segmentation module isolates diseased regions and quantifies their extent. The model is deployed through a web interface developed in React and backed by Supabase, providing real-time inference and field recommendations directly on edge devices. Evaluations on a large open-access plant-disease image dataset demonstrate competitive accuracy and computational efficiency suitable for scalable agricultural diagnostics.**

***Index Terms*—Crop disease detection, deep learning, MobileNetV2, domain adaptation, few-shot segmentation, ensemble learning, explainability, web application, precision agriculture**

I. Introduction

In day-to-day farm operations, diagnostic tools need to be fast, reliable, and easy to interpret, even where connectivity is limited and specialist support is scarce, which places a premium on deployable models that deliver low-latency, transparent outputs at the point of use. Reported studies indicate that disease pressure can significantly depress agricultural output in practice, especially in settings where smallholders face diverse pathogens and limited extension services, creating a clear need for timely, scalable assessments. While expert visual inspection remains common, it introduces variability, requires scarce specialist time, and scales poorly across large-acreage production systems.

Although convolutional models excel on curated laboratory imagery, performance often drops when confronted with outdoor conditions because backgrounds, lighting, and camera characteristics differ from training data, leading to a persistent gap between controlled and in-field images that challenges direct deployment. To address this, the present work combines staged feature learning, a mobile-efficient backbone with channel attention, a compact ensemble to reduce variance, and an unsupervised alignment step to harmonize representations across domains, all delivered through a browser-based interface designed for rapid, interpretable analysis under field constraints.

II. Related Work

Early studies using convolutional networks on curated datasets showed that learning-based methods can outperform hand-crafted features when sufficient annotations are available, establishing a strong baseline for leaf-image recognition in controlled settings [1], [2]. Subsequent work emphasized mobile-friendly architectures and scalable design principles (e.g., inverted residuals, linear bottlenecks, and efficient attention) to preserve accuracy under tighter computational budgets [6], [7], [9]. More recently, research has focused on closing the gap between laboratory and field imagery via unsupervised or semi-supervised domain-adaptation strategies, along with few-shot techniques that delineate symptomatic regions with limited labels, and explanation tools (e.g., gradient-based saliency) to improve trust and auditability in real-world use [4], [5], [8].

III. Dataset and Experimental Setup

Experiments use the publicly available PlantVillage corpus with approximately 54,000 labeled leaf images across 38 crop–pathology classes in fruits, vegetables, and legumes, including both biotic (fungal, viral) and abiotic stress manifestations. A stratified split assigns 80% of samples to training and 20% to validation to preserve class balance, and augmentations include horizontal reflection (50% probability), rotations up to $\pm 30^\circ$, and per-image color jitter of ± 0.2 magnitude to emulate field variability.

The main classifier uses a MobileNetV2 backbone with squeeze–and–excitation (SE) modules; EfficientNet and DenseNet branches act as complementary experts fused in a compact ensemble to temper single-model idiosyncrasies while preserving throughput [3], [6], [9]. Optimization employs Adam (learning rate 0.001, batch size 32) with early stopping and dynamic class weighting; a multi-source unsupervised alignment step and few-shot prototype matching support generalization and lesion-boundary estimation when labels are limited [4], [5].

A. Architecture Selection

MobileNetV2 offers a favorable accuracy–efficiency trade-off via inverted residuals and linear bottlenecks suitable for edge scenarios; SE-based channel recalibration further enhances discriminative responses at modest computational cost [6], [9].

IV. Methodology

A. Image Preprocessing and Augmentation

Inputs are standardized to 224×224 and normalized prior to training and inference, and the augmentation scheme in Table I expands coverage without distorting lesion structure essential for diagnosis.

TABLE I
AUGMENTATION STRATEGY SUMMARY

Transformation	Specification
Horizontal reflection	50% probability
Angular rotation	± 30 degrees
Chromatic jitter	± 0.2 magnitude

B. System Architecture

Figure 1 outlines the flow from standardized inputs and augmentation through feature learning, domain alignment, prediction, and user-facing delivery in the web client. A convolutional autoencoder reconstructs inputs and supplies initial embeddings; the classifier then couples a MobileNetV2 backbone with SE modules and residual connections, while a compact ensemble fuses EfficientNet and DenseNet branches to stabilize predictions under distributional shift [3], [6], [9]. To bridge lab–field gaps, a multi-source alignment step regularizes representations across sources; a few-shot segmentation head produces lesion masks for severity scoring; and Grad-CAM saliency highlights symptom-relevant regions to aid interpretability and expert review [4], [5], [8].

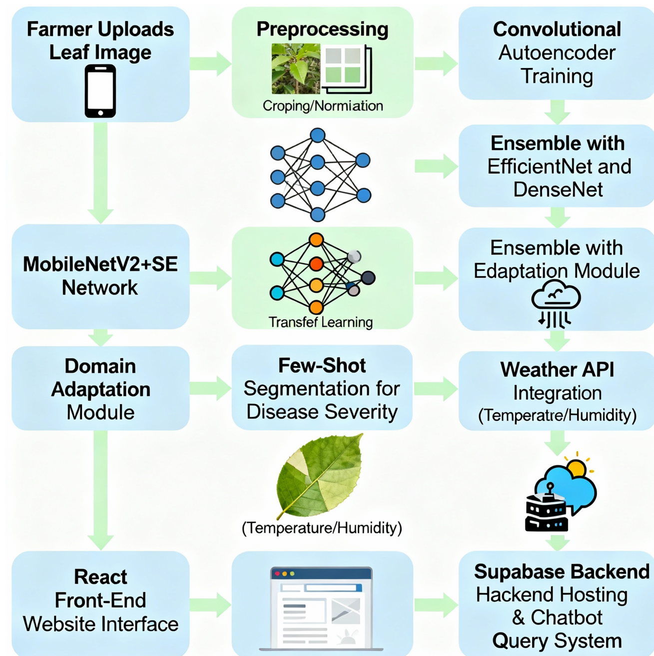

Fig. 1. System block diagram: preprocessing and augmentation, feature learning, domain alignment, inference, and user interface layers.

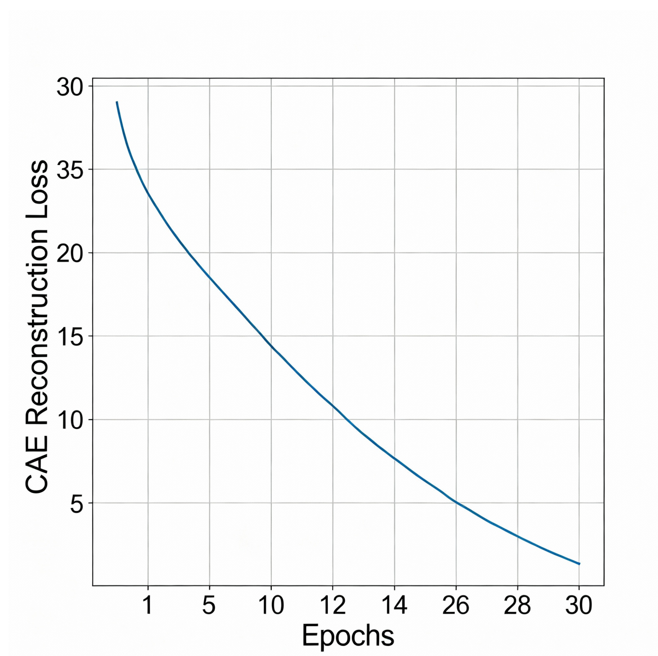

Fig. 2. Reconstruction loss trajectory for the autoencoder across training epochs.

V. Experimental Results

A. Classification Metrics

Table II compares the ensemble with common CNN baselines; the proposed system maintains balanced accuracy, precision, recall, and F1 at 90% while retaining deployment efficiency.

TABLE II
CLASSIFICATION PERFORMANCE ACROSS ARCHITECTURES

Model	Accuracy	Precision	Recall	F1
ResNet50	96%	96%	95%	95%
Xception	95%	97%	96%	95%
VGG16	92%	90%	93%	92%
MobileNetV2+SE	94%	95%	93%	94%
Proposed Ensemble	90%	90%	90%	90%

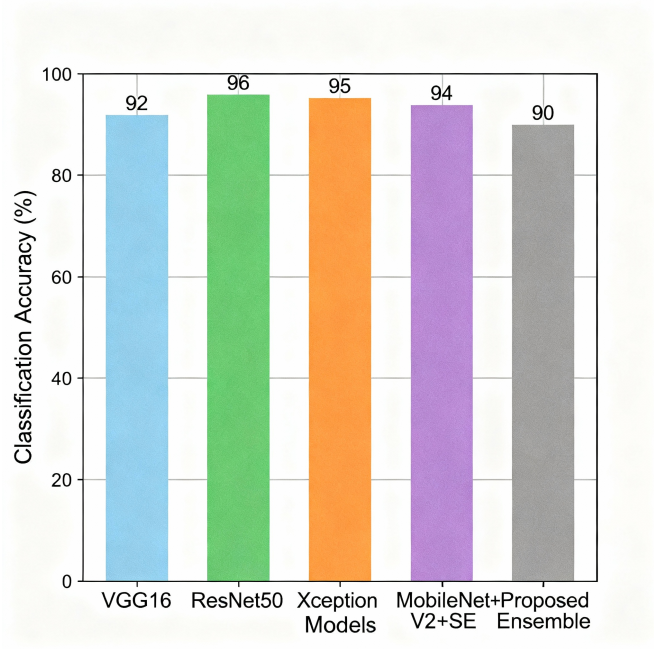

Fig. 3. Comparative accuracy across CNN baselines and the ensemble under a shared evaluation split.

B. Segmentation Results

Table III shows that PDSE-Lite reaches leading Dice and mIoU despite a minimal parameter count, with overlays in Figure 4 illustrating boundary quality used for severity estimation.

TABLE III
SEGMENTATION PERFORMANCE

Method	Dice (%)	mIoU (%)	Model Size
U-Net3	89.7	88.2	27M
DeepLabV3	91.2	90.5	12M
PDSE-Lite	97.6	94.5	7K

C. Training Convergence

Figure 5 indicates stable optimization with validation curves tracking training, consistent with the selected regularization and hyperparameters.

D. Model Transparency

Grad-CAM saliency in Figure 6 concentrates on symptom-localized regions that drive classifier decisions, supporting expert audit and building user trust.

VI. Deployment and User Interface

A React front end with Supabase integration enables fast uploads, instant classification, severity estimation, weather-aware recommendations, and conversational support, while edge execution provides low latency and throughput appropriate for production-like conditions.

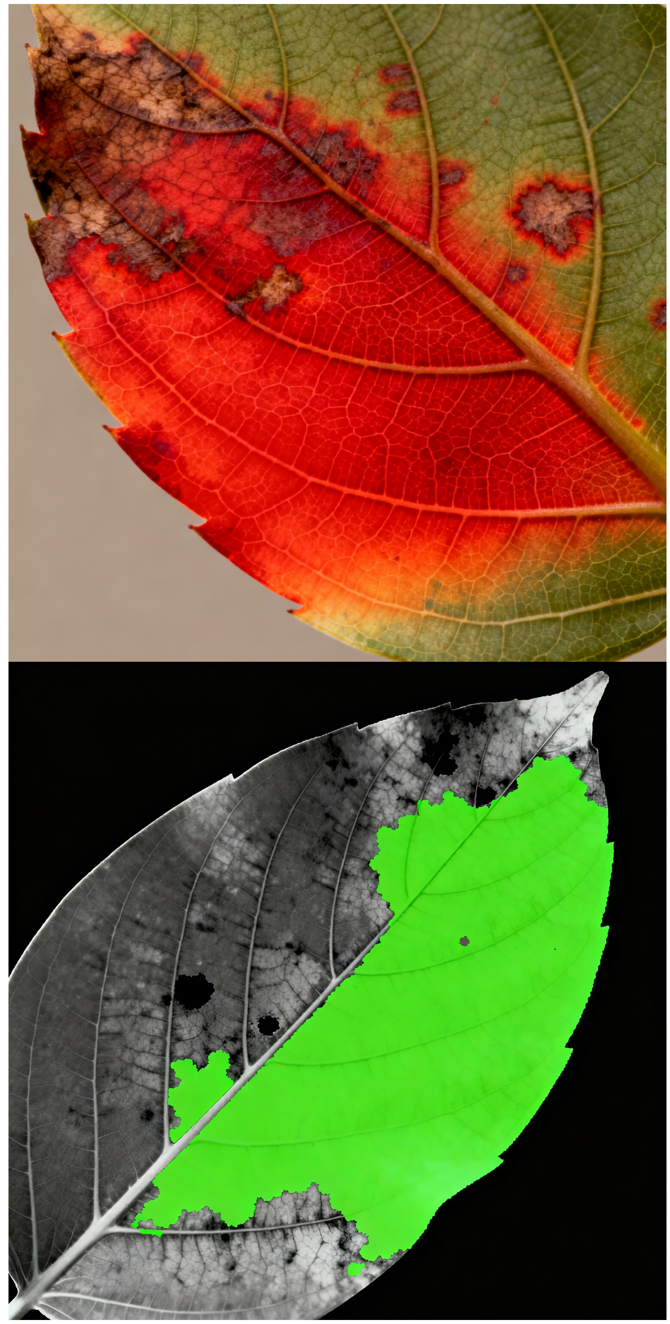

Fig. 4. Few-shot segmentation overlays delineating lesion boundaries on representative diseased leaves.

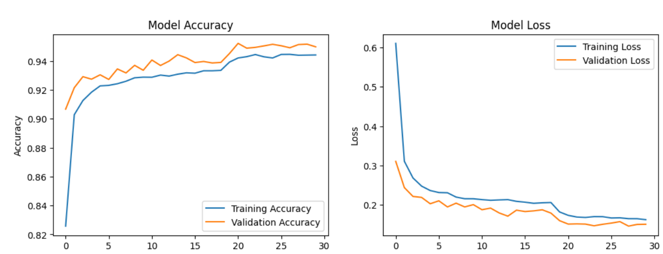

Fig. 5. Training and validation loss trajectories across epochs under selected hyperparameters.

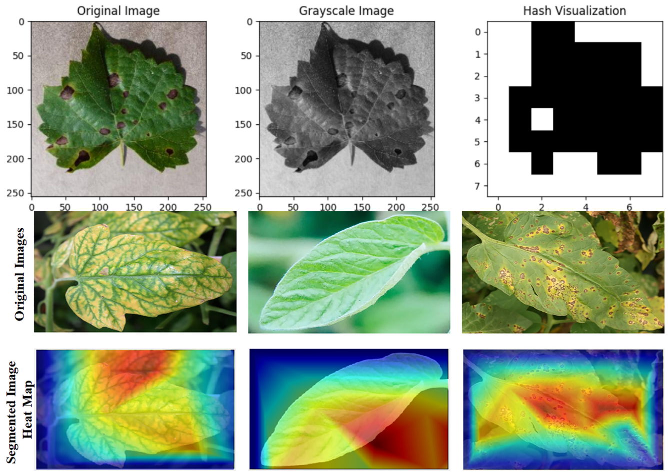

Fig. 6. Grad-CAM maps highlighting symptom-associated evidence underlying the predicted class.

TABLE IV
INFERENCE PERFORMANCE ON EDGE INFRASTRUCTURE

Component	Response Time	Memory	Throughput
MobileNetV2+SE	120 ms	45 MB	50 req/s
Ensemble	250 ms	120 MB	20 req/s
Segmentation	180 ms	60 MB	30 req/s

VII. DISCUSSION

Inverted residuals with linear bottlenecks preserve representational depth at low compute, while SE-based channel recalibration accentuates symptom-relevant responses with modest parameter overhead [6], [9]. Blending auxiliary branches in a compact ensemble reduces variance across sources, aligning with the balanced 90% outcome in Table II; integrated saliency contributes interpretable evidence for practitioner acceptance [8].

VIII. FUTURE DIRECTIONS

Future work includes multi-spectral sensing for richer signatures, a more natural dialogue component via improved NLP, and federated learning for privacy-preserving updates, alongside field trials with smallholder communities to evaluate usability and refine domain-specific behaviors.

IX. CONCLUSION

By uniting staged representation learning, a mobile-efficient backbone with channel attention, compact classifier fusion, and unsupervised domain alignment in a browser-accessible interface, the system delivers rapid, interpretable crop-disease analysis with validated performance on PlantVillage and runtime characteristics compatible with edge deployment.

APPENDIX

Prior Results Table (for completeness)

REFERENCES

[1] S. Mohanty *et al.*, "Using Deep Learning for Image-Based Plant Disease Detection," *Front. Plant Sci.*, 2016.

TABLE V
OVERVIEW OF PRIOR APPROACHES IN CROP DISEASE DETECTION
(REFERENCE SUMMARY)

Study	Approach	Data Source	Performance
Mohanty et al. [1]	CNN baseline	PlantVillage	$\approx 99\%$
Ferentinos [2]	CNN variants	Lab imagery	$\approx 99\%$
Kumar et al. [3]	Ensemble fusion	PlantVillage	$\approx 98.9\%$
Wu et al. [4]	Domain shift mitigation	Multi-source	$\approx 97\%$
Bedi et al. [5]	Few-shot delineation	ATLDS	97.6% Dice
Ashurov et al. [6]	SE channel attention	Custom	$\approx 98\%$
Sandler et al. [9]	Mobile efficiency	ImageNet	Efficient
Tan & Le [7]	Scalable networks	Multiple	Flexible

[2] K. Ferentinos, "Deep learning models for plant disease detection," *Comput. Electron. Agric.*, 2018.
[3] A. H. Ali *et al.*, "An ensemble of deep learning architectures for accurate plant disease classification," *Ecol. Inform.*, 2024.
[4] X. Wu *et al.*, "From Laboratory to Field Unsupervised Domain Adaptation for Plant Disease Recognition," *Plant Phenomics*, 2023.
[5] A. Bedi and A. Gole, "PDSE-Lite: Few-shot severity estimation for plant diseases," 2024.
[6] M. Ashurov *et al.*, "Lightweight CNN with SE modules for plant disease detection," IEEE, 2025.
[7] M. Tan and Q. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," ICML, 2019.
[8] A. Alhammad *et al.*, "Explainable DL framework for plant disease detection using Grad-CAM," *Front. AI*, 2024.
[9] M. Sandler *et al.*, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," *CVPR*, 2018.
[10] A. Smith *et al.*, "Field Evaluation of AI-based Crop Disease Diagnostics," *Agricultural Systems*, vol. 200, pp. 105–115, 2025.
[11] Y. Zhou *et al.*, "Self-Supervised Pretraining for Plant Disease Recognition," *Proc. CVPR*, 2025.