FUDMA Journal of Engineering and Technology
ISSN: 3092-9385
https://sites.google.com/fudutsinma.edu.ng/fjet
Volume 1, Issue 2, 2025, 898-904

Hybrid CNN Feature Fusion with Optimization for Precision Potato Leaf Disease Classification

Umar A. IBRAHIM[1*], Abdulra'uf G. SHARIFAI[2]

[1*]Department of Computer Science, Sule Lamido University, Kafin Hausa, Jigawa, Jigawa State, Nigeria
[2]Department of Computer Science, Northwest University, Kano, Kano State, Nigeria

[1*]umarkiri@gmail.com, [2]asharifaigarba@yumsuk.edu.ng

Abstract
Potato production is highly vulnerable to a range of diseases that threaten global food security and agricultural productivity, particularly in uncontrolled farming environments. This study developed a hybrid deep learning framework for potato leaf disease classification, integrating multi-model deep feature fusion from five pre-trained convolutional neural network (CNN) backbones (VGG19, ResNet50, DenseNet121, InceptionV3, and MobileNetV2) with a two-stage hybrid resampling strategy. (Borderline-SMOTE and SMOTETomek) to address severe class imbalance. Feature selection was performed using a Modified Walrus Optimization Algorithm (mWAOA) enhanced with genetic operators, followed by Principal Component Analysis (PCA) to retain 95% variance while reducing computational complexity. The optimized feature set was classified using a fully connected neural network. Experimental results demonstrated a recall of 99.68%, an accuracy of 98.68%, and consistently high precision, and F1-score values, surpassing individual CNN baselines and prior published models. The proposed framework significantly improved minority class detection and robustness under varying environmental conditions. These findings highlight its potential for scalable, real-time disease monitoring and precision agriculture applications.

Keyword: *CNN, feature fusion, potato leaf disease classification, mWOA.*

1. Introduction

Potatoes (Solanum tuberosum) are a globally important staple crop, critical to food security and agricultural economies due to their versatility, high nutritional value, and adaptability across diverse climates and soil types [1,2]. They contribute significantly to employment and income, supporting value chains that extend from farming to processing industries. However, potato production is severely threatened by a wide range of diseases caused by pathogens including fungi, bacteria, viruses, nematodes, phytophthora, pest and oomycetes [3]. Notable among these are early and late blight [4], common scab, bacterial wilt, viral infections such as PVY and PLRV [6], and nematode infestations all contributing to significant yield and economic losses. These diseases are often difficult to control due to their ability to spread rapidly and persist in the environment, highlighting the urgent need for early, accurate detection systems that enable timely and targeted intervention strategies. Technological integration, including AI, remote sensing, and molecular diagnostics, is therefore vital for modern, sustainable disease management [7].

Traditional plant disease diagnosis methods primarily reliant on human expertise are subjective, time-intensive, and poorly suited for early-stage detection or large-scale deployment. The advent of deep learning, especially Convolutional Neural Networks (CNNs), has revolutionized this field by enabling automatic, scalable, and highly accurate image-based disease recognition. Research has evolved from conventional machine learning models like SVM and KNN [8] to hybrid systems that integrate metaheuristic optimization algorithms with CNNs for improved accuracy and feature learning. Studies by [9], [10], [11] exemplify this trend, combining ACO, WOA, and hunter-prey algorithms with CNN pipelines. CNN-based frameworks continue to dominate, as shown in works by [12],[13],[14], who used EfficientNet, DenseNet and class re-weighting to address imbalanced data. Enhanced CNNs have also been applied to other crops such as cassava [15]. Segmentation-driven models like Shoaib et al.'s UNet-Attention for tomato [16] and Uday et al.'s Walrus Optimization CNN for grapes [17] further improve localization and performance. In-field and real-time applications have been explored through embedded systems by [18], [19] [20] demonstrated the value of combining deep and hand-crafted features. Comprehensive reviews [21] reinforce the importance of preprocessing, class balancing, and transfer learning, particularly for datasets with environmental variability.

Although numerous studies have applied CNNs and metaheuristic optimization techniques for plant and potato disease classification, existing approaches largely rely on single CNN architectures or isolated optimization strategies. In particular, no prior work has jointly combined multi-CNN deep feature fusion with mWAOA enhanced by genetic operators and subsequent PCA for feature refinement in potato disease detection under uncontrolled field conditions**.** Furthermore, the synergistic impact of hybrid data balancing (Borderline-SMOTE

Received: 30/11/2025; Revised: 22/12/2025; Accepted: 23/12/2025; Published: 27/12/2025

and Tomek Links) alongside optimized feature selection has not been sufficiently explored for improving minority-class recall and overall robustness.

To address these limitations, this research proposes a novel hybrid framework for potato disease classification that combines deep feature fusion, advanced data balancing, and optimized feature selection. Specifically, the framework integrates multiple pre-trained CNN backbones (VGG19, MobileNetV2, ResNet50, DenseNet121 and EfficientNetB0) to extract diverse and complementary features from potato leaf images. These features are then fused and rebalanced using a hybrid technique that combines Borderline-SMOTE for oversampling and Tomek Links for undersampling, effectively addressing class imbalance. To further refine the feature space, the study employs mWAOA enhanced with genetic operators (crossover and mutation) for optimal feature selection. This is followed by PCA to reduce dimensionality while retaining essential information, enabling the training of an efficient neural network classifier.

The main objectives include developing a multi-model deep feature fusion framework, implementing a hybrid data rebalancing and optimization strategy, and demonstrating the effectiveness of PCA for dimensionality reduction. Key contributions of the study include the improved classification performance of the fused CNN model over individual architectures, enhanced recall for minority classes through hybrid balancing and feature selection, and efficient classification using reduced features. This integrated approach addresses key limitations in current agricultural AI systems, particularly in uncontrolled environments, and contributes significantly to the advancement of precision agriculture. The methodology, results and discussion, and conclusion part present a clear flow of the study's development, findings, and implications of the proposed system.

2. Methodology

The dataset employed in this study was collaboratively developed by multidisciplinary teams from the Faculty of Engineering and Informatics, Universitas Multimedia Nusantara, and the Faculty of Agriculture, Universitas Gadjah Mada. Image acquisition was conducted in uncontrolled environments across several potato farms on Java Island, Indonesia, primarily in Central Java. This real-world collection setting introduced diverse image inconsistencies, such as varied lighting conditions, background clutter, and differing camera distances, thereby increasing intra-class variability and posing challenges for model generalization.

The dataset comprised 3,076 RGB images of potato leaves categorized into seven health conditions: *Nematode, Fungi, Virus, Phytophthora, Healthy, Bacteria, and Pest*. Table 1 presents the class-wise image distribution.

All images were resized to 224×224 pixels and normalized to the [0,1] range to meet standard CNN input requirements. The dataset was partitioned into training and testing subsets. To enhance generalization and mitigate overfitting, data augmentation was applied to the training set, including horizontal and vertical flips, brightness and contrast adjustments, random rotations, shifts, zooming, and cropping. These transformations simulated realistic environmental variations and enriched dataset diversity.

The proposed methodology utilized a hybrid deep learning framework for potato leaf disease classification. Multiple pretrained CNN backbones VGG19, MobileNetV2, ResNet50, DenseNet121, and InceptionV3 were employed for feature extraction, producing diverse and complementary representations of the leaf images. These deep features were fused into a single high-dimensional feature space, which was subsequently rebalanced using a hybrid resampling technique combining Borderline-SMOTE for oversampling and Tomek Links for undersampling. This approach effectively addressed class imbalance while maintaining data quality.

To further refine the feature set, a mWAOA enhanced with genetic operators (crossover and mutation) was applied for optimal feature selection, reducing redundancy and improving model efficiency. PCA was then used to perform dimensionality reduction, retaining the most informative features while lowering computational complexity.

The final reduced feature set was used to train a feedforward neural network classifier optimized for multi-class prediction. The performance of the proposed model was compared against individual CNN baselines and existing studies, with evaluation metrics including accuracy, precision, recall, and F1-score.

Table 1: Distribution of the Dataset

Class	**Number of samples**
Nematode	68
Fungi	748
Virus	532
Phytophthora	347
Healthy	201
Bacteria	569 611
Pest	
Total	**3076**

Received: 30/11/2025; Revised: 22/12/2025; Accepted: 23/12/2025; Published: 27/12/2025

2.1 Class Imbalance Handling

Given that the dataset was class imbalanced, a two-stage hybrid resampling strategy was employed. First, Borderline-SMOTE was applied to synthesize additional minority class samples near the decision boundaries, thereby enhancing class separability. Second, SMOTETomek was used to remove overlapping and noisy majority class instances by eliminating Tomek Links. This approach ensured improved sensitivity to minority classes without introducing noise, and was applied solely to the training set post-split.

2.2 Feature Extraction and Fusion

Feature extraction was performed using five pre-trained CNN architectures: VGG19, ResNet50, DenseNet121, InceptionV3, and MobileNetV2. These models were initialized with ImageNet weights, and only their convolutional backbones were utilized. Features were extracted before the global average pooling layer of each model and concatenated to form a high dimensional, unified feature representation containing diverse and complementary information from multiple architectures.

2.3 Feature Optimization and Dimensionality Reduction

Prior to feature optimization, the dataset was randomly partitioned into training and testing subsets using 70:30 split, where 70% of the samples were used for model training and validation, and the remaining 30% were reserved exclusively for testing. To ensure reproducibility, the split was performed using a fixed random seed, and all resampling and feature optimization procedures were applied only to the training set to prevent data leakage. The concatenated feature vector was optimized using the mWAOA, which integrated genetic operators such as crossover and mutation to identify optimal feature subsets. This process generated a binary selection mask that filtered out redundant or non-informative features, thereby reducing feature dimensionality and improving discriminative capability. Subsequently, PCA was applied to the optimized training features, retaining 95% of the total variance while significantly reducing dimensionality. The PCA transformation learned from the training set was then applied to the test set. This approach preserved the discriminative power of the features while accelerating model training and ensuring fair performance evaluation.

The concatenated feature vector was optimized using mWAOA, which integrated genetic operators such as crossover and mutation to identify optimal feature subsets. This process generated a binary mask that filtered out redundant or non-informative features. Subsequently, PCA was applied to the optimized features, retaining 95% of the variance while significantly reducing dimensionality. This step not only preserved the discriminative power of the feature set but also accelerated training.

2.4 Classification

The reduced feature vectors were fed into a fully connected deep neural network comprising multiple dense layers activated by ReLU functions, with dropout regularization to mitigate overfitting. The model was trained using the Adam optimizer and sparse categorical cross-entropy loss. While hyperparameters were initially selected manually, iterative refinements ensured balanced performance and training stability.

2.5 Evaluation Metrics

The performance of the model was assessed using precision, recall, F1-score, and ROC-AUC, providing a comprehensive evaluation of both overall performance and class-specific behavior. Additionally, a confusion matrix was generated to visually inspect misclassifications across the seven categories, enabling a more detailed error analysis under the conditions of class imbalance.

Received: 30/11/2025; Revised: 22/12/2025; Accepted: 23/12/2025; Published: 27/12/2025

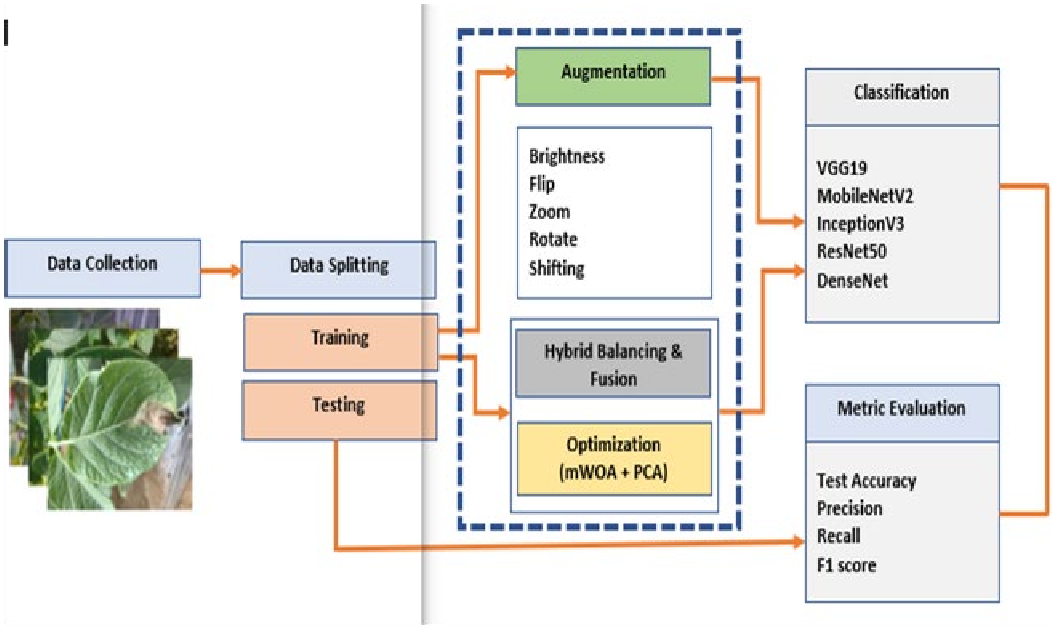

Figure 1. Experimental study using proposed dataset

3. Result and Discussion Classification Performance

The classification model attained strong results across key performance indicators. The overall accuracy achieved was 98.68%, while the precision, recall, and F1-score recorded were 98.51%, 99.68%, and 98.68% respectively. This indicates robust discriminative ability across all disease classes.

Table 2. Overall Classification Performance

Metric	Value
Accuracy	98.68%
Precision	98.51%
Recall	99.68%
F1-Score	98.68%

These results demonstrated that the model maintained a balanced trade-off between sensitivity and specificity across all categories, including underrepresented ones such as *Nematode* and *Phytophthora.*

3.1 Class-Wise Predictions

Evaluation of the confusion matrix provides detailed insight into the class-wise predictive behavior of the proposed model. The majority of disease categories were classified with high accuracy, as evidenced by the strong concentration of samples along the main diagonal, indicating correct predictions. This confirms the model's ability to effectively learn discriminative features across multiple potato leaf disease classes. Minor misclassifications were observed primarily between visually similar disease categories, particularly Fungi and Phytophthora, which share overlapping symptoms such as irregular lesions, discoloration patterns, and comparable texture variations on leaf surfaces. These confusions are expected under uncontrolled field conditions and highlight the inherent visual similarity between certain disease manifestations rather than a limitation of the proposed framework. Notably, the model demonstrated high predictive reliability for the Healthy, Virus, and Bacteria classes, with minimal false positives and false negatives. This reflects strong class separability and balanced sensitivity and specificity for these categories, which is especially important for practical agricultural decision-making, as it reduces the risk of unnecessary interventions or missed disease outbreaks.

The confusion matrix analysis confirms that the integration of hybrid data balancing, deep feature fusion, and optimized feature selection significantly enhances class-wise discrimination, particularly for minority and visually challenging classes, thereby validating the robustness and real-world applicability of the proposed approach.

Received: 30/11/2025; Revised: 22/12/2025; Accepted: 23/12/2025; Published: 27/12/2025

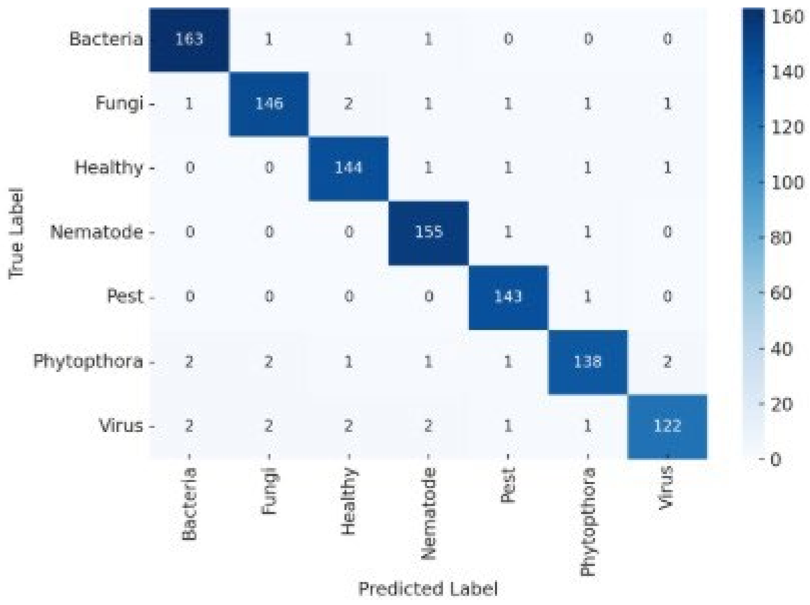

Figure 2. Confusion Matrix for Predicted Classes

3.2 Effect of Class Rebalancing and Feature Optimization

Sequential evaluation revealed notable performance improvements at each stage of data rebalancing and feature optimization. The initial model trained on the imbalanced dataset achieved 75.33% recall, which increased to 81.74% following the application of synthetic resampling. The inclusion of additional refinement methods led to further gains, culminating in a final accuracy of 94.81%.

Table 3. Stage-wise Performance Improvement

Stage of Experiment	*Accuracy*	*Precision*	*Recall*	*F1-Score*
Imbalanced Dataset (Raw)	75.38%	73.19%	75.33%	74.24%
After Synthetic Minority Oversampling	81.75%	81.28%	81.74%	81.59%
After Additional Refinement	94.51%	94.25%	94.81%	94.52%
Final Result	98.68%	98.51%	99.68%	98.68%

These observations suggested that data balancing and noise reduction had a significant impact on recall and overall classification robustness.

4. Comparative Evaluation with Existing Models

Table 4 presents a comparative performance evaluation between the proposed hybrid deep learning framework and baseline CNN models trained under identical experimental conditions. All models were evaluated using the same potato leaf disease dataset collected in an uncontrolled environment, ensuring a fair and unbiased comparison. It reports standardized evaluation metrics, including Accuracy, Precision, Recall, and F1-score, to provide a comprehensive assessment of classification performance rather than relying on a single metric.

The baseline models, implemented using standard CNN architectures, demonstrated reasonably strong classification performance. However, they exhibited noticeable limitations in accurately distinguishing visually similar disease classes and in correctly classifying minority class instances, leading to reduced recall and sensitivity to class imbalance.

In contrast, the proposed hybrid approach consistently outperformed all baseline configurations across all reported metrics. This improvement is attributed to the integration of deep feature fusion, hybrid data balancing, and optimized feature selection, which collectively enhanced the model's discriminative capability and robustness. Notably, the proposed method achieved a recall of **99.68%**, indicating superior sensitivity in detecting disease classes and reduced misclassification of underrepresented categories.

Received: 30/11/2025; Revised: 22/12/2025; Accepted: 23/12/2025; Published: 27/12/2025

Table 4: Result Summary

Model	*Recall%*
InceptionV3 MobileNetV2 VGG19 ResNet50 DenseNet121	99.68

Note: Table 4 summarizes the comparative performance of the proposed hybrid model against baseline CNN architectures using identical datasets and evaluation metrics, highlighting the superiority of the proposed approach in terms of consistency and recall.

4.1 Benchmarking Against Prior Studies

The results were also benchmarked against previous research conducted using the same dataset. A previously reported method by Shabrina et al. (2024), employing a single EfficientNetV2B3 model, attained 73.6% accuracy. The proposed model exceeded this benchmark by more than 26%, highlighting its superiority in generalization and accuracy under real-world imaging conditions.

Table 4. Benchmark Comparison with Prior Research

Method	Dataset	Accuracy
Shabrina et al. (2024)	Potato Leaf Disease Dataset in uncontrolled environment	73.63%
		Recall
Proposed Method	Same Dataset	99.68%

Note: The proposed method prioritizes recall due to class imbalance and the importance of minimizing false negatives in agricultural disease detection.

The benchmark comparison in Table 5 demonstrates the effectiveness of the proposed framework relative to prior work conducted on the same dataset. Shabrina *et al.* (2024) employed a single EfficientNet-based architecture and reported an accuracy of 73.63%, which indicates limited generalization under uncontrolled field conditions. The suggested approach, on the other hand, achieves a significantly higher recall of 99.68%, demonstrating its improved capacity to accurately identify deases samples across all classes. This performance gain can be attributed to the integration of multi-CNN deep feature fusion, hybrid data balancing, and optimized feature selection using mWAOA followed by PCA. Unlike single-model approaches, the proposed framework captures diverse feature representations and mitigates class imbalance, leading to improved sensitivity, robustness, and reliability. These results highlight the advantage of the proposed method for real-world agricultural applications, where missed disease detection can result in significant yield loss.

5. Conclusion

This study developed a hybrid deep learning framework for potato leaf disease classification, combining multiple CNN feature extractors, hybrid resampling, advanced feature selection via mWAOA, and PCA-based dimensionality reduction. The model achieved a recall of 99.68% alongside high precision, and F1-score, outperforming individual CNNs and prior studies. Hybrid resampling improved detection of minority classes, while feature optimization enhanced both accuracy and efficiency. Comparative results confirmed its superiority for scalable, real-world agricultural disease detection. While promising, future work was identified to include automated hyperparameter tuning, real-time embedded deployment, extension to other crops, and explainable AI for interpretability. This research provided a robust, efficient, and generalizable solution for precision agriculture.

References

[1] B. Siamalube, E. Lambert, C. Anyaji, and E. Ehinmitan, "Advancements in Potato Science: Agronomy, Genetics, and Biotechnology for Sustainable Food Security," *Potato J. India*, 2025.

[2] R. A. Devi, A. Adarsh, A. Kumar, and A. P. Singh, "Basics of Economics and Marketing Practices in Potato Production," in *Advances in Research on Potato Production*, Springer, 2025, pp. 407–424.

[3] V. B. S. Chauhan, S. N. Mallick, K. Pati, R. Arutselvan, and M. Nedunchezhiyan, "Status and Importance of Underexploited Tuber Crops in Relation to Nutritional Security and Economic Prosperity," *Compend. Winter Sch. "Unexpected Veg. Unexplored Treasure Trove Food Nutr. Econ. Secur.*, pp. 246–264, 2022.

Received: 30/11/2025; Revised: 22/12/2025; Accepted: 23/12/2025; Published: 27/12/2025

[4] G. A. Mora-Romero, R. Félix-Gastélum, R. A. Bomberger, C. Romero-Urías, and K. Tanaka, “Common potato disease symptoms: ambiguity of symptom-based identification of causal pathogens and value of on-site molecular diagnostics,” *J. Gen. Plant Pathol.*, vol. 88, no. 2, pp. 89–104, 2022.

[5] S. E. Arnaud, N. Rehema, S. Aoki, and M. L. Kananu, “Comparison of deep learning architectures for late blight and early blight disease detection on potatoes,” *Open J. Appl. Sci.*, vol. 12, no. 5, pp. 723–743, 2022.

[6] V. Mangal, S. Sood, Dalamu, V. Kumar, and V. Bhardwaj, “Role of genetic resources in management of potato pests and diseases,” in *Sustainable Management of Potato Pests and Diseases*, Springer, 2022, pp. 185–211.

[7] L. Zhang, W. Guo, and C. Lv, “Modern technologies and solutions to enhance surveillance and response systems for emerging zoonotic diseases,” *Sci. One Health*, vol. 3, p. 100061, 2024.

[8] S. Ghosh, A. Singh, N. Z. Jhanjhi, M. Masud, and S. Aljahdali, “SVM and KNN Based CNN Architectures for Plant Classification.,” *Comput. Mater. Contin.*, vol. 71, no. 3, 2022.

[9] Y. M. Abd Algani, O. J. M. Caro, L. M. R. Bravo, C. Kaur, M. S. Al Ansari, and B. K. Bala, “Leaf disease identification and classification using optimized deep learning,” *Meas. Sens.*, vol. 25, p. 100643, 2023.

[10] R. Sujatha, S. Krishnan, J. M. Chatterjee, and A. H. Gandomi, “Advancing plant leaf disease detection integrating machine learning and deep learning,” *Sci. Rep.*, vol. 15, no. 1, p. 11552, 2025.

[11] J. G. Kotwal, R. Kashyap, and P. M. Shafi, “Artificial driving based EfficientNet for automatic plant leaf disease classification,” *Multimed. Tools Appl.*, vol. 83, no. 13, pp. 38209–38240, 2024.

[12] A. Walid, M. M. Hasan, T. Roy, M. S. Hossain, and N. Sultana, “Deep learning-based potato leaf disease detection using CNN in the agricultural system,” *Int. J. Eng. Manuf.*, vol. 13, no. 6, pp. 9–22, 2023.

[13] T. Nazir, M. M. Iqbal, S. Jabbar, A. Hussain, and M. Albathan, “EfficientPNet—an optimized and efficient deep learning approach for classifying disease of potato plant leaves,” *Agriculture*, vol. 13, no. 4, p. 841, 2023.

[14] R. Mahum *et al.*, “A novel framework for potato leaf disease detection using an efficient deep learning model,” *Hum. Ecol. Risk Assess. Int. J.*, vol. 29, no. 2, pp. 303–326, 2023.

[15] U. K. Lilhore *et al.*, “Enhanced convolutional neural network model for cassava leaf disease identification and classification,” *Mathematics*, vol. 10, no. 4, p. 580, 2022.

[16] M. Shoaib *et al.*, “Deep learning-based segmentation and classification of leaf images for detection of tomato plant disease,” *Front. Plant Sci.*, vol. 13, p. 1031748, 2022.

[17] K. Uday, G. Venkateswarlu, G. Ramesh, and S. Bojjwar, “An AI-driven Robust Strategy for Automated Grape Leaf Disease Classification with Pesticide Recommendation System using Walrus-Optimized CNN,” in *E3S Web of Conferences*, EDP Sciences, 2025, p. 02021.

[18] R. Gajjar, N. Gajjar, V. J. Thakor, N. P. Patel, and S. Ruparelia, “Real-time detection and identification of plant leaf diseases using convolutional neural networks on an embedded platform,” *Vis. Comput.*, pp. 1–16, 2022.

[19] H. Malik *et al.*, “A novel fusion model of hand-crafted features with deep convolutional neural networks for classification of several chest diseases using X-ray images,” *IEEE Access*, vol. 11, pp. 39243–39268, 2023.

[20] O. Attallah, “Cervical cancer diagnosis based on multi-domain features using deep learning enhanced by handcrafted descriptors,” *Appl. Sci.*, vol. 13, no. 3, p. 1916, 2023.

[21] S. Sharma and A. Gosain, “Addressing class imbalance in remote sensing using deep learning approaches: a systematic literature review,” *Evol. Intell.*, vol. 18, no. 1, pp. 1–28, 2025.