IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 15, No. 2, April 2026, pp. 1783~1789
ISSN: 2252-8938, DOI: 10.11591/ijai.v15.i2.pp1783-1789

Hybrid machine learning for imbalanced lettuce disease classification

Fazlur Ihzanurahman, Wayan Firdaus Mahmudy
Department of Informatics Engineering, Faculty of Computer Science, Brawijaya University, Malang, Indonesia

Article Info

Article history:

Received Nov 28, 2024
Revised Feb 9, 2026
Accepted Mar 5, 2026

Keywords:

Deep feature extraction
EfficientNet-B3
Hybrid machine learning
Lettuce disease
Support vector machine

ABSTRACT

This study investigates a hybrid machine learning framework combining EfficientNet-B3 feature extraction with classical classifiers for lettuce disease classification under conditions of extreme class imbalance. The system utilizes EfficientNet-B3 to extract high-dimensional feature embeddings from 2,337 images, which are subsequently classified using support vector machine (SVM), random forest (RF), and k-nearest neighbors (KNN). Although the proposed SVM-based model achieves a high overall accuracy of 94.01%, experimental results reveal a substantial performance discrepancy compared to the macro F1-score of 37.94%. This critical gap indicates that while the model successfully identifies the majority classes, it fails to recognize rare disease categories with limited samples. Theoretical analysis suggests that while SVM handles high-dimensional feature spaces more effectively than RF and KNN, the deep features extracted are biased toward majority class characteristics. These findings highlight the severe limitations of accuracy-centric evaluation in agricultural diagnostics and demonstrate that deep feature extraction alone is insufficient to guarantee robust detection for minority pathologies. The study concludes that relying on aggregate accuracy can mask diagnostic failures, emphasizing the urgent need for per-class performance analysis and data-level mitigation strategies in future research.

Corresponding Author:

Wayan Firdaus Mahmudy
Department of Informatics Engineering, Faculty of Computer Science, Brawijaya University
St. Veteran No.10-11, Ketawanggede, Lowokwaru, Malang 65145, Indonesia
Email: wayanfm@ub.ac.id

1. INTRODUCTION

Agricultural productivity is often compromised by plant diseases, which negatively affect both the quality and quantity of crop yields. Lettuce (*Lactuca sativa* L.), in particular, is highly susceptible to a range of pathogens, including bacteria, fungi, and viruses [1]. These infections can result in substantial financial losses for farmers and pose broader risks to agricultural sector. As such, the ability to detect plant diseases early and accurately is essential to effectively manage their spread and reduce associated damages [2].

Recent advancements in machine learning and image processing have opened up new possibilities for automating plant disease detection, with encouraging outcomes [3]. These techniques are capable of handling large datasets and conducting intricate analyses to uncover patterns that are often undetectable by human observation. Among the various approaches, convolutional neural networks (CNNs) stand out as highly effective for image classification [4], [5]. CNNs excel at autonomously learning image features through stacked convolutional layers, which makes them particularly well-suited for tasks involving pattern recognition and visual categorization [6], [7].

Although CNN are highly effective, their classification accuracy can be enhanced by combining them with conventional machine learning classifiers like support vector machine (SVM), random forests (RF), and k-nearest neighbors (KNN) [8]. SVM aims to construct an optimal hyperplane that separates data points belonging to different classes, and it has shown strong performance across numerous classification tasks [9], [10]. RF, a type of ensemble method, leverages multiple decision trees to enhance accuracy and minimize overfitting, offering a robust solution for classification problems [11]. Meanwhile, KNN classifies data based on the proximity of neighboring instances and is widely adopted due to its ease of implementation and effectiveness [12].

Numerous CNN architectures have been developed, each tailored to offer specific advantages in image classification tasks. EfficientNet architectures have demonstrated superior performance in various plant disease classification tasks, such as in potato and mango crops [13]. Motivated by this success, this research investigates the performance of the EfficientNet-B3 model in combination with three machine learning classifiers—SVM, RF, and KNN—for disease detection in lettuce plants. EfficientNet represents a modern CNN design that balances computational cost and predictive accuracy by employing compound scaling, which simultaneously adjusts the model's depth, width, and input resolution [14]. Through this hybrid approach, the study aims to determine whether the integrated model delivers superior classification performance compared to using EfficientNet-B3 alone. The findings of this study are intended to support the advancement of more accurate and efficient techniques for detecting plant diseases. Moreover, the proposed approach holds potential for practical application in agriculture, offering a tool that could assist farmers in monitoring crop health and improving disease management strategies [15], [16].

Several recent studies have highlighted the potential of combining EfficientNet with advanced deep learning strategies to enhance the precision of plant disease detection [17], [18]. Extensive prior work also supports the effectiveness of CNN-based models for this task [19]. CNNs have been broadly adopted in the classification of plant diseases, particularly using leaf imagery as the primary input. Reviews in the field consistently show that CNNs outperform traditional classification techniques, largely because of their superior capability to capture intricate patterns and features from images [14].

Architectures such as VGG16, ResNet, and EfficientNet have been widely utilized in plant disease classification tasks. While VGG16 and ResNet are recognized for their high classification accuracy, their implementation often demands considerable computational power [20]. In contrast, EfficientNet presents a more resource-efficient alternative. Recent studies have also shown that combining EfficientNet with attention mechanisms can further enhance both accuracy and interpretability in plant disease detection tasks [21]. By utilizing compound scaling to adjust network dimensions in a balanced manner, EfficientNet achieves improved performance in terms of both efficiency and accuracy compared to earlier CNN models [15].

The application of EfficientNet in plant disease identification has shown consistently high levels of accuracy across various studies [22]. One notable example is its use in classifying apple leaf diseases, where it demonstrated exceptional performance, further validating its reliability and suitability for disease detection in agricultural contexts [23], [24]. Furthermore, a recent study demonstrated that EfficientNet performed robustly in multi-crop disease classification scenarios, highlighting its scalability and accuracy across diverse plant types [25].

This research explores the integration of the EfficientNet-B3 architecture with three traditional classification algorithms SVM, RF, and KNN, to assess potential gains in classification accuracy. These methods are widely employed in machine learning for their robust performance. SVM functions by constructing an optimal decision boundary to distinguish between classes [9]. RF, which aggregates multiple decision trees, is effective in boosting accuracy and mitigating overfitting issues [11]. It has also been previously applied in plant disease classification, serving as a solid benchmark for evaluating the performance of more recent models like EfficientNet [26]. Meanwhile, KNN determines class membership based on proximity to neighboring samples and remains a straightforward yet effective technique for classification tasks [12].

Despite the widespread adoption of CNN-based models for plant disease classification, most existing studies rely on balanced datasets and emphasize overall accuracy, with limited attention to minority disease classes. Consequently, the reliability of these models in real-world, highly imbalanced agricultural scenarios remains insufficiently explored. Current literature often overlooks the risk that high accuracy metrics may mask the model's complete failure in detecting rare but critical pathologies.

This study provides an empirical analysis of the limitations of hybrid deep feature extraction models when applied to highly imbalanced lettuce disease datasets. It offers critical insights into the mismatch between reported accuracy and actual diagnostic reliability. Specifically, it evaluates the performance of EfficientNet-B3 features combined with SVM, RF, and KNN classifiers, analyzing how extreme class imbalance distorts evaluation metrics and affects the decision boundaries of different classical algorithms.

2. METHOD

This study employs the EfficientNet-B3 architecture as a feature extractor, which is then integrated with three classical machine learning classifiers SVM, RF, and KNN to construct hybrid models. A stratified 5-fold cross-validation scheme is used to ensure that the proportional distribution of classes is preserved in every fold. This approach reduces the risk of overfitting and provides a more reliable estimation of model performance, particularly given the imbalanced nature of the dataset.

2.1. Dataset

This study utilized the lettuce disease dataset sourced from Kaggle, which includes a diverse collection of images representing both healthy and diseased lettuce leaves. The dataset contains 2,337 images of lettuce diseases categorized into eight classes: healthy (1,123 images), Shepherd's purse weeds (1,106), downy mildew on lettuce (30), bacterial infection (20), septoria blight (19), powdery mildew (18), viral infection (15), and wilt and leaf blight (6). The distribution is highly imbalanced, with two classes accounting for more than 95% of all samples, while several disease categories contain fewer than 20 images. This imbalance introduces a significant risk of biased learning, where models may achieve high accuracy by prioritizing majority classes while performing poorly on minority disease categories. The dataset includes variations in leaf texture, lighting, and background conditions, further increasing complexity of classification.

2.2. Preprocessing

All images were resized to 224×224 pixels to match EfficientNet-B3 input requirements and normalized to a [0,1] range. No aggressive augmentation was applied to avoid creating unrealistic samples for minority classes. Instead, data balancing was handled through cross-validation and class weights.

2.3. Feature extraction

EfficientNet-B3 was used as a fixed feature extractor. The model was loaded with ImageNet weights, and all layers were frozen. The output was taken from the final convolutional block followed by global average pooling, resulting in a 1,536-dimensional feature vector for each image. These feature vectors were then normalized using StandardScaler before being forwarded into the classification models.

2.4. Hybrid classification models (SVM, RF, and KNN)

Three classical machine learning classifiers were integrated with EfficientNet-B3 features: SVM (radial basis function or RBF kernel), RF (200 trees, maximum depth =none), and KNN (k =5, Euclidean distance). The implementation was conducted using the Scikit-learn library in Python to ensure reproducibility. These models were selected due to their ability to handle high-dimensional feature embeddings and provide strong baseline performance with limited training data.

2.5. Handling dataset imbalance

To mitigate the impact of severe class imbalance, stratified sampling was applied in all training and validation splits to preserve the original label distribution. Weighted loss functions were implemented for SVM, and class-balanced sampling was used during feature extraction. This strategy ensures that minority classes contribute proportionally during training and evaluation.

2.6. Stratified 5-fold cross-validation

A stratified 5-fold cross-validation procedure was implemented to ensure fair and robust evaluation. In each fold, 80% of the data was used for training and 20% for testing, preserving the distribution of each class. No image from the same class subset was allowed to appear in both training and testing partitions, preventing data leakage. Model performance was reported as mean and standard deviation across all folds.

2.7. Evaluation metrics

Due to the highly imbalanced dataset, multiple evaluation metrics were used, including accuracy, macro F1, and weighted F1. Accuracy alone is insufficient for imbalanced datasets because it may be dominated by majority classes. Macro F1 treats all classes equally and is therefore a more reliable indicator of performance on minority disease categories.

3. RESULTS AND DISCUSSION

The experimental results indicate that SVM achieves the highest performance among the three classifiers, based on both accuracy and F1-score. However, a substantial discrepancy is observed between accuracy and macro F1-scores across all models, highlighting the strong impact of class imbalance within the dataset. Consequently, the interpretation of results emphasizes not only accuracy but also class-level performance, supported by confusion matrices and detailed F1-score analysis.

3.1. EfficientNet-B3 feature extraction result

EfficientNet-B3 successfully extracted 1,536-dimensional feature vectors for each of 2,337 images. The embeddings captured leaf texture, color variation, and lesion patterns, forming a robust representation for downstream classification. After normalization using StandardScaler, the feature distributions across folds were consistent, indicating that EfficientNet-B3 provided stable high-level representations regardless of data split. This feature extraction approach reduced training time significantly while maintaining the expressive power of deep convolutional encoders.

3.2. Performance comparison across hybrid models

The hybrid classifiers SVM, RF, and KNN were evaluated using 5-fold stratified cross-validation to ensure a fair assessment on the highly imbalanced dataset. The detailed performance comparison of these models is presented in Table 1. The results demonstrate that SVM outperformed both RF and KNN. Theoretically, this can be attributed to the nature of the high-dimensional feature space (1,536 dimensions) generated by EfficientNet-B3. SVM is particularly effective in such scenarios because it constructs optimal hyperplanes that maximize the margin between classes, allowing it to handle high-dimensional data without succumbing easily to overfitting. In contrast, KNN showed the lowest performance, likely due to the "curse of dimensionality," where distance metrics (such as Euclidean distance) lose their discriminative power as the volume of the space increases, making minority samples indistinguishable from the majority neighbors. Similarly, RF struggled because, in the presence of extreme imbalance, the individual decision trees tend to be biased toward the majority class to minimize overall impurity.

Table 1. Performance comparison of models on imbalanced datasets

Classification method	Accuracy	Weighted F1	Macro F1
SVM	94.01%±0.0078	93.91%±0.0097	37.94%±0.0309
RF	88.79%±0.0163	87.19%±0.0158	32.18%±0.0551
KNN	87.93%±0.0095	86.62%±0.0098	29.39%±0.0416

3.3. Accuracy macro F1 discrepancy and class imbalance

A substantial discrepancy appears between accuracy (94%) and macro F1 (only 37%). This behavior is expected given the extreme class imbalance in the dataset:

- Healthy: 1,123 images
- Shepherd’s purse: 1,106 images
- Minority classes: 6–30 images per class

Accuracy is, therefore, misleading, as a classifier can achieve high accuracy by predicting majority classes correctly while failing almost entirely on rare disease categories. Macro F1, which weights all classes equally, reveals this weakness clearly.

3.4. Confusion matrix analysis

Figure 1 illustrates the combined 5-fold confusion matrix for the SVM classifier. The matrix clearly shows:

- Healthy (1,107/1,123) and Shepherd’s purse (1,092/1,106) are detected with near-perfect precision.
- Minority diseases such as wilt and leaf blight, viral, powdery mildew, and septoria blight exhibit severe misclassification, mostly predicted as the two majority classes.
- Diseases with fewer than 20 samples show almost zero diagonal intensity, confirming failure to learn discriminative patterns.

The confusion matrix as shown in Figure 1 provides deeper insight into the classification failures. It reveals that minority disease samples (e.g., wilt, viral, septoria) are frequently misclassified as "healthy" or "weeds." This suggests that while EfficientNet-B3 successfully extracts robust visual features for general leaf structures (texture, shape), it fails to capture the subtle, fine-grained visual patterns required to distinguish rare diseases. This is not merely a classifier failure but a data distribution issue; the feature extractor learns to prioritize the visual characteristics of the majority classes. Consequently, the decision boundaries formed by the classifiers are overwhelmed by the density of the majority class features. As shown in the confusion matrix, minority classes such as viral and wilt yielded nearly zero correct predictions, resulting in an F1-score of ~0.0 for these categories.

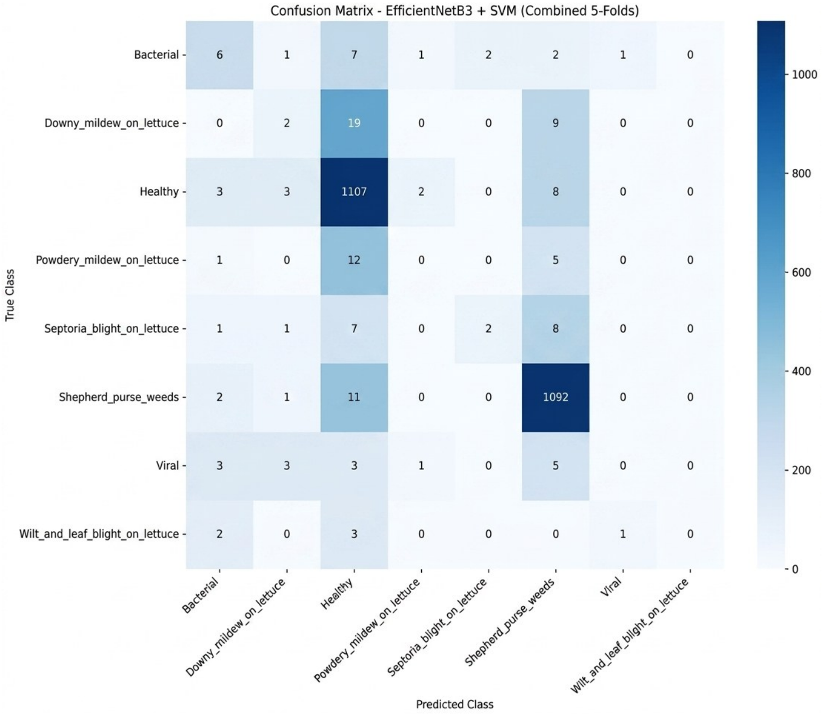

Figure 1. Confusion matrix for the SVM classifier

3.4. Why SVM outperforms RF and KNN

SVM outperformed RF and KNN because EfficientNet-B3 features are high-dimensional (1536-D). SVM with an RBF kernel is known to perform well in such spaces by learning flexible nonlinear decision boundaries. RF, on the other hand, struggles with sparse and high-dimensional feature distributions, leading to overfitting on majority classes and underfitting on rare ones. KNN suffers from the curse of dimensionality, where distance metrics lose discriminative power, making minority class samples effectively indistinguishable in feature space. Therefore, SVM remains the most appropriate choice for hybrid deep feature agricultural disease classification.

3.5. Comparison with previous studies

Previous CNN-based studies, such as the work by Liu *et al.* [23] on apple leaf diseases and Sunil *et al.* [24] on cardamom plants reported accuracies exceeding 95% using end-to-end deep learning models. While highly accurate, these methods typically require substantial computational power for gradient updates across all layers and often rely on balanced datasets. In contrast, this study presents a hybrid approach that utilizes EfficientNet-B3 solely as a fixed feature extractor combined with SVM. This method achieved comparable accuracy (94.01%) to fully trained models but with significantly lower training complexity. This highlights the strength of combining deep feature extraction with classical machine learning classifiers as an efficient alternative pipeline for low-resource agricultural environments where large, balanced datasets are unavailable.

3.6. Limitations and future work

The primary limitation of this study is the extremely imbalanced dataset, where several disease classes have fewer than 20 samples. This restricts the model’s ability to generalize minority disease patterns. Future efforts should incorporate synthetic oversampling (e.g., synthetic minority oversampling technique (SMOTE)), focal loss, generative augmentation, or additional field data acquisition to address this issue. Further improvements can also be achieved by exploring end to end fine tuning of EfficientNet-B3, hierarchical classification, or ensemble hybrid architectures to enhance rare disease detection.

4. CONCLUSION

This study demonstrated that combining EfficientNet-B3 feature extraction with classical machine learning classifiers provides an effective and computationally efficient approach for lettuce disease classification. EfficientNet-B3 successfully extracted high-dimensional representations that captured essential visual characteristics of the leaves, while the SVM achieved the best overall performance with 94%

accuracy and 93.9% weighted F1-score across five folds. However, the low macro F1-score (37.9%) and the confusion matrix analysis revealed that the model struggled to recognize minority disease classes due to extreme dataset imbalance. These findings highlight that accuracy alone is insufficient for evaluating imbalanced agricultural datasets and that macro F1 and per-class performance are critical for assessing true model reliability. Although the hybrid approach performed well for dominant classes such as healthy and Shepherd's purse weeds, it remains limited in detecting rare diseases with fewer available samples. Future work should focus on improving minority class recognition through synthetic oversampling, advanced augmentation strategies, or acquiring additional field data. Further exploration of end-to-end fine-tuning, imbalance-aware loss functions, and larger-scale benchmarks may also enhance generalization. Overall, this study establishes a strong baseline for lightweight agricultural disease classification using deep feature extraction combined with classical classifiers.

ACKNOWLEDGMENTS

The authors would like to thank the Faculty of Computer Science, Brawijaya University, for providing the facilities and support necessary to conduct this research.

FUNDING INFORMATION

This research did not receive direct financial support from any funding agency. However, the research was fully supported through facilities and infrastructure provided by the Intelligent Systems Laboratory, Faculty of Computer Science, Brawijaya University.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	**C**	**M**	**So**	**Va**	**Fo**	**I**	**R**	**D**	**O**	**E**	**Vi**	**Su**	**P**	**Fu**
Fazlur Ihzanurahman	✓	✓	✓	✓	✓	✓		✓	✓		✓			
Wayan Firdaus Mahmudy	✓	✓				✓				✓		✓	✓	

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

The data that support the findings of this study are openly available in Kaggle at https://www.kaggle.com/datasets/ashishjstar/lettuce-diseases/data.

REFERENCES

[1] L. Huang, G. Liu, Y. Wang, H. Yuan, and T. Chen, "Fire detection in video surveillances using convolutional neural networks and wavelet transform," *Engineering Applications of Artificial Intelligence*, vol. 110, 2022, doi: 10.1016/j.engappai.2022.104737.

[2] T. Kalaiselvi, S. T. Padmapriya, K. Somasundaram, and S. Praveenkumar, "E-Tanh: a novel activation function for image processing neural network models," *Neural Computing and Applications*, vol. 34, no. 19, pp. 16563–16575, 2022, doi: 10.1007/s00521-022-07245-x.

[3] P. Tripathi, N. Kumar, M. Rai, and A. Khan, "Applications of deep learning in agriculture," in *Artificial Intelligence Applications in Agriculture and Food Quality Improvement*, 2022, pp. 17–28, doi: 10.4018/978-1-6684-5141-0.ch002.

[4] O. A. D. Ammar *et al.*, "Plant leaf disease detection using ensemble learning and explainable AI," *IEEE Access*, vol. 12, pp. 156038–156049, Oct. 2024, doi: 10.1109/ACCESS.2024.3484574.

[5] S. M. Hassan, A. K. Maji, M. Jasiński, Z. Leonowicz, and E. Jasińska, "Identification of plant-leaf diseases using CNN and transfer-learning approach," *Electronics*, vol. 10, no. 12, 2021, doi: 10.3390/electronics10121388.

[6] L. Liu *et al.*, "Deep learning based automatic multiclass wild pest monitoring approach using hybrid global and local activated features," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 11, pp. 7589–7598, 2021, doi: 10.1109/TII.2020.2995208.

[7] A. Sharma, U. S. Aswal, A. Rana, V. D. Vani, A. Sankhyan, and Shekhar, "Real time plant disease detection model using deep learning," in *2023 6th International Conference on Contemporary Computing and Informatics (IC3I)*, Nagar, India, 2023, pp. 2695–2699, doi: 10.1109/IC3I59117.2023.10398070.

[8] R. Ashtagi, “Fusion of AI techniques: a hybrid approach for precise plant leaf disease classification,” *Journal of Electrical Systems*, vol. 20, no. 1s, pp. 850–861, 2024, doi: 10.52783/jes.836.

[9] L. D. A.-Valencia and S. D. Fassois, “Natural vibration response based damage detection for an operating wind turbine via random coefficient linear parameter varying AR modelling,” *Journal of Physics: Conference Series*, vol. 628, no. 1, pp. 273–297, 2015, doi: 10.1088/1742-6596/628/1/012073.

[10] J. Yao, S. N. Tran, S. Sawyer, and S. Garg, “Machine learning for leaf disease classification: data, techniques and applications,” *Artificial Intelligence Review*, vol. 56, no. s3, pp. 3571–3616, 2023, doi: 10.1007/s10462-023-10610-4.

[11] Z. Jin, J. Shang, Q. Zhu, C. Ling, W. Xie, and B. Qiang, “RFRSF: employee turnover prediction based on random forests and survival analysis,” in *Web Information Systems Engineering – WISE 2020*, Cham, Switzerland: Springer, 2020, pp. 503–515, doi: 10.1007/978-3-030-62008-0_35.

[12] T. M. Cover and P. E. Hart, “Nearest neighbor pattern classification,” *IEEE Transactions on Information Theory,* vol. 13, no. 1, pp. 22-27, 1967, doi: 10.1109/TIT.1967.1053964.

[13] U. Mishra, A. Pandey, G. Logeswari, and K. Tamilarasi, “Deep learning-based disease detection in potato and mango leaves: a comparative study of CNN, AlexNet, ResNet, and EfficientNet,” *Scientific Reports*, pp. 1–21, 2025, doi: 10.1038/s41598-025-32607-5.

[14] J. Andrew, J. Eunice, D. E. Popescu, M. K. Chowdary, and J. Hemanth, “Deep learning-based leaf disease detection in crops using images for agricultural applications,” *Agronomy*, vol. 12, no. 10, pp. 1–20, 2022, doi: 10.3390/agronomy12102395.

[15] M. Tan and Q. V. Le, “EfficientNet: rethinking model scaling for convolutional neural networks,” in *36th International Conference on Machine Learning, ICML 2019*, 2019, pp. 10691–10700.

[16] M. H. Saleem, J. Potgieter, and K. M. Arif, “A performance-optimized deep learning-based plant disease detection approach for horticultural crops of New Zealand,” *IEEE Access*, vol. 10, pp. 89798–89822, 2022, doi: 10.1109/ACCESS.2022.3201104.

[17] A. Lakshmanarao, M. R. Babu, and T. S. R. Kiran, “Plant disease prediction and classification using deep learning ConvNets,” *2021 International Conference on Artificial Intelligence and Machine Vision (AIMV),* Gandhinagar, India, 2021, pp. 1-6, doi: 10.1109/AIMV53313.2021.9670918.

[18] G. Sangar and V. Rajasekar, “Optimized classification of potato leaf disease using EfficientNet-LITE and KE-SVM in diverse environments,” *Frontiers in Plant Science*, vol. 16, no. May, pp. 1–17, 2025, doi: 10.3389/fpls.2025.1499909.

[19] J. O. Adeola, J. Degila, and M. Zennaro, “Recent advances in plant diseases detection with machine learning: solution for developing countries,” in *2022 IEEE International Conference on Smart Computing, SMARTCOMP 2022*, 2022, pp. 374–380, doi: 10.1109/SMARTCOMP55677.2022.00083.

[20] R. Chen, H. Qi, Y. Liang, and M. Yang, “Identification of plant leaf diseases by deep learning based on channel attention and channel pruning,” *Frontiers in Plant Science*, vol. 13, 2022, doi: 10.3389/fpls.2022.1023515.

[21] A. G.-Briones, S. L. Florez, P. Chamoso, L. F. Castillo-Ossa, and E. S. Corchado, “Enhancing plant disease detection: incorporating advanced cnn architectures for better accuracy and interpretability,” *International Journal of Computational Intelligence Systems*, vol. 18, no. 1, 2025, doi: 10.1007/s44196-025-00835-2.

[22] L. Li, S. Zhang, and B. Wang, “Plant disease detection and classification by deep learning - a review,” *IEEE Access*, vol. 9, pp. 56683–56698, 2021, doi: 10.1109/ACCESS.2021.3069646.

[23] B. Liu, Y. Zhang, D. J. He, and Y. Li, “Identification of apple leaf diseases based on deep convolutional neural networks,” *Symmetry*, vol. 10, no. 1, 2018, doi: 10.3390/sym10010011.

[24] C. K. Sunil, C. D. Jaidhar, and N. Patil, “Cardamom plant disease detection approach using EfficientNetV2,” *IEEE Access*, vol. 10, pp. 789–804, 2022, doi: 10.1109/ACCESS.2021.3138920.

[25] A. B. Sidhique, A. Gopakumar, and A. R. Bushara, “Efficient net-based deep learning model for accurate plant disease classification and diagnosis,” *International Journal of Science and Research Archive*, vol. 14, no. 1, pp. 1264–1270, 2025, doi: 10.30574/ijsra.2025.14.1.0170.

[26] J. Liu, F. Lv, and D. Penghui, “Identification of sunflower leaf diseases based on random forest algorithm,” in *Proceedings - 2019 International Conference on Intelligent Computing, Automation and Systems, ICICAS 2019*, 2019, vol. 11, pp. 459–463, doi: 10.1109/ICICAS48597.2019.00102.

BIOGRAPHIES OF AUTHORS

Fazlur Ihzanurahman received his bachelor's degree in Economics from the Department of Development Economics, Universitas Gajayana Malang, in 2022. He obtained his master's degree in Computer Science from the Faculty of Computer Science, Brawijaya University, in 2025. He is currently active in research in the fields of computer vision, deep learning, and data science. His research interests include digital image processing, convolutional neural network (CNN)-based classification, and the application of artificial intelligence in agriculture. He can be contacted at email: fazlurihza@student.ub.ac.id.

Wayan Firdaus Mahmudy obtained a Bachelor of Science degree from the Department of Mathematics, Brawijaya University in 1995. His master in Informatics Engineering degree was obtained from the Institut Teknologi Sepuluh Nopember, Surabaya in 1999 while a Ph.D. in Manufacturing Engineering was obtained from the University of South Australia in 2014. He is a professor at Department of Informatics Engineering, Brawijaya University, Indonesia. His research interests include optimization of combinatorial problems and machine learning. He can be contacted at email: wayanfm@ub.ac.id.