A Hybrid EfficientNetV2S–ResNet50 Model for the Accurate Classification of Plant Leaf Diseases

Fatima-Zahra Fandi
Department of Mathematics and Computer Science, Faculty of Sciences Ben M'Sik, Hassan II University of Casablanca, Morocco
Fatimazahra.fandi10@gmail.com (corresponding author)

Mohamed Ghazouani
Department of Mathematics and Computer Science, Faculty of Sciences Ben M'Sik, Hassan II University of Casablanca, Morocco
ghzouani.fsbm@gmail.com

Mohamed Azzouazi
Department of Mathematics and Computer Science, Faculty of Sciences Ben M'Sik, Hassan II University of Casablanca, Morocco
azouazii@gmail.com

Received: 15 December 2025 | Revised: 5 January 2026, 25 January 2026, 15 February 2026, 20 February 2026, and 26 February 2026 | Accepted: 27 February 2026

 DOI: https://doi.org/10.48084/etasr.16975

ABSTRACT

Accurate classification of plant leaf diseases is significant for precision agriculture. This study details the development and evaluation of a hybrid Deep Learning (DL) architecture integrating EfficientNetV2S and ResNet50 to balance accuracy and computational efficiency. The original 8,000-image PlantVillage dataset was augmented to 28,000 images to address class imbalance and enable robust training. The dataset was partitioned into 20,000 images for training, 4,000 for validation, and 4,000 for testing. The hybrid model successfully leveraged the complementary strengths of both components, achieving a final classification accuracy of 97.16%, thereby demonstrating its suitability for real-time agricultural monitoring. This work confirms the potential of feature-fusion strategies for practical, early-stage disease screening (healthy versus diseased).

Keywords-plant disease classification; deep learning; EfficientNetV2S; ResNet50; hybrid model; precision agriculture

I. INTRODUCTION

Plant diseases result in great economic losses and threaten food security; however, traditional visual inspection is slow, error-prone, and difficult to scale. Consequently, Artificial Intelligence (AI) and Deep Learning (DL) are used to automate and improve the accuracy of plant disease detection. Convolutional Neural Networks (CNNs) are widely deployed for visual recognition in smart agriculture [1, 2], but have limited suitability to real-time applications in low-resource environments. The automated detection of plant diseases via DL has evolved from basic classification to optimized deployment. Authors in [3, 4] showed that CNN architectures, particularly those with residual blocks like ResNet, capture complex hierarchical features effectively. To address the constraints of real-time agricultural applications, research shifted toward efficiency. Authors in [2] reviewed hybrid architectures, combining multiple CNN backbones, for efficient real-time monitoring. Thus, the current study develops a hybrid EfficientNetV2S-ResNet50 model to bridge the gap between high-precision DL and the resource limitations of precision agriculture.

The present study employs binary classification (healthy versus diseased) as a crucial first-stage screening task. EfficientNetV2S and ResNet50 were used for their computational efficiency and efficient feature extraction, aiming to develop a hybrid strategy. This approach is designed to deliver accurate, timely predictions with minimal resource usage, which is significant for modern agricultural systems operating under time-sensitive conditions.

II. METHODOLOGY

A. Model Definitions

1) EfficientNetV2S

The EfficientNet family utilizes a compound scaling method to proportionally adjust network depth, width, and resolution [5]. EfficientNetV2S further optimizes this via fused MBConv layers to accelerate convergence [6].

2) ResNet50

ResNet50 employs residual learning to address the vanishing gradient and degradation problems inherent in very deep networks. By utilizing identity shortcut connections, the model ensures stable gradient propagation across 50 layers [7].

3) Proposed Hybrid Model (EfficientNetV2S + ResNet50)

The proposed hybrid model utilizes feature-level fusion to integrate the complementary strengths of both backbones: the computational efficiency of EfficientNetV2S and the deep discriminative representations of ResNet50. Such fusion strategies have demonstrated superior efficiency and generalization compared to single-architecture models [8], providing a balanced solution suitable for real-time agricultural monitoring systems.

B. Dataset Description

1) PlantVillage Dataset

This study utilizes the PlantVillage dataset [9, 10], which contains high-quality images of healthy and diseased leaves captured under controlled conditions. To align the data with this work's binary classification objective, the original multi-class specimens across various crop species were reorganized into two categories:

- Diseased: All specific disease types consolidated into a single class to facilitate early-stage screening.
- Healthy: Images of asymptomatic leaves, maintained as the control group.

2) Class Imbalance Handling

The initial dataset exhibited class imbalance, with the 'Healthy' class being underrepresented. To prevent biased learning and improve generalization, a targeted minority class augmentation strategy was applied using classical image processing techniques. While advanced diffusion-based methods, such as DDPM, are reserved for future work, the current augmentation steps enhance the dataset's representativeness and the model's efficiency against variations in lighting, orientation, and scale.

TABLE I. BALANCED DATASET COMPOSITION

Class	Number of images
Diseased	14,000
Healthy	14,000
Total	28,000

3) Final Dataset Composition and Splitting

After preprocessing and augmentation, the final dataset consisted of 28,000 images, evenly distributed across the two considered classes.

C. Dataset Splitting

The dataset was partitioned into training, validation, and testing sets using a 70/15/15 ratio. Following class balancing and preprocessing, the final test set comprised 4,000 images. This balanced distribution ensures that the model evaluation is not biased toward a specific class and reflects performance across a statistically significant sample.

D. Preprocessing and Training

To ensure architectural compatibility and numerical stability, images were rescaled to 224×224 pixels and their intensity values were normalized to the [0,1] interval. RGB format was retained to preserve diagnostic color features, essential for disease identification. During training, online augmentation, including horizontal flipping and random rotations, was employed to mitigate overfitting and enhance model generalization.

TABLE II. TRAINING CONFIGURATION

Parameter	Value
Loss function	Binary cross-entropy
Batch size	32
Epochs	20
Early stopping	Yes
Learning rate reduction	Yes

III. RESULTS

A. EfficientNetV2S Model Implementation

EfficientNetV2S was initialized with pretrained ImageNet weights, employing a transfer learning strategy in which the convolutional backbone was frozen and fine-tuning was restricted to the terminal classification layers. This approach adapted the architecture for binary classification while preserving the pre-learned feature representations. Optimization was performed using the Adam algorithm, integrated with a learning rate scheduler and an early stopping mechanism to ensure stable convergence and mitigate overfitting.

Evaluation Metrics (Test Set):

- Test Loss: 0.3034
- Test Accuracy: 0.8976
- AUC-ROC: 0.9606

TABLE III. CLASSIFICATION REPORT

Class	Precision	Recall	F1-score	Support
Healthy	0.89	0.91	0.90	2000
Diseased	0.91	0.88	0.90	2000
Macro avg	0.90	0.90	0.90	4000
Weighted avg	0.90	0.90	0.90	4000

B. ResNet50 Model Implementation

ResNet50 was utilized as a static feature extractor by freezing its pretrained ImageNet weights. The architecture was augmented with a custom classification head comprising a Global Average Pooling (GAP) layer, a 128-unit Fully Connected (FC) layer with ReLU activation, and a dropout layer (p=0.5) for regularization. Final class probabilities were generated via a single-unit sigmoid output. To ensure effective performance, the training process incorporated early stopping and an automated learning rate reduction policy, optimizing the trade-off between convergence speed and generalization.

Evaluation Metrics (Test Set):

- Test Loss: 0.1931
- Test Accuracy: 0.9412
- AUC-ROC: 0.9872

TABLE IV. CLASSIFICATION REPORT:

Class	Precision	Recall	F1-score	Support
Healthy	0.98	0.90	0.94	2000
Diseased	0.91	0.98	0.94	2000
Macro avg	0.94	0.94	0.94	4000
Weighted avg	0.94	0.94	0.94	4000

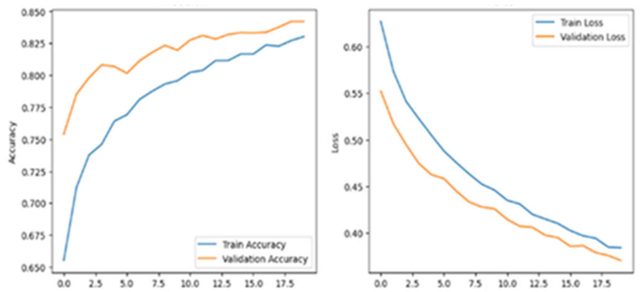

Fig. 1. Training and validation accuracy and loss vs epochs curves for the EfficientNetV2S Model.

C. Proposed Hybrid Model (EfficientNetV2S + ResNet50)

This study introduces a hybrid DL architecture integrating EfficientNetV2S and ResNet50 to optimize the efficiency of plant disease detection. This dual-stream approach leverages the synergistic relationship between the parameter-efficient feature extraction of EfficientNetV2S and the deep, discriminative representations of ResNet50. Architecturally, input images are processed through parallel backbones; the resulting latent feature maps are compressed via GAP and projected into a unified embedding space through FC layers. These feature vectors are subsequently concatenated and processed by a unified classification head to generate the final binary output. The main steps of the hybrid design are:

Feature extraction:

$$f_E = \varphi_E(x),\ f_R = \varphi_R(x) \quad (1)$$

where φ_E and φ_R denote EfficientNetV2S and ResNet50 feature extractors, respectively, and x is the input leaf image. Dimensionality reduction (GAP and projection):

$$z_E = FC_{E(GAP(f_E))}, z_R = FC_{R(GAP(f_R))} \quad (2)$$

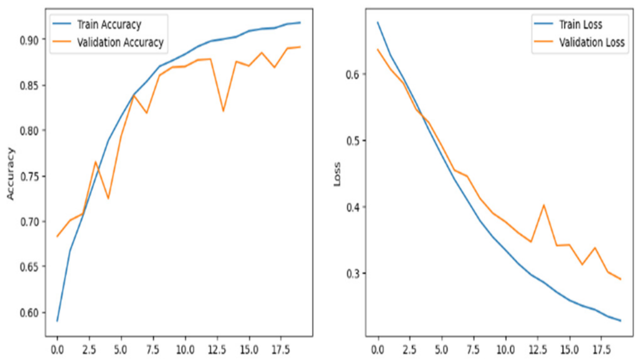

Fig. 2. Training/validation accuracy and loss curves for ResNet50.

These transformations compress the deep feature maps into lower-dimensional vectors while preserving discriminative information.

Feature fusion:

$$z = [z_E;\ z_R] \quad (3)$$

The two vectors are concatenated to form a joint representation that captures complementary patterns from both networks.

Final classification:

$$\{\widehat{y}\} = \sigma(Wz + b) \quad (4)$$

where (σ) is the sigmoid activation producing the probability that a leaf is diseased.

The concatenation of feature vectors (z_E and z_R) utilizes both lightweight and deep-level information, enhancing the separability [11] between healthy and diseased samples, especially in visually ambiguous cases.

The proposed hybrid model achieved superior performance metrics compared to the individual backbones. This improvement likely stems from the complementary nature of the fused features, which capture both global context and fine-grained texture.

$$\text{Accuracy} = \frac{(\text{TP} + \text{TN})}{(\text{TP} + \text{TN} + \text{FP} + \text{FN})} \quad (5)$$

$$\text{Precision} = \frac{\text{TP}}{(\text{TP} + \text{FP})} \quad (6)$$

$$\text{Recall} = \frac{\text{TP}}{(\text{TP} + \text{FN})} \quad (7)$$

$$\text{F1} = 2 * \frac{(\text{Precision} * \text{Recall})}{(\text{Precision} + \text{Recall})} \quad (7)$$

where:

- TP = True Positives
- TN = True Negatives
- FP = False Positives
- FN = False Negatives

The hybrid model achieved a test accuracy of 0.9716 and an AUC-ROC of 0.9956, outperforming both standalone EfficientNetV2S (0.8976) and ResNet50 (0.9412). This confirms the effectiveness of combining multi-scale and deep residual features for plant disease detection.

Evaluation Metrics (Test Set):

- Test Loss: 0.0869
- Test Accuracy: 0.9716
- AUC-ROC: 0.9956

TABLE V. CLASSIFICATION REPORT

Class	Precision	Recall	F1-score	Support
Healthy	0.98	0.97	0.97	2000
Diseased	0.97	0.98	0.97	2000
Macro avg	0.97	0.97	0.97	4000
Weighted avg	0.97	0.97	0.97	4000

IV. DISCUSSION

The combination of EfficientNetV2S's low computational cost and ResNet50's strengthened feature discrimination significantly enhances model robustness against noise and variability. This outperforms individual networks by over 3% in accuracy and 2% in F1-score. Compared to prior studies [4, 12], which primarily addressed multi-class categorization, the proposed binary screening framework focuses on a distinct, practical objective: early-stage detection (healthy versus diseased). Although direct comparison with multi-class benchmarks is not equivalent due to task differences, the 97.16% accuracy demonstrates competitive performance for binary screening.

The obtained results highlight the advantage of the feature fusion strategy. By concatenating the feature vectors (z_E and z_R), the hybrid model successfully combines both lightweight feature extraction and deep-level information. This synergy results in improved efficiency against variations in color, lighting, and leaf texture while maintaining computational efficiency, which is necessary for scalable, real-time agricultural monitoring systems.

Prediction errors mainly occur in visually ambiguous cases, specifically where disease symptoms are subtle or mimic healthy foliage patterns.

TABLE VI. PERFORMANCE COMPARISON SUMMARY

Model	Test accuracy	AUC-ROC	F1-score
EfficientNetV2S	0.8976	0.9606	0.90
ResNet50	0.9412	0.9872	0.94
Hybrid (proposed)	0.9716	0.9956	0.97

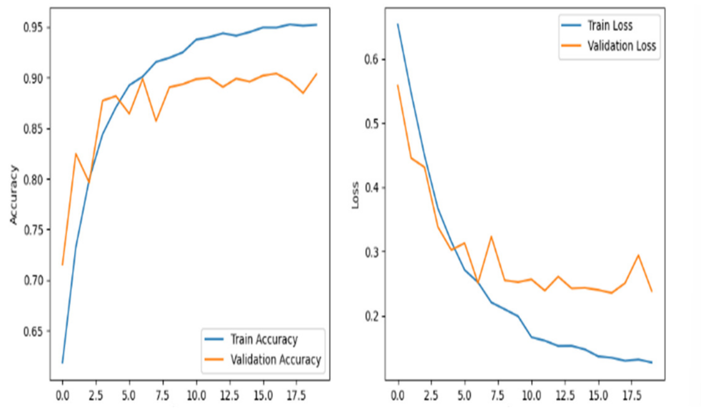

Fig. 3. Training/validation accuracy and loss curves for the hybrid model.

V. CONCLUSION AND FUTURE WORK

This study developed and validated a hybrid EfficientNetV2S–ResNet50 architecture for binary plant leaf disease classification. By employing feature-level fusion, the model achieved superior performance compared to its standalone components, confirming the benefit of combining parameter-efficient and deep residual feature extraction. This framework prioritizes deployment feasibility, offering a balanced solution that maintains high predictive performance for edge deployment in resource-constrained agricultural settings. Future research will address the study's primary limitation: the use of controlled laboratory images. This will be achieved by focusing on three key areas: extending the model to multi-class disease categorization, performing cross-domain validation using real-field imagery, and integrating Explainable Artificial Intelligence (AI) techniques, such as Grad-CAM, to enhance model transparency.

REFERENCES

[1] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, Feb. 2017, https://doi.org/10.1145/3065386.

[2] F. F. Zahra, G. Mohamed, and A. Mohamed, "Plant Health—Detecting Leaf Diseases: A Systematic Review of the Literature," in *Engineering Applications of Artificial Intelligence*, A. Chakir, J. F. Andry, A. Ullah, R. Bansal, and M. Ghazouani, Eds. Springer Nature Switzerland, 2024, pp. 219–235.

[3] S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using Deep Learning for Image-Based Plant Disease Detection," *Frontiers in Plant Science*, vol. 7, Sept. 2016, https://doi.org/10.3389/fpls.2016.01419.

[4] K. P. Ferentinos, "Deep learning models for plant disease detection and diagnosis," *Computers and Electronics in Agriculture*, vol. 145, pp. 311–318, Feb. 2018, https://doi.org/10.1016/j.compag.2018.01.009.

[5] M. Tan and Q. V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks." arXiv, Sept. 11, 2020, https://doi.org/10.48550/arXiv.1905.11946.

[6] M. Tan and Q. Le, "EfficientNetV2: Smaller Models and Faster Training," in *Proceedings of the 38th International Conference on Machine Learning*, July 2021, pp. 10096–10106.

[7] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016, pp. 770–778, https://doi.org/10.1109/CVPR.2016.90.

[8] Z.-H. Zhou, *Ensemble Methods: Foundations and Algorithms*. New York, NY, USA: Chapman and Hall/CRC, 2012.

[9] S. P. Mohanty, "PlantVillage-Dataset", Available: https://github.com/spMohanty/PlantVillage-Dataset/tree/master/raw/color.

[10] S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using Deep Learning for Image-Based Plant Disease Detection," *Frontiers in Plant Science*, vol. 7, Sep. 2016, https://doi.org/10.3389/fpls.2016.01419.

[11] M. Sokolova and G. Lapalme, "A systematic analysis of performance measures for classification tasks," *Information Processing & Management*, vol. 45, no. 4, pp. 427–437, July 2009, https://doi.org/10.1016/j.ipm.2009.03.002.

[12] E. C. Too, L. Yujian, S. Njuki, and L. Yingchun, "A comparative study of fine-tuning deep learning models for plant disease identification," *Computers and Electronics in Agriculture*, vol. 161, pp. 272–279, June 2019, https://doi.org/10.1016/j.compag.2018.03.032.