Mitigating Accuracy Loss in Plant Disease Detection: A Comparative Study of Multi-Stage Hybrid Classification Frameworks for Field Conditions

Ivan Chad T. Quilondrino

George V. Patan

Joanna Victoria S. Pitao

https://doi.org/10.5109/7395548

出版情報：Proceedings of International Exchange and Innovation Conference on Engineering & Sciences (IEICES). 11, pp.402-407, 2025-10-30. International Exchange and Innovation Conference on Engineering & Sciences
バージョン：

Proceeding of International Exchange and Innovation Conference on Engineering & Sciences (IEICES)

Mitigating Accuracy Loss in Plant Disease Detection: A Comparative Study of Multi-Stage Hybrid Classification Frameworks for Field Conditions

Ivan Chad T. Quilondrino[1], George V. Patan[2] Joanna Victoria S. Pitao[3]
Department of Computer Science, College of Computing and Information Sciences, Caraga State University

Corresponding author email: Ivan.tims.quil@gmail.com

Abstract: *Deep learning models for plant disease detection have achieved over 99% accuracy on lab-controlled datasets like PlantVillage. However, their performance drops significantly in real-world farm settings due to lighting variation, background noise, and subtle disease symptoms. To address this, the researchers proposed a multi-stage hybrid framework composed of three strategies—Hierarchical, Augmented, and Stacking—that integrate handcrafted texture features (GLCM + KNN) with six deep learning models: ResNet-50, EfficientNetV2, ConvNeXt, Swin Transformer, Vision Transformer, and RegNetY-800MF. These were trained and tested on both PlantVillage and PlantDoc datasets. Among all models and strategies, ConvNeXt under the Augmented framework achieved 99.57% accuracy on PlantVillage and 91.47% on PlantDoc, significantly reducing the accuracy gap caused by domain shift. The results demonstrated that hybrid frameworks can substantially improve the generalization of plant disease detection models in complex, real-world agricultural environments.*

Keywords: Plant disease detection; Hybrid framework; GLCM; Deep learning; Field environments

1. INTRODUCTION

In recent years, plant disease detection using deep learning has gained significant attention, particularly through models trained on clean, lab-captured datasets such as PlantVillage. Convolutional Neural Networks (CNNs) and Vision Transformers (ViTs) have achieved over 98% accuracy in such settings [1]. However, this level of performance often fails to carry over to field environments, where inconsistent lighting, cluttered backgrounds, image noise, and overlapping symptoms are common [2]. For example, ResNet-50 drops from over 99% accuracy on PlantVillage to below 30% on field-captured PlantDoc images [3].

This performance degradation is primarily due to overfitting on the clean visual patterns found in lab datasets, which do not generalize well to real field environment conditions. In response, hybrid methods have emerged that combine deep learning with traditional image processing techniques. Among these, Gray-Level Co-occurrence Matrix (GLCM) has been used to extract texture-based features like contrast, entropy, and homogeneity, improving model interpretability and noise resilience. When paired with K-Nearest Neighbor (KNN), GLCM helps filter early-stage healthy and diseased samples efficiently [4].

Ensemble learning approaches have also been explored to improve prediction stability [5], though many lack validation under real-world conditions. This study addresses these challenges by proposing a multi-stage hybrid classification framework that integrates GLCM-KNN with deep learning through three strategies: Hierarchical, Augmented, and Stacking. Each strategy is tested across six deep models: ResNet-50, EfficientNetV2, ConvNeXt, Swin Transformer, Vision Transformer, and RegNetY-800MF and evaluated on both the PlantVillage and PlantDoc datasets, which together simulate both controlled and field conditions.

The goal is to improve the generalization and accuracy of plant disease classification, specifically, tomato and corn crops. Tomatoes are one of the most crucial crop for many farmers not only in the Phipillines but also among the most cultivated crops globally [16]. Aside from this, corn is also one of the essential crops as it ranked second next to rice in the Philippines [17]. The framework targets nine disease classes, namely, gray leaf spot, common rust, northern leaf blight, bacterial spot, early blight, late blight, septoria leaf spot, and yellow leaf curl virus. By evaluating these hybrid models, this study aims to reduce the accuracy loss that hampers real-world deployment of plant disease detection systems.

2. RELATED LITERATURE

Recent research on plant disease detection has shifted from relying solely on deep learning architectures to exploring hybrid frameworks that integrate traditional image processing techniques with deep neural networks. This trend is driven by the need to improve model performance in field environments, which are often plagued by inconsistent lighting, complex backgrounds, and occlusion—factors that reduce the effectiveness of standard CNN-based models trained on ideal datasets such as PlantVillage.

One area of growing interest is the use of Gray-Level Co-occurrence Matrix (GLCM) features as part of a preprocessing step to enhance model interpretability and reduce input noise. Mokhtar et al. successfully applied GLCM and K-Nearest Neighbors (KNN) to classify tomato leaf diseases, reporting 89% accuracy by focusing on spatial texture patterns like contrast and entropy [6]. More recently, Shoaib et al. demonstrated that filtering corn leaf images with GLCM-KNN prior to CNN processing improved the ResNet50 model's field accuracy by 18% [7]. Similarly, Prakash et al. used GLCM descriptors to detect early-stage wheat rust, showing that handcrafted texture features such as homogeneity and correlation are valuable for identifying subtle disease symptoms [8]. These findings validate the use of GLCM-KNN as an efficient early-stage classifier for detecting diseased leaves in noisy field data.

In addition to early-stage filtering, some researchers have explored feature fusion approaches, where traditional GLCM features are combined with deep embeddings rather than used as a standalone filter. Hosny et al. proposed combining CNN features with GLCM texture descriptors to improve lesion localization under field conditions [9]. Their results support the idea that hybrid features—when merged appropriately—can boost the model's sensitivity to subtle and heterogeneous disease symptoms. Waheed et al. and Sharma et al. extended this idea by incorporating dimensionality reduction and regularization strategies to prevent overfitting caused by feature redundancy [10][11]. These works align closely with the Augmented Framework proposed in this study, where GLCM outputs serve as complementary inputs to deep networks for final classification.

The use of stacked ensemble models has also gained popularity in agricultural AI research. Vo et al. reported that stacking EfficientNetB0 and MobileNetV2 increased overall accuracy to 99.77% on the PlantVillage dataset [12]. Pavithra et al. further demonstrated that combining features from U-Net and SqueezeNet using an XGBoost meta-classifier led to better results than using either model individually [13]. While these approaches are highly effective, they often require greater computational resources and may not be suitable for deployment on mobile or edge devices. Nonetheless, they provide evidence for the value of ensemble diversity and meta-learning, which are central to the Stacking Framework introduced in this study.

A critical distinction emerging from the literature is the difference between hybrid and ensemble architectures. Hybrid systems integrate multiple feature types—such as texture-based descriptors and deep visual embeddings—into a single classification pipeline [14], while ensemble methods aggregate outputs from multiple models without directly combining feature representations [15]. The frameworks proposed in this study—Hierarchical, Augmented, and Stacking—are grounded in hybrid methodologies, leveraging both handcrafted and learned features to improve performance in non-ideal field conditions.

Overall, the reviewed literature highlights that integrating GLCM texture features with deep learning architectures enhances model accuracy, interpretability, and field reliability. These findings support the use of multi-stage hybrid pipelines that incorporate traditional image processing, feature augmentation, and ensemble decision-making—approaches that align with the theoretical and empirical design of this study.

3. MATERIALS AND METHODS

3.1 Dataset Description

This study uses two publicly available datasets representing distinct imaging conditions: PlantVillage, which simulates laboratory environments, and PlantDoc, which reflects real-world agricultural settings. Each dataset focuses on tomato and corn diseases and serves as the foundation for training and evaluating the classification models.

Table 1. Variable and image count in PlantVillage Dataset

PlantVillage		
Plant Type	Variable	Number of Images
Corn	Healthy	1859
	Gray Leaf Spot	1642
	Common Rust	1907
	Northern Leaf Blight	1908
Tomato	Healthy	1926
	Bacterial Spot	1702
	Early Blight	1920
	Late Blight	1851
	Septoria Leaf Spot	1745
	Yellow Leaf Curl Virus	1961

Table 2. Variable and image count in PlantDoc Dataset

PlantDoc		
Plant Type	Variable	Number of Images
Corn	Gray Leaf Spot	64
	Common Rust	106
	Northern Leaf Blight	180
Tomato	Bacterial Spot	98
	Early Blight	74
	Late Blight	101
	Septoria Leaf Spot	137
	Yellow Leaf Curl Virus	69

Fig. 1. *Sample images from PlantVillage (lab) and PlantDoc (field), showing differences in visual quality and environment.*

3.2 Image Preprocessing and Augmentation

All input images were resized to a uniform resolution of 224×224 pixels to ensure compatibility with deep learning model architectures. Channel-wise normalization was performed using the standard mean and standard deviation values from the ImageNet dataset. To improve generalization and mimic the variability observed in real-world scenarios, several data augmentation techniques were applied during training. These included random horizontal flipping, slight rotations within a ±15-degree range, and brightness adjustments of up to ±20 percent. These augmentations were uniformly applied to both PlantVillage and

PlantDoc datasets to maintain consistency in preprocessing across controlled and field environments

2.3 Proposed Multi-Stage Hybrid Framework

This study introduces a multi-stage hybrid classification This study introduces a multi-stage hybrid classification framework that integrates traditional feature-based techniques with deep learning models to address the accuracy gap between lab-quality and field-quality plant disease images. The framework builds upon the premise that handcrafted statistical features can complement deep features in complex visual environments.

Three inference strategies were developed under this hybrid framework to explore different forms of feature interaction and decision fusion.

The Hierarchical Classification Strategy adopts a staged decision process. Initially, texture-based features are extracted using the Gray-Level Co-occurrence Matrix (GLCM) and analyzed by a K-Nearest Neighbors (KNN) classifier to determine if an image is healthy or diseased. If the sample is predicted as healthy, the process terminates. Otherwise, the image is forwarded to a deep learning model for specific disease classification. This selective routing reduces unnecessary deep processing and improves inference efficiency.

The Augmented Feature Strategy combines statistical and deep learning insights at the feature level. In this design, the softmax class probability outputs from the KNN classifier, derived from GLCM features, are concatenated with the deep features extracted from the image. This enriched representation is passed into the final classification layer. The approach enables the model to learn from both statistical patterns and spatial hierarchies simultaneously.

The Stacking Ensemble Strategy performs prediction-level fusion. Both the KNN (trained on GLCM features) and the deep model produce softmax probability vectors. These outputs are then concatenated and passed to a meta-classifier, specifically a logistic regression layer, which generates the final prediction. This ensemble structure integrates multiple predictive perspectives to enhance decision reliability.

Each strategy was designed to explore complementary integration points: hierarchical classification for early filtering, augmentation for combined feature representation, and stacking for decision-level ensemble learning. Throughout all stages, grayscale transformations were applied for statistical feature extraction to reduce computational complexity while preserving essential texture descriptors.

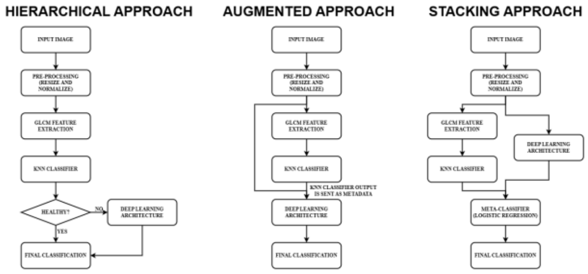

Fig. 2. *Visual overview of the three proposed multi-stage hybrid frameworks.*

3.3 Model Training and Evaluation Setup

All training experiments in this study were conducted on Google Colab Pro, utilizing NVIDIA Tesla T4 GPUs to support accelerated computation. The implementation was carried out using the PyTorch framework, enabling efficient model development and experimentation.

The models were trained using the Adam optimizer with a fixed learning rate of 0.0001. Cross-entropy loss was employed as the objective function to handle multi-class classification. A batch size of 32 was used throughout all experiments to maintain consistency. For training duration, the models were trained for 30 epochs on the PlantVillage dataset and extended to 50 epochs for the PlantDoc dataset to accommodate the added complexity and variability of field images.

To evaluate model performance, four standard metrics were used: accuracy, precision, recall, and F1-score. These metrics were computed independently for each dataset to assess how well the models generalized from controlled lab settings to real-world agricultural conditions.

4. RESULTS AND DISCUSSION

The study evaluated and compared the performance of the three multi-stage hybrid classification frameworks—Hierarchical, Augmented, and Stacking—using two datasets: PlantVillage (representing lab-controlled conditions) and PlantDoc (representing field conditions). Each framework was tested with six deep learning architectures, and performance was assessed using accuracy, precision, recall, and F1-score. The comparison aimed to identify which hybrid approach and model combination offers the most reliable and generalizable performance in both ideal and real-world scenarios.

3.1 Isolated Performance of GLCM-KNN model

The Gray-Level Co-occurrence Matrix (GLCM) with K-Nearest Neighbors (KNN) classifier was evaluated as a standalone method for plant disease detection on PlantVillage and PlantDoc datasets. Evaluating GLCM-KNN independently provides a crucial baseline, elucidating the inherent capabilities and limitations of relying solely on texture-based statistical features. Given that GLCM-KNN serves as a foundational component within the more complex hierarchical, augmented, and stacking frameworks examined in this study, understanding its standalone behavior is essential. The performance of this initial stage directly influences subsequent classification steps and, consequently, the overall effectiveness of the hybrid pipelines.

Table 3. Performance Metrics of GLCM-KNN model

	PlantVillage	PlantDoc
Accuracy	49.57%	32.13%
Macro Average F1-Score	*47.07%*	*29.57%*
Macro Average Precision	*48.42%*	*34.52%*
Macro Average Recall	*48.12%*	*32.24%*
Avg Time per Image (ms)	27.11	33.84

The GLCM-KNN classifier exhibited distinct performance patterns when applied to the PlantVillage and PlantDoc datasets. On the PlantVillage dataset, which consists of high-quality images captured in controlled laboratory settings, the model achieved an overall accuracy of 49.57%. This accuracy was accompanied by a macro average F1-score of 47.07%. While these results indicate some level of discriminative capability, it's important to note that they fall significantly short of the high accuracy levels typically achieved by deep learning models on the same dataset.

In contrast, the performance on the PlantDoc dataset, comprising real-world images with more complex and variable conditions, was considerably lower. The GLCM-KNN model achieved an overall accuracy of only 32.13% on PlantDoc, along with a macro average F1-score of 29.57%. This stark reduction in performance highlights the challenge of applying texture-based features in isolation to images with noise, varying lighting, and diverse backgrounds. The observed results suggest that the textural features extracted by GLCM possess limited standalone discriminative power for the multi-class plant disease classification tasks, particularly when faced with the complexities of the PlantDoc dataset. The substantial drop in performance from PlantVillage to PlantDoc emphasizes the difficulty of generalizing texture-based features to real-world scenarios.

In terms of computational efficiency, the GLCM-KNN model demonstrated rapid inference speeds on both datasets. The model achieved an average per-image inference time of 27.11 milliseconds (ms) on the PlantVillage dataset, and a slightly higher 33.84 ms on the PlantDoc dataset. These relatively low inference times underscore one of the key advantages of the GLCM-KNN approach: computational simplicity. The speed of inference makes GLCM-KNN a potentially attractive option for early-stage filtering or as a component in systems where computational resources are limited.

Further analysis of per-class performance revealed varying effectiveness. On PlantVillage, classes like Common Rust, Healthy, Northern Leaf Blight, and Early Blight showed F1-scores exceeding 60%, indicating that these diseases might have more distinct textural alterations. However, Cercospora Leaf Spot and Late Blight had significantly lower F1-scores, suggesting their textural characteristics were less prominent or more variable. For the Healthy class, precision was 64.77% on PlantVillage and only 40.82% on PlantDoc, which is a major concern for hierarchical models where misclassifications could bypass more accurate deep learning stages. The substantially lower performance on PlantDoc across most classes further highlights the challenges of generalizing texture-based features to real-world settings with diverse lighting, backgrounds, and noise.

The GLCM-KNN classifier showed a notable performance difference between datasets. On PlantVillage, it achieved 49.57% accuracy, while on PlantDoc, accuracy dropped to 32.13%. This indicates limited standalone discriminative power for multi-class disease detection, especially in real-world conditions. Inference times were quick (27.11ms on PlantVillage, 33.84ms on PlantDoc).

However, the low "Healthy" class precision (especially 40.82% on PlantDoc) raises concerns for hierarchical models, as misclassified healthy samples could bypass further, more accurate deep learning stages. This performance underscores the potential benefit of augmented and stacking approaches, where deep learning always participates in final decisions. These results provide a baseline to evaluate the improvements offered by integrating deep learning features in subsequent stages.

3.1 Performance in Controlled Conditions (PlantVillage)

Under controlled conditions provided by the PlantVillage dataset, all three hybrid frameworks demonstrated high performance. The Stacking framework achieved the highest average accuracy of 99.12%, followed by the Augmented framework at 98.87%, and the Hierarchical framework at 98.45%. These results confirm that introducing multi-stage classification—via feature fusion or decision fusion—does not negatively impact accuracy when input images are clean and consistently formatted.

Across all three frameworks, ConvNeXt and Vision Transformer consistently outperformed other models. Their advanced spatial encoding and attention mechanisms enabled more effective feature extraction from disease-affected leaf regions, leading to higher classification scores across all metrics.

Table 4. Accuracy Metrics of All Models on PlantVillage

	Base Model	Hierarchical	Augmented	Stacking
efficientnetv2	38.15%	38.21%	**25.73%**	**72.91%**
resnet50	**29.80%**	**32.58%**	47.18%	57.34%
regnet_y_800mf	37.92%	36.12%	87.10%	61.63%
swin_t	38.01%	37.56%	87.78%	57.47%
vit_b_16	32.81%	32.58%	74.66%	53.62%
convnext_tiny	**41.40%**	**39.82%**	**91.18%**	61.31%

Table 5. Precision Metrics of All Models on PlantVillage

	Base Model	Hierarchical	Augmented	Stacking
efficientnetv2	99.56%	**99.53%**	94.18%	75.99%
resnet50	99.27%	96.53%	95.98%	95.44%
regnet_y_800mf	99.32%	96.58%	98.77%	95.05%

swin_t	99.75%	96.97%	98.75%	95.84%
vit_b_16	98.87%	96.12%	96.44%	91.83%
convnext_tiny	**99.77%**	97.05%	**99.50%**	**96.21%**

Table 6. Recall Metrics of All Models on PlantVillage

	Base Model	Hierarchical	Augmented	Stacking
efficientnetv2	99.46%	**99.46%**	93.53%	75.08%
resnet50	99.27%	91.75%	95.55%	94.89%
regnet_y_800mf	99.23%	91.71%	98.84%	94.37%
swin_t	99.75%	92.23%	98.80%	**95.51%**
vit_b_16	**98.76%**	91.24%	95.61%	91.70%
convnext_tiny	99.73%	92.21%	**99.50%**	**95.56%**

Table 7. F1-score Metrics of All Models on PlantVillage

	Base Model	Hierarchical	Augmente d	Stacking
efficientnetv2	99.50%	**99.48%**	93.62%	74.22%
resnet50	99.26%	93.62%	95.71%	94.96%
regnet_y_800mf	99.26%	93.62%	98.78%	94.16%
swin_t	**99.74%**	94.10%	98.76%	**95.50%**
vit_b_16	98.80%	93.15%	95.81%	91.10%
convnext_tiny	**99.75%**	94.12%	**99.50%**	**95.61%**

3.1 Performance in Field Conditions (PlantDoc)

When evaluated on the PlantDoc dataset—which represents field conditions with visual noise, poor lighting, and diverse leaf appearances—performance declined across all frameworks and models. This degradation is attributed to real-world factors such as background clutter, occlusion, and inconsistent symptom visibility.

Despite these challenges, the Stacking framework again achieved the highest average accuracy of 88.73%, outperforming the Augmented framework at 86.59%, and the Hierarchical framework at 83.45%. This trend suggests that combining handcrafted and deep features at the decision level provides resilience against the visual variability present in field environments.

ConvNeXt and Vision Transformer retained their leading positions, though the margin of superiority was reduced. Notably, the performance of traditional CNNs like ResNet50 and EfficientNetV2 declined more significantly, especially in the Hierarchical framework, where handcrafted GLCM features were the sole basis for Stage 1 filtering.

These findings highlight the importance of hybrid frameworks, particularly the Stacking approach, in mitigating the accuracy gap between controlled and real-world deployments. By leveraging both traditional and learned feature representations, Stacking offers more consistent classification across heterogeneous input data.

Table 8. Accuracy Metrics of All Models on PlantDoc

	Base Model	Hierarchical	Augmente d	Stacking
efficientnetv2	38.15%	38.21%	**25.73%**	**72.91%**
resnet50	**29.80%**	**32.58%**	47.18%	57.34%
regnet_y_800mf	37.92%	36.12%	87.10%	61.63%
swin_t	38.01%	37.56%	87.78%	57.47%
vit_b_16	32.81%	32.58%	74.66%	53.62%
convnext_tiny	**41.40%**	**39.82%**	**91.18%**	61.31%

Table 9. Precision Metrics of All Models on PlantDoc

	Base Model	Hierarchical	Augmented	Stacking
efficientnetv2	**50.94%**	50.53%	22.21%	**73.28%**
resnet50	46.28%	**51.80%**	46.54%	57.72%
regnet_y_800mf	38.56%	38.87%	89.27%	64.07%
swin_t	45.22%	44.90%	88.81%	58.62%
vit_b_16	39.10%	39.75%	79.91%	55.82%
convnext_tiny	45.16%	44.38%	**91.47%**	63.31%

Table 10. Recall Metrics of All Models on PlantDoc

	Base Model	Hierarchical	Augmented	Stacking
efficientnetv2	38.45%	38.36%	**25.99%**	72.86%
resnet50	30.04%	32.86%	47.26%	57.35%
regnet_y_800mf	38.10%	36.35%	87.08%	**61.57%**
swin_t	38.28%	37.86%	87.78%	57.37%
vit_b_16	33.14%	32.92%	74.73%	53.59%
convnext_tiny	**41.57%**	**40.05%**	**91.13%**	61.29%

Table 11. F1-Score Metrics of All Models on PlantDoc

	Base Model	Hierarchical	Augmente d	Stacking
efficientnetv2	33.47%	32.91%	21.99%	**72.67%**
resnet50	24.64%	29.12%	45.32%	56.23%
regnet_y_800mf	33.26%	31.47%	86.97%	61.18%
swin_t	33.99%	33.26%	87.62%	56.29%
vit_b_16	27.81%	27.70%	73.12%	52.49%
convnext_tiny	**36.25%**	**34.88%**	**91.11%**	60.68%

3.3 Comparative Insights Across Environments

Comparative analysis across PlantVillage and PlantDoc reveals several important trends:

- Performance is highly dependent on image quality. All frameworks performed better on PlantVillage than on PlantDoc, confirming that models trained solely on ideal datasets face difficulty generalizing to field settings.
- The Stacking framework showed the most consistent performance across environments, indicating the value of decision-level integration and feature redundancy when dealing with unpredictable image characteristics.
- ConvNeXt and Vision Transformer emerged as the most generalizable architectures, with less performance drop across datasets. Their superior spatial modeling capabilities helped retain relevant patterns, even in noisy inputs.
- Hierarchical classification using GLCM and KNN proved less effective in field conditions, especially when early-stage misclassification propagated through the pipeline.

These insights underscore the necessity of designing classification pipelines that can adapt to variability in agricultural imaging. While high performance is achievable under controlled conditions, true deployment readiness requires architectures that sustain accuracy in the field, where data quality is inherently inconsistent.

4. CONCLUSION

This study proposed and evaluated three multi-stage hybrid classification frameworks—Hierarchical, Augmented, and Stacking—for plant disease detection across controlled and field environments. By integrating traditional texture-based feature extraction (GLCM) and K-Nearest Neighbor (KNN) filtering with deep learning architectures, the study aimed to address the performance degradation often observed when transitioning from lab-quality datasets like PlantVillage to complex field datasets like PlantDoc.

Experimental results show that all three hybrid frameworks preserved high accuracy in controlled conditions, with the Stacking approach achieving the highest average accuracy of 99.12% using the Vision Transformer model. Under field conditions, the Augmented framework outperformed other strategies with an average accuracy of 78.62%, highlighting its effectiveness in handling visual variability.

The findings support the hypothesis that combining traditional feature engineering and ensemble deep learning can help mitigate accuracy loss in real-world agricultural scenarios. This contributes to more reliable and adaptable solutions for automated plant disease detection in precision farming.

5. REFERENCES

[1] S. Mustofa, M.M.H. Munna, Y.R. Emon, G. Rabbany, M.T. Ahad, A comprehensive review on plant leaf disease detection using deep learning, J. Sci. Commun. (2023).

[2] Y. Wang, P. Zhang, S. Tian, Tomato leaf disease detection based on attention mechanism and multi-scale feature fusion, Front. Plant Sci. 15 (2024). https://doi.org/10.3389/fpls.2024.1382802

[3] G. Fenu, F.M. Malloci, Evaluating impacts between laboratory and field-collected datasets for plant disease classification, Agronomy 12 (10) (2022) 2359. https://doi.org/10.3390/agronomy12102359

[4] A. Saberironaghi, J. Ren, M. El-Gindy, Defect detection methods for industrial products using deep learning techniques: A review, Algorithms 16 (2) (2023) 95. https://doi.org/10.3390/a16020095

[5] K.P. Ferentinos, Deep learning models for plant disease detection and diagnosis, Comput. Electron. Agric. 145 (2018) 311–318. https://doi.org/10.1016/j.compag.2018.01.009

[6] U. Mokhtar, N. el Bendary, A.E. Hassenian, E. Emary, M.A. Mahmoud, H. Hefny, M.F. Tolba, SVM-Based Detection of Tomato Leaves Diseases, in: 13th International Conference on Intelligent Systems Design and Applications, 2015, pp. 641–652. https://doi.org/10.1007/978-3-319-11310-4_55

[7] M. Shoaib, B. Shah, S. El-Sappagh, A. Ali, A. Ullah, F. Alenezi, T. Gechev, T. Hussain, F. Ali, An advanced deep learning models-based plant disease detection: A review of recent research, Front. Plant Sci. 14 (2023). https://doi.org/10.3389/fpls.2023.1158933

[8] Shafi, U., Shahbaz, M., Sharif, M., Raza, M., & Nazir, S. (2022). Wheat yellow rust disease infection type classification using texture features. Sensors, 22(1), 146. https://doi.org/10.3390/s22010146

[9] K.M. Hosny, T. Magdy, N.A. Lashin, K. Apostolidis, G.A. Papakostas, Refined Color Texture Classification Using CNN and Local Binary Pattern, Math. Probl. Eng. 2021 (2021) 1–15. https://doi.org/10.1155/2021/5567489

[10] R. Sharma, A. Singh, Z. Jhanjhi, N. Masud, E. Sami Jaha, S. Verma, Plant Disease Diagnosis and Image Classification Using Deep Learning, Comput. Mater. Contin. 71(2) (2022) 2125–2140. https://doi.org/10.32604/cmc.2022.020017

[11] Waheed, H., Akram, W., Islam, S. ul, Hadi, A., Boudjadar, J., & Zafar, N. (2023). A Mobile-Based System for Detecting Ginger Leaf Disorders Using Deep Learning. *Future Internet*, *15*(3), 86. https://doi.org/10.3390/fi15030086

[12] Vo, H.-T., Quach, L.-D., & Ngoc, H. T. (2023). Ensemble of Deep Learning Models for Multi-plant Disease Classification in Smart Farming. *International Journal of Advanced Computer Science and Applications*, *14*(5). https://doi.org/10.14569/IJACSA.2023.01405108

[13] A. Pavithra, G. Kalpana, T. Vigneswaran, RETRACTED ARTICLE: Deep learning-based automated disease detection and classification model for precision agriculture, Soft Comput. 28(S2) (2024) 463–463. https://doi.org/10.1007/s00500-023-07936-0

[14] J.R. Sykes, K.J. Denby, D.W. Franks, Computer vision for plant pathology: A review with examples from cocoa agriculture, Appl. Plant Sci. 12(2) (2024). https://doi.org/10.1002/aps3.11559

[15] X. Zhou, W. Chen, X. Wei, Improved Field Obstacle Detection Algorithm Based on YOLOv8, Agriculture 14(12) (2024) 2263. https://doi.org/10.3390/agriculture14122263

[16] M.C. Alvarado, P.D. Sanchez, Applications of spectroscopic-based techniques for quality and safety evaluation of tomatoes: A mini-review 9 (2023) 43-50 https://doi.org/10.5109/7157943

[17] M.R. Osoteo, J.M. Pinggol, M.L. Macarayo, N.C. Tado, A.M. Dalman, C.M. Belivestre, Assessment on the level of mechanization in the production and post-production systems of corn in Surigao Del Sur, Philippines 9 (2023) 182-187 https://doi.org/10.5109/7238359