

A Hybrid Lightweight Convolutional Neural Network Model for Image-Based Robusta Coffee Leaf Diseases Classification

Purwohudiono¹, Subiyanto^{2*}, Nektar Cahayasabda³, Galang Putra Prasetya⁴,
Dhitsa Anggara Ari Samasta⁵

^{1,2,3,4,5}Departement of Electrical Engineering, Faculty of Engineering, Universitas Negeri Semarang,
Jawa Tengah, Indonesia

ABSTRACT

Open Access

Type: Research Paper

Article:

Received: March 31, 2026

Revised: June 18, 2026

Accepted: August 15, 2026

Issued: October 30, 2026

© Purwohudiono et al. (2026).



This is an open-access article
under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license

*Correspondence Address:

subiyanto@mail.unnes.ac.id

Convolutional Neural Networks (CNNs) have been widely used for the classification of coffee leaf diseases. However, the resource-intensive nature of conventional CNNs limits their feasibility for coffee leaf disease identification on edge devices in farming scenarios. To achieve a lightweight, high-accuracy CNN model, this paper develops an hybrid MobileNetV2-NASNetMobile for classifying Robusta coffee leaf diseases. The model integrates two lightweight CNNs: MobileNetV2, which employs depthwise separable convolutions for efficiency, and NASNetMobile, which optimizes discriminative feature extraction. Both branches receive input images resized to 224×224 pixels. A feature-level integration via concatenation is then used to combine their independent feature representations, thereby enhancing model robustness and generalization. The training and evaluation of the implemented method were conducted using a dataset of 3,678 Robusta coffee leaf images, sourced from publicly available datasets and field photographs. Model performance was assessed using 5-fold cross-validation to obtain a reliable estimate of generalization and to mitigate bias from a single train-test split. The implemented method attained an accuracy of 99.11%, with corresponding precision of 99.05%, recall of 98.97%, and F1-score of 99.00%. Furthermore, expert validation yielded 85% agreement, a discrepancy from the 99.11% test accuracy attributable to curated test conditions versus real-world variability in lighting, symptom severity, and diagnostic subjectivity. These results nonetheless demonstrate the model's effectiveness as a scalable solution for Robusta coffee leaf disease classification on resource-limited devices.

Keywords: *Coffee leaf disease, Deep learning, Image classification, Lightweight model*

1. INTRODUCTION

Convolutional Neural Networks (CNNs) have emerged as a powerful tool for plant disease detection due to their capacity to automatically learn hierarchical features directly from images, ranging from simple edges and textures to complex disease specific patterns [1], [2]. In the context of coffee cultivation, this capability is particularly valuable, as variations in leaf color, texture, and morphology associated with diseases such as rust, leaf miner, and Phoma can be effectively captured by CNN architectures [3]. Several studies have demonstrated the successful application of CNN based models for coffee leaf disease classification, achieving high accuracy under controlled conditions [4].

However, deploying conventional CNN models for real-world coffee disease detection poses practical constraints, as farmers in remote areas often run them on resource-constrained mobile devices. Additionally, existing solutions exhibit limited diagnostic capabilities, often identifying only a narrow range of disease types and failing to cover the full spectrum of leaf conditions found in diverse farming environments. The high computational demands of standard CNN architectures make them impractical for such edge devices, creating a gap between laboratory proven technologies and their actual adoption on coffee plantations [5]. Consequently, farmers continue relying on time consuming manual inspection methods that delay critical treatment decisions [6].

The high memory and computational demands of conventional CNNs (e.g., VGG16, ResNet50) limit their applicability in real-world agricultural settings, particularly on mobile and edge platforms [7]. MobileNetV2, NASNetMobile, and EfficientNetV2B0 are lightweight CNNs that address this by balancing efficiency and accuracy, making them ideal for field-level diagnostics [8].

Lightweight CNN models are effective for plant disease detection across multiple studies. MobileNetV2, for instance, achieved a classification accuracy of up to 96.67% in soybean leaf disease detection [9], while an enhanced EfficientNetV2-S model attained 98.1% accuracy in coffee disease classification with fewer parameters than DenseNet-121 [10]. NASNetMobile has also demonstrated strong generalization capability through transfer learning, maintaining accuracy above 92.6%

across maize, tomato, and rice datasets [11]. Beyond single-model optimization, recent research trends indicate a shift toward hybrid architecture. Hybrid CNN models combine local feature extraction and long-range attention to enhance feature representation, consistently outperforming conventional architectures [12].

The hybrid CNN architectures integrate multiple models to exploit complementary feature representations have consistently demonstrated superior performance in plant disease classification [13]. Several studies provide evidence for this trend. For example, the integration of ResNet-50, DenseNet-121, and Xception achieved 96.83% accuracy in sesame disease classification [14], while a feature fusion approach using VGG16 and ResNet-50 reached 98.22% accuracy for potato leaf disease detection [15]. Other research combined MobileNetV2 and NASNetMobile for tomato leaf disease classification, achieving accuracies exceeding 97% [16]. Furthermore, a MobileNetV2 and NASNetMobile ensemble optimized through Neural Architecture Search achieved high accuracy in identifying banana nutrient deficiencies [17]. An ensemble approach integrating NASNetMobile and InceptionResNetV2 also attained 98.91% accuracy in cucumber leaf disease classification [18].

Study [3] is limited by a small, imbalanced dataset and poor multi-class performance, both of which reduce generalization and increase the risk of overfitting. In [10], the evaluation lacks cross-validation, statistical testing, and independent validation, which weakens the reliability of the reported accuracy. The method in [19] focuses solely on binary classification and relies heavily on augmented data, with limited robust real-world testing, limiting its applicability to complex scenarios. This highlights the need for an ensemble-based approach to improve generalization and consistency in performance.

To address these limitations, a hybrid lightweight CNN model specifically designed for Robusta coffee leaf disease classification is implemented, integrating MobileNetV2 and NASNetMobile at the feature level. This integration is theoretically advantageous, as MobileNetV2 provides efficient feature extraction via depthwise separable convolutions, while NASNetMobile captures complex patterns and produces refined, discriminative representations via NAS-

optimized cells. The features from both complementary architectures are combined to create a holistic representation, which is then passed to the final softmax classifier. This hybrid approach results in stable, multi-class predictions that maintain high accuracy while remaining suitable for practical edge deployment.

2. METHODS

This research adopts a deep learning framework grounded in Convolutional Neural Networks (CNNs), integrating an hybrid learning strategy with two pre-trained model architectures: MobileNetV2 and NASNetMobile. These models are further optimized through transfer learning to enhance classification performance for common Robusta coffee leaf diseases. In the methodological design, this study refers to prior research [17] that implemented a hybrid framework combining NASNetMobile and MobileNetV2 for plant leaf disease detection. The referenced study demonstrated the effectiveness of integrating these architectures, particularly on datasets comprising coffee leaf and banana leaf images, thereby providing a relevant empirical foundation for the hybrid strategy adopted in the present research.

2.1. Data Collection

Data for this study were drawn from two distinct sources, resulting in a diverse dataset that closely aligns with the research context. The primary source comprises data provided in [20]. To capture domain-specific variations and local environmental conditions, an additional private dataset was collected through direct image acquisition at a Robusta coffee plantation in Pagergunung Village, Pringsurat District, Temanggung Regency, Indonesia. This combined approach enriches the dataset by incorporating variations in lighting conditions, image acquisition angles, and regional agricultural characteristics, thereby providing a robust, representative foundation for model development.

2.2. Data Description

The data description begins with image acquisition, including images from the Mendeley dataset and a private dataset. The dataset concentrates on three categories: healthy leaves, leaf rust, and phoma. A total of

3,678 images were utilized, with the distribution of data shown in Table 1. The combination of public and private sources introduces beneficial variability in capture conditions, ranging from controlled laboratory settings to natural field environments, thereby reducing the risk of acquisition bias. Nevertheless, differences in class proportions may introduce imbalance bias, which is addressed through source stratification and targeted data augmentation to bolster minority class representation.

Table 1. *Dataset of Robusta coffee leaves*

No	Class	Amount of images
1	Healthy	1241
2	Leaf rust	1156
3	Phoma	1281
	Total	3678

Sample images of Robusta coffee leaves used in this study for the three disease classes, namely healthy leaves, leaf rust, and phoma, are presented in Figure 1.

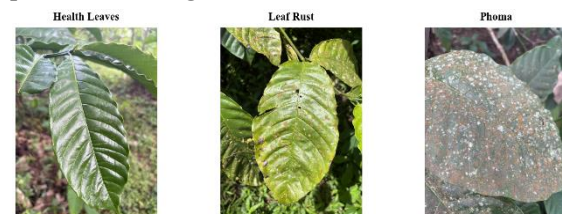


Figure 1. *Visual sample examples from three classes in the dataset*

2.3. Data Preprocessing

The initial phase of data preprocessing involves filtering to select only relevant, high-quality images for this study [21]. This step eliminates unsuitable data, consequently enhancing the CNN model's classification performance [22]. After data splitting, augmentation is performed. Rotation, flipping, zooming, and shifting transformations are introduced to artificially expand the training set's diversity without affecting the validation or test distributions. Figure 2 presents examples of augmented images. This strategy is intended to improve the model's adaptability and generalization capability when encountering data variations, while also reducing the risk of overfitting [23].

The subsequent stage involves image scaling to obtain uniform dimensions compatible with the employed CNN architecture. This normalization process plays a crucial role in ensuring input data consistency

and supporting stable model training [24]. As a final step, all images are resized to 224×224 pixels, a dimension favoured by state-of-the-art CNN architectures (e.g., ResNet, VGG) for its ability to balance computational efficiency with classification accuracy [25].

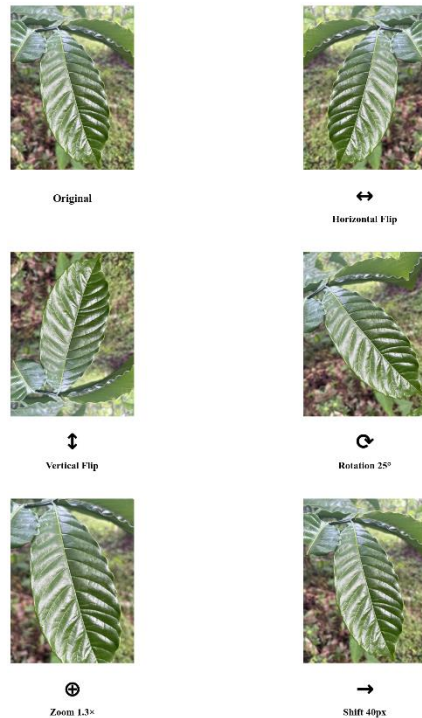


Figure 2. Data augmentation techniques applied to Robusta coffee leaf images

2.4. Data Splitting

Following standard practice, model performance was assessed using stratified 5-fold cross-validation, with the dataset partitioned into five equal folds and the average validation accuracy across all folds used for evaluation [26]. This configuration allows the cross-validation folds to guide hyperparameter selection and mitigate overfitting. At the same time, a separate held-out test set is reserved exclusively for assessing the model's generalization performance on unseen data [27].

2.5. Classification Model

A total of four CNN architectures were employed in this study for Robusta coffee leaf disease classification, including one hybrid model and three standalone models. The hybrid model combines features extracted from MobileNetV2 and NASNetMobile via feature concatenation, while the single models include MobileNetV2, EfficientNetV2B0, and NASNetMobile. All models employ pre-trained weights from the ImageNet dataset, with

training, validation, and evaluation conducted on the respective training, validation and testing datasets to assess model performance and generalization.

a. MobileNetV2

MobileNetV2 is a lightweight CNN architecture purpose-built for mobile and embedded scenarios [28]. It uses depthwise separable convolutions, which split the conventional convolution process into depthwise and pointwise components, resulting in lower parameter counts and reduced computational complexity while retaining robust feature extraction [19]. Furthermore, the incorporation of inverted residuals and linear bottlenecks improves memory efficiency and the quality of feature representations [29]. In this study, the pre-trained MobileNetV2 model was adapted for Robusta coffee leaf disease classification by employing it as a feature extractor and appending task-specific classification layers.

b. EfficientNetV2B0

EfficientNetV2 is a series of Convolutional Neural Network (CNN) models created to deliver high accuracy while enhancing computational efficiency compared to EfficientNet [30]. This architecture integrates compound scaling and Fused-MBConv to reduce latency and enhance performance [31]. EfficientNetV2B0, the smallest variant, delivers fast inference without compromising accuracy, making it suitable for resource-constrained deployment scenarios [32]. This study employs EfficientNetV2B0 as a single model via transfer learning, leveraging pre-trained weights from the ImageNet dataset.

c. NASNetMobile

NASNetMobile is a Convolutional Neural Network (CNN) architecture generated using the Neural Architecture Search (NAS) approach to automatically design optimal network structures [17]. This model is composed of normal cells and reduction cells that are learned automatically and stacked to form an efficient network [17]. NASNetMobile reduces computational demands by limiting its parameter count and floating-point operations (FLOPs), making it well-suited for mobile environments while maintaining strong classification results [33]. This study adopts NASNetMobile as an individual model, applying transfer learning from pre-trained ImageNet weights to extract features from Robusta coffee leaf images.

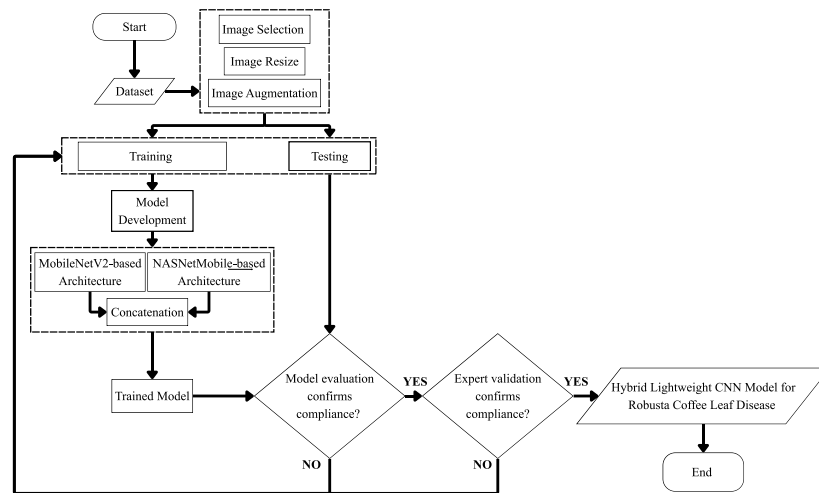


Figure 3. General diagram of the Robusta coffee leaf disease detection process

2.6. System Design

The Robusta coffee leaf disease detection system is designed as a comprehensive workflow that starts with a Robusta coffee leaf image dataset and progresses through preprocessing, dataset partitioning, model development, evaluation, and expert validation, as shown in Figure 3. The dataset undergoes preprocessing, including image selection, resizing to a consistent input size, and data augmentation to increase variability and improve generalization. Model development and evaluation employed 5-fold cross-validation, in which the data were partitioned into 5 stratified folds for iterative training and validation, with an independent test subset retained for final performance assessment. During model development, MobileNetV2 and NASNetMobile are selected as complementary lightweight feature extractors to balance efficiency and representational capacity. Their outputs are concatenated at the representation level and subsequently reduced to 64 dimensions, a strategy that improves generalization and limits memory usage for edge deployment without additional computational overhead.

Evaluation of the trained model is first conducted on the test dataset to determine whether it meets the predefined performance criteria. If the evaluation indicates that the criteria are not satisfied, the workflow returns to the model development stage for further improvement. However, if the model evaluation confirms compliance, the process proceeds to the expert validation stage to verify the consistency between the model predictions and expert assessments. If the expert validation also

confirms satisfactory performance, the final output is established as a hybrid lightweight CNN model for Robusta coffee leaf disease classification integrating the MobileNetV2 and NASNetMobile architectures via a concatenation mechanism. Conversely, if the expert validation does not meet the required criteria, the process returns to the model development phase for further refinement until the desired performance level is attained.

2.7. Implemented Method

Following ensemble learning paradigms from previous studies [17], a hybrid ensemble-based deep learning approach for coffee leaf disease classification is implemented by combining two lightweight CNN architectures, namely MobileNetV2 and NASNetMobile. The hybrid model is designed to exploit the complementary characteristics of the two architectures: MobileNetV2 efficiently extracts features via depthwise separable convolutions [34], while NASNetMobile extracts features from architectures generated via the Neural Architecture Search (NAS) approach [35], yielding more optimal feature representations. The lightweight CNN hybrid for Robusta coffee leaf disease classification, along with its architectural design, is presented in Figure 4.

The architecture froze the base layers of both pre-trained models while retraining the classification head. The feature extraction process for MobileNetV2, which produces a feature vector F_m via global average pooling, is represented in Equation (1). Similarly, the feature extraction process for NASNetMobile, yielding F_n , is represented in Equation (2).

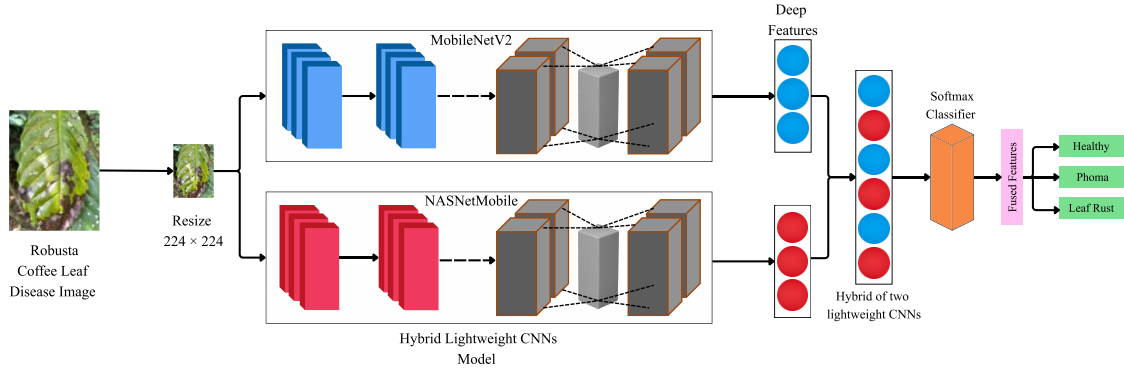


Figure 4. Architecture of the implemented hybrid lightweight CNN for Robusta coffee leaf disease classification

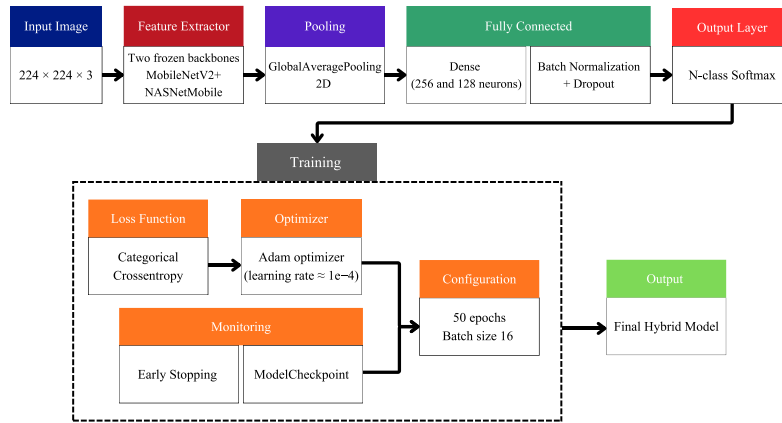


Figure 5. Overview of model architecture and training process using dual transfer learning backbones (MobileNetV2 and NASNetMobile)

$$f_{m,k} = \frac{1}{h \cdot w} \sum_{i=1}^h \sum_{j=1}^w A_{i,j,k}, \quad k = 1, \dots, d_m \quad (1)$$

$$f_{n,k} = \frac{1}{h' \cdot w'} \sum_{i=1}^{h'} \sum_{j=1}^{w'} B_{i,j,k}, \quad k = 1, \dots, d_n \quad (2)$$

These feature vectors are then passed through separate fully-connected layers to reduce dimensionality and introduce non-linearity, as represented in Equation (3)(4), following the ensemble design in [16]:

$$V_m = \text{ReLU}(W_m f_m + b_m) \quad (3)$$

$$V_n = \text{ReLU}(W_n f_n + b_n) \quad (4)$$

where $W_m \in \mathbb{R}^{64 \times d_m}$, $b_m \in \mathbb{R}^{64}$ and $W_n \in \mathbb{R}^{64 \times d_n}$, $b_n \in \mathbb{R}^{64}$. The resulting 64-dimensional vectors are subsequently combined via concatenation, as represented in Equation (5), adopting the feature fusion strategy from [16]:

$$V_{fused} = [V_m, V_n] \in \mathbb{R}^{128} \quad (5)$$

The fused representation V_{fused} is a 128-dimensional vector that integrates complementary deep features extracted from both MobileNetV2 and NASNetMobile. This combined feature set captures a richer, more discriminative representation of the input image, which is then processed by additional fully connected layers and a softmax classifier for final disease classification. The individual models were constructed analogously, utilizing Dense layers with 256 and 128 neurons. All models were trained using the Adam optimizer (initial learning rate of 1e-4) and categorical cross-entropy loss, with 50 epochs and a batch size of 16. Early stopping and ModelCheckpoint were employed to monitor and preserve optimal model states. The model architecture consists of 6,528,660 trainable parameters (24.90 MB in size).

In the hybrid model, freezing the base layers preserves the knowledge acquired from pre-trained models. The spatial feature maps are summarized into a feature vector via Global Average Pooling. Subsequently, the hybrid model processes this vector through a 128-neuron Dense layer, whereas the individual

models utilize Dense layers of 256 and 128 neurons as fully connected layers for pattern learning. Batch Normalization standardizes activations, ensuring training stability, and Dropout randomly deactivates neurons to improve regularization.

The models were implemented in Python programming language using the TensorFlow package. All experiments were conducted on a system equipped with an Intel Core i7-11370H CPU, Intel Iris Xe Graphics, and 16GB of LPDDR4x-4266 MHz RAM. Training is conducted using the Adam optimizer with an initial learning rate of 0.0001, and the categorical cross-entropy loss is used to evaluate classification error. Early stopping tracks validation performance and halts training upon convergence, while ModelCheckpoint saves the model with the best validation accuracy. A batch size of 16 is used for up to 50 epochs. Figure 5 depicts the entire training workflow.

2.8. Model Evaluation

The effectiveness of the coffee leaf disease classification model was measured using a rigorously withheld independent test dataset, with performance evaluated via confusion matrices and the metrics of accuracy, precision, recall, and F1-score [17], [19], [29]. This analysis provided detailed class-wise insights for each Robusta coffee leaf condition (healthy, leaf rust, and phoma). The model was evaluated via 5-fold cross-validation, with the final metric calculated as the mean validation accuracy over all folds. The hybrid model outperformed three individual models and aligned with prior studies, consistently achieving the highest metrics and being selected as the optimal solution.

2.9. Expert-Based Validation

The practical reliability of the classification results was validated through expert assessment [36]. A total of 120 images from the testing dataset were independently evaluated by experts from the Balai Perlindungan Tanaman Perkebunan dan Hortikultura (BPTPHP) Salatiga. The diagnoses based on the expert's field experience were then compared with the model predictions to assess agreement. This validation process was conducted without any involvement in the model training stage, thereby enhancing the credibility of the research findings and

affirming the relevance of the implemented method for practical agricultural deployment.

3. RESULTS AND DISCUSSION

3.1 Five-Fold Cross-Validation Performance of the Hybrid Model

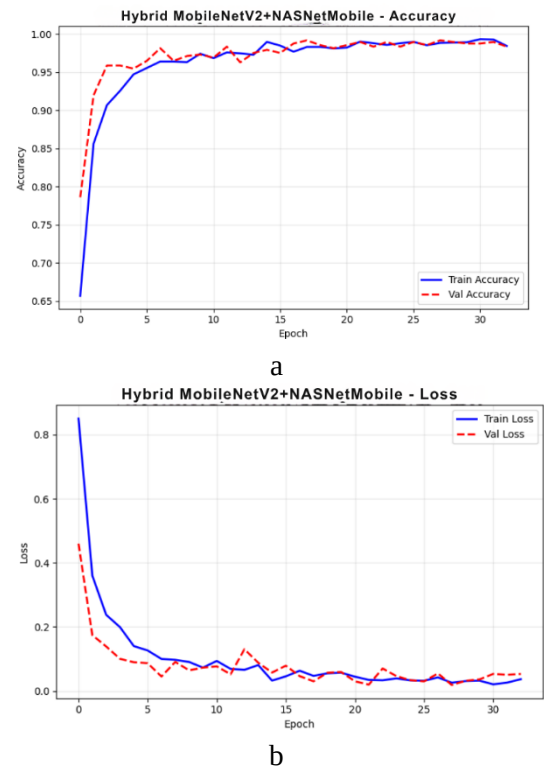


Figure 6. Performance evaluation of hybrid MobileNetV2 + NASNetMobile via 5-fold cross-validation: (a) Accuracy curves, (b) Loss curves

Figure 6(a) shows the mean training and validation accuracy curves. Accuracy rises sharply within the first ten epochs from approximately 0.65 to above 0.90, then stabilizes near 0.95–1.00 through epoch 30. The close correspondence between training and validation accuracy indicates the model does not overfit and generalizes reliably across folds.

Figure 6(b) depicts the corresponding loss trajectories. Both training and validation loss decline steeply during the initial epochs and plateau at a low value thereafter, with no divergence observed. This behavior confirms stable convergence and robust generalization.

Overall, the curves demonstrate that the hybrid lightweight model converges efficiently to a high-performance optimum, validating its suitability for classification tasks under computational constraints.

Based on the results of the 5-fold cross-validation evaluation illustrated in Figure 7, the hybrid model combining MobileNetV2 and NASNetMobile architectures demonstrated robust, consistently high performance across all data splits. The detailed accuracy for each fold is as follows: Fold 1 achieved 98.61%, Fold 2 achieved 98.61%, Fold 3 achieved 99.72%, Fold 4 achieved 98.61%, and Fold 5 achieved 100.00%. Consequently, the overall average accuracy derived from the five folds is 99.11%.

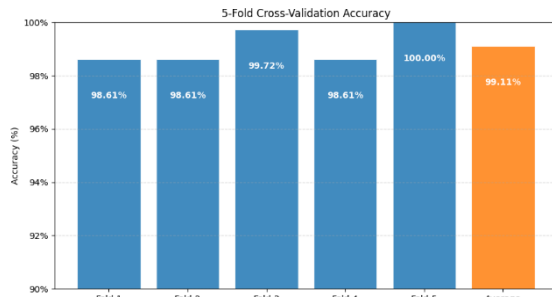


Figure 7. Five fold cross validation accuracy of the hybrid MobileNetV2 and NASNetMobile model

The high average accuracy, coupled with the narrow range of variation across the folds (spanning from 98.61% to 100.0%), indicates excellent model stability and an absence of overfitting to the training partitions. The attainment of perfect classification in Fold 5, alongside consistent performance across the other three folds, substantiates the assertion that

integrating feature representations from two distinct architectures, MobileNetV2 and NASNetMobile, yields a richer feature set and more reliable classification decisions. These findings confirm that the hybrid approach exhibits superior generalization capability when rigorously assessed using a 5-fold cross-validation framework.

3.2 Confusion Matrix Analysis of the Model

The confusion matrices in Figure 8 present the classification performance of each model across three classes: Healthy leaves, leaf rust, and phoma. These matrices provide the basis for the accuracy, precision, recall, and F1 score metrics reported in Table 2.

MobileNetV2 (Figure 8a) achieves perfect classification for Healthy leaves (1.00), correctly identifies 97% of leaf rust samples, and attains 95% accuracy for phoma with minor misclassifications. EfficientNetV2B0 (Figure 8b) also yields perfect Healthy leaves classification, with leaf rust and phoma accuracies of 98% and 97% respectively, showing minimal interclass confusion. NASNetMobile (Figure 8c) exhibits lower performance, with class accuracies of 90% for Healthy leaves, 91% for leaf rust, and 92% for phoma, along with noticeable confusion between the two disease classes.

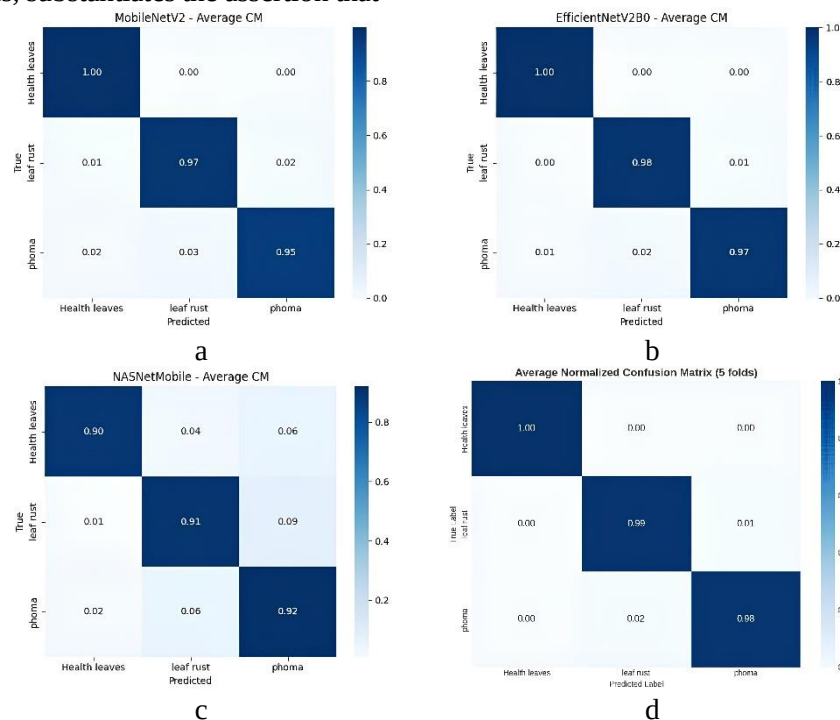


Figure 8. Confusion matrices of the experiments showing the classification performance of the CNN models (a) MobileNetV2, (b) EfficientNetV2B0, (c) NASNetMobile, and (d) The implemented hybrid lightweight CNN model

Table 2. Accuracy, precision, recall, and F1-score of three individual CNN models and the implemented method (MobileNetV2 + NASNetMobile)

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
MobileNetV2	97.79% $\pm$ 0.0094	97.75% $\pm$ 0.0101	97.43% $\pm$ 0.0114	97.56% $\pm$ 0.0109
EfficientNetV2B0	98.78% $\pm$ 0.0064	98.73% $\pm$ 0.0066	98.58 % $\pm$ 0.0078	98.65% $\pm$ 0.0072
NASNetMobile	90.82% $\pm$ 0.0056	90.12 % $\pm$ 0.0056	90.97% $\pm$ 0.0067	90.32% $\pm$ 0.0053
Hybrid MobileNetV2+ NASNetMobile (Implemented Method)	99.11% $\pm$ 0.0062	99.05 % $\pm$ 0.0069	98.97% $\pm$ 0.0071	99.00% $\pm$ 0.0069

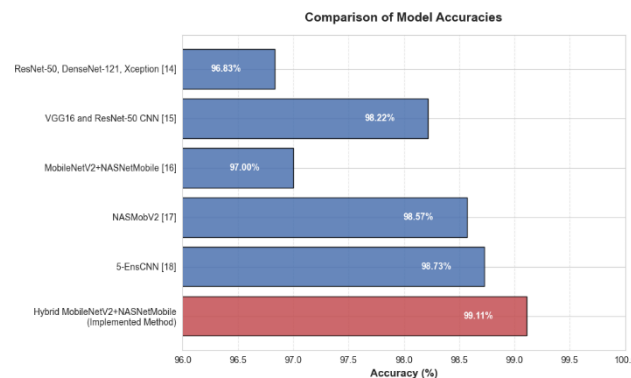


Figure 9. Comparative accuracy of hybrid and ensemble CNN architectures for leaf disease classification

The hybrid model combining MobileNetV2 and NASNetMobile (Figure 8d) demonstrates strong diagonal dominance with minimal off-diagonal errors. It achieves perfect classification for Healthy leaves (1.00) and correctly identifies leaf rust and phoma at 99% and 98%, respectively. No Healthy leaves samples are misclassified, and confusion between leaf rust and phoma is the lowest among all evaluated models. This confirms that feature fusion enhances discriminative capability and yields more consistent predictions for visually similar diseases.

A small degree of misclassification persists between leaf rust and phoma in the hybrid model, attributable to their visual resemblance in early infection stages and variability in illumination and symptom progression. Nonetheless, the marked reduction in off-diagonal errors relative to single models indicates that the hybrid approach substantially mitigates this inherent ambiguity.

Compared with ensemble models reported in previous studies that were also applied to coffee leaf datasets, particularly the NASMobileNetV2 (NasMobV2) approach [17], the earlier ensemble method still exhibits higher misclassification rates in certain classes. This comparison suggests that the hybrid model implemented in this study provides more consistent classification performance than

previous ensemble approaches, both in terms of accuracy and prediction stability across all coffee leaf disease classes.

3.3 Evaluation of Model Performance

The quantitative performance metrics of accuracy, precision, recall, and F1 score are reported as the mean and standard deviation from a 5-fold cross-validation, comparing three individual convolutional neural network models, MobileNetV2, EfficientNetV2B0, and NASNetMobile, with the implemented hybrid approach combining MobileNetV2 and NASNetMobile, as presented in Table 2. These metrics were calculated from the confusion matrices of each model, which provide the numbers of true positives, true negatives, false positives, and false negatives, as shown in Figure 8.

This comparison also serves as an ablation study, directly contrasting the hybrid model with its constituent single-architecture components. By evaluating the individual models under identical experimental conditions and then assessing their combined output, the contribution of feature integration across two distinct architectures can be isolated and quantified.

As shown in Table 2, the hybrid model achieves the highest average performance across all metrics, attaining an accuracy of

99.11% $\pm$ 0.0062, precision of 99.05% $\pm$ 0.0069, recall of 98.97% $\pm$ 0.0071, and F1-score of 99.00% $\pm$ 0.0069, thereby surpassing all individual CNN architectures. The low standard deviations further indicate consistent, stable classification across data splits, confirming that the fusion of MobileNetV2 and NASNetMobile enhances both predictive accuracy and generalizability.

The bar chart in Figure 9 compares the classification accuracies reported by several prior studies that employed hybrid or ensemble Convolutional Neural Network (CNN) strategies with those of the method implemented in this work. The implemented hybrid model (MobileNetV2 + NASNetMobile) achieves the highest accuracy of 99.11%, outperforming ResNet-50 + DenseNet-121 + Xception (96.83%)[14], VGG16 + ResNet-50 (98.22%)[15], MobileNetV2 + NASNetMobile (97.00%)[16], NASMobV2 (98.57%)[17], and 5-EnsCNN (98.73%)[18]. The quantitative comparison demonstrates the superior classification capability of the implemented lightweight hybrid model within the evaluated group of architectures.

3.4. Expert-Based Model Validation

The implemented hybrid lightweight CNN model was evaluated through expert-based validation in collaboration with the Balai Perlindungan Tanaman Perkebunan dan Hortikultura (BPTPHP) in Salatiga. A total of 120 images from the testing dataset were independently assessed by plant disease experts and compared with predictions generated by the hybrid model (MobileNetV2 + NASNetMobile), with no expert involvement in the model development process.

Table 3. Summary of model prediction outcomes compared to expert diagnosis

Validation Outcome	Number of Images	Percentage (%)
Consistent with expert diagnosis	102	85
Inconsistent with expert diagnosis	18	15
Total	120	100

The expert validation results in Table 3 show that the hybrid model's predictions align with expert diagnoses for 85% of the images, indicating a high level of practical reliability.

The remaining 15% discrepancies stem primarily from inherent variability in symptom interpretation and visual similarities among disease classes, which also pose challenges for human experts. This strong agreement underscores the hybrid model's ability to capture complex disease patterns, reinforcing its potential as a robust decision-support tool for Robusta coffee leaf disease identification in real-world plantation settings.

3.5. Comparative Summary

The implemented hybrid model outperforms individual CNN models with greater consistency across the 5-fold cross-validation procedure, as evidenced by the mean validation and test accuracy profiles, indicating reduced overfitting. The results confirm that integrating MobileNetV2 and NASNetMobile yields a more balanced and stable performance profile than either model alone.

Quantitative analysis shows the hybrid reduces inter-class misclassification, achieving superior precision, recall, and F1-score. Compared to NASMobV2 [17], which achieved 98.57% validation accuracy but addressed a broader range of deficiencies across multiple crops, the implemented approach achieves higher accuracy specifically for three Robusta coffee disease classes. Nevertheless, expert validation yielded an 85% agreement rate, confirming the model's reliability as a decision-support system for Robusta coffee leaf disease detection. The discrepancy between the 99.11% test accuracy and the 85% expert agreement stems from the controlled, curated nature of the test dataset versus the inherent variability of real-world field conditions, including lighting, symptom severity, and subjective diagnostic thresholds. This gap is expected in agricultural machine learning applications and highlights the value of expert-in-the-loop validation for operational deployment.

CONCLUSION

A hybrid lightweight CNN model was successfully developed for the classification of Robusta coffee leaf conditions, including healthy leaves, leaf rust, and phoma. With 99.11% accuracy, 99.05% precision, 98.97% recall, and 99.00% F1-score, the hybrid MobileNetV2-NASNetMobile model outperformed the individual architectures.

These results confirm that the hybrid approach enhances inter-class discrimination without additional computational cost, and expert validation (85% agreement) supports its field reliability. However, limitations include training data restricted to Robusta coffee leaf diseases, with unverified performance on Arabica varieties. Deployment risks involve accuracy degradation under variable lighting, motion blur, and leaf occlusion. Expert oversight remains necessary for ambiguous or early-stage infections. Future work should expand the dataset diversity and conduct field validation across varied agronomic conditions.

DECLARATIONS

Author Contributions: The first author contributed to conceptualization, methodology, investigation, software, data curation, and original draft preparation. The second author contributed to conceptualization, methodology, investigation, supervision, visualization, final review, and manuscript approval. The third author contributed to methodology, conceptualization, formal analysis, and manuscript review. The fourth and fifth authors contributed to conceptualization, data curation, and manuscript review. All authors reviewed and approved the final manuscript.

Funding: This research received no external funding. All research activities were supported through personal funds from the authors.

Conflict of Interest: The authors declare no conflict of interest.

Generative Artificial Intelligence Use Statement: Generative AI tools were used only for language improvement and manuscript editing. All AI-generated suggestions were reviewed and selectively revised by the authors. No Generative AI was used for generating scientific content, research data, analysis, interpretation, or conclusions.

Ethics Approval and Informed Consent: Not applicable. This study did not require ethical approval or informed consent.

Data Availability: The data supporting this study were obtained from two datasets. The public dataset is available at

(<https://data.mendeley.com/datasets/k36wnd6knb/1>), while the private dataset is not publicly available and can be obtained from the corresponding author upon reasonable request.

ACKNOWLEDGMENTS

The authors offer their sincere thanks to the UNNES Electrical Engineering Students Research Group (UEESRG) and the Department of Electrical Engineering, Universitas Negeri Semarang, for their ongoing support and for making resources available that facilitated this research.

REFERENCES

- [1] B. Tugrul, E. Elfatimi, and R. Eryigit, "Convolutional Neural Networks in Detection of Plant Leaf Diseases: A Review," Aug. 01, 2022, *MDPI*. doi: 10.3390/agriculture12081192.
- [2] T. D. Salka, M. B. Hanafi, S. M. S. A. A. Rahman, D. B. M. Zulperi, and Z. Omar, "Plant leaf disease detection and classification using convolution neural networks model: a review," *Artif. Intell. Rev.*, vol. 58, no. 10, Oct. 2025, doi: 10.1007/s10462-025-11234-6.
- [3] W. Ode Asriyani *et al.*, "Expert System for Coffee Leaf Disease Classification With Convolutional Neural Networks Sistem Pakar Klasifikasi Penyakit Pada Daun Kopi Dengan Metode CNN," *Jurnal Informatika dan Teknologi Informatika*, vol. 22, no. 3, pp. 1–12, 2025, doi: 10.31515/telematika.v22i3.14323.
- [4] O. Adelaja and B. Pranggono, "Leveraging Deep Learning for Real-Time Coffee Leaf Disease Identification," *AgriEngineering*, vol. 7, no. 13, 2025, doi: 10.3390/agriengineering7010013.
- [5] I. V. C. Motta, N. Vuillerme, H. H. Pham, and F. A. P. de Figueiredo, "Machine learning techniques for coffee classification: a comprehensive review of scientific research," *Artif. Intell. Rev.*, vol. 58, no. 15, 2024, doi: 10.1007/s10462-024-11004-w.
- [6] C. Calderón-Mosilot, U. Tapia-Gálvez, J. Arcila-Díaz, and H. I. Mejía-Cabrera, "Diagnosis of Nutritional Deficiencies in Coffee Plants Through Automated Analysis of Digital Images Using Deep Learning in Uncontrolled Agricultural Environments," *AgriEngineering*, vol. 7, no. 421, 2025, doi: 10.3390/agriengineering7120421.

- [7] M. D. Islam *et al.*, “Towards real-time weed detection and segmentation with lightweight CNN models on edge devices,” *Comput. Electron. Agric.*, vol. 237, Jun. 2025, doi: 10.1016/j.compag.2025.110600.
- [8] S. Khan, F. Siddiqui, and M. A. Ahad, “Bridging efficiency and interpretability: Explainable AI for multi-classification of pulmonary diseases utilizing modified lightweight CNNs,” *Image Vis. Comput.*, vol. 158, May 2025, doi: 10.1016/j.imavis.2025.105553.
- [9] I Gede Rian Lanang Oka, Anak Agung Gede Bagus Ariana, Wayan Sauri Peradhayana, Ni Luh Wiwik Sri Rahayu Ginantra, and I Ketut Sutarwiyasa, “Performance Comparison of MobileNetV2 and NASNetMobile Architectures in Soybean Leaf Disease Classification,” *Indonesian Journal of Data and Science*, vol. 6, no. 2, pp. 251–259, Jul. 2025, doi: 10.56705/ijodas.v6i2.243.
- [10] S. Kulkarni *et al.*, “Coffee Plant Disease Identification using Enhanced Short Learning EfficientNetV2,” in *2023 IEEE 20th India Council International Conference, INDICON 2023*, Institute of Electrical and Electronics Engineers Inc., 2023, pp. 91–96. doi: 10.1109/INDICON59947.2023.10440883.
- [11] W. Shafik, A. Tufail, C. De Silva Liyanage, and R. A. A. H. M. Apong, “Using transfer learning-based plant disease classification and detection for sustainable agriculture,” *BMC Plant Biol.*, vol. 24, no. 1, Dec. 2024, doi: 10.1186/s12870-024-04825-y.
- [12] K. Dhakyanakabdel and A. J., “A Hybrid CNN-Vision Transformer Framework for Optimized Leaf Disease Detection Using Feature Fusion and Transfer Learning,” *Journal of Robotics and Control (JRC)*, vol. 6, no. 6, p. 2025, doi: 10.18196/jrc.v6i6.28426.
- [13] K. Taji, A. Sohail, T. Shahzad, B. S. Khan, M. A. Khan, and K. Ouahada, “An Ensemble Hybrid Framework: A Comparative Analysis of Metaheuristic Algorithms for Ensemble Hybrid CNN Features for Plants Disease Classification,” *IEEE Access*, vol. 12, pp. 61886–61906, Apr. 2024, doi: 10.1109/ACCESS.2024.3389648.
- [14] A. A. Hailu, B. C. Kassa, E. A. Desta, F. B. Adugna, and A. O. Salau, “Ensemble-based sesame disease detection and classification using deep convolutional neural networks (CNN),” *Sci. Rep.*, vol. 15, no. 1, Dec. 2025, doi: 10.1038/s41598-025-08076-1.
- [15] A. K. Trivedi, T. Mahajan, T. Maheshwari, R. Mehta, and S. Tiwari, “Leveraging feature fusion ensemble of VGG16 and ResNet-50 for automated potato leaf abnormality detection in precision agriculture,” *Soft comput.*, vol. 29, no. 4, pp. 2263–2277, Feb. 2025, doi: 10.1007/s00500-025-10523-0.
- [16] M. S. A. M. Al-gaashani, F. Shang, M. S. A. Muthanna, M. Khayyat, and A. A. Abd El-Latif, “Tomato leaf disease classification by exploiting transfer learning and feature concatenation,” *IET Image Process.*, vol. 16, no. 3, pp. 913–925, Feb. 2022, doi: 10.1049/ipr2.12397.
- [17] S. Muthusamy and S. P. Ramu, “A neural architecture search optimized lightweight attention ensemble model for nutrient deficiency and severity assessment in diverse crop leaves,” *Sci. Rep.*, vol. 15, no. 1, Dec. 2025, doi: 10.1038/s41598-025-20124-4.
- [18] M. Emin Sahin, U. Özkaya, C. Arisoy, H. İ. Coşar, and H. Ulutaş, “CucuNetCNNs: Application of novel ensemble deep neural networks for classification of cucumber leaf disease,” *Ain Shams Engineering Journal*, vol. 16, no. 5, Apr. 2025, doi: 10.1016/j.asej.2025.103380.
- [19] Y. Aufar and T. P. Kaloka, “Robusta coffee leaf diseases detection based on MobileNetV2 model,” *International Journal of Electrical and Computer Engineering*, vol. 12, no. 6, pp. 6675–6683, Dec. 2022, doi: 10.11591/ijece.v12i6.pp6675-6683.
- [20] S. Chelangat, R. Anirwoth, K. N. Mayanja, and A. Sserwadda, “A Machine Learning Dataset for Classification of Common Coffee Leaf Diseases in Uganda,” vol. 1, Feb. 2025, doi: 10.17632/K36WND6KNB.1.
- [21] T. B. Nguyen-Tat, T. Q. Hung, P. T. Nam, and V. M. Ngo, “Evaluating pre-processing and deep learning methods in medical imaging: Combined effectiveness across multiple modalities,” *Alexandria Engineering Journal*, vol. 119, pp. 558–586, Apr. 2025, doi: 10.1016/j.aej.2025.01.090.
- [22] R. Zhang, B. Zhou, C. Lu, and M. Ma, “The Performance Research of the Data Augmentation Method for Image Classification,” *Math. Probl. Eng.*, vol. 2022, 2022, doi: 10.1155/2022/2964829.
- [23] T. Kumar, R. Brennan, A. Mileo, and M. Bendechache, “Image Data Augmentation Approaches: A Comprehensive Survey and Future Directions,” *IEEE Access*, vol. 12, pp. 187536–187571, Sep. 2024, doi: 10.1109/ACCESS.2024.3470122.
- [24] M. M. Musthafa, I. Manimozhi, T. R. Mahesh, and S. Guluwadi, “Optimizing double-layered convolutional neural networks for efficient lung cancer classification through hyperparameter optimization and advanced image pre-processing techniques,” *BMC Med. Inform.*

- Decis. Mak.*, vol. 24, no. 1, Dec. 2024, doi: 10.1186/s12911-024-02553-9.
- [25] I. U. W. Mulyono, E. H. Rachmawanto, C. A. Sari, and M. K. Sarker, "A high accuracy of deep learning based CNN architecture: classic, VGGNet, and RestNet50 for Covid-19 image classification," *Telkomnika (Telecommunication Computing Electronics and Control)*, vol. 22, no. 5, pp. 1187–1195, Oct. 2024, doi: 10.12928/TELKOMNIKA.v22i5.26017.
- [26] D. Rastogi, P. Johri, V. Tiwari, and A. A. Elngar, "Multi-class classification of brain tumour magnetic resonance images using multi-branch network with inception block and five-fold cross validation deep learning framework," *Biomed. Signal Process. Control*, vol. 88, p. 105602, Feb. 2024, doi: 10.1016/J.BSPC.2023.105602.
- [27] T. M. Lam, S. N. Prova, T. e. Mustafa, and N. A. Khan, "Improving Lung Cancer Prediction Using Advanced Hyperparameter Optimization and Nested Cross-Validation," *Engineering Reports*, vol. 7, no. 11, p. e70486, Nov. 2025, doi: 10.1002/ENG2.70486;WGROU:STRING: PUBLICATION.
- [28] C. E. S. E. Silva, J. B. Fragoso, T. Paixao, A. B. Alvarez, and F. Palomino-Quispe, "A Low Computational Cost Deep Learning Approach for Localization and Classification of Diseases and Pests in Coffee Leaves," *IEEE Access*, vol. 13, pp. 71943–71964, 2025, doi: 10.1109/ACCESS.2025.3562832.
- [29] B. Mehta, S. Bharany, D. H. Elkamchouchi, A. U. Rehman, R. Singh, and S. H. Adem, "EffResViT-SE FusionNet: A Hybrid Deep Learning Framework for Accurate Classification of Coffee Leaf Diseases," *Food Sci. Nutr.*, vol. 13, no. 12, Dec. 2025, doi: 10.1002/fsn3.71311.
- [30] Z. Zhao, E. B. A. Bakar, N. B. A. Razak, and M. N. Akhtar, "Progressive optimization of EfficientNetV2 for classification based on images of corroded objects," *Computational and Applied Mathematics*, vol. 44, no. 419, Aug. 2025, doi: 10.1007/s40314-025-03382-w.
- [31] J. R. Korsipati, R. M. R. Yanamala, A. Pallakonda, R. D. A. Raj, and K. K. Prakasha, "Multi-resolution transfer learning for tampered image classification using SE-enhanced fused-MBConv and optimized CNN heads," *Sci. Rep.*, vol. 15, no. 1, Dec. 2025, doi: 10.1038/s41598-025-17799-0.
- [32] R. Tobiasz, G. Wilczyński, P. Graszka, N. Czechowski, and S. Łuczak, "Edge Devices Inference Performance Comparison," Jun. 2023, doi: 10.5626/JCSE.2023.17.2.51.
- [33] M. Reda, R. Suwwan, S. Alkafri, Y. Rashed, and T. Shanableh, "AgroAid: A Mobile App System for Visual Classification of Plant Species and Diseases Using Deep Learning and TensorFlow Lite," *Informatics*, vol. 9, no. 3, Sep. 2022, doi: 10.3390/informatics9030055.
- [34] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," Mar. 2019, [Online]. Available: <http://arxiv.org/abs/1801.04381>
- [35] B. Zoph, V. Vasudevan, J. Shlens, and Q. V. Le, "Learning Transferable Architectures for Scalable Image Recognition," Apr. 2018, [Online]. Available: <http://arxiv.org/abs/1707.07012>
- [36] W. M. Alawad, A. Alharbi, S. Alanazi, T. Almotairy, A. I. Alhujaylan, and K. E. Hamed, "Development of a Certainty Factor-Based Mobile Expert System for Palm Tree Pest and Disease Classification," *Univers. J. Agric. Res.*, vol. 13, no. 5, pp. 68–80, 2025, doi: 10.13189/ujar.2025.130502.