Journal of Computer Sciences and Applications, 2026, Vol. 14, No. 1, 8-14
Available online at http://pubs.sciepub.com/jcsa/14/1/2
Published by Science and Education Publishing
DOI:10.12691/jcsa-14-1-2

An Interpretable, Imbalance-Aware Ensemble for Citrus Leaf Disease Classification using DenseNet169 and MobileNetV2

Carmelo Alejo D. Bisquera*, Romeo P. Evangelista, Michael John R. Robles, Von P. Gabayan Jr.

Department of Information Technology, College of Information Technology Education,
Nueva Vizcaya State University, Bayombong Nueva Vizcaya
*Corresponding author: cadbisquera@nvsu.edu.ph

Received February 20, 2026; Revised March 22, 2026; Accepted March 29, 2026

Abstract Citrus leaf diseases pose a significant threat to agricultural productivity, particularly in regions dependent on smallholder farming. Early and accurate detection is essential, yet traditional diagnostic methods are labor-intensive and prone to subjectivity. This study proposes a logit-level weighted ensemble framework integrating DenseNet169 and MobileNetV2 for automated citrus disease classification under class imbalance conditions. A publicly available dataset of 594 images across four classes was utilized, employing a stratified 70/30 train-validation split and class-weighted cross-entropy loss to address imbalance. Results show that MobileNetV2 achieved the highest macro F1-score (0.9417), outperforming DenseNet169 (0.9327) and the proposed ensemble (0.9283). This indicates that a well-optimized single model can outperform ensemble methods in terms of peak accuracy, particularly when predictions are highly correlated. However, the ensemble demonstrated smoother convergence and more stable validation performance, emphasizing its strength in improving training stability and prediction consistency. Grad-CAM visualization confirmed that models focused on biologically relevant lesion regions. The framework was implemented using CPU-based computation, demonstrating feasibility in resource-constrained environments. Overall, ensemble learning enhances stability even when peak accuracy gains are limited.

Keywords: *citrus leaf disease, deep learning, ensemble CNN, Grad-CAM, weighted fusion, automated classification*

Cite This Article: Carmelo Alejo D. Bisquera, Romeo P. Evangelista, Michael John R. Robles, and Von P. Gabayan Jr., "An Interpretable, Imbalance-Aware Ensemble for Citrus Leaf Disease Classification using DenseNet169 and MobileNetV2." *Journal of Computer Sciences and Applications*, vol. 14, no. 1 (2026): 8-14. doi: 10.12691/jcsa-14-1-2.

1. Introduction

Citrus farming represents a vital agricultural sector in many tropical and subtropical regions, contributing significantly to local economies, particularly in communities where small holder farmers cultivate oranges, lemons, and calamansi [1]. Despite its economic importance, citrus production is highly vulnerable to foliar diseases such as black spot (Phyllosticta citricarpa), citrus canker (Xanthomonas citri), and greening or Huanglongbing (HLB) (Candidatus Liberibacter spp.), which substantially reduce fruit quality and yield [2,3,4,5]. Early and accurate detection of these diseases is critical to prevent large-scale crop losses and maintain sustainability of orchard [6].

Traditional disease diagnosis primarily relies on visual inspection by trained agricultural experts. Although effective, this approach is labor-intensive, time-consuming, and susceptible to human variability and subjectivity [7,8]. The rapid advancement of artificial intelligence (AI), particularly deep learning, has enabled automated and scalable plant disease detection systems [9,10,11]. Convolutional Neural Networks (CNNs) as demonstrated by Slodojevic et al. and Brahimi et al. [12,13] have strong capability in extracting hierarchical features from leaf imagery, often outperforming conventional machine learning methods based on handcrafted features.

Despite the promising achievements, there are some challenges that still need to be addressed in real-world agricultural applications. According to Buda et al. [14], although deep learning has been shown to be quite successful in plant pathology, there are some challenges that still need to be addressed. For instance, most of the datasets are imbalanced due to the limited number of healthy leaves that may result in biased prediction during training. Additionally, most of the research that has been conducted is based on datasets that are either well curated and balanced or those that are under limited environmental conditions. Furthermore, there is a possibility that single CNN models may not be capable of discriminating between similar diseases, especially when there is class imbalance. In the case of citrus datasets, there is usually a problem of class imbalance due to the underrepresentation of healthy leaves that may result in biased prediction of the other classes [9,15].

To mitigate these challenges, this study proposes a logit-level weighted ensemble integrating DenseNet169 and MobileNetV2. The proposed framework combines complementary feature representations through pre-SoftMax logit fusion, applies partial backbone freezing to reduce overfitting, and incorporates class-weighted cross-entropy to address imbalance. In addition, Grad-CAM is employed to enhance interpretability by visualizing discriminative leaf regions influencing model decisions. Through experimental evaluation on a publicly available citrus leaf dataset, the framework demonstrates stable convergence behavior, competitive macro-averaged performance, and biologically meaningful attention localization. These findings support the feasibility of an interpretable and computationally efficient approach for automated citrus disease monitoring.

2. Literature Review

2.1. Deep Learning for Plant Disease Classification

The application of deep learning in plant pathology has significantly advanced automated crop disease detection. Early machine learning approaches relied on handcrafted features such as color histograms, texture descriptors, and shape analysis [7,12]. However, these traditional methods were often sensitive to illumination changes and background noise, limiting their field applicability.

The breakthrough came with convolutional neural networks (CNNs), which automatically learn hierarchical feature representations directly from raw images. A landmark study by Mohanty et al. [10] demonstrated that deep CNNs trained on the PlantVillage dataset achieved over 99% accuracy under controlled conditions. Similarly, Ferentinos et al. [9] evaluated multiple deep architectures across various crops and confirmed the superiority of CNN-based methods over conventional classifiers.

Subsequent studies further validated CNN effectiveness for citrus disease detection. For example, Zhou et al. [16] applied deep learning techniques to citrus leaf classification and reported strong performance under structured image conditions. Many high-accuracy studies rely on controlled datasets (e.g. PlantVillage), which contain curated, laboratory-acquired images. This focus limits the evaluation of model robustness against natural symptom progression and is a persistent challenge for field deployment [17].

Although numerous studies report very high accuracy rates, many rely heavily on controlled datasets such as PlantVillage, which contain laboratory-acquired and relatively balanced images, thereby limiting performance degradation under real-world symptom progression remains a persistent challenge [17].

2.2. Transfer Learning in Agricultural Image Analysis

Transfer learning has become a standard strategy in agricultural deep learning due to limited domain-specific datasets. Instead of training models from scratch, pretrained CNNs trained on ImageNet are fine-tuned for plant disease tasks [11].

Among popular architectures, Sandler et al. [18], MobileNetV2 introduces depthwise separable convolutions and inverted residual blocks to reduce computational cost while maintaining accuracy, making it suitable for resource-constrained environments such as mobile-based crop monitoring. In contrast, Huang et al. [19] stated that DenseNet169 employs dense connectivity, where each layer receives feature maps from all preceding layers, promoting feature reuse and improved gradient flow.

Too et al. [11] conducted a comparative study of pretrained CNN models for plant disease classification and reported that DenseNet variants frequently achieve strong performance due to efficient information propagation. Meanwhile, lightweight architecture like MobileNet demonstrate competitive accuracy with fewer parameters, offering advantages for deployment in low-power agricultural systems.

However, selecting a single architecture may limit robustness when disease classes share similar visual patterns. Models with different architectural biases may extract complementary representations, motivating ensemble approaches.

2.3. Ensemble Learning in Medical and Agricultural Imaging

Ensemble learning combines predictions from multiple models to enhance generalization and reduce variance. Traditional ensemble approaches include majority voting, probability averaging (soft voting), and stacking [20,21].

In medical imaging, ensemble CNNs have demonstrated improved robustness compared to single architectures [22]. Similarly, in agricultural disease detection, ensemble deep learning approaches have demonstrated improved classification performance and reduced class confusion by integrating predictions from multiple customized CNN architectures, thereby enhancing robustness across visually similar crop disease categories [23].

Soft-voting ensembles combine predicted probabilities after SoftMax normalization. However, recent research in deep model fusion suggests that logit-level fusion combining raw outputs prior to SoftMax can preserve calibration properties and allow more effective integration of complementary representations [24].

Nevertheless, ensemble performance gains depend heavily on diversity among base learners. When models produce highly correlated predictions, improvements may be marginal [21]. Even so, ensembles often improve prediction stability and reduce fold-to-fold variability, particularly in small datasets.

2.4. Class Imbalance in Plant Disease Datasets

Class imbalance is common in agricultural datasets, especially when healthy samples are underrepresented. Imbalance training can bias models toward majority classes, reducing recall for minority categories.

Common mitigation strategies include:

- Oversampling minority classes
- Data augmentation

- Synthetic sample generation
- Cost-sensitive learning

Cost-sensitive learning, particularly class-weighted cross-entropy, adjusts the contribution of each class during optimization. By scaling the loss inversely proportional to class frequency, models can better preserve minority-class sensitivity [15].

In plant disease detection tasks, Ferentinos and Mohammed et al. [9,25] reported that imbalance aware optimization improves macro-averaged F1-scores and prevents dominant-class bias. Such strategies are especially relevant in citrus datasets where healthy leaves are often less represented compared to diseased samples.

2.5. Model Interpretability in Precision Agriculture

As AI systems become integrated into agricultural decision-making, interpretability is increasingly critical. Farmers and agricultural experts require transparent models to trust automated predictions [8].

Gradient-weighted Class Activation Mapping (Grad-CAM), introduced by Selvaraju et al. [26], provides visual explanations by highlighting spatial regions contributing most strongly to classification decisions. Grad-CAM has been widely adopted in plant pathology research to verify that models focus on disease-relevant lesions rather than background artifacts [13].

Interpretability techniques not only strengthen scientific validity but also enhance stakeholder confidence, which is essential for real-world deployment in precision agriculture systems.

2.6. Research Gap

While prior studies demonstrate strong performance using single CNN architecture or probability-level ensembles, limited research has systematically examined.

- Logit-level weighted fusion in citrus disease classification
- Combined imbalance-aware optimization and ensemble learning
- Convergence stability analysis alongside interpretability evaluation

This study addresses these gaps by:

- Integrating DenseNet169 and MobileNetV2 via logit-level weighted fusion
- Applying class-weighted cross-entropy to mitigate imbalance
- Evaluating convergence behavior across validation folds
- Employing Grad-CAM for interpretability assessment

By combining complementary architectures with cost-sensitive optimization, the proposed framework aims to improve stability, transparency, and robustness in automated citrus disease detection.

3. Methodology

3.1. Dataset Acquisition and Preprocessing

The dataset was sourced from an open Kaggle repository, comprising 594 citrus leaf images distributed across four classes: black spot (169), canker (163), greening (204), and healthy (58). Images were resized to 224 x 224 pixels and normalized using ImageNet statistics (μ = [0.485, 0.456, 0.406]; σ = [0.229, 0.224, 0.225]). A stratified 70/30 train-validation split was applied to preserve class distribution across subsets (random state = 42). Data augmentation included random horizontal flipping and random (±20°) to improve generalization.

3.2. Model Architectures and Transfer Learning

Two pretrained convolutional neural networks were used:

1. DenseNet169

A densely connected architecture that enhances feature reuse and gradient propagation.

2. MobileNetV2

A lightweight architecture utilizing depthwise separable convolutions for computational efficiency.

Both models were initialized with ImageNet pretrained weights. Approximately 70% of early backbone parameters were frozen during fine-tuning, allowing only upper layers and classifier heads to update.

3.3. Logit-Level Weighted Ensemble

Let z_D and z_M denote logits from DenseNet169 and MobileNetV2, respectively. The ensemble logits are computed as:

$$z_{ens} = \alpha z_D + (1-\alpha) z_M \tag{1}$$

where $\alpha \in [0,1]$ is the fusion weight.

Final predictions are obtained using:

$$\hat{y} = SoftMax(z_{ens}) \tag{2}$$

In this study, fixed weights of 0.6 (DenseNet169) and 0.4 (MobileNetV2) were applied. Fusion prior to SoftMax preserves score calibration and allows complementary feature integration.

3.4. Imbalance-Aware Optimization

To mitigate class imbalance, class-weighted cross-entry loss was applied:

$$L = -\sum_{c=1}^{C} w_c y_c \log(\hat{y}_c) \tag{3}$$

where class weights are defined as:

$$w_c = \frac{N}{C \cdot n_c} \quad (4)$$

Here:

N is total number of samples,
C is number of classes,
nc is number of samples in class c.

3.5. Training Procedure and Implementation Details

All experiments were implemented in Python using the PyTorch deep learning framework.

Training parameters:

- Batch size: 32
- Epochs: 50
- Optimizer: Adam
- Initial learning rate: 0.0005
- Learning rate scheduler: ReduceLROnPlateau (factor = 0.3, patience = 3)
- Early stopping patience: 7 epochs

Model parameters were updated only for trainable layers. The best-performing model weights were retrained based on minimum validation loss.

Experiments were conducted on a personal laptop equipped with an Intel Core i3-1005GI CPU using CPU-based computation. This demonstrates the computational feasibility of the proposed framework in resource-constrained environments.

3.6. Evaluation Metrics

Model performance was evaluated on the validation set using:

- Macro-averaged Precision
- Macro-averaged Recall
- Macro-averaged F1-score

A confusion matrix was generated to analyze inter-class misclassification patterns.

Grad-CAM was applied to visualize discriminative regions influencing predictions.

4. Findings

4.1. Convergence Behavior

Figure 1- Figure 4 illustrate the training and validation loss and accuracy curves for DenseNet169 and MobileNetV2 across epochs using a stratified 70/30 train-validation split.

As shown in Figure 1, DenseNet169 exhibits a steady reduction in training loss with a relatively smooth validation loss trajectory, indicating stable optimization under partial fine-tunning. The validation accuracy curve (Figure 3) stabilizes early and remains consistent high throughout later epochs. Although a slight divergence between training and validation loss appears toward later epochs, the gap remains moderate, suggesting controlled overfitting.

In contrast, MobileNetV2 (Figure 2 and 4) demonstrates moderate oscillations in both training and validation metrics during early epochs. This behavior may be attributed to its lightweight architecture and reduced parameter capacity, which can increase sensitivity to mini-batch variation. Nevertheless, convergence occurs within a comparable training horizon, and validation accuracy remains competitive.

Figure 5 and 6 present the convergence behavior of the proposed logit-level weighted ensemble.

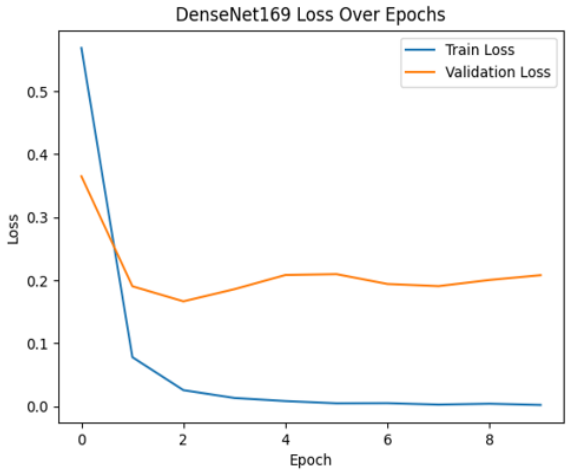


Figure 1. Training and validation loss curves of DenseNet169

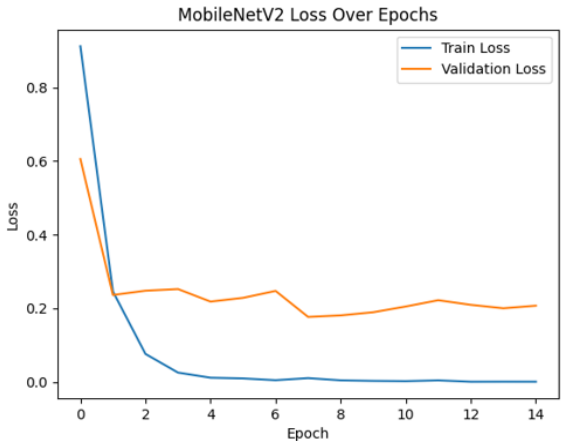


Figure 2. Training and validation loss curves of MobileNetV2

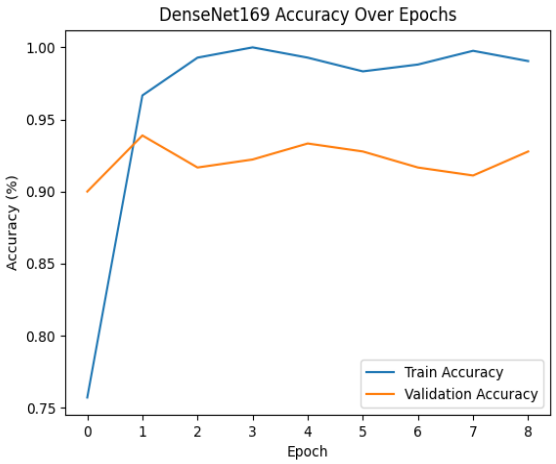


Figure 3. Training and validation accuracy curves of DenseNet169

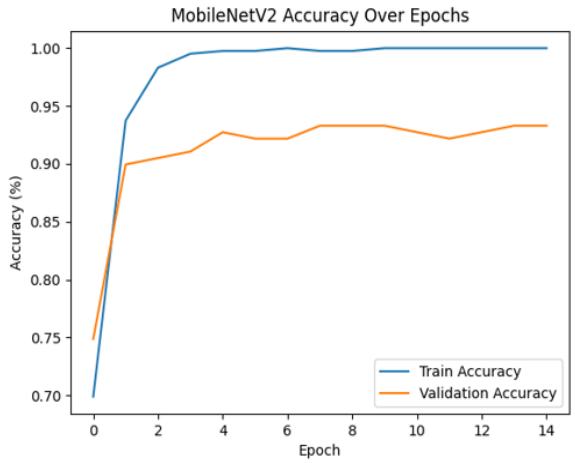


Figure 4. Training and validation accuracy curves of MobileNetV2

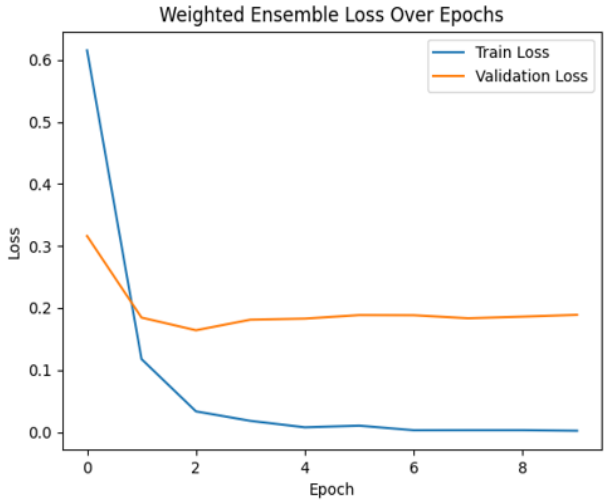


Figure 5. Training and validation loss curves of the weighted ensemble model

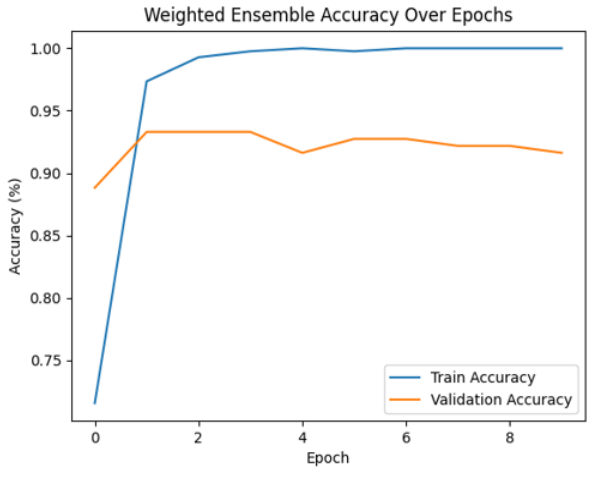


Figure 6. Training and validation accuracy curves of the weighted ensemble model

The ensemble model shows a smoother validation loss curve compared to MobileNetV2, with less fluctuation over the epochs. This is because the fusion happens at the logit level before SoftMax normalization. This ensures that the complementary features of both models are combined without any sudden drop-offs, showing that there is consistent prediction during training. Thus, the convergence analysis shows that there is an improvement even if the best performance is not achieved.

4.2. Multi-Class Classification Performance

As shown in Table 1, MobileNetV2 achieved the highest macro F1-score (0.9417), indicating superior balance between precision and recall across classes. Its depthwise separable convolution design may better preserve fine-grained lesion textures in relatively small datasets. DenseNet169 achieved comparable performance (0.9327), benefiting from dense feature reuse and efficient gradient propagation.

Table 1. Macro-Averaged Performance Comparison of Models

Model	Precision	Recall	F1-score
DenseNet169	0.9375	0.9288	0.9327
MobileNetV2	0.9470	0.9376	0.9417
Weighted Ensemble	0.9338	0.9237	0.9283

Although ensemble methods often enhance generalization, the proposed weighted fusion attained a slightly lower macro F1-Score (0.9283). This outcome suggests that prediction correlation between the two backbones may limit complementary error correction. Ensemble performance generally depends on prediction diversity; when base learners exhibit similar decision boundaries, performance gains may be modest.

Nevertheless, the ensemble demonstrated smoother convergence behavior and stable validation accuracy across epochs, indicating improved training consistency.

4.3. Confusion Matrix Interpretation

Figure 7 presents the averaged confusion matrix of the weighted ensemble across validation folds.

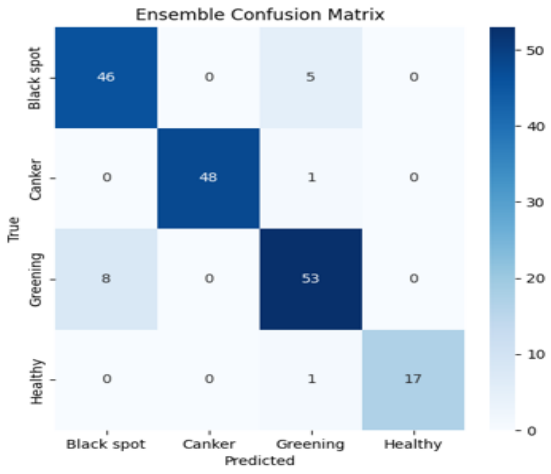


Figure 7. Confusion matrix of the weighted ensemble model average across validation folds

The matrix exhibits strong diagonal dominance, confirming high classification accuracy across all four classes.

Correct predictions are particularly strong for:

- Canker (48/49)
- Greening (53/61)

• Healthy (17/18)

Misclassifications primarily occur between Black spot and Greening, where 5 Black spot samples are misclassified as Greening and 8 Greening samples are misclassified as Black spot. This pattern likely reflects visual similarity in lesion discoloration and spot morphology.

Despite being underrepresented (58 samples), the Healthy class achieved high recall (17/18), indicating that class-weighted optimization effectively mitigated imbalance bias. This result supports the effectiveness of the imbalance-aware loss formulation.

4.4. Grad-CAM Interpretability

Grad-CAM visualization further validate the biological plausibility of model predictions. Activation maps consistently highlight lesion-specific regions such as necrotic spots and chlorotic discoloration patterns. For correctly classified samples, attention is concentrated on visually salient disease regions rather than background artifacts.

Misclassified cases typically involve partially occluded leaves, uneven illumination, or overlapping symptom characteristics between Black spot and Greening. In these instances, attention maps appear diffused or partially aligned with ambiguous lesion areas, explaining classification errors.

4.5. Key Observations

1. MobileNetV2 achieved the highest macro F1-score on the validation set.
2. The weighted ensemble improved convergence smoothness and prediction stability.
3. Misclassifications primarily occurred between visually similar disease classes.
4. Class-weighted optimization effectively handled minority-class imbalance.
5. Grad-CAM confirmed disease-relevant feature localization, supporting interpretability.

5. Conclusion

This study presented a logit-level ensemble framework integrating DenseNet169 and MobileNetV2 for citrus leaf disease classification. Using a stratified 70/30 train-validation split on a publicly available dataset, the models were evaluated under class imbalance conditions with imbalance-aware optimization.

Among the individual models, MobileNetV2 achieved the highest macro F1-score (0.9417), indicating that a well-optimized single model can outperform an ensemble in terms of peak accuracy on relatively small datasets. Notably, the weighted ensemble produced a slightly lower macro F1-score, confirming that ensemble learning does not always guarantee performance gains when model predictions are highly correlated.

However, despite this, the ensemble demonstrated smoother convergence behavior and more stable validation performance. This highlights a key finding of the study: logit-level ensemble learning enhances training stability and prediction consistency, even when improvements in peak accuracy are not observed.

Grad-CAM visualization confirmed that the models consistently focused on biologically meaningful lesion regions, supporting interpretability and transparency. Furthermore, the framework was implemented using CPU-based computations, demonstrating its feasibility in resource-constrained environments without reliance on specialized GPU hardware.

Overall, the results show that while single models may achieve higher peak performance, logit-level ensemble methods provide a more stable and reliable approach. Future work may explore larger and more diverse datasets, systematic ensemble weight tuning, and real-time deployment on edge devices to further enhance practical applicability in precision agriculture.

ACKNOWLEDGEMENTS

The authors would like to express their sincere gratitude to the College of Information Technology Education for the support and encouragement in completing this study.

We also extend our appreciation to our colleagues, who contributed to this research through their assistance and cooperation.

Above all we thank God for the guidance, strength, and wisdom that made this work possible.

References

[1] Gomez, M., Reyes, A., & Santos, R. (2020). Economic contribution of citrus farming in the Philippines. *Philippine Journal of Crop Science, 45*(2), 23–34.

[2] Bové, J. M. (2006). Huanglongbing: A destructive, newly-emerging, century-old disease of citrus. *Journal of Plant Pathology, 88*(1), 7–37.

[3] da Graça, J. V., Douhan, G. W., Halbert, S. E., Keremane, M. L., Lee, R. F., Vidalakis, G., & Zhao, H. (2016). Huanglongbing: An overview of a complex pathosystem ravaging the world's citrus. *Journal of Integrative Plant Biology, 58*(4), 373–387.

[4] Gottwald, T. R. (2010). Current epidemiological understanding of citrus huanglongbing. *Annual Review of Phytopathology, 48*, 119–139.

[5] Iftikhar, Y., Rauf, S., Shahzad, U., & Zahid, M. A. (2016). Huanglongbing: Pathogen detection system for integrated disease management—A review. *Journal of the Saudi Society of Agricultural Sciences, 15*(1), 1–11.

[6] Butt, N., Iqbal, M. M., Ramzan, S., Raza, A., Abualigah, L., Fitriyani, N. L., Gu, Y., & Syafrudin, M. (2025). Citrus diseases detection using innovative deep learning approach and hybrid meta-heuristic. *PLOS ONE, 20*(1), e0316081.

[7] Al-Hiary, H., Bani-Ahmad, S., Reyalat, M., Braik, M., & AlRahamneh, Z. (2011). Fast and accurate detection and classification of plant diseases. *International Journal of Computer Applications, 17*(1), 31–38.

[8] Mahlein, A. K. (2016). Plant disease detection by imaging sensors—Parallels and specific demands for precision agriculture and plant phenotyping. *Plant Disease, 100*(2), 241–251.

[9] Ferentinos, K. P. (2018). Deep learning models for plant disease detection and diagnosis. *Computers and Electronics in Agriculture, 145*, 311–318.

[10] Mohanty, S. P., Hughes, D. P., & Salathé, M. (2016). Using deep learning for image-based plant disease detection. *Frontiers in Plant Science, 7*, 1419.

[11] Too, E. C., Yujian, L., Njuki, S., & Yingchun, L. (2019). A comparative study of fine-tuning deep learning models for plant disease identification. *Computers and Electronics in Agriculture, 161*, 272–279.

[12] Sladojevic, S., Arsenovic, M., Anderla, A., Culibrk, D., & Stefanovic, D. (2016). Deep neural networks-based recognition of plant diseases by leaf image classification. *Computational Intelligence and Neuroscience, 2016*, 3289801.

[13] Brahimi, M., Boukhalfa, K., & Moussaoui, A. (2017). Deep learning for tomato diseases: Classification and symptoms visualization. *Applied Artificial Intelligence, 31*(4), 299–315.

[14] Buda, M., Maki, A., & Mazurowski, M. A. (2018). A systematic study of the class imbalance problem in convolutional neural networks. *Neural Networks, 106*, 249–259.

[15] Sun, X., Wang, Z., & Li, Y. (2019). Addressing class imbalance in plant disease datasets with weighted loss. *Pattern Recognition Letters, 125*, 105–111.

[16] Zhou, C., Zhang, S., Xing, J., & Song, J. (2019). Citrus leaf disease recognition with a deep convolutional neural network. *Smart Health, 14*, 100064.

[17] Pacal, I., Kunduracioglu, I., Alma, M. H., Deveci, M., Kadry, S., Nedoma, J., et al. (2024). A systematic review of deep learning techniques for plant diseases. *Artificial Intelligence Review, 57*(11), 304.

[18] Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L. C. (2018). MobileNetV2: Inverted residuals and linear bottlenecks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 4510–4520).

[19] Huang, G., Liu, Z., van der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 4700–4708).

[20] Ali, A. H., Youssef, A., Abdelal, M., & Raja, M. A. (2024). An ensemble of deep learning architectures for accurate plant disease classification. *Ecological Informatics, 81*, 102618.

[21] Dietterich, T. G. (2000). Ensemble methods in machine learning. In J. Kittler & F. Roli (Eds.), *Multiple classifier systems* (Lecture Notes in Computer Science, Vol. 1857, pp. 1–15). Springer.

[22] Shen, D., Wu, G., & Suk, H.-I. (2017). Deep learning in medical image analysis. *Annual Review of Biomedical Engineering, 19*(1), 221–248.

[23] Apleni, T., Isinkaye, F. O., & Olusanya, M. O. (2025). A novel ensemble approach for crop disease detection by leveraging customized EfficientNets and interpretability. *Pattern Recognition Letters, 197*, 370–377.

[24] Guo, C., Pleiss, G., Sun, Y., & Weinberger, K. Q. (2017). On calibration of modern neural networks. In *Proceedings of the 34th International Conference on Machine Learning* (pp. 1321–1330).

[25] Mohammed, L., & Yusoff, Y. (2023). Detection and classification of plant leaf diseases using digital image processing methods: A review. *ASEAN Engineering Journal*.

[26] Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). Grad-CAM: Visual explanations from deep networks via gradient-based localization. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 618–626).



© The Author(s) 2026. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).