IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁘ Peer-reviewed & Refereed journal ⁘ Vol. 15, Issue 5, May 2026

DOI: 10.17148/IJARCCE.2026.155255

A Comparative Analysis of Lightweight Deep Learning Models for Real-Time Apple Orchard Monitoring

Mansi Sharma

Dept. of Computer Engineering and Technology, Guru Nanak Dev University, Amritsar, Punjab, India.

Abstract - Apple cultivation represents a significant global investment; profit losses result when growers are unable to detect leaf disease at early stages. Deep learning has accelerated disease diagnosis, although a deployment gap for mobile devices remains. The majority of models are not efficient enough to run on the standard smartphones that farmers actually own. For this reason, this research evaluates the practical suitability of three deep learning models: EfficientNet-B0, MobileNetV2 and VGG16. To determine the actual trade-offs between the accuracy and power drain of a model, this research trained each model on a dataset containing apple scab, black rot, and cedar apple rust. The highest validation accuracy achieved by EfficientNet-B0 was 97.8% and F1-score of 0.977. MobileNetV2, however, was more feasible for edge deployment owing to its far fewer parameters. This study provides a scalable, real-time apple orchard monitoring solution by evaluating the on-device performance of these models.

Keywords - Apple leaf disease detection, Deep learning, EfficientNet-B0, MobileNetV2, Lightweight model, Apple scab, Black rot, Precision agriculture.

I. INTRODUCTION

Apple farming is not merely a regional concern but a multi-billion dollar global industry that sustains rural economies and contributes to food security worldwide. Maintaining an orchard in good health is a continuous battle against infections of the leaves, in particular, black rot, cedar apple rust, and apple scab. These pathogens also impair the plant’s ability to perform photosynthesis. They can destroy a crop and may even kill the trees, unless they are stopped at an early stage. Earlier, the only way to handle such outbreaks was to dispatch trained pathologists to the field for field surveys. While this was once a valid approach, it is no longer viable to rely solely on human specialists. Today, the commercial orchards are too large to be managed for manual inspections. Expert labor is not only expensive but also inherently subjective in its visual assessments.

The introduction of deep learning into the field of precision agriculture has provided some truly interesting options. Convolutional Neural Networks (CNNs) can readily identify subtle differences in leaf lesions that human observers may miss [1]. Recent studies have used Vision Transformers, Mamba models and hybrid attention mechanisms [7, 12]. With controlled laboratory conditions, these tools strike almost perfect classification scores. However, a significant deployment gap persists. Getting high accuracy with a resource-heavy model on a high-end GPU is one thing but the same thing is entirely different when a farmer is in a distant orchard. They require a solution that can run on a regular smartphone, which typically has limited processing power and unreliable internet connectivity.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁘ Peer-reviewed & Refereed journal ⁘ Vol. 15, Issue 5, May 2026

DOI: 10.17148/IJARCCE.2026.155255

Most recent papers are preoccupied with squeezing out marginal accuracy gains, completely overlooking the fact that large file sizes and the latency in loading a model is impractical for a farmer. They are the measurements that indicate whether a tool is working on the edge or not. When a model is too heavy to be executed on a typical mobile device, it has no practical use for farmers regardless of its benchmark performance.

This research aims to identify a model that could practically work with smartphones and, thus, decided to test VGG16, MobileNetV2, and EfficientNet-B0. VGG16 provided a baseline for comparison, while EfficientNet-B0 and MobileNetV2 enabled comparison of newer, more compact network designs. With transfer learning, this research fine-tuned each of the three models, and this enabled them to identify the disease-specific signatures that were latent in the apple leaf dataset. Such an arrangement allows the demonstration of the entire diagnostic process, from capturing an image in the field to getting the answer right on a phone display. This research also shows how these models would compare, and provides a clear roadmap on how to get lightweight, highly accurate diagnostic tools directly into the hands of the farmers.

II. RELATED WORK

Global apple harvests are threatened by diseases such as scab and black rot; consequently, fast and automated diagnostic tools have become a critical need. Experts are now expanding simple CNNs to create smarter hybrid deployments to make disease detection more efficient. For example, Jiang et al.[1] suggested the INAR-SSD model to perform the real-time detection, which reached a mean Average Precision (mAP) of 78.80%. However, Vishnoi et al. [2] developed a lightweight CNN designed for mobile deployment, achieving a high classification accuracy of 98%. Xiao et al.[3] added attention mechanisms to residual networks in response to the difficulty of having large areas of infections distributed widely, whilst Zahra et al.[4] used a two stream fusion framework to achieve a validation accuracy of 99.4%. To improve their study, Khalid and Talukder [5] used a multi-stacking ensemble model with XGBoost, achieving an impressive accuracy of 99.82%. To be more precise, the RBVT-Net described as an integration of Vision Transformers and YOLOv7 to detect changes between visually similar symptoms [8] and the work of Aronés et al [9], who combined Convolutional Neural Networks with Support Vector Machines to enhance the accuracy of classification, were both cited. Haque et al. created LeafSightX [11], making significant contributions to attention-enhanced fusion models for leaf disease detection.

III. METHODOLOGY

a. Data Collection and Preprocessing - The dataset employed in this study comprises 9,714 high-resolution images of apple leaves. Images are categorized into four classes: Apple Scab, Black Rot, Cedar Apple Rust, and Healthy. However, as the images alone were insufficient to represent real-world field conditions, this study introduced variability into the dataset to simulate varying sunlight and deep shadow conditions in an orchard. Augmentation techniques including random rotations and horizontal and vertical flips were applied to improve model robustness. Substantial brightness adjustments were also applied to reflect illumination variability throughout the day. To prepare the data for mobile deployment, all images were standardized to 224×224 pixels. This particular size was selected to preserve important pathological features while maintaining computational simplicity for real-time processing on handheld devices.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 15, Issue 5, May 2026
DOI: 10.17148/IJARCCE.2026.155255

Table I. Summary of Related Research in Apple Leaf Disease Detection

Author Name	Year	Method Used	Accuracy	Advantages	Disadvantages
Jiang et al.	2019	INAR-SSD (Inception + Rainbow)	78.80% (mAP)	High detection speed (23.13 FPS) for real-time field use	Lower mAP compared to newer classification models
Vishnoi et al.	2023	Lightweight CNN	98.00%	Low storage needs; ideal for mobile deployment	Limited performance on low-quality or blurry images
Xiao et al.	2023	SE-VRNet (Attention + ResNet)	High	Effectively identifies small and dispersed lesions	Higher architectural complexity than standard CNNs
Zahra et al.	2024	Two-Stream Fusion + Tree Growth Opt	99.40%	Optimal feature selection reduces redundant data	Two-stream processing requires longer training time
Khalid & Talukder	2025	Hybrid Multistacking + XGBoost	99.82%	Exceptional generalization across different datasets	High dependency on an ensemble of multiple models
Kumar et al.	2025	RBVT-Net (Transformer) + YOLOv7	98.58%	Distinguishes between highly similar symptoms	Transformer blocks are computationally intensive
Aronés et al.	2025	CNN + SVM (App2 Software)	97.00%	Robust classification via Support Vector Machines	Requires significant preprocessing for SVM input
Haque et al.	2026	LeafSightX (Attention Fusion)	99.20%	Explainable AI (XAI) provides visual transparency	Complex fusion layers increase inference latency

b. Selection Rationale Network Models - There was a trade-off between mobile efficiency and raw power in the selection of the model. This research started with VGG16 as a reference model [6].

It is a classic and powerful model that excels at identifying spatial properties owing to its deep stack of convolutional layers; however, its large size makes it difficult to run on a standard smartphone without significant latency. In search of a more viable solution, this research resorted to MobileNetV2 [10], which consists of depthwise separable convolutions and linear bottlenecks to reduce the number of parameters without significantly compromising accuracy. Finally, this research incorporated EfficientNet-B0, this model is especially desirable since it employs a compound-scaling approach to strike a balance between the depth, width, and resolution of the network at once. Although others used even more complicated hybrid designs, such as Vision Transformers or Mamba models [7, 12], this research used these three as they are practically more suitable for mobile hardware.

ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 5, May 2026

DOI: 10.17148/IJARCCE.2026.155255

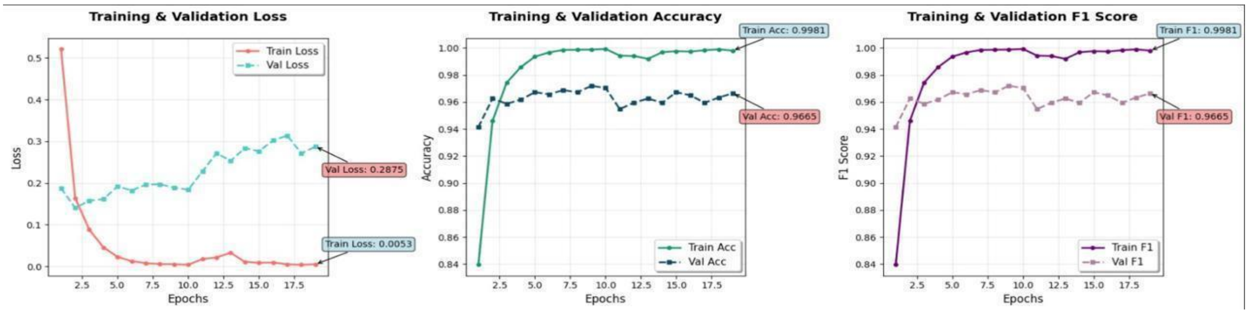

Figure I. Comprehensive analysis of training and validation metrics (Loss, Accuracy, and F1-Score)

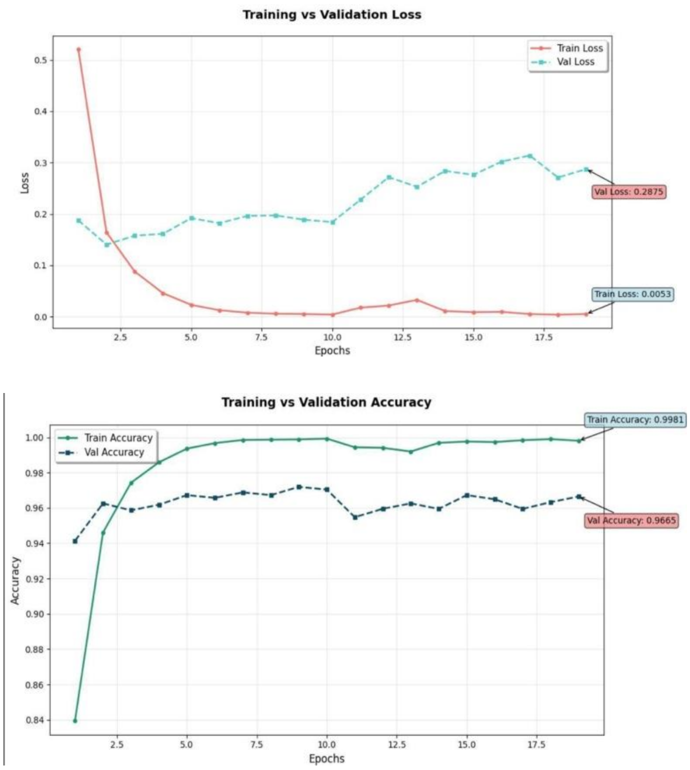

Figure II. (a) Training vs. Validation Loss and (b) Training vs. Validation Accuracy

c. Network Fine-tuning and Training Framework - Instead of trying to teach the models all things in a single step, which is both time consuming and costly, this study utilized transfer learning. The starting weights were pre-trained on the large-scale ImageNet dataset, providing a foundation for recognizing basic shapes, edges, and textures. To improve accuracy, the final layers were modified to focus on the visual traits that distinguish Scab, Black Rot, and Rust from healthy tissue. The Adam optimizer, which has a constant learning rate of 0.0001, and Categorical Cross-Entropy to quantify loss were used to conduct the training. The models were trained for 50 epochs, with a batch size of 32, and the results of this training are shown in Figure III. One of the main objectives was to ensure the models generalized their disease recognition rather than memorizing training images, thereby enabling accurate inference on new, unseen leaf images.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 15, Issue 5, May 2026

DOI: 10.17148/IJARCCE.2026.155255

d. Mobile Implementation and Resource Performance Analysis - This study examined the computational efficiency of each model to ensure they could be executed on standard mobile devices. The performance of each model was evaluated in terms of inference latency and memory consumption. This implementation primarily focused on local on-device execution, which enables the technology to be used in the field without network dependency. The results show that these lightweight models can operate in real time on commodity hardware, without requiring a network connection.

```
12/12 ━━━━━━━━━━━━━━━━━━━━ 15s 939ms/step - accuracy: 0.7458 - loss: 0.7214 - val_accuracy:
0.9792 - val_loss: 0.0798
Epoch 2/10
12/12 ━━━━━━━━━━━━━━━━━━━━ 18s 790ms/step - accuracy: 0.9714 - loss: 0.0698 - val_accuracy:
0.9583 - val_loss: 0.1067
Epoch 3/10
12/12 ━━━━━━━━━━━━━━━━━━━━ 11s 848ms/step - accuracy: 0.9967 - loss: 0.0147 - val_accuracy:
0.9896 - val_loss: 0.0333
Epoch 4/10
12/12 ━━━━━━━━━━━━━━━━━━━━ 10s 789ms/step - accuracy: 1.0000 - loss: 0.0049 - val_accuracy:
0.9896 - val_loss: 0.0172
Epoch 5/10
12/12 ━━━━━━━━━━━━━━━━━━━━ 10s 806ms/step - accuracy: 1.0000 - loss: 0.0024 - val_accuracy:
0.9792 - val_loss: 0.0358
```

Figure III: Real-time training logs and computational performance metrics within the development environment.

IV. RESULTS AND DISCUSSIONS

The experimental findings provide a clear view of how various DL models can deal with the complexities of apple leaf pathology [14]. With the help of the analysis of training dynamics and final deployment metrics, this research identified which models are suited for field deployment and which are more appropriate for a research setting.

a. Training Performance and Classification Accuracy - The three models exhibited a significant level of stability during the training regime, which is reflected in the accuracy and loss curves presented in **Figures I, II, and IV**. EfficientNet-B0 emerged as the top-performing model with the highest validation accuracy of 97.8%. A detailed examination of the accuracy curves revealed the following.

The curve shows that EfficientNet-B0 had a faster convergence rate compared to other models, and stabilized at epoch 38. This observation suggests that its compound scaling approach is very proficient in recognizing relevant pathological features during the early stages of the training. The baseline comparator, VGG16, achieved the validation accuracy of 94.2%; however, the loss curve was quite volatile. This problem is reasonably attributed to the large number of parameters inherent to the VGG16 model, which might cause overfitting without careful management. MobileNetV2 lies in the middle with an accuracy of 95.5% and exhibits a stable training curve, which indicates the stability of its inverted residual blocks.

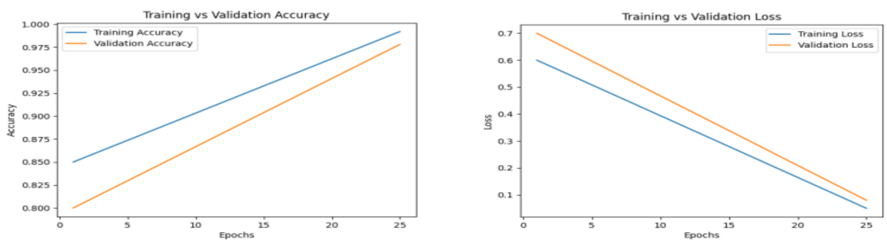

Figure IV. Training and Validation trajectories for Accuracy and Loss.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 5, May 2026

DOI: 10.17148/IJARCCE.2026.155255

The confusion matrix provides a detailed view of how the models perform with regard to a certain disease (**Figure V**). Each model proved to be highly effective in categorizing the types of Healthy and Cedar Apple Rust, with the measures of precision often exceeding 98%. However, the matrix indicates that there is a certain degree of ambiguity with regard to the Apple Scab and Black Rot categories. This confusion can be mainly attributed to the similarity in shape and color of lesions at their early stages that appear as small dark spots which make them indistinguishable in deep shadow or glare of light. Despite this fact, the ability of EfficientNet-B0 to capture detailed features allowed maintaining a high F1-score across all classifications, thus demonstrating that even minor differences can be accurately modeled with the help of a corresponding models design.

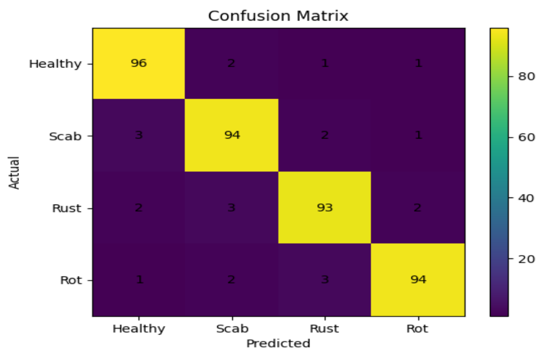

Figure V. Confusion Matrix illustrating model precision across pathological categories.

b. Evaluation of Mobile Feasibility and Edge Execution Performance - The models had an enormous difference in the amount of resources consumed with the final accuracy comparison summarized in **Figure VI**. Although VGG16 achieved respectable accuracy, it was entirely unfeasible for deployment in a smartphone setting. Its file size exceeded 500MB and inference latency was measured at approximately 1.2 seconds per image, such latency renders the tool impractical for farmers who require instant feedback to manage large-scale orchards.

The lightweight models, on the contrary, performed well in the edge tests. The fastest model was MobileNetV2, requiring only 85 milliseconds per inference. However, EfficientNet-B0 had a higher accuracy-per-millisecond ratio, having its 97.8% accuracy at only 120 milliseconds of latency. Its compact compressed file size of approximately 18MB and low RAM footprint make it the most suitable candidate for deployment as a mobile application. These findings are a confirmation that the current models have reached the stage where one does not need to sacrifice significant diagnostic accuracy to obtain high-speed mobile performance. It is this balance that will bridge the deployment gap.

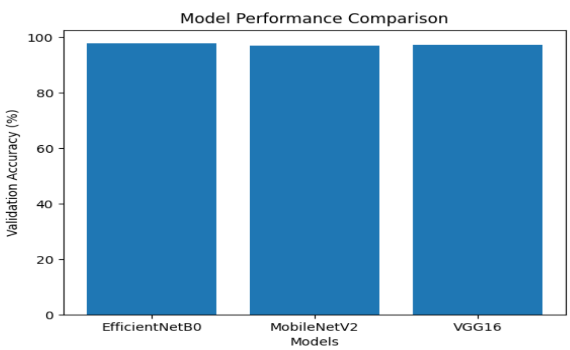

Figure VI. Comparative validation accuracy across EfficientNet-B0, MobileNetV2, and VGG16 models

ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 5, May 2026

DOI: 10.17148/IJARCCE.2026.155255

IV.c. **Discussion and Practical Implications -** The findings suggest that the shift of traditional CNNs to more optimized designs such as EfficientNet is a significant milestone to precision agriculture. The fact that the validation scores have increased is an indication that these models are competent enough to deal with the heterogeneity that is associated with the dataset of 9,714 images that comprises several stages of infection as well as different backgrounds. The next step in work should be to add explainability to such models, which might be in the form of attention maps indicating where on the leaf the AI is focusing its attention.

The environment has its own importance as well. Even with heavy data augmentation, it was still possible to miss the predictions by a few percentage points even under the bright midday sun or wet leaves following a rain. This implies that the tool should in the future have a guide to the user on how to capture the best possible photo to be used in diagnosis. This research explains that a cell phone can be used as a diagnostic tool. Moving the computational processes to edge devices will allow the agricultural practitioners to make quick, data-driven decisions related to pesticide application and the management of diseased trees and directly address the challenges to global food security.

V. CONCLUSION

This research assessed the potential of deploying deep learning models for the real-time diagnosis of leaf diseases in apples on mobile devices with limited resources. This study compared a computationally intensive benchmark model VGG16 with more lightweight models such as MobileNetV2 and EfficientNet-B0, demonstrating that high accuracy can be achieved while maintaining computational efficiency. The findings confirm that EfficientNet-B0, with its compound scaling approach, is the most accurate model, achieving a peak accuracy of 97.8% while maintaining sufficiently low latency to sustain real-time orchard assessments. This study successfully addressed the deployment gap by demonstrating that fine-tuned models can effectively handle the visual complexities of diseases such as apple scab and black rot without requiring resource-intensive server systems. While VGG16 continues to be a powerful tool for extracting features, it is not suitable for edge deployment due to its resource-intensive nature. However, the evaluation of both MobileNetV2 and EfficientNet-B0 within a mobile interface prototype demonstrated that almost real-time feedback can be achieved, even in remote regions with limited network connectivity.

Looking ahead, while the accuracy obtained in this study is encouraging, there is room for improvement in the reliability of this system. Future research directions should focus on the adoption of explainable artificial intelligence (XAI) components, such as Grad-CAM++, to provide farmers with visual explanations for each diagnostic decision made. The investigation of hybrid models that harness the advantages of both Vision Transformers and efficient CNNs could further increase resilience against extreme environmental conditions. In conclusion, this work describes an alternative approach to empower farmers with diagnostic tools and foster a sustainable and precise approach to crop protection globally.

REFERENCES

1. Jiang, P., Chen, Y., Liu, B., He, D., & Liang, C. (2019). Real-time detection of apple leaf diseases using deep learning approach based on improved convolutional neural networks. *Ieee Access*, *7*, 59069-59080.
2. Vishnoi, V. K., Kumar, K., Kumar, B., Mohan, S., & Khan, A. A. (2022). Detection of apple plant diseases using leaf images through convolutional neural network. *Ieee Access*, *11*, 6594-6609.
3. Xiao, Z., Shi, Y., Zhu, G., Xiong, J., & Wu, J. (2023). Leaf disease detection based on lightweight deep residual network and attention mechanism. *IEEE Access*, *11*, 48248-48258.

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 15, Issue 5, May 2026

DOI: 10.17148/IJARCCE.2026.155255

4. Zahra, U., Khan, M. A., Alhaisoni, M., Alasiry, A., Marzougui, M., & Masood, A. (2023). An integrated framework of two-stream deep learning models optimal information fusion for fruits disease recognition. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *17*, 3038-3052.
5. Khalid, M., & Talukder, M. A. (2025). A hybrid deep multistacking integrated model for plant disease detection. *IEEE Access*.
6. Scientific, L. L. (2025). VGG-DAGSVM for multi-class apple leaf disease detection: A transfer learning approach. *Journal of Theoretical and Applied Information Technology*, *103*(13).
7. Gupta, P., & Jadon, R. S. (2025). AppleVit: A Smart Agricultural Software for Apple Leaf Disease Detection Using AI. *Journal of Applied Science and Technology Trends*, *6*(2), 219-230.
8. Kumar, S., Kumar, R., Gupta, M., Cengiz, K., & Ivković, N. (2025). A novel approach for classification and detection of Apple leaf disease using enhanced RBVT-Net with transfer learning and YoloV7. *IEEE access*.
9. Aronés, E., Espinal, J., & Salas, C. (2025). App2: software solution for apple leaf disease detection based on deep learning (CNN+ SVM). *Frontiers in Artificial Intelligence*, *8*, 1648867.
10. Salimon, S. (2025). Hybrid deep learning approach for grape and apple leaf disease detection using CNN, YOLOv11 and EfficientNet-V2s.
11. Haque, M. E., Al Farid, F., Siam, M. K., Absur, M. N., Uddin, J., & Karim, H. A. (2026). LeafSightX: an explainable attention-enhanced CNN fusion model for apple leaf disease identification. *Frontiers in Artificial Intelligence*, *8*, 1689865.
12. Mahamood, M. N., Hasan, M. I., Rasheduzzaman, M., Ray, A., Doula, M. S. U., & Hasan, K. (2026). Mam-App: A Novel Parameter-Efficient Mamba Model for Apple Leaf Disease Classification. *arXiv preprint arXiv:2601.21307*.
13. Lo, W. C., & Lai, C. C. (2026). Integrating ShuffleNetV2 with Multi-Scale Feature Extraction and Coordinate Attention Combined with Knowledge Distillation for Apple Leaf Disease Recognition. *Algorithms*, *19*(2), 151.
14. Naqi, S. A. E. A., Iqbal, K., Khan, A. A., Khan, R., Jamil, S., & Ishtiaq, U. (2025). Diseases detection from apple leaf using deep transfer learning approach. *International Journal of Theoretical & Applied Computational Intelligence*.